Information Content in Organic Molecules: Reaction Pathway Analysis via Brownian Processing

Daniel J. Graham* and Matthew V. Schulmerich

Department of Chemistry, Loyola University of Chicago, 6525 North Sheridan Road, Chicago, Illinois 60626

Received February 16, 2004

Carbon chemistry offers infinite possibilities for molecules as information carriers. Moreover, there is no boundary on the number of ways in which a carrier's information can change via a chemical reaction. Organic reaction pathways thus pose new types of informatic variables which compel characterization using Brownian methods. We apply the tools of the preceding paper to these variable types for select reactions and classes. Along the way, geometric descriptors are formulated which complement the structure graph sequences of chemistry texts and journals. In addition, the statistical structure underpinning carbon transformations is explored more deeply. Overall, this work brings to light several informatic principles of organic reactions. Knowledge of these can assist in synthetic designs on both large- and small-scales. Brownian methods are able to address the pathway structures of all organic reactions: those reported in the literature, explored in the lab, or in the developmental stage.

I. INTRODUCTION

Figure 1 illustrates a sampling of chemistry lab events described by two-dimensional (2D) structure graphs. In each case, the information expressed by the left-hand-side compound(s) is not conserved but is rather enhanced or diminished via a transformation leading to the right-hand-side. In more than one respect, organic molecules share characteristics with the logic gates (e.g. NAND, Figure 2) of everyday computing. To name a few, the systems represented in both Figures 1 and 2 transform electronic information in a discontinuous manner. Molecules and logic circuitry both support input/output (I/O) channels whose flux is governed by material composition and environment. Logic gates and molecules offer serial and combinatorial assemblies with unlimited diversity; these are able to effect information changes over short time scales at exceedingly low error rates.[1−3] As a physical quantity, information enjoys many storage and communication settings. Its beauty and power, molecular or otherwise, lie arguably in its mutability.

For an organic molecule, the information expressed via the covalent bond network is intimately tied to chemical reactivity. This latter property can be viewed as the collective response functions wired into a given molecule and dependent on neighboring molecules. For example, Figure 3 illustrates the correlated information (CI) obtained by Brownian processing versus structure order for benzene and toluene molecules. The CI disparities, both in scaling and magnitude, reflect the functional differences of these and related aromatics. As is well-known, benzene and toluene respond in markedly different ways to compounds such as alkylating agents.[4] It is the network-encoded information of these ring systems which directs the different chemical responses.

The preceding paper focused on organic molecules as information carriers.[5] Brownian/random walk processing

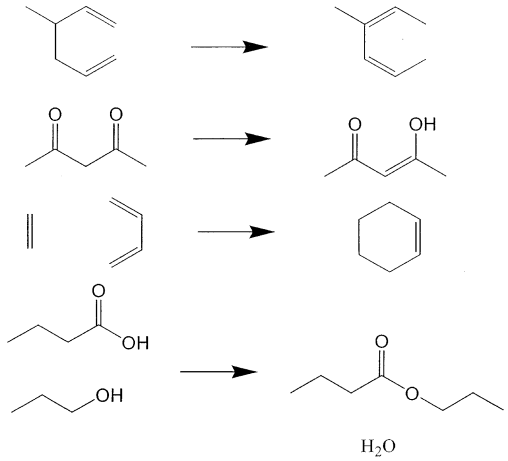

Figure 1. Structure graphs for organic reaction sequences. Four sample reactions are represented via molecular graphs in a 2D format.

enabled CI-quantification to several structure orders according to two distinct processing modes. Such a methodology was exercised for both individual molecules, as articulated by structure graphs and several functional libraries. In so doing, new insights were obtained regarding the information content of carbon-based systems at the covalent bond level. A molecule's information content is of critical importance because it predicates a capacity for work control at the Angstrom scale.

Information-wise, organic molecules offer infinite possibilities. Moreover, there is no boundary on the number of ways by which this quantity can change during chemical reactions. Accordingly, reactions pose new types of informatic variables which compel characterization using Brownian methods. In this paper, we address these variables for several well-known organic reactions and classes. Along the way, geometric descriptors are introduced which complement graph sequences such as in Figure 1, chemistry texts, and

* Corresponding author fax: (773)508-3086; e-mail: dgraha1@luc.edu.

10.1021/ci040022v CCC: $27.50
Published on Web 07/22/2004

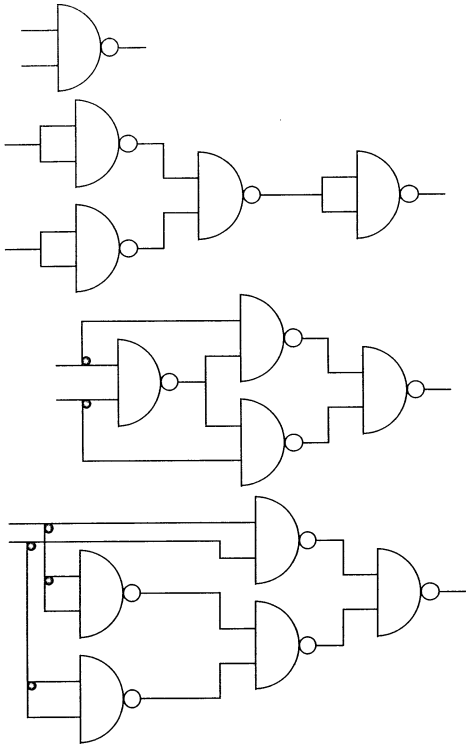

Figure 2. NAND logic gate assemblies. Four sample assemblies of NAND gates demonstrating series and combination structures.

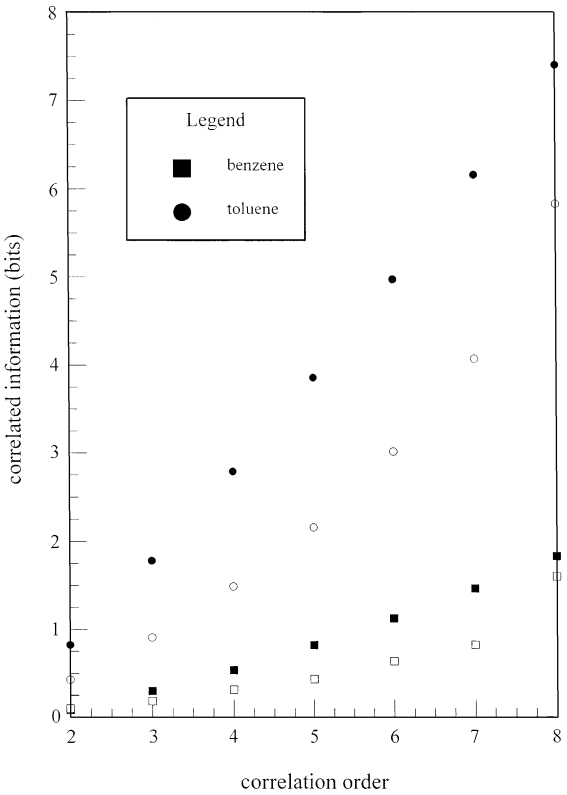

Figure 3. Correlated information for benzene and toluene molecules. Illustrated are CI-values versus structure order. Data for Brownian serial and parallel processing are marked using filled and open symbols, respectively.

journals. In addition, the statistical structure underpinning carbon chemistry is probed more deeply, as it is tied closely to the mutability of information. Synthetic methods are inspired by nature, the literature, and individual exploration. Like the preceding work, this paper focuses on the probabilistic framework surrounding these lab events, with important principles identified along the way. Knowledge of the informatic principles can assist organic reaction design on a variety of scales.

Not surprisingly, the connections between molecules and information processing command a diverse, fascinating literature. To cite a few representative works, Conrad has designed a variety of laboratory interfaces involving the information processing of enzymes.[6] Conrad and Zuner have, in turn, examined important features of molecular computing and pattern processing by enzymes.[7] Along the same lines, Adelman, Schneider, and quite a few other researchers have explored the computational capabilities of nucleic acids.[8−10] Using reaction−diffusion media, Ross, Showalter, and their co-workers have effected a variety of computational operations involving finite state machines and logic gates.[11,12] These are but a sampling of pivotal studies which have extended the informatic territory explored by Turing and others in the 1940s and 1950s.[13] The reader is directed to Bennett's review of the thermodynamics of computation for an in-depth, physical perspective of this subject. For our part, we concentrate here on the informatic territory at the covalent bond level.

II. INFORMATION PATHWAYS AND ORGANIC MOLECULES

Molecules in a reaction embody a type of I/O pathway by which information is transformed. The physical pathway is materialized in real solvents and portrayed (as with the logic gates of Figure 2) using 2D graphs. Yet there exists as well several abstract methods for pathway delineation. Accordingly, each graph such as in Figure 1 can be affiliated with a point in a space defined by the state variables of reactant and product compounds. Structure graphs convey what happens at the Angstrom scale; thermodynamic ideas offer coordinate systems in which to describe the transformations.[14]

For systems in general, the initial and terminal states of any transformation correspond to different points in planes such as enthalpy−entropy, free energy-temperature, pressure−volume, etc.[14] A particular locus of points can depict a continuous pathway connecting two states. Alternatively, a pair of points mediated by a line can portray a discontinuous pathway such as the case of certain phase transitions. These ideas are visualized in the upper panel of Figure 4. Open and filled boxes respectively represent the initial and terminal states of two sample pathways in the enthalpy−entropy (H,S) plane. The representations are abstract in the sense that they are well removed from the actual material systems. Even so, this portrait genre gives root to several transformation fundamentals, including a direction sense, operational distance measure, and scaling functionality. A variety of thermodynamic metrics have been investigated in recent years by Weinhold and Berry, Salomon, and coworkers.[15−17]

Things work in an analogous way for informatic pathways, the most day-to-day familiar of these involving computer electronics. For example, the input lines of each Figure 2 configuration express 2.00 information bits and an energy dependent on current flow. Each logic assembly then effects an information reduction of 1.00 bit over the pathway and

concomitant thermal dissipation. Figure 2 conveys the physical pathways of NAND logic; the lower Figure 4 panel represents the pathways in one of several possible abstract spaces. Here the state variables are enthalpy and information, the latter closely related to entropy.[18] As with thermodynamic representations in general, the abstract pathway structure evinces a certain direction, operational length, and functionality.

How does this analysis apply to reactions? Molecules express many of the transition characteristics of Figure 2; however, their information measures are not integer powers of 2. As shown in the preceding paper, the information for a given compound can be ascertained via Brownian/random walk processing applied to the structure graph.[5] In applying the same methods to chemical reactions, the mechanics work briefly as follows.

In the simplest reactions involving single molecules, one constructs an array/matrix for each I/O state. For example, a sigmatropic reaction (Figure 1, top sequence) offers two such graphs which can be labeled according to the Figure 5 scheme. Note that only the graph skeletons need to be coded directly as they are readily expanded to include hydrogen symbols. The sigmatropic arrays and matrices are represented as follows:

array entries	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	matrix column indices
C 1	0	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	
C 2	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	
C 3	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	
C 4	0	0	1	0	2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	
C 5	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	
C 6	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	1	0	
C 7	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	
H 8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 9	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 10	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	*(input state)*
H 11	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 12	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 13	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 14	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 15	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 16	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 17	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 18	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 19	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	

C 1	0	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	
C 2	1	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	
C 3	0	1	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
C 4	0	0	2	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
C 5	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	
C 6	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	0	
C 7	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	
H 8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 9	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 10	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	*(output state)*
H 11	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 12	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 13	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 14	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 15	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 16	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 17	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 18	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 19	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	

The above enable Brownian/random walks and tape recordings which can be analyzed for information content. The results include the CI-measures (serial or parallel processing) for the input and terminal states averaged over several structure orders.[5] When these values are combined with energy data such as bond enthalpy, one obtains an information pathway structure such as illustrated in the lower left panel of Figure 5. Just as in Figure 4, the open and filled boxes denote the input and terminal states, respectively; these indicate the bias or flux direction along the pathway. In this simple case, a single-ray asserts a well-defined length and slope/derivative functionality. One thereby observes the transition to be a constant-enthalpy one in which the average CI is diminished only by 0.16 bits. Evidently, the response

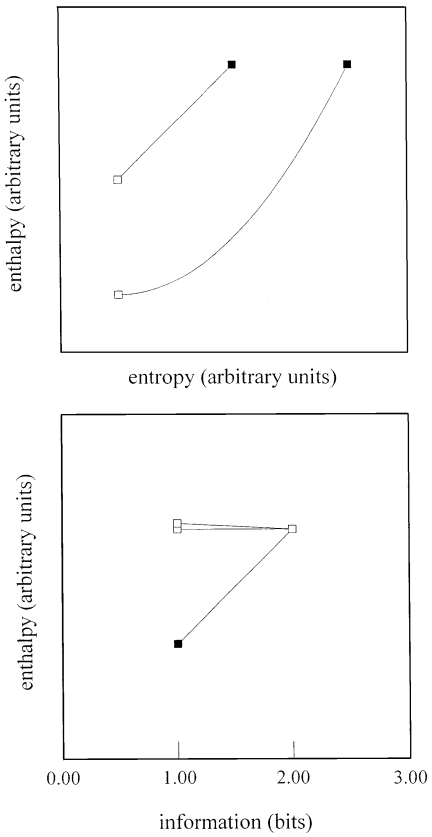

Figure 4. Pathway structures for sample I/O transformations. Upper panel illustrates pathway structures for two transformations in the enthalpy/entropy plane. Input and output state points have been marked using open and filled squares, respectively. Lower panel illustrates a pathway structure which is applicable to a NAND-type configuration of Figure 3. Input and output state points are marked, respectively, by open and filled boxes.

encoded by 3-methyl-1,5-hexadiene effects only a fractional adjustment of information when transiting to 3-methyl-2,4-hexadiene.

The pathway delineations are only slightly more involved for bimolecular and higher order reactions. Here one initiates the process by applying Brownian methods to all individual graphs of a given I/O channel. For example, the Diels−Alder transition (Figure 1, third sequence) poses two input graphs (ethylene and butadiene) which can be labeled according to the left-most scheme of Figure 5. Coding and expansion for each array/matrix yield

expansion for each array/matrix yield:

array entries	1	2	3	4	5	6	matrix column indices
C 1	0	2	1	1	0	0	
C 2	2	0	1	0	1	1	
H 3	1	1	0	0	0	0	
H 4	1	0	0	0	0	0	*(ethylene input state)*
H 5	0	1	0	0	0	0	
H 6	0	1	0	0	0	0	

array entries	1	2	3	4	5	6	7	8	9	10	matrix column indices
C 1	0	2	0	0	1	1	0	0	0	0	
C 2	2	0	1	0	0	0	1	0	0	0	
C 3	0	1	0	2	0	0	0	1	0	0	
C 4	0	0	2	0	0	0	0	0	1	1	
H 5	1	0	0	0	0	0	0	0	0	0	*(butadiene input state)*
H 6	1	0	0	0	0	0	0	0	0	0	
H 7	0	1	0	0	0	0	0	0	0	0	
H 8	0	0	1	0	0	0	0	0	0	0	
H 9	0	0	0	1	0	0	0	0	0	0	
H 10	0	0	0	1	0	0	0	0	0	0	

Brownian processing is then exercised for each encoded lattice. The next one combines the molecular graphs, the result being a single array/matrix albeit expressing no covalent linkage between the input-state compounds. Fol-

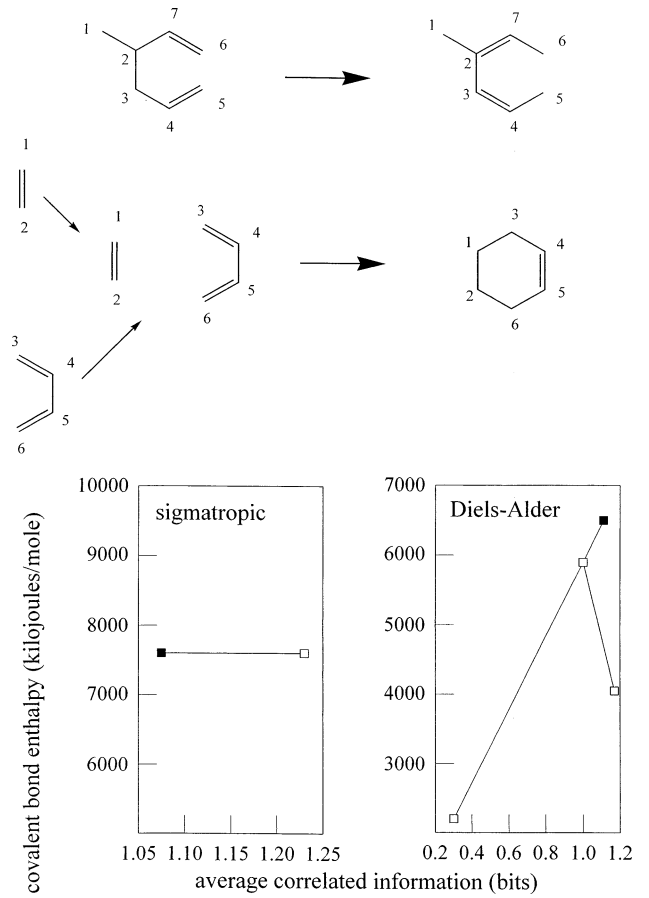

Figure 5. Graph sequences, coding, and informatic pathway structures. The examples of the sigmatropic and Diels−Alder sequences of Figure 1 are illustrated. The numbering used for constructing the array/matrix pairs is arbitrary. The lower panels illustrate the pathway structures obtained by Brownian processing. The information quantities have been obtained using serial processing and have been averaged over orders 2−8 with equal weight coefficients applied. Open and filled symbols distinguish the input and output states, respectively.

lowing the Figure 5 notation, one arrives at

array entries	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	matrix column indices
C 1	0	2	0	0	0	0	1	1	0	0	0	0	0	0	0	0	
C 2	2	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	
C 3	0	0	0	2	0	0	0	0	0	0	1	1	0	0	0	0	
C 4	0	0	2	0	1	0	0	0	0	0	0	0	1	0	0	0	
C 5	0	0	0	1	0	2	0	0	0	0	0	0	0	1	0	0	
C 6	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	1	
H 7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	*(butadiene/ethylene combination)*
H 9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 10	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 11	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 12	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 13	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
H 14	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
H 15	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
H 16	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	

for the ethylene/butadiene combination state. The output state pertains to a single molecule, cyclohexene, with the following array/matrix pair:

array entries	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	matrix column indices
C 1	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	
C 2	1	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	
C 3	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	
C 4	0	0	1	0	2	0	0	0	0	0	0	0	1	0	0	0	
C 5	0	0	0	2	0	1	0	0	0	0	0	0	0	1	0	0	
C 6	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1	
H 7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 10	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	*(output state)*
H 11	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 12	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
H 13	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
H 14	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
H 15	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
H 16	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	

Brownian methods are applied to both the combination and product structures with recorded message tapes established along the way. It is straightforward to adapt this procedure to reactions featuring multiple reactants and products, for example, the alkylation of substituted aromatics.

The results are abstract pathway structures as illustrated in the lower right-hand panel of Figure 5. Note how three intersecting rays, each asserting length and slope convey the informatic transition in the H/CI plane. The state points for ethylene and butadiene molecules (open squares) mark the input channels, while cyclohexene molecule (filled square) locates the output. The ethylene/butadiene combination state (middle open square) expresses CI midway between the two input channels. This state apparently connects to the output in a transition which enhances CI. Informatically speaking, a Diels−Alder transition offers one way to enhance (by about a factor of 5) the work control capacity of an ethylene molecule. Concomitantly, this transition enhances the bond enthalpy of butadiene by about 60%, with only modest diminution of CI, < 0.1 bit.

In organic reaction analysis, several descriptors follow naturally. These include (1) the length L over a single ray of the pathway, (2) the operational distance D equivalent to the sum total of the ray lengths, (3) the changes ΔH, ΔCI which incur for a given ray, and (4) the slope τ for a ray gauged by $\Delta H/\Delta CI$. L and D are positive definite, while ΔH, ΔCI, and τ assert both positive and negative values. These quantities are readily measured for all possible organic reactions following Brownian processing and informatic pathway delineation; they augment familiar graph sequences (e.g. Figure 1) in a new and simple way.

For pathway analysis to proceed further, however, certain dimensional issues must be reconciled. In particular, molecular CI is evaluated in bits, while bond enthalpies are measured (as in Figure 5) in kilojoules or kilocalories per mole. How can one establish the information pathways in a manner free of unit details?

The use of dimensionless/reduced variables proves beneficial in representing thermodynamic transforms. For example, reduced variables are time-honored when representing the corresponding states and transforms of different systems.[14] We find an analogous tack applicable to informatic pathways. Accordingly, enthalpy values can be scaled in units of k_BT per molecule; k_B and T refer in their usual way to Boltzmann's constant and temperature. For example, the **C**−**H** bond enthalpy averages 413 kilojoules per mole.[19] At $T = 298$ K, this corresponds to 413×10^3 joules/$6.02 \times 10^{23}/(1.38 \times 10^{-23}$ j/K $\times$ 298 K) = 167.

Further, CI for a molecule measures a departure from noise behavior during Brownian/random walk processing. Molecules offering noise-plagued signals assert little CI and sparse capacity to direct meaningful (nontrivial) work. Bennett has advanced the idea that correlated information can be viewed in fuel-equivalent terms: a highly correlated message tape can direct substantial work production by an engine (e.g. a computer) compared with a random tape.[1,2] In a similar vein, Brillouin has compared the correlated information within a system with a deficit of entropy and a corresponding work availability proportional to temperature.[20] Putting these thermodynamic ideas together, one can connect measures of CI with free energy or work equivalents: 1.00 correlated information bit signifies $k_BT\log_e 2$ of work control

Table 1. Informatic Pathway Descriptors for Figure 1 Reactions

	$\Delta\boldsymbol{CI}$	$\Delta\boldsymbol{H}$	$\boldsymbol{D}$	τ
sigmatropic	+0.0711	0.00	0.0711	0.00
tautomerization	+1.039	−26.6	26.6	−25.6
Diels−Alder	+0.122	+65.0	65.0	+266.4
condensation	−0.783	0.00	0.783	0.00

potential.[1,2,20] For our purpose, molecular CI can be rescaled in terms of reduced work equivalents (RWE) of k_BT. When this is combined with the enthalpy rescaling, all of the pathway variables become "reduced" or dimensionless. Among other tangibles, the rescaling enables ready discrimination of the enthalpic versus informatic contributions to pathway distance. Pathways based on reduced variables will be featured in the remaining figures of this paper. Morever the rescaled descriptors will be denoted using bold italics: $\boldsymbol{L}$, $\boldsymbol{D}$, $\Delta\boldsymbol{H}$, $\Delta\boldsymbol{CI}$, and $\tau = \Delta\boldsymbol{H}/\Delta\boldsymbol{CI}$. As examples, Figure 7 illustrates rescaled pathways for the Figure 1 reactions; Table 1 lists values for the pathway descriptors. These data show the Diels−Alder reaction to be the one of greatest informatic pathway length; the tautomerization of Figure 1 asserts the largest information change.

Real laboratory chemistry depends on conditions such as temperature, pressure, and solvent structure. A Brownian approach views organic reactions quite apart from these details. The focus is instead on the information properties intrinsic to chemical reactions, the results being pathway structures on a universal scale. The structures intrigue in several ways, even for elementary chemistry. Note, for example, the disparity of the ray-lengths and geometries for the Figure 1 reactions in Figure 6, all plotted using identical $\boldsymbol{H}/\boldsymbol{CI}$ scales. Note the contrast between the minuscule length in the sigmatropic transition and the length affiliated with tautomerization. Observe as well the diversity of pathway lengths and angles for the Diels−Alder and condensation reactions. However simple, the informatics offer fresh insights for chemical reactions. In particular, when Brownian methods are applied to reaction libraries, the informatic statistical structure is brought to light. This is vital to comprehending the norms, outliers, and limitations of carbon-based transformations.

III. SAMPLE RESULTS

As with their constituent molecules, organic reactions admit to ready classification—sigmatropic, tautomerizations, Diels−Alder, and so forth.[21] A reaction of a given class plus its variations can be explored in great depth in order to identify new synthetic methods. The reaction procedures of disparate classes can be combined sequentially in total synthesis projects. In this section, we examine the informatic pathway structures of a few well-known classes and sequences. It goes without saying that the results are not exhaustive. Rather the aim is to capture several informatic principles of carbon-based transformations. To this end, we confine attention to the informatic changes gauged by Brownian serial processing; the trends are reflected equally well using parallel methods.[5] Further, the information averages have been computed uniformly over correlation orders 2−8 as in the preceding paper; the trends are reflected equally well at higher orders. Finally, no attempt has been made to identify intermediate states via mechanistic-type structure graphs.[22] Rather all the attention is on the stable input and output states allowed by chemical structure theory.

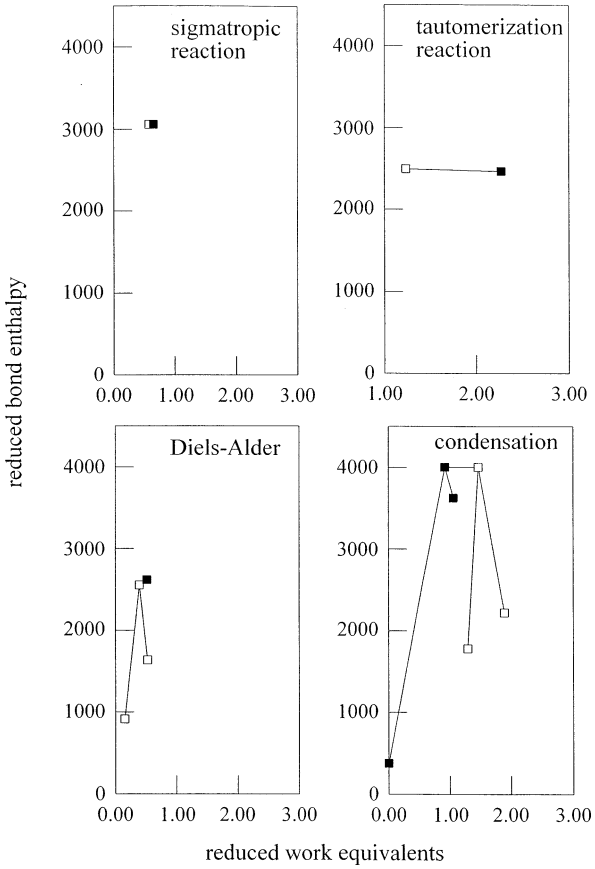

Figure 6. Rescaled informatic pathway structures. Structures for the Figure 1 reactions are illustrated. Open and filled symbols distinguish the input and output states, respectively. The ordinate and abscissa scales in each panel are dimensionless.

We begin with tautomers or valence isomers which interconvert readily under favorable conditions.[21−23] An example of a tautomerization was illustrated in the second graph sequence of Figure 1. Because of interconversion, it is virtually impossible to isolate a given tautomeric form in a condensed phase; a pathway always connects one input state to a neighboring output. A variety of techniques have been employed over the years to deduce the structures and equilibrium ratios of tautomeric molecules.[23] Figure 7 offers an assortment of transitions belonging to this class. Obviously this reaction type asserts itself frequently in nature given its ties to carbonyl, vinyl alcohol and amine, and amide functionality.

Figure 8 illustrates the informatic pathways of tautomerizations following Brownian computations. The upper panel illustrates the path-coupled state points of the Figure 7 transitions, with open and filled boxes marking the input and output states, respectively. Figure 7 describes an abbreviated library of various molecular sizes and functionalities. This diversity is reflected in measured $\boldsymbol{D}$, $\Delta\boldsymbol{H}$, $\Delta\boldsymbol{CI}$, and τ. A few observations are noteworthy. For example, the $\boldsymbol{H}$-values span a range of 1000 (owing to molecular size variations), whereas $\Delta\boldsymbol{H}$ lies near zero in several cases. In these latter cases, τ is also approximately zero, and the pathway length $\boldsymbol{D}$ is determined almost completely by $\Delta\boldsymbol{CI}$. There are additional ramifications in that the distributions for $\boldsymbol{D}$, $\Delta\boldsymbol{H}$,

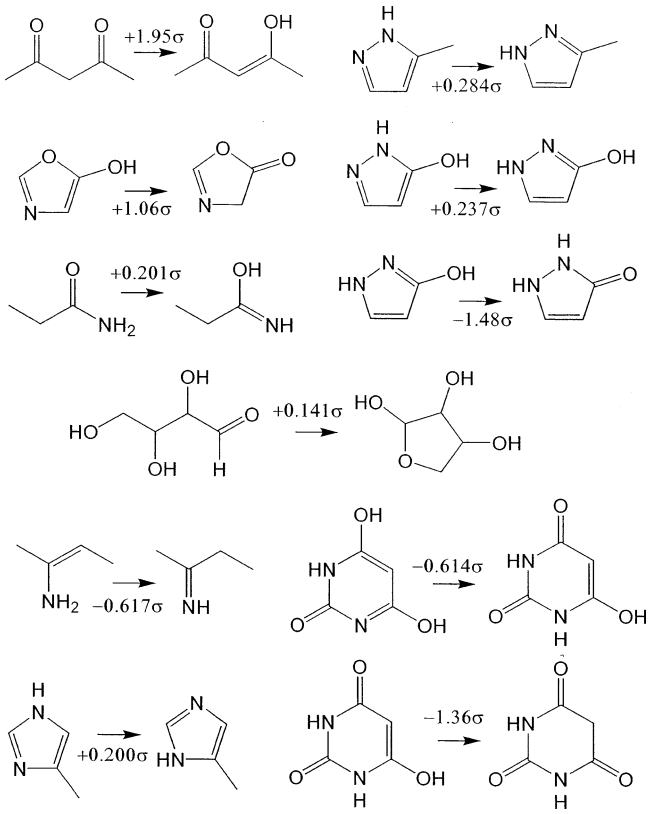

Figure 7. Tautomerization reaction library. A variety of molecular sizes and functionalities is represented. The reactions have been labeled using Z-values in library σ-units from statistical structure analysis.

and τ are distinctly bimodal; only the $\Delta\boldsymbol{CI}$-distribution expresses a single mode nature.

For the Figure 7 library, the $\Delta\boldsymbol{CI}$ span a range of $-1.41 - +1.48$; the $\Delta\boldsymbol{CI}$-average ($= \langle\Delta\boldsymbol{CI}\rangle$) and standard deviation σ are, respectively, -0.159 and 0.842. The $\Delta\boldsymbol{CI}$-distribution is represented in the lower panel of Figure 8 along with the standard normal distribution with zero mean and unity standard deviation. Here the variable Z equates with the number of standard deviation (σ) units contained in the difference between measured $\Delta\boldsymbol{CI}$ and the mean library value. F(z) measures the fraction of library members with $z \leq Z$.[24]

Nature poses an infinite number of possible tautomerizations and collections thereof. In our experience, the data of Figures 7 and 8 are representative for this brand of carbon transformation. Other tautomerization collections expressing a mix of C, H, O-functionalities yield similar results in the statistical sense. As with all molecular graph sequences, each Figure 7 transition can be characterized by informatic descriptors $\boldsymbol{D}$, $\Delta\boldsymbol{H}$, τ, $\Delta\boldsymbol{CI}$. Further, each pathway of a given collection is readily labeled by a Z-value (using σ units) following from the $\Delta\boldsymbol{CI}$-distribution. Figure 7 includes such Z-labels, while Table 2 reports data for the library set {$\boldsymbol{D}$, $\Delta\boldsymbol{H}$, τ, $\Delta\boldsymbol{CI}$}. For tautomerizations in general, the most significant $\boldsymbol{CI}$-modifications occur during the creation or loss of a vinyl alcohol (enol) functionality. Interestingly, ring closures such as the center sequence of Figure 7 ($+0.141\sigma$) entail only modest changes in molecular information.

Like tautomerizations, sigmatropic reactions also compose an infinite size class.[21,25,26] Sigmatropic reactions can be activated by both thermal and photochemical means. As with all isomerizations, the reactions are subject to orbital symmetry rules as, for example, in the topmost sequence of Figure 1.[25,26]

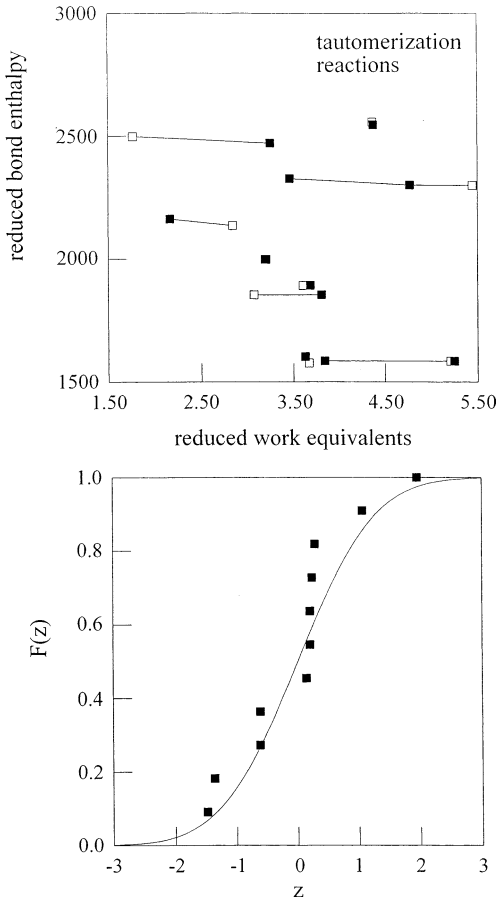

Figure 8. Informatic pathway structures for tautomerizations. Structures of the upper panel pertain to the Figure 7 reactions. Lower panel offers a comparison between library $\Delta\boldsymbol{CI}$ (filled squares) and the standard normal distribution (solid curve). In constructing this plot, $\langle CI_n^{(S)}\rangle$-values have been scaled relative to the library mean in units of the library standard deviation. Specifically, $Z = [\langle CI_n^{(S)}\rangle - \langle CI_n^{(S)}\rangle_{avg}]/\sigma$ with F(z) quantifying the fraction of library compounds with $z \leq Z$.

Table 2. Informatic Pathway Descriptors for Tautomerization Library

	average	standard deviation
$\Delta\boldsymbol{CI}$	-0.159	0.842
$\Delta\boldsymbol{H}$	$+3.78$	16.7
$\boldsymbol{D}$	11.0	12.8
τ	-180	405

Figure 9 presents an abbreviated library of [1,5]-diene isomerizations. Figure 10 elaborates by showing the results of pathway analysis in the upper panel. As with Figure 7, this collection hosts a variety of functionalities and thus diversity in $\boldsymbol{D}$, $\Delta\boldsymbol{H}$, $\Delta\boldsymbol{CI}$, and τ. The $\boldsymbol{H}$-values span a range of several thousand, while $\Delta\boldsymbol{H}$, $\tau \approx 0$ for several transitions. Further, the distributions for $\boldsymbol{D}$, $\Delta\boldsymbol{H}$, and τ are bimodal, whereas the $\Delta\boldsymbol{CI}$-distribution is monomodal. For the Figure 9 library, the $\Delta\boldsymbol{CI}$ span a range of $-1.34 - +0.477$; $\langle\Delta\boldsymbol{CI}\rangle$ and σ are -0.139 and 0.203, respectively. The $\Delta\boldsymbol{CI}$-distribution is represented in the lower panel of Figure 10 and compared with the standard normal distribution. Figure 9 describes a much abbreviated library, yet in our experience the informatic pathways are representative for [1,5]-diene

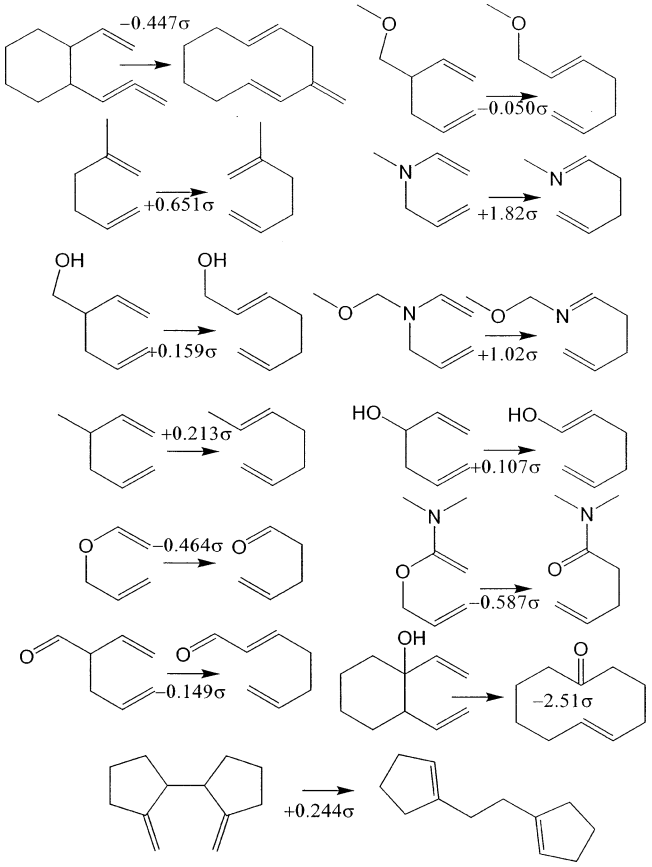

Figure 9. Sigmatropic reaction library. As in Figure 7, a variety of molecular sizes and functionalities has been represented; the reactions have been labeled using computed Z-values. The library is confined to [1,5]-diene rearrangements.

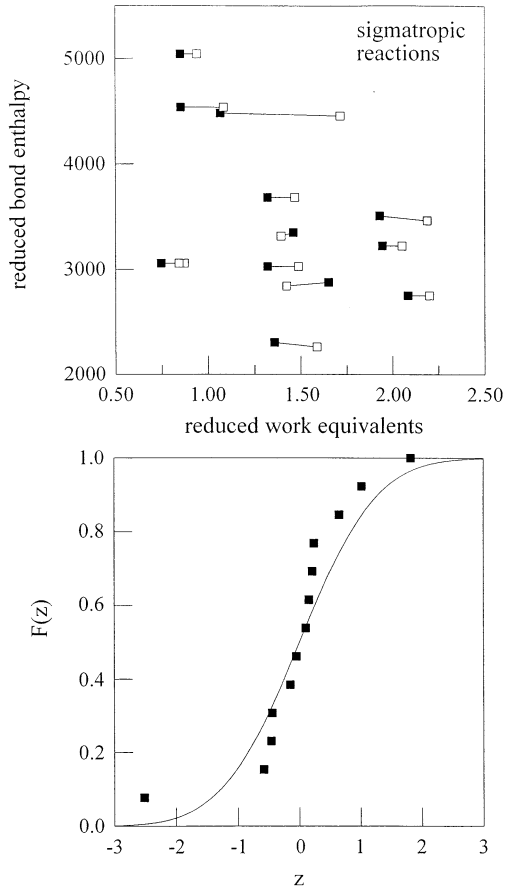

Figure 10. Pathway structures for sigmatropic reactions. Structures (upper panel) pertain to the Figure 9 reactions. Plotted in the lower panel are library Δ***CI*** data (filled squares) to be compared with the standard normal distribution (solid curve).

rearrangements. As in Figure 7, each Figure 9 transition has been labeled by its Z-value according to the Δ***CI***-distribution. Table 3 reports statistical parameters for the library {***D***, Δ***H***, τ, Δ***CI***}. Clearly, the most significant CI-modifications take place when the rearrangements involve heteroatoms O and N; these transitions are indeed affiliated with the extremum σ-values. Interestingly, Figure 9 includes an identity operation by way of the second graph sequence from top, left-hand-side. Here $\Delta CI = 0$ contrasts with affiliated $Z = +0.615\sigma$. While identity transitions are commonplace in digital electronics, they are to be considered somewhat unusual for reactions such as sigmatropic.

The Diels−Alder reaction was illustrated via the third sequence of Figure 1; nature allows infinite possible variations on this theme. We again offer data for an abbreviated library in Figures 11 and 12, such as can be obtained in combinatorial designs.[21,27,28]

The upper portion of Figure 11 portrays a Diels−Alder sequence, with variations indicated via 12 substituent (R-) groups: $-CH_3$, −OH, etc. Each transition involving ethylene and butadiene moieties can result in two products, ignoring stereochemical considerations. The pictorials thus describe a reaction library with $12 \times 12 \times 2 = 288$ members. The information pathway for the classic Diels−Alder (R− = H) was illustrated in Figure 6. Figure 12 elaborates by showing the pathway structures for all 288 library members. One is struck by the overlap complexity of the collection. While the Δ***H*** are positive-definite, the Δ***CI*** assert both positive and negative values. One learns that depending on the substituent identities, a Diels−Alder reaction can enhance or diminish the reagent ***CI***.

Table 3. Informatic Pathway Descriptors for Sigmatropic Reaction Library

	average	standard deviation
Δ***CI***	−0.138	0.203
Δ***H***	+14.1	19.1
D	14.2	18.9
τ	+19.9	169

Information pathway wise, the Diels−Alder set is complex. However, statistical methods help to identify structure within the pathway library. Figure 11 includes some notable results of this analysis. Shown in the middle and lower-most sections are several product graphs (ring compounds) from which it is straightforward to deduce the precursor combination. These graphs address the questions: What are the terminal points of the library reactions showing maximum Δ***CI***? Minimum? Which molecules terminate the pathways of greatest distance? Least distance? For the molecules illustrated, the extremum Z-values exceed 2.00σ or are less than -2.00σ.

For the Diels−Alder library, the Δ***CI*** span a range of −0.365 − +1.068 with respective mean and σ of +0.344 and 0.276. The total pathway ***D*** span a range of 2621−4577 with a mean and σ of 3667 and 436, respectively. The Δ***CI***-distribution is illustrated in the lower panel of Figure 12; one notes library ***D***, Δ***H***, and τ to adhere nominally to normal law statistics as well. This represents a departure from the

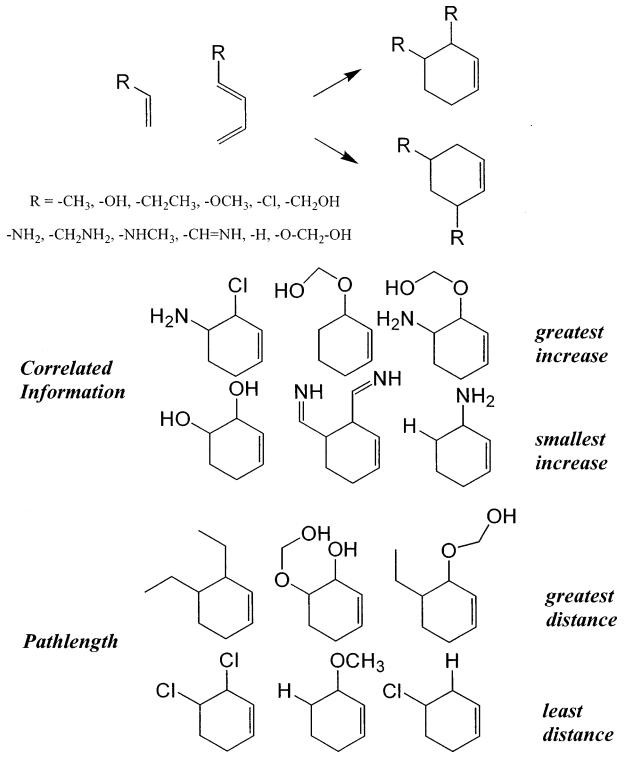

Figure 11. Diels−Alder library. Graphs in upper third of figure communicate a library of 288 reactions. Middle graphs illustrate product molecules whose informatic pathway structures demonstrated extremum Δ***CI*** values. Lower-most graphs illustrate product molecules whose informatic pathway structures demonstrated extremum ***D***.

unimolecular pathways where only the Δ***CI***-distributions were Gaussian in nature.

Additional chemical reaction perspective is offered via two classic studies in natural product synthesis. The topmost graphs of Figure 13 represent the initial and terminal states in the synthesis of morphine by Gates and Tschudi in the 1950s.[29] The enlarged arrow is shorthand for the multiple transformations which take a naphthalene derivative (left-hand-side graph) into the complex bridge structure on the right (morphine). With five chiral centers, there are $2^5 = 64$ versions of morphine but only one with the stereochemistry expressed by nature.

Equally impressive was the synthesis of cholesterol by Woodward and co-workers.[30] Here a Diels−Alder variant commenced a multistep process targeting a steroid with eight chiral centers. There are $2^8 = 256$ stereoisomers of cholesterol; the Woodward group obtained nature's version in chemistry initiated by achiral materials. Via chemical structure graphs, the literature documents the physical pathways of synthesis projects, such as for morphine and cholesterol.[31] What are the informatic structures for this chemistry?

Figure 14 illustrates the pathway structures for the reactions targeting morphine and cholesterol. The upper panels show the path-coupled state points, with open and filled boxes denoting the input and output states, respectively. The intermediate state points have been indicated using open circles. For clarity, the details involving ancillary reagent materials have been omitted. For both cases, the sequential reactions compose a type of library. A zigzag complexity is

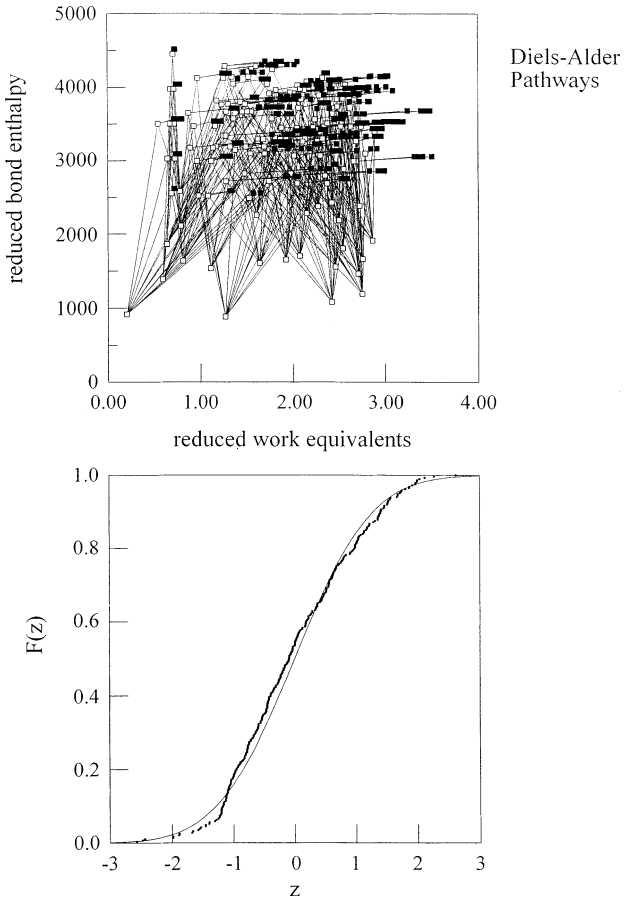

Figure 12. Informatic pathway structures for Diels−Alder library. Structures pertain to the combinatoric library of Figure 11. Open and filled symbols distinguish the input and output states, respectively. Plotted in the lower panel are library Δ***CI*** (filled squares) to be compared with the standard normal distribution (solid line).

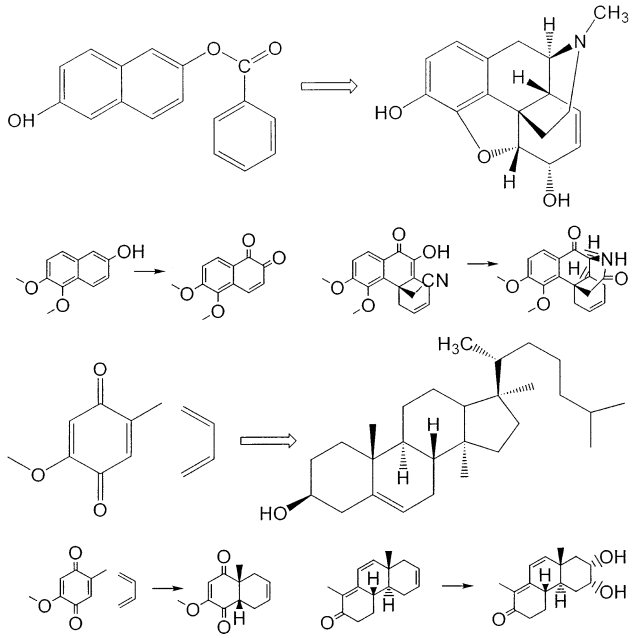

Figure 13. I/O states for morphine and cholesterol synthesis. The graph sequences illustrate reactions whose informatic pathway structures demonstrated extremum Δ***CI*** values within each synthetic program.

observed which reflects the diversity of the intermediate states and transformations.

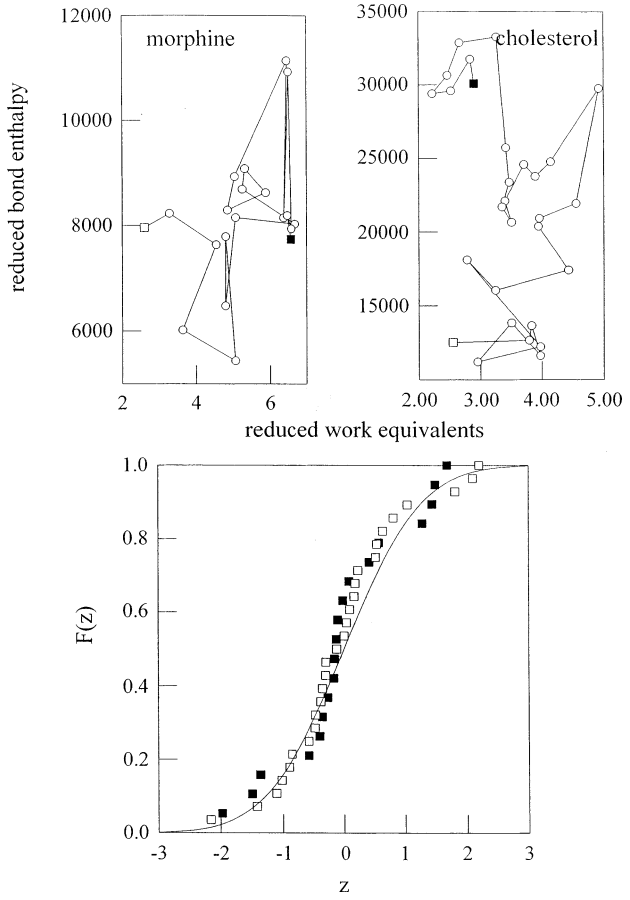

Figure 14. Informatic pathway structures for morphine and cholesterol synthesis. Structures of the upper panels pertain to the synthetic programs detailed in refs 29−31. Plotted in the lower panel are Δ***CI*** data to be compared with the standard normal distribution. The morphine and cholesterol synthesis data have been distinguished using filled and open squares, respectively.

Table 4. Informatic Pathway Descriptors for Morphine and Cholesterol Syntheses

	average	standard deviation
	Morphine	
Δ***CI***	+0.434	1.714
Δ***H***	−11.6	1588
D	1193	1010
τ	+4379	17589
	Cholesterol	
Δ***CI***	+0.026	1.16
Δ***H***	+253	1187
D	893	806
τ	−196	3473

The informatic pathways of natural product syntheses are complicated as a rule. Yet each step is readily characterized by ***D***, Δ***H***, τ, Δ***CI*** in a way which augments the molecular graph sequences. Further, each step can be labeled by a Z-value following from the library Δ***CI***-distribution. Figure 13 includes transformations which are noteworthy in the statistical sense, while Table 4 reports parameters for the library set {***D***, Δ***H***, τ, Δ***CI***}. For example, in the synthesis of morphine, an early transition involved a significant reduction of ***CI*** with Z value of -1.37σ. The bridging step brought about a significant CI increase with $Z \approx 1.68\sigma$. The molecular graphs for these informatically critical transformations are illustrated under the pathways connecting the input and output states. Note that these two synthetic literature examples represent markedly different informatic strategies. For the morphine synthesis, the chemistry is initiated via low-CI materials; molecular CI increases incrementally along the route to the target molecule. In synthesizing morphine, there appear four unusual compounds midway in the journey with CI equivalent to that of the target. In the total synthesis of cholesterol, the chemistry was initiated via materials with CI already comparable to the target value. Interesting is that the compounds with highest CI along the synthetic pathway are well removed from the target compound.

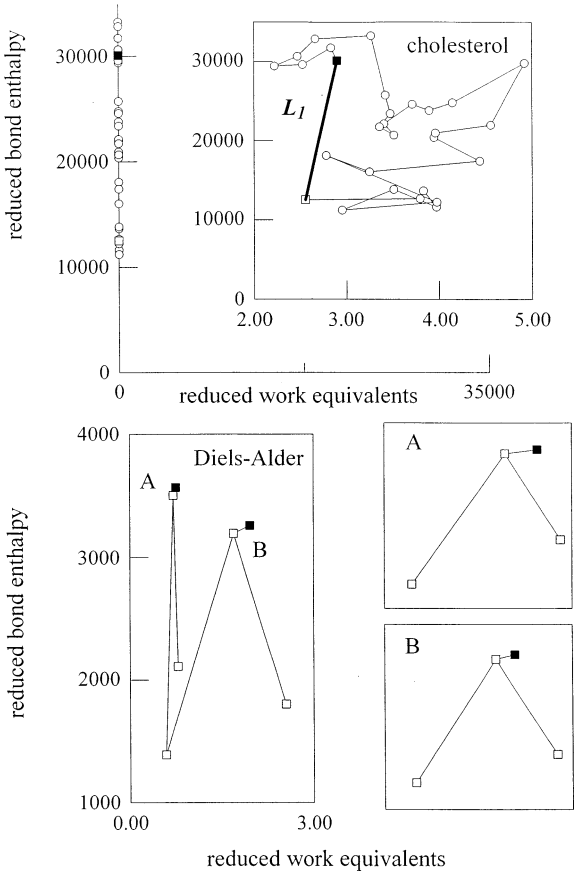

Figure 15. Informatic pathway structures for cholesterol synthesis and Diels−Alder reactions revisited. The upper panel illustrates the cholesterol pathway structure using identical scales for ***H*** and ***CI***. The inset plot shows the pathway structure along with quantity L_1 discussed in text. The lower panels illustrate pathway structures for two Diels−Alder reactions discussed in text.

IV. DISCUSSION

This paper has focused on reaction informatics at the covalent bond level, the results being new types of pathway structures. The latter do not serve in lieu of molecular graph sequences which communicate chemistry. Rather they connect with the random variable nature of molecules, and the information response functions they encode. The section III data offer only a small sampling of informatic pathway structures; however, three principles and two design tools are illustrated in the process. To assist the discussion, Figure 15 revisits the pathway structures for the cholesterol synthesis by Woodward and two Diels−Alder reactions of the Figure 11 library.

The first principle is that information changes along single reaction steps are by and large small: less than 2 on the average ***CI***-scale (eight correlation orders) for all of the

reactions investigated for this paper. This principle emphasizes certain limitations which confront synthetic practice, as molecules are able to modify information only incrementally. Thus for a chemist to alter ***CI*** for a given compound by several units, the logistics surrounding information storage are as critical as the reaction procedures themselves. A molecule's information can be augmented or decreased only in stages, with storable intermediates necessary to the reaction design and execution. Total synthesis programs such as for morphine and cholesterol illustrate beautifully what can be accomplished by sequential transformations. This paper adds that from an informatic perspective (Figure 15, upper panel), an extended reaction sequence is not unlike a mathematical operation, such as calculating the square roots of numbers, which requires the intermittent storage of data between the input and output states. However sophisticated the chemistry leading to cholesterol, the maximum $|\Delta \boldsymbol{CI}|$ demonstrated in the Woodward synthesis is -1.5; several intermediates are moreover necessary to the journey with ***CI*** more than double the input and output state values.

Note that the informatic path lengths for organic reactions are ordinarily dictated by enthalpic quantities. This is emphasized in the upper panel of Figure 15 (cholesterol synthesis) which includes a plot of ***H***-versus-***CI*** on the exact same scale, 0−35000; one finds all of the state points situated to the left-most side of the graph under these conditions. Although certain isomerizations (cf. Figures 7−10) offer notable exceptions to this principle, in our experience, we find $|\tau| = |(\Delta \boldsymbol{H}/\Delta \boldsymbol{CI})|$ to fall in the range twenty − several hundred for individual pathway segments. Clearly organic molecules alter information typically with sizable heat exchanges with the environment. On one hand, this is not surprising given typical reaction values of $|\Delta H°/T\Delta S°|$, the analogue of informatic $|\tau|$. Yet there is further insight available from an informatic standpoint as pathway structures identify another important trait that molecules share with logic gates (Figure 2). Both molecules and logic gates modify information with concomitant heat exchanges, with logic gates significantly more dissipative in their operation.[1,2] In small-signal transistors, for example, path length τ range from 10^4-10^8. Organic molecules lack the synchronous switching capabilities of solid-state elements; however, they effect information changes at orders-of-magnitude lower thermal expense.

A second principle is that the $\Delta \boldsymbol{CI}$-distributions of typical reaction libraries adhere to well-known probability laws. This statistical structure, while elementary, is important for estimation purposes in synthetic designs. It is impossible to identify all the transformation scenarios for a given molecule or derivative class, however advanced the substructure analytical strategies and chemical literature search-engines.[32] Nonetheless, it is a quick study to identify representative libraries of reactions (cf. Figures 7, 9, and 11) and, via Brownian computations, assemble collections of pathway structures. Via these, one is provided statistical guidelines surrounding the mutability of information. Pathway structure collections help to address the questions: What are typical $\Delta \boldsymbol{CI}$ for a given reaction class? What is informatically feasible for a class?

A third principle is that organic reactions evince a range of pathway lengths in total synthesis projects. Projects such as for morphine and cholesterol offer landmark examples of best-practice, custom-designed chemistry. Yet they demonstrate an informatic signature common to synthetic programs, namely that the connecting pathways express a random walk nature. Note how the zigzag structures in Figure 14 assert a Brownian-type process. In our experience, we have found this statistical behavior typical for synthetic designs, however rational. A Brownian nature is part and parcel to chemical transformation sequences.

The random walk attributes of sequential reactions pose a practical tool provided by information pathway analysis, namely a figure of merit Λ for comparing two or more synthetic programs. Let any two points of the ***H***/***CI*** plane be separated by a distance $\boldsymbol{L_1}$; for example, the upper panel of Figure 15 indicates this distance between the initial and final states of the Woodward cholesterol synthesis. Accordingly, for a reaction sequence with N number of transformations, a universal figure of merit can be defined as follows:

$$\Lambda = (1/\mathrm{N})\sum_{i=1}^{i=n} \boldsymbol{D_i}/\boldsymbol{L_1}$$

Clearly Λ equals unity for any single reaction step while for any reaction sequence $\Lambda \leq 1$. It is thus interesting to compute Λ for various landmark syntheses following pathway analysis. For example, one finds Λ to equal 5.1×10^{-3} and 1.0×10^{-2} for morphine and cholesterol syntheses, respectively. Informatically speaking, the cholesterol synthesis program offers a more direct pathway structure—by a factor of 2—connecting the input and output states. Chemists design reaction sequences with issues of cost, yield, and number of steps in mind. This paper would add that the pathway structures computed by Brownian methods offer another tool for weighing the informatic merits of different programs. Given the choice of two reaction sequences in a synthetic challenge, the one with the higher Λ offers the more direct route.

A second tool offered by information pathway analysis is conveyed via two of the Diels−Alder reactions of Figures 11 and 12. Illustrated in the lower panels of Figure 15 are the pathway structures for the reactions leading to 3,4-dimethylcyclohexene **(A)** and 3-ol-4-methylcyclohexene **(B)**. The left-most panel shows the two structures on the same ***H***/***CI*** scale. One observes the structures to be qualitatively similar in that each expresses three intersecting rays. The pathway structures are otherwise different in terms of operational lengths, angles, and ***H***/***CI*** positions. Different chemical reactions demonstrate different pathway structures. Where is the useful tool?

The lower, right-most panels of Figure 15 illustrate matters following a linear translation and change of scale. Note how the redrawn pathway structures of the two Diels−Alder reactions express many more similarities than differences; they are indeed nearly superimposable or operationally congruent in the informatic sense. Exact operational congruence is demonstrated by pathway structures expressing identical ***T*** and length ratios involving the individual rays.

A second tool of informatic pathway analysis is thus a test for operational congruence. Pathway structures are similar if they demonstrate the same number and relative configuration of intersecting rays and markedly different otherwise. Yet two reactions are informatically congruent if

their pathway structures are exactly superimposable following a linear translation and scale change in the ***H***/***CI*** plane; such reactions express the same degree of dissipation and relative distances between the I/O states. Congruence tests assist a chemist in adapting synthetic methods which have been detailed in the literature to new, unexplored territory. If there is a choice of more than one literature procedure to apply in a new synthetic challenge, the preferred one should be that which demonstrates the greatest degree of congruency with the unfamiliar. Congruency tests offer geometric tools applicable to all brands of molecular functionality and reaction classes.

V. SUMMARY AND CLOSING

Informatic pathway analysis is brought about by Brownian processing applied to the molecular graphs involved in a chemical reaction. By this analysis, one converts one set of 2D graphics into an alternative one. Reactions illustrated via this alternative geometry offer the informatic changes brought about in the covalent bond transformations. The pathway structures complement the graph sequences of texts and journals. Several examples have been discussed in this paper so as to illustrate the key principles and tools. A follow-up study of solvent effects on pathway informatics is currently underway. Brownian information processing figures prominently in all instances of molecular recognition and chemical action.

ACKNOWLEDGMENT

The authors are grateful to Professor Gordon Ramsey for technical assistance and to Professor James Babler for helpful discussions. One of the authors (M.V.S.) is grateful for the support via the Research Experiences for Undergraduates program sponsored by the National Science Foundation. The authors appreciate the helpful criticism of anonymous referees.

REFERENCES AND NOTES

(1) Bennett, C. H. Thermodynamics of Computation—A Review. *Intl. J. Theo. Phys.* **1982**, *21*, 905.
(2) Feynman, R. P. *Feynman Lectures on Computation*; Hey, A. J. G., Allen, R. W., Eds.; Addison-Wesley: Reading, MA, 1996.
(3) Pohl, I.; Shaw, A. *The Nature of Computation: An Introduction to Computer Science*; Computer Science Press: Potomac, MD, 1981.
(4) Rodd's Chemistry of Carbon Compounds. In *Aromatic Compounds*; Coffey, S., Ed.; Elsevier: Amsterdam, 1976; Vol. III, Part D.
(5) Graham, D. J.; Malarkey, C.; Schulmerich, M. V. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1601−1611.
(6) Conrad, M. Microscopic-Macroscopic Interface in Biological Information Processing. *BioSystems* **1984**, *16*, 345. Conrad, M. Molecular Computing. In *Advances in Computers*; 1990; Vol. 31, p 235. Conrad, M. On Design Principles for a Molecular Computer. *Commun. ACM* Yovits, M. C., Ed.; **1985**, *28*, 464.
(7) Zauner, K. P.; Conrad, M. Enzymatic Pattern Processing. *Naturwissenschaften* **2000**, *87*, 360.
(8) Adelman, L. M. Molecular Computation Solutions to Combinatorial Problems. *Science* **1994**, *266*, 1021.
(9) Hengen, P. N.; Lyakhov, I. G.; Stewart, L. E.; Schneider, T. D. Molecular Flip-Flops Formed by Overlapping FIS Sites. *Nucl. Acid. Res.* **2003**, *31*, 6663.
(10) Oliver, J. S. Matrix Multiplication With DNA. See: Oliver, J. S. *J. Mol. Evol.* **1996**, *45*, 161. Lipton, R. J. DNA Solution of Hard Computational Problems. *Science* **1995**, *268*, 542.
(11) Hjelmfelt, A.; Weinberger, E. D.; Ross, J. Chemical Implementation of Finite-State Machines. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 383. Hjelmfelt, A.; Ross, J. Implementation of Logic Functions and Computations by Chemical Kinetics. *Physica D* **1995**, *84*, 180. Hjelmfelt, A.; Ross, J. Mass-Coupled Systems with Computational Properties. *J. Phys. Chem.* **1993**, *97*, 7988.
(12) Steinbock, O.; Kettunen, P.; Showalter, K. Chemical Wave Logic Gates. *J. Phys. Chem.* **1996**, *100*, 18970.
(13) Turing, A. M. The Chemical Basis of Morphogenesis. *Philos. Trans. R. Soc. London Biol. Sci.* **1952**, *327*, 37.
(14) Fermi, E. *Thermodynamics*; Dover: New York, 1936.
(15) Weinhold, F. Geometric Representation of Equilibrium Thermodynamics. *Acc. Chem. Res.* **1976**, *9*, 6.
(16) Salamon, P.; Nulton, J. D.; Berry, R. S. Length in Statistical Thermodynamics. *J. Chem. Phys.* **1985**, *82*, 2433.
(17) Salamon, P.; Nulton, J.; Ihrig, E. On the Relation Between Entropy and Energy Versions of Thermodynamic Length. *J. Chem. Phys.* **1984**, *80*, 436.
(18) Hamming, R. W. *Coding and Information Theory*; Prentice-Hall: Englewood Cliffs, 1980; Chapter 6.
(19) Cox, J. D.; Pilcher, G. *Thermochemistry of Organic and Organometallic Compounds*; Academic Press: New York, 1970.
(20) Brillouin, L. *Science and Information Theory*, 2nd ed.; Academic: New York, 1962.
(21) March, J. *Advanced Organic Chemistry*, 3rd ed.; Wiley: New York, 1985.
(22) le Noble, W. J. *Highlights of Organic Chemistry*; Dekker: New York, 1974; Chapter 1.
(23) Wheland, G. W. *The Theory of Resonance*; Wiley: New York, 1944; Chapter 8.
(24) Freund, J. E. *Mathematical Statistics*; Prentice Hall: Englewood Cliffs, 1962.
(25) Woodward, R. B.; Hoffmann, R. *The Conservation of Orbital Symmetry*; Academic Press: 1970.
(26) Borden, W. T. *Modern Molecular Orbital Theory for Organic Chemists*; Prentice-Hall: Englewood Cliffs, 1975; Chapter 4.
(27) Fringuelli, F.; Taticchi, A. *The Diels−Alder Reaction. Selected Practical Methods*; Wiley: New York, 2002.
(28) Olah, G.; Molnar, A. *Hydrocarbon Chemistry*; Wiley: New York, 1995; Chapter 6.
(29) Gates, M.; Tschudi G. The Synthesis of Morphine. *J. Am. Chem. Soc.* **1952**, *74*, 1109. Gates, M.; Tschudi, G. The Synthesis of Morphine. *J. Am. Chem. Soc.* **1956**, *78*, 1380.
(30) Woodward, R. B.; Sondheimer, F.; Taub, D.; Heusler, K.; McLamore, W. M. The Total Synthesis of Steroids. *J. Am. Chem. Soc.* **1952**, *74*, 4223.
(31) Fleming, I. *Selected Organic Syntheses*; Wiley: London, 1973.
(32) Ho, T.-L. *Tactics of Organic Synthesis*; Wiley: New York, 1994.

CI040022V