Cahn – Ingold – Prelog Descriptors of Absolute Configuration for Carbon Cages

André Rassat,*[a] Patrick W. Fowler,[b] and Benoît de La Vaissière[b]

Abstract: A simplified procedure is described for assigning Cahn – Ingold – Prelog descriptors to stereocentres in spheroalkanes (the C_nH_n molecules, with n even, based on trivalent, polyhedral carbon frameworks, a class which subsumes the fulleranes). By extension, similar descriptors can be found for the atoms of fullerenes and related carbon-only molecules. Assignment maps are given for chiral fullerenes C_{28}, C_{76}, C_{78}, C_{84} and C_{140}, and for a number of spheroalkanes. Cases of breakdown of the simple procedure for triangle-rich spheroalkane molecular graphs are discussed.

Keywords: cage compounds · Cahn – Ingold – Prelog rules · chirality · fullerenes · polyhedra

Introduction

Assignment of configurational descriptors for specification of stereogenic centres in molecules is, in principle, a solved problem. The powerful and general set of sequence rules constructed by Cahn, Ingold and Prelog (CIP) provides the standard description.[1] These rules, possibly with some further refinement,[2] are applicable to all molecular structures, including, therefore, the polyhedral carbon and hydrocarbon cages that are the subjects of this paper. Application of the rules to fullerenes and their derivatives is, however, perceived to be cumbersome, to the extent that alternative descriptions have been devised specifically for use in this area.[3] The purpose of the present paper is to demonstrate that, at least for many fully hydrogenated such cages (spheroalkanes[4] and fulleranes), the CIP assignment is readily computed from purely graph-theoretical considerations. An algorithm is described and applied to some representative spheroalkanes and fulleranes, with a simple expedient for extension of the procedure to the fullerenes themselves. Although the simple procedure works well for fulleranes and fullerenes, some problems remain for triangle-rich spheroalkanes.

[a] Prof. A. Rassat
UMR CNRS 8640, Département de Chimie
Ecole Normale Supérieure
24 rue Lhomond, 75231 Paris Cedex 05 (France)
Fax: (+33) 1-44-32-33-25
E-mail: andre.rassat@ens.fr

[b] Prof. P. W. Fowler, B. de La Vaissière
School of Chemistry, University of Exeter, Stocker Road
Exeter, EX4 4QD (UK)
Fax: (+44) 1392-263434
E-mail: p.w.fowler@exeter.ac.uk

Definitions: A *spheroalkane*[4] is the fully hydrogenated form of a trivalent, polyhedral carbon framework. When the faces of the polyhedron are all either pentagonal or hexagonal, the spheroalkane is a *fullerane*, the fully hydrogenated form of a *fullerene*.[5] Some sources extend the definition of fulleranes to cover all spheroalkanes of 20 or more vertices, irrespective of face size.[6] The carbon centre in a spheroalkane is directly bonded to three carbon neighbours and to one hydrogen atom. Each hydrogen atom may lie inside or outside the cage; this leads to many isomeric possibilities, but here we consider spheroalkanes in their "*normal*" form, in which all CH bonds are *exo* to the cage. On this understanding, the graph may be "pruned" of hydrogens, which can be restored later without ambiguity, and the only criterion for distinguishing carbon centres (CH units) is then the connectivity of the underlying framework. A pruned spheroalkane graph is identical to the graph of a *spheroarene*.[4] An atlas of all possible spheroalkanes with up to 16 carbon atoms is given in [7] and adjacency information for larger structures is readily generated, for example with the *plantri* program.[8]

Cubic graphs are those in which each vertex is joined by edges to three neighbours (in a chemical context these are known as *trivalent* graphs); *three-connected* graphs are those for which the removal of at least three vertices is required to disconnect the graph; *planar* graphs are those which can be drawn in the plane without the crossing of edges; *polyhedral* graphs are those which are both planar and three-connected.[9] Thus, the problem of assignment of descriptors in these C_nH_n frameworks is concerned with cubic, polyhedral graphs embedded in three-dimensional space.

At first sight, fulleranes and spheroalkanes are rather unpromising subjects for the application of the CIP rules, as each carbon to be labelled is surrounded by a sea of chemically equivalent near and distant neighbours. There are indeed

  0947-6539/01/0718-3985 $ 17.50+.50/0

special problems in distinguishing such "nearly equivalent" sites. The uniformity of the structure, however, implies that the assignment process will largely depend on graph-theoretical considerations, and may in fact be more readily programmed than for some more chemically varied structures.

The process of attaching CIP labels in these arrays of chemically identical subunits can be separated into two logically distinct steps. The first is purely *combinatorial*, in which priorities one to four are assigned to the four ligands around each stereogenic centre. As H has a lower priority than any carbon atom (Rule 1[1c]), this task for C_nH_n reduces to the assignment of large, medium and small priorities to the three carbon neighbours of each centre. The second is *geometrical* or *affine* in nature: once the labels are found, the choice of descriptor from the set {*R*, *S*} is determined by the "steering-wheel rule",[1a] which depends on the relative positions in space of the ligands to which we have attached priorities L(arge) > M(edium) > S(mall) > (H), that is, on the *embedding* of the graph in space.

A polyhedron can be drawn in the plane as a *Schlegel diagram*.[10] Several definitions of this construction exist in the literature.[11] Here we take the point of view that a Schlegel diagram is constructed by seating a polyhedron on a face and flattening the whole object so that the bottom face expands to contain the rest.[5] It may sometimes be convenient to stand the polyhedron on a vertex or an edge, in which case the Schlegel diagram will have a vertex or edge midpoint at infinity. Notice that we started by looking at the three-dimensional object from outside and above, so that the central portion of the diagram represents components of structure that were initially closest to the viewer. With this convention, the drawing encodes both the connectivity and, in the case of a chiral polyhedron, the enantiomeric identity; that is, both the combinatorial and geometrical aspects. Mirror-image Schlegel diagrams represent opposite enantiomers and may be interconverted by lifting them out of the drawing plane and turning them over, or by breaking and re-forming edges, but not by using solely operations that preserve connectivity and constrain all vertices to the plane (see Figure 1).

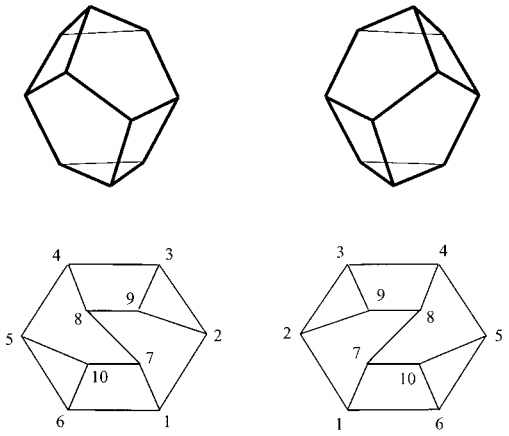


Figure 1. Enantiomers of a C_2-symmetric spheroalkane, represented as three-dimensional and Schlegel diagrams of the underlying carbon framework.

An algorithm: An algorithm is readily devised from a perusal of the CIP rules,[1] and it turns out that the cases of interest are small enough to be handled without any special attention to optimisation of the program. In fact, the specific case of descriptor assignment for spheroalkanes was considered by Dreiding and Prelog in correspondence a decade ago.[12]

Note that the simplified rules described below are a condensation of the general CIP rules for the specific case of polyhedral arrays of equivalent units such as C_nH_n. They are easily adapted to derivatives such as C_nF_n and so on, where *R* (*S*) centres in C_nH_n become *S* (*R*) in C_nF_n. They do not apply to more general molecules.

1) Starting from a Schlegel diagram or other representation of the molecular connectivity, construct an adjacency list for the pruned polyhedron in some freely chosen vertex-numbering scheme.
2) For each stereocentre, C_1, say, construct a *rooted tree* based on the atom. (Figure 2) The *nodes* of the tree carry vertex labels from the graph and are arranged on concentric "s*pheres*" S_i. S_0 is C_1 itself; S_1 contains the three *neighbours* of C_1, *u*, *v* and *w*, each joined to C_1 by an edge of the graph. The three sub-trees starting from *u*, *v*, *w* are

Abstract in French: *On décrit un procédé simplifié par rapport à la méthode générale de Cahn, Ingold et Prelog (CIP), et spécialement adapté à l'attribution des descripteurs stéréochimiques R ou S aux stéréocentres des sphéroalcanes (molécules C_nH_n (n pair), dont le squelette carboné forme un polyèdre trivalent, et dont les fulleranes constituent une classe particulière). On étend la définition de ces descripteurs aux atomes des fullerènes et autres cages polyédriques de carbone (sphéroarènes). Les figures donnent les attributions pour les fullerènes chiraux C_{28}, C_{76}, C_{78}, C_{84} et C_{140}, ainsi que pour d'autres sphéroalcanes. On présente enfin certains sphéroalcanes riches en triangles, dans lesquels il est impossible d'attribuer un descripteur à tous les stéréocentres, et où on atteint les limites du procédé décrit ici.*

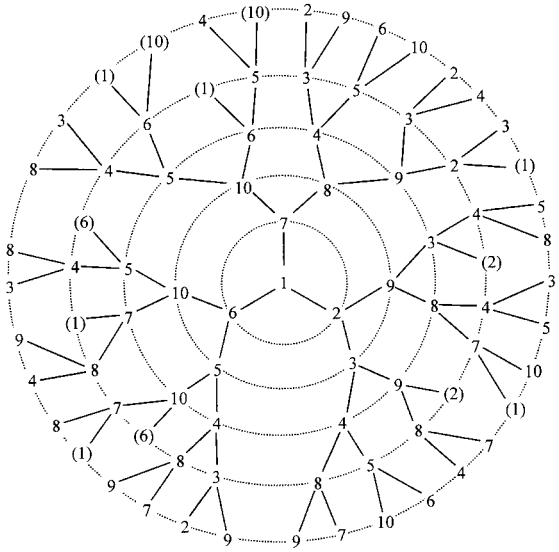


Figure 2. Rooted tree based on vertex 1 of the Schlegel diagrams of Figure 1.[12]

0947-6539/01/0718-3986 $ 17.50+.50/0 *Chem. Eur. J.* **2001**, *7*, No. 18

the *branches* B_u, B_v, B_w The tree grows outwards by bifurcation at each node, so that each node x on sphere S_i ($i > 0$) is joined by an edge to one predecessor on S_{i-1} and (if x is not a duplicate node) to each of two successors on S_{i+1} The concept of a duplicate node is defined in the following way: each x on sphere S_k defines a path of length k leading from C_1; if this path closes at x, then x is a *duplicate node* and is either a copy of the root atom itself (closure as a *cycle*) or of a vertex already encountered on the path (closure as a "*tadpole*"—a cycle with a tail). S_1 and S_2 contain no duplicate nodes, as all cycles in polyhedra are of size three or more. In principle, the tree grows radially out from the root, until every path has terminated with a duplicate node. In practice the assignment of priorities to branches B_u, B_v, B_w normally requires explicit construction of only a part of the whole tree.

3) The CIP ordering of branches B_u, B_v, B_w as of L(arge), M(edium) and S(mall) rank (L > M > S) is done on the basis of the number of duplicate nodes found in each branch on successive spheres.
   Let the first sphere on which at least one duplicate node is found be S_j, and let $N_j(u)$, $N_j(v)$ and $N_j(w)$ be the numbers of duplicate nodes on this sphere in the branches B_u, B_v and B_w, respectively:
   a) If $N_j(u) < N_j(v) < N_j(w)$, then u is L, v is M and w is S;
   b) If $N_j(u) = N_j(v) < N_j(w)$, then w is S. To fix the other priorities, one has to go further, to sphere m ($m > j$), the first sphere where a difference between B_u and B_v is found, with, say, $N_m(u) < N_m(v)$, and then u is L and v is M;
   c) If $N_j(u) < N_j(v) = N_j(w)$, then u is L. To fix the other priorities, one has to go further to sphere m ($m > j$), the first sphere where a difference between B_v and B_w is found, with, say, $N_m(v) < N_m(w)$, and then v is M and w is S.

Remark: The symmetry of the rooted tree is the same as the site symmetry of the root atom, which, in a trivalent polyhedron, is C_1 or C_3 (*chiral* sites in a chiral molecule, or *prochiral* sites in an achiral molecule), and C_s or C_{3v} (*achiral* sites in an achiral molecule). Trivially, the process described must fail to terminate with an LMS ordering of the three branches when C_1 is at an achiral site, but it may also fail to deliver an ordering in some cases to be discussed below.

4) Now go on to atom C_2, and so on. If the molecule has nontrivial symmetry, it is only necessary to treat one atom from each *orbit* (set of symmetry-equivalent atoms), and, of course, only chiral and prochiral orbits need to be considered. In chiral molecules, all members of an orbit carry the same descriptor. In achiral molecules, prochiral atoms that exchange under proper operations have the same label *R* or *S*, and prochiral atoms that exchange under improper operations have opposite labels. Thus, in an achiral molecule, each orbit of prochiral sites is equally balanced between *R* and *S*.

Remark: Up to this point no information beyond the adjacency list of the graph of the carbon skeleton has been used. Steps 1 to 4 of the algorithm are purely combinatorial. In the next step, the LMS triple is converted to an {*R*, *S*} descriptor of absolute configuration.

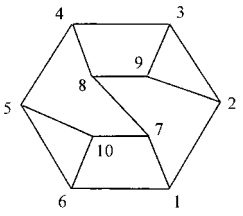
5) From the coordinates, the Schlegel diagram, a three-dimensional model or a picture, determine the clockwise or anticlockwise orientation of the LMS triple viewed from C_1 out to its attached *exo* H atom. A convenient mnemonic for this is the CIP "steering-wheel rule".[1a] A more formal device for the same purpose is the (pseudo)-scalar triple product:
   $P_{LMS} = (\mathbf{r}_L \times \mathbf{r}_M) \cdot \mathbf{r}_S$
   in which $\mathbf{r}_L$, $\mathbf{r}_M$ and $\mathbf{r}_S$ are position vectors of the L, M and S ligands, drawn from a common origin *inside* the cage; the sign of P_{LMS} determines the descriptor, positive for *R* and negative for *S*.

Remark: A useful way to obtain plausible three-dimensional coordinates for spherical polyhedral structures directly from adjacency information has been described in the fullerene literature under the name of "*topological coordinates*",[13] Cartesian coordinates deduced directly from the *p*-orbital-like eigenvectors of the adjacency matrix. They may then be scaled to give chemically reasonable bond lengths, but are already sufficient for the purpose of computing the sign of the triple product P_{LMS} as they give the bonded partners of each atom as its physically nearest neighbours. Topological coordinates are used in all the examples quoted in the present paper.

Extension to fully unsaturated cages (spheroarenes): In spheroarenes (and fullerenes), the carbon atoms occupy three-coordinate sites and the CIP rules for tetrahedral stereocentres, strictly speaking, do not apply. An obvious and simple expedient is to take the spheroarene as a pruned, normal-form (fully *exo*) spheroalkane and to apply the descriptors for corresponding carbon sites directly. Descriptors obtained in this way would not distinguish between the different possible coordination *geometries* at a spheroarene carbon, pyramidal pimple or inward dimple, but they would encode the sense of the triangle of neighbours. Incidentally, the same descriptors would be retained for endohedral fullerenes X@C_n, in which X is a high-priority "neighbour" on the inside of the cage and replaces the low-priority *exo*-hydrogen atoms of the normal fullerane.

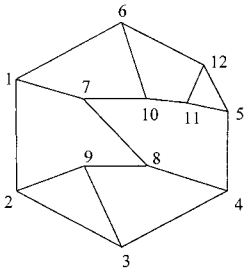
Proposed IUPAC conventions for chiral fullerenes[5b] allow for the assignment of an overall descriptor *C* or *A* to different enantiomers. Given a standard numbering scheme for the fullerene vertices, also envisaged in the conventions, knowledge of the CIP {*R*, *S*} label assigned as above for any one given numbered vertex is tantamount to fixing the {*C*, *A*} label.

Complexity: The above description gives the essentials of a programmable algorithm of CIP assignment for spheroalkanes. Computation of the numbers of nonbacktracking, non-self-crossing paths in a graph has the potential to grow alarmingly with the total number of vertices, and if resolution of branch priority requires construction of a substantial portion of the whole rooted tree for each vertex, the process is bound to be costly. Initially the number of nodes on sphere



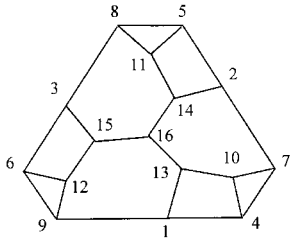
	L	M	S	
1	7	2	6	*S*
2	1	3	9	*R*
3	4	2	9	*S*
4	8	5	3	*S*
5	4	6	10	*R*
6	1	5	10	*S*
7	8	1	10	*S*
8	7	4	9	*S*
9	8	2	3	*R*
10	7	5	6	*R*

Figure 3. Assignment of chiral descriptors for an enantiomer of the C_2-symmetric $C_{10}H_{10}$, barettane, shown in Figure 1.



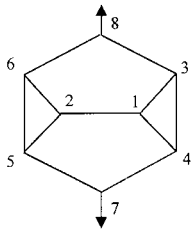
	L	M	S	
1	7	6	2	*R*
2	1	3	9	*R*
3	4	2	9	*S*
4	8	5	3	*S*
5	4	12	11	*R*
6	1	10	12	*R*
7	8	1	10	*S*
8	7	4	9	*S*
9	8	2	3	*R*
10	7	6	11	*S*
11	10	5	12	*R*
12	6	5	11	*S*

Figure 4. Assignment of chiral descriptors for an enantiomer of the C_1-symmetric $C_{12}H_{12}$ spheroalkane.

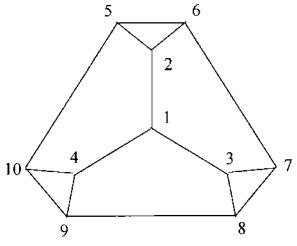


	L	M	S	
1	13	9	4	*R*
2	14	7	5	*R*
3	15	8	6	*R*
4	1	7	10	*R*
5	2	8	11	*R*
6	3	9	12	*R*
7	2	10	4	*R*
8	3	11	5	*R*
9	1	12	6	*R*
10	13	7	4	*S*
11	14	8	5	*S*
12	15	9	6	*S*
13	16	1	10	*R*
14	16	2	11	*R*
15	16	3	12	*R*
16	15	14	13	*S*

Figure 5. Assignment of chiral descriptors for an enantiomer of the C_3-symmetric $C_{16}H_{16}$ spheroalkane.

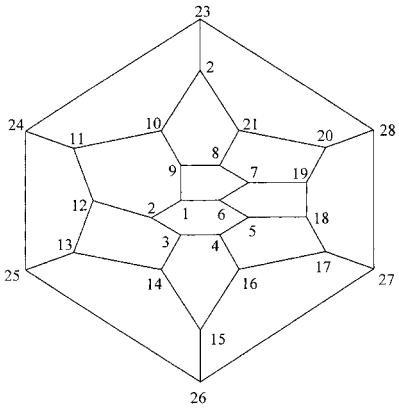


	L	M	S	
3	8	1	4	*S*
4	7	1	3	*R*
5	7	2	6	*S*
6	8	2	5	*R*



	L	M	S	
1				
2	1			
3	1			
4	1			
5	10	2	6	*R*
6	7	2	5	*S*
7	6	3	8	*R*
8	9	3	7	*S*
9	8	4	10	*R*
10	5	4	9	*S*

Figure 6. Assignment of chiral descriptors for the C_{2v}-symmetric C_8H_8, cuneane, and the C_{3v}-symmetric $C_{10}H_{10}$, diademane or mitrane.



	L	M	S	
1	2	6	9	*S*
2	1	12	3	*S*
3	2	4	14	*R*
4	5	3	16	*S*
5	6	4	18	*S*
6	1	5	7	*S*
7	6	19	8	*S*
8	9	7	21	*S*
9	1	10	8	*R*
10	11	9	22	*S*
11	12	10	24	*S*
12	2	11	13	*S*
13	12	25	14	*S*
14	3	13	15	*S*
15	26	16	14	*S*
16	17	4	15	*S*
17	27	18	16	*S*
18	17	19	5	*S*
19	18	20	7	*S*
20	28	19	21	*R*
21	20	22	8	*S*
22	23	10	21	*S*
23	28	24	22	*S*
24	23	25	11	*S*
25	24	26	13	*S*
26	27	25	15	*R*
27	28	17	26	*S*
28	27	23	20	*S*

Figure 7. Assignment of chiral descriptors for the hypothetical D_2-symmetric C_{28} fullerene.

S_k grows as 3.2^{k-1}, and for a Hamiltonian graph (one with a cycle containing all vertices)[9] for example, the rooted tree will certainly have as many spheres as the graph has vertices. A priori, it is not possible to rule out a case where it would be necessary to construct the whole tree to complete the assignment, at least for some awkward vertex.

It is already clear that the decision process may sometimes require consideration of a very large number of spheres. As an example, consider a large icosahedral fullerene framework, initially of full I_h symmetry, in which the twelve pentagons are separated by large triangular graphitic regions. Now perturb the framework by a Stone–Wales[14] bond rotation at a prochiral site in one of the large triangular faces, introducing paired pentagonal and heptagonal rings. Sites which were achiral in the unperturbed cage now become chiral, but for those which are approximately antipodal to the site of the perturbation, the length of path required to fix the {*R*, *S*} assignments will be of the order of the *diameter*[9] of the graph.

Another specific pattern that is likely to lead to long paths is one where the root vertex and two of its neighbours, *u* and *v*, say, all belong to the same orbit, that is, are all equivalent in the point group of the polyhedral framework. Equivalence of *u* and *v* is lifted in the rooted tree, but the distinction between the two branches involved will probably remain "small" and become apparent only at path lengths that span a large part of the polyhedron, if indeed the distinction appears at all.

It has already been noted in the description of the algorithm that the process of searching for differences between branches in their numbers of nonduplicate nodes, sphere by sphere, does not terminate in a decision for all stereocentres.

The first case of nontermination is the trivial one in which the root of the tree is an achiral site.

The next case of nontermination is when all three branches are equivalent by symmetry. When the site symmetry of the root vertex is C_3, the three branches are of equal weight on all spheres S_k, and it is necessary to invoke a further sequence

0947-6539/01/0718-3988 $ 17.50+.50/0 *Chem. Eur. J.* **2001**, *7*, No. 18

rule (see the final sequence rule, §6.3 of [1c]). In this case, assignment proceeds by arbitrarily giving highest rank to one of the three rotationally equivalent neighbours (u, say) and then deciding between the other two branches on the basis of the more frequent of occurrence of u on the first sphere at which there is a difference in numbers of copies of u. This usually gives a decision, but there are examples of molecular graphs where this simple enumeration of copies of u fails to give an assignment (see Figure 13, later).

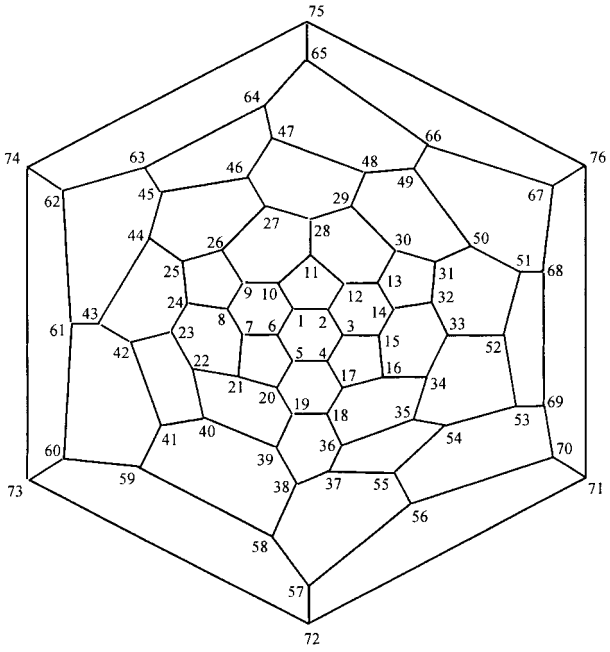
A third possibility for nontermination is that a rooted tree could have two or even three branches that coincided in numbers of nonduplicate nodes at every level, even though not forced to do so by symmetry equivalence. In such cases an auxiliary rule is needed (see Figure 12 and 13, later). At least for the isomers of fullerenes and fulleranes considered here, this problem does not arise, and complete assignment is possible.

Results

The algorithm is implemented in a FORTRAN program which takes adjacency data as input. Framework coordinates are calculated by diagonalisation of the adjacency matrix and selection of eigenvectors. Rooted trees are constructed from the centre out to a sufficiently remote sphere, LMS priorities assigned and triple products used to convert them to {R, S} descriptors as described above. The numbering schemes are shown on the Schlegel diagrams. In the case of the larger fulleranes/fullerenes these are the IUPAC standard schemes.[5b] Descriptor assignments are shown as accompanying tabulations. A bracketed entry denotes the use of the extra rule for a C_3 site.

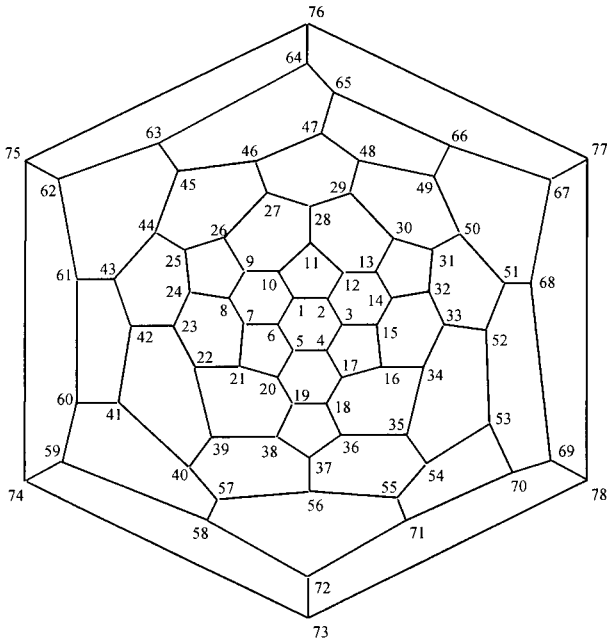
Figures 3 to 5 show specific enantiomers of small spheroalkanes with {R, S} attributions. The C_2-symmetric "barettane"[15] (Figure 3) was the example assigned by Dreiding and Prelog.[12] Paths of length of at most five are needed in this case. Figure 4 shows a C_1-symmetric spheroalkane obtained by truncating one vertex of the previous structure. Figure 5 shows a C_3-symmetric spheroalkane for which all vertices but one fall into isochiral orbits of size three, and the unique vertex (16) requires invocation of the special procedure for sites with all ligands equivalent. Figure 6 includes the orbits of prochiral sites in C_{2v} cuneane[16] with two R and two S sites, and C_{3v} diademane[17] or mitrane[18] with three R and three S sites. These examples serve to illustrate the procedure.

The hypothetical D_2 isomer of C_{28} (Figure 7) is the smallest intrinsically chiral fullerene.[5a] Figures 8 to 10 show three experimentally characterised chiral fullerenes,[19] isomers of C_{76} (D_2), C_{78} (D_3) and C_{84} (D_2). Comparison with [5b] shows



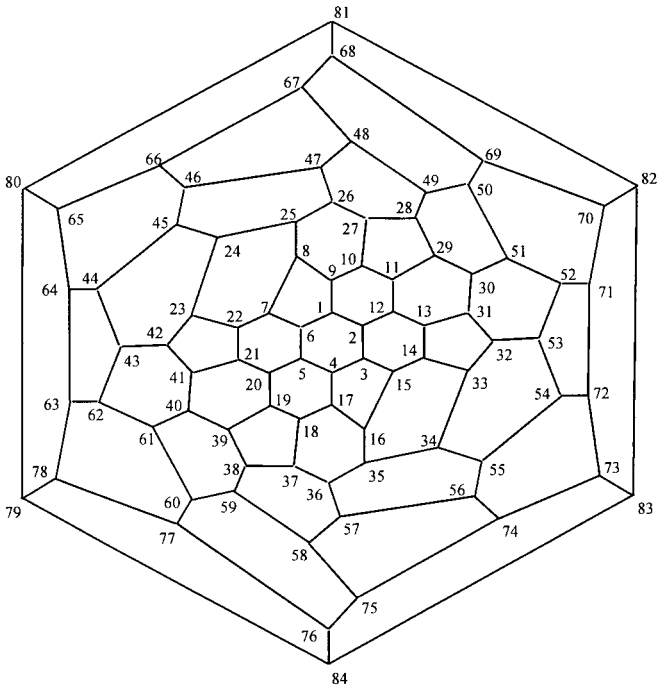
	L	M	S	
1	6	10	2	*S*
2	3	12	1	*R*
3	2	15	4	*S*
4	5	17	3	*R*
5	4	20	6	*S*
11	28	12	10	*S*
12	13	11	2	*R*
13	12	30	14	*S*
14	15	32	13	*R*
15	14	16	3	*S*
16	34	17	15	*S*
17	18	16	4	*R*
18	19	36	17	*R*
19	39	18	20	*R*
29	48	28	30	*R*
30	29	31	13	*S*
31	50	30	32	*R*
32	33	31	14	*R*
33	34	32	52	*S*

Figure 8. Assignment of chiral descriptors for the D_2-symmetric C_{76} isolated-pentagon fullerene (*A* enantiomer); here and on the following four figures, one representative is shown for each orbit of atomic sites.



	L	M	S	
1	6	10	2	*R*
2	3	12	1	*S*
7	8	21	6	*S*
9	10	26	8	*S*
10	9	11	1	*R*
11	28	10	12	*S*
22	39	23	21	*R*
23	22	42	24	*R*
24	23	25	8	*R*
25	44	26	24	*R*
26	27	25	9	*S*
27	46	28	26	*R*

Figure 9. Assignment of chiral descriptors for the D_3-symmetric C_{78} isolated-pentagon fullerene (*C* enantiomer).



	L	M	S	
1	2	6	9	*R*
2	12	1	3	*S*
3	2	4	15	*S*
7	22	8	6	*R*
8	25	9	7	*R*
9	10	8	1	*S*
10	9	27	11	*R*
11	12	29	10	*S*
12	2	11	13	*R*
13	12	31	14	*R*
14	15	33	13	*S*
23	24	42	22	*S*
24	25	45	23	*S*
25	24	26	8	*R*
26	25	47	27	*R*
27	26	28	10	*R*
28	49	29	27	*R*
29	30	28	11	*S*
30	51	29	31	*S*
31	30	13	32	*S*
32	53	31	33	*S*

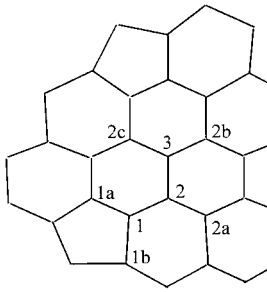
Figure 10. Assignment of chiral descriptors for the D_2-symmetric C_{84} (*A* enantiomer), isolated-pentagon fullerene 84:22 (spiral nomenclature[5a]).

that the forms illustrated here are, respectively, the *A*, *C* and *A* enantiomers.

Figure 11 shows the repeat unit of the smallest chiral icosahedral fullerene C_{140},[20, 21] viewed from outside the carbon sphere. The 140 vertices can be broken down into two sets of 60 (C_1 site symmetry) and one of 20 (C_3 site symmetry). Each of the 60 pentagon vertices has two neighbours in the same orbit, and resolution requires paths of length 18. For the orbit of the 60 vertices *exo* to the pentagons, the required paths are of length seven, and for the remaining 20 vertices the C_3 rule is invoked. Again the {*R*, *S*} label of any one of the carbons of this structure is sufficient to fix the enantiomer of both fullerane (in normal form) and fullerene.

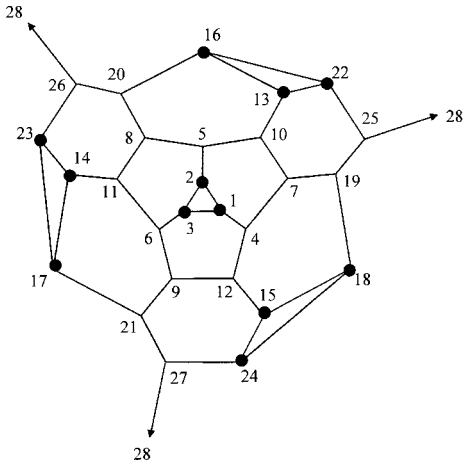
Problem cases: Figure 12 shows a bad case for step 3 of the algorithm. This is a 28-vertex *T*-symmetric polyhedral cage, in fact the smallest spheroalkane of this symmetry.[4, 21] It has four triangular and twelve hexagonal faces, and the vertices fall into two orbits of twelve (site-symmetry C_1) and one of four (site-symmetry C_3). Application of step 3 of the algorithm gives immediate assignment for the twelve vertices *exo* to the triangles. The special procedure for C_3 sites gives an assignment for the set of four vertices at the centres of the tetrahedral faces. Each member of the 12-orbit of triangle vertices has two neighbours in its own orbit and one in the *exo* orbit. It is clear that the *exo* branch has L(arge) priority. However, computation shows that two branches of the rooted tree remain of equal weight at each sphere from the first all the way to the last (S_{27}), even though these branches are *not* equivalent by symmetry. These vertices remain unassigned with the simple algorithm.

Systematic truncation of this graph yields further problem cases. Figure 13 shows 10 truncations that preserve C_3 symmetry and lead to C_{34} spheroalkane skeletons. In all 10



	L	M	S	
1	2	1a	1b	*R*
2	3	2a	1	*S*
3	2	2b	2c	*R*

Figure 11. Assignment of chiral descriptors for the *I*-symmetric C_{140} isolated-pentagon fullerene. Only a portion of the surface sufficient to show the environment of the sites belonging to the three orbits is shown, as seen from outside the molecule.



	L	M	S	
1	4			
2	5			
3	6			
4	7	12	1	*S*
5	8	10	2	*S*
6	9	11	3	*S*
7	19	10	4	*R*
8	20	11	5	*R*
9	21	12	6	*R*
10	7	5	13	*S*
11	8	6	14	*S*
12	9	4	15	*S*
13	10			
14	11			
15	12			
16	20			
17	21			
18	19			
19	7	25	18	*S*
20	8	26	16	*S*
21	9	27	17	*S*
22	25			
23	26			
24	27			
25	28	19	22	*S*
26	28	20	23	*S*
27	28	21	24	*S*
28	27	26	25	*R*

Figure 12. A chiral, tetrahedrally symmetric $C_{28}H_{28}$ (smallest spheroalkane of *T* symmetry) spheroalkane, in which some sites (marked with a black spot) of C_1 symmetry are not fully assigned by step three of the algorithm.

0947-6539/01/0718-3990 $ 17.50+.50/0 *Chem. Eur. J.* **2001**, *7*, No. 18

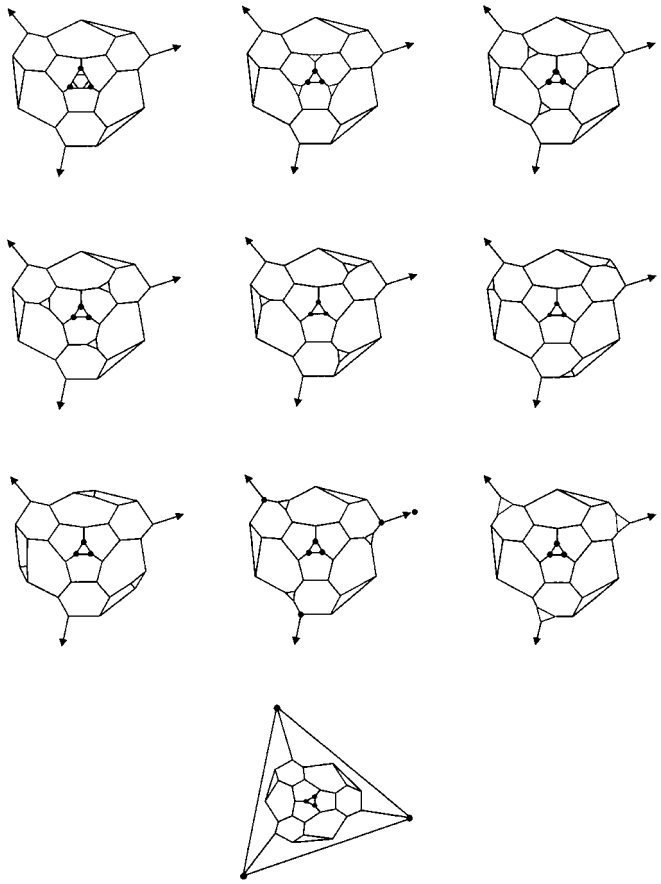


Figure 13. Further C_3-symmetric polyhedra, created by truncation of the tetrahedral $C_{28}H_{28}$ graph, all of which give incomplete assignments with the simple algorithm. Unassigned vertices are marked with a black spot.

there are sets of vertices in triangles that are unassigned even when step three of the algorithm is taken to the last sphere, S_{33}, and in one case the atom on the C_3 axis itself and its nearest neighbours are also unassigned, notwithstanding the application of the weighted counting rule for C_3 site symmetries. Although it is risky to generalise from such a small sample, it is perhaps noteworthy that all these failures occur in graphs with many triangles. The equivalent of Figure 12 with six squares (*O*-symmetric C_{56}) or twelve pentagons (the aforementioned *I*-symmetric C_{140}), the smallest spheroalkanes of these symmetries,[4, 21, 22] are both fully assignable by a combination of step three and the C_3-rule. Further investigation of the problem cases is indicated, though the simple algorithm suffices for our main purpose of assigning CIP descriptors to fullerenes.

Acknowledgement

We thank Prof. A. Dreiding for helpful comments and a detailed explanation of his own implementation of the CIP rules for spheroalkanes. This work was supported by the European Union, TMR Network "BIOFULLERENES"; Contract grant number: FMRX-CT98-0192.

[1] a) R. S. Cahn, C. K. Ingold, V. Prelog, *Experientia* **1956**, *12*, 81–124; b) R. S. Cahn, C. K. Ingold, V. Prelog, *Angew. Chem.* **1966**, *78*, 413–447; *Angew. Chem. Int. Ed. Engl.* **1966**, *5*, 385–415; c) V. Prelog, G. Helmchen, *Angew. Chem.* **1982**, *94*, 614–631; *Angew. Chem. Int. Ed. Engl.* **1982**, *21*, 567–583.
[2] a) P. Mata, A. M. Lobo, C. Marshall, A. P. Johnson, *Tetrahedron: Asymmetry* **1993**, *4*, 657–668; b) R. H. Custer, *Match* **1986**, *21*, 3–31.
[3] a) Y. Nakamura, M. Taki, J. Nishimura, *Chem. Lett.* **1995**, 703–704; b) C. Thilgen, A. Herman, F. Diederich, *Helv. Chem. Acta* **1997**, *80*, 183–199.
[4] a) A. Rassat, D. Seroussi, C. Coulombeau, *J. Chim. Phys. Phys.–Chim. Biol.* **1994**, *91*, 1683–1710; b) A. Rassat, G. Scalmani, D. Seroussi, G. Berthier, *THEOCHEM* (*J. Mol. Struct.*) **1995**, *338*, 31–41; c) A. Rassat in *Crystallography of Supramolecular Compounds* (Eds.: G. Tsoucaris, J. L. Atwood, J. Lipkowski), Kluwer, Dordrecht, **1996**, pp. 181–201; d) A. Rassat, *Chirality* **2001**, in press.
[5] a) P. W. Fowler, D. E. Manolopoulos, *An Atlas of Fullerenes*, Oxford University Press, Oxford, **1995**; b) E. W. Godly, R. Taylor, *Pure Appl. Chem.* **1997**, *69*, 1411–1434.
[6] A. L. Goodson, C. L. Gladys, D. E. Worst, *J. Chem. Inf. Comp. Sci.* **1995**, *35*, 969–978.
[7] P. W. Fowler, D. Mitchell, *J. Chem. Soc. Faraday Trans.* **1996**, *92*, 4145–4150.
[8] G. Brinkmann, *J. Graph Theory* **1996**, *23*, 139–149. The *plantri* program is available from http://cs.anu.edu.au/people/bdm/plantri/.
[9] R. J. Wilson, *Introduction to Graph Theory*, 3rd ed., Longman Scientific, Harlow, **1995**.
[10] V. Schlegel, *Verhandlungen (Nova Acta) der Kaiserlichen Leopoldinisch-Carolinischen Deutschen Akademie der Naturforscher* **1883**, *44*, 343–359.
[11] a) H. M. S. Coxeter, *Regular Polytopes*, 3rd ed., Dover Publications, New York, **1973**; b) A. J. Wells, *The Third Dimension in Chemistry*, Oxford University Press, Oxford, **1956**.
[12] Personal communication from V. Prelog and A. Dreiding to A. Rassat, December 17 **1991**.
[13] D. E. Manolopoulos, P. W. Fowler, *J. Chem. Phys.* **1992**, *96*, 7603–7614.
[14] A. J. Stone, D. J. Wales, *Chem. Phys. Lett.* **1986**, *128*, 501–503.
[15] a) D. Bosse, A. De Meijere, *Tetrahedron Lett.* **1977**, 1155–1158; b) D. Bosse, A. De Meijere, *Chem. Ber.* **1978**, *111*, 2223–2242.
[16] L. Cassar, P. E. Eaton, J. Halpern, *J. Am. Chem. Soc.* **1970**, *92*, 6336–6368.
[17] D. Kaufmann, H. H. Fick, O. Schallner, W. Spielmann, L. U. Meyer, P. Gölitz, A. De Meijere, *Chem. Ber.* **1983**, *116*, 587–609.
[18] H. Prinzbach, D. Stusche, *Helv. Chim. Acta* **1971**, *54*, 755–759.
[19] C. Thilgen, F. Diederich, *Top. Curr. Chem.* **1999**, *199*, 135–171.
[20] P. W. Fowler, *Chem. Phys. Lett.* **1986**, *131*, 444–450.
[21] M. Goldberg, *Tôhoku Math. J.* **1937**, *43*, 104–108.
[22] B. Grünbaum, T. S. Motzkin, *Can. J. Math.* **1953**, *15*, 744–751.

Received: February 12, 2001 [F3064]