In search of simplification

Bertz, Steven H.; Rücker, Christoph

Published in:
Croatica Chemica Acta

Publication date:
2004

Document Version
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for pulished version (APA):
Bertz, S. H., & Rücker, C. (2004). In search of simplification: The use of topological complexity indices to guide retrosynthetic analysis. *Croatica Chemica Acta*, *77*(1-2), 221-235. https://hrcak.srce.hr/102667

In Search of Simplification: the Use of Topological Complexity Indices to Guide Retrosynthetic Analysis*, **

Steven H. Bertz[a],*** and Christoph Rücker[b],***

[a]*Complexity Study Center, Mendham, NJ 07945, USA*

[b]*Department of Mathematics, University of Bayreuth, D-95440 Bayreuth, Germany*

Key words
bicyclo[2.2.1]heptane
complexity indices
disconnections
retrosynthetic analysis
simplification principles
spiro[3.3]heptane
strategic bonds
synthetic analysis
topological indices

Topological complexity indices N_S, N_T, $N_S(lpe)$, $N_T(lpe)$, *twc* and *wcx* are used to rank the one-bond disconnections of bicyclo[2.2.1]heptane, spiro[3.3]heptane and their aza derivatives with respect to the degree of simplification they afford. Selected two-bond disconnections of bicyclo[2.2.1]heptane and its aza derivatives are also evaluated. Simplification principles are derived which are useful for guiding the retrosynthetic analysis of complex target molecules. Comparison with the LHASA rules for strategic bonds reveals a great deal of similarity and some important differences.

INTRODUCTION

Topological indices, pioneered by Wiener,[2] Trinajstić,[3] Randić,[4] and others,[5–8] have been used for diverse applications, including quantitative structure-activity relations (QSAR),[9] polymer properties,[10] and resonance energy.[11] As part of a program to provide new tools for synthetic analysis, we have been investigating the use of topological complexity indices to evaluate synthetic reactions and series of reactions, *i.e.*, synthetic routes to target molecules.[12] We have chosen to operate within the general framework of retrosynthetic analysis,[13] which is the prevailing paradigm for synthesis planning.

The first thing a synthetic chemist does when he or she looks at a target is to search for recognizable patterns (*cf.* retrons[13b]), which will suggest precursors and the reactions to assemble them. For example, if there is a six-membered ring that contains a double bond, then the synthetic chemist's real personal computer – his or her brain – will instantly think of the Diels-Alder reaction. Even when there is no double bond, the chemist's brain will immediately insert one in all six positions and rapidly imagine the precursors for all possible Diels-Alder reactions. The chemist will then apply the same 'backwards' reasoning to the precursors to find all possible ways to make them, and the process will be repeated until simple starting materials are recognized. For example, Watt and coworkers began this process for the natural product, quassin, with an intermolecular Diels-Alder reaction and ended it when they arrived at the Wieland-Miescher ketone.[14]

With the advent of silicon chip computing devices ('computers'), it became possible to store large knowledge bases and to search them rapidly. As part of CAOS (Computer-Aided Organic Synthesis),[15] several groups developed synthesis planning algorithms and programs, which were generally faster and more thorough than humans, but we would argue, not more creative. For example, the LHASA (Logic and Heuristics Applied to Synthetic Analysis)[13] retrosynthetic scheme for the anticancer agent, taxol, features an intramolecular Diels-Alder reaction as the key step.[16] However, a team led by Nicolaou, who came out of the same group as LHASA,[17] has synthesized taxol *via* a highly convergent intermolecular Diels-Alder strategy,[18] ultimately based on disjoint isomorphic substructures.[12b,c]

Our goal is to give synthetic chemists additional conceptual tools, derived from discrete mathematics, *e.g.*, graph theory, information theory and combinatorics,[19] which we call MASA (Mathematics Applied to Synthetic Analysis). The practicing chemist can then apply the mathematical tools directly, apply the results (principles, guidelines, heuristics, *etc*.) derived from them, or use computer programs that incorporate the results of MASA.[20] In this paper we develop simplification principles for polycyclic skeletons based on topological complexity indices that we have introduced previously.[7,8]

RESULTS

In the Discussion section (*vide infra*) we succinctly review the topological complexity indices used in this study before analyzing the results obtained with them. In this section we will proceed directly to demonstrate their application, after brief digressions to introduce the indices of complexity, the change in complexity for a reaction, and branching in a molecule.

We use two families of indices, one based on subgraphs and the other on walk counts. In the former are the number of kinds of subgraphs, N_S, the total number of subgraphs, N_T, and the corresponding quantities, $N_S(lpe)$ and $N_T(lpe)$, when lone pairs of electrons (lpe) are included.[7] In the latter are the total walk count, *twc*, and the walk complexity, *wcx*.[8] Formal definitions are deferred to the Discussion.

The change in complexity ΔC for the general case of a reaction in which the *reactants* (or *precursors*) are converted into *products* (or *targets*) is calculated by applying equation (1). The way the total complexity C is calculated for an ensemble of n molecules depends on the index, and the original literature should be consulted.[7,8] In retrosynthetic analysis a transform is the exact reverse of a synthetic reaction and *vice versa*. Thus, the result of a retrosynthetic transform applied to a target molecule comprises the starting materials for the corresponding synthetic reaction, and we can recast equation (1) as equation (2), where the desired product of a synthetic re-

action is the target and the starting materials for it are the precursors. Equation (3) holds for any reaction-transform pair. In the typical transform the precursors are simpler than the target, and ΔC is negative.[22] The larger the absolute value, the greater the simplification. Then, for alternate disconnections of a target, the best one has the simplest precursors, *i.e.*, the lowest value of C(precursors). Therefore, in order to save space, we do not list values of ΔC in the tables. Instead, we list the rank in parentheses, where the lowest value of C(precursors) corresponds to the highest rank (#1).

$$\Delta C(\text{reaction}) = C(\text{products}) - C(\text{reactants}) \quad (1)$$

$$\Delta C(\text{transform}) = C(\text{precursors}) - C(\text{target}) \quad (2)$$

$$\Delta C(\text{transform}) = -\Delta C(\text{reaction}) \quad (3)$$

An important aspect of complexity is branching.[2–8] For our purposes a branched atom is one having two or more bonds incident to it,[23] not counting bonds to H. The more bonds, *i.e.*, the higher the degree of the atom, the more branched it is. The branching in a molecule is reduced whenever a bond to a branched atom is broken.

We begin by considering one-bond disconnections, which result in a single precursor for the cyclic structures discussed here. Table I contains index values for the three possible disconnections of bicyclo[2.2.1]heptane (**1**) and 7-azabicyclo[2.2.1]heptane (**5**), which has the same symmetry. All six indices identify **1** $\Rightarrow$ **2** as the best disconnection of **1**, *i.e.*, the most effective in terms of simplification. It reduces branching by breaking one of the equivalent bonds (*e.g.*, C1–C7) in the shorter bridge, thereby preserving the larger (six-membered) ring. Disconnection **1** $\Rightarrow$ **3**, which breaks one of four equivalent bonds (*e.g.*, C1–C2) in the longer bridges, reduces branching, but does not preserve the larger ring. It is alone in second place according to N_T, $N_T(lpe)$ and *twc*, which are independent of symmetry. It is tied for second/third by N_S and $N_S(lpe)$ and placed third (last) by *wcx*, the indices sensitive to symmetry. The symmetry-independent indices rank **1** $\Rightarrow$ **4** last (#3); it neither reduces branching nor preserves the larger ring.

For 7-azabicyclo[2.2.1]heptane (**5**) (Table I), where the N atom is at the unique position, there are again three possible disconnections. Breaking either C_α–N bond reduces branching and preserves the six-membered ring, and all six indices rank **5** $\Rightarrow$ **6** as the best disconnection. The degeneracy in N_S and $N_S(lpe)$ observed for the #2 and #3 disconnections of **1** (*vide supra*) is resolved by the introduction of N. According to N_T, $N_T(lpe)$ and *twc*, the #2 disconnection is **5** $\Rightarrow$ **7**, which reduces branching and preserves a five-membered ring by breaking C_α–C_β. According to N_S, $N_S(lpe)$ and *wcx*, #2 is **5** $\Rightarrow$ **8**, which breaks C_β–C_β and preserves both tertiary centers and a five-membered ring. This dichotomy

TABLE I. One-bond disconnections of bicyclo[2.2.1]heptane (**1**) and 7-azabicyclo[2.2.1]heptane (**5**)
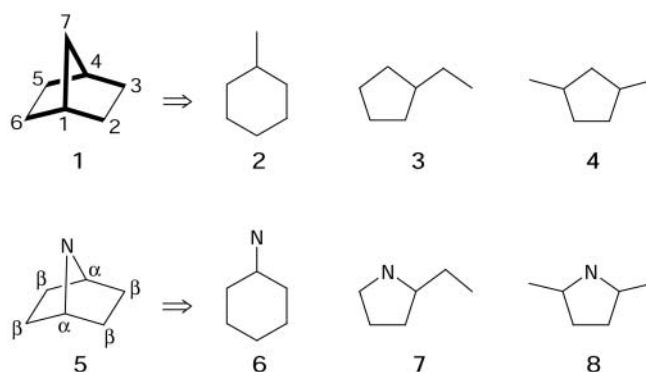
ID	N_S	N_T	N_S(lpe)	N_T(lpe)	*twc*	*wcx*
1	22	131	22	131	1974	880
2	15 (1)	60 (1)	15 (1)	60 (1)	1094 (1)	773 (1)
3	16 (2/3)	61 (2)	16 (2/3)	61 (2)	1180 (2)	822 (3)
4	16 (2/3)	70 (3)	16 (2/3)	70 (3)	1340 (3)	784 (2)
5	38	131	69	225	5660	3276
6	21 (1)	60 (1)	35 (1)	83 (1)	2400 (1)	1933 (1)
7	32 (3)	61 (2)	58 (3)	101 (2)	4208 (2)	4208 (3)
8	29 (2)	70 (3)	52 (2)	119 (3)	4722 (3)	3206 (2)

TABLE II. One-bond disconnections of 1-azabicyclo[2.2.1]heptane (**9**)
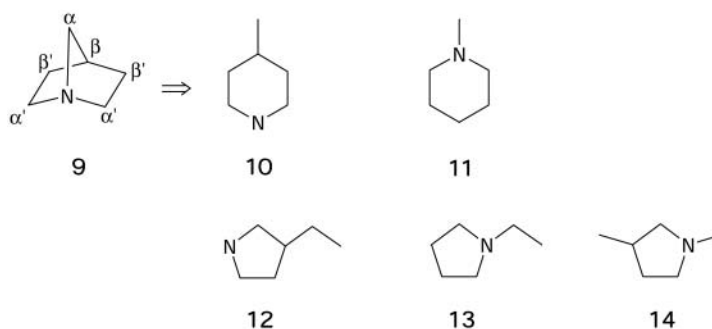
ID	N_S	N_T	N_S(lpe)	N_T(lpe)	*twc*	*wcx*
9	48	131	88	237	7254	5746
10	26 (2/3)	60 (1/2)	44 (1)	95 (1)	3604 (1)	2685 (1)
11	26 (2/3)	60 (1/2)	47 (3)	104 (3)	5440 (3)	4263 (3)
12	32 (4)	61 (3/4)	56 (4)	97 (2)	3810 (2)	3810 (2)
13	25 (1)	61 (3/4)	46 (2)	109 (4)	6052 (5)	4736 (4)
14	34 (5)	70 (5)	62 (5)	122 (5)	5854 (4)	5854 (5)

clearly reflects the difference between the indices that are independent of symmetry and those that are influenced by it, respectively, as a result of the greater symmetry in **8**. We tentatively put these disconnections in a tie, which will be resolved later (see Discussion).

In 1-azabicyclo[2.2.1]heptane (**9**) in Table II, the N atom is at one of the bridgehead positions, and the mirror plane is still present, but the C_2 axis through the 7-position in molecules **1** and **5** no longer exists; consequently, there are five possible disconnections. According to all indices except N_S, the best one is **9** ⇒ **10**, which reduces branching by breaking the C_α–N bond in the short bridge, thereby preserving the six-membered ring. It is analogous to **5** ⇒ **6** (*vide supra*); the polarity of the disconnected bond is merely reversed (N1–C7 instead of C1–N7).

The #2 disconnection according to N_T(lpe), *twc* and *wcx* is $\mathbf{9 \Rightarrow 12}$, which reduces branching by breaking a $C_{\alpha'}$–N bond in one of the equivalent long bridges, thereby preserving one of the five-membered rings. The 'kinds of' indices N_S and N_S(lpe) rank this disconnection significantly lower (#4), because of the lack of symmetry in **12**. All indices rank $\mathbf{9 \Rightarrow 11}$ third except N_T, which does not take account of heteroatoms. (*N.B.*, N_S has it tied for second/third.) It reduces branching and preserves the larger ring by breaking the C_α–C_β bond. The solid cases for #1 and #3 help cement the #2 position.

While N_T(lpe) and *wcx* place it fourth, at least one index puts $\mathbf{9 \Rightarrow 13}$ in every position (counting ties). Nevertheless, this disconnection can be assigned fourth place with confidence, since the assignments for third (*vide supra*) and fifth (*vide infra*) are very secure. It reduces branching by breaking one of the C_β–$C_{\beta'}$ bonds, thereby preserving a five-membered ring. The #1 and #2 rankings by N_S and N_S(lpe), respectively, are a consequence of symmetry in **13** and its subgraphs. (*N.B.*, breaking the ring bond farthest from N gives a subgraph with threefold symmetry.) All indices rank $\mathbf{9 \Rightarrow 14}$ in fifth (last) place except *twc*, which puts it fourth. It breaks a $C_{\alpha'}$–$C_{\beta'}$ bond rather than a C–N bond, preserves one of the smaller rings rather than the larger one, and preserves both tertiary centers rather than eliminating one of them.

The order of simplification (by precursor), $\mathbf{10 > 12 > 11 > 13 > 14}$, can be explained by three factors: (i) reducing branching puts **10–13** ahead of **14**; (ii) the kind of bond broken (C–N *vs.* C–C) puts **10** and **12** ahead of **11** and **13**; and (iii) preserving the larger ring gives **10 >**

12 and **11 > 13**. (*N.B.*, proximity to N gives **11 > 13**, but not **10 > 12**.)

The 2-azabicyclo[2.2.1]heptane skeleton (**15**), which has neither mirror plane nor C_2 axis, is the most interesting, as it has the largest number of disconnections, eight (Table III). In this case it is not possible to simultaneously break a C–N bond and preserve the larger ring, and the indices must determine which process is more important. According to all the indices that are sensitive to heteroatoms, the best disconnection is $\mathbf{15 \Rightarrow 18}$, which reduces branching by breaking C_α–N, thereby preserving a five-membered ring. Thus, in this case breaking a C–N bond is more important than preserving the six-membered ring by breaking either C–C bond in the short bridge (*cf.* $\mathbf{15 \Rightarrow 16}$ or $\mathbf{15 \Rightarrow 17}$ below).

Alone in second place according to N_S and N_T(lpe) is $\mathbf{15 \Rightarrow 19}$, which reduces branching by breaking the $C_{\alpha'}$–$C_{\beta'}$ bond, but does not preserve the larger ring. Alone in second place according to *twc* and *wcx* is $\mathbf{15 \Rightarrow 20}$, which breaks the $C_{\alpha'}$–N bond, but does not reduce branching or preserve the larger ring. Interestingly, N_S(lpe) spans this divide by putting them in a tie for #2/#3, and we tentatively consider them tied (see Discussion), since the symmetry-independent indices N_T(lpe) and *twc* are split on which is more important. (A caveat is the fact that the difference between N_T(lpe) values for **19** and **20** is relatively small, 95 *vs.* 97, respectively.) The important thing to note is that each of the top three disconnections of **15** breaks one of the three bonds in the bridge containing the N atom.

Disconnection $\mathbf{15 \Rightarrow 17}$, which breaks C_α–$C_{\beta''}$ in the short bridge, is ranked third by symmetry-independent

TABLE III. One-bond disconnections of 2-azabicyclo[2.2.1]heptane (**15**)
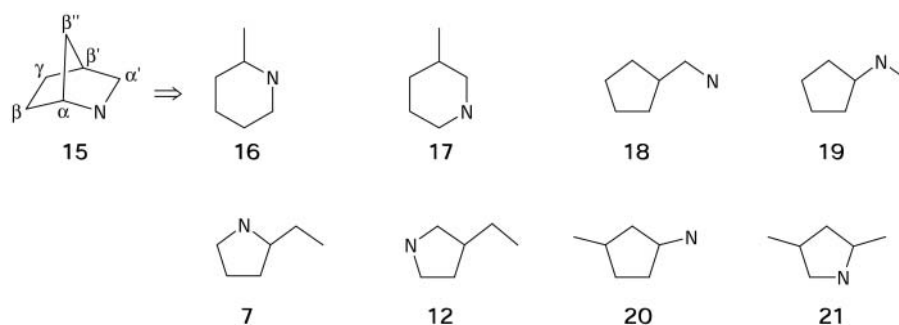
ID	N_S	N_T	N_S(lpe)	N_T(lpe)	*twc*	*wcx*
15	66	131	120	219	5210	5210
16	32 (5/6/7)	60 (1/2)	58 (6/7)	99 (6)	4002 (6)	4002 (6)
17	31 (4)	60 (1/2)	54 (4)	96 (3)	3676 (3)	3676 (4)
18	24 (1)	61 (3/4/5/6)	36 (1)	79 (1)	2188 (1)	1786 (1)
19	27 (2)	61 (3/4/5/6)	48 (2/3)	95 (2)	3698 (4)	3126 (3)
7	32 (5/6/7)	61 (3/4/5/6)	58 (6/7)	101 (7)	4208 (7)	4208 (7)
12	32 (5/6/7)	61 (3/4/5/6)	56 (5)	97 (4/5)	3810 (5)	3810 (5)
20	30 (3)	70 (7/8)	48 (2/3)	97 (4/5)	2670 (2)	2670 (2)
21	39 (8)	70 (7/8)	70 (8)	116 (8)	4396 (8)	4396 (8)

indices N_T(lpe) and *twc*, and it is ranked fourth by symmetry-dependent indices N_S, N_S(lpe) and *wcx*. Overall, it must be ranked fourth, since there is a tie for second/third, as discussed above. Breaking the other bond in the short bridge, $C_{\beta'}$–$C_{\beta''}$, also reduces branching and preserves the larger ring, but this bond is farther from N, and the corresponding disconnection, $15 \Rightarrow 16$, is ranked #6 by all indices sensitive to heteroatoms. (*N.B.*, N_S and N_S(lpe) are involved in ties.)

Disconnection $15 \Rightarrow 12$ is ranked #5, and $15 \Rightarrow 7$ is ranked #7 by all indices sensitive to heteroatoms. (*N.B.*, there are two ties in each case.) They reduce branching by breaking C_α–C_β and $C_{\beta'}$–C_γ, respectively, in the long bridge that does not contain N, thereby preserving five-membered rings. All indices indicate that $15 \Rightarrow 21$ is in last place (#8): it breaks C_β–C_γ rather than a C–N, α,β or β,β-bond, preserves the smaller ring rather than the larger one, breaks a bond in a bridge without N, and does not reduce branching.

The top three disconnections of **15** break bonds in the bridge containing N (factor (i)), and their order will be settled in the Discussion section. As far as the disconnections that break C–C bonds in the two bridges without N are concerned, the order of simplification (by precursor), **17** > **12** > **16** > **7** > **21**, can be explained by three additional factors: (ii) reducing branching puts the first four ahead of **21**; (iii) proximity to N puts **17** and **12** ahead of **16** and **7**; and (iv) preserving the larger ring gives **17** > **12** and **16** > **7**. (*N.B.*, proximity to N also gives **16** > **7**, but not **17** > **12**.)

In the previous targets, it was always possible to simultaneously break a C–N bond and reduce branching. In the next series of examples, the number of atoms is the same, but a quaternary center is present instead of two tertiary ones, and in one case neither C–N bond cleavage reduces branching.

Spiro[3.3]heptane (**22**) and 4-azoniaspiro[3.3]heptane (**25**) in Table IV have the same symmetry, since the N atom is at the unique center position, and in each case there are two possible disconnections. According to all the relevant indices, the best ones are $22 \Rightarrow 23$ and $25 \Rightarrow 26$, respectively, both of which reduce branching by breaking a bond to the quaternary center. In the latter the bond broken is one of four equivalent C–N bonds, and there is no conflict between the goal of reducing branching and breaking a C–N bond. The quaternary N in **25** does not have a lpe, and consequently N_S(lpe) = N_S and N_T(lpe) = N_T for this structure. When one N in a transform has a lpe and another does not, the order of disconnections may change because of this fact alone, as in this case. The 'lpe' indices are not useful here, but N_S, *twc* and *wcx* are relevant (see Discussion).

For 2-azaspiro[3.3]heptane (**28**) in Table V, there are four possible disconnections, and a choice must be made between breaking a C–N bond that does not reduce branching and a C–C bond that does. There is a three-way tie in N_S and a pair of two-way ties in N_T, which are all resolved by the more sophisticated 'lpe' indices; however, the orders induced by N_S(lpe) and N_T(lpe) are different. Three indices, N_S(lpe), *twc* and *wcx* (the 'troika'), indicate that $28 \Rightarrow 30$ is #1, $28 \Rightarrow 29$ is #2 and $28 \Rightarrow 31$ is #3. The first breaks one of the equivalent C_α–N bonds, but does not reduce branching. The second and third reduce branching by breaking C_α–C_β or C_β–C_γ, respectively, the former in the ring with N and the latter in the ring without it.

TABLE IV. One-bond disconnections of spiro[3.3]heptane (**22**) and 4-azoniaspiro[3.3]heptane (**25**)
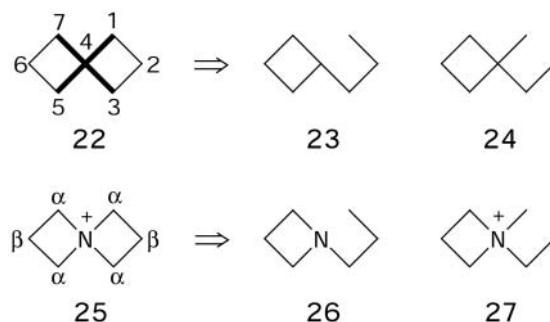
ID	N_S	N_T	N_S(lpe)	N_T(lpe)	*twc*	*wcx*
22	22	133	22	133	2408	1118
23	16 (1)	56 (1)	16 (1)	56 (1)	1214 (1)	1001 (1)
24	19 (2)	76 (2)	19 (2)	76 (2)	1698 (2)	1407 (2)
25	28	133	28	133	11302	6226
26	22 (1)	56 (1)	41 (2)	100 (2)	6238 (1)	5277 (1)
27	25 (2)	76 (2)	25 (1)	76 (1)	9474 (2)	8179 (2)

On the other hand, N_T(lpe) indicates that the order of simplification (by precursor) is $29 > 31 > 30$, $i.e.$, it demotes $28 \Rightarrow 30$ to third place, but keeps $28 \Rightarrow 29$ ahead of $28 \Rightarrow 31$. Thus, N_T(lpe) and the troika agree that the best C–C bond to disconnect is the one closest to N, but they disagree as to which is more important, reducing branching or breaking a C–N bond, respectively. All indices put $28 \Rightarrow 32$ last; it does not break a C–N bond or reduce branching, and the C_γ–C_δ bond broken is the farthest from N.

In this case the order of simplification (by precursor), $30 > 29 > 31 > 32$, can be explained by a single factor, proximity to N.

For 1-azaspiro[3.3]heptane (33) in Table VI, there are more kinds of disconnections, six, owing to lower symmetry. All indices agree that the #1 disconnection is $33 \Rightarrow 34$, which breaks the C_α–N bond to the quaternary carbon, thereby reducing branching. While the bond polarity is reversed, the $C_{\alpha'}$–N bond in 33 is in the same position as the C_α–N bonds in 28 ($vide\ supra$), and the

same choice must be confronted, but for #2 this time. According to the troika (N_S(lpe), twc and wcx), the preferred disconnection again breaks the C–N bond to the secondary carbon rather than a C–C bond to the quaternary carbon, $i.e.$, $33 \Rightarrow 36$ is #2. And once again, N_T(lpe) favors breaking a C–C bond that reduces branching, although this time it is in the ring that does not contain N ($vide\ infra$). According to both N_T(lpe) and the troika, $33 \Rightarrow 35$ is the #3 disconnection; it reduces branching by breaking C_α–$C_{\beta'}$ in the ring that contains N.

There is not as much consensus about fourth place as third and fifth ($cf.$ 9). According to N_S(lpe) and twc, it is $33 \Rightarrow 38$, which reduces branching by breaking C_α–C_β in the ring that does not contain N. It is ranked significantly higher (#2) by N_T(lpe) and a bit lower (#5) by wcx; however, the former ranking is tenuous, as the difference between the #2 and #3 disconnections is only ΔN_T(lpe) = 1 subgraph. Fifth, according to N_S(lpe), N_T(lpe) and twc, is $33 \Rightarrow 37$, which breaks $C_{\alpha'}$–$C_{\beta'}$ in the ring that contains

TABLE V. One-bond disconnections of 2-azaspiro[3.3]heptane (28)
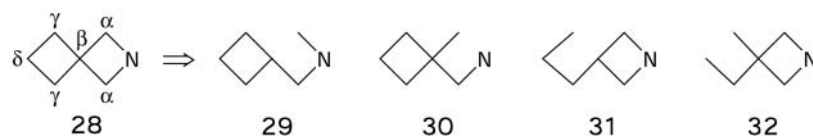
ID	N_S	N_T	N_S(lpe)	N_T(lpe)	twc	wcx
28	41	133	69	214	5470	4152
29	26 (1/2/3)	56 (1/2)	43 (2)	82 (1)	3310 (2)	3051 (2)
30	26 (1/2/3)	76 (3/4)	38 (1)	100 (3)	2802 (1)	2473 (1)
31	26 (1/2/3)	56 (1/2)	44 (3)	88 (2)	3972 (3)	3253 (3)
32	28 (4)	76 (3/4)	48 (4)	122 (4)	4696 (4)	3879 (4)

TABLE VI. One-bond disconnections of 1-azaspiro[3.3]heptane (33)
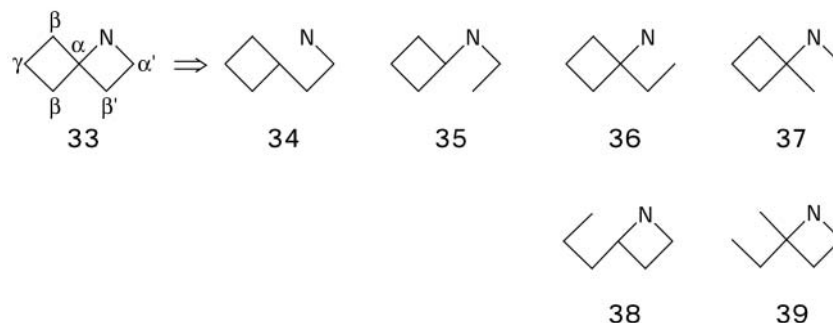
ID	N_S	N_T	N_S(lpe)	N_T(lpe)	twc	wcx
33	57	133	102	224	6198	5648
34	22 (1)	56 (1/2/3)	32 (1)	70 (1)	2134 (1)	1913 (1)
35	26 (2)	56 (1/2/3)	47 (3)	92 (3)	4126 (3)	3739 (3)
36	28 (3)	76 (4/5/6)	44 (2)	110 (4)	3466 (2)	3041 (2)
37	32 (4/5)	76 (4/5/6)	55 (5)	122 (5)	4714 (5)	4239 (4)
38	32 (4/5)	56 (1/2/3)	49 (4)	91 (2)	4298 (4)	4298 (5)
39	38 (6)	76 (4/5/6)	68 (6)	127 (6)	5304 (6)	5304 (6)

N, but does not reduce branching. (*N.B.*, N_S has them tied for #4/#5.) As with first place, there is unanimity as to last (#6): $33 \Rightarrow 39$ does not break a C–N bond or reduce branching, and moreover it breaks a C_β–C_γ bond, which is farthest from N and in the ring without it.

The order of simplification (by precursor), $34 > 36 > 35 > 38 > 37 > 39$, can be explained by three factors: (i) the kind of bond broken (C–N *vs*. C–C) puts **34** and **36** ahead of the rest; (ii) reducing branching gives **34** > **36** and puts **35** and **38** ahead of **37** and **39**; and (iii) breaking a bond in the same ring as N gives **35** > **38** and **37** > **39**. (*N.B.*, proximity to N gives **37** > **39**, but not **35** > **38**, as the disconnections that afford **35**, **37** and **38** all break α,β-bonds.)

The two-bond disconnections that yield two components from bicyclo[2.2.1]heptane **1** and its aza derivatives **5**, **9** and **15** are summarized in Table VII. Choosing the best one for each target, we have $1 \Rightarrow 42$, $5 \Rightarrow 43$, $9 \Rightarrow 45$ and $15 \Rightarrow 51$. In every case at least five of the six indices agree as to #1. Ignoring N_T for the aza derivatives, there are but two genuine exceptions: only *wcx* ranks $1 \Rightarrow 40$ above $1 \Rightarrow 42$, and only *twc* ranks $15 \Rightarrow 49$ above $15 \Rightarrow 51$. The top ranking of $1 \Rightarrow 40$ by *wcx* is undoubtedly a result of its unusually high (six-fold) symmetry. The only disconnection of **15** that breaks two C–N bonds is $15 \Rightarrow 49$, and its top ranking by *twc* is a manifestation of the heteroatom effect (see Discussion).

Since the number of components resulting from each disconnection is the same, the most important simplifying factor appears to be making the complexities of the two precursors as equal as possible. Taking the two-bond disconnections of **9** for example, ethane (*cf*. **45**) is more complex than methane (*cf*. **46**–**48**), and azacyclopentane (*cf*. **45**) is less complex than azacyclohexane (*cf*. **46**), 2-methylazacyclopentane (*cf*. **47**) or *N*-methylazacyclopentane (*cf*. **48**), so that the gap between components in **45** is smaller than in **46**–**48**. The smaller the gap, the more equal the complexities of the components must be, and all indices agree that $9 \Rightarrow 45$ is the best two-bond disconnection of target **9**. As will be shown below, the two-bond disconnections for each target can be ordered in the same manner as the one-bond disconnections.

DISCUSSION

The kinds of complexity and the criteria for a topological index to also be a complexity index have been discussed in detail,[3a,7a–c] and these aspects will not be repeated here. A succinct Prolegomenon to Chemical Graph Theory is available as Electronic supplementary information,[12a] and a compendious book on chemical graph theory has been published by Trinajstić.[21] We précis the essential concepts here.

A *graph* is a collection of points, joined by lines, which in turn are defined as unordered pairs of points.[19a]

Some of the points are labeled with 'N', the chemical symbol for nitrogen. As in our examples, the points in chemical graph theory usually represent atoms and the lines, bonds,[21] although there are many other possibilities for a concept as general as this one, *e.g.*, the points can represent molecules and the lines, reactions that convert one molecule into another (*cf.* reaction graphs[24] and synthesis graphs[25]). *Hydrogen-suppressed* molecular graphs are usually used, in which H atoms and the bonds to them are omitted. As in the abbreviated structural formulas commonly used by organic chemists, a carbon atom is assumed to be present wherever two or more lines intersect. The structures drawn in Tables I–VII are in fact graphs, since no explicit information is contained in them concerning bond lengths or angles. They are drawn in a manner that is suggestive of their three-dimensional architecture to someone with a knowledge of organic chemistry, but this information is not necessary to understand our results, which are based solely on changes in constitution.

Isomorphic graphs M and N, M $\cong$ N, have the same adjacency matrix for some labeling. Therefore, it does not matter when a graph is redrawn with different point positions or line lengths. A *subgraph* $S_i(M)$ of graph M has all its points and lines in M; thus, a graph is a subgraph of itself, *e.g.*, when M is methane, $S_1 = M$ is the only subgraph. A subgraph is usually named after the stable molecule with the same skeleton. The *number of kinds of subgraphs* $N_S = N_S(M)$ counts non-isomorphic connected subgraphs of the molecular graph M, and the *total number of subgraphs* $N_T = N_T(M)$ counts all possible connected subgraphs, isomorphic and non-isomorphic.[7a–e] Lone pairs of electrons (lpe) are not included in N_S and N_T. When lpe are added, *e.g.*, to a heteroatom, the indices are called $N_S(lpe)$ and $N_T(lpe)$, respectively, as illustrated in the examples below. It can be a daunting task to find all possible subgraphs in a structure of even moderate complexity, and the computer program written for this purpose is very helpful.[8d–f]

A *walk* in a graph is a series of points $p_1, p_2, p_3, \ldots, p_n$, connected by lines $l_{1,2}, l_{2,3}, l_{3,4}, \ldots, l_{n-1, n}$. The points and hence the lines are not necessarily unique, *i.e.*, they can occur more than once in a given walk. We now use directed walks;[26] thus, the walk from p_1 to p_n is distinct from the walk from p_n to p_1 that traverses the intermediate points in the reverse order. The *total walk count (twc)* of a graph on n points is defined as the total number of (directed) walks of length 1 through $n - 1$, which is easily obtained from the adjacency matrix.[3a,8a–c] The *walk complexity (wcx)* is similar to *twc*, but only includes walks beginning at one point from each set of equivalent points based on symmetry.[8a] Consequently, it is decreased by symmetry, as are N_S and $N_S(lpe)$. On the other hand, *twc*, N_T and $N_T(lpe)$ are not sensitive to symmetry, and we prefer them for determining the order of simplification, as symmetry is not always a simplifying factor in synthetic analysis.[12f,13] More to the point, the symmetry present in a disconnection is often ab-

TABLE VII. Selected two-bond disconnections of **1** and its aza derivatives
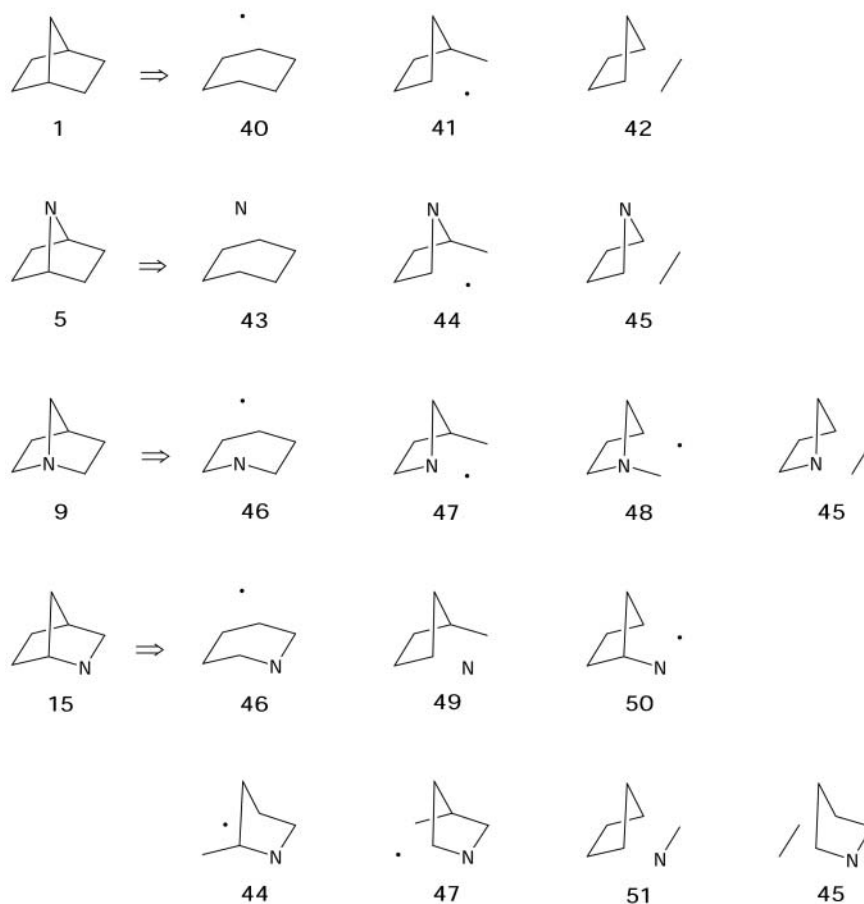
ID	N_S	N_T	N_S(lpe)	N_T(lpe)	*twc*	*wcx*
1	22	131	22	131	1974	880
40	7 (2)	38 (2)	7 (2)	38 (2)	756 (2)	126 (1)
41	12 (3)	44 (3)	12 (3)	44 (3)	968 (3)	638 (3)
42	6 (1)	29 (1)	6 (1)	29 (1)	642 (1)	132 (2)
5	38	131	69	225	5660	3276
43	8 (1)	38 (2)	9 (1)	39 (1)	882 (1)	252 (1)
44	24 (3)	44 (3)	43 (3)	72 (3)	3880 (3)	3880 (3)
45	14 (2)	29 (1)	24 (2)	45 (2)	3104 (2)	2210 (2)
9	48	131	88	237	7254	5746
45	14 (1)	29 (1)	24 (1)	45 (1)	3104 (1)	2210 (1)
46	18 (2)	38 (2)	31 (2)	60 (2)	3218 (2)	2372 (2)
47	23 (4)	44 (3/4)	40 (4)	70 (3)	3570 (3)	3570 (3)
48	20 (3)	44 (3/4)	36 (3)	76 (4)	5314 (4)	4076 (4)
15	66	131	120	219	5210	5210
44	24 (7)	44 (4/5/6/7)	43 (7)	72 (7)	3880 (7)	3880 (7)
45	14 (3)	29 (1/2)	24 (3)	45 (2/3)	3104 (4)	2210 (4)
46	18 (5)	38 (3)	31 (5)	60 (4)	3218 (5)	2372 (5)
47	23 (6)	44 (4/5/6/7)	40 (6)	70 (6)	3570 (6)	3570 (6)
49	13 (2)	44 (4/5/6/7)	14 (2)	45 (2/3)	1094 (1)	764 (2)
50	17 (4)	44 (4/5/6/7)	28 (4)	61 (5)	2274 (3)	1790 (3)
51	8 (1)	29 (1/2)	10 (1)	31 (1)	1204 (2)	700 (1)

sent from the transform chosen to accomplish the corresponding reaction in a synthesis plan.

In the case of *twc* and *wcx*, heteroatoms are included by weighting points, resulting in diagonal elements $a_{ii} > 0$ in the adjacency matrix. Customarily, $a_{ii} = 2$ for nitrogen, but we also discuss selected results with $a_{ii} = 1$ or 3 for comparison. For carbon $a_{ii} = 0$ throughout, although it could be set at non-zero values, in particular, for carbanions or carbenes. (Likewise, bond multiplicity is included by weighting lines, so that $a_{ij} = 1$ for a single bond, $a_{ij} = 2$ for a double bond and $a_{ij} = 3$ for a triple bond.) When a heteroatom is present, the result of weighting a_{ii} on *twc* is referred to as the *heteroatom effect*.

We illustrate the calculation of these indices for the graphs of ethane (C1–C2) and methylamine (N1–C2), where one of the points has been labeled with 'N' in the latter. Ethane has two methane subgraphs (C1 and C2) and one ethane (C1–C2), consequently $N_S = 2$ and $N_T = 3$. Methylamine has one ammonia (N1), one methane (C2) and one methylamine subgraph (N1–C2), and $N_S = N_T = 3$. When the lpe on nitrogen is added, methylamine has two additional subgraphs, :N1 and :N1–C2, and $N_S(\text{lpe}) = N_T(\text{lpe}) = 5$. For ethane there are two walks of length 1, C1–C2 and C2–C1, so that *twc* = 2, and since C1 and C2 represent equivalent points, *wcx* = 1. For methylamine with

$a_{ii} = 2$, the walks are N1–C2 (1), C2–N1 (1) and N1–N1 (2); therefore, *twc* = *wcx* = 4.

Occasionally, the order of simplification based on *twc* changes upon going from $a_{ii} = 2$ to $a_{ii} = 1$, *i.e.*, towards the order for the corresponding hydrocarbon ($a_{ii} = 0$). For example, making this change with the disconnections of **28**, the top three precursors go from **30 > 29 > 31** to **29 > 31 > 30**, which is the same as the order for $N_T(\text{lpe})$. When $a_{ii} = 1$, the #4 and #5 disconnections of **9** according to *twc* are reversed, whereupon they match the order induced by $N_T(\text{lpe})$. Going from $a_{ii} = 2$ to $a_{ii} = 3$ may also affect the results of *twc*; *e.g.*, it interchanges third and fourth place for **15**. The resulting order of simplification does not match that given by $N_T(\text{lpe})$, nor does it match upon going to $a_{ii} = 1$; thus, a lower value of a_{ii} does not always bring *twc* into alignment with $N_T(\text{lpe})$. All things considered, $a_{ii} = 2$ appears to be a very reasonable choice for applications such as this one; we do not see a reason to use $a_{ii} = 1$, since it does not fully reveal the heteroatom effect. We believe that it is better to have indices that respond differently to various structural features, so that the best one can be chosen for a particular application. Based on our results, *twc* appears to be more influenced by proximity to N and $N_T(\text{lpe})$ by branching.

Table VIII summarizes the one-bond disconnections from Tables I–III, Table IX those from Tables IV–VI, and

TABLE VIII. One-bond disconnections of **1** and its aza derivatives ordered by topological complexity indices

Rank[a]	Disconnection	Bond broken	Indices[b]	SP
1	**1 ⇒ 2**	C1–C7	N_S, N_T, $N_S(\text{lpe})$, $N_T(\text{lpe})$, *twc*, *wcx*	4, 5
2	**1 ⇒ 3**	C1–C2	$[N_S]$, N_T, $[N_S(\text{lpe})]$, $N_T(\text{lpe})$, *twc*	5
3	**1 ⇒ 4**	C2–C3	$[N_S]$, N_T, $[N_S(\text{lpe})]$, $N_T(\text{lpe})$, *twc*	φ
1	**5 ⇒ 6**	C_α–N	N_S, $N_S(\text{lpe})$, $N_T(\text{lpe})$, *twc*, *wcx*	3, 4, 5
2/3	**5 ⇒ 7**	C_α–C_β	$N_T(\text{lpe})$, *twc*	5, (7)[c],[d]
2/3	**5 ⇒ 8**	C_β–C_β	N_S, $N_S(\text{lpe})$, *wcx*	φ
1	**9 ⇒ 10**	C_α–N	$N_S(\text{lpe})$, $N_T(\text{lpe})$, *twc*, *wcx*	3, 4, 5
2	**9 ⇒ 12**	$C_{\alpha'}$–N	$N_T(\text{lpe})$, *twc*, *wcx*	3, 5
3	**9 ⇒ 11**	C_α–C_β	$[N_S]$, $N_S(\text{lpe})$, $N_T(\text{lpe})$, *twc*, *wcx*	4, 5, (7)[c],[e]
4	**9 ⇒ 13**	C_β–$C_{\beta'}$	$N_T(\text{lpe})$, *wcx*	5
5	**9 ⇒ 14**	$C_{\alpha'}$–$C_{\beta'}$	N_S, $N_S(\text{lpe})$, $N_T(\text{lpe})$, *wcx*	(7)[c],[e]
1	**15 ⇒ 18**	C_α–N	N_S, $N_S(\text{lpe})$, $N_T(\text{lpe})$, *twc*, *wcx*	3, 5
2/3	**15 ⇒ 20**	$C_{\alpha'}$–N	$[N_S(\text{lpe})]$, *twc*, *wcx*	3
2/3	**15 ⇒ 19**	$C_{\alpha'}$–$C_{\beta'}$	N_S, $[N_S(\text{lpe})]$, $N_T(\text{lpe})$	5, 6, (7)[c],[f]
4	**15 ⇒ 17**	C_α–$C_{\beta''}$	N_S, $N_S(\text{lpe})$, *wcx*	4, 5, (7)[c],[f]
5	**15 ⇒ 12**	C_α–C_β	$[N_S]$, $N_S(\text{lpe})$, $[N_T(\text{lpe})]$, *twc*, *wcx*	5, (7)[c],[f]
6	**15 ⇒ 16**	$C_{\beta'}$–$C_{\beta''}$	$[N_S]$, $[N_S(\text{lpe})]$, $N_T(\text{lpe})$, *twc*, *wcx*	4, 5, (7)[c],[g]
7	**15 ⇒ 7**	$C_{\beta'}$–C_γ	$[N_S]$, $[N_S(\text{lpe})]$, $N_T(\text{lpe})$, *twc*, *wcx*	5
8	**15 ⇒ 21**	C_β–C_γ	N_S, $N_S(\text{lpe})$, $N_T(\text{lpe})$, *twc*, *wcx*	φ

[a] A 2/3 indicates a tie for #2 and #3 with indices given for #2. [b] Brackets indicate a tie. [c] Only applies when comparing this disconnection to one of lower rank according to SP7. [d] Breaks an α,β-bond, which is better than the β,β-bond in **5 ⇒ 8**. [e] Breaks an α,β-bond, which is better than the β,β-bond in #4. [f] Breaks an α,β-bond, which is better than the β,β or β,γ-bonds in #6–8. [g] Breaks a β,β-bond, which is better than the β,γ-bonds in #7, 8.

TABLE IX. One-bond disconnections of **22** and its aza derivatives ordered by topological complexity indices

Rank	Disconnection	Bond broken	Indices[a]	SP
1	**22 ⇒ 23**	C1–C4	N_S, N_T, N_S(lpe), N_T(lpe), twc, wcx	5
2	**22 ⇒ 24**	C1–C2	N_S, N_T, N_S(lpe), N_T(lpe), twc, wcx	φ
1	**25 ⇒ 26**	C_α–N	N_S, twc, wcx	3, 5
2	**25 ⇒ 27**	C_α–C_β	N_S, twc, wcx	φ
1	**28 ⇒ 30**	C_α–N	[N_S], N_S(lpe), twc, wcx	3
2	**28 ⇒ 29**	C_α–C_β	[N_S], N_S(lpe), twc, wcx	5, 6, (7)[b],[c]
3	**28 ⇒ 31**	C_β–C_γ	[N_S], N_S(lpe), twc, wcx	5, (7)[b],[d]
4	**28 ⇒ 32**	C_γ–C_δ	N_S, N_S(lpe), N_T(lpe), twc, wcx	φ
1	**33 ⇒ 34**	C_α–N	N_S, N_S(lpe), N_T(lpe), twc, wcx	3, 5
2	**33 ⇒ 36**	$C_{\alpha'}$–N	N_S(lpe), twc, wcx	3
3	**33 ⇒ 35**	C_α–$C_{\beta'}$	N_S(lpe), N_T(lpe), twc, wcx	5, 6, (7)[b],[e]
4	**33 ⇒ 38**	C_α–C_β	[N_S], N_S(lpe), twc	5, (7)[b],[e]
5	**33 ⇒ 37**	$C_{\alpha'}$–$C_{\beta'}$	[N_S], N_S(lpe), N_T(lpe), twc	6, (7)[b],[e]
6	**33 ⇒ 39**	C_β–C_γ	N_S, N_S(lpe), N_T(lpe), twc, wcx	φ

[a] Brackets indicate a tie. [b] Only applies when comparing this disconnection to one of lower rank according to SP7. [c] Breaks an α,β-bond, which is better than the β,γ-bond in #3 or the γ,δ-bond in #4. [d] Breaks a β,γ-bond, which is better than the γ,δ-bond in #4. [e] Breaks an α,β-bond, which is better than the β,γ-bond in #6.

TABLE X. Two-bond disconnections of **1** and its aza derivatives ordered by topological complexity indices

Rank[a]	Disconnection	Bonds broken	Indices[b]	SP
1	**1 ⇒ 42**	C1–C2, C3–C4	N_S, N_T, N_S(lpe), N_T(lpe), twc	2, 5, 5
2	**1 ⇒ 40**	C1–C7, C4–C7	N_S, N_T, N_S(lpe), N_T(lpe), twc	(2),[c] 4, 5, 5
3	**1 ⇒ 41**	C1–C2, C2–C3	N_S, N_T, N_S(lpe), N_T(lpe), twc, wcx	5
1	**5 ⇒ 43**	C_α–N, C_α–N	N_S, N_S(lpe), N_T(lpe), twc, wcx	2, 3, 3, 4, 5, 5
2	**5 ⇒ 45**	C_α–C_β, C_α–C_β	N_S, N_S(lpe), N_T(lpe), twc, wcx	(2),[c] 5, 5, (7)[d],[e]
3	**5 ⇒ 44**	C_α–C_β, C_β–C_β	N_S, N_S(lpe), N_T(lpe), twc, wcx	5
1	**9 ⇒ 45**	$C_{\alpha'}$–N, C_β–$C_{\beta'}$	N_S, N_S(lpe), N_T(lpe), twc, wcx	2, 3, 5, 5
2	**9 ⇒ 46**	C_α–N, C_α–C_β	N_S, N_S(lpe), N_T(lpe), twc, wcx	(2),[c] 3, 4, 5, 5, (7)[d],[f]
3	**9 ⇒ 47**	$C_{\alpha'}$–N, $C_{\alpha'}$–$C_{\beta'}$	N_T(lpe), twc, wcx	(2),[c] 3, 5, (7)[d],[f]
4	**9 ⇒ 48**	$C_{\alpha'}$–$C_{\beta'}$, C_β–$C_{\beta'}$	N_T(lpe), twc, wcx	5
1	**15 ⇒ 51**	C_α–N, $C_{\alpha'}$–$C_{\beta'}$	N_S, N_S(lpe), N_T(lpe), wcx	2, 3, 5, 5, 6, (7)[d],[g]
2	**15 ⇒ 49**	C_α–N, $C_{\alpha'}$–N	N_S, N_S(lpe), [N_T(lpe)], wcx	(2),[c] 3, 3, 5
3/4	**15 ⇒ 50**	$C_{\alpha'}$–N, $C_{\alpha'}$–$C_{\beta'}$	twc, wcx	(2),[h] 3, 5, 6, (7)[d],[g]
3/4	**15 ⇒ 45**	C_α–C_β, $C_{\beta'}$–C_γ	N_S, N_S(lpe), [N_T(lpe)]	(2),[c] 5, 5, (7)[d],[i]
5	**15 ⇒ 46**	C_α–$C_{\beta''}$, $C_{\beta'}$–$C_{\beta''}$	N_S, N_S(lpe), twc, wcx	(2),[c] 4, 5, 5, (7)[d],[j]
6	**15 ⇒ 47**	C_α–C_β, C_β–C_γ	N_S, N_S(lpe), N_T(lpe), twc, wcx	(2),[c] 5, (7)[d],[i]
7	**15 ⇒ 44**	C_β–C_γ, $C_{\beta'}$–C_γ	N_S, N_S(lpe), N_T(lpe), twc, wcx	5

[a] A 3/4 indicates a tie for #3 and #4 with indices given for #3. [b] Brackets indicate a tie. [c] The precursors have more equal complexities in this case than those of lower rank according to SP2, but not the most equal overall, which is indicated by 2. [d] Only applies when comparing this disconnection to one of lower rank according to SP7. [e] Breaks two α,β-bonds and is better than #3, which breaks one α,β and one β,β-bond. [f] Breaks a C–N and an α,β-bond and is better than #1, which breaks a C–N and a β,β-bond, and #4, which breaks an α,β and a β,β-bond. [g] Breaks a C–N and an α,β-bond and is better than **15 ⇒ 45** and #5–7, which break at least one bond more distant from N. [h] According to SP2, **15 ⇒ 50** is ranked above #5–7, but not above **15 ⇒ 45** (see Discussion). [i] Breaks an α,β and a β,γ-bond and is better than #7, which breaks two β,γ-bonds. [j] Breaks an α,β and a β,β-bond and is better than **15 ⇒ 45** and #6, 7, which break at least one bond more distant from N.

Table X the two-bond disconnections from Table VII. For each target the disconnections are listed in order of decreasing simplification, along with the indices upon which the ranks are based and the relevant simplification principles (SP), which are explained below. All six indices are used with hydrocarbons **1** and **22**; however, N_T is omitted when heteroatoms are present, as it is not sensitive to them.

Looking over all of the results, the most remarkable observation is the high level of agreement among the various indices as far as the #1 disconnection of each target is concerned. For parent compounds **1** and **22** all six indices point to the best one-bond disconnections, **1** $\Rightarrow$ **2** and **22** $\Rightarrow$ **23**, respectively. For the aza derivatives all five of the heteroatom-sensitive indices select **5** $\Rightarrow$ **6**, **15** $\Rightarrow$ **18** and **33** $\Rightarrow$ **34** as the best one-bond disconnections, and four of five pick **9** $\Rightarrow$ **10** and **28** $\Rightarrow$ **30**, counting a three-way tie in N_S for the latter. The lack of lpe in **25** means $N_S(\text{lpe}) = N_S$ and $N_T(\text{lpe}) = N_T$ for this structure, so that there are only three valid indices (N_S, *twc* and *wcx*) in this example, and all of them agree that **25** $\Rightarrow$ **26** is #1. The best two-bond disconnections are also identified unanimously or with a single exception.

The other positions are not as clear-cut; nevertheless, most of them can be assigned with a high degree of certainty. They are usually supported by at least five of six indices for the hydrocarbons and at least three of five for the aza derivatives. There are two entries in Table VIII and one in Table X that are supported by two indices. Disconnections **5** $\Rightarrow$ **7** and **15** $\Rightarrow$ **50** are involved in ties, and **9** $\Rightarrow$ **13** is ranked fourth essentially by default, since the assignments for third and fifth are very secure (see Results). No ranking is supported by only one index.

In all but two cases, at least one of the indices is independent of symmetry. One exception is **5** $\Rightarrow$ **8**, which we tentatively put in a tie with **5** $\Rightarrow$ **7** for #2/#3, since all the symmetry-dependent indices favor the former and all the symmetry-independent ones the latter. Since there are three indices that depend on symmetry (N_S, $N_S(\text{lpe})$ and *wcx*) and two that do not ($N_T(\text{lpe})$ and *twc*) when a heteroatom is present, it is not valid to average them or to take a majority vote. The other exception with this kind of dichotomy is **15** $\Rightarrow$ **17**, which is ranked #3 by the symmetry-independent indices and #4 by the symmetry-dependent ones.

Recently, simplifying principles SP1-3 were derived for monocyclic systems based on subgraph counts.[12a] Essentially the same conclusions can be drawn from the behavior of *twc* observed previously.[8a] Four additional SP can be deduced from the above examples, as well as previous ones.[8a,12a] Then, we can collect SP1-7 in the list below.

SP1: Simplification increases with the number of components of equal complexity that result from a disconnection.

SP2: For alternative disconnections of a target into a given number of components, simplification increases as their complexities approach equality.

SP3: Breaking a heteroatom-carbon bond affords greater simplification than breaking a corresponding carbon-carbon bond.

SP4: Preserving the larger ring affords the greater simplification.

SP5: Breaking a bond to the more branched atom affords the greater simplification.

SP6: Breaking a C–C bond in a bridge with a heteroatom gives greater simplification than breaking an analogous C–C bond in a bridge without one.

SP7: Breaking the C–C bond closer to a heteroatom affords the greater simplification.

As far as SP1 is concerned, a smaller number of components of equal complexity can be simpler than a larger number of unequal complexity.[12a] For a given number of components, the maximum simplification obtains when they are equal in complexity, which is the logical extension of SP2. SP1 and SP2 encompass the heuristic of convergence, which as Warren has pointed out, is based on 'all other things being equal.'[13c] When breaking a bond to the less branched atom gives a symmetrical precursor and breaking a bond to the more branched one does not (*cf*. **5** $\Rightarrow$ **8** and **5** $\Rightarrow$ **7**, respectively), indices that are sensitive to symmetry may give a different result from symmetry-independent ones. Furthermore, if one of the atoms is a heteroatom, then the results can depend on the sensitivity of the index to the heteroatom effect (*cf*. **28** $\Rightarrow$ **29** and **28** $\Rightarrow$ **30**). Hence, we understand SP5 in the context of 'all other things being equal.' Similarly, the bridges in SP6 must be the same length for the bonds in them to be strictly analogous. For the purposes of SP6, the rings in spiro systems such as **33** are considered to be bridges. This SP is necessary for cases in which there are bonds in different bridges an equal distance from a N atom (*cf*. **33** $\Rightarrow$ **35** and **33** $\Rightarrow$ **38**).

We do not have an example of it here, but SP3 can be extended to heteroatom-heteroatom bonds, which give greater simplification upon disconnection than the corresponding heteroatom-carbon or carbon-carbon bonds. SP7 can be thought of as an extension of SP3 from bonds incident to a heteroatom to those more distant from it. Thus, while not as favorable as C–N bonds, C–C bonds can be ranked according to their proximity to N: an α,β-bond is more favorable than a β,β-bond, which is more favorable than a β,γ-bond, and so on. In order to be as general as possible, the bonds compared in SP7 do not have to be in the same ring or even in the same bridge. When comparing two disconnections, SP7 is only relevant if each of them breaks a C–C bond, *e.g*., if one of them breaks a C–C and the other a C–N bond, then SP7 does not apply (*cf*. **15** $\Rightarrow$ **19** and **15** $\Rightarrow$ **20**). To promote ease of application, if two disconnections break

the same kind of bond, preserve the larger ring, reduce branching by the same degree, *etc*., then the relevant SP can be listed for both, since it cancels out when they are compared (*vide infra*).

It is possible to establish a partial order for the SP based on the disconnections in Tables VIII–X.[27] Taking the three one-bond disconnections of **1** as an example, #1 is consistent with SP4 + SP5, #2 with SP5 and #3 with none of the SP ($\phi \equiv$ the null set). The 'plus sign' (+) stands for »and«, the 'greater than sign' (>) for »is more important than« or »outweighs«, and the 'equal sign' (=) for »is of equal importance to«. For the purpose of evaluation, the SP can be treated like physical quantities with magnitudes > 0.

In the case of **1**, it is easy to see that the SP comport with the order induced by the indices: the first and second entries give SP4 + SP5 > SP5, which reduces to SP4 > ϕ; the first and third give SP4 + SP5 > ϕ; and the second and third give SP5 > ϕ, all three of which are obviously true statements. While they were formulated with the aid of indices of complexity, the SP are independent of them, and inequalities (>) with the null set on the right side confirm the order of simplification induced by the indices.

The first and second entries for **5** give SP3 + SP4 + SP5 > SP5, which reduces to SP3 + SP4 > ϕ. The 7 in parentheses in the second entry is not comparable to the SP in the first, but is to those in the third. Thus, the second and third entries give SP5 + SP7 > ϕ, assuming that **5** $\Rightarrow$ **7** is more simplifying than **5** $\Rightarrow$ **8**. This inequality is obviously true, which confirms the assumption and breaks the tie in favor of the former. The first and third entries give SP3 + SP4 + SP5 > ϕ. Thus, the SP order all three disconnections, even though the indices fail to do so.

Taking the five one-bond disconnections of **9** as a more sophisticated example, there are ten (5 choose 2) possible pairs, each leading to an inequality, as there are no ties, *e.g*., SP3 + SP4 + SP5 > SP3 + SP5, which reduces to SP4 > ϕ; SP3 + SP4 + SP5 > SP4 + SP5, which reduces to SP3 > ϕ; SP3 + SP4 + SP5 > SP5, which reduces to SP3 + SP4 > ϕ; *etc*. The last two entries for **9** give SP5 > SP7, which is the only source of this pairwise ordering. It is important to note that SP5 indicates **9** $\Rightarrow$ **13** is more simplifying than **9** $\Rightarrow$ **14**, but SP7 indicates the opposite. Thus SP5 trumps SP7, provided that the order induced by the indices is correct. All things considered, the case for the order listed is a very strong one (*vide supra*).

For **15** there are 28 (8 choose 2) pairs of disconnections and (owing to a tie) 27 inequalities, *e.g*., SP3 + SP5 > SP3, which reduces to SP5 > ϕ; SP3 + SP5 > SP5 + SP6, which reduces to SP3 > SP6; SP3 + SP5 > SP4 + SP5, which reduces to SP3 > SP4; *etc*. The tie between **15** $\Rightarrow$ **20** and **15** $\Rightarrow$ **19** does not signify that SP3 = SP5 + SP6; it merely means that, based on the indices used, we are not able to decide which precursor is simpler in this case. However, based on the results in Table IX, we have SP3 > SP5 + SP6 (*cf*. first and second entries for **28** and

second and third for **33**), which breaks the tie in favor of the former. Also, from this inequality we can conclude that SP3 > SP5 and SP3 > SP6. Furthermore, from the fourth and fifth entries for **33**, we derive SP5 > SP6, which is the only source of this inequality.

Continuing with **15**, the second and fourth entries give SP3 > SP4 + SP5, which reduces to SP3 > SP4 and SP3 > SP5. The third and fourth entries give SP5 + SP6 > SP4 + SP5, which reduces to SP6 > SP4, the only source of this inequality. From this result and SP5 > SP6 (previous paragraph) we have SP5 > SP4. Finally, the fifth and sixth entries give SP5 + SP7 > SP4 + SP5, which reduces to SP7 > SP4, the only source of this inequality. From this result and those for **9**, *viz*. SP5 > SP7 (*vide supra*), we have a duplicate derivation of SP5 > SP4.

In the two-bond disconnections it is possible for a SP to come into play twice, *e.g*., from the first two entries for **9** (Table X), we have SP2 + SP3 + SP5 + SP5 > SP3 + SP4 + SP5 + SP5 + SP7, which reduces to SP2 > SP4 + SP7 and then to SP2 > SP4 and SP2 > SP7. In this case SP2 favors **9** $\Rightarrow$ **45** and SP7 favors **9** $\Rightarrow$ **46**; nevertheless, all five indices that are sensitive to heteroatoms agree that the former is more simplifying, which allows us to conclude with confidence that SP2 trumps SP7. In the case of **15** $\Rightarrow$ **50** and **15** $\Rightarrow$ **45**, the indices based on walk counts all favor the former, and those based on subgraph counts all favor the latter, so that it is not possible to objectively apply SP2. Assuming that the former is better, we have SP3 + SP5 + SP6 + SP7 > SP5 + SP5, which reduces to SP3 + SP6 + SP7 > SP5. Since we have established that SP3 > SP5 (*vide supra*), this statement is true and the assumption is valid; thus, we can break the tie in favor of the former.

Collecting the orders established for pairs of SP, we have at least two examples to guarantee that SP3 > SP4-6 and SP5 > SP4. There are single derivations of SP5 > SP6, SP5 > SP7, SP6 > SP4 and SP7 > SP4, which are not as secure and require more study. Since SP7 is an extension of SP3 to bonds farther from a heteroatom (*vide supra*), we have SP3 > SP7. (*N.B*., they could be combined into one SP.) Given these pairwise orderings, we have the hierarchy SP3 > SP5 > SP6, SP7 > SP4. In general, SP1 and SP2 cannot be ordered (*vide supra*), and from the information available here, there is no way to order SP6 and SP7. We conjecture that SP1, SP2 > SP3, but cannot prove it from the present examples, since whenever N is present, the two most equal components in terms of complexity are obtained by breaking at least one C–N bond. While it is of primary importance for parent compound **1**, SP4 tends to be secondary when a heteroatom is present; nevertheless, it serves as the distinguishing factor in several cases (*cf*. Table VIII). Most importantly, the SP can be applied even when the indices fail to order two disconnections, as illustrated by breaking all three ties in Tables VIII–X.

As part of LHASA,[13] a number of rules have been promulgated to guide the selection of *strategic bonds* for the disconnection of polycyclic systems. Ideally, these bonds would be broken first in the retrosynthetic analysis of a target, or equivalently they would be formed last in the synthetic scheme:

Rule 1: A strategic bond must be in a four, five, six, or seven-membered primary ring. (A primary ring is one that cannot be expressed as the envelope of two or more smaller rings bridged or fused together.)

Rule 2: A strategic bond must be directly attached (*exo*) to another ring (Rule 2A), but not *exo* to a three-membered ring (Rule 2B).

Rule 3: Strategic bonds should be in rings that exhibit the greatest degree of bridging, provided they are less than eight membered.

Rule 4: Any bond common to a pair of bridged or fused primary rings whose envelope is an eight-membered or larger ring is not considered strategic.

Rule 5: Bonds within aromatic rings are not considered strategic.

Rule 6: If a cyclic arc linking a pair of common atoms contains an asymmetric carbon atom, then none of the bonds in the arc is strategic (Rule 6A); however, the bonds directly attached to the asymmetric atom are strategic when it is the only one in the arc (Rule 6B).

Rule 7: To the set of strategic bonds determined by application of rules 1–6 above is added the collection of bonds in the cyclic network between C and O, N or S which satisfy rules 2B, 4, 5, and 6.

These rules are taken from the original paper,[13a] and it should be noted that they are presented somewhat differently in the subsequent book.[13b] Rules 1 and 2 are combined into a new criterion 1, rule 4 is renumbered as criterion 2, rules 3 and 5 are preserved essentially unchanged in criteria 3 and 5, respectively, rule 7 is recast as criterion 6, and rule 6 is replaced by a new criterion 4: »If the disconnection of a bond found to be strategic by criteria 1–3 [rules 1–4 above] produces a new ring appendage bearing stereocenters, those centers should be removed if possible (by stereocontrolled transforms) before the disconnection is made.« No bonds are forbidden by the new criterion 4, as they were by the old rule 6, presumably because of progress in stereocontrolled synthesis. Also, the new criterion 6, »heterobonds involving O, N and S which span across or otherwise join fused, spiro or bridged rings are strategic for disconnection, whether or not in a ring of maximum bridging,«[13b] appears to be less restrictive than the old rule 7.

Rule 7 was added to the six constitutional rules because of the relative ease of formation of heteroatom-carbon bonds compared to carbon-carbon bonds. It fulfills the same function as our SP3, but is limited by the codicil, »which satisfy rules 2B, 4, 5, and 6.« It may be noted

that rules 1, 2B, 4, 6 and 7 were based on the state of the art of synthetic chemistry when they were written, *e.g.*, eight-membered rings were more difficult to prepare at that time than they are today (*cf.* rule 4).[28] Rule 5 affords aromatic compounds special status out of consideration for their ready availability, and it also recognizes the importance of the convergent assembly of subunits. Our SP2 is, in fact, a statement of convergence (*vide supra*). Rule 2A is effective because it reduces branching, which is the explicit goal of SP5. We analyze the LHASA rules in more detail in a companion paper.[29]

The strategic bonds in bicyclo[2.2.1]heptane (**1**) (bold lines in Table I) are C1–C2, C1–C6, C1–C7, C3–C4, C4–C5 and C4–C7. Since they are not *exo* to another ring (rule 2), C2–C3 and C5–C6 are not strategic. The LHASA rules allow only a 'yes or no' answer to the question, »Is this bond strategic?« Where more than one strategic bond is identified, they do not tell us which is the most strategic. Our method allows us to assign a numerical value to the degree of simplification and thus to rank strategic bonds, *e.g.*, in the case of **1**, the strategic bonds are stratified into #1 (C1–C7, C4–C7) > #2 (C1–C2, C1–C6, C3–C4, C4–C5). The non-strategic bonds (C2–C3, C5–C6) are last (#3).

In the case of **5**, no new strategic bonds are added by rule 7, as C1–N7 and C4–N7 are already strategic by rules 1–6. They are #1 and the remaining strategic bonds (C1–C2, C1–C6, C3–C4, C4–C5) are #2. Likewise, the three C–N bonds in **9** are already strategic. Our method stratifies them #1 (N1–C7) > #2 (N1–C2, N1–C6) ahead of the strategic C–C bonds #3 (C4–C7) > #4 (C3–C4, C4–C5). Again, the non-strategic bonds (C2–C3, C5–C6) are last (#5). In the case of **15**, one of the C–N bonds (C1–N2) is strategic by rules 1–6 and the other (N2–C3) is made strategic by rule 7. Our method differentiates all the strategic bonds: #1 (C1–N2) > #2 (N2–C3) > #3 (C3–C4) > #4 (C1–C7) > #5 (C1–C6) > #6 (C4–C7) > #7 (C4–C5), and the non-strategic bond (C5–C6) is last (#8).

In the case of spiro[3.3]heptane (**22**), the four equivalent bonds to the quaternary center (C4) are strategic (bold lines in Table IV); therefore, the C–N bonds in its 4-azonia derivative **25** are strategic without rule 7. In both cases the strategic bonds are rated #1 by our method, and the non-strategic ones are #2. In **28** both C–N bonds are made strategic by rule 7. They are rated #1 (C1–N2, N2–C3), the strategic C–C bonds are #2 (C1–C4, C3–C4) > #3 (C4-C5, C4-C7), and the non-strategic bonds are #4 (C5-C6, C6-C7). LHASA rules 1–6 elevate **33** ⇒ **34**, **33** ⇒ **35** and **33** ⇒ **38** equally as strategic bond disconnections, and rule 7 adds **33** ⇒ **36** on an equal basis. They are differentiated **34** > **36** > **35** > **38** by our mathematical method, as the strategic bonds are stratified #1 (N1–C4) > #2 (N1–C2) > #3 (C3–C4) > #4 (C4–C5, C4–C7), and the non-strategic bonds are #5 (C5–C6, C6–C7). In this case the C–N bond made strategic by rule 7 is #2.

As far as the best two-bond disconnections are concerned (Tables VII and X), three of the four are Diels-Alder disconnections, *viz*. **1 ⇒ 42**, **9 ⇒ 45** and **15 ⇒ 51**. The exception is **5 ⇒ 43**, which breaks two C–N bonds, and in this case the Diels-Alder disconnection, **5 ⇒ 45**, is ranked second. Two Diels-Alder disconnections are possible for **15**; the more favorable one, **15 ⇒ 51**, is ranked first, and the less favorable, **15 ⇒ 45**, is tied for third/fourth. The Diels-Alder reaction has been used to synthesize a number of bicyclo[2.2.1]heptane derivatives.[30]

CONCLUSION

Indices of molecular complexity can be used to calculate the degrees of simplification resulting from alternative disconnections of a target, and thus they can be used to order the disconnections. Simplification principles can be derived, which are consistent with the orders of disconnections induced by the indices, and they are very useful generalizations, since they can be applied even when the indices fail to yield a consensus. For example, they allow us to break the tie between **5 ⇒ 7** and **5 ⇒ 8** in favor of the former and the tie between **15 ⇒ 19** and **15 ⇒ 20** in favor of the latter. Furthermore, the simplification principles do not depend on the state of the art of synthetic chemistry and will not require modification over time like some of the LHASA rules, *e.g.*, rule 4 as eight-membered rings become easier to prepare and rule 6 as progress is made in asymmetric synthesis. All things considered, the agreement between our results, which are based on pure mathematics, and those based on 'logic and heuristics' is remarkably good. This observation suggests that the LHASA rules and the synthesis experience that produced them may have a mathematical (topological) basis. The same is true for the heuristic of convergence.

In 1997, David Bradley observed,[31] »It is rather astonishing how little chemists really know about designing the shortest and simplest synthesis of a new molecule.« In this regard, not much has changed since 1997 – or even since 1956,[32] when R. B. Woodward gave synthetic chemists their marching orders: »Synthesis must always be carried out by plan, and the synthetic frontier can be defined only in terms of the degree to which realistic planning is possible, utilizing all of the intellectual and physical tools available.« With the perfection of chromatographic methods for the purification of chemical compounds and spectroscopic methods for their identification in the second half of the twentieth century, synthetic chemists now have an impressive armamentarium of physical tools. The first half of the twenty-first century will be the era of intellectual tool development, and we believe that topology will play a key role in this endeavor.

REFERENCES

1. For part 22, see Ref. 29a. For part 21, see Ref. 29b.
2. H. Wiener, *J. Am. Chem. Soc*. **69** (1947) 17–20.
3. (a) S. Nikolić, N. Trinajstić, I. M. Tolić, G. Rücker, and C. Rücker, in: D. Bonchev and D. H. Rouvray (Eds.), *Complexity in Chemistry: Introduction and Fundamentals*, Taylor and Francis, London, 2003, pp. 29–89; (b) S. Nikolić, I. M. Tolić, N. Trinajstić, and I. Baučić, *Croat. Chem. Acta* **73** (2000) 909–921; (c) D. Bonchev and N. Trinajstić, *J. Chem. Phys*. **67** (1977) 4517–4533; (d) I. Gutman, B. Ruščić, N. Trinajstić, and C. F. Wilcox, *J. Chem. Phys*. **62** (1975) 3399–3405; (e) See also Ref. 21.
4. (a) M. Randić and D. Plavsić, *Int. J. Quant. Chem*. **91** (2003) 20–31; (b) M. Randić, *J. Mol. Graph. Modell*. **20** (2001) 19–35; (c) M. Randić, *J. Chem. Inf. Comput. Sci*. **37** (1997) 672–687; (d) see also Ref. 11.
5. (a) A. T. Balaban, *Chem. Phys. Lett*. **89** (1982) 399–404; (b) A. T. Balaban, *Pure Appl. Chem*. **55** (1983) 199–206; (c) O. Mekenyan, D. Bonchev, and A. T. Balaban, *J. Math. Chem*. **2** (1988) 347–375; (d) see also Ref. 24.
6. (a) D. Bonchev, *J. Chem. Inf. Comput. Sci*. **40** (2000) 934–941; (b) D. Bonchev, in: J. Devillers and A. T. Balaban (Eds.), *Topological Indices and Related Descriptors*, Gordon and Breach, Reading, UK, 1999, pp. 361–401; (c) A. V. Zeigarnik, O. N. Temkin, and D. Bonchev, *J. Chem. Inf. Comput. Sci*. **36** (1996) 973–981; (d) D. Bonchev, V. Kamenska, and O. Mekenyan, *J. Math. Chem*. **5** (1990) 43–72; (e) see also Refs. 3c, 5c and 9b.
7. (a) S. H. Bertz, in: D. Bonchev and D. H. Rouvray (Eds.), *Complexity in Chemistry: Introduction and Fundamentals*, Taylor and Francis, London, 2003, pp. 91–156; (b) S. H. Bertz and C. M. Zamfirescu, *MATCH – Commun. Math. Comput. Chem*. **42** (2000) 39–70; (c) S. H. Bertz and W. F. Wright, *Graph Theory Notes of New York*, NY Acad. Sci. **XXXV** (1998) 32–48; (d) S. H. Bertz and T. J. Sommer, *Chem. Commun.* (1997) 2409–2410; (e) S. H. Bertz and W. C. Herndon, in: T. H. Pierce and B. A. Hohne (Eds.), *Artificial Intelligence Applications in Chemistry*, ACS Symp. Ser., Vol. 306, Am. Chem. Soc., Washington, DC, 1986, pp. 169–175; (f) S. H. Bertz, *Discrete Appl. Math*. **19** (1988) 65–83; (g) S. H. Bertz, *Bull. Math. Biol*. **45** (1983) 849–855; (h) S. H. Bertz, *J. Chem. Soc., Chem. Commun.* (1981) 818–820; (i) S. H. Bertz, *Chem. Commun.* (2001) 2516–2517.
8. (a) G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci*. **40** (2000) 99–106; (b) I. Gutman, C. Rücker, and G. Rücker, *J. Chem. Inf. Comput. Sci*. **41** (2001) 739–745; (c) G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci*. **33** (1993) 683–695; (d) G. Rücker and C. Rücker, *MATCH – Commun. Math. Comput. Chem*. **41** (2000) 145–149; (e) G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci*. **41** (2001) 314–320, 865; (f) G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci*. **41** (2001) 1457–1462; (g) see also Ref. 3a.
9. (a) S. C. Basak, S. Nikolić, N. Trinajstić, D. Amić, and D. Beslo, *J. Chem. Inf. Comput. Sci*. **40** (2000) 927–933; (b) D. Bonchev, *SAR & QSAR Environ. Res*. **7** (1997) 23–43; (c) D. H. Rouvray, *J. Comput. Chem*. **8** (1987) 470–480.
10. (a) M. Gordon and J. W. Kennedy, *J. Chem. Soc., Faraday Trans. II* **69** (1973) 484–504; (b) J. W. Essam, J. W. Kennedy, M. Gordon, and P. Whittle, *J. Chem. Soc., Faraday Trans. II* **73** (1977) 1289–1307; (c) A. Graovac, I. Gutman, M. Randić, and N. Trinajstić, *Colloid Polym. Sci*. **255** (1977) 480–487.

11. (a) M. Randić, *J. Am. Chem. Soc.* **99** (1977) 444–450; (b) M. Randić, *Tetrahedron* **33** (1977) 1905–1920; (c) see also Ref. 10c.

12. (a) S. H. Bertz, *New J. Chem.* **27** (2003) 860–869; (b) S. H. Bertz, *New J. Chem.* **27** (2003) 870–879; (c) S. H. Bertz and T. J. Sommer, *Chem. Commun.* (2003) 1000–1001; (d) S. H. Bertz and T. J. Sommer, in: T. Hudlicky (Ed.), *Organic Synthesis: Theory and Applications*, Vol. 2, JAI Press, Greenwich, CT, 1993, pp. 67–92; (e) S. H. Bertz, *J. Chem. Soc., Chem. Commun.* (1986) 1627–1628; (f) S. H. Bertz, *J. Chem. Soc., Chem. Commun.* (1984) 218–219; (g) S. H. Bertz, *J. Am. Chem. Soc.* **104** (1982) 5801–5803; (h) S. H. Bertz, *J. Am. Chem. Soc.* **103** (1981) 3599–3601; (i) see also Refs. 7a–d.

13. (a) E. J. Corey, W. J. Howe, H. W. Orf, D. A. Pensak, and G. Petersson, *J. Am. Chem. Soc.* **97** (1975) 6116–6124; (b) E. J. Corey and X.-M. Cheng, *The Logic of Chemical Synthesis,* Wiley, New York, 1989; see also (c) S. Warren, *Organic Synthesis: The Disconnection Approach*, Wiley, Chichester, 1982, p. 352.

14. (a) M. Kim, K. Kawada, R. S. Gross, and D. S. Watt, *J. Org. Chem.* **55** (1990) 504–511; for other Diels-Alder approaches, see (b) T. K. M. Shing and Q. Jiang, *J. Org. Chem.* **65** (2000) 7059–7069 (intramolecular); (c) N. Stojanac and Z. Valenta, *Can. J. Chem.* **69** (1991) 853–855; (d) G. Vidari, S. Ferriño, and P. A. Grieco, *J. Am. Chem. Soc.* **106** (1984), 3539–3548 (first total synthesis).

15. W.-D. Ihlenfeldt and J. Gasteiger, *Angew. Chem., Int. Ed. Engl.* **34** (1995) 2613–2633.

16. E. L. M. van Rozendaal, M. A. Ott, and H. W. Scheeren, *Recl. Trav. Chim. Pays-Bas* **113** (1994) 297–303.

17. N. Hall, *Chem. Commun.* (2003) 661–664.

18. K. C. Nicolaou, P. G. Nantermet, H. Ueno, R. K. Guy, E. A. Couladouros, and E. J. Sorensen, *J. Am. Chem. Soc.* **117** (1995) 624–633.

19. (a) F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1969; (b) C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, 1949; (c) D.I.A. Cohen, *Basic Techniques of Combinatorial Theory* (Wiley, New York, 1978).

20. J. B. Hendrickson and P. Huang, *J. Chem. Inf. Comput. Sci.* **29** (1989) 145–151.

21. N. Trinajstić, *Chemical Graph Theory*, 2nd edn., CRC Press, Boca Raton, FL, 1992.

22. Sometimes, bonds must be formed in the retrosynthetic direction for strategic reasons, *e.g.*, the retro-Diels-Alder reaction has been used synthetically, and its transform is the Diels-Alder reaction, which has a relatively large positive value of ΔC. See also Ref. 12a.

23. An analogy is cities, represented by points on a map, and the roads that connect them, represented by lines. When there is one road to a city, there is no choice of routes; however, when there are two roads, a decision must be made as to which route to take, making the problem more complex. The number of choices increases factorially as roads are added.

24. A. T. Balaban, D. Farcasiu, and R. Bănică, *Rev. Roum. Chim.* **11** (1966) 1205–1227.

25. J. B. Hendrickson, *J. Am. Chem. Soc.* **99** (1977) 5439–5450. See also Refs. 12e, f.

26. Originally, a factor of ½ was included in the formula for *twc* to compensate for the fact that each walk is counted twice without it (Refs. 8a, c). It was later dropped for other reasons (Refs. 3a and 8b).

27. For a 'Prolegomenon on Partial Orderings in Chemistry,' see D. J. Klein, *MATCH – Commun. Math. Comput. Chem.* **42** (2000) 7–21.

28. (a) G. Mehta and V. Singh, *Chem. Rev.* **99** (1999) 881–930; (b) P. A. Wender and B. L. Miller, in: T. Hudlicky (Ed.), *Organic Synthesis: Theory and Applications*, Vol. 2, JAI Press, Greenwich, CT, 1993, pp. 27–66.

29. (a) C. Rücker, G. Rücker, and S. H. Bertz, *J. Chem. Inf. Comput. Sci.* **44** (2004) 378–386; (b) S. H. Bertz, C. Rücker, G. Rücker, and T. J. Sommer, *Eur. J. Org. Chem.* (2003) 4737–4740.

30. (a) D. R. Wilson, US Patent 3,459,684 (1-aza); (b) D. Blondet and C. Morin, *Heterocycles* **19** (1982) 2155–2182 (2-aza); (c) H.-J. Altenbach, D. Constant, H.-D. Martin, B. Mayer, M. Müller, and E. Vogel, *Chem. Ber.* **124** (1991) 791–801 (7-aza).

31. D. Bradley, *New Scientist*, November 15 (1997) 40–43.

32. R. B. Woodward, in: A. Todd (Ed.), *Perspectives in Organic Chemistry*, Interscience, New York, 1956, pp. 155–184.

SAŽETAK

U potrazi za pojednostavljenjem: uporaba topologijskih indeksa kompleksnosti za vođenje retrosintetičke analize

Steven H. Bertz i Christoph Rücker

Topologijski indeksi kompleksnosti N_S, N_T, N_S(lpe), N_T(lpe), *twc* i *wcx* upotrebljeni su za rangiranje struktura, koje se dobiju uklanjanjem jedne veze iz biciklo[2.2.1]heptana, spiro[3.3]heptana i njihovih aza derivata, prema stupnju pojednostavljenja. Također su ispitane strukture, koje se dobiju kada se uklone dvije veze iz biciklo[2.2.1]heptana i njegovih aza derivata. Izvedena su načela pojednostavljenja, koja su upotrebljiva za vođenje retrosintetičke analize ciljanih kompleksnih molekula. Usporedba s pravilima LHASA za pronalaženje strateških veza pokazala je da su predložena metoda i LHASA u velikoj mjeri slične, ali da među njima postoje neke značajne razlike.