

Proposed Extension to the Natta Projection Notation System for Enabling an Indication of Relative Stereochemistry and the Stereochemical State

Karel D. Klika

Department of Chemistry, University of Turku, Turku, Finland

E-mail: klikakd@yahoo.co.uk

Received October 1, 2011; revised November 6, 2011; accepted November 18, 2011

Abstract

A system of structure depiction, as an extension of the wedge and hashed wedge bonds (Natta projection), and text notation is herein suggested that embodies more explicit information—or reduced over-statement as circumstances warrant—on the stereochemical nature of the system at hand, in particular, for those cases where only the relative stereochemistry of a compound is known.

Keywords: Stereochemistry, Stereochemical Descriptors, Natta Projection, Racemate, Enantiomer

The Natta projection system¹ for indicating stereochemistry is familiar to all organic chemists, but surprisingly, there are limitations in the depictions. For example, in **Figure 1**, the structure presented for 1,2,4-trihydroxycyclohexane is very explicit: it represents the (1*R*,2*S*,4*R*)-enantiomer. If this is the intended information to be conveyed, then it is not only appropriate, it is also clearly unambiguous. At the opposite end of the scale, the structure shown in **Figure 2** is equally profound in its meaning: nothing at all is known regarding the stereochemistry of 1,2,4-trihydroxycyclohexane and the state could be one or all or any mixture whatsoever of the eight possible stereoisomers.

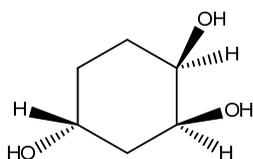


Figure 1. The structural depiction of (1*R*,2*S*,4*R*)-1,2,4-trihydroxycyclohexane.

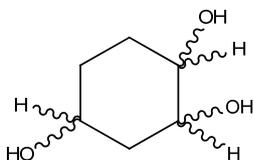


Figure 2. The structural depiction of 1,2,4-trihydroxycyclohexane without any indication of stereochemistry.

¹Giulio Natta, 1903-1979.

Problems arise however, when a known mixture of stereoisomers is present or the relative stereochemistry is known but not the absolute configuration. How does one then represent the stereochemical state of the system? For example, for a sample of known relative configuration but unknown absolute configuration (irrespective of whether one or two stereoisomers were present, presumably in the latter case a pair of enantiomers constituting a racemic sample), then clearly both **Figures 1** and **2** do not suffice to accurately convey this information. For **Figure 1**, because it can be both lacking² (if two enantiomers were present) and overstating (since the absolute configuration is not known); for **Figure 2**, because it does not convey all the information that is known. It thus seems curious that there is not a widely used convention in place to conveniently address what is a rather common occurrence. To alleviate this limitation and to enable one to ascertain at a relative glance the stereochemical state in more systems without the need to scour the text for corroborating or qualifying information, the following extension to the Natta system is therefore suggested.

For systems where something is known about the stereochemistry but the information is incomplete or there is a need amalgamate data, the solid wedge and hashed wedge bonds used to represent forward and backward

²Even if mention of this is made alongside the structural figure, in the figure caption, or in the main body of the text one still has to locate and retrieve it and, regrettably, all too often mention of this important aspect may even be neglected altogether.

oriented bonds, respectively, are replaced analogously by a hollow wedge bond and a hashed wedge bond with a line through it, respectively³. For example, if the relative stereochemistry of 1,2,4-trihydroxycyclohexane is known but the absolute configuration is unknown, then representing the structure as depicted in **Figure 3** conveys this notion succinctly.

Moreover, if the number of stereoisomers present in the sample is known, this can be indicated by an integer number in italics near to the questionable stereochemical bond associated with the lowest carbon atom number. If the number of stereoisomers is unknown, this can be indicated by the abbreviation “*unk*”; if the sample is racemic, then the abbreviation “*rac*” can be applied (due the presence of three stereogenic centers in the particular example at hand, a “2” does not unambiguously equate to a racemate); and if the sample is holemic (*i.e.* enantiopure), this can be indicated by the abbreviation “*hol*” (needless to say, a “1” would necessarily be equivalent). For the example chosen, “*cis*” and “*trans*” can also indicate relative stereochemistry, but do not infer anything regarding the number of stereoisomers, and the notation is limited in the main to only two or three stereogenic centers. Furthermore, the enantiomeric content⁴ (*ec*) of the sample can be indicated by a real number *in lieu* of an integer if the sample is scalemic (or if the use of enantiomeric excess (*ee*) is preferred, this can be alternatively indicated by a percentage to distinguish it from the *ec*). For cases other than scalemates where only two stereoisomers are unambiguously present, the real number will indicate the proportion of the two stereoisomers. If the

major stereoisomer is known, this can be indicated by use of the new bond indicators (a hollow wedge bond and a hashed wedge bond with a line through it).

The number system can also be applied to wavy bonds to eliminate ambiguity. In **Figure 4**, the replacement of the stereochemically explicit bonds for the groups at position 1 for the example at hand may alternatively be used to indicate either a pair of C-1 epimers or just one stereoisomer of unknown stereochemistry at C-1. The ambiguity is simply resolved by the incorporation of a number. Other examples of the notation are also given in **Figure 4**.

This notational system is readily useable in exactly the same manner within the general nomenclature, so in addition to the conventional “(±)-” or “(*rac*)-” preceding a name to indicate a racemic mixture, so too can “(1)-”, “(2)-”, “(0.8)-”, or “(80%)-” *etc.* be used to indicate the presence of a single stereoisomer, two stereoisomers, an 80% *ec*, or an 80% *ee* mixture, respectively. If the major isomer is unknown, this can appropriately be indicated

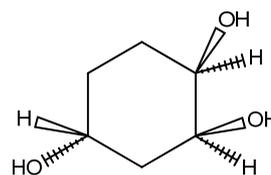


Figure 3. Depiction of the extension of the Natta projection system for a sample of 1,2,4-trihydroxycyclohexane of known relative stereochemistry but unknown absolute configuration.

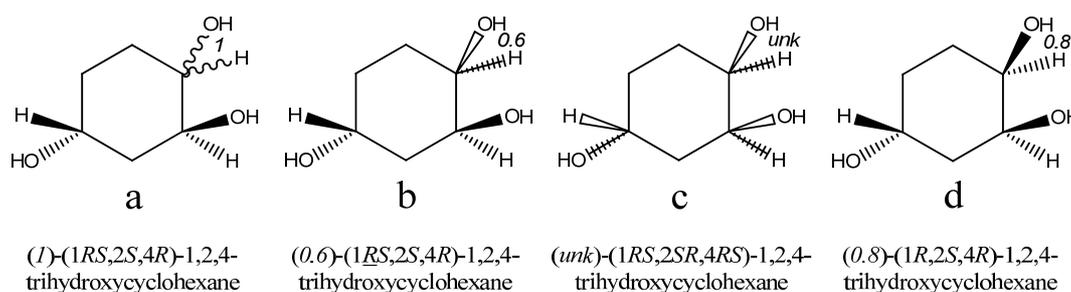


Figure 4. For 1,2,4-trihydroxycyclohexane, the structure depicted in (a) explicitly indicates the presence of a single stereoisomer of unknown configuration at C-1 (alternatively, a “2” would indicate the presence of a pair of C-1 epimers), whilst the structure depicted in (b) explicitly indicates the presence of a pair of C-1 epimers consisting of the main stereoisomer, (1*R*,2*S*,4*R*)-1,2,4-trihydroxycyclohexane, in 60% content and the minor stereoisomer, (1*S*,2*R*,4*S*)-1,2,4-trihydroxycyclohexane, in 40% content. For the structure depicted in (c), the relative stereochemistry is known but the number of stereoisomers is unknown (*i.e.* the sample could be holemic, racemic, or scalemic). For the structure depicted in (d), the sample is scalemic and is composed of 80% (1*R*,2*S*,4*R*)-1,2,4-trihydroxycyclohexane and 20% (1*S*,2*R*,4*S*)-1,2,4-trihydroxycyclohexane. The names beneath the structures reflect the stereochemical states depicted and are a direct extension of the notation to the general nomenclature. Additional explanation on the nomenclature notation is given in the ensuing paragraph.

³Selection was based on avoiding confusion or contradiction with established notation for other purposes, or similar purposes especially, and with notations already available in drawing programs. Similarity, though with clear distinction, to the Natta projection bonds was also sought. Thus bold, hashed, and dashed bonds were not chosen whilst

color options were not considered since color is not yet ubiquitous and half-tone was considered too indistinct and dependent on print quality.

⁴The recommendation of Gawley [1] is preferred with regards to the use of enantiomeric composition (*ec* as the percentage of one enantiomer) over enantiomeric excess (*ee*).

by asterisks appended to the *R* and *S* descriptors; if known, then this can be indicated by the use of under-scores. Additional descriptors can include “(hol)-” for holemic, “(scl)-” for scalemic (where the *ec* is between 50% - 100%), and “(unk)-” for an unknown number of stereoisomers. Of note, it has always been ambiguous when “*R**” *etc.* (or *RS etc.*) has been used as to whether one or more stereoisomers is present based on the name alone and the use of numbers readily eliminates this limitation. For the examples given in **Figure 4**, the corresponding names are given within the figure.

In summary, an extension of the conventional Natta projection system to indicate stereochemistry has been formulated as an aid to better describing the stereochemical state of a system. Adoption of the notation may reduce the level of confusion or misunderstanding in the mind of a reader since there is an ever increasing need to rapidly peruse the output of database searches or voluminous compilations of papers with efficacy. In light of

the importance of stereochemistry in biological systems and the stereochemical dependence of drug activity [2], this is of paramount significance.

Acknowledgements

Financial support of this work by Turun Yliopistosäätiö is gratefully acknowledged.

References

- [1] R. E. Gawley, “Do the Terms ‘% ee’ and ‘% de’ Make Sense as Expressions of Stereoisomer Composition or Stereoselectivity?” *The Journal of Organic Chemistry*, Vol. 71, No. 6, 2006, pp. 2411-2416. [doi:10.1021/jo052554w](https://doi.org/10.1021/jo052554w)
- [2] J. Caldwell, “The Importance of Stereochemistry in Drug Action and Disposition,” *The Journal of Clinical Pharmacology*, Vol. 32, No. 10, 1992, pp. 925-929.