CROATICA CHEMICA ACTA CCACAA **69** (4) 1385–1403 (1996)

ISSN-0011-1643
CCA–2383

Original Scientific Paper

Knotted Structures in Chemistry, Biochemistry, and Molecular Biology*

*Kurt Mislow*** *and Chengzhi Liang*

Department of Chemistry, Princeton University, Princeton
New Jersey 08544, USA

Received April 19, 1996; accepted July 1, 1996

The term »knot« in the mathematical sense, as a closed curve in 3-space that does not intersect itself but that cannot be embedded in the plane without intersections, has been used to describe a wide variety of molecular structures – except in the protein literature, where it has been used to describe structural motifs that are neither closed knots in the above sense nor even knots in the everyday sense. We argue in this paper that description of these motifs as »knots« constitutes »improper« use of scientific terminology, and we propose a definition of »molecular knot«.

> *»Ne faites grace à aucune dénomination impropre. Ceux qui savent déjà, entendront toujours; ceux qui ne savent pas encore, entendront plutôt«.*[1]

INTRODUCTION

In the preface to their classic treatise on entomology, William Kirby and William Spence pointed out that they had introduced a »great number of new terms, and alterations of old ones« in order to achieve »a greater degree of precision and concinnity [*i.e.*, harmony or consistency]«.[1] Concerned that critics might object to these changes, Kirby and Spence cited »the advice of Bergman to Morveau, when reforming the nomenclature of Chemistry,« that

---

* Dedicated to the memory of Stanko Borčić.

** Author to whom correspondence should be addressed.

there is no room for »improper« nomenclature in science. We have chosen this advice as the epigraph of the present paper.

The present study is an outgrowth of earlier work dealing with topological chirality and achirality of knots and links.[2] In what follows we discuss the use of scientific terminology, specifically with reference to the different ways in which the term »knot« has been applied to molecular structures and structural motifs.

MOLECULAR STRUCTURES THAT ARE KNOTS IN THE MATHEMATICAL SENSE

As defined in topology, a knot is a polygonal or smooth closed curve in 3-space that does not intersect itself anywhere but that cannot be embedded in the plane without intersections (called »crossings«).[3] The word »closed« is crucial to this definition, for a curve that is knotted but not closed can be unknotted, like a tied shoelace, by continuous deformation, whereas a knot as defined above cannot be unknotted without cutting the curve. Thus, from the point of view of topology, only a closed curve can be »truly knotted«.[4] We shall refer to such a curve as a *closed knot*.

There is no dearth of molecules whose constitutions are represented by graphs that are closed knots. The rational synthesis of a molecular trefoil knot was first achieved by Dietrich-Buchecker and Sauvage in 1989 (Figure 1).[5] Soon thereafter, Seeman and coworkers reported the synthesis of trefoil and figure-eight knots from single-stranded DNA (Figure 2).[6] A great variety of single- and double-stranded DNA closed knots have been observed in diverse biological systems, and by now have become a commonplace in

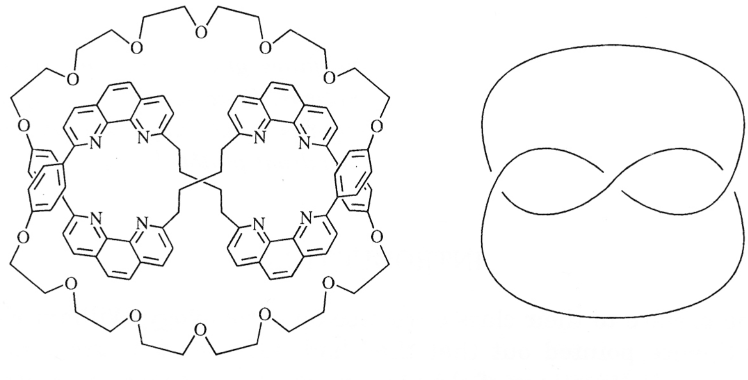
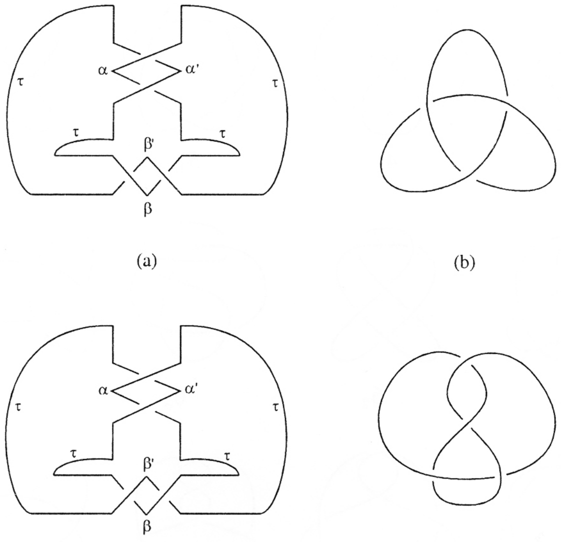


Figure 1. Left: Synthetic trefoil molecule K-86.[5] Unmarked vertices represent C, CH, or CH_2 groups. Right: An abstract version of the same knot.



(a) (b)

(c) (d)

Figure 2. (a) Synthetic single-stranded DNA trefoil knots[6] in which α = (A–C–T–G–G–A–C–C–T–C–T), β = (C–G–T–A–G–C–C–G–C–A–T) or $(dCpdGp)_6$, α' and β' refer to sequences that complement α and β by Watson-Crick hydrogen bonding, respectively, and $\tau = dT_{15}$ symbolizes a single-stranded linkage between α (α') and β (β'). (b) An abstract version of (a) under denaturing conditions. (c) A synthetic single-stranded DNA figure-eight knot[6] in which $\beta = (dCpdGp)_6$, and α, α', β', and τ have the same significance as in (a). (d) An abstract version of (c) under denaturing conditions.

»biochemical topology« (Figure 3).[7,8] Closed knots, consisting of polypeptide chain segments combined with cofactors and disulfide intrachain cross-links, have recently been observed among some metalloproteins (Figure 4).[2b]

MOLECULAR STRUCTURES THAT ARE CALLED KNOTS BUT ARE NOT KNOTS IN THE MATHEMATICAL SENSE

Mansfield has pointed out that »we may be unable to define a knot in an open path, but we know one when we see one. Most reasonable people would agree that [Figure 5a] is knotted while [Figure 5c] is not.«[9] The crucial difference lies in the fact that when the free ends of the curve are glued together, Figure 5a becomes a trefoil knot (Figure 5b), *i.e.* a »true« knot as defined above, whereas Figure 5c becomes an »unknot« (Figure 5d), *i.e.,* a closed curve that can be embedded in the plane without crossings. Thus, the curve in Figure 5a, while not a knot in the mathematical sense, has the fa-

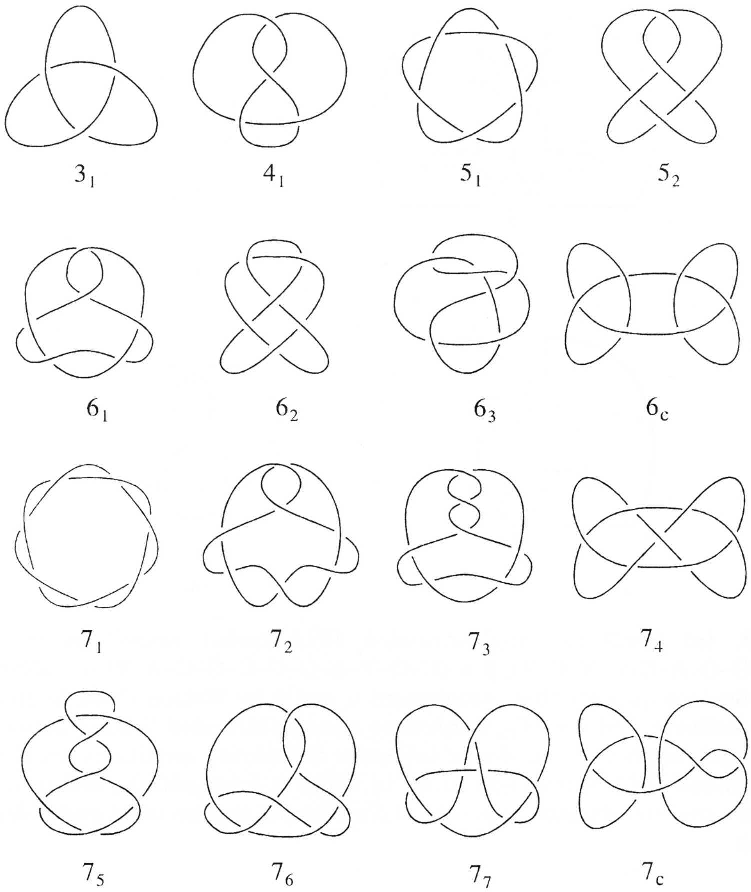


Figure 3. Standard diagrams of duplex DNA knots produced by topoisomerase I.[7b]

miliar appearance of a knot »in the every-day sense«;[10] it is also called a »quasi-knot«.[11] Such a knot, which we shall refer to as an *open knot*, may be defined as a closed knot that has been cut once but has not yet been unknotted – like a tied shoelace. An open knot may be subjectively identified by a »reasonable-person test«.[9] Using this test, Mansfield (»a reasonable person (namely myself)«) found that human carbonic anhydrase B forms an »incipient knot« because in some projections a few residues at one end are tucked through a loop that passes through the exterior of the molecule.[9] Carbonic anhydrase is thus an open knot – a »loosely knotted« polypeptide chain.

There exist, however, a great many proteins with structural motifs that are neither open nor closed knots, as defined above, but that are nevertheless characterized as »knots«.[12–45] The »knots« in these motifs are entangle-

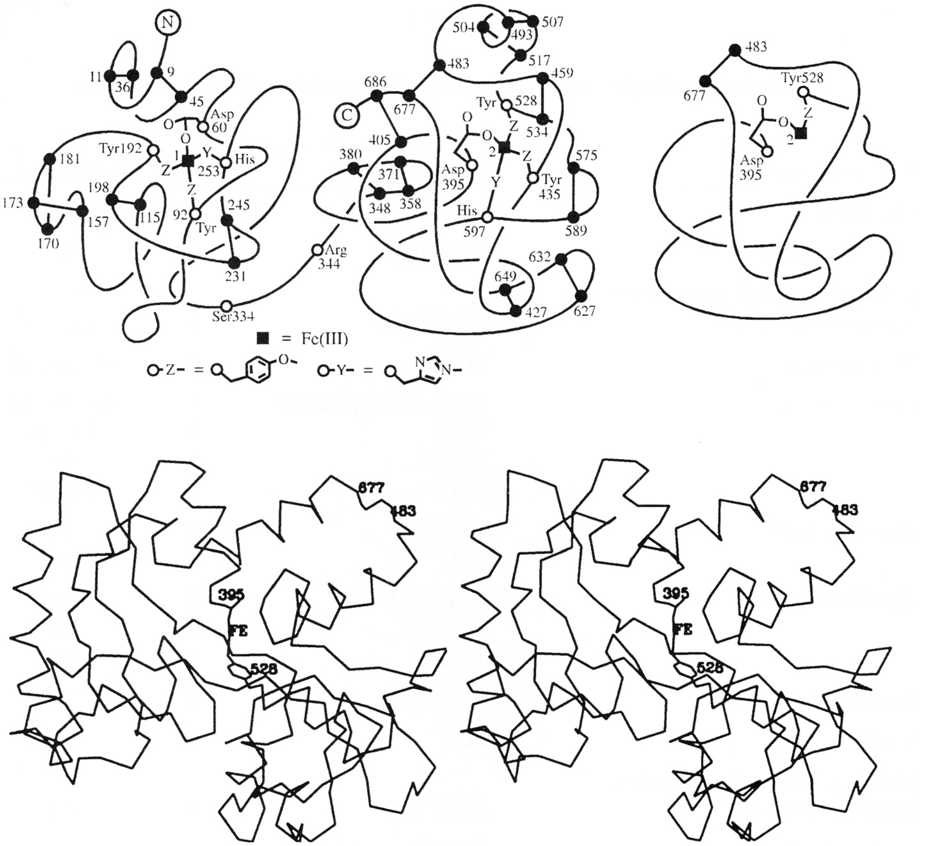


Figure 4. Top left: Condensed schematic diagram of human lactoferrin (hLf), with the C-lobe shown on the right and the N-lobe on the left. The α-carbons of cysteine and selected non-cysteine residues are symbolized by filled and open circles, respectively. Unlabeled vertices symbolize carbon atoms, and hydrogen atoms are suppressed for clarity. Top right: The trefoil knot derived from the C-lobe of hLf. Bottom: Stereoview (Cα trace plus cross-links) of the knot shown on the top right. (Reprinted with permission from Ref. 2b. Copyright 1995, American Chemical Society).

ments that bear little or no resemblance to the knots described above. In one type of protein (Table I), a bond or a segment of the chain is threaded through a closed loop, in a topology reminiscent of a rotaxane[2b] (Figures 6,[12] 7,[13,46] 8,[14–28] and 9[29–32]). In three other types (Table II) even this minimal interlacement is absent (Figures 10,[33] 11,[34–37,47] and 12[38]). In a number of instances, the term »knot« has been applied to structural motifs in proteins even in the absence of three-dimensional structural information. These include the following: references to »knots« in fibrinogen and human TIMP-1 (Table II); the expression »E-knot«, which has been extensively used in a

TABLE I

Threaded topologies in proteins and the corresponding expressions

Expression(s)	Molecule(s)[a]	Figure	Reference(s)
slip-knot	BPTI	6	12
knot structure	RES-701-1	7	13
knotted arrangement, knot-like (cystine) topology	PDGF BB	8	14, 20
TGF-β knot	TGFβ	8	15, 20, 21
cystine knot	NGF, TGFβ, PDGF	8	16, 20–22
	NDP	8	17, 23
	PTTH	8	18
	hCG	8	19–21
cysteine knot	βNGF	8	24
	hCG	8	25
disulfide knot	Human C3a	8	26
	TGF-β2	8	27
	βNGF	8	24
knotted topology	kalata B1	8	28
knot	PCI	9	29
knottin, knotted structure, knotted protein	EETI II	9	30
cystine knot	ω-CgTx, kalata BI, and CMTI-I	9	31
cysteine knot	PMP-C, PMP-D2, HI	9	32

[a] The abbreviations used are: BPTI, bovine pancreatic trypsin inhibitor; RES-701-1, endothelin B ($\mathrm{ET_B}$) selective antagonist; PDGF BB, platelet-derived growth factor BB; TGFβ and TGF-β2, transforming growth factor-β2; NGF, nerve growth factor; NDP, norrie disease protein; PTTH, prothoracicotropic hormone; hCG, human chorionic gonadotrophin; Human C3a, human anaphylatoxin; βNGF, murine β-nerve growth factor; PCI, carboxypeptidase inhibitor from potatoes; EETI II, *Ecballium elaterium* trypsin inhibitor II; ω-CgTx, ω-conotoxin GVIA; CMTI-I, *Cucurbita maxima* trypsin inhibitor I.

study of the relation between the fragment E and the N-DSK (NH_2-terminal disulfide knot) of fibrinogen;[39] the term »disulfide knot«, which appears in connection with human C4 binding protein,[40] the fragment D of fibrinogen,[41] the basement membrane glycoprotein laminin,[42] the basement-membrane type IV collagen,[43] and the *Drosophila* snake protease;[44] and the terms »hydrophobic disulfide knot« and »hydrophilic disulfide knot«, which can be found in discussions of human fibrinogen.[45]

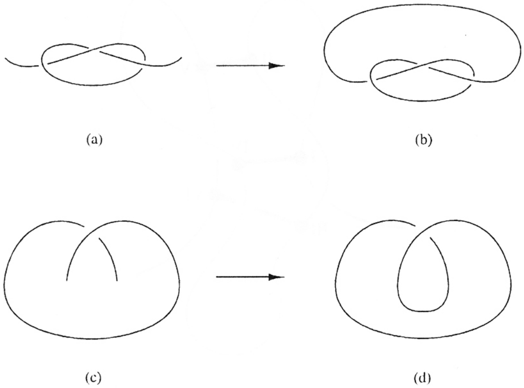


Figure 5. (a) An open knot. (b) A closed knot (trefoil). (c) A presentation of a curve in space. (d) An unknot.

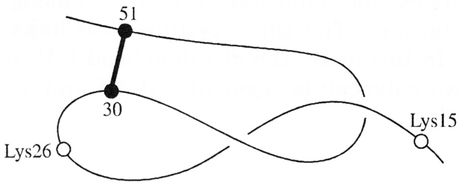


Figure 6. Condensed schematic diagram of the »slip-knot« motif in BPTI.[12] The α-carbons of cysteine residues and selected non-cysteine residues are symbolized by filled and open circles, respectively. An internal cystine cross-link (disulfide bond) is shown as a heavy line. In this motif, the polypeptide chain of residues Lys15 through Lys26 passes through the closed loop defined by the Cys30-Cys51 disulfide bond and the intervening polypeptide chain.

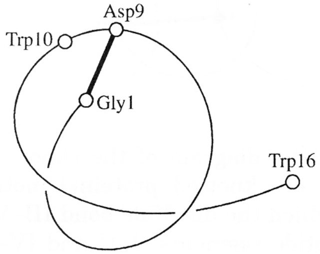


Figure 7. Condensed schematic diagram of RES-701-1.[13] Some selected non-cysteine residues are symbolized by open circles. An internal amide linkage (bond) between the β-carboxy group of Asp9 and the α-amino group of Gly1 is shown as a heavy line. The »tail« (polypeptide chain Trp10–Trp16) of this »knot structure« penetrates the »ring« (polypeptide chain Gly1 – Asp9 plus the amide bond between Asp9 and Gly1). This structural feature of RES-701-1 is the same as that of a tricyclic peptide RP-71955.[13,46]

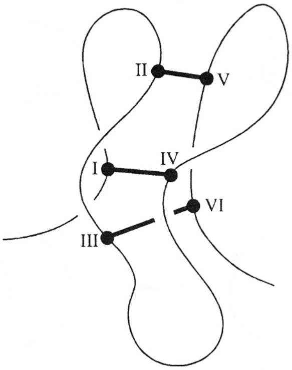


Figure 8. Condensed schematic diagram of the »cystine knot« (a. k. a. »cysteine knot«, »disulfide knot«, »TGF-β knot«, »knot-like (cystine) topology«, »knotted arrangement«, or »knotted topology«) motif of proteins in a growth-factor superfamily,[14–27] and cyclic uterotonic polypeptide kalata B1.[28] The α-carbons of cysteine residues are symbolized by filled circles, and numbered (I, II, III, ···) along the polypeptide chain from NH_2 to COOH termini. The three cystine cross-links (disulfide bonds) are shown as heavy lines. In this motif, the disulfide bond I–IV passes through the closed loop formed by two polypeptide segments (II–III and V–VI) and two disulfide bonds (II–V and III–VI).

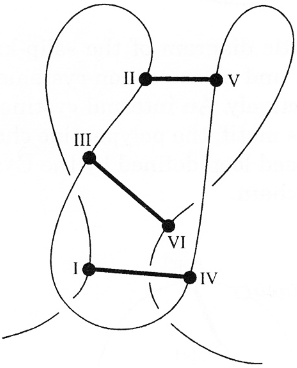


Figure 9. Condensed schematic diagram of the »knot« (a. k. a. »knottin«, »cystine knot«, »knotted structure«, or »knotted protein«) motif of PCI[29] and other proteins[30–32] (see Table I), in which the disulfide bond III–VI passes through the closed loop formed by two polypeptide segments (I–II and IV–V) and two disulfide bonds (II–V and I–IV).

Because the plethora of various structural motifs referred to as »knots« in the protein literature have little if anything in common with the open or closed knots described above, it is natural to wonder whether there is any basis for limiting the meanings that can be attached to this term when used

TABLE II

Some structural motifs in proteins and the corresponding expressions

Motif	Expression(s)	Molecule(s)[a]	Figure	Reference(s)
type 1	T-knot scaffold	MCTI-A, CPI, ω-Aga-IVB, FIX, EGF, and TGF-α	10	33
type 2	disulfide knot	Fibrinogen	11	34
	disulfide knot (DSK)	Fibrinogen	11	35
	NH_2-terminal disulfide knot (N-DSK)	Fibrinogen	11	36, 37
type 3	knotlike structure, knot	Human TIMP-1	12	38

[a] The abbreviations used are: MCTI-A, trypsin inhibitor from bitter gourd; CPI, carboxypeptidase A inhibitor from potato; ω-Aga-IVB, ω-toxin from funnel web spider venom; FIX, human factor IX; EGF, mouse epidermal growth factor; TGF-α, transforming growth factor-α; TIMP-1, tissue inhibitor of metalloproteinases.

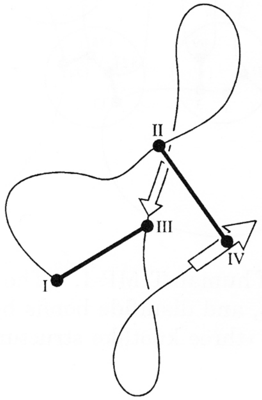


Figure 10. The »T-knot scaffold« shared by the EGF-like proteins, ω-toxins and proteinase inhibitors from plants.[33] The two arrows denote the two β-strands in the scaffold. The α-carbons of cysteine residues are symbolized by filled circles, and disulfide bonds by heavy lines.

with reference to molecular structures. More generally, what, if any, are the constraints imposed on the number and kind of meanings that can be attached to a given word?

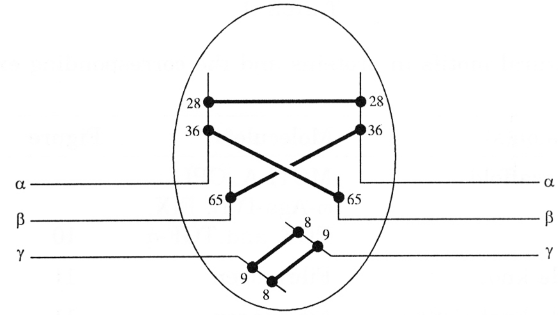


Figure 11. A model of the NH_2-terminal portion of fibrinogen based on Figure 4 of Ref. 47. The area encompassed by an oval is referred to as the »NH_2-terminal disulfide knot«.[34–37] The α-carbons of cysteine residues are symbolized by filled circles, and disulfide bonds by heavy lines. The earliest model of this motif can be found in Figure 5 of Ref. 36.

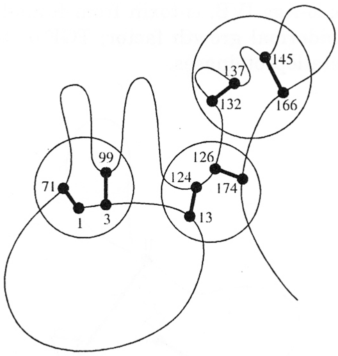


Figure 12. Primary structure of human TIMP-1.[38] The α-carbons of cysteine residues are symbolized by filled circles, and disulfide bonds by heavy lines. The three areas encompassed by circles indicate »three knotlike structures«.

WHAT'S IN A NAME?

In discussions dealing with the semantic content of words, it is customary to cite the famous exchange between Alice and Humpty Dumpty, »a philologist and philosopher skilled primarily in linguistic matters«:[48]

> »When *I* use a word,« Humpty Dumpty said, in rather a scornful tone, »it means just what I choose it to mean – neither more nor less.«
>
> »The question is,« said Alice, »whether you *can* make words mean so many different things.«
>
> »The question is,« said Humpty Dumpty, »which is to be master – that's all.«

Humpty Dumpty is right: words mean »neither more nor less« than what they are intended to mean. According to Martin Gardner, »Lewis Carroll was fully aware of the profundity in Humpty Dumpty's whimsical discourse on semantics. Humpty Dumpty takes the point of view ... that universal terms do not refer to objective existences but are nothing more than *flatus vocis*, verbal utterances. This view ... is now held by almost all contemporary logical empiricists.«[48]

Thus, Humpty Dumpty has no trouble explaining the meaning of some curious words in the poem »Jabberwocky«: »brillig« means »four o'clock in the afternoon«, »toves« are »something like badgers – they're something like lizards – and they're something like corkscrews ... also they make their nests under sundials – also they live on cheese«, and so forth. Invented or »nonsense« words like these present no problem, because it is we and not the words who are master: the semantic space of the neologism is empty, and we can fill it arbitrarily with any meaning we choose, limited only by the boldness of our imagination. Humpty Dumpty's denotation of the word »brillig« to mean »4 p.m.« in the context of Jabberwocky may appear bizarre, but it is no more unreasonable than the equally arbitrary denotation of the letter N to signify, say, the number of moles per liter of a gas at STP in the context of chemistry. Conversely, there is nothing wrong with calling four o'clock in the afternoon »brillig«. Juliet, inviting Romeo to »doff thy name ... which is not part of thee«, argues: »What's in a name? That which we call a rose / By any other name would smell as sweet«.[49]

It is quite another matter, however, to give totally new meanings to words that are already in common use and whose semantic content is accepted by general consensus: except in a poetic or metaphoric context, such a reassignment of meaning leads to a breakdown in communication. A case in point is Humpty Dumpty's response to Alice's question: »I don't know what you mean by 'glory'« – a response that precipitated the exchange quoted above:

> Humpty Dumpty smiled contemptuously. »Of course you don't – till I tell you. I meant 'there's a nice knock-down argument for you!'«
>
> «But 'glory' doesn't mean 'a nice knock-down argument,'« Alice objected.

Alice's objection is justified. If we admit that an utterance is merely *flatus vocis*, then Humpty Dumpty is entitled to call a nice knock-down argument anything he pleases, including »glory«. As a practical matter, however, the choice of this word is bound to cause extreme bewilderment because there already exists another, previously established and totally different meaning of the same word. Similarly, it would create utter confusion if Humpty Dumpty were to call a nice knock-down argument a »tove« because he previously described a tove as some sort of corkscrew-shaped cheese-eating lizardy badger. Because Humpty Dumpty invested his newly coined word with a meaning, his creation has now become master.

Lexical Definitions

Unlike »brillig«, the word »knot« does not have a precise and narrowly defined meaning but is part of the general vocabulary and carries a broad spectrum of meanings. Definitions of words, such as »knot«, that are in general use, and whose semantic content has been well-established over time, are found in standard dictionaries. These »lexical definitions«[50] are authoritative statements concerning the actual use of a word in the present, recent, or distant past.

Before we move on to consider the lexical definition of »knot«, we briefly digress to comment on the title of a recent series of papers by Elk, »Orismology (the science of defining words) and the geometrical foundations of chemistry«.[51] According to Elk,[51a] »it is only in very recent times that a special nuance of meaning, distinct from that implied by the word 'terminology', has arisen in science that has resurrected this otherwise neglected word«. For the reason given below, a strong case can be made for continued neglect.

In the preface to their treatise on entomology that contained the advice of Bergman to Morveau, Kirby and Spence stated: »In the *Terminology*, or what, to avoid the barbarism of a word compounded of Latin and Greek, they [*i.e.,* the authors] would beg to call the *Orismology* of the science [*i.e.*, entomology] ...«.[1] Words, as we saw, mean neither more nor less than what they are intended to mean, and the authors' sole and clearly stated intention was to replace »terminology«, a word of mixed etymology, by a neologism with exactly the same meaning but with etymologically matching roots. Thus, orismology is synonymous with terminology. The correspondence between the two terms is as complete as one could wish: the Latin *terminus* means »boundary« or »limit«, and the Greek *horismos*, meaning »definition«, is derived from *horos*, which also means »boundary« or »limit« [whence: horizon]. The synonymic relation between the two terms is confirmed by the contemporary lexical definition of orismology simply as »terminology«,[52,53] in harmony with the originators' intention; alternatively, »terminology« may be defined as a »scientific study of terms«[52] and »orismology« equivalently as »the science of defining technical terms«.[53] We thus fail to see any distinction in meaning.

In contradistinction to the rarely used and obsolete »orismology«, the synonymous »terminology« enjoys universal recognition. We live in a less fastidious age than Kirby and Spence, and hybrid words compounded of Latin and Greek roots, such as »terminology«, »sociology«, »mineralogy« and the like no longer cause offence. We therefore see no reason to rescue »orismology« from its present obscurity.

We now turn to the lexical definition of the noun »knot«. There are two central definitions,[52,53] knot_a and knot_b. Knot_a is a word of Old Norse origin that has been around for at least a millenium and that has, in the course of time, acquired numerous different meanings. All, however, are related, in one way or another, to the idea of an entanglement. Knot$_b$, references to which are found dating from the 15th c., is a stocky bird of the snipe family also called Red-breasted Sandpiper. »Highly sociable, knots stand almost body-to-body on the shore, moving like a carpet of birds as they feed.«[54] A truly fascinating scene, but our concern is with knot$_a$.

Knot$_a$ is associated with a variety of senses, knot$_{a1}$, knot$_{a2}$, *etc.* Knot$_{a1}$ is a compact interlacement or intertwinement of a string or rope. Knots$_{a1}$ are of immense practical importance and have existed from the time humans first used vines and cordlike fibers to bind stone heads to wood in primitive axes and to construct nets and traps. Knots$_{a1}$ are indispensable to the sailor: as Ashley observes in his richly illustrated book of nearly 4000 knots, »most important knots owe both their origin and their names to the requirements of a ship at sea.«[55] And the decorative aspect of knots$_{a1}$ is nowhere better illustrated than in the intricately interlaced and convoluted Celtic knotwork.[56]

The other senses of knot$_a$ may all be thought of as special cases, extensions, or connotations of knot$_{a1}$.

Knot$_{a2}$. The closed knot [a knot in the mathematical sense; see above].

Knot$_{a3}$. A tie, bond, linkage, connection, nexus.

Knot$_{a4}$. A tight cluster of persons or things, a bunch, an assemblage.

Knot$_{a5}$. A perplexing difficulty [as in »knotty problem« or in »cutting the Gordian knot«].

Knot$_{a6}$. A unit of velocity: one nautical mile per hour [from the rate at which a string, knotted at regular intervals, runs out as a ship moves forward].

Knot$_{a7}$. A lump, knob, knurl, node, bulge, protuberance, growth, excrescence [on the stem of a plant], the cross section on a tree trunk from which a branch growth out [whence: knothole].

Knot$_{a8}$. An elevated land region formed by the juncture of several mountain regions.

The best sense is the one that most aptly fits the context in which the word is used; in the present case the context is molecular structure. The choice among the senses can accordingly be narrowed by eliminating manifestly unsuitable candidates, such as knots$_{a5-a8}$. Knots$_{a3-a4}$ are on the borderline of acceptability. We are thus left with just two viable candidates: knot$_{a1}$ and knot$_{a2}$. In the next section we present the basis for a choice between these two.

Evolution of Scientific Terminology

In chemistry, as in all the sciences, the semantic content of technical terms changes as our understanding of concepts and phenomena deepens: meanings that are seen to be inappropriate in the light of new discoveries are pruned away, new subsenses are born, and so forth. That is, scientific terminology is dynamic and evolves along with the flowering of the subject.

An example is the change over time in the meaning of the term »acid«:[51a] from (1) the original meaning (a substance with a sour taste), to (2) a molecule with an easily separable hydrogen atom, to (3) a molecule that yields hydrogen ions when dissolved in water (Arrhenius' definition), to (4) a molecule that donates a proton to another substance (Bronsted and Lowry's definition), to (5) any species that can accept a pair of electrons (Lewis's definition).

A more complex example of terminological evolution is provided by the expression »aromatic compound«.[57] The term »aromatic« refers to a molecular property (»aromatic character« or »aromaticity«) that »is remarkable in organic chemistry for the variety of meanings and interpretations which have been ascribed to it in more than a century of use«.[58] The expression was originally applied by chemists in the early part of the 19th c. to describe compounds with an aromatic odor. In 1865, Kekulé used the terms »aromatic« and »aromaticity« with reference to the similarity in structure (cyclic conjugation) of benzene derivatives. Shortly thereafter (1866), Erlenmeyer proposed that the same terms be used with reference to the similarity in reactivity of the same class of compounds. In 1925, Armit and Robinson formulated the concept of the aromatic sextet and associated it with a common tendency of aromatic compounds to undergo substitution rather than addition reactions; this was followed by Hückel's $(4n + 2)$ π-electron rule (1931) for aromaticity. More recent criteria for aromaticity, based on early theories of Pauling (1936) and Pascal (1910), are the molecule's ability to sustain an induced ring current (Elvidge and Jackman, 1961) and diamagnetic susceptibility exaltation (Dauben, 1968). In a general way, »aromaticity« is thus associated with cyclic delocalization of electrons and with certain physical and/or chemical properties. There is, however, no crisp definition of »aromatic character« because there exists no consensus on how to measure it.

What these examples reveal is that whenever a word that is in general use, such as »acid« and »aromatic«, is adopted for use in a specific scientific context, its originally broad and often vague meaning acquires a narrowly defined technical significance. This redefinition is sometimes accompanied by a radical change in meaning: from a substance with a sour taste to a species that can accept a pair of electrons, or from a substance with an aromatic odor to a molecule capable of sustaining an induced ring current. The word »knot« has undergone a similar, if less dramatic, transition into the technical domain. In its sense of knot_{a1}, this word had been in circulation among the general public for centuries and had been broadly applied to any compact entanglement of a string or rope. The transition from this word, with its elastic meaning, to a sharply defined scientific term occurred in 1876, when the closed knot (knot$_{a2}$) was introduced as a mathematical object by the Scottish physicist Peter Guthrie Tait[59] (Tait's work on knots had been stimulated by the theory of vortex atoms advanced by another Scottish physicist, Sir William Thomson,[60] later Lord Kelvin). The definition of a closed knot, thanks to its mathematical underpinning, is the only well-defined sense of the word »knot« in a scientific context.

CONCLUSION

With one conspicuous exception, discussed in further detail below, whenever the term »knot« has been applied in chemistry, biochemistry, or molecular biology to describe the structure of molecules or molecular segments, it has been used in the rigorously defined mathematical sense of a closed knot. It would be unreasonable, however, to insist that a knot in a molecular structure must necessarily be closed; as Mansfield rightly points out, although we may be unable to define a knot in an open path, »we know one when we see one«.[9] Thus, »reasonable people« would agree that while an open knot, like the one in Figure 5a, is not a knot in the mathematical sense, it is a knot in the every-day sense. For the purpose of describing molecular structures, we therefore believe that it would be judicious to stretch the mathematical definition slightly by including open knots – closed knots that have been cut once but have not yet been unknotted. The »loosely knotted« carbonic anhydrase is such an example. [*Note added in proof:* A more recent example is the open knot observed in the polypeptide chain of MAT.[61]]

We therefore propose the following DEFINITION of a MOLECULAR KNOT: A molecular knot is a closed or open knot in a molecular graph, where »closed knot« and »open knot« are defined as above.

The exception mentioned above is exclusively limited to the protein literature, where it is widely spread and deeply entrenched. It consists of the use of »knot« in its generalized, nebulous sense of »entanglement« to describe a wide variety of structural motifs, sometimes even without benefit of direct evidence. Not only are the motifs depicted in Figures 7–12 not knots in the mathematical sense, but they even fail Mansfield's »reasonable-person test«.[9] They obviously do not qualify as molecular knots. Their description as »knots« therefore constitutes »improper« use of scientific terminology.

Acknowledgement – Support by the National Science Foundation is gratefully acknowledged.

REFERENCES

1. W. Kirby and W. Spence, *An Introduction to Entomology*, 2nd. ed., Longman *et al.*, London, 1816, Vol. 1, pp xii–xiii.
2. (a) C. Liang and K. Mislow, *J. Math. Chem.* **15** (1994) 1, 35, 245. C. Liang and K. Mislow, *ibid.* **16** (1994) 27. C. Liang and K. Mislow, *ibid.* **18** (1995) 1. C. Liang and K. Mislow, *J. Am. Chem. Soc.* **116** (1994) 3588, 11189. C. Liang and K. Mislow, *Biopolymers* **35** (1995) 343. (b) C. Liang and K. Mislow, *J. Am. Chem. Soc.* **117** (1995) 4201.
3. R. H. Crowell and R. H. Fox, *Introduction to Knot Theory*, Blaisdell Publ. Co., New York, 1963. C. Livingston, *Knot Theory*, Mathematics Association of America, Washington, DC, 1993.
4. C. C. Adams, *The Knot Book*, W. H. Freeman, New York, 1994.
5. C. O. Dietrich-Buchecker and J. -P. Sauvage, *Angew. Chem., Int. Ed. Engl.* **28** (1989) 189. See also: C. O. Dietrich-Buchecker, J. Guilhem, C. Pascard, and J. -P. Sauvage, *ibid.* **29** (1990) 1154. J. -P. Sauvage, *Acc. Chem. Res.* **23** (1990) 319. Ch. Dietrich-Buchecker and J. -P. Sauvage, *New J. Chem.* **16** (1992) 277. C. O. Dietrich-Buchecker, J. -P. Sauvage, J. -P. Kintzinger, P. Maltèse, C. Pascard, and J. Guilhem, *New J. Chem.* **16** (1992) 931. C. Dietrich-Buchecker and J. -P. Sauvage, *Bull. Soc. Chim. Fr.* **129** (1992) 113. C. O. Dietrich-Buchecker and J. -P. Sauvage, *Synthetic Molecular Knots*, in: V. Balzani and L. De Cola (Eds.), *Supramolecular Chemistry*, Kluwer Acad. Publ., Dordrecht, 1992, pp. 259–277. C. O. Dietrich-Buchecker, J. -F. Nierengarten, J. -P. Sauvage, N. Armaroli, V. Balzani, and L. De Cola, *J. Am. Chem. Soc.* **115** (1993) 11237. A. M. Albrecht-Gary, C. O. Dietrich-Buchecker, J. Guilhem, M. Meyer, C. Pascard, and J. P. Sauvage, *Recl. Trav. Chim. Pays-Bas.* **112** (1993) 427.
6. J. E. Mueller, S. M. Du, and N. C. Seeman, *J. Am. Chem. Soc.* **113** (1991) 6306. N. C. Seeman, *Mol. Eng.* **2** (1992) 297. S. M. Du and N. C. Seeman, *J. Am. Chem. Soc.* **114** (1992) 9652. H. Wang, S. M. Du, and N. C. Seeman, *J. Biomol. Struct. Dyn.* **10** (1993) 853. N. C. Seeman, J. Chen, S. M. Du, J. E. Mueller, Y. Zhang, T. -J. Fu, Y. Wang, H. Wang, and S. Zhang, *New J. Chem.* **17** (1993) 739. S. M. Du and N. C. Seeman, *Biopolymers* **34** (1994) 31. S. M. Du, H. Wang, Y. -C. Tse-Dinh, and N. C. Seeman, *Biochemistry* **34** (1995) 673. S. M. Du, B. D. Stollar, and N. C. Seeman, *J. Am. Chem. Soc.* **117** (1995) 1194.
7. (a) J. C. Wang, *Sci. Am.* **247/7** (1982) 94. M. A. Krasnow, A. Stasiak, S. J. Spengler, F. Dean, T. Koller, and N. R. Cozzarelli, *Nature* **304** (1983) 559. S. A. Wasserman and N. R. Cozzarelli, *Proc. Nat. Acad. Sci. U. S. A.* **82** (1985) 1079. S. A. Wasserman, J. M. Dungan, and N. R. Cozzarelli, *Science* **229** (1985) 171. (b) F. B. Dean, A. Stasiak, T. Koller, and N. R. Cozzarelli, *J. Biol. Chem.* **260** (1985) 4975. (c) S. A. Wasserman and N. R. Cozzarelli, *Science* **232** (1986) 951, and references therein. J. H. White, K. C. Millett, and N. R. Cozzarelli, *J. Mol. Biol.* **197** (1987) 585.
8. D. W. Sumners, *The Role of Knot Theory in DNA Research*, in: C. McCrory and T. Schifrin (Eds.), *Geometry and Topology*, Marcel Dekker, New York, 1987, pp. 297–318. C. O. Dietrich-Buchecker and J. -P. Sauvage, *Interlocked and Knotted Rings in Biology and Chemistry*, in: H. Dugas (Ed.), *Bioorganic Chemistry Frontiers*, Springer-Verlag, Berlin, 1991, Vol. 2, pp. 195–248. A. D. Bates and A. Maxwell, *DNA Topology*, Oxford University Press, New York, 1993.
9. M. L. Mansfield, *Nature Struct. Biol.* **1** (1994) 213.

10. G. E. Schulz and R. H. Schirmer, *Principles of Protein Structure*, Springer-Verlag, New York, 1979, pp. 86–87.
11. A. Grosberg and S. Nechaev, *Adv. Polymer. Sci.* **106** (1993) 1.
12. T. E. Creighton, *Prog. Biophys. Molec. Biol.* **33** (1978) 231.
13. R. Katahira, K. Shibata, M. Yamasaki, Y. Matsuda, and M. Yoshida, *Bioorgan. Med. Chem.* **3** (1995) 1273.
14. C. Oefner, A. D'Arcy, F. K. Winkler, B. Eggimann, and M. Hosang, *EMBO J.* **11** (1992) 3921.
15. M. P. Schlunegger and M. G. Grütter, *Nature* **358** (1992) 430. M. P. Schlunegger and M. G. Grütter, *J. Mol. Biol.* **231** (1993) 445.
16. N. Q. McDonald and W. A. Hendrickson, *Cell* **73** (1993) 421.
17. T. Meitinger, A. Meindl, P. Bork, B. Rost, C. Sander, M. Haasemann, and J. Murken, *Nature Genetics* **5** (1993) 376.
18. T. Noguti, T. Adachi-Yamada, T. Katagiri, A. Kawakami, M. Iwami, J. Ishibashi, H. Kataoka, A. Suzuki, M. Go, and H. Ishizaki, *LEBS Letters* **376** (1995) 251.
19. A. J. Lapthorn, D. C. Harris, A. Littlejohn, J. W. Lustbader, R. E. Canfield, K. J. Machin, F. J. Morgan, and N. W. Isaacs, *Nature* **369** (1994) 455. H. Wu, J. W. Lustbader, Y. Liu, R. E. Canfield, and W. A. Hendrickson, *Structure* **2** (1994) 545.
20. N. W. Isaacs, *Current Opin. Struct. Biol.* **5** (1995) 391.
21. P. D. Sun and D. R. Davies, *Annu. Rev. Biophys. Biomol. Struct.* **24** (1995) 269.
22. G. Venkataraman, V. Sasisekharan, C. L. Cooney, R. Langer, and R. Sasisekharan, *Proc. Natl. Acad. Sci. USA* **92** (1995) 5406.
23. P. Strasberg, H. A. Liede, T. Stein, I. Warren, J. Sutherland, and P. N. Ray, *Hum. Mol. Genet.* **4** (1995) 2179. D. E. Schuback, Z. Y. Chen, I. W. Craig, X. O. Breakefield, and K. B. Sims, *Hum. Mutation* **5** (1995) 285.
24. D. R. Holland, L. S. Cousens, W. Meng, and B. W. Matthews, *J. Mol. Biol.* **239** (1994) 385.
25. W. R. Moyle, R. K. Campbell, S. N. V. Rao, N. G. Ayad, M. P. Bernard, Y. Han, and Y. Wang, *J. Biol. Chem.* **270** (1995) 20020.
26. T. E. Hugli, *J. Biol. Chem.* **250** (1975) 8293.
27. S. Daopin, M. Li, and D. R. Davies, *Proteins: Struct., Funct., Genet.* **17** (1993) 176.
28. O. Saether, D. J. Craik, I. D. Campbell, K. Sletten, J. Juul, and D. G. Norman, *Biochemistry* **34** (1995) 4147.
29. D. C. Rees and W. N. Lipscomb, *Proc. Natl. Acad. Sci. USA* **77** (1980) 4633. D. C. Rees and W. N. Lipscomb, *J. Mol. Biol.* **160** (1982) 475.
30. D. Le Nguyen, A. Heitz, L. Chiche, B. Castro, R. A. Boigegrain, A. Favel, and M. A. Coletti-Previero, *Biochimie* **72** (1990) 431.
31. P. K. Pallaghy, K. J. Nielsen, D. J. Craik, and R. S. Norton, *Protein Science* **3** (1994) 1833.
32. C. Kellenberger, C. Boudier, I. Bermudez, J. G. Bieth, B. Luu, and H. Hietter, *J. Biol. Chem.* **270** (1995) 25514.
33. S. L. Lin and R. Nussinov, *Nature Struct. Biol.* **2** (1995) 835.
34. B. Blombäck, M. Blombäck, B. Hessel, and S. Iwanaga, *Nature* **215** (1967) 1445. S. A. Olexa and A. Z. Budzynski, *J. Biol. Chem.* **256** (1981) 3544.
35. J. A. Nagy, Y. C. Meinwald, and H. A. Scheraga, *Biochemistry* **24** (1985) 882.
36. B. Blombäck, M. Blombäck, A. Henschen, B. Hessel, S. Iwanaga, and K. R. Woods, *Nature* **218** (1968) 130.
37. V. J. Marder, A. Z. Budzynski, and H. L. James, *J. Biol. Chem.* **247** (1972) 4775. B. Blombäck, N. J. Gröndahl, B. Hessel, S. Iwanaga, and P. Wallén, *J. Biol. Chem.*

248 (1973) 5806. B. Kudryk, J. Reuterby, and B. Blombäck, *Eur. J. Biochem.* **46** (1974) 141. D. Collen, B. Kudryk, B. Hessel, and B. Blombäck, *J. Biol. Chem.* **250** (1975) 5808. A. Z. Budzynski, V. J. Marder, and S. Sherry, *Blood* **45** (1975) 757. G. D. Qureshi, V. P. Butler, Jr., E. J. Harfenist, R. E. Canfield, and H. L. Nossel, *Thromb. Res.* **6** (1975) 357. B. Kudryk and B. Blombäck, *Thromb. Diath. Haemorrh.* **34** (1975) 325. T. Takagi and R. F. Doolittle, *Thromb. Res.* **7** (1975) 813. R. E. Canfield, J. Dean, H. L. Nossel, V. P. Butler, Jr., and G. D. Wilner, *Biochemistry* **15** (1976) 1203. H. L. Nossel, V. P. Butler, Jr., G. D. Wilner, R. E. Canfield, and E. J. Harfenist, *Thromb. Haemost.* **35** (1976) 101. J. L. York and B. Blombäck, *Thromb. Res.* **8** (1976) 607. S. A. Olexa and A. Z. Budzynski, *J. Biol. Chem.* **254** (1979) 4925. J. N. Telford, J. A. Nagy, P. A. Hatcher, and H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **77** (1980) 2372. V. P. Butler, Jr., D. A. Weber, H. L. Nossel, D. Tse-Eng, K. S. LaGamma, and R. E. Canfield, *Blood* **59** (1982) 1006. T. M. Pozdnyakova, V. N. Rybachuk, and E. V. Vovk, *Ukr. Biokhim. Zh.* **55** (1983) 614. A. Shimizu, Y. Saito, A. Matsushima, and Y. Inada, *J. Biol. Chem.* **258** (1983) 7915. J. L. York, *Thromb. Res.* **31** (1983) 203. S. Birken, G. Agosto, B. Lahiri, and R. Canfield, *Thromb. Haemost.* **51** (1984) 16. C. Y. Liu, J. A. Koehn, and F. J. Morgan, *J. Biol. Chem.* **260** (1985) 4390. A. Bini, J. Fenoglio, Jr., J. Sobel, J. Owen, M. Fejgl, and K. L. Kaplan, *Blood* **69** (1987) 1038. J. W. Fenton II, T. A. Olson, M. P. Zabinski, and G. D. Wilner, *Biochemistry* **27** (1988) 7106. J. D. Loike, B. Sodeik, L. Cao, S. Leucona, J. I. Weitz, P. A. Detmers, S. D. Wright, and S. C. Silverstein, *Proc. Natl. Acad. Sci. USA* **88** (1991) 1044. J. Koopman, F. Haverkate, S. T. Lord, J. Grimbergen, and P. M. Mannucci, *J. Clin. Invest.* **90** (1992) 238. Y. I. Veklich, O. V. Gorkun, L. V. Medved, W. Nieuwenhuizen, and J. W. Weisel, *J. Biol. Chem.* **268** (1993) 13577. D. A. Meh, K. R. Siebenlist, G. Bergtrom, and M. W. Mosesson, *J. Lab. Clin. Med.* **125** (1995) 384.

38. M. K. Bodden, G. J. Harber, B. Birkedal-Hansen, L. J. Windsor, N. C. M. Caterina, J. A. Engler, and H. Birkedal-Hansen, *J. Biol. Chem.* **269** (1994) 18943. M. K. Bodden, L. J. Windsor, N. C. M. Caterina, G. J. Harber, B. Birkedal-Hansen, and H. Birkedal-Hansen, *Annal. N. Y. Acad. Sci.* **732** (1994) 403.
39. B. Kowalska-Loth, B. Gardlund, N. Egberg, and B. Blombäck, *Thromb. Res.* **2** (1973) 423.
40. S. Nagasawa, K. Mizuguchi, C. Ichihara, and J. Koyama, *J. Biochem.* **92** (1982) 1329.
41. M. Furlan, G. Kemp, and E. A. Beck, *Biochim. Biophys. Acta* **400** (1975) 95.
42. H. Rohde, G. Wick, and R. Timpl, *Eur. J. Biochem.* **102** (1979) 195.
43. B. Siebold, R. Deutzmann, and K. Kühn, *Eur. J. Biochem.* **176** (1988) 617.
44. C. Smith, H. Giordano, and R. DeLotto, *Genetics* **136** (1994) 1355.
45. B. Gardlund, B. Kowalska-Loth, N. J. Grondahl, and B. Blombäck, *Thromb. Res.* **1** (1972) 371. J. L. York and B. Blombäck, *J. Biol. Chem.* **254** (1979) 8786.
46. D. Frechet, J. D. Guitton, F. Herman, D. Faucher, G. Helynck, B. M. du Sorbier, J. P. Ridoux, E. James-Surcouf, and M. Vuilhorgne, *Biochemistry* **33** (1994) 42.
47. S. Huang, Z. Cao, and E. W. Davie, *Biochem. Biophys. Res. Commun.* **190** (1993) 488.
48. L. Carroll, *Alice's Adventures in Wonderland & Through the Looking Glass*, in: M. Gardner, *The Annotated Alice*, New American Library, 1960, pp. 268–271.
49. W. Shakespeare, *Romeo and Juliet*, Act II, ii, 43 (1596).
50. D. B. Zilversmit, *Persp. Biol. Med.* **7** (1964) 227.

51. (a) S. B. Elk, *J. Mol. Struct.* (*Theochem*) **313** (1994) 199. (b) S. B. Elk, *ibid.* **358** (1995) 119. S. B. Elk, *J. Chem. Inf. Comput. Sci.* **34** (1994) 325. S. B. Elk, *ibid.* **36** (1996) 385.
52. *The Oxford English Dictionary*, 2nd ed., Clarendon Press, Oxford, 1989.
53. *Webster's Third New International Dictionary*, G. & C. Merriam Co., Springfield, MA, 1964.
54. *Encylopaedia Britannica*, 15th ed., Chicago, 1978, Micropaedia vol. V, p. 859.
55. C. W. Ashley, *The Ashley Book of Knots*, Doubleday, New York, 1944.
56. G. Bain, *Celtic art: The methods of construction*, Dover, New York, 1973. P. R. Cromwell, *The Mathematical Intelligencer* **15** (1993) 36, and references therein.
57. J. March, *Advanced Organic Chemistry*, 4th ed., Wiley, New York, 1992, pp. 40–42 and references therein.
58. D. Lloyd and D. R. Marshall, *Angew. Chem., Int. Ed. Engl.* **11** (1972) 404. See also: D. Lloyd, *J. Chem. Inf. Comput. Sci.* **36** (1996) 442.
59. P. G. Tait, *Trans. Roy. Soc. Edin.* **28** (1876–77) 145.
60. W. Thomson, *Phil. Mag.* **34** (1867) 15.
61. F. Takusagawa and S. Kamitori, *J. Am. Chem. Soc.* **118** (1996) 8945.

SAŽETAK

Uzlaste strukture u kemiji, biokemiji i molekularnoj biologiji

Kurt Mislow i Chengzhi Liang

Pojam »uzao« u matematičkom smislu, kao zatvorena krivulja u trodimenzijskom prostoru koja ne presijeca samu sebe, ali koja se ne može položiti u ravninu bez presijecanja, iskorišten je da se opiše mnoštvo molekulskih struktura – izuzevši literaturu o proteinima, gdje je korišten za opisivanje strukturnih motiva koji nisu niti zatvoreni »uzlovi« u gornjem smislu niti pravi »uzao« u svakodnevnom smislu. U ovom radu se obrazlaže da opis ovih motiva kao »uzlima« predstavlja »neprikladno« korištenje znanstvene terminologije, pa se predlaže definicija »molekulski uzao«.