The nomenclature of fused-ring arenes and heterocycles: a guide to an increasingly important dialect of organic chemistry

Seth C. Rasmussen¹

Received: 21 July 2016 / Accepted: 6 September 2016 / Published online: 17 September 2016
© Springer International Publishing Switzerland 2016

Abstract As a single system of chemical nomenclature is used worldwide, it is the universal language of chemistry and thus a critical component to effectively communicate with others in the field. Of course, in the same way that chemistry as a discipline covers a wide range of topics and classes of chemical compounds, the associated nomenclature of the field can also be subdivided into separate ‘dialects’, each of which addresses different classes of chemical species and contains its own set of specialized rules for the systemic naming of the included compounds. Polycyclic fused-ring arenes and heterocycles are classes of organic compounds that are finding growing importance in polymer chemistry, materials science, and pharmaceutical chemistry, yet the nomenclature of these compounds is rarely covered even in graduate texts and students are thus not taught how to apply this nomenclature as needed. As such, the current report aims to present the basic rules and application of fused-ring nomenclature such that one can apply it to most common cases.

Keywords Organic nomenclature · Polycyclic arenes · Fused-ring heterocycles

Introduction

When introducing the subject of chemistry to new students, it is often acknowledged that the study of the subject has many similarities to learning a new language. Not only must one become fluent in the symbolic language of the subject (symbols, chemical formulas, equations, etc.) that relates the microscopic to the macroscopic, one must ultimately become proficient in the use of chemical nomenclature, the formal systemic process for naming chemical compounds. As a single system of chemical nomenclature is used worldwide, it is thus a critical component to effectively communicate with others in the field [1]. Of course, in the same way that chemistry as a discipline covers a wide range of topics and classes of chemical compounds, the associated nomenclature of the field can also be subdivided into separate ‘dialects,’ each of which addresses different classes of chemical species and contains its own set of specialized rules for the systemic naming of the included compounds. It is one such subset of modern chemical nomenclature, that of polycyclic fused-ring arenes and heterocycles, which is the focus of the current publication.

Chemical nomenclature can be traced back to the beginning of the modern period of chemistry at the end of the eighteenth century. Although Torbern Bergman (1735–1784) published a limited nomenclature in 1782, most associate the formal beginnings of systemic chemical nomenclature with the efforts led by Guyton de Morveau (1737–1816) beginning in that same year [1–4]. These efforts resulted in the May 1787 publication of a 300-page book titled *Methode de Nomenclature Chimique* (Method of Chemical Nomenclature), which was coauthored by de Morveau, Antoine Lavoisier (1743–1794), Claude Berthollet (1748–1822), and Antoine de Fourcroy

✉ Seth C. Rasmussen
seth.rasmussen@ndsu.edu

¹ Department of Chemistry and Biochemistry, North Dakota State University, NDSU Dept. 2735, P.O. Box 6050, Fargo, ND 58108-6050, USA

(1755–1809) [1–3]. This publication established a rational system that allowed one to know the composition of a chemical species directly from its name and is the system still in use today for the naming of simple inorganic compounds.

As the study of organic chemistry did not really become a focus until the beginning of the nineteenth century, systemic organic nomenclature had to wait until efforts in 1889, beginning with focused discussions at the *International Congress of Chemistry* held in Paris. Although no rules of nomenclature were determined at the conference, it did establish a continuing *International Commission of Chemical Nomenclature* and these efforts ultimately led to the *Geneva Nomenclature* of 1892, which became the initial basis of modern organic nomenclature [1, 4–7]. The *Geneva Nomenclature*, however, was not comprehensive [1, 5] and did not contain rules for either polycyclic or heterocyclic systems, which led to various scholars publishing personal works aimed at addressing the lack of nomenclature for these systems [8]. Notable examples of such efforts included publications by Max M. Richter (b. 1861), Constantin I. Istrati (1850–1918), and Louis Bouveault (1864–1909) [8, 9]. Little of these early proposals found much acceptance, however, and much of the modern nomenclature of fused-ring arenes and heterocycles can be traced to the work of Austin M. Patterson (1876–1956) (Fig. 1) [10–15].

Fig. 1 Austin M. Patterson (1876–1956). Courtesy of Antiochiana, Antioch College

Professor and Chair at Antioch College, Austin Patterson (Ph.D. Johns Hopkins 1900), was an early force at *Chemical Abstracts* and was also a member of the working party of the *Commission for the Reform of the Nomenclature in Organic Chemistry* of the *International Union of Pure and Applied Chemistry* (IUPAC) [7, 10–12, 16]. Patterson began working on the numbering and nomenclature of cyclic systems as early as 1917 [13] and proposed to the editorial committee of the *Journal of the American Chemical Society* the publication of a catalog detailing the formula of cyclic organic systems, along with their names and numberings [10]. The committee agreed, but it was felt that it would be favorable to first establish commitment to Patterson's numbering systems. As such, an American commission was formed, of which Patterson was a member, which established a final set of numbering rules to be submitted to IUPAC [7, 10]. Patterson's numbering system for polycyclic systems [14] was given provisional status by IUPAC in 1930 [16] and was quickly adopted by *Chemical Abstracts*, with the initially proposed catalog published as *The Ring Index* by the American Chemical Society in 1940 [10]. However, Patterson's numbering system and his following systemic nomenclature of fused-ring systems [15] were not officially codified by IUPAC until the release of the 'IUPAC 1957 Rules' [10, 17]. As Patterson died in February of 1956 [11, 12], he unfortunately did not live to see this occur. This now established nomenclature of fused-ring systems has continued to be modified and revised with the most current rules published in 1998 [18].

Although the importance of chemical nomenclature is generally recognized, particularly as part of the chemistry curriculum, the subject does not generally receive much interest among practicing chemists beyond its most basic application [19]. This is particularly true of more complex systems, such as the fused-ring species which are the focus of the current discussion, which require a bit more effort and knowledge of systemic nomenclature beyond that covered in the first two years of the undergraduate curriculum. The practicing chemist typically regards formal chemical nomenclature as a necessary evil during the publication process, something imposed by more critical referees and journal editors to keep at least some structure and uniformity in the primary literature, as well as to facilitate communication or the retrieval of information [19]. In fact, it is now common for many authors to bypass the complication of correct formal nomenclature by reporting new species solely by structure and compound numbers without even an attempt to include any associated chemical names. A growing number of chemists also tend to rely on software programs to provide systematic names from structural formulas and while such programs perform reasonably well for simple systems, they often have more

difficulty with fused-ring systems and cannot be relied upon to give the desired name without errors.

Of course, although everyone learns the basic chemical nomenclature used to name simple inorganic species and common organic compounds, more complex forms of nomenclature are typically only used by chemists in specific fields and are not taught as part of the general undergraduate curriculum. This is especially true for the nomenclature of fused-ring arenes and heterocycles, which is not even mentioned in undergraduate texts and can only be found in a very limited number of graduate texts [20–24]. Even for those texts that do include this nomenclature, the coverage is extremely brief and not adequate for one to develop the ability to adequately apply this nomenclature with any competence or mastery. As a consequence, incorrect application of fusion nomenclature is somewhat commonplace in the scientific literature. Because of this reality and the growing importance of fused-ring organic species in polymer chemistry, materials science, and pharmaceutical chemistry, it was felt that an educational guide to the application of this nomenclature is long overdue. The aim of the following presentation is thus to present the basic rules of fused-ring nomenclature, along with their application via examples such that one can competently apply this nomenclature to most commonly encountered cases. The scope of this discussion, however, will be limited to conventional fused-ring systems without covering the more complicated bridged fused-ring analogs. It should also be pointed out that although the fused-ring nomenclature outlined by IUPAC and that of *Chemical Abstracts Service* (CAS) are very similar, there are some minor differences that date back to the initial development of this nomenclature in 1930 [16]. Whenever possible, such differences will be highlighted in the current discussion.

Nomenclature of simple fused polycyclic arenes

A number of fused aromatic ring systems have trivial names, as shown in Figs. 2 and 3. The structures given in Figs. 2 and 3 also illustrate the numbering system used throughout fused-ring nomenclature and thus the discussion will begin with an explanation of its correct application. However, it should first be pointed out that both phenanthrene and anthracene (Fig. 2) have non-standard numbering that does not follow the systematic conventions discussed here.

To correctly apply the systematic numbering of fused-ring systems, the compound is first oriented such that the greatest number of rings lies in a horizontal row. The structure is then flipped such that the greater number of rings is positioned above and to the right of this horizontal row. If multiple orientations give similar numbers of rings

to the upper right, the one that has as few rings as possible to the lower left is chosen. Numbering then begins at the most counterclockwise position not involved in fusion of the top ring furthest to the right and proceeds clockwise around the exterior of the fused-ring compound. Only those carbons not participating in ring fusions are numbered, with bridgehead positions given the number of the preceding non-fused position followed by a letter beginning with “a” [14, 17, 18, 21]. Thus, the two bridgehead positions of pentalene would be 3a and 6a.

For fused-ring aromatic systems not included in Figs. 2 and 3, systematic fusion nomenclature must then be applied to determine its formal name. The correct application of this systematic nomenclature begins with first identifying the trivially named structure contained within the polycyclic compound of interest that has the greatest number of rings [17, 18, 25–27]. If multiple structures with the same number of rings can be used, the structure containing the larger ring at the first point of differentiation is used. For the example given in Fig. 4, fluoranthene has a maximum of two consecutive six-membered rings before the smaller five-membered ring, while the remaining two polycycles contain three consecutive six-membered rings. Thus, fluoranthene would be of lower priority than the other two polycycles. If multiple structures still exist, then the structure with the lowest numbered bridgehead carbon is selected [18]. Thus, for the two remaining choices given in Fig. 4, the first bridgehead of aceanthrylene (2a) is lower than that of acephenanthrylene (3a), and thus aceanthrylene would be the correct parent in this example. The correct application of these principles to the species of Figs. 2 and 3 then results in the priority rankings given in Table 1.

This trivially named compound then becomes what is referred to as the *parent* species and its trivial name becomes the root of the final compound's name. Any remaining ring systems not described by the parent must then also be identified and given prefix names. Prefix names for monocyclic components with the maximum number of alternating double bonds are constructed through the combination of ‘cyclo-’ with the appropriate numerical prefix to describe the size of the ring (i.e., cyclobuta, cyclopenta, etc.). The exception to this is benzene, which is given the prefix ‘benzo-’ [17]. Prefix names for polycyclic arenes are formed by replacing the ‘-ene’ suffix with ‘-eno’ [18]. There are several exceptions to this, however, which use abbreviated forms as outlined in Table 2 [17]. In practice, the terminal vowel of the prefix is omitted if the parent name also begins with a vowel (i.e., benzanthracene, rather than benzoanthracene) [20, 26].

If there are two or more potential locations for the parent within the complete compound, the location that gives the simplest name via fusion nomenclature is preferred. If this guideline alone cannot distinguish between possible

Fig. 2 Trivial names of fused aromatic ring systems containing up to 20 carbons

choices, the location that results in the greatest number of directly fused (i.e., first-order) components is selected [18, 20]. An example of this is illustrated in Fig. 5.

Once the parent and 'substituent' components have been identified, the exact location and orientation of the fusions between the parent and the various components are described by what is called the *fusion locant*. The fusion position on the parent arene is described by a letter representing the face, or two-atom bond, that is the point of fusion. The first face described by the numbered positions 1 and 2 is lettered 'a', with the letters continuing alphabetically clockwise around the exterior of the parent (Fig. 6) [15, 17, 18, 20–25]. The fused atoms of the attached substituent are then described using the lowest possible position numbers of the substituent component and are always cited in the same order as the letter locant direction of the parent [18, 25]. The complete fusion descriptor is finally

given in brackets with the position numbers of the substituent first, followed by a hyphen and the italicized face letter of the parent [15, 17, 18, 23]. Thus, in the first example of Fig. 6, the 1- and 2-positions of naphthalene are fused to the 'j' face of aceanthrylene to give the complete fusion descriptor '[1,2-j]'. The complete fusion descriptor is then placed in the compound name directly following the prefix name of the substituent. In cases in which there is no ambiguity, such as a fused benzene, the position numbers are omitted and only the italicized face letter is given in brackets.

Finally, the positions of the fully named fused structure are renumbered, ignoring the original numbering of either the parent or components (Fig. 7) [14, 17, 18, 20, 21]. This is accomplished as previously described above for the parents in Figs. 2 and 3, by orienting the structure such that the greatest number of rings lies in along the horizontal and

Fig. 3 Trivial names of larger fused aromatic ring systems

	# of Rings	Ring Order	First Bridgehead
fluoranthene	4	6-5-6-6	3a
acephenanthrylene	4	6-6-6-5	3a
aceanthrylene	4	6-6-6-5	2a

1. All have same exact number of rings
2. Fluoranthene has smaller ring at first fusion, other two are the same
3. The first bridgehead carbon of aceanthrylene is lower than acephenanthrylene

aceanthrylene is the correct parent

Fig. 4 Example of determining the correct parent structure

to the upper right [14, 17, 18, 21, 25]. Numbering begins at the most counterclockwise position not involved in fusion of the top ring furthest to the right and proceeds clockwise around the exterior of the fused-ring compound. As previously discussed, bridgehead positions are not formally numbered, but are given the number of the preceding non-fused position followed by a letter beginning with “a” [14, 17, 18, 21, 25].

Extension of all-carbon polycyclic systems to fused-ring heterocycles: basic rules

The nomenclature of fused-ring heterocycles follows many of the same rules and methods as described above for polycyclic arenes. The most critical difference is in the correct selection of the parent which will be the foundation of the final name of the fused-ring heterocycle. This process is a bit more complicated than the previously discussed polycyclic arenes, and it is important to consider as heterocycles always have higher priority than arenes

Table 1 Priority ranking of parent polycyclic arenes

Rank	Polycyclic arene	Rank	Polycyclic arene
1	ovalene	19	triphenylene
2	pyranthrene	20	aceanthrylene
3	heptacene	21	acephenanthrylene
4	heptaphene	22	fluoranthene
5	trinaphthylene	23	anthracene
6	coronene	24	phenanthrene
7	rubicene	25	phenalene
8	hexacene	26	fluorene
9	hexaphene	27	acenaphthylene
10	tetraphenylene	28	<i>s</i> -indacene
11	pentacene	29	<i>as</i> -indacene
12	pentaphene	30	biphenylene
13	perylene	31	heptalene
14	picene	32	azulene
15	pleiadene	33	naphthalene
16	naphthacene/tetracene	34	indene
17	chrysene	35	pentalene
18	pyrene		

Rank of 1 is the highest priority parent, Refs. [17, 18, 25–27]

Table 2 Polycyclic arenes with abbreviated prefix names

Polycyclic arene	Full prefix	Abbreviated prefix
perylene	perylene-	perylo-
anthracene	anthraceno-	anthra-
phenanthrene	phenanthreno-	phenanthro-
acenaphthylene	acenaphthylene-	acenaphtho-
naphthalene	naphthaleno-	naphtho-

Refs. [17, 20, 25]

Fig. 5 Choosing between variable parent placement

[15, 18, 20, 27, 28]. Thus, if the fused-ring system only contains one heterocycle, it will always be the parent. As with the previously discussed arenes, many simple heterocycles also have trivial names as illustrated in Fig. 8.

Non-fused heterocycles of ten atoms or less that do not have trivial names are named using Hantzsch–Widman nomenclature [27–29]. This nomenclature involves combinations of the element prefixes given in Table 3 with the

Fig. 6 Determining the fusion locants and constructing the final fused-ring name **Fig. 7** Determining the final position numbers

appropriate Hantzsch–Widman suffixes for the corresponding ring size and saturation (Table 4). The final ‘a’ of the prefix is dropped when followed by a vowel. Numerical locants for the heteroatoms are then placed together in front of the replacement prefixes [20, 21, 27–29]. For heterocycles containing multiple heteroatoms, the number of similar heteroatoms are indicated by the appropriate numerical prefixes (i.e., di-, tri-, tetra-, etc.) and dissimilar heteroatom prefixes are listed in decreasing order of heteroatom priority [26, 28]. Priority is determined according to a descending order of groups in the periodic table and an ascending order of atomic numbers in each group as shown in Fig. 9 [14, 15]. Application of these two trends thus gives the following order of priority for the commonly encountered heteroatoms: O > S > Se > Te > N > P > As > Sb > Bi > Si > Ge > Sn > Pb > B [14, 15, 17, 18].

Finally, numbering of the ring begins with the most senior heteroatom and proceeds around the heterocycle to provide the lowest position numbers to the remaining heteroatoms. If multiple options exist, preference is given to assigning the lowest position numbers to the highest priority heteroatoms. The numerical locants listed in front of the replacement prefixes are also given in decreasing

Fig. 8 Names of simple heterocycles utilized in fusion nomenclature **Table 3** Element replacement prefixes

Element	Prefix	Element	Prefix
oxygen	oxa-	antimony	stiba-
sulfur	thia-	bismuth	bimutha-
selenium	sena-	silicon	sila-
tellurium	tellura-	germanium	germa-
nitrogen	aza-	tin	stanna-
phosphorus	phospha-	lead	plumba-
arsenic	arsa-	boron	bora-

Refs. [27–29]

Table 4 Hantzsch–Widman suffixes

Ring size	Unsaturated	Saturated
3	-irene ^a	-irane ^b
4	-ete	-etane ^b
5	-ole	-olane ^b
6A (O, S, Se, Te, Bi, Hg)	-ine ^c	-ane
6B (N, Si, Ge, Sn, Pb)	-ine ^c	-inane ^d
6C (B, F, Cl, Br, I, P, As, Sb)	-inine ^c	-inane ^e
7	-epine ^c	-epane ^e
8	-ocine ^c	-ocane ^e
9	-onine ^c	-onane ^e
10	-ecine ^c	-ecane ^e

Refs. [26–28]

^a The suffix ‘-irene’ may be used for rings containing only nitrogen^b The suffixes ‘-iridine’, ‘-etidine’, ‘-olidine’ are preferred for nitrogen-containing rings^c CAS drops the final ‘e’ on these suffixes for rings not containing nitrogen^d Used by CAS only for nitrogen systems. Other heteroatoms use the unsaturated suffix and saturation is expressed by hydro prefixes^e For nitrogen systems, CAS uses the unsaturated suffix and saturation is expressed by hydro prefixes

order of heteroatom priority [20, 27, 29]. Examples of various common heterocycles having Hantzsch–Widman names are given in Fig. 8.

Some fused-ring heterocycles also have trivial names and common examples of these are given in Figs. 10 and 11. As with the previous polycyclic arenes, several fused-ring heterocycles have non-standard numbering for various historical reasons. These include purine in Fig. 10 and the tricyclic heterocycles acridine, xanthene, and various chalcogen analogs of xanthene in Fig. 11.

For systems with multiple heterocycles, a number of additional rules are used to determine priority, beginning with ranking nitrogen-containing systems over other heterocycles [15, 27, 28]. If none of the heterocycles contain nitrogen, then priority is given to the ring with the most senior heteroatom. The heteroatom ranking used here is the same as described above in the discussion of

Hantzsch–Widman nomenclature. If priority still cannot be determined via these first two principles, then the following additional criteria are used in descending order: (1) the parent with the greatest number of rings; (2) the parent with the largest ring; (3) the parent with the most heteroatoms; (4) the parent with the greater variety of heteroatoms; (5)

1. Descending order of group			
13	14	15	16
5 B Boron	6 C Carbon	7 N Nitrogen	8 O Oxygen
13 Al Aluminum	14 Si Silicon	15 P Phosphorus	16 S Sulfur
31 Ga Gallium	32 Ge Germanium	33 As Arsenic	34 Se Selenium
49 In Indium	50 Sn Tin	51 Sb Antimony	52 Te Tellurium
81 Tl Thallium	82 Pb Lead	83 Bi Bismuth	84 Po Polonium

2. Ascending order of atomic number

Fig. 9 Rules for ranking the priority of heteroatoms

the parent with the greatest number of more senior heteroatoms [17, 18, 27, 28].

Once the parent is determined, the names of the remaining substituents are constructed by modifying the accepted radical name of the substituents (i.e., the name of the ring if it was a substituent) by replacing the terminating ‘-yl’ with ‘-o’ [15]. In most cases, the radical name to be modified is formed by replacing the terminal ‘-e’ of the heterocycle name with ‘-yl’. However, there are a few exceptions as shown in Table 5 [17, 27]. In addition, some of the examples given in Table 5 use abbreviated fusion prefixes [17]. The fusion locants are then determined such that the complete fusion descriptor can be constructed. This process is essentially identical to that previously described above for polycyclic arenes. The fusion position on the parent heterocycle is described by a letter representing the

face, or two-atom bond, that is the point of fusion. The first face described by the numbered positions 1 and 2 of the parent is lettered ‘a’, with the letters then continuing alphabetically clockwise around the exterior [15, 17, 18, 20–25]. The fused atoms of the attached substituent are then described using the lowest possible position numbers of the substituent component and are always cited in the same direction as the letter locants [18, 25].

Finally, the positions of the complete fused-ring heterocyclic structure are renumbered, ignoring the original numbering of either the parent or components. This is accomplished in the same way as with the previously discussed polycyclic arenes, although with additional considerations for heteroatom position numbers. As before, the structure is oriented such that the greatest number of rings lies along the horizontal and to the upper right [14, 17, 18, 21]. However, if multiple orientations are still possible, choose the one that would give the heteroatoms the lowest possible position numbers. Lastly, if the various rules above still allow multiple orientations, choose the one that gives the highest priority heteroatom the lowest position number [14, 17, 18, 27].

As with polycyclic arenes, numbering begins at the most counterclockwise position not involved in fusion of the top ring furthest to the right and proceeds clockwise around the exterior of the fused-ring compound. Fused-ring heterocycles differ from arenes, however, in that any heteroatoms located at bridgehead positions are given position numbers (see Fig. 10 for examples) [17, 18, 27]. As with arenes, bridgehead carbons are not formally numbered and are given the number of the preceding non-fused position followed by a letter beginning with ‘a’ [14, 17, 18, 21]. The remaining sections below will apply

Fig. 10 Trivial names of parent bicyclic heterocycles

Fig. 11 Trivial names of parent tricyclic heterocycles **Table 5** Radical and fusion prefix names of common heterocycles

Heterocycle	Radical name	Fusion prefix
furan	furyl-	furo-
imidazole	imidazolyl-	imidazo-
indole	indolyl-	indolo-
isoindole	isoindolyl-	isoindolo-
isoquinoline	isoquinolyl-	isoquino-
oxazole	oxazolyl-	oxazolo-
purine	purinyl-	purino-
pyran	pyranyl-	pyrano-
pyrazine	pyrazinyl-	pyrazino-
pyrazole	pyrazolyl-	pyrazolo-
pyridine	pyridyl-	pyrido-
pyrimidine	pyrimidinyl-	pyrimido-
pyrrole	pyrrolyl-	pyrrolo-
quinoline	quinolyl-	quino-
quinoxaline	quinoxalyl-	quinoxalino-
thiazole	thiazolyl-	thiazolo-
thiophene	thienyl-	thieno-

Refs. [17, 27]

these basic rules of fused-ring heterocyclic nomenclature to commonly encountered situations.

The simplest example of a fused-ring heterocycle: a single heterocycle fused to benzene

In beginning to apply the above rules to working examples, those fused-ring heterocycles involving the fusion of benzene to a single heterocyclic system are without doubt the

simplest, but are also perhaps the most commonly encountered cases of fused-ring systems. As can be seen in Fig. 10, many examples of such benzo-fused heterocycles have trivial names, such as the ubiquitous systems quinoline, quinoxaline, or indole. Although the common chalcogen analogs of indole were once also referred to by the trivial names benzofuran and thianaphthene (Fig. 12), these are no longer approved by IUPAC. Benzofuran and the analogous isobenzofuran, however, are still utilized by CAS [18].

As heterocycles always have higher priority than arenes, the parent of these examples is always the heterocycle. In addition, as there is no ambiguity in the fusion of the benzene, the fusion descriptor consists of only the italicized face letter of the parent heterocycle. Thus, as illustrated by the examples given in Fig. 12, the nomenclature of these species is typically quite straightforward and simple.

In addition to the fusion nomenclature discussed above, an alternate nomenclature system can be used specifically for fused-ring compounds consisting only of benzene fused to a single heterocyclic ring. In this alternate nomenclature, the prefix 'benzo-' is added to the Hantzsch–Widman or trivial name of the heterocycle, with the use of the trivial name preferred. The resulting composite name is then prefixed by numerical locants giving the positions of the heteroatoms in the final system (Fig. 12) [18, 27, 28]. These benzoheterocycle names (i.e., "benzo" names) are sometimes employed as parent components in fusion nomenclature [28].

When the heterocycle fused to benzene does not have a trivial name (i.e., is named via Hantzsch–Widman nomenclature), the numerical locants of the original

Fig. 12 Simple examples of benzoheterocycles

Hantzsch–Widman name must be included in the final name constructed via fusion nomenclature. In such a case, the original numerical locants are included in a separate set of brackets directly following the fusion descriptor (Fig. 13). Thus, while the numerical locants of the benzoheterocycle name describe the placement of the heteroatoms in the final fused-ring structure, the analogous numerical locants in fusion nomenclature describe the placement of the heteroatoms in the original non-fused heterocycle. As such, these numbers can differ between the two systems (Fig. 13) and the correct numbering system needs to be carefully applied in the two different nomenclature systems.

More complicated fused-ring heterocycles consisting of only two systems

While still limiting ourselves to systems resulting from the fusion of only one substituent to the parent, but involving more complex combinations, the most critical factor is the

Fig. 13 The incorporation of Hantzsch–Widman numerical locants in fusion nomenclature

correct identification of the parent. In addition, as more complicated substituents will almost always involve ambiguity in the fusion, these systems will also require a full fusion descriptor. Examples illustrating both the correct determination of the parent and the correct construction of the fusion descriptor are given in Fig. 14.

Fused-ring tricyclic heterocycles consisting of only two systems: symmetric examples

For fused-ring systems consisting of three rings, but only two components, the basic nomenclature is fairly similar to that discussed above, although with a few additional factors. One such factor is that as no ring should be expressed more than once in the final name of the compound [15], numerical prefixes must be added to indicate any multiple occurrences of a particular ring system. In addition, the fusion descriptor now becomes a bit more complex as it must adequately describe the multiple points of fusion. Various examples of fused-ring tricyclic heterocycles found in the conjugated polymer literature and their corresponding fusion names are given in Fig. 15.

The first example given in Fig. 15 is the most simple, in which two thiophenes are fused to a central pyrrole ring. As a nitrogen-based heterocycle, pyrrole has priority and is thus the parent here. To specify that two separate thiophenes are fused to the central pyrrole, the numerical prefix ‘di-’ is added to the thiophene substituents [18] to give the base name ‘dithienopyrrole’. As with previous cases, the bracketed fusion descriptor is then inserted between the substituents and parent. The difference here is that the fusion descriptor must now detail two separate fusions, rather than just a single fusion. Both sets of fusion locants are included in the same set of brackets, although separated by a colon with the fusion containing the lowest locant letter given first [18]. To differentiate between the two identical substituents, unprimed numbers are used to describe the first substituent and primed numbers are used to describe the second.

The related phosphorus and arsenic compounds have identical structures and thus are commonly named in the literature in the same way as dithieno[3,2-*b*:2',3'-*d*]pyrrole with either phosphole or arsole as the parent. Not only does this simplify the nomenclature, but also highlights the central heterocycle of interest in these compounds. However, as none of the heterocycles contain nitrogen, the parent is then determined by heteroatom priority and sulfur has higher priority than either phosphorus or arsenic. As such, while structurally identical, the thiophene is the parent in both of these analogs and thus the correct names are not analogous to the pyrrole compound. As the parent occurs twice in these examples, the numerical prefix is still

Fig. 14 Examples of fused-ring systems consisting of two heterocycles

Fig. 15 Examples of fused-ring tricyclic heterocycles consisting of only two components

added, but is now applied to the parent. Thus, the correct base name of the phosphorus example is 'phosphodithiophene'. In cases such as the thiazole examples given in Fig. 15, the prefix 'di-' would give 'dithiazole' which itself is the name of another heterocycle. In such cases where the application of the standard type 1 numerical prefix (i.e., di-, tri-, etc.) can lead to confusion, then type 2 prefixes (i.e., bis-, tris-, etc.) are used instead [18]. Thus, the correct base name of the thiazole examples here are 'benzobisthiazole'.

For such examples that contain two identical parents and a shared substituent, the fusion descriptor must again detail both separate ring fusions. However, the difference here is that the use of primes is now applied to the parent locant letters as the two letters refer to

two different parents. In contrast, all of the substituent numbers now refer to the single substituent and thus all numbers are unprimed. As with all fusion descriptors, the fused atoms of the attached substituent are described using the lowest possible position numbers of the substituent component and are always cited in the same direction as the letter locants [18]. However, in cases in which two identical parents share a substituent, the numbering used in the first fusion descriptor sets the overall numbering of the substituent and the position numbers for the second fusion descriptor are thus determined by the first fusion descriptor. In cases where both locant letters are the same, the fusion with the lowest substituent numbers is listed first with the corresponding locant letter unprimed.

Fused-ring heterocycles consisting of three different systems

The final set of examples is the most complicated to be discussed. The more straightforward of these cases are illustrated by the first two examples in Fig. 16. In both of these compounds, two different substituent rings are fused to the central parent. In this case each substituent is cited in alphabetical order [18] with the corresponding fusion descriptor listed directly after each substituent. The first component cited is given the lowest locant letter [18]. Therefore, in the first example, furan is cited before thiophene and thus the pyrrole face fused to furan is 'b' rather than 'd'.

The final two examples illustrate compounds in which the parent is a terminal ring with one substituent directly fused to the parent and a secondary substituent fused to the primary substituent. Unlike the first two examples discussed above, the substituents here are not cited alphabetically, but are given in the order of fusion. Thus, the secondary substituent is cited first, followed by the primary substituent and then the parent.

The two fusions are then described by two separate fusion descriptors. The primary fusion descriptor is described in exactly the same way as for all of the other examples already discussed above and this descriptor then sets the numbering of the first substituent. The procedure for describing the fusion between the primary and secondary substituent is then similar, but consists of two sets of numerical locants separated by a colon rather than

numbers and a letter locant separated by a hyphen [17, 18, 26]. The first two numbers describe the fusion positions of the secondary substituent and are given as primed numbers. The second two numbers describe the fusion positions of the primary substituent to the secondary [17], with the position numbers already determined by the fusion descriptor describing the fusion of the primary substituent to the parent.

Conclusion

In conclusion, as fused-ring organic species are continuing to grow in importance for applications in polymer chemistry, materials science, and pharmaceutical chemistry, basic knowledge of the nomenclature of these compounds is necessary to be able to effectively communicate with other scientists working in these fields. In fact, the proliferation of incorrect names in the chemical literature is due to the current lack of basic training in this nomenclature and contributes to unnecessary confusion in reports of these compounds. As with all dialects of nomenclature, mastery requires learning a number of rules, as well as suitable practice in the application of these rules. However, if one can become proficient in the basic tasks of determining the correct parent ring system and then constructing the corresponding fusion descriptor, the process becomes fairly straightforward with most commonly encountered cases. It is hoped that the overall discussion presented here should adequately prepare those that wish to learn the fundamentals of this system of nomenclature.

References

- Loening KL (1970) International cooperation on scientific nomenclature. *J Chem Docum* 10:231–236
- Oesper RE (1945) The birth of modern chemical nomenclature. *J Chem Ed* 22:290–292
- Partington JR (1989) A short history of chemistry, 3rd edn. Dover, New York, pp 135–136
- Evieux EU, Oesper RE (1954) The Geneva congress on organic nomenclature, 1892. *J Chem Soc* 31:326–327
- Traynham JG (1987) The official rules for organic chemical nomenclature: emergence, evolution, emphases, and errors. *J Chem Ed* 64:325–326
- Verkade PE (1985) A history of the nomenclature of organic chemistry. D. Reidel Publishing Company, Dordrecht, pp 1–2
- Smith HA Jr (1992) The centennial of systematic organic nomenclature. *J Chem Ed* 69:863–865
- Verkade PE (1985) A history of the nomenclature of organic chemistry. D. Reidel Publishing Company, Dordrecht, pp 62–70
- Richter MM (1896) Ein Beitrag zur Nomenclatur. *Ber* 29:586–608
- Verkade PE (1985) A history of the nomenclature of organic chemistry. D. Reidel Publishing Company, Dordrecht, pp 108–115

Fig. 16 Examples of fused-ring tricyclic heterocycles consisting of three components

11. Anonymous (1949) Patterson wins new award for service to chemical literature. *Chem Eng News* 27:1363
12. Corwin JF (1956) Austin McDowell Patterson (1816–1956). *J Chem Ed* 33:396
13. Patterson AM, Curran CE (1917) A system of organic nomenclature. *J Am Chem Soc* 39:1623–1638
14. Patterson AM (1925) Proposed international rules for numbering organic ring systems. *J Am Chem Soc* 47:543–561
15. Patterson AM (1928) The nomenclature of parent ring systems. *J Am Chem Soc* 50:3074–3087
16. Patterson AM (1933) Definitive report of the commission on the reform of the nomenclature of organic chemistry. *J Am Chem Soc* 55:3905–3925
17. IUPAC (1960) Definitive rules for nomenclature of organic chemistry. *J Am Chem Soc* 82:5545–5574
18. Moss GP (1998) Nomenclature of fused and bridged fused ring systems. *Pure Appl Chem* 70:143–216
19. Loening KL (1985) Foreward. In: Verkade PE (ed) *A history of the nomenclature of organic chemistry*. D. Reidel Publishing Company, Dordrecht, pp ix–x
20. March J (1968) *Advanced organic chemistry: reactions, mechanisms, and structure*. McGraw-Hill Inc, New York, pp 926–930
21. Stowell JC (1988) *Intermediate organic chemistry*. Wiley, New York, pp 6–16
22. Gupta RR, Kumar M, Gupta V (1998) *Heterocyclic chemistry*. Springer, Heidelberg, pp 18–24
23. Joule JA, Mills K (2010) *Heterocyclic chemistry*, 4th edn. Wiley, New York, pp 1–3
24. Bansal RK (2014) *Heterocyclic chemistry*, 5th edn. New Age International, New Delhi, pp 1–8
25. Fletcher JH, Dermer OC, Fox RB (1974) *Nomenclature of organic compounds. principles and practice*. *Advances in chemistry series 126*. American Chemical Society, Washington, DC, pp 29–42
26. Fox RB, Powell WH (2001) *Nomenclature of organic compounds. Principles and practice*, 2nd edn. American Chemical Society, Washington DC, pp 78–85
27. McNaught AD (1976) The nomenclature of heterocycles. *Adv Heterocycl Chem* 20:175–319
28. Fox RB, Powell WH (2001) *Nomenclature of organic compounds. Principles and practice*, 2nd edn. American Chemical Society, Washington DC, pp 101–117
29. Powell WH (1983) Revision of the extended Hantzsch–Widman system of nomenclature for heteromonocycles. *Pure Appl Chem* 55:409–416