Manual Construction and Mathematics- and Computer-Aided Counting of Stereoisomers. The Example of Oligoinositols

Christoph Rücker,* Ralf Gugisch, Adalbert Kerber
Department of Mathematics, Universität Bayreuth,
D-95440 Bayreuth, Germany

Received May 12, 2004

Two methods to obtain numbers of stereoisomers and of achiral stereoisomers of a given molecular structure are detailed on the example of di- and triinositols. The first method is manual exhaustive construction free of redundance of all stereoisomers, which is rendered feasible by symmetry considerations despite the large number of isomeric triinositols (82176). The second method is counting without constructing, made possible by use of a mathematical tool, the Cauchy-Frobenius lemma, which actually is a formalized manner of considering symmetry. The results are compared to those obtained by computer-aided stereoisomer generation using the program MOLGEN 3.5. It is demonstrated that in their results all three methods agree.

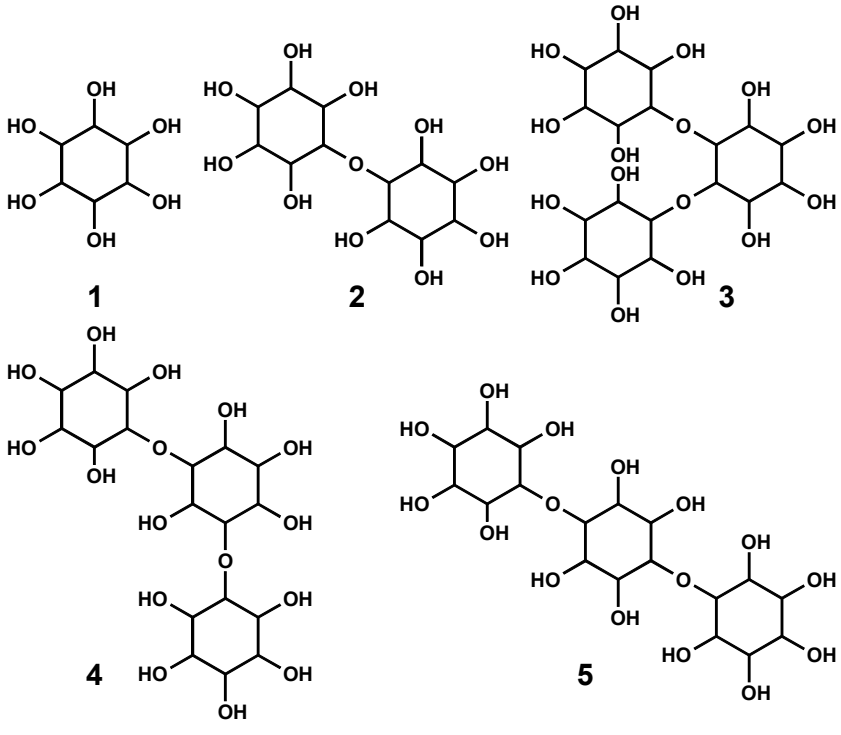
INTRODUCTION

Software for solving chemical problems should be tested on cases of intermediate complexity: For simple problems that are easily treated by hand a computer program is not needed, for very complex cases a solution may be produced by a program, but it cannot manually be checked for correctness. In cases of intermediate complexity only, a nontrivial computer result can be compared to a result obtained in a necessarily lengthy and error-prone manual procedure, thus demonstrating both the necessity of having a computer program and, in the best case, the correctness of a particular program for at least the particular case at hand.

A few years ago, T. Hudlicky *et al.* raised the question how many stereoisomers exist of O-linked diinositols and triinositols of structures **2** and **3**.[1] They gave an answer for **2** (990 stereoisomers), but did not explain how this number was arrived at. Later Hönig *et al.* derived 528 as the number of stereoisomers of **2** using their program ISOMERS.[2] The same authors, when publishing a detailed compilation of

numbers of inositol oligomers, still left unanswered the question for the number of stereoisomers of **3**.[3] They gave, however, the number of all stereoisomeric linear triinositols, i.e. the sum of stereoisomer counts for **3** and its constitutional isomers **4** and **5**. They also asked for the number of achiral stereoisomers among all stereoisomeric di- and triinositols, and treated this problem by inspection of many thousands of stereoisomers.

The chemist's manual approach is often limited by the size of the problem at hand and by human deficiencies. In our opinion, mathematical methods and already existing computer programs for treating such stereochemical problems can be very helpful and deserve to be more recognized by chemists than is actually the case. Therefore in the present paper we compare methods to solve the exemplary problem of the oligo-inositol stereoisomers. First we construct (in principle) all stereoisomers of **2**, **3**, **4**, and **5** free of redundance manually using symmetry considerations and a simple classification scheme, then we use a mathematical theorem known as the Cauchy-Frobenius lemma for counting the stereoisomers. In both cases we obtain not only the number of all stereoisomers, but also the numbers of achiral stereoisomers and thus of pairs of enantiomers. Finally we compare the results to those obtained by using the computer program MOLGEN 3.5 developed in Bayreuth,[4] which is able to exhaustively construct stereoisomers as far as stereocenters and *E*/*Z* double bonds only are involved.



Stereoisomerism in **1** - **5** may be considered to be due solely to tetrahedral stereocenters (chiral centers), the number of stereoisomers exclusively depends on the possibilities to distribute C-O single bonds above and below the average planes of the cyclohexane rings. Due to rapid interconversion of conformers, we are allowed to approximate each cyclohexane ring as a planar hexagon (Haworth projection). The numbers of stereoisomers in the present cases are not trivial since **1** - **5** exhibit high constitutional symmetry. Therefore many of the configurational patterns will turn out to be identical, so that the numbers of truly distinct patterns presumably are well below 2^n, where n is the number of stereocenters.

RESULTS AND DISCUSSION

1. Manual construction of all stereoisomers.

1.1 Stereoisomeric inositols 1

The inositol (**1**) stereoisomers are easily constructed manually, by systematically placing none, one, two and three substituent OH groups at one side of a regular hexagon, with all others on the opposite side. These isomers have been known for a long time:[5] There are exactly 9 stereoisomers called *scyllo*-, *myo*-, (+)-*chiro*-, (-)-*chiro*-, *neo*-, *epi*-, *cis*-, *muco*-, and *allo*-inositol, all shown in Figure 1 together with their point group symbols, in the order of appearence in reference 5.[6]

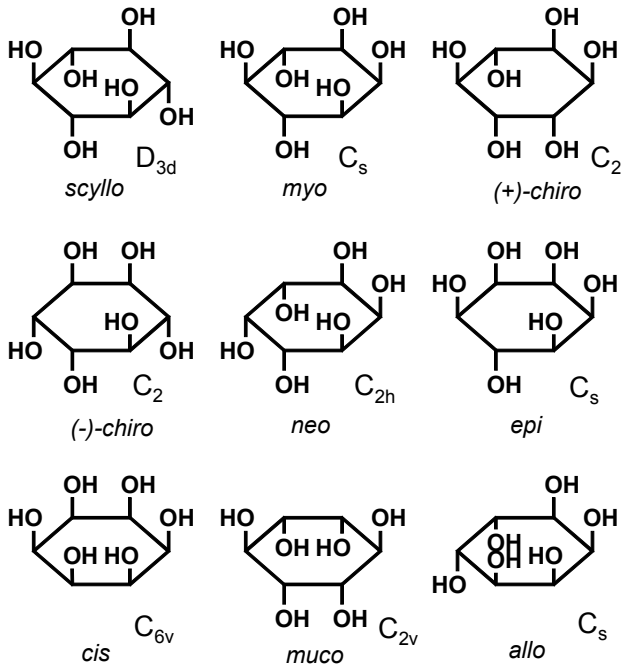


Figure 1. The nine inositols.

1.2. Stereoisomeric diinositols 2

We dissect molecule **2** into two halves, each of which is a monosubstituted inositol. We construct all possible such building blocks and finally combine these in all possible ways.

By replacing in turn each *distinct* OH group in an inositol (as determined by its symmetry) with an OX group, from each inositol between one and six isomeric monosubstituted inositols are obtained. Thus, in *scyllo*-inositol all six OH groups are homotopic[7] and therefore substitutions lead to a single monosubstituted product. In contrast, in *myo*-inositol no two OH groups are homotopic (they are either enantiotopic or diastereotopic[7]) so that six distinct monosubstituted products are formed. Enantiomers are distinguished and counted individually by this procedure.

In Figure 2 all 32 general monosubstituted inositols are shown. The OH symbols are left out for clarity, OH groups are represented by vertical lines.

In this paper we assume X to represent an achiral substituent or a free valence. Then, of the 32 monosubstituted inositols, 8 are achiral (*meso* compounds: *scyllo1*, *myo1*, *myo6*, *neo1*, *epi5*, *epi6*, *cis1*, *muco1*) and 24 are chiral, forming twelve pairs of enantiomers (*myo2*/*myo3*, *myo4*/*myo5*, (+)-*chiro1*/(-)-*chiro1*, (+)-*chiro2*/(-)-*chiro2*, (+)-*chiro3*/(-)-*chiro3*, *neo2*/*neo3*, *epi1*/*epi2*, *epi3*/*epi4*, *muco2*/*muco3*, *allo1*/*allo2*, *allo3*/*allo4*, *allo5*/*allo6*). Enantiomers are linked by a bracket in Figure 2. For isomers of nontrivial symmetry the point group symbols are also given. The numbering *myo1*, *myo2*, …, is ours and arbitrary.

Forming dimers from these 32 monomers, we obtain 32 homodimers and 32 ·31/2 = 496 heterodimers. Thus there is a total of 528 stereoisomers of formula **2**.

Dolhaine and Hönig further raised the question how many of the diinositols are achiral. In their first paper[2] the answer given was 48 achiral isomers, while in the later paper they found only 46 achiral isomers.[3]

Our reasoning with respect to this problem is simple: Diinositols made up of two achiral halves (either identical or different) obviously are

achiral. Of the former kind there are 8 (*scyllo1–scyllo1*, …), while there are 8·7/2 = 28 of the latter kind (*scyllo1–myo1*, …).

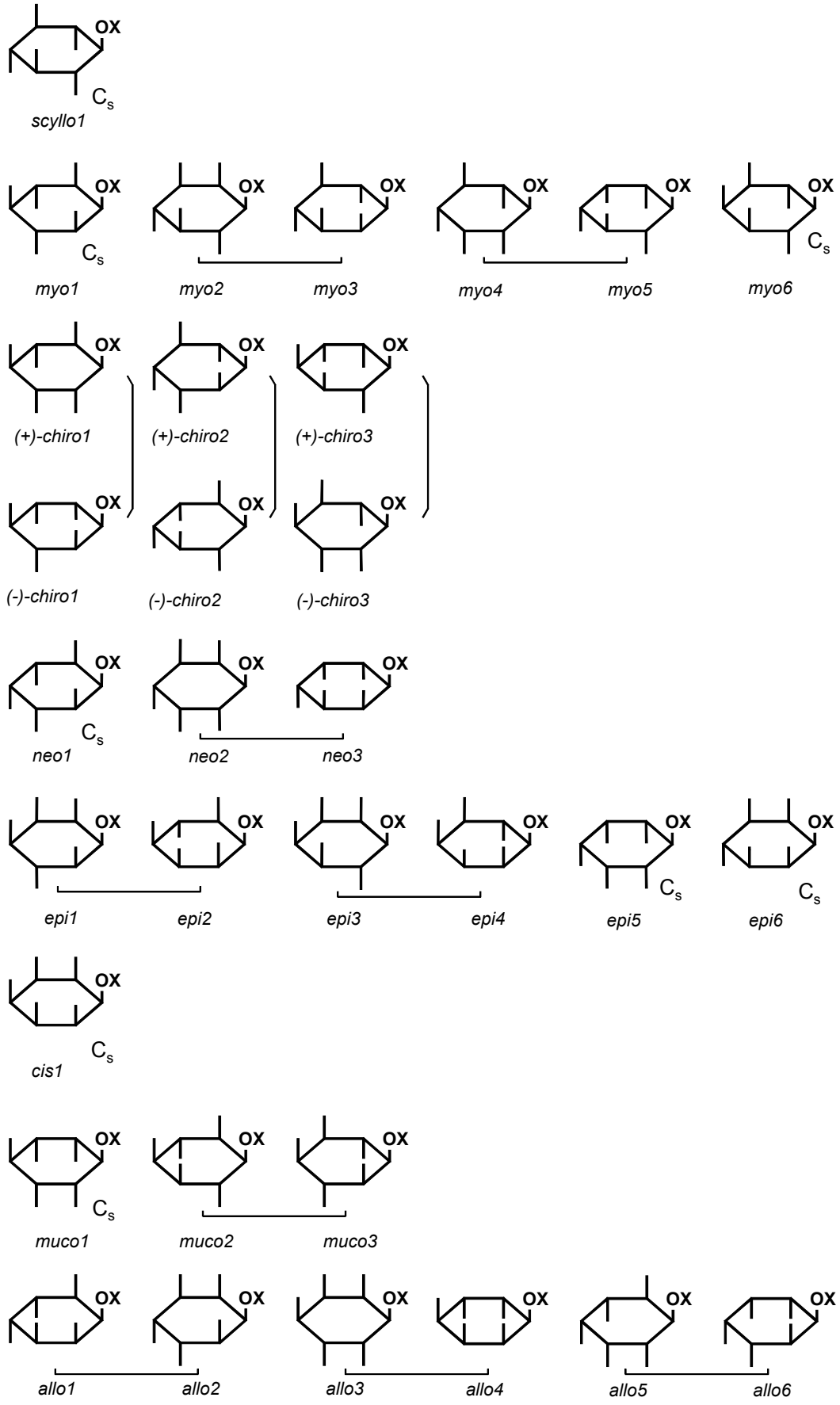


Figure 2. All general monosubstituted inositols.

Further, a diinositol made of two chiral halves is achiral if these are enantiomeric. There are 12 diinositols of this kind (*myo2-myo3*, …). So there exist 48 achiral diinositols altogether. In fact in Table II of ref. 3 two achiral diinositols were not identified as such (*epi5-epi6* and *myo1-myo6* in our numbering, presumably).

The derivations given above are essentially the same as those given by the earlier authors.[2] They are detailed here because in the following sections we require the list of the 32 monosubstituted inositols.

1.3. Stereoisomeric triinositols 3 - 5

1.3.1 Stereoisomers of **3** *(1,2-disubstituted central unit)*

The procedure is similar to that for **2**, though necessarily a bit more involved. In addition to all monosubstituted inositols (Figure 2) we now require all general 1,2-disubstituted inositols as the central building blocks in **3**. The isomers shown in Figure 3 were constructed by replacing for each *distinct* CC bond in an inositol (as determined by its symmetry) the two OH groups attached to this bond with two OX groups. In this manner from each inositol between one and six products are obtained. The labeling *myoa*, *myob*, …, is ours and arbitrary.

Altogether there are 36 distinct general 1,2-disubstituted inositols. While it is possible to write down, for each of the 36, all products of substituting each OX independently by one of the 32 monosubstituted inositols, and then to check for pairwise identities, this would obviously be a tedious procedure. Instead we partition the 36 into classes according to the topicity of the OX groups (which is determined by the symmetry of the molecule) and then treat one representative of each class:

1) The substituents OX are enantiotopic in a mirror symmetric (C_s) and therefore achiral molecule. In this class are *cisa*, *mucoc*, *alloa*, *allof*.

2) The substituents OX are diastereotopic in an unsymmetric (C_1) and therefore chiral molecule. Here we have *myoa*/*myob*, (+)-*chirob*/(−)-*chirob*, *neoa*/*neob*, *epia*/*epib*, *epie*/*epif*, *allob*/*alloc*, *myoc*/*myod*, *myoe*/*myof*, (+)-*chiroc*/(−)-*chiroc*, *epic*/*epid*, *mucoa*/*mucob*, *allod*/*alloe*.

3) The substituents OX are homotopic in a C_2 symmetric and therefore chiral molecule. In this category are *scylloa*/*scyllob*, (+)-*chiroa*/(-)-*chiroa*, (+)-*chirod*/(-)-*chirod*, *neoc*/*neod*.

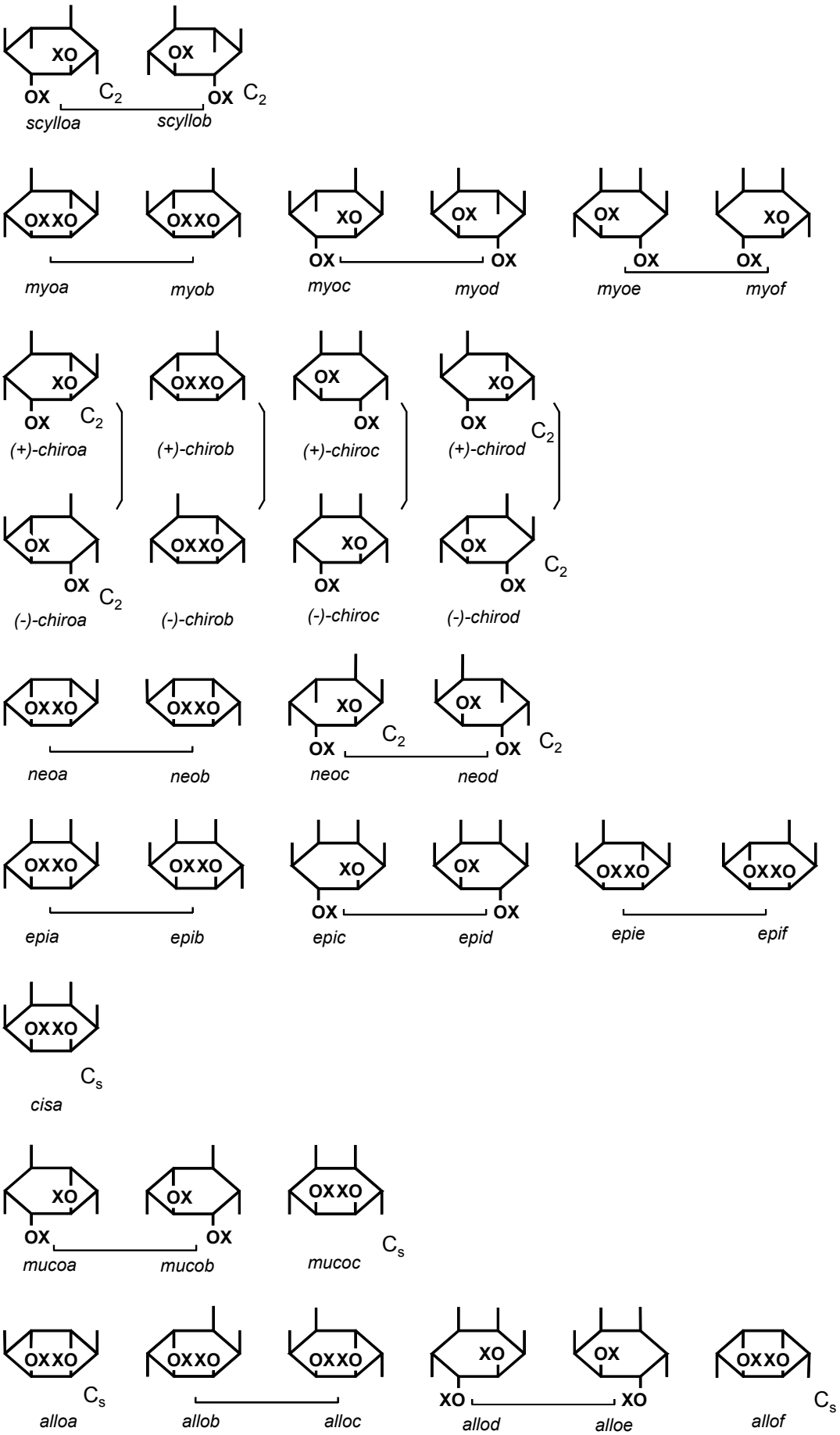

XO
OX
C2
scylloa
scyllob
OXXO
myoa
myob
myoc
myod
myoe
myof
(+)-chiroa
(+)-chirob
(+)-chiroc
(+)-chirod
(-)-chiroa
(-)-chirob
(-)-chiroc
(-)-chirod
neoa
neob
neoc
neod
epia
epib
epic
epid
epie
epif
Cs
cisa
mucoa
mucob
mucoc
alloa
allob
alloc
allod
alloe
allof

Figure 3. All general 1,2-disubstituted inositols.

We choose as representatives *mucoc*, (+)-*chirob*/(-)-*chirob*, and *neoc*/*neod*. In Figure 4, the isomers are listed that result from replacing, in each representative, the two substituents OX by two like substituents,[8] either achiral (symbolized by achiral letter A), or chiral (symbolized by chiral letter R, and –R for its enantiomorph[9]). We are interested, of course, in substituents A, R representing monosubstituted inositols.

(Figure 4)

Figure 4. All inositols 1,2-disubstituted by two like achiral (A) or chiral (R) substituents. For the three classes see text.

In Figure 5 all isomers are listed that are obtained by replacing, in each representative, the two substituents OX by two unlike substituents,[8] either both achiral (achiral symbols A and M), or one achiral and one chiral, or both chiral (chiral symbols F, G, and –F and –G[9]).

(Figure 5)

Figure 5. All inositols 1,2-disubstituted by two unlike achiral (A, M) or chiral (F, G) substituents. For the three classes see text.

Finally, in Table 1, parts A and B, the results from Figures 4 and 5 are gathered, respectively. Summation of all isomer counts results in a total number of 32896 stereoisomers for **3**.

(Table 1)

1.3.2 Stereoisomers of **4** *(1,3-disubstituted central unit)*

The procedure is analogous to that for **3**. In Figure 6 all general 1,3-disubstituted inositols are listed. They were obtained by substituting OX for OH in all *distinct* 1,3-relations of two OH groups in all inositols, as determined by symmetry. We partition these 32 isomers into two classes according to the molecule's symmetry and the topicity of the OX groups. 1) The substituents OX are enantiotopic in a mirror symmetric (C_s) and therefore achiral molecule. In this category are *scylloc*, *myog*, *myol*, *neoe*, *epik*, *epil*, *cisb*, *mucod*.

2) The substituents OX are diastereotopic in an unsymmetric (C_1) and therefore chiral molecule. These are *myoj*/*myok*, (+)-*chirog*/(-)-*chirog*, *epig*/*epih*, *allog*/*alloh*, *myoh*/*myoi*, (+)-*chiroe*/(-)-*chiroe*, (+)-*chirof*/(-)-*chirof*, *neof*/*neog*, *epii*/*epij*, *mucoe*/*mucof*, *alloi*/*alloj*, *allok*/*allol*.

Note that there are no 1,3-disubstituted inositols with the OX groups related by a C_2 axis. Such isomers cannot exist since the central C atom bears one OH group which points either up or down.

(Figure 6)

Figure 6. All general 1,3-disubstituted inositols.

Figures 7 and 8 show, for a representative from either class, *scylloc* and *myoj/myok*, all isomers bearing two like and two unlike substituents, respectively. Table 2 summarizes the results of Figures 7 and 8. There are 32768 stereoisomers of structure **4**.[10]

(Figure 7)

Figure 7. All inositols 1,3-disubstituted by two like achiral (A) or chiral (R) substituents. For the two classes see text.

(Figure 8)

Figure 8. All inositols 1,3-disubstituted by two unlike achiral (A, M) or chiral (F, G) substituents. For the two classes see text.

(Table 2)

1.3.3 Stereoisomers of **5** *(1,4-disubstituted central unit)*

In Figure 9 all general 1,4-disubstituted inositols are listed. They were obtained by substituting OX for OH in all *distinct* 1,4-relations of two OH groups in all inositols, as determined by symmetry. We partition these 20 isomers into 4 classes according to the molecule's symmetry and the topicity of the OX groups.

1) The substituents OX are enantiotopic or diastereotopic in a mirror symmetric (C_s) or centrosymmetric (C_i) achiral molecule: *myom*, *alloo*, *neoi*, *epio*.

2) The substituents OX are homotopic in a C_2 and mirror symmetric (C_{2v} or C_{2h}) achiral molecule: *cisc*, *mucog*, *scyllod*, *neoh*.

3) The substituents OX are homotopic in a C_2 symmetric chiral molecule: *mucoh/mucoi*, (+)-*chiroi*/(-)-*chiroi*.

4) The substituents OX are diastereotopic in an unsymmetric (C_1) chiral molecule: (+)-*chiroh*/(-)-*chiroh*, *epim/epin*, *myon/myoo*, *allom/allon.*

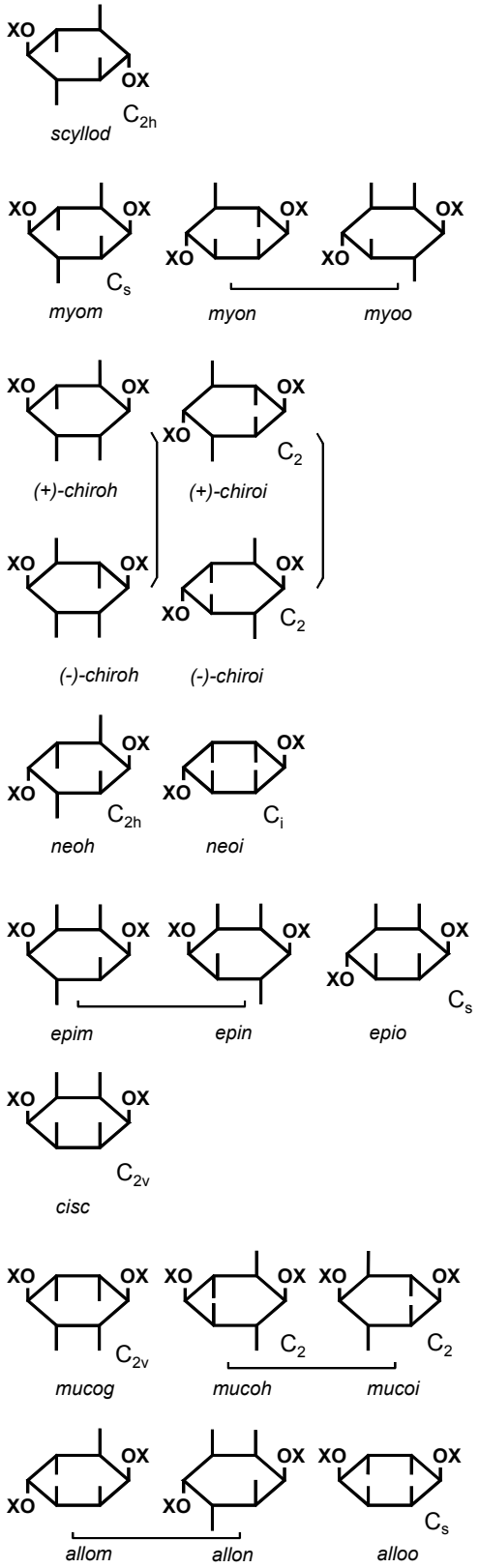


Figure 9. All general 1,4-disubstituted inositols.

Figures 10 and 11 show, for a representative from each class (*alloo*, *mucog*, *mucoh*/*mucoi*, *epim*/*epin*), all isomers bearing two like and two unlike substituents, respectively. Table 3 summarizes the results of Figures 10 and 11. There are 16512 stereoisomers of structure **5**.

(Figure 10)

Figure 10. All inositols 1,4-disubstituted by two like achiral (A) or chiral (R) substituents. For the four classes see text.

(Figure 11)

Figure 11. All inositols 1,4-disubstituted by two unlike achiral (A, M) or chiral (F, G) substituents. For the four classes see text.

(Table 3)

The sum of stereoisomer counts for **3**, **4**, and **5**, 82176, equals the number given by Dolhaine and Hönig for all stereoisomeric tri-inositols.[3]

1.3.4 Achiral stereoisomers of triinositols **3**, **4**, **5**

With respect to the question of how many triinositols are achiral, Dolhaine and Hönig remarked "The achiral isomers again have to be evaluated individually by carefully examining the symmetrical isomers as well as those with **d**-**l** – pairs." These authors in a presumably lengthy manual procedure obtained 630 achiral triinositols.[3]

We derived the number of achiral stereoisomers of **3** - **5** by simple reasoning. For an achiral triinositol it is necessary that the central unit (considered in isolation, a disubstituted inositol) exhibits an element of improper rotation symmetry, S_n, in particular a mirror plane or a center of inversion. So we extract all such disubstituted central units from Figures 3, 6, and 9. We obtain 20 species, of which 4 are 1,2-disubstituted, of C_s symmetry, bearing two enantiotopic OX groups (*cisa*, *mucoc*, *alloa*, *allof*), 8 are 1,3-disubstituted, of C_s symmetry, bearing enantiotopic OX groups (*scylloc*, *myog*, *myol*, *neoe*, *epik*, *epil*, *cisb*, *mucod*), and 8 are 1,4-disubstituted, of various symmetries, bearing pairs of OX groups of various topicity (C_s, enantiotopic OX groups: *alloo*; C_i, enantiotopic OX groups: *neoi*; C_s, diastereotopic OX groups: *myom*, *epio*;

C_{2v}, homotopic OX groups: *cisc*, *mucog*; C_{2h}, homotopic OX groups: *scyllod*, *neoh*).

Of these each species of C_s symmetry bearing enantiotopic OX groups (e.g. *mucoc* or *scylloc*) is transformed into an achiral triinositol by replacing the OX groups with either two identical achiral monoinositols (8 possibilities, **a1** or **a4**, Figures 4 and 7), or with two chiral enantiomeric monoinositols (2·12 = 24 possibilities, **a2**, **a3**, **a5**, **a6**), resulting in 32 achiral triinositols. So there are 4·32 = 128 achiral stereoisomers of **3**, and 8·32 = 256 achiral stereoisomers of **4**.

For an 1,4-disubstituted central unit of C_s or C_i symmetry with enantiotopic OX groups (*alloo* or *neoi*), exactly the same reasoning leads to 32 achiral triinositols (**a7**, **a8**, **a9**, Figure 10), resulting in a total of 64 achiral such stereoisomers.

From a C_s symmetric 1,4-disubstituted species with diastereotopic OX groups (e.g. *myom*) an achiral triinositol is obtained if both OX groups are replaced with (identical or different) achiral monosubstituted inositols (8·8 = 64 possibilities, **a10**, **a11**, **a12**). In this case, chiral monosubstituted inositols as terminal units are incompatible with the required C_s symmetry and therefore cannot lead to further achiral triinositols. So there is a total of 128 achiral triinositols with a *myom* or *epio* central unit.

(Formulae **a10**–**a12**)

From a C_{2h} or C_{2v} species with homotopic OX groups (e.g. *mucog*) an achiral triinositol is obtained by replacing the OX groups with two achiral (identical or different) monosubstituted inositols (8 and 8·7/2 = 28 possibilities, **a13** and **a15**, Figures 10 and 11), or by using two chiral enantiomeric monosubstituted inositols (12 possibilities, **a14**). So we have 4·48 = 192 achiral triinositols with a *scyllod*, *neoh*, *cisc*, or *mucog* central unit. Altogether these are 384 achiral stereoisomers of **5**.

From the numbers of achiral **3**, **4**, and **5** stereoisomers a total of 768 achiral triinositols is obtained. The remarkable discrepancy between this number and the number 630 found by the earlier authors[3] is easily explained by several erroneous entries in the "achirals" columns in Table III in reference 3.

In summary, among 9 monoinositols (**1**) 7 are achiral (78%), among 528 diinositols (**2**) 48 are achiral (9.1%), and among 82176 triinositols (**3**–**5**) 768 are achiral (0.93%). This is in accord with the general trend that the larger and complexer a system is, the more rarely is symmetry observed.

2. Results obtained using the Cauchy-Frobenius lemma

The same numerical results as above should be obtained using a mathematical tool known as the Cauchy-Frobenius lemma (or Burnside's lemma).[11] This theorem was used for solving a large variety of combinatorial problems, including Pólya's approach to questions such as how many distinct possibilities exist to attach certain numbers of substituents of certain types to a given molecular core.[12] Our stereochemical problem is similar to such a so-called coloration problem: We could ask for the number of distinct possibilities to attach exactly one substituent (out of several possible types, say methyl and hydrogen) to each carbon atom of a benzene ring, or equivalently, to color the atoms of the aromatic ring black and blue (for carbon and nitrogen). Or we can ask for the number of distinct possibilities to attach a substituent in a specific manner (out of several such manners, say up or down) to each atom of a six-membered carbocycle. The difference between these two problems will become clear below.

Before considering the Cauchy-Frobenius lemma itself we have to understand an important concept used therein, the concept of automorphism. An automorphism is a permutation of atoms such that the neigborhood relations are maintained, i.e. a symmetry operation. For example, in a planar regular hexagon permutations g_1, g_2 and g_3 (Figure 12) are automorphisms, while permutation p_4 is not an automorphism.

Here g_1 can be interpreted as a 60° rotation about the sixfold symmetry axis of the hexagon, while g_2 may be interpreted as a 180° rotation about the same axis, or as an inversion at the hexagon's center. Permutation g_3 is a rotation about an axis passing through the midpoints of lines 1-2 and 4-5, or a reflection at a plane perpendicular to the hexagon plane and dissecting lines 1-2 and 4-5. An automorphism as such is independent of its geometric interpretation. In more complex

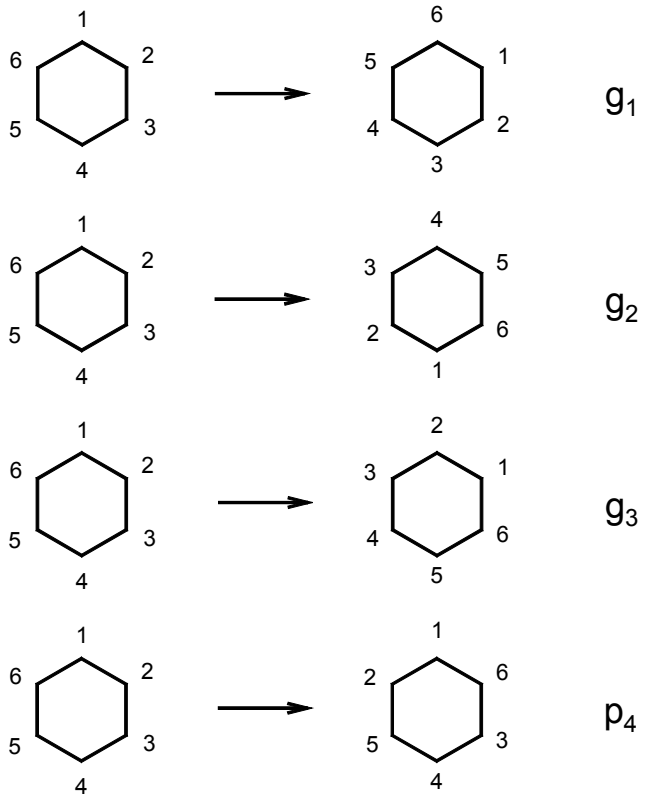


Figure 12. Some permutations of vertices in the planar regular hexagon.

structures, there may exist automorphisms that correspond to more complex symmetry operations not interpretable as simple rigid-body rotations or reflections.

Each permutation can be represented by a cycle notation, where a cycle is a pair of parentheses containing the labels of permuted atoms in the order in which they are moved. Thus,

g_1: (123456), 1 → 2, 2 → 3, … , 6 → 1;

g_2: (14)(25)(36), 1 → 4, 4 → 1, 2 → 5, 5 → 2, 3 → 6, 6 → 3;

g_3: (12)(36)(45), 1 → 2, 2 → 1, 3 → 6, 6 → 3, 4 → 5, 5 → 4;

p_4: (1)(26)(3)(4)(5), 2 → 6, 6 → 2, all others unchanged.

The symbol c(g) is used for the number of cycles in the cycle representation of g. Thus, $c(g_1) = 1$, $c(g_2) = 3$, $c(g_3) = 3$, $c(p_4) = 5$.

The null permutation ("do nothing"), called the identity, id, (1)(2)(3)(4)(5)(6), is also an automorphism, c(id) = 6.

The Cauchy-Frobenius lemma reads (we here take it for granted):

$$N = \frac{1}{|G|} \sum_{g \in G} |X_g|$$

where N is the number we are looking for, G is the set of all automorphisms of the core structure (its automorphism group), and $|X_g|$ is the number of fixed patterns of the automorphism $g \in G$. A fixed pattern of automorphism g is a substitution or coloring pattern that is not changed by g. Now, what is the number $|X_g|$ of such patterns? Each cycle in the representation of g contains those atoms that are interchanged by g. For a substitution or coloring pattern not to be changed by g, all atoms within a cycle have to bear the same kind of substituent, or have to be "of the same color". This applies to each cycle: Each cycle has to be "monochromatic". So the number of fixed patterns of g is the number of possibilities to color c(g) cycles with m colors (in this paper always m = 2):

$$|X_g| = m^{c(g)} .$$

We insert this into the above formula and obtain

$$N = \frac{1}{|G|} \sum_{g \in G} m^{c(g)}$$

We are now able to tackle the first of the above problems, the question for the number of methyl-substituted benzenes. The automorphisms of a regular hexagon are the following, easily found manually as rigid-body rotations:

id, (123456), (135)(246), (14)(25)(36), (153)(264), (165432),
(12)(36)(45), (14)(23)(56), (16)(25)(34),
(1)(26)(35)(4), (13)(2)(46)(5), (15)(24)(3)(6).

The five automorphisms listed after the identity are rotations about the sixfold axis, next in the list are the 180° rotations about the three axes in the hexagon plane dissecting opposite lines (also interpretable as reflections), finally followed by 180° rotations about the three axes in the hexagon plane passing through opposite vertices (also interpretable as reflections).

Now using the information on the automorphisms we obtain the number of methyl-substituted benzenes,

$$N = (1/12)\cdot(1\cdot 2^6 + 2\cdot 2^1 + 2\cdot 2^2 + 1\cdot 2^3 + 3\cdot 2^3 + 3\cdot 2^4) = 13 \ .$$

In fact there are thirteen benzenes bearing H or methyl substituents, as shown in Figure 13, upper row.

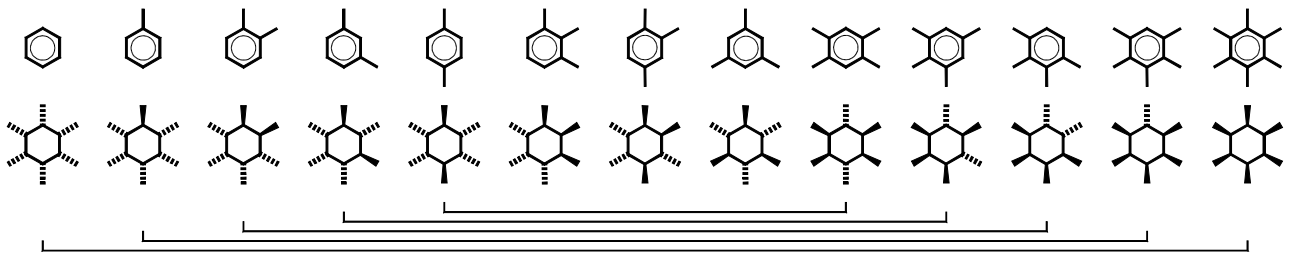


Figure 13. Parallelism and difference between the vertex coloring for a planar hexagon (upper row) and the stereochemical problem in inositols (lower row).

2.1 Stereoisomeric inositols 1

Let us now treat the stereochemical problem of attaching up and down substituents to a cyclohexane. Replacing each methyl (appearing in the former problem) quite formally with an up substituent and each non-methyl with a down substituent, we obtain the lower row in Figure 13. However, this treatment obviously is not appropriate for this problem, since now five species were constructed twice, as indicated by brackets in Figure 13, where the picture on the right is obtained from the picture on the left by turning it upside-down. Thus a decisive difference between the two kinds of problems is that a substituent's up or down property is affected by turning the molecule around, while an atom's black or blue property is not. Another problem is that the seventh species in the lower row is chiral, so that both it and its enantiomer should be counted.

As above, we interpret an automorphism g as a rotation or a coupling of rotations of the molecule or parts of it. If an atom (or equivalently its substituent) is turned upside-down by g, then we assign a minus sign to that atom's label within the cycle representation of g. Conversely, a minus sign in a cycle says that an atom's substituent turns from up to down or vice versa. If an automorphism has a fixed pattern at all, for each substituent turning from up to down there certainly is an equivalent substituent turning from down to up. In other words, counting fixed patterns by the Cauchy-Frobenius lemma, we have to consider those automorphisms only that have an even number of minus signs in each cycle:

$$N=\frac{1}{|G|}\sum_{g\in G^{-}} m^{c(g)}$$

where the summation is over the set G^- of automorphisms having an even number of minus signs in each cycle (zero is, of course, an even number).

Thus, the automorphisms of the regular hexagon are to be listed as follows

id, (123456), (135)(246), (14)(25)(36), (153)(264), (165432),
(-1-2)(-3-6)(-4-5), (-1-4)(-2-3)(-5-6), (-1-6)(-2-5)(-3-4),
(-1)(-2-6)(-3-5)(-4), (-1-3)(-2)(-4-6)(-5), (-1-5)(-2-4)(-3)(-6).

The last three of these do not meet the criterion and thus are not included in the summation for stereoisomers. So for inositol **1** we obtain the number of stereoisomers

$$N = (1/12)\cdot(1\cdot 2^6 + 2\cdot 2^1 + 2\cdot 2^2 + 1\cdot 2^3 + 3\cdot 2^3) = 9 ,$$

which as we know is the correct result.

Since we are interested not only in the total number of stereoisomers (N), but also in the numbers of achiral stereoisomers (N_a) and of pairs of enantiomers (N_c), in addition to

$$N = N_a + 2\cdot N_c$$

we need a second equation. We will obtain an expression for

$$N' = N_a + N_c ,$$

so that $N_a = 2\cdot N' - N$.

We obtain N′ by identifying the individual enantiomers in a pair of enantiomers. To this end we construct a set of automorphisms G′ that is very similar to G. The elements of G′ are similar to the automorphisms in G, with the only difference that in G′ each automorphism of G is coupled to (immediately followed by) a reflection at a plane. Thereby all up substituents are turned down and vice versa, and correspondingly all plus and minus signs in the cycle representations are interchanged. Now we work with the union of G and G′, $G \cup G'$.[13] This union contains twice as many automorphisms as G, and in the summation we now include both those automorphisms in G that have an even number of minus signs in each cycle, and those in G′ that have an even number of minus signs in each cycle. The latter exactly correspond to those in G that have an even number of plus signs in each cycle. Thus the Cauchy-Frobenius lemma now reads

$$N' = \frac{1}{2|G|} \sum_{g \in (G \cup G')^-} m^{c(g)} .$$

In practice, we only have to test each automorphism $g \in G$ for an even number of minus signs in each cycle, and for an even number of plus signs in each cycle. For either criterion fulfilled, g is included once.

In the case of the regular hexagon, the 2nd, 4th, 6th, as well as the last six automorphisms have an even number of plus signs in each cycle, thus

$$N' = (1/24)\cdot(1\cdot 2^6 + 2\cdot 2^1 + 2\cdot 2^2 + 1\cdot 2^3 + 3\cdot 2^3 + 2\cdot 2^1 + 1\cdot 2^3 + 3\cdot 2^3 + 3\cdot 2^4) = 8 .$$

Finally, the number of achiral stereoisomers is, as we expected,

$$N_a = 2\cdot 8 - 9 = 7 .$$

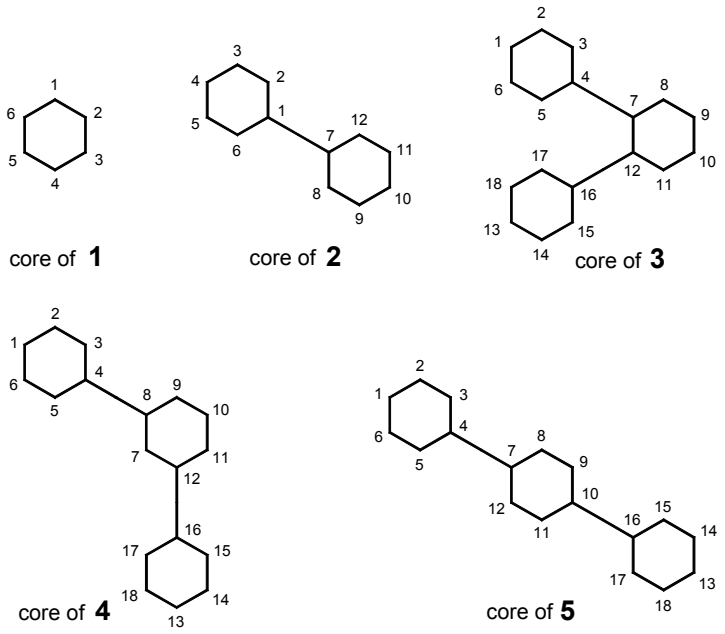


Figure 14. Atom numbering used in the text for the cores of mono-, di-, and triinositols.

2.2 Stereoisomeric diinositols 2

Automorphisms that are rotations of the core of diinositol **2** or of parts of it (torsions) are found manually. The complete list is (see Figure 14 for atom labeling)

g_1: id, ←

g_2: (1 7)(2 8)(3 9)(4 10)(5 11)(6 12), ← ⇐

g_3: (-1)(-2 -6)(-3 -5)(-4)(-10)(-9 -11)(-8 -12)(-7), ⇐

g_4: (-1 -7)(-2 -12)(-3 -11)(-4 -10)(-5 -9)(-6 -8), ← ⇐

g_5: (-1)(-2 -6)(-3 -5)(-4)(7)(8)(9)(10)(11)(12),

g_6: (1)(2)(3)(4)(5)(6)(-7)(-8 -12)(-9 -11)(-10),

g_7: (-1 7)(-2 12 -6 8)(-3 11 -5 9)(-4 10),

g_8: (1 -7)(2 -8 6 -12)(3 -9 5 -11)(4 -10).

Here g_2 through g_4 are 180° solid-body rotations, g_5 and g_6 are 180° torsions of one or the other half of the molecule, and both g_7 and g_8 are a 180° torsion coupled to a 180° rotation.

The automorphisms marked with a simple arrow fulfill the criterion for minus signs, so we obtain for the stereoisomers of **2**

$$N = (1/8)\cdot(2^{12} + 2^6 + 2^6) = 528.$$

The automorphisms marked with a double arrow fulfill the criterion for plus signs, so we obtain

$$N' = (1/16)\cdot(2^{12} + 2^6 + 2^6 + 2^6 + 2^8 + 2^6) = 288 .$$

The number of achiral stereoisomers is

$$N_a = 2\cdot 288 - 528 = 48 .$$

2.3 Stereoisomeric triinositols 3 - 5

The automorphisms of the core of **3** are (see Figure 14 for atom labeling)

g_1: id, ←

g_2: (-1 -13)(-2 -14)(-3 -15)(-4 -16)(-5 -17)(-6 -18)(-7 -12)(-8 -11)(-9 -10), ← ⇐

g_3: (-1)(-2 -6)(-3 -5)(-4)(7)(8)(9)(10)(11)(12)(13)(14)(15)(16)(17)(18),

g_4: (1)(2)(3)(4)(5)(6)(7)(8)(9)(10)(11)(12)(-13)(-14 -18)(-15 -17)(-16),

g_5: (-1)(-2 -6)(-3 -5)(-4)(7)(8)(9)(10)(11)(12)(-13)(-14 -18)(-15 -17)(-16)

g_6: (1 -13)(2 -18 6 -14)(3 -17 5 -15)(4 -16)(-7 -12)(-8 -11)(-9 -10),

g_7: (-1 13)(-2 14 -6 18)(-3 15 -5 17)(-4 16)(-7 -12)(-8 -11)(-9 -10),

g_8: (1 13)(2 18)(3 17)(4 16)(5 15)(6 14)(-7 -12)(-8 -11)(-9 -10). ← ⇐

Here g_2 is a solid-body 180° rotation, g_3 and g_4 are torsions of one or the other terminal hexagon, g_5 is a torsion of both terminal hexagons, g_6 and g_7 are torsions of one terminal hexagon coupled to a solid-body 180° rotation, and g_8 is a torsion of both terminal hexagons coupled to a 180° rotation. For the number of all stereoisomers of **3** the automorphisms marked with a simple arrow are included in the summation,

$$N = (1/8)\cdot(2^{18} + 2^9 + 2^9) = 32896 .$$

For N' the automorphisms marked with a double arrow are included additionally

$$N' = (1/16)\cdot(2^{18} + 2^9 + 2^9 + 2^9 + 2^9) = 16512 ,$$

and $N_a = 2\cdot 16512 - 32896 = 128$.

The automorphisms of the core of **4** are

g_1: id, ←

g_2: (-1 −13)(-2 −14)(-3 −15)(-4 −16)(-5 −17)(-6 −18)(-7)(-8 −12)(-9 −11)(-10), ⇐

g_3: (-1)(-2 −6)(-3 −5)(-4)(7)(8)(9)(10)(11)(12)(13)(14)(15)(16)(17)(18),

g_4: (1)(2)(3)(4)(5)(6)(7)(8)(9)(10)(11)(12)(-13)(-14 −18)(-15 −17)(-16),

g_5: (-1)(-2 −6)(-3 −5)(-4)(7)(8)(9)(10)(11)(12)(-13)(-14 −18)(-15 −17)(-16),

g_6: (1 -13)(2 -18 6 -14)(3 -17 5 -15)(4 -16)(-7)(-8 −12)(-9 −11)(-10),

g_7: (-1 13)(-2 14 -6 18)(-3 15 -5 17)(-4 16)(-7)(-8 −12)(-9 −11)(-10),

g_8: (1 13)(2 18)(3 17)(4 16)(5 15)(6 14)(-7)(-8 −12)(-9 −11)(-10). ⇐

These automorphisms are found and interpreted exactly as in the case of **3**. Here the identity only has an even number of minus signs in each cycle, whereas there are two automorphisms with an even number of plus signs in each cycle, so that the numbers of stereoisomers of **4** are

$$N = (1/8)\cdot 2^{18} = 32768 ,$$

$$N' = (1/16)\cdot(2^{18} + 2^{10} + 2^{10}) = 16512 ,$$

$$N_a = 2\cdot 16512 - 32768 = 256 .$$

The automorphisms of the core of **5** are

g_1: id, ←

g_2: (-1)(-2 −6)(-3 −5)(-4)(7)(8)(9)(10)(11)(12)(13)(14)(15)(16)(17)(18),

g_3: (1)(2)(3)(4)(5)(6)(-7)(-8 −12)(-9 −11)(-10)(13)(14)(15)(16)(17)(18),

g_4: (1)(2)(3)(4)(5)(6)(7)(8)(9)(10)(11)(12)(-13)(-14 −18)(-15 −17)(-16),

g_5: (-1)(-2 −6)(-3 −5)(-4)(7)(8)(9)(10)(11)(12)(-13)(-14 −18)(-15 −17)(-16),

g_6: (-1)(-2 −6)(-3 −5)(-4)(-7)(-8 −12)(-9 −11)(-10)(13)(14)(15)(16)(17)(18),

g_7: (1)(2)(3)(4)(5)(6)(-7)(-8 −12)(-9 −11)(-10)(-13)(-14 −18)(-15 −17)(-16),

g_8: (-1)(-2 −6)(-3 −5)(-4)(-7)(-8 −12)(-9 −11)(-10)(-13)(-14 −18)(-15 −17)(-16), ⇐

g_9: (-1 −13)(-2 −14)(-3 −15)(-4 −16)(-5 −17)(-6 −18)(-7 −10)(-8 −9)(-11 −12), ← ⇐

g_{10}: (1 -13)(2 -18 6 -14)(3 -17 5 -15)(4 -16)(-7 −10)(-8 −9)(-11 −12),

g_{11}: (-1 13)(-2 14 -6 18)(-3 15 -5 17)(-4 16)(-7 −10)(-8 −9)(-11 −12),

g_{12}: (1 13)(2 18)(3 17)(4 16)(5 15)(6 14)(-7 −10)(-8 −9)(-11 −12), ← ⇐

g_{13}: (-1 −13)(-2 −14)(-3 −15)(-4 −16)(-5 −17)(-6 −18)(7 10)(8 11)(9 12), ← ⇐

g_{14}: (1 -13)(2 -18 6 -14)(3 -17 5 -15)(4 -16)(7 10)(8 11)(9 12),
g_{15}: (-1 13)(-2 14 -6 18)(-3 15 -5 17)(-4 16)(7 10)(8 11)(9 12),
g_{16}: (1 13)(2 18)(3 17)(4 16)(5 15)(6 14)(7 10)(8 11)(9 12). ← ⇐

Here g_8, g_9, g_{16} are 180° solid-body rotations about the three principal axes, g_2 through g_7 are 180° torsions of various parts of the molecule, g_{10} through g_{15} are torsions coupled to a rotation. For **5** the numbers of stereoisomers are

$$N = (1/16)\cdot(2^{18} + 4\cdot2^{9}) = 16512 ,$$
$$N' = (1/32)\cdot(2^{18} + 4\cdot2^{9} + 4\cdot2^{9} + 2^{12}) = 8448 ,$$
$$N_a = 2\cdot8448 - 16512 = 384 .$$

3. Results obtained using MOLGEN 3.5

In contrast to using the Cauchy-Frobenius lemma, treatment of stereochemistry in MOLGEN 3.5 crucially depends on the notion of stereocenter, well-known to every chemist. MOLGEN primarily is a program for constructing all constitutional isomers of a given molecular formula.[4] So it works on the topological, not on the geometrical level. However, stereocenters and stereogenic double bonds in a given constitutional isomer can be identified on the topological level.[14] Therefore T. Wieland implemented in MOLGEN 3.5 a module that constructs and displays all stereoisomers for a given structure, as far as no stereogenic features beyond stereocenters and double bonds are involved. The method was described in detail elsewhere,[4b,15] here it may suffice to state that it is based on group-theoretical methods that take into account both the distribution of ligands around a stereocenter and the overall symmetry of the molecule (the configuration symmetry group[12,14]).

For **1**, its mono-, 1,2-, 1,3-, and 1,4-disubstituted derivatives, and for **2** – **5** MOLGEN 3.5 produces the same stereoisomer counts as were obtained above manually or using the Cauchy-Frobenius lemma. On the other hand, a differentiation between chiral and achiral stereoisomers is not implemented in this version of MOLGEN.

REFERENCES AND NOTES

(1) Hudlicky, T.; Abboud, K.A.; Entwistle, D.A.; Fan, R.; Maurya, R.; Thorpe, A.J.; Bolonick, J.; Myers, B. Toluene-Dioxygenase-Mediated *cis*-Dihydroxylation of Aromatics in Enantioselective Synthesis. Iterative Glycoconjugate Coupling Strategy and Combinatorial Design for the Synthesis of Oligomers of nor-Saccharides, Inositols and Pseudosugars with Interesting Molecular Properties. *Synthesis* **1996**, 897-911.

(2) Dolhaine, H.; Hönig, H.; van Almsick, M. Sample Applications of an Algorithm for the Calculation of the Number of Isomers with more than one Type of Achiral Substituent. *MATCH – Commun. Math. Comput. Chem*. **1999**, *39*, 21-37.

(3) Dolhaine, H.; Hönig, H. Full Isomer-Tables of Inositol-Oligomers up to Tetramers. *MATCH – Commun. Math. Comput. Chem*. **2002**, *46*, 91-119.

(4)(a) Grund, R.; Kerber, A.; Laue, R. MOLGEN, ein Computeralgebra-System für die Konstruktion molekularer Graphen. *MATCH – Commun. Math. Comput. Chem.* **1992**, *27*, 87-131. (b) Wieland, T.; Kerber, A.; Laue, R. Principles of the Generation of Constitutional and Configurational Isomers. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 413-419. (c) Benecke, C.; Grüner, T.; Kerber, A.; Laue, R.; Wieland, T. MOLecular Structure GENeration with MOLGEN, New Features and Future Developments. *Fresenius' J. Analyt. Chem.* **1997**, *359*, 23-32. (d) Kerber, A.; Laue, R.; Grüner, T.; Meringer, M. MOLGEN 4.0. *MATCH – Commun. Math. Comput. Chem.* **1998**, *37*, 205-208.

(5) Eliel, E. L. *Stereochemistry of Organic Compounds*, Wiley: New York, 1994.

(6) Signs (+) and (-) in the names of the *chiro*-inositols refer to the sense of optical rotation. In the present work, however, we use these descriptors to identify the absolute configuration of a *chiro*-inositol derivative: The (+) sign denotes the absolute configuration of (+)-*chiro*-inositol.

(7) Two features in a molecule (atoms, bonds, substituents, a pair of nonbonded atoms, etc.) are homotopic ("in the same place") if they are related by a proper rotation symmetry, the action of a C_n axis of symmetry. Any manipulation at one or the other of two such substituents results in identical products, i.e. one product only.

Two features in a molecule are enantiotopic ("in mirror-related places") if they are related by an improper rotation symmetry only, the action of an S_n axis, in particular a mirror plane ($\sigma = S_1$) or a center of inversion ($i = S_2$). Any manipulation at one or the other of two such

substituents results in two distinct products, two enantiomers in the simplest case.

Two features in a molecule are diastereotopic if they are not related by any symmetry. Any manipulation at one or the other results in two diastereomers.

For more information see reference 5 or the following: Rücker, C.; Braun, J. *UNIMOLIS – A Computer-aided Course on Molecular Symmetry and Isomerism*, http://unimolis.uni-bayreuth.de (*C. A.* 140:235199(2004)).

(8) Two like substituents here are either identical or enantiomeric, unlike substituents are diastereomeric or not stereoisomeric.

(9) Since the enantiomorphs of letters F, G, R are not available in common word processors nor in common molecule drawing programs, we here use instead the letters preceeded by a minus sign.

(10) As seen in Figures 7, 8 and in Table 2, a class 2 entry results in exactly twice as many isomers as a class 1 entry. This factor of 2 is fully explained by the difference in chirality: A pair of enantiomers (class 2) results in two products where a single species (class 1) gives one product. In other words, there is no difference in the number of products obtained from two enantiotopic or from two diastereotopic OX groups. This is also seen in Figures 4, 5, and Table 1 (and Figures 10, 11, and Table 3). This is an immediate consequence of the definitions of enantiotopic and diastereotopic molecular features, see Note 7. If we are interested in the number of products only, we are therefore allowed to combine cases of enantiotopic and of diastereotopic OX groups in a single class, as it was done in class 1 for **5**. In fact the classifications of disubstituted inositols made above are somewhat arbitrary, except for the distinction between homotopic OX groups on the one hand and enantiotopic/ diastereotopic OX groups on the other, which is compulsory.

(11) Kerber, A., *Applied Finite Group Actions*. 2nd ed., Springer: Berlin, 1999, Chapter 2.

(12)(a) Pólya, G. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Math.* **1937**, *68*, 145-254. (b) Pólya, G.; Read, R. C. *Combinatorial Enumeration of of Groups, Graphs, and Chemical Compounds*, Springer: Berlin, 1987.

(13) Mathematically speaking, we consider $S_2 \times G = G \cup G'$, where the elements g ∈ G are identified with the (1,g) ∈ S_2xG, and the elements g’∈ G’ are identified with the ((12),g) ∈ S_2xG.

(14)(a) Nourse, J. G. The Configuration Symmetry Group and Its Application to Stereoisomer Generation, Specification, and Enumeration. *J. Am. Chem. Soc.* **1979**, *101*, 1210-1216. (b) Nourse, J. G.; Carhart, R. E.; Smith, D. H.; Djerassi, C. Exhaustive Generation of Stereoisomers for Structure Elucidation. *J. Am. Chem. Soc.* **1979**, *101*, 1216-1223.
(15)(a) Wieland, T. Erzeugung, Abzählung und Konstruktion von Stereoisomeren. *MATCH – Commun. Math. Comput. Chem.* **1994**, *31*, 153-203. (b) Wieland, T. Enumeration, Generation, and Construction of Stereoisomers of High-Valence Stereocenters. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 220-225.

Footnote for page 1

*Corresponding author phone: +49 921 553386; fax: +049 921 553385; e-mail: christoph.ruecker@uni-bayeuth.de

Further Figures

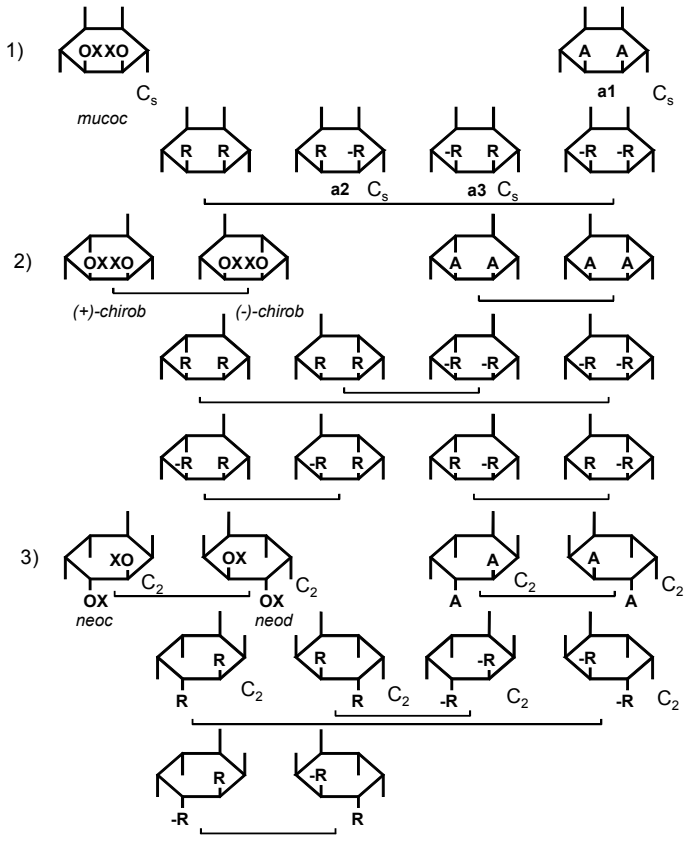


Figure 4. All inositols 1,2-disubstituted by two like achiral (A) or chiral (R) substituents. For the three classes see text.

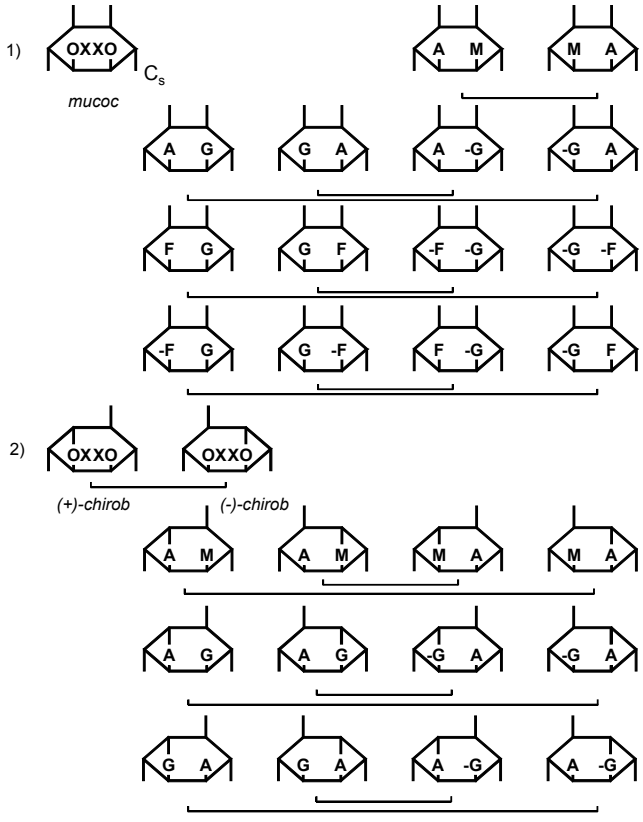


Figure 5(Part 1). All inositols 1,2-disubstituted by two unlike achiral (A, M) or chiral (F, G) substituents. For the three classes see text.

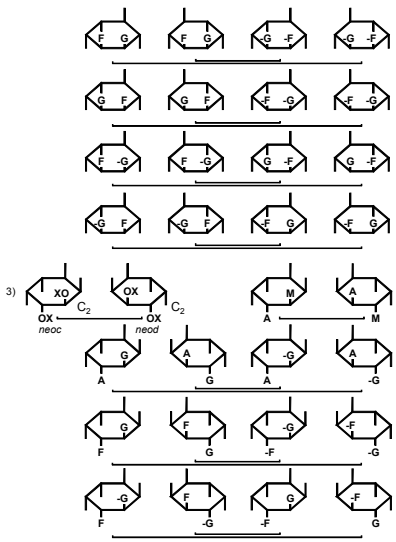


Figure 5(Part 2). All inositols 1,2-disubstituted by two unlike achiral (A, M) or chiral (F, G) substituents. For the three classes see text.

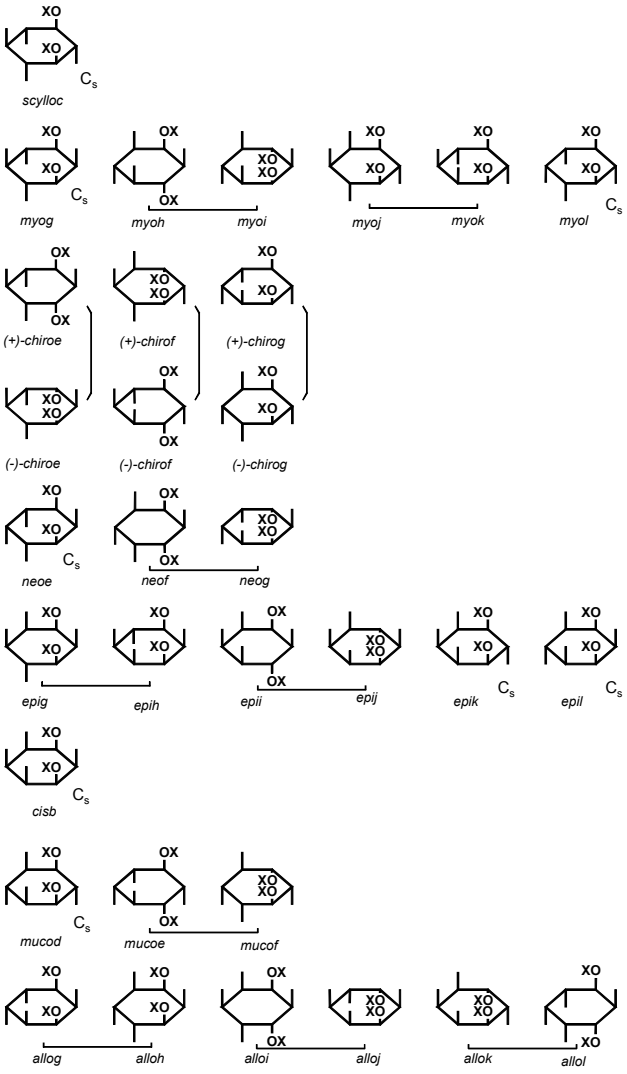

XO
XO
Cs
scylloc
myog
myoh
myoi
myoj
myok
myol
OX
(+)-chiroe
(+)-chirof
(+)-chirog
(-)-chiroe
(-)-chirof
(-)-chirog
neoe
neof
neog
epig
epih
epii
epij
epik
epil
cisb
mucod
mucoe
mucof
allog
alloh
alloi
alloj
allok
allol

Figure 6. All general 1,3-disubstituted inositols.

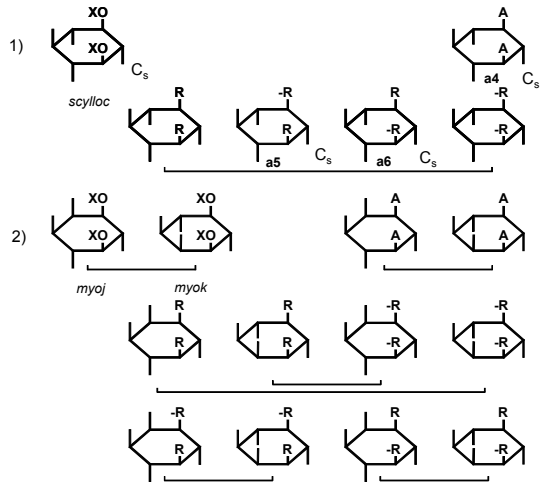


Figure 7. All inositols 1,3-disubstituted by two like achiral (A) or chiral (R) substituents. For the two classes see text.

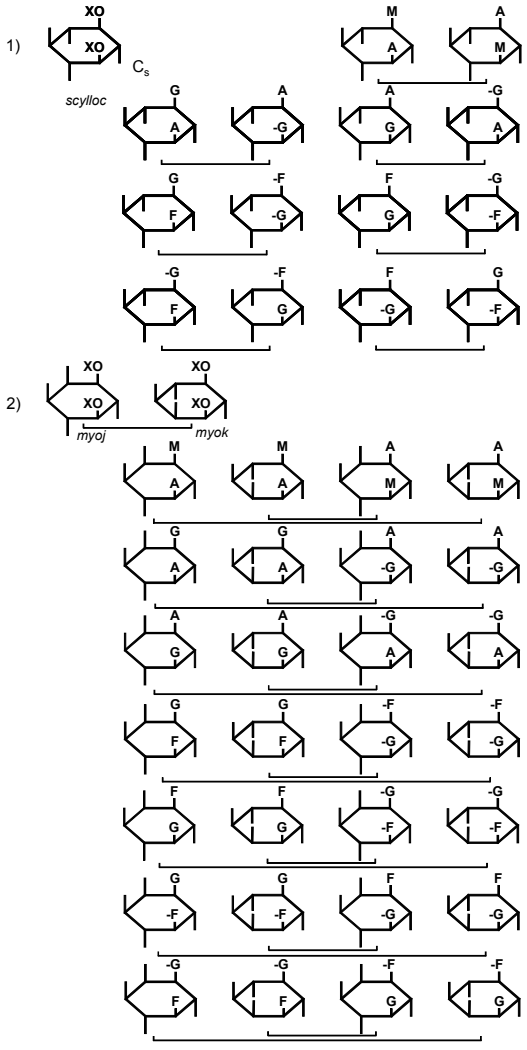

1)
XO
XO
Cs
scylloc
G
A
-G
F
-F
M
2)
myoj
myok

Figure 8. All inositols 1,3-disubstituted by two unlike achiral (A, M) or chiral (F, G) substituents. For the two classes see text.

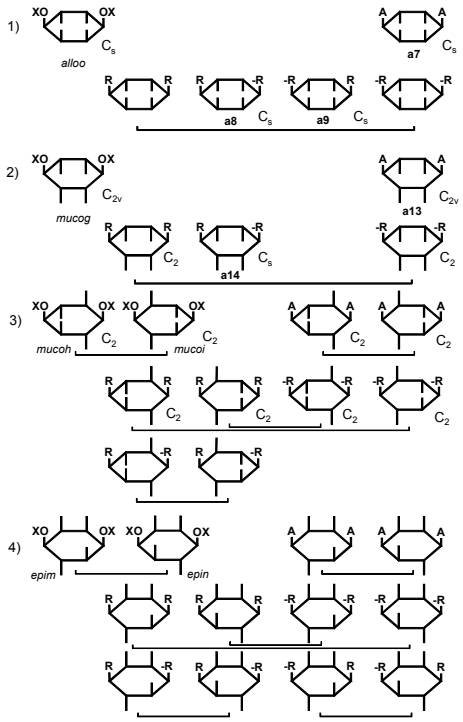


Figure 10. All inositols 1,4-disubstituted by two like achiral (A) or chiral (R) substituents. For the four classes see text.

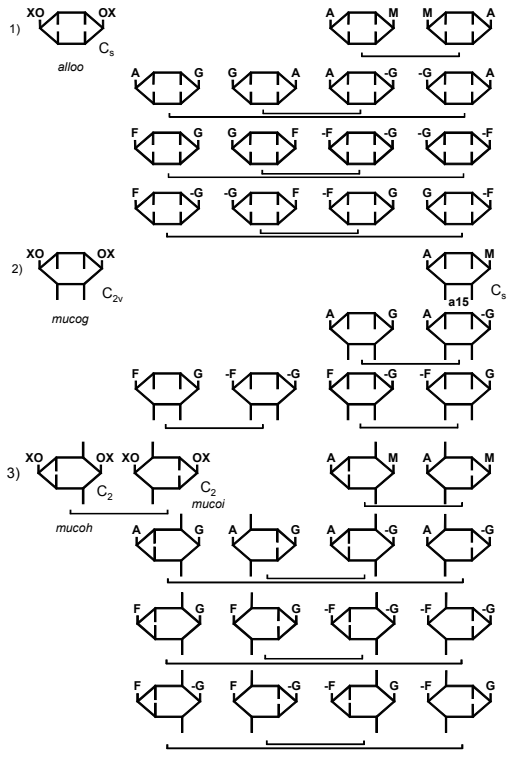


Figure 11(Part1). All inositols 1,4-disubstituted by two unlike achiral (A, M) or chiral (F, G) substituents. For the four classes see text.

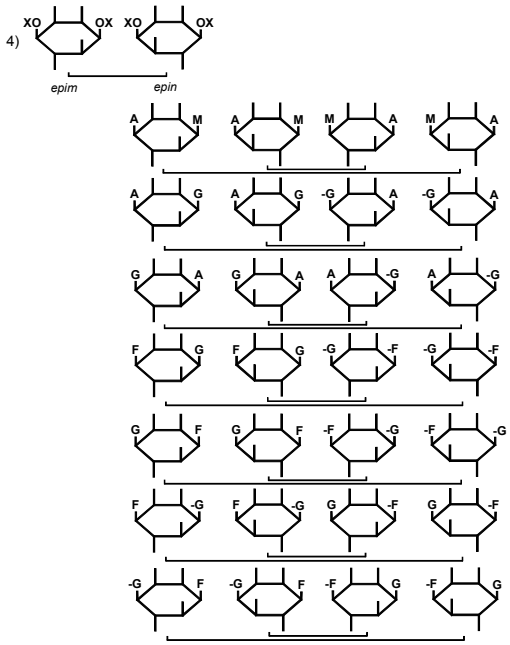
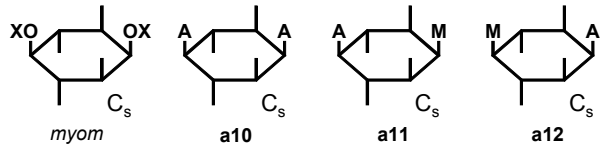


Figure 11(Part2). All inositols 1,4-disubstituted by two unlike achiral (A, M) or chiral (F, G) substituents. For the four classes see text.



myom **a10** **a11** **a12**

Formulae a10 – a12

Table 1. Stereoisomer numbers for tri-inositols with 1,2-disubstituted central moiety (**3**).

Central Moiety	terminal moieties	#Isomers per combination[a]	#Combinations[b]	total #Isomers
A. Two like terminal moieties				
Class 1	achiral	1	4 x 8	32
	chiral	4	4 x 12	192
Class 2	achiral	2	12 x 8	192
	chiral	8	12 x 12	1152
Class 3	achiral	2	4 x 8	64
	chiral	6	4 x 12	288
B. Two unlike terminal moieties				
Class 1	both achiral	2	4 x 28	224
	one chiral	4	4 x 96	1536
	both chiral	8	4 x 66	2112
Class 2	both achiral	4	12 x 28	1344
	one chiral	8	12 x 96	9216
	both chiral	16	12 x 66	12672
Class 3	both achiral	2	4 x 28	224
	one chiral	4	4 x 96	1536
	both chiral	8	4 x 66	2112
			total	32896

[a]from Figures 4, 5.

[b]the first figure is the number of individual 1,2-disubstituted inositols or pairs of enantiomers in a class; the second figure is the number of distinct achiral monosubstituted inositols or distinct pairs of enantiomeric monosubstituted inositols (Part A), or the number of combinations of two different achiral monosubstituted inositols, or the number of combinations of an achiral and a chiral monosubstituted inositol, or the number of combinations of two different chiral monosubstituted inositols (Part B).

Table 2. Stereoisomer numbers for tri-inositols with 1,3-disubstituted central moiety (**4**).

Central Moiety	terminal moieties	#Isomers per combination[a]	#Combinations[b]	total #Isomers
A. Two like terminal moieties				
Class 1	achiral	1	8 x 8	64
	chiral	4	8 x 12	384
Class 2	achiral	2	12 x 8	192
	chiral	8	12 x 12	1152
B. Two unlike terminal moieties				
Class 1	both achiral	2	8 x 28	448
	one chiral	4	8 x 96	3072
	both chiral	8	8 x 66	4224
Class 2	both achiral	4	12 x 28	1344
	one chiral	8	12 x 96	9216
	both chiral	16	12 x 66	12672
			total	32768

[a]from Figures 7, 8.

[b]the first figure is the number of individual 1,3-disubstituted inositols or pairs of enantiomers in a class; the second figure is the number of distinct achiral monosubstituted inositols or distinct pairs of enantiomeric monosubstituted inositols (Part A), or the number of combinations of two different achiral monosubstituted inositols, or the number of combinations of an achiral and a chiral monosubstituted inositol, or the number of combinations of two different chiral monosubstituted inositols (Part B).

Table 3. Stereoisomer numbers for tri-inositols with 1,4-disubstituted central moiety (**5**).

Central Moiety	terminal moieties	#Isomers per combination[a]	#Combinations[b]	total #Isomers
A. Two like terminal moieties				
Class 1	achiral	1	4 x 8	32
	chiral	4	4 x 12	192
Class 2	achiral	1	4 x 8	32
	chiral	3	4 x 12	144
Class 3	achiral	2	2 x 8	32
	chiral	6	2 x 12	144
Class 4	achiral	2	4 x 8	64
	chiral	8	4 x 12	384
B. Two unlike terminal moieties				
Class 1	both achiral	2	4 x 28	224
	one chiral	4	4 x 96	1536
	both chiral	8	4 x 66	2112
Class 2	both achiral	1	4 x 28	112
	one chiral	2	4 x 96	768
	both chiral	4	4 x 66	1056
Class 3	both achiral	2	2 x 28	112
	one chiral	4	2 x 96	768
	both chiral	8	2 x 66	1056
Class 4	both achiral	4	4 x 28	448
	one chiral	8	4 x 96	3072
	both chiral	16	4 x 66	4224
			total	16512

[a]from Figures 10, 11.

[b]the first figure is the number of individual 1,4-disubstituted inositols or pairs of enantiomers in a class; the second figure is the number of distinct achiral monosubstituted inositols or distinct pairs of enantiomeric monosubstituted inositols (Part A), or the number of combinations of two different achiral monosubstituted inositols, or the number of combinations of an achiral and a chiral monosubstituted

inositol, or the number of combinations of two different chiral monosubstituted inositols (Part B).