



OPEN

# Curvilinear regression analysis of benzenoid hydrocarbons and computation of some reduced reverse degree based topological indices for hyaluronic acid-paclitaxel conjugates

Vignesh Ravi &amp; Kalyani Desikan✉

Graph theoretical molecular descriptors alias topological indices are a convenient means for expressing in numerical form the chemical structure encoded in a molecular graph. The structure descriptors derived from molecular graphs are widely used in quantitative structure-property relationship (QSPR) and quantitative structure-activity relationship (QSAR) studies. The reason for introducing new indices is to obtain predictions of target properties of considered molecules that are better than the predictions obtained using already known indices. In this paper, we apply the reduced reverse degree based indices introduced in 2021 by Vignesh et al. In the QSPR analysis, we first compute the reduced reverse degree based indices for a family of benzenoid hydrocarbon molecules and then we obtain the correlation with the Physico-chemical properties of the considered molecules. We show that all the properties taken into consideration for the benzenoid hydrocarbons can be very effectively predicted by the reduced reverse degree based indices. Also, we have compared the predictive capability of reduced reverse degree based topological descriptors against 16 existing degree based indices. Further, we compute the defined reduced reverse degree based topological indices for Hyaluronic Acid-Paclitaxel Conjugates ( $HAP)_n$ ,  $n \geq 1$ .

The most important part of the drug discovery process is finding and improving promising lead compounds as quickly and cheaply as possible. Drug design methodologies can be classified into two primary groups: ligand based and ligand-receptor interaction based approaches. Ligand based approaches depend only on the structure of ligands. Ligand-receptor interaction based approaches are often referred to as structure-based methods (docking, molecular dynamics, etc.).

The past decade has seen tremendous progress in computational approaches to drug design, and this has played a crucial part in the creation of several currently available medications. Molecular topology<sup>1</sup> is a new approach used to design and select new molecules, particularly new pharmaceuticals. It uses topological factors rather than physical or geometric dimensions. Molecular topology not only provides an alternate method but also a different paradigm compared to conventional drug design methodologies. Molecular topology has proven to be useful in molecular design despite its inability to account for the physical or chemical properties of ligands and receptors or the ligand-receptor interaction. In drug development and design, molecular topology has been used to find new hits and leads. The discovery of novel lead compounds using molecular topology has enabled the interpretation of results in terms of their structural and physicochemical properties. The toxicity profile of a drug can be affected by its physicochemical properties, which are thought to play a role in its absorption, distribution, metabolism, and elimination. Physicochemical descriptors form the basis of most current approaches to molecular and drug design<sup>2,3</sup>. In this manner, a direct connection can be established between the experimental properties and chemical structures.

Division of Mathematics, School of Advanced Sciences, Vellore Institute of Technology, Chennai, India. ✉email: kalyanidesikan@vit.ac.in

A common critique of molecular topology, however, is that it functions like a “black box,” bypassing the mechanism of action that offers neither a physical explanation nor a chemical explanation of the processes<sup>4</sup>. However, this critique may be answered by the fact that topology determines physical and geometrical magnitudes, a claim backed by both experimental and theoretical physicists. This results in molecular topology becoming a self-consistent and independent paradigm describing the molecular behavior. The results of molecular topology can be understood both in terms of structure and in terms of physicochemical characteristics. This leads to both the mechanisms of action and the underlying physical chemistry to be deduced from the mathematical topological pattern. This is the most rational explanation, as it is the physical variable that is dependent on mathematics, not the other way around.

Molecular topology is independent of physicochemical molecular descriptors. This fact explains why it is feasible to find new lead compounds using only the knowledge gleaned from mathematical-topological patterns and to interpret the findings in terms of their structural and physicochemical properties. Molecular Topology has transformed from a potential intriguing option to a cornerstone of the drug discovery process. One of them is Quantitative Structure Analysis Relationship (QSAR), which is still effective and common and is also coupled with machine learning<sup>5–8</sup>. Additionally, molecular topology has become more relevant recently in terms of novel or unexpected uses. For instance, Speck-Planche et al. developed the multi-scale de novo drug design paradigm, which provides instructions to create novel compounds with desirable drug-like properties<sup>9</sup>.

In general, descriptors are classified into two types, namely, Experimental (logP, aqueous solubility etc.) and Theoretical (assessed in Silico from 1D, 2D or 3D molecular structures) descriptors<sup>1</sup>. Based on the parameterization type, the descriptors are sorted into distinct categories: geometric, quantum-chemical, electrostatic, constitutional, and topological<sup>10</sup>. Descriptors based on the molecular graph representation are widely used because they incorporate precious chemical information like size, degree of branching, neighbourhood of atoms, flexibility and overall shape. Molecular topological indices can be broken down into structure-explicit descriptors like quantum chemical ones and structure-implicit ones like hydrophobicity or electronic constants ones<sup>11</sup>. Incorporating ideas from molecular topology makes computing these indices straightforward. In fact, many physicochemical and biological properties can be quickly and accurately estimated using molecular topology<sup>12</sup>.

Based on the known features of existing pharmaceuticals, topological indices can be used to design new lead compounds. These medications have superior pharmacokinetic, pharmacodynamic and toxicological qualities compared to those already available. This can be accomplished at a very low computational cost and economic cost within a short period of time. Due to the need to reduce the costs associated with both the synthesis and clinical testing of pharmaceuticals, the pharmaceutical industry has contributed to the increased interest in molecular descriptors. The development of the predictive quantitative structure-property relationship (QSPR) and quantitative structure-activity relationship<sup>13</sup> (QSAR) models is particularly important for the design of fine chemicals and pharmaceuticals that are made for a specific purpose. For Structure Activity Relationship (SAR) connections to be developed for medicinal compounds using computational or theoretical methods, accurate representations of molecular structure are essential. The molecular descriptor is the final outcome of applying a logical and mathematical procedure that enables the transformation of chemical information encoded in the symbolic representation of a molecule into a meaningful number.

Lead discovery and lead optimization are only two examples of how topological indices are put to use in the drug design and discovery process. Topological indices play an important part in the process of developing any QSPR or QSAR<sup>14</sup> model because they quantitatively represent the chemical information that is encoded. They allow researchers to delve into the mechanical aspects of a biochemical process and aid in uncovering the mathematical correlation that exists between chemical structure information and the response of interest. For instance, natural bio-polymer monomers (nucleotides and amino acids) can act as nodes in a graph, with the edges representing covalent bonds, hydrogen bridges, electrostatic interactions, and van der Waals bonds. As a result, the structure of complex biopolymers can be reduced to the topology of a graph, revealing useful information about these molecular systems. Topological indices can be used to characterize graphs that represent molecular systems. Topological indices store information about the connections between molecular atoms and the properties of those atoms. Natural biopolymers such as DNA, RNA, and protein sequences can also be characterized in this manner.

Considering that a theoretical physics starting hypothesis was proved for the first time in experimental physics, the following result is pertinent<sup>4</sup>. Traditionally it has been believed that topology is dependent on energy. This false assumption gave rise to the criticism that topological descriptors serve as a “black box” since they lack a physical interpretation. Now, if topology (a mathematical magnitude) is what determines energy (a physical magnitude) rather than the other way around, as advanced physicists and chemists claim, then mathematical descriptors can be used to predict the energy of any system along with other physical or experimental properties dependent on it. This is precisely what happens when novel active molecules are created or found utilising topological indices.

Quantitative structure-property relationship (QSPR) and Quantitative structure-activity relationship (QSAR)<sup>15</sup> studies rely heavily on molecular descriptors, many of which are based on topological indices.

In a chemical graph, nodes represent atoms or molecules and the links denote the chemical bonding between the atoms or molecules. Graph theoretical molecular descriptors are topological indices. These are graph invariants that play an important role in pharmaceutical science, chemistry, materials science and engineering, etc. The value of a molecular descriptor is not dependent on the particular molecular representation, such as atom numbering or labeling. Molecules of hydrocarbons are modeled using the corresponding molecular graph. Here the vertices represent the carbon atoms and the edges represent the bonds between them. The degree of a vertex represents the number of edges that are incident on that vertex and it is denoted by  $d_u$  or  $d(u)$ <sup>16</sup>.

Benzenoid hydrocarbons are represented through benzenoid graphs comprising of hexagons. Here there are only two types of vertices, (*i.e.*), vertices of degree 2 or 3<sup>17,18</sup>. In benzenoids, there are only 3 types of edges,

(i.e.), (2, 2), (2, 3) and (3, 3) where the numbers in each tuple denote the degree of the end vertices of the corresponding edge. Twelve physico-chemical properties of benzenoid hydrocarbons have been selected based on the availability of data: Boiling Point (BP), Critical Temperature (CT), Critical Volume (CV), Critical Pressure (CP), Exact Mass, Heavy Atom Count, HL, GE, Log P, MR, PI and Molecular Weight (MW). The data for the benzenoid hydrocarbon molecules is presented in Table 1. Energy of benzenoid hydrocarbons are taken from<sup>19</sup>. The experimental values of boiling points were taken from Basak et al.<sup>20</sup>.

Vignesh et al.<sup>21</sup> defined the reduced reverse degree as

$$\mathcal{RR}(v) = \Delta(G) - d(v) + 2 \quad (1)$$

Vignesh et al.<sup>21</sup> proposed some reduced reverse degree-based topological indices. These indices are

$$\mathcal{RRM}_1(G) = \sum_{uv \in E} [\mathcal{RR}(u) + \mathcal{RR}(v)] \quad (2)$$

$$\mathcal{RRM}_2(G) = \sum_{uv \in E} [\mathcal{RR}(u) * \mathcal{RR}(v)] \quad (3)$$

$$\mathcal{RRHM}_1(G) = \sum_{uv \in E} [\mathcal{RR}(u) + \mathcal{RR}(v)]^2 \quad (4)$$

$$\mathcal{RRHM}_2(G) = \sum_{uv \in E} [\mathcal{RR}(u) * \mathcal{RR}(v)]^2 \quad (5)$$

$$\mathcal{RRF}(G) = \sum_{uv \in E} [\mathcal{RR}(u)^2 + \mathcal{RR}(v)^2] \quad (6)$$

S. No.	BP	CT	CP	CV	GE	LOG P	MR	HL	PI	MW	EM	HAC
1	80.1	323.79	47.69	263.5	121.68	2.03	25.28	0.66	8	78.11	78.04695	6
2	218	484.95	38.97	409.5	252.38	3.03	42.45	1.67	13.6832	128.17	128.0626	10
3	340	586.11	32.43	555.5	383.08	4.03	59.62	2.68	19.3137	178.23	178.0783	14
4	340	586.11	32.43	555.5	383.08	4.03	59.62	2.68	19.4483	178.23	178.0783	14
5	432.02	650.8	27.41	701.5	513.78	5.9	76.79	3.69	24.9308	228.3	228.0909	18
6	436.7	650.8	27.41	701.5	513.78	5.03	76.79	3.69	25.1875	228.3	228.0909	18
7	438	650.8	27.41	701.5	513.78	5.03	76.79	3.69	25.1012	228.3	228.0909	18
8	448	650.8	27.41	701.5	513.78	5.03	76.79	3.69	25.1922	228.3	228.0909	18
9	438	650.8	27.41	701.5	513.78	5.03	76.79	3.69	25.2745	228.3	228.0909	18
10	404	625.65	30.73	619.5	491.18	5.08	68.36	3.47	22.5055	202.25	202.0783	16
11	524.6	714.53	23.47	847.5	644.48	6.02	93.96	4.7	30.544	278.3	278.1096	22
12	547.5	714.53	23.47	847.5	644.48	6.02	93.96	4.7	30.7255	278.3	278.1096	22
13	524	714.53	23.47	847.5	644.48	6.02	93.96	4.7	30.8805	278.3	278.1096	22
14	524.7	714.53	23.47	847.5	644.48	6.02	93.96	4.7	30.8795	278.3	278.1096	22
15	547.5	714.53	23.47	847.5	644.48	6.02	93.96	4.7	30.7627	278.3	278.1096	22
16	525	714.53	23.47	847.5	644.48	6.02	93.96	4.7	30.999	278.3	278.1096	22
17	524.7	714.53	23.47	847.5	644.48	6.02	93.96	4.7	30.9386	278.3	278.1096	22
18	520	714.53	23.47	847.5	644.48	6.02	93.96	4.7	30.9432	278.3	278.1096	22
19	524.7	714.53	23.47	847.5	644.48	6.02	93.96	4.7	30.839	278.3	278.1096	22
20	518	714.53	23.47	847.5	644.48	6.02	93.96	4.7	30.9418	278.3	278.1096	22
21	400	689.17	26.08	765.5	621.88	5.34	85.53	4.48	28.2453	252.3	252.0939	20
22	467.5	689.17	26.08	765.5	621.88	5.34	85.53	4.48	28.3361	252.3	252.0939	20
23	495	689.17	26.08	765.5	621.88	5.34	85.53	4.48	28.222	252.3	252.0939	20
24	500	728.06	24.85	829.5	729.98	5.66	94.28	5.27	31.4251	276.3	276.0939	22
25	604	779.67	20.33	993.5	775.18	7.02	111.13	5.71	36.1557	328.4	328.1252	26
26	525	767.68	23.7	893.5	838.08	5.98	103.02	6.06	34.5718	300.4	300.0939	24

**Table 1.** Data for benzenoid hydrocarbons.

$$\mathcal{RR}ABC(G) = \sum_{uv \in E} \left[ \sqrt{\frac{\mathcal{RR}(u) + \mathcal{RR}(v) - 2}{\mathcal{RR}(u) * \mathcal{RR}(v)}}} \right] \quad (7)$$

$$\mathcal{RR}GA(G) = \sum_{uv \in E} \left[ \frac{2\sqrt{\mathcal{RR}(u) * \mathcal{RR}(v)}}{\mathcal{RR}(u) + \mathcal{RR}(v)} \right] \quad (8)$$

$$\mathcal{RR}R_{\alpha}(G) = \sum_{uv \in E} \left[ \mathcal{RR}(u) * \mathcal{RR}(v) \right]^{\alpha} \quad (9)$$

In Eq. (9) we consider  $\alpha = 1, -1, \frac{1}{2}, \frac{-1}{2}$ .

In this proposed work, we consider 26 benzenoid hydrocarbon molecules and obtain the values for the above-mentioned reduced reverse degree based topological indices to perform curvilinear analysis and ascertain the predictive capability of the indices against the physico-chemical properties of the benzenoid hydrocarbons.

### Curvilinear regression analysis of proposed indices

In this section, we analyze the aforesaid reduced reverse degree based topological indices with respect to the following physico-chemical properties of the benzenoid hydrocarbon molecules: Boiling Point (*BP*), Critical Pressure (*CP*), Critical Temperature (*CT*), Critical Volume (*CV*), Exact Mass, Heavy Atom Count, *HL*, *GE*, *Log P*, *MR*, *PI* and Molecular Weight (*MW*). The experimental values of physico-chemical properties of benzenoid hydrocarbon molecules are given in Table 1.

We analyze the topological indices vis-a-vis the physico-chemical properties using the following regression models

$$P = \alpha_1(TI) + \gamma \quad (10)$$

$$P = \alpha_1(TI)^2 + \alpha_2(TI) + \gamma \quad (11)$$

where *P* is the physical property, *TI* is the topological descriptor,  $\alpha_i$ ,  $i = 1, 2$  and  $\gamma$  represent the coefficients and constant, respectively. For the twelve physico-chemical properties we found the correlation between the properties and the twenty indices proposed by us. Based on the recommendations of the International Academy of Mathematical Chemistry (IAMC), we have only considered the topological indices for which  $R^2 \geq 0.8$ . We now present the analysis of the linear and quadratic regression models based on the  $R^2$  value and Root Mean Square Error value.

### Results and discussion

Using Eq. (10), we obtained the linear regression models for the physico-chemical properties vis-a-vis the index for which the  $R^2$  value is maximum for the property. We observe that

- $\mathcal{RR}R_{-1}$  is best suited for predicting Boiling Point (*BP*), Critical Pressure (*CP*) and *Log P* with corresponding  $R^2$  values 0.9587, 0.9539 and 0.9637, respectively.
- $\mathcal{RR}R_{\frac{-1}{2}}$  is best suited for predicting Critical Volume (*CV*), Exact Mass (*EM*), Heavy Atom Count (*HAC*), Molar Refraction (*MR*) and Molecular Weight (*MW*) with corresponding  $R^2$  values 0.9967, 0.9979, 0.9979, 0.9979 and 0.9979, respectively.
- $\mathcal{RR}HM_1$  is best suited for predicting (*GE*) with  $R^2$  value 0.9986.
- $\mathcal{RR}GA$  is best suited for predicting Critical Temperature (*CT*) with  $R^2$  value 0.9602.
- $\mathcal{RR}ABC$  is best suited for predicting (*PI*) with  $R^2$  value 0.9976.
- $\mathcal{RR}F$  is best suited for predicting Henry's Law (*HL*) with  $R^2$  value 0.9995.

Hence the linear regression models for predicting Boiling Point (*BP*), Critical Pressure (*CP*), Critical Temperature (*CT*), Critical Volume (*CV*), Exact Mass, Heavy Atom Count, *HL*, *GE*, *Log P*, *MR*, *PI* and Molecular Weight (*MW*) are:

$$\widehat{BP} = 129(\mathcal{RR}R_{-1}) - 75.26 \quad (12)$$

$$\widehat{LogP} = 1.228(\mathcal{RR}R_{-1}) + 0.296 \quad (13)$$

$$\widehat{CT} = 16.62(\mathcal{RR}GA) + 291.4 \quad (14)$$

$$\widehat{GE} = 0.9654(\mathcal{RR}HM_1) + 26.46 \quad (15)$$

$$\widehat{PI} = 1.63(\mathcal{RR}ABC) + 1.197 \quad (16)$$

$$\widehat{HL} = 0.01448(\mathcal{RRF}) - 0.03858 \quad (17)$$

$$\widehat{CP} = -6.6085(\mathcal{RRR}_{-1}) + 54.162 \quad (18)$$

$$\widehat{CV} = 73.21\left(\mathcal{RRR}_{\frac{-1}{2}}\right) + 46.92 \quad (19)$$

$$\widehat{EM} = 25.23\left(\mathcal{RRR}_{\frac{-1}{2}}\right) + 3.109 \quad (20)$$

$$\widehat{HAC} = 2.024\left(\mathcal{RRR}_{\frac{-1}{2}}\right) - 0.022 \quad (21)$$

$$\widehat{MR} = 8.699\left(\mathcal{RRR}_{\frac{-1}{2}}\right) - 0.6273 \quad (22)$$

$$\widehat{MW} = 25.25\left(\mathcal{RRR}_{\frac{-1}{2}}\right) + 3.143 \quad (23)$$

We use Eq. (12) to compute the predicted boiling point  $\widehat{BP}$  for each of the benzenoid hydrocarbon molecules, refer Table 2. Among the linear models, we have listed the best predictive models for the remaining physical properties in Eqs. (13) to (23).

We obtain the predicted values for all properties using the proposed linear models and we summarize the results in Table 3. From Table 3 we see that the average residual error of BP is 5.3008% with  $R^2$  value 0.9587.

The hypotheses which we considered for checking the goodness of fit of the regression models are:

$H_0$  : Proposed regression model is a good fit.

$H_1$  : Proposed regression model is not a good fit.

S. No.	$\mathcal{RRR}_{-1}$	BP	$\widehat{BP}$	RES	$RES^2$	$RES^2/BP$	RES %
1	1.5	80.1	118.24	38.14	1454.66	18.16054	47.61548
2	2.2778	218	218.5762	0.5762	0.332006	0.001523	0.264312
3	3.0556	340	318.9124	21.0876	444.6869	1.307903	6.202235
4	3.0833	340	322.4857	17.5143	306.7507	0.902208	5.151265
5	3.8333	432.02	419.2357	12.7843	163.4383	0.378312	2.959192
6	3.8889	436.7	426.4081	10.2919	105.9232	0.242554	2.356744
7	3.8611	438	422.8219	15.1781	230.3747	0.52597	3.46532
8	3.8889	448	426.4081	21.5919	466.2101	1.040648	4.819621
9	3.9167	438	429.9943	8.0057	64.09123	0.146327	1.827785
10	3.8889	404	426.4081	22.4081	502.1229	1.242879	5.546559
11	4.6111	524.6	519.5719	5.0281	25.28179	0.048193	0.958464
12	4.6389	547.5	523.1581	24.3419	592.5281	1.082243	4.446009
13	4.6667	524	526.7443	2.7443	7.531182	0.014372	0.523721
14	4.6667	524.7	526.7443	2.0443	4.179162	0.007965	0.389613
15	4.6389	547.5	523.1581	24.3419	592.5281	1.082243	4.446009
16	4.7222	525	533.9038	8.9038	79.27765	0.151005	1.695962
17	4.6944	524.7	530.3176	5.6176	31.55743	0.060144	1.070631
18	4.6944	520	530.3176	10.3176	106.4529	0.204717	1.984154
19	4.6667	524.7	526.7443	2.0443	4.179162	0.007965	0.389613
20	4.6944	518	530.3176	12.3176	151.7233	0.292902	2.377915
21	4.2222	400	469.4038	69.4038	4816.887	12.04222	17.35095
22	3.9444	467.5	433.5676	33.9324	1151.408	2.462904	7.258267
23	4.1944	495	465.8176	29.1824	851.6125	1.720429	5.895434
24	4.5278	500	508.8262	8.8262	77.90181	0.155804	1.76524
25	5.3889	604	619.9081	15.9081	253.0676	0.418986	2.633791
26	4.8333	525	548.2357	23.2357	539.8978	1.028377	4.425848

**Table 2.** Predicted boiling points for 26 benzenoid hydrocarbon molecules.

Property	Model number	$R^2$	$\chi^2$	Avg. residual %	RMSE
<i>BP</i>	(12)	0.9587	44.7293	5.3008	23.2957
<i>Log P</i>	(13)	0.9637	0.18095	1.9591	0.2074
<i>CT</i>	(14)	0.9608	20.3819	2.3567	18.9922
<i>GE</i>	(15)	0.9986	1.3062	0.7285	5.8237
<i>PI</i>	(16)	0.9976	0.0855	0.8416	0.3184
<i>HL</i>	(17)	0.9995	0.0035	0.2915	0.0256
<i>CP</i>	(18)	0.9539	1.2389	3.1242	1.2647
<i>CV</i>	(19)	0.9967	2.6089	0.7031	9.3284
<i>EM</i>	(20)	0.9979	0.6239	0.4748	2.5671
<i>HAC</i>	(21)	0.9979	0.0505	0.4404	0.2046
<i>MR</i>	(22)	0.9978	0.2287	0.4379	0.9010
<i>MW</i>	(23)	0.9979	5.0938	0.4796	2.5666

**Table 3.** Statistical summary of linear models for 26 benzenoid hydrocarbons.

The table value for goodness of fit ( $\chi^2$ ) with d.o.f. 25 is 37.652 for 5% level of significance.

From Table 3, we observe that the calculated  $\chi^2$  values for all the properties except *BP* are lesser than the table value with d.o.f. 25 for 5% level of significance. Thus we accept all the models except model (12).

Though the  $R^2$  values are high in the linear models, some of the models have high residual error and thus we explore the predictive capability of quadratic regression models.

### Quadratic regression models

Using Eq. (11), we obtained the quadratic regression models for the physico-chemical properties vis-a-vis each of the proposed indices. We observe that

- $\mathcal{RR}_{-1}$  is best suited for predicting Boiling Point (*BP*) and *Log P* with corresponding  $R^2$  values 0.9673 and 0.9651, respectively.
- $\mathcal{RR}_{\frac{-1}{2}}$  is best suited for predicting Critical Pressure (*CP*), Critical Volume (*CV*), Exact Mass (*EM*), Heavy Atom Count (*HAC*), Molar Refraction (*MR*) and Molecular Weight (*MW*) with corresponding  $R^2$  values 0.9916, 0.9968, 0.9980, 0.9979, 0.9980 and 0.9980, respectively.
- $\mathcal{RRHM}_1$  is best suited for predicting (*GE*) with  $R^2$  value 0.9988.
- $\mathcal{RRGA}$  is best suited for predicting Critical Temperature (*CT*) with  $R^2$  value 0.9918.
- $\mathcal{RRABC}$  is best suited for predicting (*PI*) with  $R^2$  value 0.9978.
- $\mathcal{RRF}$  is best suited for predicting Henry's Law (*HL*) with  $R^2$  value 0.9996.

Among the quadratic models, we listed the best predictive models from Eqs. (24) to (35) and the quadratic regression models for Boiling Point (*BP*), Critical Pressure (*CP*), Critical Temperature (*CT*), Critical Volume (*CV*), Exact Mass, Heavy Atom Count, *HL*, *GE*, *Log P*, *MR*, *PI* and Molecular Weight (*MW*) are:

$$\widehat{BP} = -10.47(\mathcal{RR}_{-1})^2 + 202.4(\mathcal{RR}_{-1}) - 192.6 \quad (24)$$

$$\widehat{LogP} = -0.0407(\mathcal{RR}_{\frac{-1}{2}})^2 + 1.513(\mathcal{RR}_{\frac{-1}{2}}) - 0.1598 \quad (25)$$

$$\widehat{CT} = -0.3939(\mathcal{RRGA})^2 + 31.57(\mathcal{RRGA}) - 166.2 \quad (26)$$

$$\widehat{GE} = 0.000047(\mathcal{RRHM}_1)^2 + 0.9208(\mathcal{RRHM}_1) + 35.4 \quad (27)$$

$$\widehat{PI} = -0.004891(\mathcal{RRABC})^2 + 1.761(\mathcal{RRABC}) + 0.4245 \quad (28)$$

$$\widehat{HL} = 8.815 * (10^{-8})(\mathcal{RRF})^2 + 0.01444(\mathcal{RRF}) - 0.03427 \quad (29)$$

$$\widehat{CP} = 0.1774\left(\mathcal{RR}_{\frac{-1}{2}}\right)^2 - 5.46\left(\mathcal{RR}_{\frac{-1}{2}}\right) + 62.12 \quad (30)$$

$$\widehat{CV} = -0.2004\left(\mathcal{RR}_{\frac{-1}{2}}\right)^2 + 76.46\left(\mathcal{RR}_{\frac{-1}{2}}\right) + 35.08 \quad (31)$$

$$\widehat{EM} = -0.05377\left(\mathcal{RRR}_{\frac{-1}{2}}\right)^2 + 26.1\left(\mathcal{RRR}_{\frac{-1}{2}}\right) - 0.06767 \quad (32)$$

$$\widehat{HAC} = -0.003732\left(\mathcal{RRR}_{\frac{-1}{2}}\right)^2 + 2.084\left(\mathcal{RRR}_{\frac{-1}{2}}\right) - 0.2424 \quad (33)$$

$$\widehat{MR} = -0.0148\left(\mathcal{RRR}_{\frac{-1}{2}}\right)^2 + 8.939\left(\mathcal{RRR}_{\frac{-1}{2}}\right) - 1.503 \quad (34)$$

$$\widehat{MW} = -0.0085\left(\mathcal{RRR}_{\frac{-1}{2}}\right)^2 + 3.042\left(\mathcal{RRR}_{\frac{-1}{2}}\right) - 1.157 \quad (35)$$

We use Eq. (24) to compute the predicted boiling point  $\widehat{BP}$  for each of the every benzenoid hydrocarbon molecules, refer Table 4. From Table 5 we see that the average error is 3.0962% with the  $R^2$  value is 0.9673. Similarly, we predict all properties with the help of the quadratic models which we proposed and we summarize the result in Table 5.

The table value for the goodness of fit ( $\chi^2$ ) with d.o.f. 25 is 37.652 for 5% level of significance. The same hypothesis which we considered in linear models apply for the quadratic models too.

From Table 5, we see that the calculated  $\chi^2$  value of boiling point is less than the table value with d.o.f. 25 in 5% level of significances and thus we accept all the models.

The quadratic models (second order) obtained by us are a good fit for predicting all the properties. Hence we stop with the quadratic regression models.

Figures 1, 2, 3, 4, 5, 6 show the plots of the quadratic regression models that give the best predicted values for the properties.

### Comparison of our indices against existing indices

In this section, we present the comparison of our proposed indices against 16 well known degree based indices.

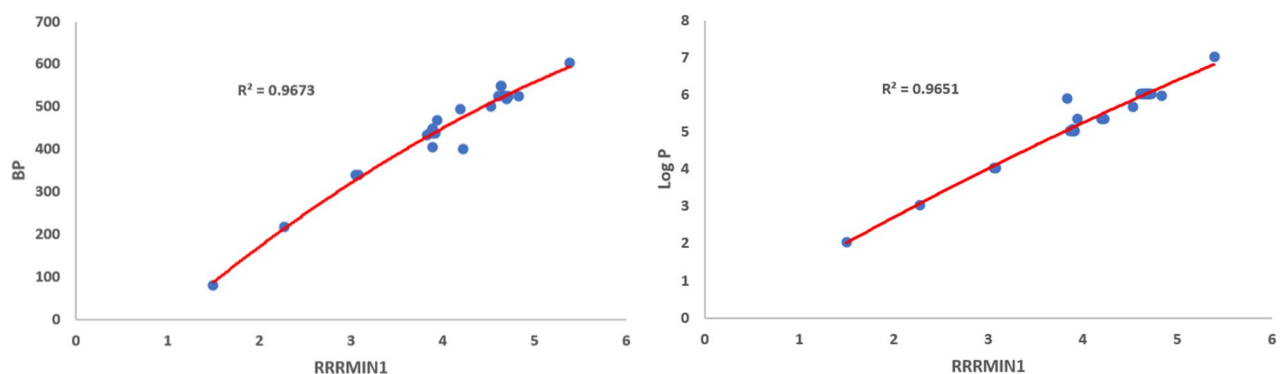
The degree based indices that we have taken into consideration include Zagreb ( $M_1, M_2$ )<sup>22,23</sup> and Hyper Zagreb ( $HM_1, HM_2$ )<sup>24,25</sup>, Forgotten index  $F$ <sup>26</sup>, Atom Bond Connectivity Index  $ABC$ <sup>27</sup>, Randić  $R$ <sup>28</sup>, Geometric Arithmetic

S. No.	$\mathcal{RRR}_{-1}$	BP	$\widehat{BP}$	RES	RES <sup>2</sup>	RES <sup>2</sup> /BP	RES %
1	1.5	80.1	87.4425	7.3425	53.91231	0.673063	9.166667
2	2.2778	218	214.1045	3.895544	15.17526	0.069611	1.786947
3	3.0556	340	328.0983	11.90172	141.6509	0.41662	3.500505
4	3.0833	340	331.9244	8.075636	65.2159	0.191811	2.375187
5	3.8333	432.02	429.4118	2.608238	6.802904	0.015747	0.603731
6	3.8889	436.7	436.1699	0.530137	0.281046	0.000644	0.121396
7	3.8611	438	432.7989	5.201096	27.0514	0.061761	1.187465
8	3.8889	448	436.1699	11.83014	139.9522	0.312393	2.640656
9	3.9167	438	439.5246	1.524638	2.32452	0.005307	0.348091
10	3.8889	404	436.1699	32.16986	1034.9	2.561634	7.962837
11	4.6111	524.6	518.071	6.529046	42.62845	0.081259	1.244576
12	4.6389	547.5	521.0053	26.49469	701.9684	1.282134	4.839212
13	4.6667	524	523.9235	0.076511	0.005854	1.12E-05	0.014601
14	4.6667	524.7	523.9235	0.776511	0.602969	0.001149	0.147991
15	4.6389	547.5	521.0053	26.49469	701.9684	1.282134	4.839212
16	4.7222	525	529.7009	4.70094	22.09884	0.042093	0.895417
17	4.6944	524.7	526.8151	2.115072	4.473532	0.008526	0.403101
18	4.6944	520	526.8151	6.815072	46.44521	0.089318	1.310591
19	4.6667	524.7	523.9235	0.776511	0.602969	0.001149	0.147991
20	4.6944	518	526.8151	8.815072	77.7055	0.150011	1.701751
21	4.2222	400	475.3249	75.32487	5673.837	14.18459	18.83122
22	3.9444	467.5	442.8512	24.64875	607.5609	1.299596	5.27246
23	4.1944	495	472.1479	22.85206	522.2166	1.054983	4.616578
24	4.5278	500	509.1815	9.181534	84.30057	0.168601	1.836307
25	5.3889	604	594.062	9.937986	98.76357	0.163516	1.645362
26	4.8333	525	541.0725	16.07246	258.324	0.492046	3.061421

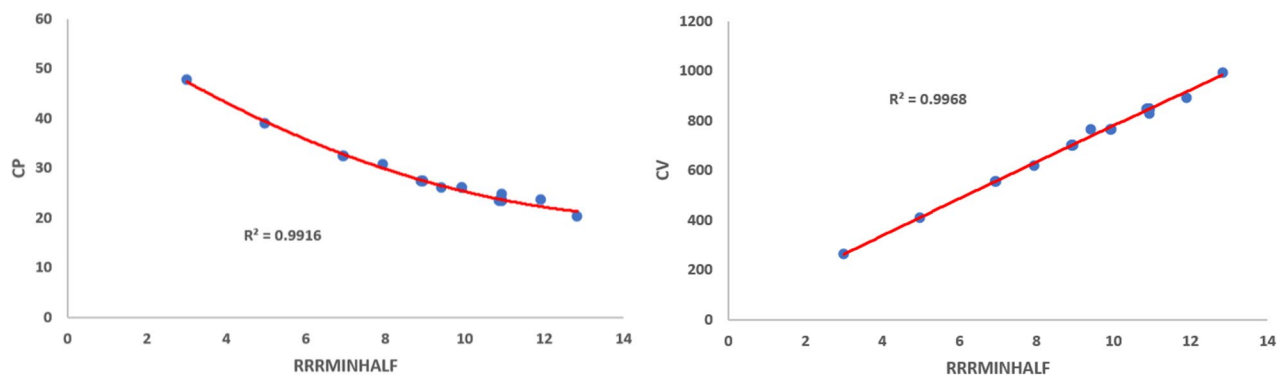
**Table 4.** Predicted boiling points for 26 benzenoid hydrocarbon molecules using quadratic model.

Property	Model number	$R^2$	$\chi^2$	Avg. residual %	RMSE
BP	(24)	0.9673	24.6097	3.0962	21.1934
Log P	(25)	0.9651	0.1703	1.9449	0.2077
CT	(26)	0.9874	3.3935	1.0299	10.9277
GE	(27)	0.9988	1.2211	0.6948	5.6693
PI	(28)	0.9978	0.0814	0.8564	0.3075
HL	(29)	0.9996	0.0035	0.2806	0.0261
CP	(30)	0.9916	0.2814	1.5133	0.5521
CV	(31)	0.9968	2.5356	0.6829	9.4338
EM	(32)	0.9980	0.6086	0.4829	2.5974
HAC	(33)	0.9980	0.0496	0.4394	0.2075
MR	(34)	0.9979	0.2255	0.4630	0.9150
MW	(35)	0.9980	0.6076	0.4819	2.5963

**Table 5.** Statistical summary of quadratic models for 26 benzenoid hydrocarbons.



**Figure 1.** Quadratic regression curves for  $\mathcal{RRR}_{-1}$  against BP and LogP.



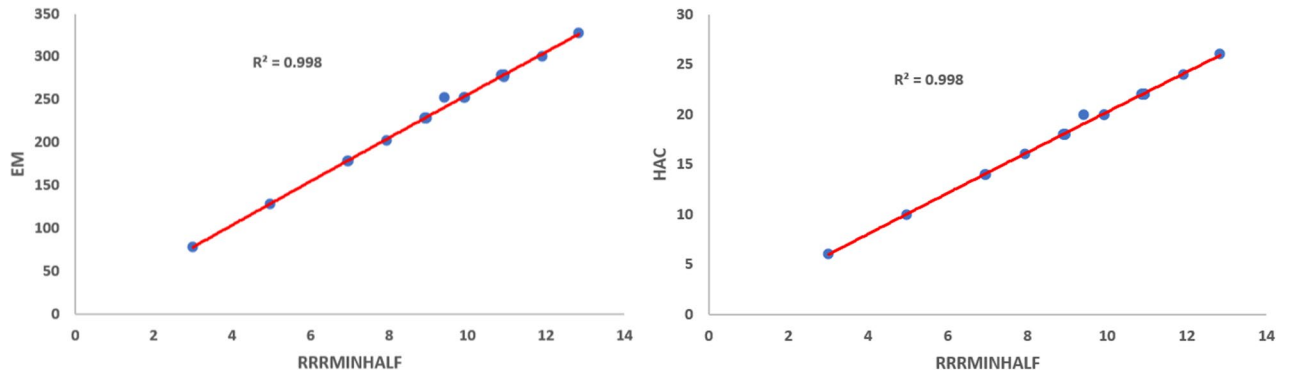
**Figure 2.** Quadratic regression curves for  $\mathcal{RRR}_{-1/2}$  against CP and CV.

index  $GA^{29}$ , Harmonic  $H^{30}$ , Randić version indices  $R_{-1}$ ,  $R_{1/2}$  and  $R_{-1/2}$ , Gourava indices<sup>31</sup> ( $GO_1$ ,  $GO_2$ ) and Hyper Gourava indices<sup>32</sup> ( $HGO_1$ ,  $HGO_2$ ).

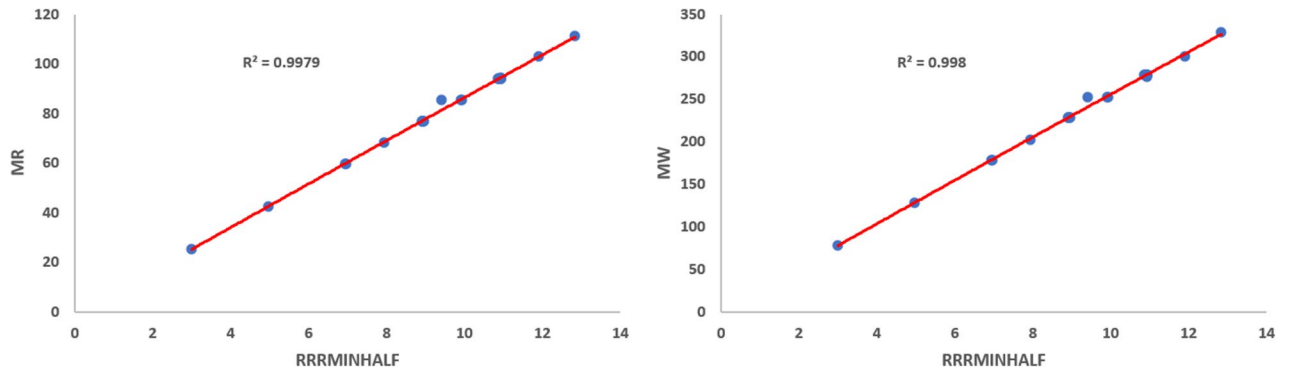
We formulated the linear regression models for the physico-chemical properties vis-a-vis these 16 degree based indices. Among these indices, we selected the best models based on the  $R^2$  value. For instance, the property Boiling Point (BP) is predicted well by the index  $R_{-1/2}$ . It has a maximum  $R^2$  value 0.8679.

In Table 6, we present the comparison of 16 degree based indices and 10 reduced reverse degree based indices against the properties of 26 benzenoid hydrocarbons along with their corresponding  $R^2$  values. Based on the  $R^2$  values corresponding to both the linear and quadratic regression models, it can be observed that the reduced reverse degree based indices are well correlated with the physico-chemical properties than the degree based indices. This implies that the reduced reverse degree-based descriptors are more suited for predicting the physico-chemical properties of benzenoid hydrocarbons compared to the existing degree based indices.

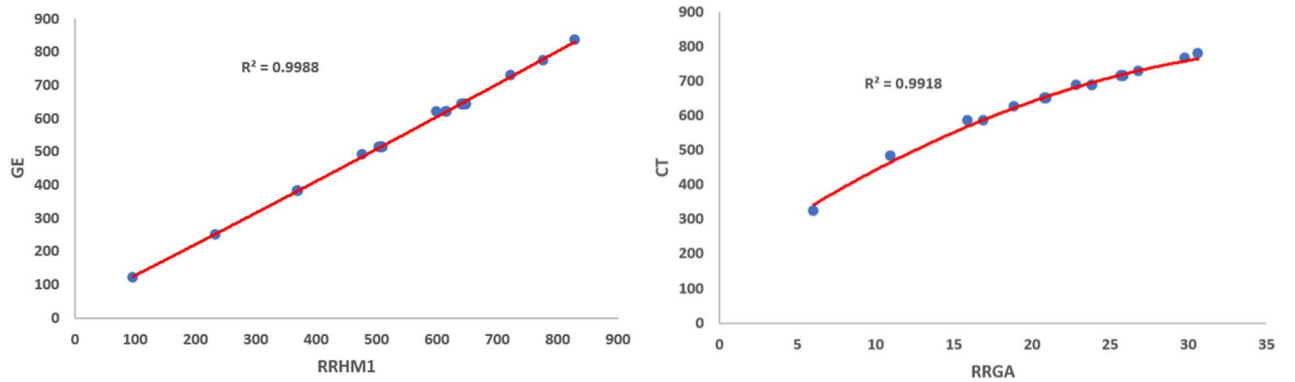




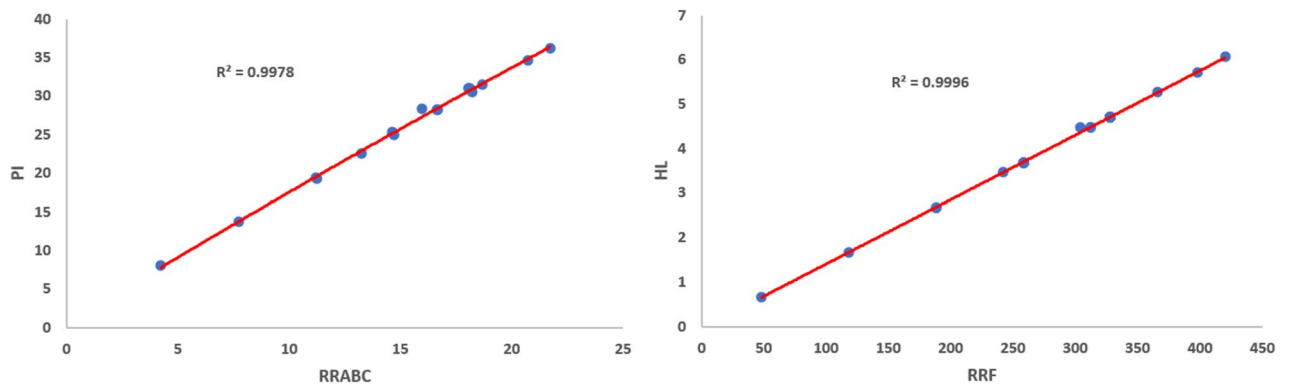
**Figure 3.** Quadratic regression curves for  $\mathcal{RRR}_{\frac{1}{2}}$  against *EM* and *HAC*.



**Figure 4.** Quadratic regression curves for  $\mathcal{RRR}_{\frac{1}{2}}$  against *MR* and *MW*.



**Figure 5.** Quadratic regression curves for  $\mathcal{RRHM}_1$  against *GE* and  $\mathcal{RRGA}$  against *CT*.



**Figure 6.** Quadratic regression curves for  $\mathcal{RRABC}$  against *PI* and  $\mathcal{RRF}$  against *HL*.

Property	Existing index	Existing index $R^2$	$\mathcal{RR}$ index/ $R^2$ (linear)	$\mathcal{RR}$ index/ $R^2$ (quadratic)
BP	$R_{\frac{-1}{2}}$	0.8679	$\mathcal{RR}_{R_{-1}} / 0.9587$	$\mathcal{RR}_{R_{-1}} / 0.9673$
Log P	$R_{\frac{-1}{2}}$	0.8629	$\mathcal{RR}_{R_{-1}} / 0.9637$	$\mathcal{RR}_{R_{-1}} / 0.9651$
CT	ABC	0.8902	$\mathcal{RR}_{GA} / 0.9608$	$\mathcal{RR}_{GA} / 0.9874$
GE	GO <sub>2</sub>	0.9401	$\mathcal{RR}_{HM_1} / 0.9986$	$\mathcal{RR}_{HM_1} / 0.9988$
PI	ABC	0.9247	$\mathcal{RR}_{ABC} / 0.9976$	$\mathcal{RR}_{ABC} / 0.9978$
HL	F	0.9392	$\mathcal{RR}_F / 0.9995$	$\mathcal{RR}_F / 0.9996$
CP	$R_{-1}$	0.8717	$\mathcal{RR}_{R_{-1}} / 0.9539$	$\mathcal{RR}_{R_{-1}} / 0.9916$
CV	$R_{\frac{-1}{2}}$	0.9095	$\mathcal{RR}_{R_{\frac{-1}{2}}} / 0.9967$	$\mathcal{RR}_{R_{\frac{-1}{2}}} / 0.9968$
EM	ABC	0.9145	$\mathcal{RR}_{R_{\frac{-1}{2}}} / 0.9979$	$\mathcal{RR}_{R_{\frac{-1}{2}}} / 0.9980$
HAC	ABC	0.9167	$\mathcal{RR}_{R_{\frac{-1}{2}}} / 0.9979$	$\mathcal{RR}_{R_{\frac{-1}{2}}} / 0.9980$
MR	ABC	0.9176	$\mathcal{RR}_{R_{\frac{-1}{2}}} / 0.9978$	$\mathcal{RR}_{R_{\frac{-1}{2}}} / 0.9979$
MW	ABC	0.9146	$\mathcal{RR}_{R_{\frac{-1}{2}}} / 0.9979$	$\mathcal{RR}_{R_{\frac{-1}{2}}} / 0.9980$

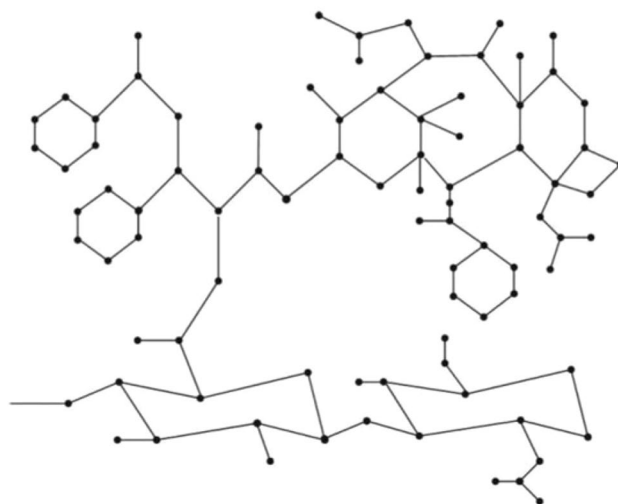
**Table 6.** Comparison of reduced reverse degree indices against 16 degree based indices.

### Application to hyaluronic acid-paclitaxel conjugates

Cancer is one of the leading factors of mortality in the world, and its fatality rate is on the rise, with breast, stomach, lung, and colon cancers responsible for the bulk of deaths. Although there have been significant advancements in cancer biology and treatments. There are still challenges in the treatment of primary and metastatic disease. Furthermore, present anticancer medications cause limited selectivity and significant toxicity, greatly restricting their usefulness. In recent years, there have been some advancements in molecularly targeted anti-cancer therapy.

Hyaluronic acid (HA) is a compound that occurs naturally. It is a glycosaminoglycan polymer composed of a linear structure of units of D-glucuronic acid and N-acetyl-D-glucosamine. These are linked via alternating  $\beta$ -1,3- and  $\beta$ -1,4-glycosidic bonds. Disaccharide, HA's primary structure is energetically stable<sup>33</sup>. Because of its unique biodegradable, biocompatible, harmless, hydrophilic, and non-immutable properties, HA is a prospective cancer treatment compound; also, HA receptors are over-expressed on numerous tumour cells. HA intends to improve antitumor therapeutics by targeting CD44-overexpressing cells, which is a rapidly developing platform nowadays<sup>34–36</sup>. HA is an excellent medication transporter and target.<sup>37</sup> Paclitaxel (PTX) is an effective medication that is prescribed for a variety of malignancies, notably ovarian, breast, lung, bladder, prostate, and esophageal tumors. Although PTX treatment has its own set of drawbacks, such as poor solubility and associated adverse effects, as well as the excipients commonly utilised in its formulation. The key benefits of HA-PTX conjugates include increased water solubility and activity retention, as well as the possibility of using it as a drug carrier to boost anti-tumor potency<sup>38–40</sup>. Figure 7 shows the structure of hyaluronic acid-paclitaxel conjugates. Numerous degree-based indices of the Hyaluronic acid (HA) conjugates<sup>41</sup> and Hyaluronic Acid-Paclitaxel conjugates<sup>42–44</sup> have been proposed by researchers.

There are still no reduced reverse degree based topological results on the molecular structures of HA-Paclitaxel conjugates. As a result of the immense pharmaceutical significance of HA-Paclitaxel conjugates, present study intends to investigate the reduced reverse degree based topological indices of the chemical structure



**Figure 7.** Molecular graph of  $(HAP)_n, n = 1$ .

of HA-Paclitaxel conjugates. Furthermore, these findings may serve as a theoretical base for pharmacological engineering.

### Methodology and results

Figure 7 depicts the structure of the HA-Paclitaxel conjugates for the values  $n = 1$ . Edge partitioning based on the reduced reverse degree sum counting is the mechanism employed here. By observing the graph structure, we get  $|V((HAP)_n)| = 87n$  and  $|E((HAP)_n)| = 96n$ .

### Results on reduced reverse degree based descriptors of hyaluronic acid-paclitaxel conjugates

Here we compute the reduced reverse degree based topological indices of  $(HAP)_n, n \geq 1$ .

Applying the reduced reverse degree based edge partitions as given in Table 7, we get

$$\begin{aligned} \mathcal{RRM}_1((HAP)_n) &= \sum_{uv \in E_1} [5+4] + \sum_{uv \in E_2} [5+3] + \sum_{uv \in E_3} [5+2] + \sum_{uv \in E_4} [4+4] \\ &\quad + \sum_{uv \in E_5} [4+3] + \sum_{uv \in E_6} [4+2] + \sum_{uv \in E_7} [3+3] + \sum_{uv \in E_8} [3+2] \\ &\quad + \sum_{uv \in E_9} [2+2] \\ &= 664n + 4. \end{aligned}$$

$$\begin{aligned} \mathcal{RRM}_2((HAP)_n) &= \sum_{uv \in E_1} [5*4] + \sum_{uv \in E_2} [5*3] + \sum_{uv \in E_3} [5*2] + \sum_{uv \in E_4} [4*4] \\ &\quad + \sum_{uv \in E_5} [4*3] + \sum_{uv \in E_6} [4*2] + \sum_{uv \in E_7} [3*3] + \sum_{uv \in E_8} [3*2] \\ &\quad + \sum_{uv \in E_9} [2*2] \\ &= 1133n + 15. \end{aligned}$$

$$\begin{aligned} \mathcal{RRHM}_1((HAP)_n) &= \sum_{uv \in E_1} [5+4]^2 + \sum_{uv \in E_2} [5+3]^2 + \sum_{uv \in E_3} [5+2]^2 \\ &\quad + \sum_{uv \in E_4} [4+4]^2 + \sum_{uv \in E_5} [4+3]^2 + \sum_{uv \in E_6} [4+2]^2 \\ &\quad + \sum_{uv \in E_7} [3+3]^2 + \sum_{uv \in E_8} [3+2]^2 + \sum_{uv \in E_9} [2+2]^2 \\ &= 4684n + 60. \end{aligned}$$

$E_i$	$(\mathcal{RRR}(u), \mathcal{RRR}(v))$	Number of edges
$E_1$	(5, 4)	$n + 1$
$E_2$	(5, 3)	$16n$
$E_3$	(5, 2)	$4n$
$E_4$	(4, 4)	$13n + 1$
$E_5$	(4, 3)	$32n - 1$
$E_6$	(4, 2)	$3n$
$E_7$	(3, 3)	$19n - 1$
$E_8$	(3, 2)	$7n$
$E_9$	(2, 2)	$n$

**Table 7.** Reduced reverse degree based edge partitions.

$$\begin{aligned} \mathcal{RRHM}_2((HAP)_n) &= \sum_{uv \in E_1} [5 * 4]^2 + \sum_{uv \in E_2} [5 * 3]^2 + \sum_{uv \in E_3} [5 * 2]^2 + \sum_{uv \in E_4} [4 * 4]^2 \\ &+ \sum_{uv \in E_5} [4 * 3]^2 + \sum_{uv \in E_6} [4 * 2]^2 + \sum_{uv \in E_7} [3 * 3]^2 + \sum_{uv \in E_8} [3 * 2]^2 \\ &+ \sum_{uv \in E_9} [2 * 2]^2 \\ &= 14335n + 431. \end{aligned}$$

$$\begin{aligned} \mathcal{RRF}((HAP)_n) &= \sum_{uv \in E_1} [5^2 + 4^2] + \sum_{uv \in E_2} [5^2 + 3^2] + \sum_{uv \in E_3} [5^2 + 2^2] \\ &+ \sum_{uv \in E_4} [4^2 + 4^2] + \sum_{uv \in E_5} [4^2 + 3^2] + \sum_{uv \in E_6} [4^2 + 2^2] \\ &+ \sum_{uv \in E_7} [3^2 + 3^2] + \sum_{uv \in E_8} [3^2 + 2^2] + \sum_{uv \in E_9} [2^2 + 2^2] \\ &= 2418n + 30. \end{aligned}$$

$$\begin{aligned} \mathcal{RRABC}((HAP)_n) &= \sum_{uv \in E_1} \left[ \sqrt{\frac{5+4-2}{5 * 4}} \right] + \sum_{uv \in E_2} \left[ \sqrt{\frac{5+3-2}{5 * 3}} \right] \\ &+ \sum_{uv \in E_3} \left[ \sqrt{\frac{5+2-2}{5 * 2}} \right] + \sum_{uv \in E_4} \left[ \sqrt{\frac{4+4-2}{4 * 4}} \right] \\ &+ \sum_{uv \in E_5} \left[ \sqrt{\frac{4+3-2}{4 * 3}} \right] + \sum_{uv \in E_6} \left[ \sqrt{\frac{4+2-2}{4 * 2}} \right] \\ &+ \sum_{uv \in E_7} \left[ \sqrt{\frac{3+3-2}{3 * 3}} \right] + \sum_{uv \in E_8} \left[ \sqrt{\frac{3+2-2}{3 * 2}} \right] + \sum_{uv \in E_9} \left[ \sqrt{\frac{2+2-2}{2 * 2}} \right] \\ &= 62.6009177191n - 0.1082. \end{aligned}$$

$$\begin{aligned} \mathcal{RRGA}((HAP)_n) &= \sum_{uv \in E_1} \left[ \frac{2 * \sqrt{5 * 4}}{5 + 4} \right] + \sum_{uv \in E_2} \left[ \frac{2 * \sqrt{5 * 3}}{5 + 3} \right] + \sum_{uv \in E_3} \left[ \frac{2 * \sqrt{5 * 2}}{5 + 2} \right] \\ &+ \sum_{uv \in E_4} \left[ \frac{2 * \sqrt{4 * 4}}{4 + 4} \right] + \sum_{uv \in E_5} \left[ \frac{2 * \sqrt{4 * 3}}{4 + 3} \right] + \sum_{uv \in E_6} \left[ \frac{2 * \sqrt{4 * 2}}{4 + 2} \right] \\ &+ \sum_{uv \in E_7} \left[ \frac{2 * \sqrt{3 * 3}}{3 + 3} \right] + \sum_{uv \in E_8} \left[ \frac{2 * \sqrt{3 * 2}}{3 + 2} \right] + \sum_{uv \in E_9} \left[ \frac{2 * \sqrt{2 * 2}}{2 + 2} \right] \\ &= 93.9919881117n + 1.9836. \end{aligned}$$

$$\begin{aligned} \mathcal{RRR}_\alpha((HAP)_n) &= \sum_{uv \in E_1} [5 * 4]^\alpha + \sum_{uv \in E_2} [5 * 3]^\alpha + \sum_{uv \in E_3} [5 * 2]^\alpha + \sum_{uv \in E_4} [4 * 4]^\alpha \\ &+ \sum_{uv \in E_5} [4 * 3]^\alpha + \sum_{uv \in E_6} [4 * 2]^\alpha + \sum_{uv \in E_7} [3 * 3]^\alpha + \sum_{uv \in E_8} [3 * 2]^\alpha \\ &+ \sum_{uv \in E_9} [2 * 2]^\alpha. \end{aligned}$$

In the above  $\mathcal{RRR}_\alpha((HAP)_n)$ ,

- If  $\alpha = \frac{1}{2}$ , then  $\mathcal{RRR}_\alpha((HAP)_n) = 326.571942n + 2.00803434$ .
- If  $\alpha = -1$ , then  $\mathcal{RRR}_\alpha((HAP)_n) = 8.8986n - 0.0819$ .
- If  $\alpha = \frac{-1}{2}$ , then  $\mathcal{RRR}_\alpha((HAP)_n) = 26.1823n - 0.0075$ .

## Conclusion

In this article, novel topological descriptors  $\mathcal{RRM}_1$ ,  $\mathcal{RRM}_2$ ,  $\mathcal{RRHM}_1$ ,  $\mathcal{RRHM}_2$ ,  $\mathcal{RRF}$ ,  $\mathcal{RRABC}$ ,  $\mathcal{RRGA}$  and  $\mathcal{RRR}_n$  have been analyzed with respect to 26 Benzenoid Hydrocarbons. From the QSPR analysis, it is evident that these descriptors are useful molecular descriptors. We tested the predictive capability of the indices with respect to 26 Benzenoid Hydrocarbons.

QSPR study using curvilinear models reveals that both linear and quadratic regression models provide good estimates for the physico-chemical properties of the 26 Benzenoid Hydrocarbons. From quadratic regression models, we observe that our proposed indices have high correlation with all the physico-chemical properties considered in the above sections.

The best models that predict the physico-chemical properties are as follows:

- 1  $\mathcal{RRR}_{-1}$  is best suited for predicting the properties, *BP*, *LogP*.
- 2  $\mathcal{RRR}_{-1}$  is best suited for predicting the properties, *CP*, *CV*, *EM*, *HAC*, *MR* and *MW*.
- 3  $\mathcal{RRHM}_1$  is best suited for predicting *GE*.
- 4  $\mathcal{RRGA}$  is best suited for predicting *CT*.
- 5  $\mathcal{RRABC}$  is best suited for predicting *PI*.
- 6  $\mathcal{RRF}$  is best suited for predicting *HL*.

On comparing the proposed reduced reverse degree-based indices with the existing degree-based topological indices, we found that our proposed indices are well correlated with all the physico-chemical properties of 26 benzenoid hydrocarbons. The errors are significantly reduced for our proposed indices. This demonstrates the significance of the reduced reverse degree-based descriptors in predicting the physico-chemical properties of benzenoid hydrocarbons over the existing degree-based descriptors.

The defined reduced reverse degree based topological indices have been determined for Hyaluronic Acid-Paclitaxel Conjugates ( $HAP$ )<sub>n</sub>,  $n \geq 1$ .

In future, these indices can be applied to various transformations of graphs and to analyze different chemical networks.

## Data availability

All data generated or analysed during this study are included in this published article. The experimental data for benzenoid hydrocarbons were taken from <https://pubchem.ncbi.nlm.nih.gov>. MS-Excel 2019 was used for statistical analyses.

Received: 2 July 2022; Accepted: 18 January 2023

Published online: 24 February 2023

## References

1. Todeschini, R. & Consonni, V. *Handbook of Molecular Descriptors* (Wiley, New York, 2008).
2. Azam, F. Ginger components as anti-alzheimer drugs: focus on drug design. *Neuroprotect. Effect. Phytochem. Neurol. Disorders*, **149** (2017).
3. Azam, F., Mohamed, N. & Alhussen, F. Molecular interaction studies of green tea catechins as multitarget drug candidates for the treatment of parkinson's disease: computational and structural insights. *Netw. Comput. Neural Syst.* **26**(3–4), 97–115 (2015).
4. Zanni, R., Galvez-Llompert, M., Garcia-Domenech, R. & Galvez, J. What place does molecular topology have in today's drug discovery?. *Expert Opin. Drug Discov.* **15**(10), 1133–1144 (2020).
5. Lavecchia, A. Machine-learning approaches in drug discovery: methods and applications. *Drug Discovery Today* **20**(3), 318–331 (2015).
6. Gertrudes, J. C. *et al.* Machine learning techniques and drug design. *Curr. Med. Chem.* **19**(25), 4289–4297 (2012).
7. Lo, Y.-C., Rensi, S. E., Tornig, W. & Altman, R. B. Machine learning in chemoinformatics and drug discovery. *Drug Discovery Today* **23**(8), 1538–1546 (2018).
8. Zhang, L., Tan, J., Han, D. & Zhu, H. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug Discovery Today* **22**(11), 1680–1685 (2017).
9. Speck-Planche, A. Recent advances in fragment-based computational drug design: Tackling simultaneous targets/biological effects. *Fut. Med. Chem.* **10**(17), 2021–2024 (2018).
10. Mauri, A., Consonni, V. & Todeschini, R. *Handbook of computational chemistry*. ed J Ileszczynski. A Kaczmarek-Kedziera, T Puzyn, MG Papadopoulos, H Reis and MK Shukla (Cham: Springer Int. Publishing) Molecular descriptors, 2065–93 (2017).
11. Estrada, E., Patlewicz, G. & Uriarte, E. From molecular graphs to drugs. a review on the use of topological indices in drug design and discovery (2003).
12. Galvez, J., Galvez-Llompert, M., Zanni, R. & Garcia-Domenech, R. Molecular topology-dissimilar similarities. *Drug Discov. Today Technol.* **10**(4), 475–481 (2013).
13. Katritzky, A. R., Lobanov, V. S. & Karelson, M. Qspr: the correlation and quantitative prediction of chemical and physical properties from structure. *Chem. Soc. Rev.* **24**(4), 279–287 (1995).
14. Katritzky, A. R., Mu, L., Lobanov, V. S. & Karelson, M. Correlation of boiling points with molecular structure. 1. A training set of 298 diverse organics and a test set of 9 simple inorganics. *J. Phys. Chem.* **100**(24), 10400–10407 (1996).
15. Doley, A., Buragohain, J. & Bharali, A. Inverse sum indeg status index of graphs and its applications to octane isomers and benzenoid hydrocarbons. *Chemom. Intell. Lab. Syst.* **203**, 104059 (2020).
16. West, D. B. *et al.* *Introduction to Graph Theory* Vol. 2 (Prentice Hall, Upper Saddle River, 2001).
17. Lučić, B., Trinajstić, N. & Zhou, B. Comparison between the sum-connectivity index and product-connectivity index for benzenoid hydrocarbons. *Chem. Phys. Lett.* **475**(1–3), 146–148 (2009).
18. Randić, M. Quantitative structure-property relationship. Boiling points of planar benzenoids. *New J. Chem.* **20**(10), 1001–1009 (1996).
19. Dias, J.R. *Handbook of polycyclic hydrocarbons*. Part a: Benzenoid hydrocarbons (1987).
20. Basak, S.C., Grunwald, G.D. & Niemi, G.J. Use of graph-theoretic and geometrical molecular descriptors in structure-activity relationships. In: *From Chemical Topology to Three-dimensional Geometry*, pp. 73–116. Springer, Berlin (2002).

21. Ravi, V., Siddiqui, M. K., Chidambaram, N. & Desikan, K. On topological descriptors and curvilinear regression analysis of antiviral drugs used in covid-19 treatment. *Polycyclic Aromatic Compound*. 1–14 (2021).
22. Gutman, I., Ružić-Ičić, B., Trinajstić, N. & Wilcox Jr, C.F. Graph theory and molecular orbitals. xii. acyclic polyenes. *J. Chem. Phys.* **62**(9), 3399–3405 (1975).
23. Balaban, A.T., Motoc, I., Bonchev, D. & Mekenyan, O. Topological indices for structure-activity correlations. In: *Steric Effects in Drug Design*, pp. 21–55. Springer, Berlin (1983).
24. Došlić, T. *et al.* On vertex-degree-based molecular structure descriptors. *MATCH Commun. Math. Comput. Chem* **66**(2), 613–626 (2011).
25. Došlić, T., Réti, T. & Vukičević, D. On the vertex degree indices of connected graphs. *Chem. Phys. Lett.* **512**(4–6), 283–286 (2011).
26. Furtula, B. & Gutman, I. A forgotten topological index. *J. Math. Chem.* **53**(4), 1184–1190 (2015).
27. Estrada, E., Torres, L., Rodriguez, L. & Gutman, I. An atom-bond connectivity index: Modelling the enthalpy of formation of alkanes (1998).
28. Randić, M. Characterization of molecular branching. *J. Am. Chem. Soc.* **97**(23), 6609–6615 (1975).
29. Vukičević, D. & Furtula, B. Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges. *J. Math. Chem.* **46**(4), 1369–1376 (2009).
30. Fajtlowicz, S. On conjectures of graffiti-ii. *Congr. Numer.* **60**, 187–197 (1987).
31. Kulli, V. The gourava indices and coincides of graphs. *Ann. Pure Appl. Math.* **14**(1), 33–38 (2017).
32. Kulli, V. On hyper-gourava indices and coincides. *Int. J. Math. Arch.* **8**(12), 116–120 (2017).
33. Mráček, A. *et al.* The influence of hofmeister series ions on hyaluronan swelling and viscosity. *Molecules* **13**(5), 1025–1034 (2008).
34. Arpicco, S., Milla, P., Stella, B. & Dosio, F. Hyaluronic acid conjugates as vectors for the active targeting of drugs, genes and nano-composites in cancer treatment. *Molecules* **19**(3), 3193–3230 (2014).
35. Dosio, F., Arpicco, S., Stella, B. & Fattal, E. Hyaluronic acid for anticancer drug and nucleic acid delivery. *Adv. Drug Deliv. Rev.* **97**, 204–236 (2016).
36. Zhang, X. *et al.* Cd44-targeted facile enzymatic activatable chitosan nanoparticles for efficient antitumor therapy and reversal of multidrug resistance. *Biomacromol* **19**(3), 883–895 (2018).
37. Markman, M. & Mekhail, T. M. Paclitaxel in cancer therapy. *Expert Opin. Pharmacother.* **3**(6), 755–766 (2002).
38. Leonelli, F., La Bella, A., Migneco, L. M. & Bettolo, R. M. Design, synthesis and applications of hyaluronic acid-paclitaxel bioconjugates. *Molecules* **13**(2), 360–378 (2008).
39. Xin, D., Wang, Y. & Xiang, J. The use of amino acid linkers in the conjugation of paclitaxel with hyaluronic acid as drug delivery system: Synthesis, self-assembled property, drug release, and in vitro efficiency. *Pharm. Res.* **27**(2), 380–389 (2010).
40. Chen, Y. *et al.* Conjugation of paclitaxel to c-6 hexanediamine-modified hyaluronic acid for targeted drug delivery to enhance antitumor efficacy. *Carbohydr. Polym.* **181**, 150–158 (2018).
41. Wang, J., Wang, Y., Wang, Y. & Zheng, L. Computation on the topological indices of hyaluronic acid. *J. Appl. Anal. Comput.* **10**(3), 1193–1198 (2020).
42. Zheng, L., Wang, Y. & Gao, W. Topological indices of hyaluronic acid-paclitaxel conjugates' molecular structure in cancer treatment. *Open Chem.* **17**(1), 81–87 (2019).
43. Jahanbani, A., Shao, Z. & Sheikholeslami, S. M. Calculating degree based multiplicative topological indices of hyaluronic acid-paclitaxel conjugates' molecular structure in cancer treatment. *J. Biomol. Struct. Dyn.* **39**(14), 5304–5313 (2021).
44. Rauf, A., Naeem, M., Maqbool, S. & Siddiqui, M. K. Irregularity indices for certain anti-tumor and anti-covid drugs. *Polycyclic Aromatic Compounds* 1–14 (2021).

## Acknowledgements

The authors would like to acknowledge the editor and reviewers for their valuable comments towards the improvement of this paper. Also, the authors sincerely thank the management of Vellore Institute of Technology, Chennai, India, for their support and encouragement to carry out this research work.

## Author contributions

V.R. was involved in conceptualization and writing-original draft preparation; K.D. was involved in supervision and validation.

## Funding

This research work is financially supported by Vellore Institute of Technology, Chennai, India.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to K.D.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023