

# کاربردهای شیمیایی نظریه گرافها

## مبانی و شاخصهای توپولوژیک

پیتر هنسن و پیتر جرس

مطالعات رابطه کمی ساختار - فعالیت<sup>۱</sup> و مطالعات رابطه کمی ساختار - خاصیت<sup>۲</sup> به عمل می‌آورد.

### مبانی نظریه گرافها

از نظر ریاضی، یک گراف، زوجی از مجموعه‌هاست: (۱) مجموعه‌ای از عناصر، و (۲) مجموعه‌ای از زوجهای این عناصر. اما، تجسم یک گراف با استفاده از نمایش تصویری آن، یعنی، (۱) مجموعه‌ای از نقطه‌ها، و (۲) مجموعه‌ای از خطها که برخی یا تمام زوج نقطه‌ها را بهم می‌پیوندند، بسیار آسانتر است. (باید یادآوری کرد که هنوز در نظریه گرافها نامگذاری استاندارد نشده است؛ از نقطه یا نام رأس، گره، و اتصال، و از خط یا نام پال، کمان، و شاخه نیز یاد می‌شود.)

بسیاری از اصطلاحات و مفاهیمی که در نظریه گراف به کار برده می‌شوند همتایی هم در واژه‌نامه شیمی یا فرد عادی دارند. در اینجا به توصیف اساسی‌ترین آنها می‌پردازیم، و سایر مفاهیم را بعداً بر حسب نیاز معرفی می‌کنیم. گراف، مفهومی توپولوژیک، و نه مفهومی هندسی است، و بنابراین، طول، زاویه، و پیکربندی سه بعدی فضایی معنی ندارند.

دو نقطه که با یک خط بهم وصل می‌شوند، نقطه‌های مجاور و دو خط دارای یک نقطه مشترک، خطهای مجاور نامیده می‌شوند. درجه یک نقطه، شمار خطهای متصل به آن است. راه، دنباله‌ای متناوب از نقطه و خط است که از نقطه‌ای آغاز و به نقطه‌ای ختم می‌شود. مسیر، راهی است که در آن، هیچ نقطه‌ای بیش از یک بار ظاهر نمی‌شود. در یک گراف همبسته، هر زوجی از نقطه‌ها با مسیری بهم وصل می‌شوند. یک گراف دودی یا حلقه‌ای، گرافی است شامل دست کم یک راه متشکل از سه نقطه یا بیشتر که نقطه آغاز و نقطه پایانش یکی باشند، هیچ نقطه‌ای (بجز نقطه اول و آخر) بیش از

نظریه گرافها، زیر شاخه‌ای از ریاضیات است که هم با توپولوژی و هم با ترکیبیات ارتباط نزدیک دارد. دانشجویان دوره کارشناسی ریاضی احتمالاً نخستین بار در درس ریاضیات گسسته با نظریه گرافها آشنا می‌شوند. خاستگاه نظریه گرافها به حدود ۲۰۰ سال پیش و به کار دانشورانی از قلمروهای مختلف می‌رسد. اولر<sup>۱</sup>، به خاطر انتشار حل مسئله پلهای کونیگسبرگ<sup>۲</sup>، معمولاً به عنوان پدر نظریه گرافها به حساب می‌آید. اثر کیرشهف<sup>۳</sup> درباره مدارهای الکتریکی و اثر کیلسی<sup>۴</sup> در زمینه شمارش ایزومرهای آلی، نیز به مثابه کشف مستقل نظریه گرافها به حساب می‌آیند.

همان طور که از ملاحظات بالا برمی‌آید، نظریه گرافها، به هیچ وجه، رشته تازه‌ای از ریاضیات یا قلمروی فاقد کاربرد در سایر زمینه‌ها نیست. نظریه گرافها، افزون بر فیزیک و شیمی، در مردم-شناسی، معماری، مهندسی راه و ساختمان، ارتباطات، علوم کامپیوتر، اقتصاد، مهندسی برق، ژنتیک، جغرافیا، مدیریت صنعتی، زبان‌شناسی، تحقیق در عملیات، علوم سیاسی، روان‌شناسی، و جامعه‌شناسی نیز به کار گرفته شده است. در قلمرو شیمی، نظریه گرافها در مسائل مربوط به گستره وسیعی از زمینه‌های تحقیقاتی اعمال شده است: شیمی سنتزی، شیمی بسپار، شیمی کوانتوم، شیمی آلی - فلزی، شیمی نفت، گرماشیمی، سینتیک شیمیایی، مکانیک آماری، تعادل‌های فازی، تحلیل طیف‌بینی، نظریه هوکل، انبارش و بازخوانی اطلاعات شیمیایی. به رغم داشتن کاربردی چنین گسترده هنوز هم بسیاری از شیمیدانان با واژگان و مفاهیم بنیادی نظریه گرافها، هیچ‌گونه آشنایی ندارند.

چندین مقاله مروری و تک‌نگاری، که منحصر آ به کاربرد نظریه گرافها در شیمی اختصاص دارند، منتشر شده است. این نوشتار، خواننده را با برخی از اصطلاحات و مفاهیم بنیادی نظریه گرافها آشنا خواهد ساخت و بحثی مجمل درباره یکی از کاربردهای فراوان آن در شیمی، یعنی، استفاده از شاخصهای توپولوژیک در

1. QSAR = quantitative structure - activity relationship.

2. QSPR = quantitative structure - property relationship.

1. Euler 2. Königsberg Bridge 3. Kirchhoff 4. Cayley

يك بار ظاهر نشود\*. گراف همبسته نادوری یا ناحلقه‌ای را ددخت می‌نامند و فاصله بین دو نقطه، شمار خطهای کوتاهترین مسیری است که آن دو نقطه را بهم وصل می‌کند. دو گراف در صورتی همریخت‌اند که تناظری يك به يك بین نقطه‌های آنها و خطهای آنها وجود داشته باشد\*\*. گرافهای همریخت ممکن است از نظر شکل ظاهری بسیار متفاوت باشند. شکل ۱ چند گراف متفاوت را نشان می‌دهد، و شکل ۲ نمونه‌هایی از مفاهیم مختلف نظریه گرافها را به دست می‌دهد.

### ساختارهای شیمیایی به صورت گراف

گرافهایی که در شیمی به کار می‌روند، به طور کلی، به یکی از دو دسته بزرگ زیرتعلق دارند: (۱) گرافهای ساختاری، که گرافهای شیمیایی، گرافهای مولکولی، یا گرافهای سرشتی نیز خوانده می‌شوند، و (۲) گرافهای واکنشی. همان طور که از نامها برمی‌آید، يك گراف ساختاری با ساختار شیمیایی معینی در تناظر است، اما يك گراف واکنشی به مجموعه‌ای از واکنشهای شیمیایی مربوط می‌شود. در جدول ۱ فهرستی از مفاهیم نظریه گرافها و اصطلاحهای متناظر مورد استفاده در شیمی، برای هر يك از این دو نوع گراف، آورده شده است. در این نوشتار، تنها مفاهیم مربوط به گرافهای ساختاری در نظر گرفته می‌شوند. کاربرد اصلی گرافهای واکنشی در سینتیک شیمیایی و در سنتز مواد آلی به کمک کامپیوتر است. در کاربردهای سینتیکی، مکانیسم واکنش با گراف مشخص می‌شود؛ در کارهای سینتیکی، گراف مشخص کننده مکانیسم يك واکنش است هر نقطه به يك ماده واکنش پذیر، يك فراورده، یا يك حد واسط، و هر خط به يك واکنش بنیادی مربوط می‌شود. در کاربردهای مربوط به سنتز مواد آلی، نقطه نمایشگر واکنش دهنده، تکه مولکولی، حد واسط، یا مولکول آماج است؛ و خطها مراحل سنتز را نمایش می‌دهند. بالابان<sup>۱</sup>، اصطلاح گراف سینتن<sup>۲</sup> را برای نوع اخیر به کار برده است.

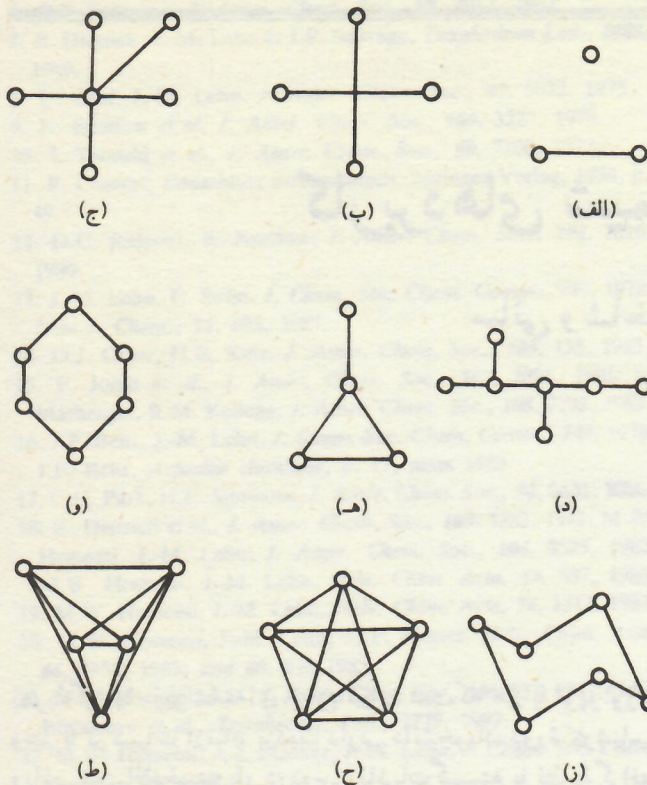
جدول ۱. مفاهیم متناظر در نظریه گرافها و شیمی

نظریه گرافها	شیمی
گراف	گراف واکنشی
نقطه	گونه‌های شیمیایی
خط	واکنش شیمیایی
مسیر	زیرساختار شیمیایی
درجه	والانس اتم
گراف دوری	ترکیب حلقه‌ای
دردخت	ساختار ناحلقه‌ای

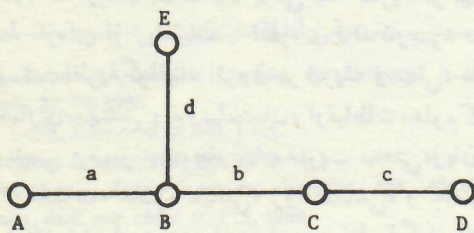
\* این تعریف نقص دارد و باید عبارت «و همه نقاط و خطوط گراف را دربر داشته باشد» به آن اضافه شود. م. ش.

\*\* این تعریف غلط است. تعریف دقیق مفهوم همریختی دو گراف بدین صورت است: دو گراف در صورتی همریخت‌اند که تابعی دوسویی بین مجموعه نقاط آنها وجود داشته باشد که مجاورت و عدم مجاورت نقطه‌ها را حفظ کند. م. ش.

1. Balaban 2. synthon graph



شکل ۱. گرافهای (الف) و (ب) ناهمبندند، اما سایر گرافها همبندند. گرافهای (ج) و (د) درخت‌اند. گرافهای (ه) تا (ط) حلقه‌ای‌اند. گرافهای (و) و (ز) همریخت‌اند. همچنین‌اند، گرافهای (ح) و (ط). گرافهای (ب) و (و) تا (ط) نمونه‌هایی از گراف منظم‌اند (تمامی نقاط آن درجه یکسان دارند). گرافهای (ح) و (ط) گرافهای کامل هستند (یعنی، تمام زوج نقطه‌ها مجاورند).



شکل ۲. زوج نقطه‌های مجاور عبارت‌اند از AB, BC, BE, CD, و زوج خطهای مجاور عبارت‌اند از ab, bd, ad, bc. نقطه‌های A, D, E درجه ۱، نقطه C درجه ۲، و نقطه B درجه ۳ دارد. نمونه‌ای از راه به صورت CbBdEdBaA است. EdBbc مسیری به طول ۲ است. نقطه‌های D و E با فاصله ۳ از هم جدا می‌شوند.

شیمی‌دانها، هنگامی که می‌خواهند درباره مواد شیمیایی و ساختار آنها اطلاعاتی منتقل کنند، از نامها و فرمولهای گوناگون استفاده می‌کنند. اما، در اغلب موارد، نامها و فرمولها معنی مستقیم، فوری، یا صریح ریاضی ندارند. از سوی دیگر، نظریه گرافها روشهای متفاوت فراوانی برای مشخص کردن عددی ساختارهای شیمیایی به دست می‌دهد. فرمول مولکولی و نام آیوپاک يك ماده شیمیایی، فرمول ساختاری آن، و گراف شیمیایی مربوط به ساختار هیدروژن زدوده آن در شکل ۳ آمده است؛ دو ماتریس ارائه شده

تشکیل ماتریس مجاورت آسان است، و ماتریس فاصله را به سهولت می توان با استفاده از یک الگوریتم بسیار کارا از ماتریس مجاورت محاسبه کرد.

مقادیر درج شده در جدول ۲ نشان می دهند که عدد وینر با افزایش اندازه مولکولی، افزایش می یابد، ولی، به طور مشهودی می توان استنتاج کرد که با افزایش شاخه ها، کاهش پیدا می کند - برای مثال، از پنج ایزومر هگزان،  $W$  ایزومر بدون شاخه برابر با ۳۵ است،  $W$  دو ایزومر با تنها یک شاخه متیل، برابر با ۳۲ و ۳۱ است، و  $W$  دو ایزومر با دو شاخه متیل برابر با ۲۹ و ۲۸ است.

جدول ۲. دمای جوش معمولی ( $^{\circ}C$ )، اعداد وینر، شاخصهای رندیج، اعداد شناسایی مولکولی برای آلکانهای  $C_6$  تا  $C_7$ .

ترکیب	دمای جوش (تجربی)	عدد وینر ( $W$ )	شاخص رندیج ( $\chi$ )	عدد شناسایی مولکولی*
۲- متیل پروپان	-۱۱۷۳	۹	۱۷۳۲	۶۷۳۲۱
۲- بوتان	-۵۵	۱۰	۱۹۱۴	۶۸۷۱۳
۲، ۲- دی متیل پروپان	۹۵	۱۶	۲۵۰۰	۸۵۰۰۰
۲- متیل بوتان	۲۷۸۵	۱۸	۲۲۷۰	۸۶۹۶۸
۲- پنتان	۳۶۰۷	۲۰	۲۴۱۴	۸۸۴۹۹
۲، ۲- دی متیل بوتان	۴۹۷۴	۲۸	۲۵۶۱	۱۰۴۶۶۰
۳، ۲- دی متیل بوتان	۵۷۹۹	۲۹	۲۶۴۳	۱۰۵۵۲۳۶
۲- متیل پنتان	۶۰۲۷	۳۲	۲۷۷۰	۱۰۶۷۹۲
۳- متیل پنتان	۶۳۲۸	۳۱	۲۸۰۸	۱۰۶۷۵۹
۲- هگزان	۶۸۷۴	۳۵	۲۹۱۴	۱۰۸۳۹۲
۲، ۲- دی متیل پنتان	۷۹۲۰	۴۶	۳۰۶۱	۱۲۴۴۹۰
۲، ۲- دی متیل پنتان	۸۰۵۰	۴۸	۳۱۲۶	۱۲۵۰۹۲
۳، ۲، ۲- تری متیل بوتان	۸۰۸۸	۴۲	۲۹۴۳	۱۲۲۹۳۱
۳، ۳- دی متیل پنتان	۸۶۰۳	۴۴	۳۱۲۱	۱۲۴۴۲۷
۳، ۲- دی متیل پنتان	۸۹۷۸	۴۶	۳۱۸۱	۱۲۵۰۵۲
۲- متیل هگزان	۹۰۰۵	۵۲	۳۲۷۰	۱۲۶۷۰۴
۳- متیل هگزان	۹۱۸۵	۵۰	۳۳۰۸	۱۲۶۶۰۰
۳- اتیل پنتان	۹۳۴۸	۴۸	۳۳۴۶	۱۲۶۶۹۲
۲- هپتان	۹۸۴۲	۵۶	۳۴۱۴	۱۲۸۳۳۸

\* برگرفته از رندیج، پس از گرد کردن.

راستی، اساس منطقی عددهای وینر چیست؟ درختهای با یک تعداد نقطه (نظیر، ایزومرهای ناحله ای) دارای شمار مساوی مسیر هستند (این نکته ممکن است از مشاهده دوتا از این درختها روشن نباشد، اما، اگر به خاطر بیاوریم که درجه ماتریس فاصله آنها یکسان است، نکته روشن خواهد شد). شاخه شاخه شدن، منجر به کاهش مسیرهای بلند و افزایش مسیرهای کوتاه می شود؛ در نتیجه مجموع تمام طول مسیرها باید کاهش یابد. پلات اظهار

کوناگونی برای سنجش اندازه - وزن مولکولی، حجم مولکولی، بلندی زنجیر، و عدد کربن - به دست آمده اند که با دماهای جوش آلکانهای نرمال همبستگی نزدیک دارند. راستی، به طور کل، چگونه می توان شکل را اندازه گرفت یا به آن کمیت نسبت داد؟ در مورد شکل ساختارهای آلی، غالباً درجه یا میزان شاخه شاخه شدن آن به نظرمان می رسد. هگزان نرمال، طبق تعریف، بی شاخه است، و بیشتر شیمی دانها در اینکه ۲، ۳-دی متیل بوتان، نسبت به ۲-متیل پنتان، شاخه های بیشتری دارد همداستانند، ولی آیا ۳-متیل پنتان شاخه های بیشتری دارد یا ۲-متیل پنتان؟ آیا می توان به شاخه شاخه شدن، کمیت نسبت داد و 'گر پاسخ مثبت است، چگونه؟ برای پاسخگویی به چنین پرسشهایی، با استفاده از چیزی که شاخصهای توپولوژیک خوانده می شود، تلاشهای بسی شماری به عمل آمده است. شاخص توپولوژیک، کمیتی عددی است که به روش ریاضی باشیوه ای مستقیم و بدون ابهام از گراف ساختاری یک مولکول به دست می آید. (برای این مفهوم اصطلاح شاخص نظریه گرافها از شاخص توپولوژیک دقیقتر است، اما اصطلاح اخیر در متون شیمی معمولتر می باشد.) چون گرافهای همریخت برای هر شاخص توپولوژیک مقادیر یکسان دارند، این شاخصها را ناوداهای گراف می نامند. شاخصهای توپولوژیک، معمولاً، نمایشگر اندازه و شکل مولکولی هستند. تریناستیج<sup>۱</sup> و همکاران، اخیراً ۳۹ شاخص توپولوژیک را که «اینک در متون شیمی کاربرد دارند» گزارش کرده اند. در بخشهایی که در پی خواهند آمد، چند مورد از جالبترین شاخصهای توپولوژیک را، همراه با مثالهای مکتوب کاربرد آنها، مورد بررسی قرار می دهیم.

### عدد وینر

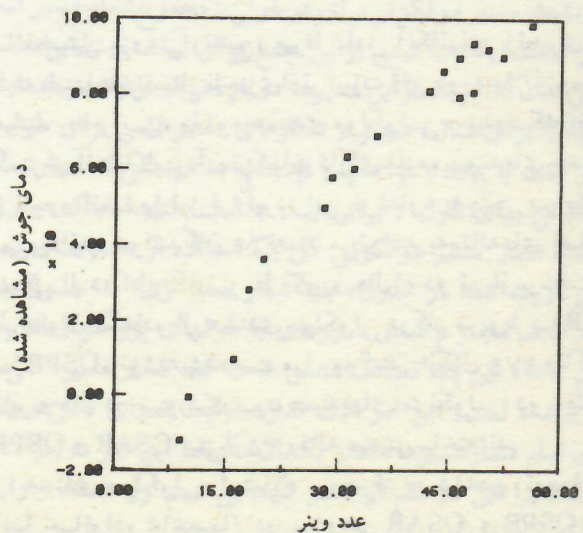
نخستین کاربرد گزارش شده یک شاخص توپولوژیک در شیمی، مربوط به مطالعات وینر درباره دماهای جوش هیدروکربنهاست. وینر از عدد قطبیت  $p$ ، که شمار تمام مسیرهای با طول ۳ است، و از عدد مسیری،  $w$ ، که «مجموع فاصله های میان هر دو اتم کربن موجود در مولکول، برحسب پیوندهای کربن - کربن می باشد»، استفاده کرده است. در اینجا، تنها به عدد مسیری که امروزه عموماً آن را عدد وینر،  $W$ ، می نامند، می پردازیم. تعریف عدد وینر، گسترش داده شده تا تمام اتمهای غیرهیدروژن را دربرگیرد؛ لذا امروزه، برای غیرهیدروکربنها نیز، مانند هیدروکربنها، به کار برده می شود. گرچه محاسبه عدد وینر برای مولکولهای کوچک، نسبتاً آسان است، شمار جملات این مجموع، تقریباً به صورت مربع شمار اتمهای غیرهیدروژن افزایش پیدا می کند، و به همین دلیل، برای محاسبه آن، معمولاً، از روشهای کامپیوتری استفاده می شود. عدد وینر، با نصف مجموع عناصر ماتریس فاصله برابر است؛ علت اینکه با نصف آن برابر است، این است که این ماتریس شامل فاصله اتم  $i$  تا اتم  $j$  و همچنین اتم  $j$  تا اتم  $i$  است.

$$\text{عدد وینر} = \frac{1}{2} \sum_{i,j} d_{i,j}$$

کرده است که ریشه سوم عدد وینر «نوعی فاصله میانگین بین اتمهای کربن در یک مولکول می باشد و تقریباً عکس انداز احتمال جذب بخشی از مولکول به بخش دیگر با نیروهای وان د والسی است».

جدول ۲ دوکاستی ظاهری عدد وینر را نشان می دهد. آلکانهای آورده شده در این جدول برحسب افزایش دمای جوش دسته بندی شده اند. اگر قرار باشد که عدد وینر به عنوان یک پیشگوی دمای جوش آلکانها به کار گرفته شود، آن گاه باید امیدوار باشیم که این اعداد به صورتی یکنواخت افزایش یا کاهش یابند. به روشنی، چنین نیست زیرا ۵ آلکان از ۱۹ آلکان نشان داده شده، خارج از روند هستند. جدول ۲ همچنین نشان می دهد که دو زوج از ایزومرها عددهای وینر یکسان دارند (۲، ۲- و ۳، ۲-دی متیل پنتان هر دو دارای عدد وینر ۴۶، و ۳- اتیل و ۴- دی متیل پنتان نیز هر دو دارای عدد وینر ۴۸ هستند). دست کم، به دو دلیل، این رخداد اندکی مزاحم است. نخست، به سهولت روشن نیست که ایزومرهای نظیر ۳- اتیل پنتان و ۴- دی متیل پنتان دارای شاخه های مساوی هستند. ثانیاً، با فرض اینکه این دو ترکیب دارای شاخه های مساوی باشند، در آن صورت آیا معقول نخواهد بود اگر انتظار داشته باشیم که هر دو ترکیب دماهای جوش نزدیک به هم داشته باشند؟ در واقع، برای گرافهای هم ریخت، داشتن شاخصهای توپولوژیک مزیت است - که در آن صورت شاخص توپولوژیک هم از خاصیت یکتایی برخوردار است و هم از خاصیت عدم ابهام.

در شکل ۵، نموداری از دماهای جوش در مقابل عددهای وینر برای آلکانهای C<sub>۷</sub> تا C<sub>۷</sub> نشان داده شده است. گرچه همبستگی بالاست (ضریب همبستگی، r = ۰.۹۷۱) پراکندگی نقطه ها چشمگیرند. به هر حال، باید به خاطر سپرد که مدل اولیه ای که توسط وینر ارائه شد شامل متغیر مستقل دیگر، یعنی عدد قطبیت، نیز بود. در مطالعات QSAR و QSPR، اغلب از مدل های چند متغیره استفاده می شود. هر چند که احتمالاً از این شکل مشهود نیست، رابطه میان این متغیرها، قطعاً غیر خطی است - مع هذا، این امر سبب کاهش کاربرد آنها نمی شود.



شکل ۵. دمای جوش معمولی (°C) مشاهده شده در مقابل عدد وینر برای آلکانهای C<sub>۷</sub> تا C<sub>۷</sub>.

عدد وینر در مطالعات مربوط به QSPR/QSAR کاربرد گسترده ای پیدا نکرده است. اما، گفته شده است که همبستگی بالای آن با شاخصهای بسیار موفق توپولوژیک، توجه بسیار ژرفتری به آن را ایجاد می کند. افزون بر تکوین مدل های پیشگویی برای دمای جوش آلکانها، وینر مدلهایی برای پیشگویی سایر خاصیت های آلکانها، از جمله: شکست مولی نور، حجم مولی، و گرمای تشکیل؛ فشار بخار به صورت تابعی از دما؛ و کشش سطحی، «پراکندگی ویژه»، و «دمای بحرانی انحلال در آنیلین» را نیز ارائه کرده است. به تازگی، ترینایستیچ، بانچف، و همکاران مدلی دو پارامتری برای پیشگویی شاخص بازداري کروماتوگرافی گازها برای مونو آلکیل و اِدتو-دی آلکیل بنزنها ساخته اند.

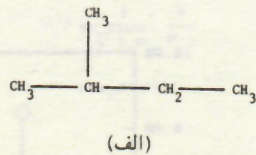
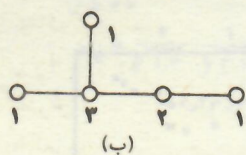
### شاخص شاخه زنی رندیچ و همبندی مولکولی

رندیچ، در سال ۱۹۷۵، شاخصی توپولوژیک پیشنهاد کرد که در مطالعات QSPR و QSAR به صورت یکی از پر مصرف ترین شاخصها درآمد. شاخص شاخه زنی رندیچ (R)، به صورت زیر تعریف می شود.

$$R = \sum_{\text{تمام پیوندها}} \left(\frac{1}{mn}\right)^{1/2}$$

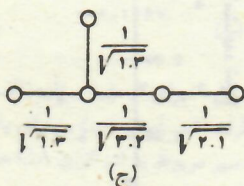
این مجموع بندی به ازای هر یال گراف ساختاری هیدروژن زدوده شامل یک جمله است، متغیرهای m و n ظرفیت نقاط مجاور می هستند که با هر یال به هم وصل می شوند. شکل ۶ دنباله گامهای رسیدن به شاخص شاخه زنی رندیچ، از فرمول ساختاری ۲- متیل بوتان، را نمایش می دهد.

با توجه به والانس ۴ که در شیمی آلی برای کربن مشاهده می شود، ظرفیت نقاط گرافهای ساختاری هیدروژن زدوده آلکانها به مقادیر ۳، ۲، ۱، ۴ و محدود می شوند. در نتیجه، شمار انواع یالها به ۱۵ زوج ممکن از این چهار والانس منحصر می شود (یعنی،



$$R = \sum X = \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}} + 0.5774 + 0.5774 + 0.4082 + 0.7071 = 2.2701$$

(د)

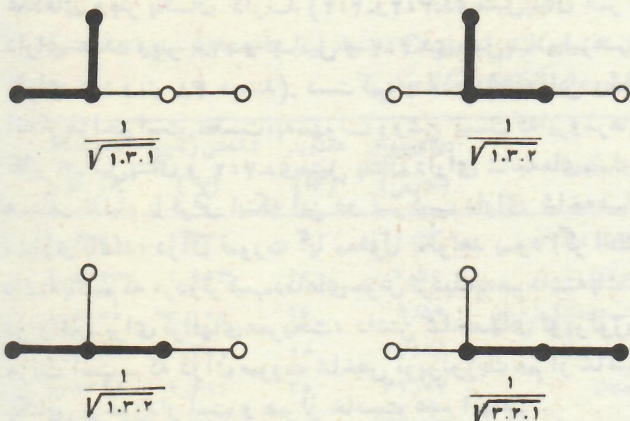


شکل ۶. شاخص شاخه زنی رندیچ (همبندی مولکولی یک - مسیر) برای ۲- متیل بوتان. (الف) فرمول ساختاری، (ب) گراف ساختاری همراه با ظرفیت هر نقطه، (ج) گراف ساختاری با  $(1/mn)^{1/2}$  برای هر یک از یالها، و (د) مجموعه  $(1/mn)^{1/2}$  که شاخص شاخه زنی رندیچ را به دست می دهد.

را عموماً، به عنوان همبندی مولکولی يك - مسیر  $\chi^1$ ، به حساب می آورند. این نامگذاری را می توان به سهولت برای نامیدن همبندیهای مولکولی گسترده نیز به کار برد؛ برای نمونه، همبندی مولکولی دو - مسیر،  $\chi^2$ ، به صورت زیر تعریف می شود،

$$\chi^2 = \sum_k \left( \frac{1}{mnp} \right)^{1/2}$$

در اینجا  $k$  مربوط است به تمام مسیرهای دارای طول  $\chi$  و  $m$  و  $n$ ، و  $p$  و  $l$  آنسهای سه نقطه واقع به این مسیرها هستند (شکل ۸ را ببینید).



$$\begin{aligned} \chi^2 &= \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}} \\ &= 0.5774 + 0.4082 + 0.4082 + 0.4082 \\ &= 1.8021 \end{aligned}$$

شکل ۸. همبندی مولکولی دو - مسیر برای ۲ - متیل بوتان. هر یک از مسیرهای دارای طول دو با خط سیاه توأم با سهم آن مسیر در  $\chi^2$  نشان داده شده است.

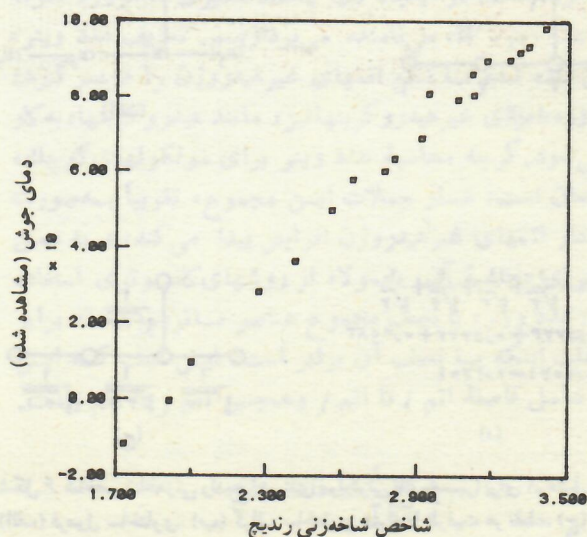
نخستین پژوهش رندیچ، صرفاً شامل آلکانهای زنجیری بود؛ البته، اینها بخش بسیار ناچیزی از ترکیبات آلی مورد علاقه شیمی دانها هستند. بنا بر این، مفهوم همبندی مولکولی، بیلرنگک تغییر و گسترش یافت تا کاربرد آن ترکیبات دارای حلقه، پیوندهای چند گانه، و ناچور اتمها را نیز فرا گیرد. این نوشتار، به چنین پیرایشهایی نمی پردازد؛ خوانندگان علاقه مند می توانند به مقاله های اصلی، یا به یکی از دو کتابی که توسط کایر و هال<sup>۱</sup>، دو تن از سرسخت ترین طرفداران استفاده از همبندی مولکولی در کار مربوط به QSAR یا QSPR، نوشته شده است مراجعه کنند. از سال ۱۹۷۵ تا کنون، کایر و هال در مورد کاربرد همبندیهای مولکولی در مطالعات QSPR و QSAR بیش از ۳۵ مقاله منتشر ساخته اند.

همبندی مولکولی، احتمالاً، پرمصرفترین شاخص توپولوژیک (یا خانواده شاخصها) در پژوهش QSAR و QSPR است.

البته یال نوع ۱-۱ تنها در اتان می تواند رخ دهد. به علاوه، چون یالهای نوع ۱-۲ و ۲-۲، هر دو، حاصل ضرب یکسان  $mn$  پدید می آورند، باید برای  $(1/mn)^{1/2}$  نیز تعدادی یکسان پدید آورند. زیرا، از لحاظ عملی، شاخص شاخه زنی رندیچ بر تجزیه یک ترکیب به هشت نوع مختلف پیوند کربن-کربن استوار است.

اگر جدول ۲ را بیازماییم می بینیم که برای ۱۹ آلکان  $C_4$  تا  $C_7$  فهرست شده، تنها تغییری در موقعیت دوتای آنها یک دنباله یکتواخت - صعودی از مقادیر شاخص رندیچ به دست می دهد. هر چند برای این ۱۹ ترکیب، چندگانگی مشاهده نمی شود، اما اگر فهرست را چنان بسط دهیم که آلکانهای با تعداد کربنهای بیشتر را شامل شود، دیگر چنین نخواهد بود (برای مثال، ۳-متیل هپتان، و ۴-متیل هپتان انواع پیوندی کاملاً یکسان به دست می دهند و در نتیجه باید مقادیر شاخص رندیچ یکسان داشته باشند). همبستگی بالای (۰.۹۹۴ =  $r$ ) دمای جوش با شاخص رندیچ از شکل ۷ مشهود است. نظیر عدد وینر، ارتباط کاری شاخص رندیچ با دمای جوش آلکانها تا حدی غیر خطی است. نشان داده شده است که برای ۲۱ آلکان  $C_4$  تا  $C_7$  استفاده از ریشه سوم شاخص رندیچ منجر به بهبود همبستگی بادمای جوش (۰.۹۹۹ =  $r$ ) و کاهش انحراف معیار به نصف می شود.

راه دیگر تلقی شاخص شاخه زنی رندیچ، این است که آن را به مثابه مجموع بندی تمام مسیرهای با طول یک گراف ساختاری به حساب آوریم. گسترش طبیعی این دیدگاه، تعریف شاخصهای بیشتری برای مسیرهای با طولهای بزرگتر از یک، و برای زیر-گرافهای دیگر (خوشه ها، مسیر - خوشه ها، و دورها) است که برخی از آنها ممکن است شامل نقطه های با درجه بالاتر از دو نیز باشند. در واقع، نخستین مقاله رندیچ به مفهوم همبندی گسترده که همسایه های دورتر را فرامی گیرد اشاره دارد. لذا اینک، در متون شیمی اصطلاح همبندی مولکولی را برای ارجاع به خانواده بزرگی از شاخصهای توپولوژیک به کار می برند. شاخص شاخه زنی رندیچ



شکل ۷. دمای جوش معمولی (°C) مشاهده شده در مقابل شاخص شاخه زنی رندیچ برای آلکانهای  $C_4$  تا  $C_7$ .

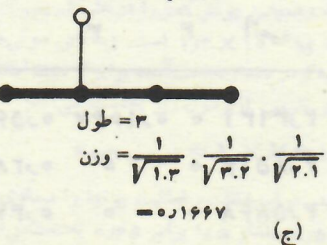
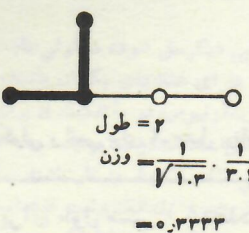
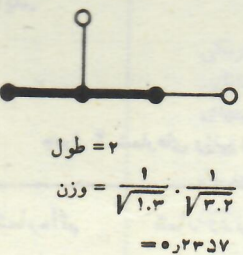
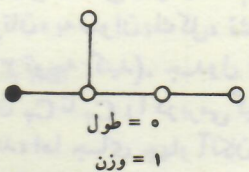
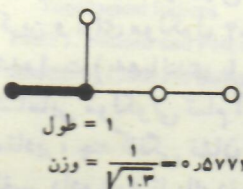
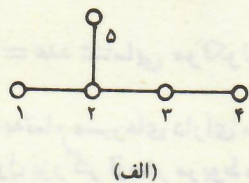
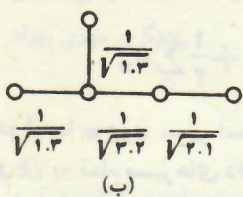
1. Kier & Hall

جدول ۳. شماره‌های مسیر اتمی و مولکولی برای ۲-متیل بوتان

شماره اتم	شمار طول مسیر				مجموع
	۰	۱	۲	۳	
۱	۱	۲	۱	۰	۵
۲	۱	۳	۱	۰	۵
۳	۱	۲	۲	۰	۵
۴	۱	۱	۱	۲	۵
۵	۱	۱	۲	۱	۵
شماره‌های ۵	۴	۴	۲	۰	

مسیر مولکولی

کل شمار مسیر مولکولی = ۱۵



شکل ۹. برای ۲-متیل بوتان: (الف) سیستم شماره گذاری اتمها که در جدول ۳ به کار گرفته شده است. (ب) وزنه‌های یالی یا پیوندی مربوط به والانسه‌های اتمها. (ج) پنج مسیر مربوط به اتم کربن C-۱ همراه با وزنها و طولهای آنها.

روند. رندیچ و ویلکینز همانندیهای کمی ساختاری مجموعه‌ای از ۲۹ مونوترپن تک حلقه‌ای را در قیاس با میزان ناهمانندی زیر مطالعه کردند.

کایر و هال، فهرستی شامل ۱۵۸ مرجع مربوط به مجلات علمی (تا سال ۱۹۸۴) ضمیمه کرده‌اند که در آنها همبندی مولکولی نقش اصلی داشته است. استفاده از همبندی مولکولی، از آن زمان به بعد، احتمالاً فزونی گرفته است برای مثال، همبندی مولکولی در مطالعات QSAR یا QSPR شاخص بازدارنده گاز کروماتوگرافی هیدرو-کربنهای آروماتیک چند حلقه‌ای نیترودار شده، انحلال پذیری آلکانها در آب، طراحی معقول داروها، هوشبرهای هالو کربن، واکنشهای آنزیمی، مشابه‌های توهم‌زای مسکالین، ضریب غلظت زیستی مواد شیمیایی خطرناک، و سرطان‌زایی شیمیایی به کار گرفته شده‌اند.

### عددهای شناسایی مولکولی

هنگام تلاش برای دستیابی به شاخصی توپولوژیک که سهل الوصول، یکتا، و از لحاظ ساختاری چشمگیر باشد، رندیچ، در سال ۱۹۸۴، عدد شناسایی مولکولی را پیشنهاد کرد. این شاخص توپولوژیک، پیوندی از همبندی مولکولی و شمار مسیر مولکولی است. به طور دقیقتر، یک شمار وزنی مسیر مولکولی است که در آن وزنها به همبندی مولکولی مربوط می‌شوند. پیش از تعریف عدد شناسایی مولکولی، بهتر است شمار مسیر اتمی و شمار مسیر مولکولی را، بازرغای بیشتری، بررسی کنیم.

در جدول ۳، شمار و مسیرهای جزئی و کلی برای تمام اتمهای کربن و برای مولکول ۲-متیل بوتان به عنوان یک کل (مربوط به گراف هیدروژن زدوده) نشان داده شده است. سیستم نامگذاری اختیار شده برای این مولکول در شکل ۹ آمده است. افزون بر این، مسیرهای وابسته و طول این مسیرها برای اتم کربن C-۱ نیز نمایش داده شده است (از سایر اطلاعات نشان داده شده، در حال حاضر می‌توان چشم پوشید). برخی از داده‌ها و رابطه‌های آورده شده در جدول ۳، ممکن است مایه سردرگمی باشند. نخست، برای اینکه از نظر ریاضی جدول کامل باشد، مسیرهای از طول صفر را شمارش کرده، و آنها را آورده‌ایم؛ این مسیرها صرفاً از یک نقطه منفرد یا اتم تشکیل می‌شوند. ثانیاً، بجز برای مسیرهای از طول صفر، شمار مسیر مولکولی برای هر طولی با مجموع شمار مسیر اتمی برابر نیست، بلکه با نیمی از این جمع برابر است. ریشه این مطلب در این است که هر یک از این مسیرها دو بار، هر بار از یک سر، شمارش شده‌اند. برای مثال، مجموع شماره‌های مسیر اتمی برای طول ۲ برابر است با  $2 + 1 + 2 + 1 + 2 = 8$ ، در حالی که شمار مسیر مولکولی برای طول ۲ برابر با ۴ است. نتیجه اینکه، برای محاسبه شمار مسیر مولکولی کل با استفاده از شماره‌های مسیر اتمی آورده شده در جدول، باید شمار مسیرهای با طول صفر را با نصف مجموع شمار مسیرهای با طولهای بزرگتر از صفر جمع کرد. ثالثاً، این یک تصادف محض است که شمار مسیر اتمی هر اتم ۵ شده است؛ این امر کلیت ندارد. رابماً اتمهای C-۱ و C-۵ برای تمام شماره‌های مقدار یکسان دارند؛ این امر دور از انتظار نیست زیرا این دو اتم از نظر شیمیایی هم ارز هستند و رابطه تقارنی دارند.

شماره‌های مسیر مولکولی (کدهای مسیر مولکولی نیز خوانده می‌شوند) خود می‌توانند در مطالعات QSAR و QSPR به کار

$$D_{ab}^1 = \sum_i (a_i - b_i)$$

در اینجا،  $a$  و  $b$  شماره‌های مسیر مولکولی برای مسیرهای به طول  $n$  از دو ساختار را نشان می‌دهند. میزان ناهمانندی،  $D$ ، تاحدودی به فاصله اقلیدسی شباهت دارد.

همان‌طور که در بالا گفته شد، عدد شناسایی مولکولی، شمار وزنی مسیر مولکولی است. چه وزنهایی به کار برده می‌شوند؟ به مسیری به طول صفر، وزن واحد داده می‌شود. برای مسیرهایی که دارای طول بزرگتر از صفر هستند، وزن برابر با حاصلضرب  $(1/mn)^{1/2}$  جمله است، یک جمله برای هر یک از یالهای مسیر؛ برای هر یک از این جملات،  $m$  و  $n$  برابر است با ظرفیت اتمهای دوسر آن یال. (رابطه نزدیک با همبندی مولکولی از اینجا ناشی می‌شود). برای مسیرهای مربوط به اتم کربن ۱-C شکل ۹ روش به کار گرفته شده برای محاسبه این وزنها را نشان می‌دهد. عدد شناسایی مولکولی دقیقاً با کل شمار مسیر مولکولی مشابه است، بجز در اینکه شمار وزنی مسیر در نظر گرفته می‌شود.

$$\text{عدد شناسایی مولکولی} = \sum_j w_{0j} + \frac{1}{4} \sum_{ik} w_{ik}$$

در اینجا  $w$  وزن مسیر است،  $0j$  به تمام مسیرهای دارای طول صفر و  $ik$  به تمام مسیرهای دارای طول بزرگتر از صفر مربوط می‌شود. در جدول ۴ شمار وزنی مسیر جزئی و کلی برای تمام اتمهای کربن و برای مولکول ۲-متیل بوتان، به عنوان یک کل، نشان داده شده است (به همانندی با جدول ۳ توجه کنید). جدول ۲ اعداد شناسایی مولکولی تمام آلکانهای  $C_4$  تا  $C_7$  را در برمی‌گیرد. این مقادیر، چندگانگی نشان نمی‌دهند، اما جای چهار آلکان بایسد تغییر داده شود تا یک دنباله یکتا وخت صعودی از اعداد به دست آید.

جدول ۴. شماره‌های وزنی مسیر مولکولی و اتمی برای ۲-متیل بوتان\*

شماره اتم	شمار وزنی مسیر برای طول مسیر					عدد شناسایی اتمی
	۰	۱	۲	۳	۴	
۱	۱	۵۷۷۴	۵۶۹۰	۱۶۶۷	۰	۲۳۱۳۱
۲	۱	۵۶۲۹	۲۸۸۷	۰	۰	۲۸۵۱۶
۳	۱	۱۱۵۴	۴۷۱۴	۰	۰	۲۵۸۶۸
۴	۱	۷۰۷۱	۲۸۸۷	۳۳۳۳	۰	۲۳۲۹۱
۵	۱	۵۷۷۴	۵۶۹۰	۱۶۶۷	۰	۲۳۱۳۱
شماره‌های = ۵ ۲۲۷۰۱ ۱۰۹۳۴ ۳۳۳۳ ۰						
وزنی مسیر مولکولی						
عدد شناسایی مولکولی = ۸۶۹۶۸۰۱۹						

\* بیشتر نتایج تا چهار رقم اعشار گرد شده‌اند.

منطق نهفته در پس عدد شناسایی مولکولی به عنوان یک شاخص توپولوژیک چیست؟ به گفته رندیچ، استفاده از وزنهایی پیوندی یا یالی کوچکتر از ۱ منتج به «تضعیف تدریجی نقش مسیرهای با طولهای بلندتر می‌شود»، بنابراین «مجموع مسیر وزنی مولکولی، نمایشگر شمار مسیر اتمی میانگینی است که در آن خصیصه‌های موضعی اهمیت بیشتری دارند.»

برای برخی افراد، این ادعا که عدد شناسایی مولکولی «سهل الوصول» اندمکن است بحث‌انگیز باشد. اما اغلب، با این نکته موافق‌اند که پیشرفت و در دسترس بودن کامپیوتر معنای عبارت «سهل الوصول» را تغییر داده است. مسلماً اعمال جبری لازم برای محاسبه اعداد شناسایی مولکولی به ریاضی پیشرفته یا پیچیده‌ای نیاز ندارد. رندیچ، ویلکینز، و همکاران بر نامهای به زبان «بیسیک» به نام ALL-PATH (شامل کمتر از ۱۰۰ سطر کد)، نوشته‌اند که می‌تواند برای محاسبه اعداد شناسایی مولکولی به شیوه‌ای نسبتاً ساده، به کار رود.

رندیچ، عدد شناسایی مولکولی بیش از ۴۰۰ ساختار (شامل تمام آلکانهای ناحلقه‌ای تا  $n=10$  و همچنین انواع تک‌حلقه‌ایها و چند حلقه‌ایها) را محاسبه کرده و چندگانگی در آنها ندیده است. او خاطر نشان ساخته است که اگر اعداد شناسایی مولکولی یکتا بودند، می‌توانستند به عنوان روشی برای کدگذاری، شاید مکملی برای اعداد ثبت سرویس کمیکال ابستراکت<sup>۱</sup> که خود هیچ اطلاعات ساختاری ندارند، به کار روند. البته، این کار مستلزم محاسبه دقیق اعداد شناسایی مولکولی است، زیرا در صورتی که این اعداد تنها به چند رقم اعشار گرد شوند، حداکثر تعداد ترکیباتی که بدون روبروشدن با چندگانگی، قابل کدگذاری هستند مسلماً، بسیار محدود خواهد بود.

در سال ۱۹۸۵، ترینایستیچ و همکاران گزارش دادند که آزمون اسلوبمند آنها برای ۶۱۸۰۰۰ ایزومر آلکانها تا  $n=7$  وجود ۱۲۲ زوج و یک سه‌تایی از ساختارهای ناهم‌ریخت با اعداد شناسایی یکسان را نشان داده است. برای مثال، ۳،۲،۲-دی‌متیل-۶-اتیل-۵-ایزوپروپیل اکتان و ۶-دی‌متیل-۵-اتیل-۳-ایزوپروپیل اکتان هر دو دارای عدد شناسایی یکسان هستند (وقتی اعداد شناسایی به صورت عبارات صحیح، بدون گرد کردن بیان شدند، باز این نکته صادق بود). پژوهش آنان نشان داد که عدد شناسایی مولکولی یکتا نیست. در نتیجه، رندیچ عدد شناسایی مولکولی دیگری را پیشنهاد کرد که در آن از ریشه دوم عکس نخستین نه عدد اول برای نوع یالی که معمولاً در گرافهای شیمیایی پیش می‌آیند، به عنوان وزن (به جای وزن بر اساس والانس اتم) استفاده می‌شود. در این نوشتار بیش از این به این تعدیل نمی‌پردازیم. ترینایستیچ همچنین خاطر نشان ساخت که «برای ساختارهای پیچیده (مثلاً پلی‌هگزنها)، محاسبه اعداد شناسایی مولکولی چندان آسان نیست.»

محاسبه اعداد شناسایی مولکولی منجر به چیزی می‌شود که ممکن است آن را اعداد شناسایی اتمی بنامیم (ستون آخر جدول ۴ را ببینید). در حالی که، اغلب شاخصهای توپولوژیک کل مولکول را

1. CAS Registry Numbers

توصیف می کنند. اعداد شناسایی اتمی براتم استوارند و امیدمی رود از آنها در پژوهشهای QSAR و QSPR، که در آنها فعالیت یا ویژگی مورد نظر موضعی است (نظیر جا به جاییهای شیمیایی  $^{13}\text{C-NMR}$  یا سرطانزایی ناشی از یک موضع مولکولی خاص)، استفاده شود. همچنین، با جمع کردن اعداد شناسایی اتمی مناسب می توان برای گروههای عاملی متشکل از چند اتم، نظیر گروه نیترو یا گروه کریوکیلی اسید، اعداد شناسایی زیر گرافها را محاسبه کرد.

تکوین اعداد شناسایی مولکولی، آن قدر تازه است که نباید انتظار داشت در مطالعات QSAR و QSPR زیادی به کار گرفته شده باشند. اما، رندیچ تحلیل خوشه ای را در مورد اعداد شناسایی مولکولی مجموعه ای از ۴۱ ماده فعال دارویی، شامل آنتی هیستامینها، مواد متوقف کننده عبور موج تحریکات عصبی از اعصاب پاراسمپاتیک، ضدافسردگیها، مسکنها، و ضدپارکینسونها، به کار برده و با توفیق نسبی آنها را رده بندی کرده است.

### نتیجه گیری

نظریه گرافها، تعدادی ابزار پژوهشی بسیار سودمند در اختیار شیمیدانها قرار داده است. تلاش ما در این نوشتار، آشنا کردن خواننده با یکی از این ابزارها، یعنی با شاخص توپولوژیک بوده است. شاخصهای توپولوژیک اطلاعات ساختاری ارزشمندی را در بر دارند که توفیق آنها را در کاربردهای گسترده در مطالعات QSAR و QSPR مشاهده کردیم. اما، شاخصهای توپولوژیک داروی تمام دردها نیستند، استفاده کنندگان باید از محدودیتها و

خطاهایی که ممکن است همراه با کاربردشان باشد، آگاه باشند. برای مثال، روشهای تجربی باید چنان انتخاب شوند که احتمال همبستگیهای شانسی را به حداقل برسانند. در نظریه گرافها خصیصه ذاتی مفهومی چون شاخص توپولوژیک چنان است که نوع معینی از اطلاعات ساختاری نمی تواند قابل کدگذاری باشند؛ برای نمونه، اطلاعات ساختاری دارای ماهیت فضایی، نظیر ایزومری سیس/ترانس یا کایرالیت، را نمی توان در مقادیر یک از سه شاخص توپولوژیک آورده شده در این نوشتار، منعکس ساخت.

جستجو برای شاخصهای توپولوژیک بهتر ادامه دارد، زیرا دلیلی وجود ندارد که معتقد باشیم شاخصهای بهتری نمی توان یافت. سه تن از پیشگامان نظریه گرافهای شیمیایی در روزگار ما، الکساندرو بالابان (انستیتوی پلسی تکنیک، بخارست، رومانی)، میلان رندیچ (یوگسلاوی الاصل، اکنون در دانشگاه دراک، دمنوس، آیووا، آمریکا) و ننادتسیرینایستیچ (انستیتوی روجر باسکوویچ، زاگرب، کروات، یوگسلاوی) هستند.

ترجمه عیسی یآوری

بازنگری مهدی بهزاد

### ● Chemical Applications of Graph Theory, Fundamentals and Topological Indices

Peter J. Hansen and Peter C. Jurs

Journal of Chemical Education, July 1988

### چه وقت يك الكتروليت قوی، قوی است؟

الكتروليت را می توان به ماده ای اطلاق کرد که اگر در آب یا هر محلول قطبی دیگر حل شود، محلولی با رسانایی الكتریکی تولید کند. هر گاه چند ماده حل شدنی حل شده باشند و محلولهای حاصل، دارای غلظتهای یکسان باشند، هر قدر رسانایی الكتریکی يك الكتروليت زیادتر باشد، آن الكتروليت قویتر است. بنا به نظریه آرنیوس، درجه تفكیک  $\alpha$  برای الكتروليتهای قوی، برابر با يك است. اکنون این سؤال پیش می آید که آیا يك ماده حل شده صد درصد یونیده، يك الكتروليت قوی نیز هست؟ این سؤال ممکن است هنگامی پیش آید که نمکهای کم محلولی مانند  $\text{AgCl}$  مدنظر باشند. ما به دانشجویان می آموزیم که  $\text{AgCl}$  موجود در محلول صد درصد یونیده است. اما آیا  $\text{AgCl}$  يك الكتروليت قوی است؟ برای پاسخ به این سؤال، باید رسانندگی الكتریکی محلولهای نقره کلرید و همچنین یونش کامل را در نظر بگیریم.

انحلال پذیری  $\text{AgCl}$  در آب که از  $K_{sp}$  محاسبه می شود،  $1.3 \times 10^{-5} \text{ M}$  است. رسانایی مولی حدی  $\text{AgCl}$  را می توان از رساناییهای یونی مولی حدی  $\text{Ag}^+$  و  $\text{Cl}^-$  محاسبه کرد. مقدار به دست آمده برابر  $138 \text{ ohm}^{-1} \text{ cm}^2 \text{ mol}^{-1}$  می باشد. رسانندگی الكتریکی محلول سیر شده  $\text{AgCl}$  به طور تقریبی از فرمول زیر به دست می آید:

$$\kappa_{\text{AgCl}} \text{ سیر شده} = \Lambda_0 \times M = 138 \times 10^{-6} \text{ ohm}^{-1} \text{ cm}^{-1}$$

محلول استیک اسید  $1.3 \times 10^{-5} \text{ M}$  را در نظر بگیرید. در این غلظت، فعالیت یونها را مساوی غلظتها در نظر می گیریم و از غلظت یونهای هیدروژن و استات،  $1.3 \times 10^{-5} \text{ M}$ ، مقدار  $\alpha$  را برابر  $0.65$  به دست می آوریم. با استفاده از رساناییهای یونی مولی حدی  $\text{H}^+$  و  $\text{CH}_3\text{CO}_2^-$ ،  $\Lambda_0 = 391 \text{ ohm}^{-1} \text{ cm}^2 \text{ mol}^{-1}$ ، به دست می آید. با استفاده از همان تقریب مربوط به  $\text{AgCl}$ ، رسانندگی محلول استیک اسید برابر  $5.1 \times 10^{-6} \text{ ohm}^{-1} \text{ cm}^{-1}$ ، در مقایسه با  $1.38 \times 10^{-6}$  برای  $\text{AgCl}$  خواهد بود. آیا استیک اسید الكترولیتی قویتر از  $\text{AgCl}$  است؟ این محاسبات نشان می دهد که وقتی درباره قدرت الكتروليتها بحث می کنیم، تأکید بر غلظت اهمیت زیادی دارد.

تعریف مناسب برای الكتروليتهای قوی وضعیف، باید شامل رسانندگی حداقل و یا غلظت و یا انحلال پذیری و همچنین درجه یونش باشد. می توان الكتروليت قوی را به عنوان ماده ای تعریف کرد که اگر در آب حل شود،  $\alpha = 1$  گردد و دارای رسانندگی حداقلی نیز باشد. یا می توانیم الكتروليت قوی را به ماده ای اطلاق کنیم که بتوان از آن محلولهای آبی  $0.001 \text{ M}$  یا رسانندگی بزرگتر از يك محلول مرجع تهیه کرد.

ترجمه زهرا حاج سید جوادی

● Journal of Chemical Education, July 1988