IUPAC, Systematic Nomenclature for CIPAC Documentation – an Analysis

Seventh JOINT CIPAC/FAO/WHO Meeting - Symposium
(54th CIPAC Meeting and 9th JMPS Meeting)
Ljubljana, June 8th 2010

Dr. Reiner Kober (BASF SE, ECPA/Specifications Expert Group) **&**
Dr. Ursula Bünzli-Trepp
(Member of IUPAC's Chemical Nomenclature Advisory Subcommittee)






IUPAC, Systematic Nomenclature for CIPAC Documentation

Agenda:

1. Introduction, definitions & basics
2. About IUPAC, CA & other name construction systems
3. IUPAC 2010: What's new in the last 17 years
   - # About IUPAC names – significant changes
   - # About chemical-structure drawing
4. Systematic name construction – an example
5. CIPAC-doc checks – still close to IUPAC?
6. References (links, nomenclature software & book ref.)
7. Summary

Missing Awareness for Nomenclature Compliance

Some main reasons:

- **Unawareness of the importance of unequivocal chemical names for registration, indexing, patents, publications**
- **Limited expertise and education at universities & industries**
- **Missing time, money & willingness**
- **Chemical nomenclature: complex & confusing (e.g., through Internet)**
- **No (product) premium for nomenclature compliance**

About the Language of Chemistry

- Signs (Science), Codes & Names -

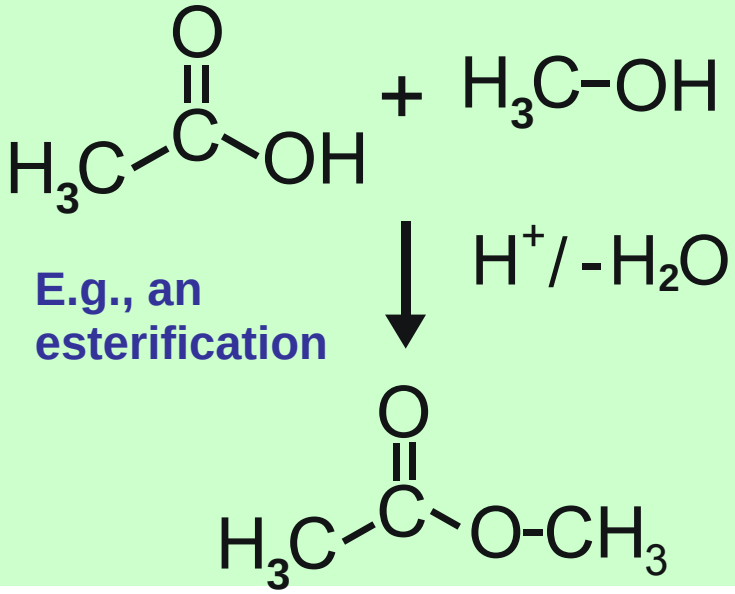
Letter codes - for atoms

…B,C,F,H,K,N,P,S…

Words - for molecules & substances

Sentences - for reactions

Texts - for synthetic pathways,
e.g., on an active ingredient



Until the end of the 19th century, alchemists invented nonsystematic, more source-related or trivial names for empirically found substances, **e.g.: phosphorus, carbinol (methanol), urea**

Chemical-Name Construction Systems

- More than 100 Years of Overlapping Traditions -

Nomenclature type, e.g.	Examples
Substitutive nom.	2-chloropropane
Conjunctive nom.	benzeneethanol
Fusion nom.	1,3-benzothiazole
Subtractive nom.	deoxy-, anhydro-
Additive nom.	tetrahydrofuran
Hantzsch-Widman nom.	1,3-oxazole
Replacement nom.	silacyclotetradecane

Chemical-Name Construction Systems
- More than 100 Years of Overlapping Traditions -

Nomenclature type, e.g.	Examples
Functional-class nom.	diethyl ether
Systematic nom.	
IUPAC name (PIN*)	ethoxyethane
CA name	1,1´-oxybis[ethane]
CA name (inverted for index)	ethane, 1,1´-oxybis-

*Preferred IUPAC Name

That's IUPAC Nomenclature

International Union of Pure and Applied Chemistry

- In 1892, Geneva Convention with the first approach to develop a systematic nomenclature for chemical substances
- IUPAC was established in 1919 in Geneva (Switzerland)
- IUPAC's Blue Book on Org. Chem. Names:

  # 1979: last printed version (out of print)
  # 1993: last but incomplete update (printed guide)
  # 2004: complete but still provisional revision (inconsistencies; www. – see ref. page)

- In 2006 (confirmed in 2009), English was adopted as IUPAC's official language
- Translations into >20 other languages by national chemical organizations like GDCh

That's CAS

Chemical Abstrats Service

In many cases, CA follows the IUPAC recommendations

- Established in 1907 as Division of the Am. Chem. Soc.
- Generation of abstracts of published documentations (chemical journals, patents) incl.:
  - CAS Registry Numbers (RN) for chemical substances
  - >50 Mio substance codes
  - >32 Mio document records
  - CA index names (inverted names based on an Index Guide 2004, now www. – see ref.) & corresponding noninverted CA names

About IUPAC Names

IUPAC-Related names	Definitions	Notes
PINs	**P**referred **I**UPAC **N**ames	IUPAC – highest tier **2004 draft on PINs as provisional recommendations**
General-use names **(general IUPAC names)**	substituents (prefixes) allowed	**still IUPAC-accepted, but only in discussions or, e.g., to bridge to older references**
	substituents (prefixes) not allowed	
Trivial names	typically used for only one single chemical compound	no or only very limited IUPAC support

Trivial Names – Still Important for Natural Products

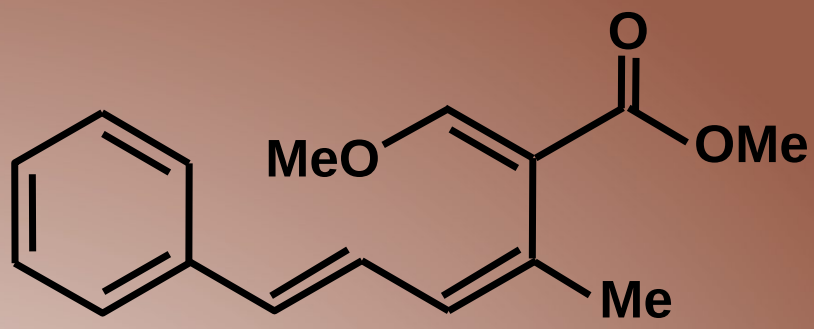
E.g., Strobilurin A – Lead Molecule for Powerful Fungicides

Initial paper on Strobilurin A:

Antibiotics from Basidiomycetes, III. Strobilurin A and B, Antifungal Metabolites from Strobilurus tenacellus.

Schramm, Steglich, Anke, and Oberwinkler

Chem. Ber. 1978, vol. 111, p. 2779 ff.



Strobilurin A

Strobilurus tenacellus

CLP (=GHS) Regulation (EC) No 1272/2008

(44) The International Union of Pure and Applied Chemistry (IUPAC) is a long-standing global authority on chemical nomenclature and terminology. Identification of substances by their IUPAC name is widespread practice worldwide and provides the standard basis for identifying substances in an international and multilingual context. It is therefore appropriate to use these names for the purposes of this Regulation.

(45) The Chemical Abstracts Service (CAS) provides a system whereby substances are added to the CAS Registry and are assigned a unique CAS Registry Number. Those CAS numbers are used in reference works, databases, and regulatory compliance documents throughout the world to identify substances without the ambiguity of chemical nomenclature. It is therefore appropriate to use the CAS numbers for the purposes of this Regulation.

EU CLP standard:

IUPAC Names & CAS registry numbers (RN)

IUPAC Names – What's new (1)
- Alkanes -

PINs (IUPAC)	Not accepted	IUPAC-Accepted use names (no substituents)	CA Names (inverted)
methane			methane
ethane			ethane
propane			propane
butane	*n*-butane		butane
pentane	*n*-pentane		pentane
hexane	*n*-hexane		hexane
2-methylpropane		isobutane	propane, 2-methyl-
2,2-dimethylpropane		neopentane	propane, 2,2-dimethyl-

Forget *n*-alkanes because, e.g., butane means already an unbranched C_4-chain

IUPAC Names - What's new (2)
- Alkanols -

PINs (IUPAC)	Not accepted	IUPAC-Accepted use names (no substituents)	CA Names (noninverted)
methanol			methanol
ethanol			ethanol
propan-1-ol	*n*-propanol		1-propanol
propan-2-ol	isopropanol	isopropyl alcohol	2-propanol
butan-1-ol	*n*-butanol		1-butanol
butan-2-ol	*sec*-butanol	*sec*-butyl alcohol	2-butanol
2-methylpropan-1-ol	isobutanol	isobutyl alcohol	2-methyl-1-propanol
2-methylpropan-2-ol	*tert*-butanol	*tert*-butyl alcohol	2-methyl-2-propanol
pentan-1-ol	*n*-pentanol		1-pentanol

Use PINs in CIPAC texts for alkan-1-ol or …2-ol solvents

IUPAC Names - What's New (3)
- Alkanoic Acids -

PINs (IUPAC)	IUPAC-Accepted use names	CA Names
formic acid	methanoic acid	formic acid
acetic acid	ethanoic acid	acetic acid
	IUPAC-Accepted trivial names *- no substituents -*	
propanoic acid	propionic acid	propanoic acid
butanoic acid	butyric acid	butanoic acid
pentanoic acid	(valeric acid*)	pentanoic acid
hexanoic acid	(capronic acid*)	hexanoic acid

*IUPAC removed many trivial acid names already in 1993 (marked by an asterisk (*)). Moreover, other traditionally used substituent names are now only accepted for general use, e.g., vinyl- (PIN: ethenyl-)

About IUPAC Chem.-Structure Drawing

Reference: http://www.iupac.org/publications/pac/2008/pdf/8002x0277.pdf

Pure Appl. Chem., Vol. 80, No. 2, pp. 277–410, 2008.
doi:10.1351/pac200880020277

INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY

CHEMICAL NOMENCLATURE AND STRUCTURE REPRESENTATION DIVISION*

GRAPHICAL REPRESENTATION STANDARDS FOR CHEMICAL STRUCTURE DIAGRAMS**

(IUPAC Recommendations 2008)

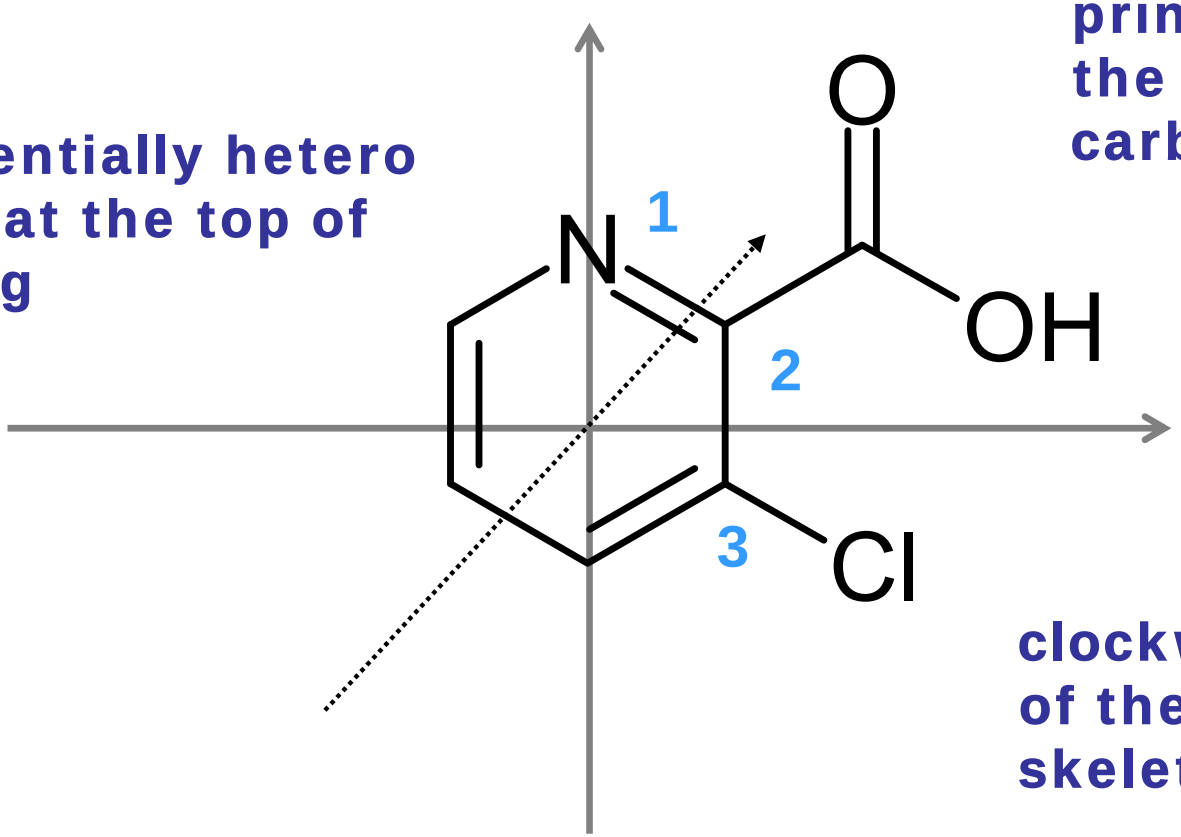
Prepared for publication by
JONATHAN BRECHER

CambridgeSoft Corporation, 100 CambridgePark Drive, Cambridge, MA 02140, USA

About IUPAC Chem.-Structure Drawing
- What's New -

About heterocycles

preferentially hetero atoms at the top of the ring



principal group to the right – here a carboxylic acid

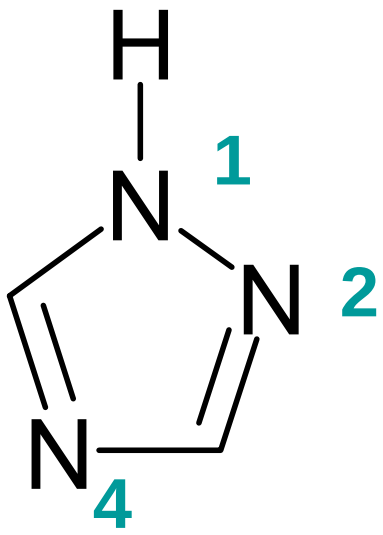
clockwise numbering of the molecular-skeleton parent

Note: For some heterocycles, other preferential orientations are accepted, e.g., due to biochemical traditions

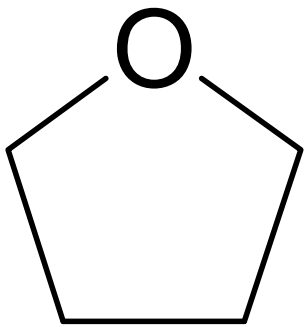
IUPAC Chem.-Structure Drawing

- Heterocycles -

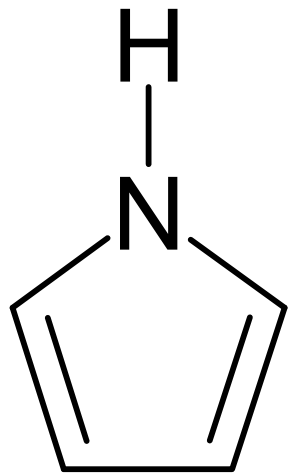
upwards orientation



PIN and CA:
1*H*-1,2,4-triazole



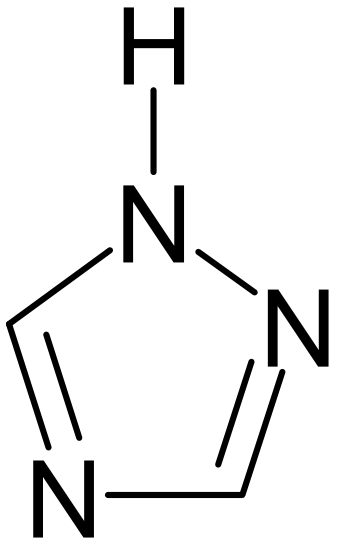
PIN: **oxolane**
tetrahydrofuran
(CA and IUPAC's general-use name)



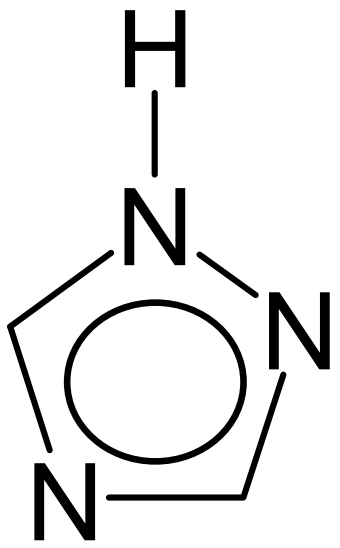
PIN and CA:
1*H*-pyrrole

IUPAC Chem.-Structure Drawing

- Heterocycles -



preferred

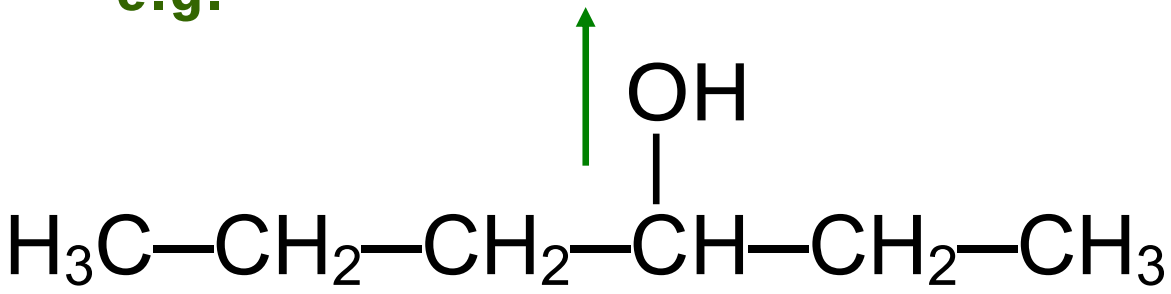


acceptable

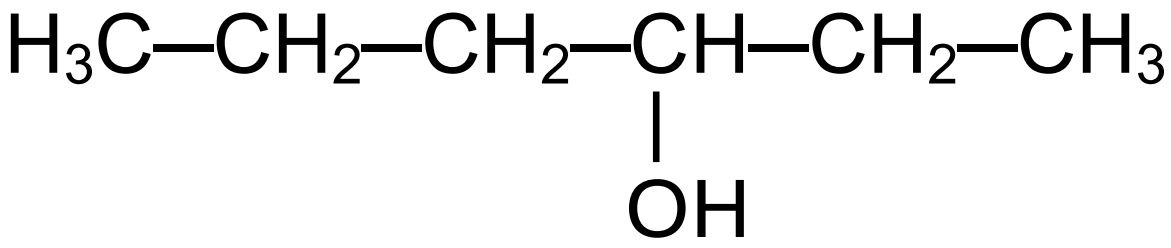
IUPAC Chem.-Structure Drawing

- Horizontal & Upwards Orientation -

Substituents should preferentially point upwards, e.g.



preferred

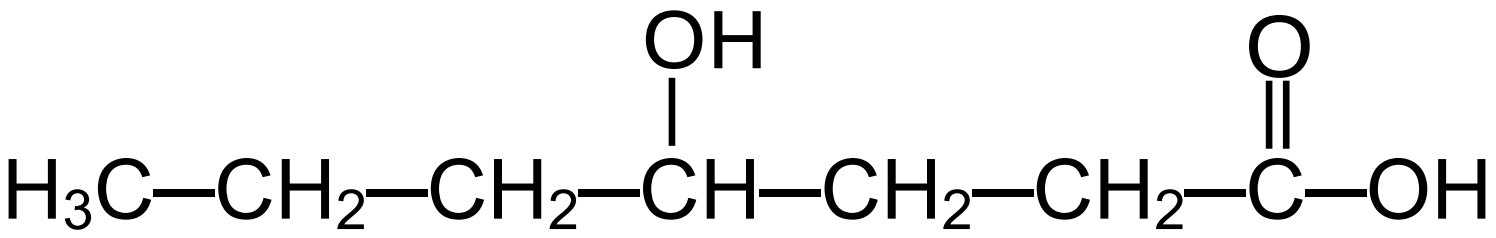


acceptable

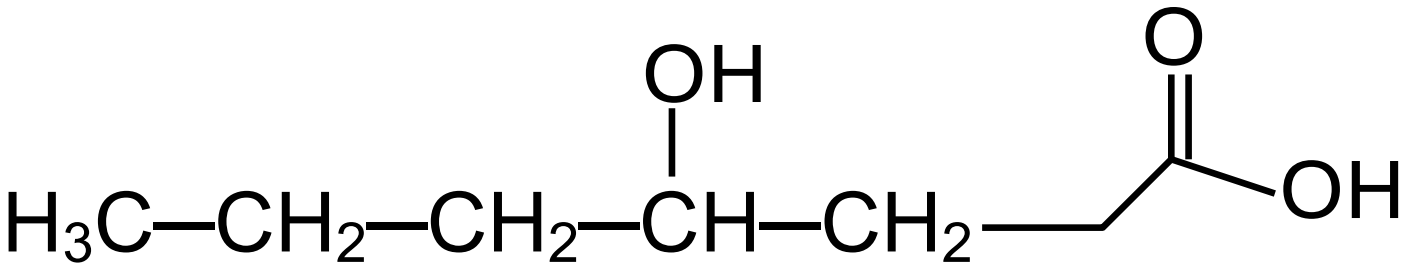
IUPAC Chem.-Structure Drawing

- Drawing Style -

Drawing style should be consistent, e.g.



preferred

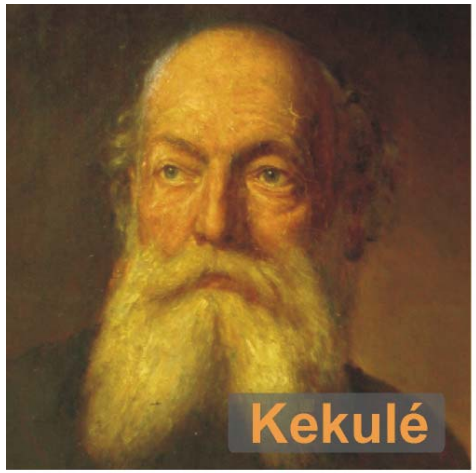


not acceptable

IUPAC Chem.-Structure Drawing

- Drawing Style -

Friedrich August Kekulé von Stradonitz **(1829-1896) proposed the cyclic structure of benzene and pioneered symbolic representation of organic structures.**



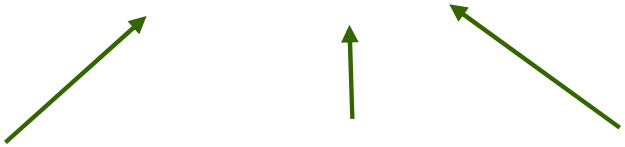
preferred **acceptable** **not acceptable**

About Enclosing Marks (1)

- In Names -

curved-square-curly

IUPAC: { [(…)] }



parentheses - brackets - braces

curved-square-square

CA: [[(…)]]

About Enclosing Marks (2)

- In Formulae -

Substituent with 2 free valencies:

IUPAC recom.: $-[CH_2]_n-$

Scarce acceptance in the chemical community

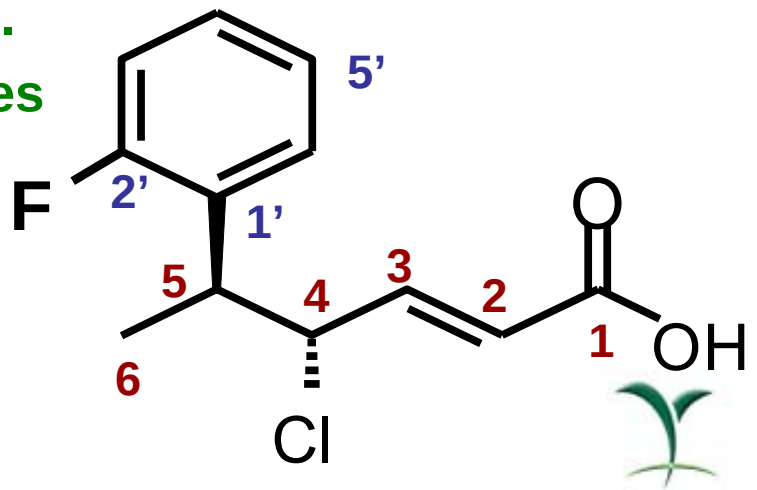
In practice: $-(CH_2)_n-$ **Mostly, only the curved enclosing marks are in use**

Accepted for larger units, e.g.: $-[CH_2-CH_2-O]_n-$

Systematic Name Construction: an Example

The Principles, esp.:

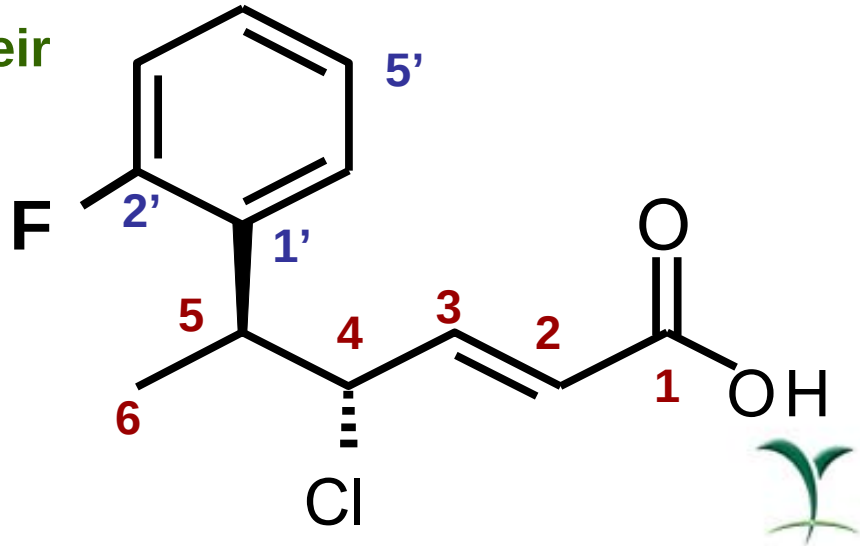
1. Identify the senior compound class by choosing the principal group among all characteristic groups (= functional groups). The principal group furnishes the ending of the name (suffix), the remaining functional groups are named as prefixes.
2. Identify the senior molecular-skeleton parent (here: hexane)
3. Number the molecular-skeleton parent by attributing locants.
4. Give prefixes to possibly present side chains or nonsenior rings and number them (lowest locant for free valency); give prefixes to all other substituents and nonsenior functional groups.
5. Determine the alphabetical order of the prefixes and arrange them with their locants.
6. Determine the configurations of stereogenic units and complete the name with the corresponding stereodescriptors.



Systematic Name Construction: an Example

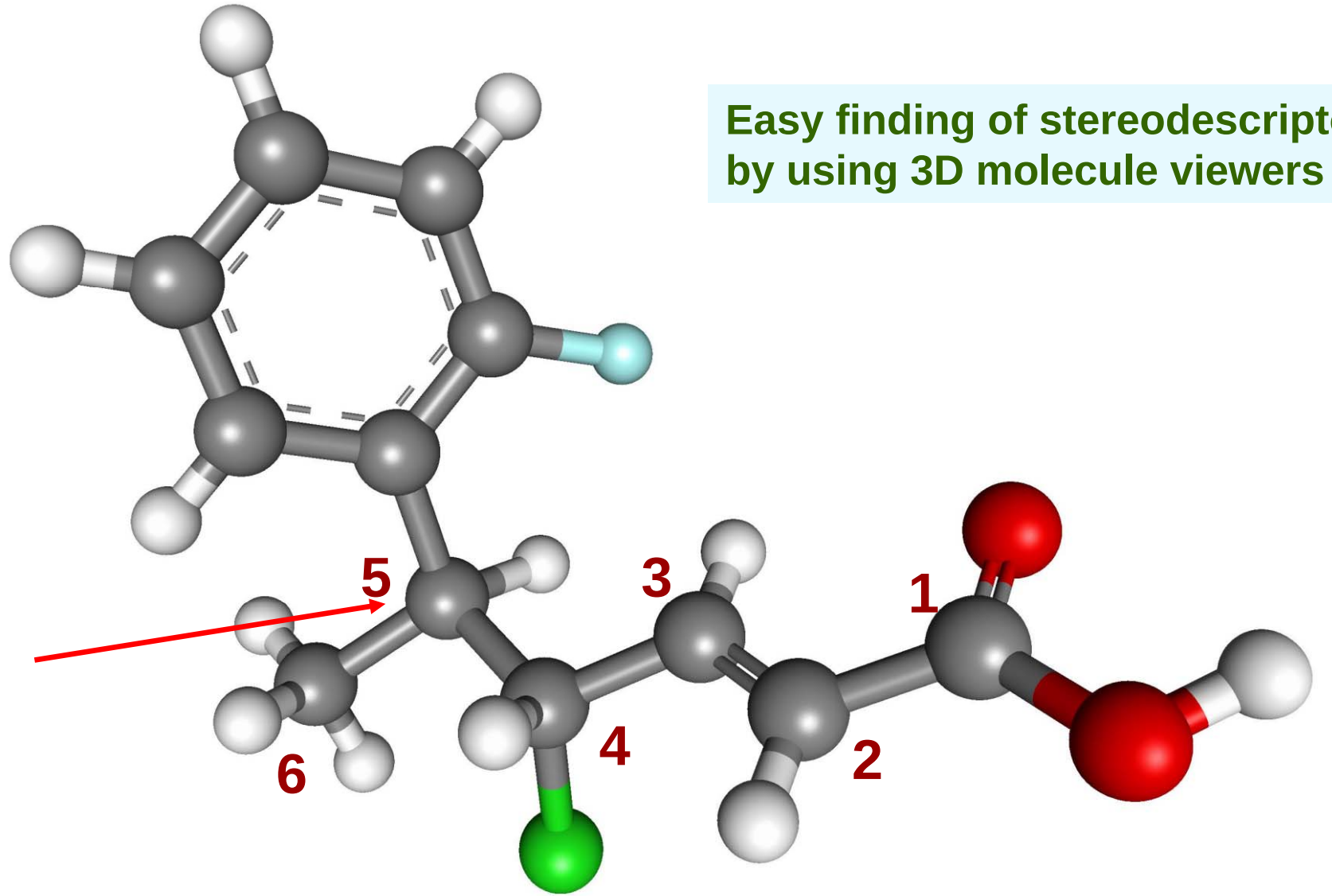
Resolution:

1. **Senior compound class, determined by the principal group (acid):**
   -oic acid
2. **Senior molecular-skeleton parent: hexane → hexene**
1. **Locants of the molecular-skeleton parent: 1 to 6**
2. **Prefixes for nonsenior rings or side chains with their own locants, prefixes for other substituents or nonsenior functional groups:**
   (2-fluorophenyl)-, chloro-
3. **Alphabetical order of the prefixes with their locants of the molecular-skeleton parent:**
   4-chloro-5-(2-fluorophenyl)-
4. **Determine stereodescriptors:**
   - see next slides



Systematic Name Construction: an Example

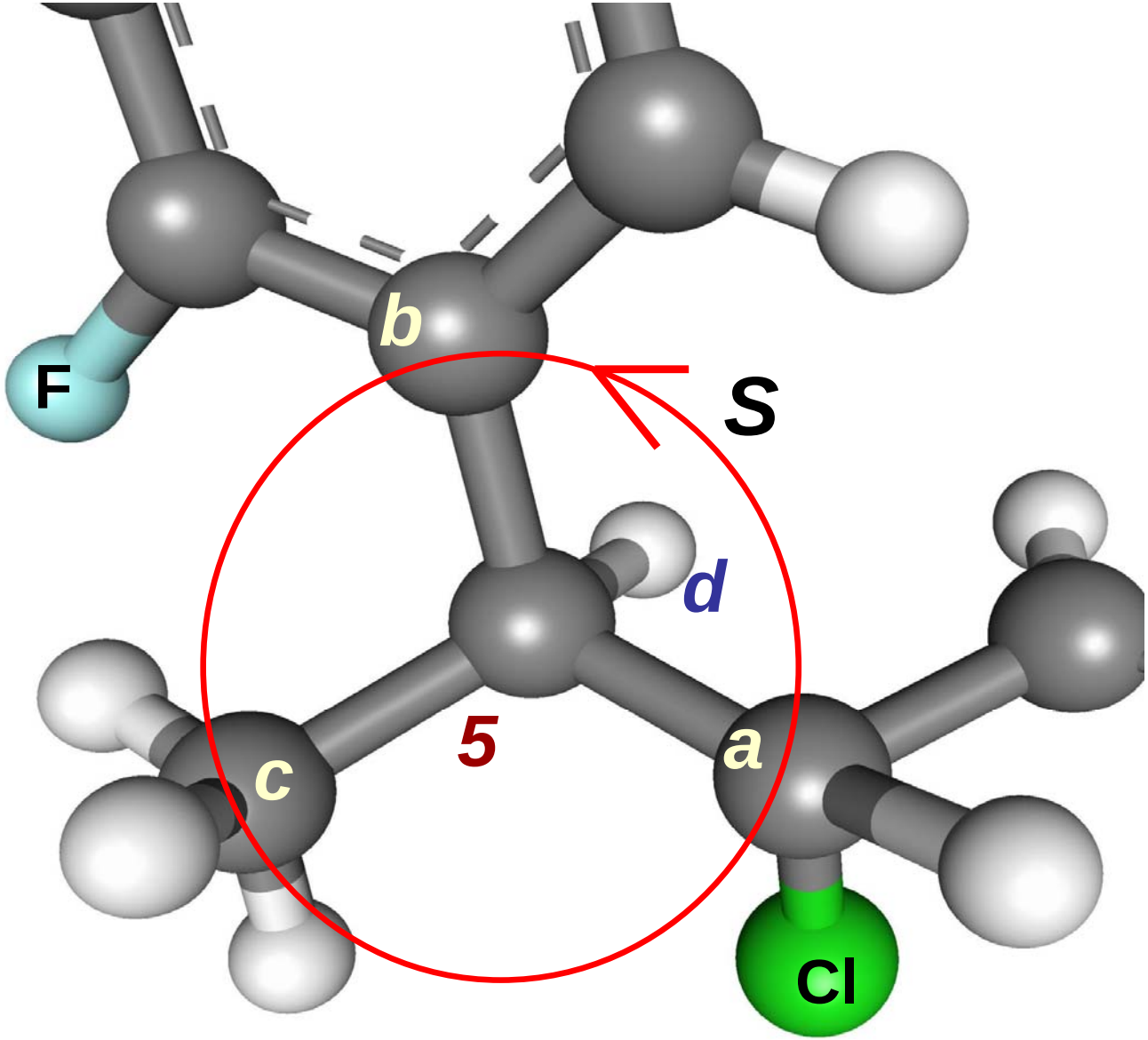
- Stereodescriptors by the *Cahn-Ingold-Prelog* System -



Easy finding of stereodescriptors by using 3D molecule viewers

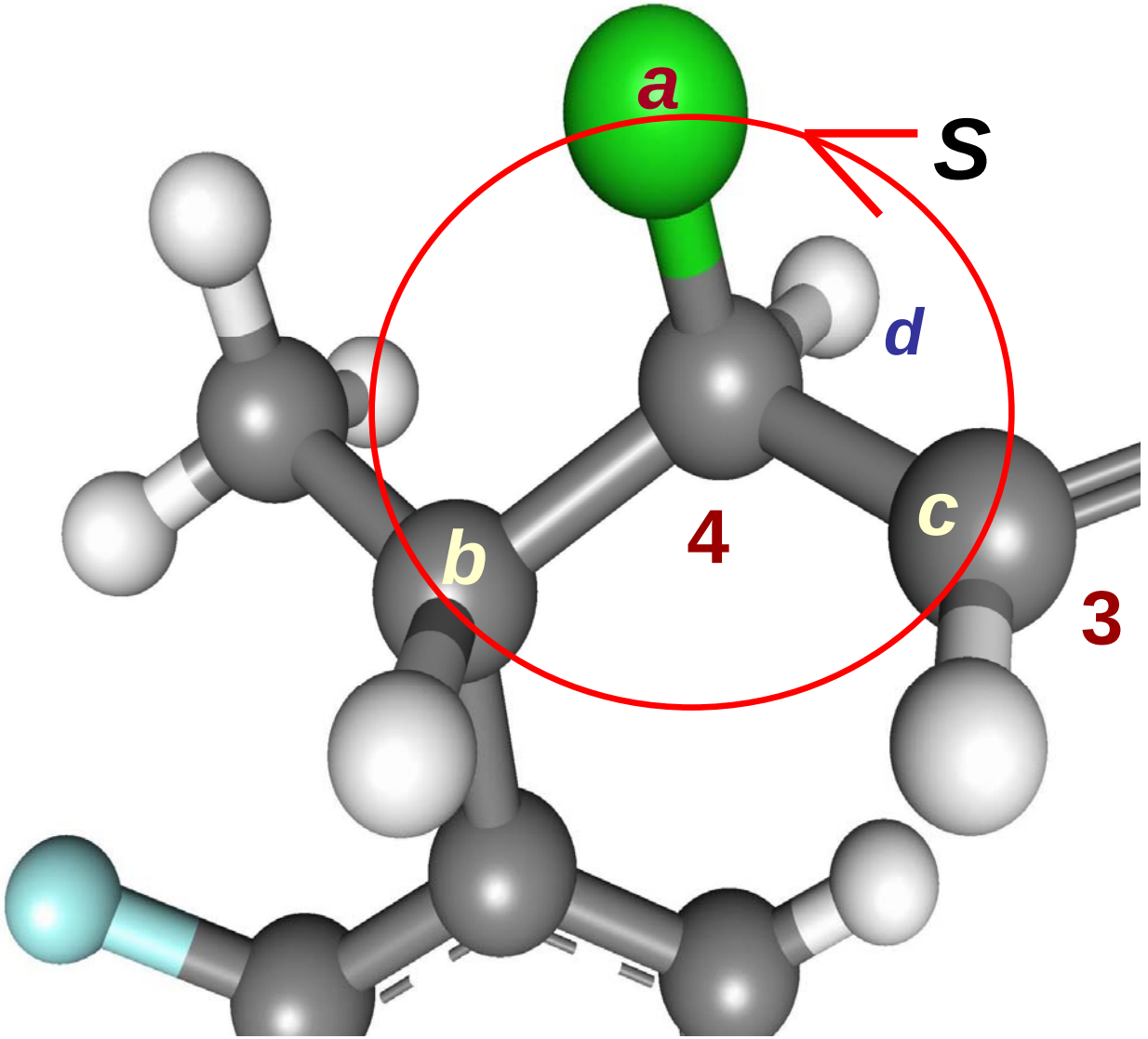
Systematic Name Construction: an Example

- Stereodescriptors by the *Cahn-Ingold-Prelog* System -



Systematic Name Construction: an Example

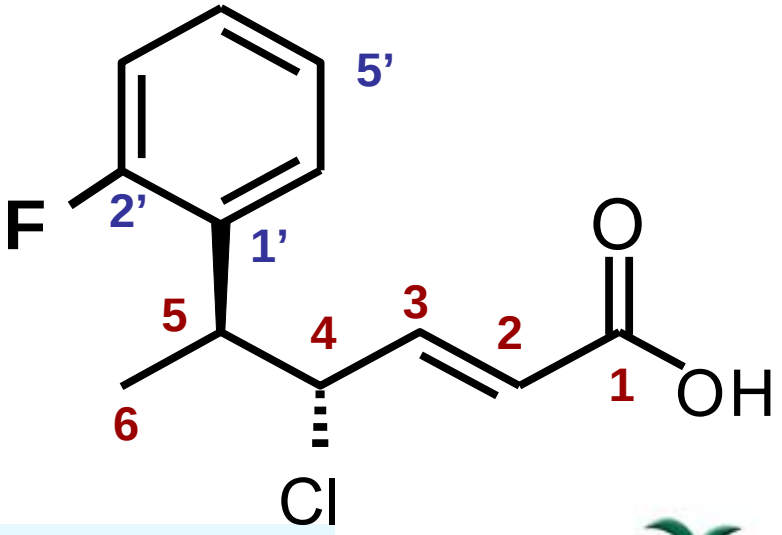
- Stereodescriptors by the *Cahn-Ingold-Prelog* System -



Systematic Name Construction: an Example

Resolution:

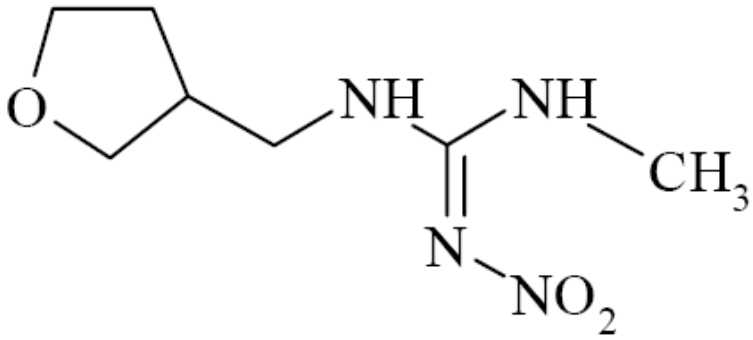
1. Senior compound class, determined by the principal group (acid):
   -oic acid
2. Senior molecular-skeleton parent: **hexane → hexene**
3. Locants of the molecular-skeleton parent: **1 to 6**
4. Prefixes for nonsenior rings or side chains with their own locants, prefixes for other substituents or nonsenior functional groups:
   (2-fluorophenyl)-, chloro-
5. Alphabetical order of the prefixes with their locants of the molecular-skeleton parent:
   4-chloro-5-(2-fluorophenyl)-
6. Determine stereodescriptors:
   (2*E*,4*S*,5*S*)



(2*E*,4*S*,5*S*)-4-chloro-5-(2-fluorophenyl)hex-2-enoic acid

CIPAC-Doc Checks on IUPAC/CA '2010' Accuracy

Here, a print screen out of CIPAC Handbook L, p. 67 on Dinotefuran



ISO common name	Dinotefuran
Chemical name	(RS)-1-methyl-2-nitro-3-(tetrahydro-3-furylmethyl)-guanidine (IUPAC); N-methyl-N'-nitro-N''-[(tetrahydro 3-furyl)methyl]guanidine (CA; 165252-70-0)

9 Ai.s were checked (structure drawing, IUPAC and CA names)

CIPAC-Doc Checks on IUPAC/CA '2010' Accuracy

- Results from 9 Doc.s (Handbook K and L)-

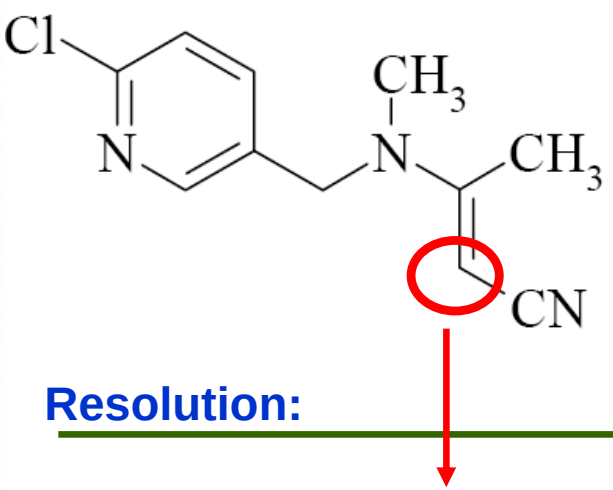
Page	Ai. CIPAC No.	ISO Name	Correctness according to:			
			Formula drawing	PIN (IUPAC)	CA Name	
4	649	acetamiprid*				*
23	584	azimsulfuron*				*
38	750	*S*-bioallethrin				
45	333	deltamethrin				
61	734	diflovidazin				
67	749	dinotefuran				
104	174	picloram				
38	609	epoxiconazol*				*
77	568	kresoxim-methyl				

*3 detailed analysis – see next slides

OK | Slight discrepancies | Strong discrepancies

CIPAC-Doc Checks on IUPAC/CA '2010' Accuracy

- Acetamiprid doc. / Handbook L, page 4 ff. -



Acetamiprid

(*E*)-N^1-[(6-chloro-3-pyridyl)methyl]-N^2-cyano-N^1-methyl-acetamidine (IUPAC);

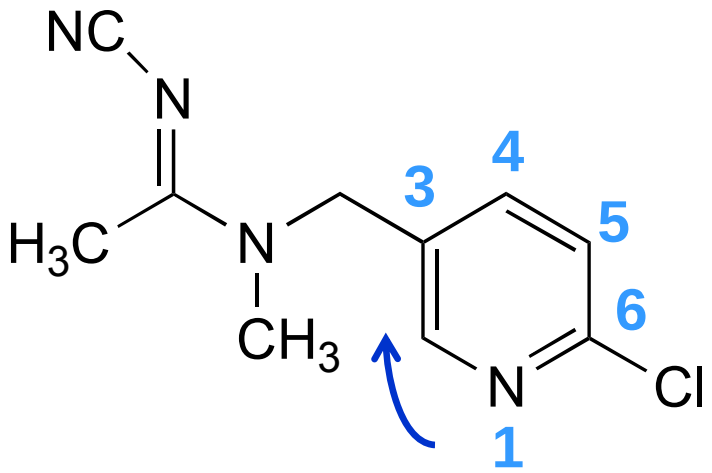
(*E*)-*N*-[(6-chloro-3-pyridinyl)methyl]-*N*′-cyano-*N*-methyl-ethanamidine (CA; *135410-20-7*)

$C_{10}H_{11}ClN_4$

Resolution:

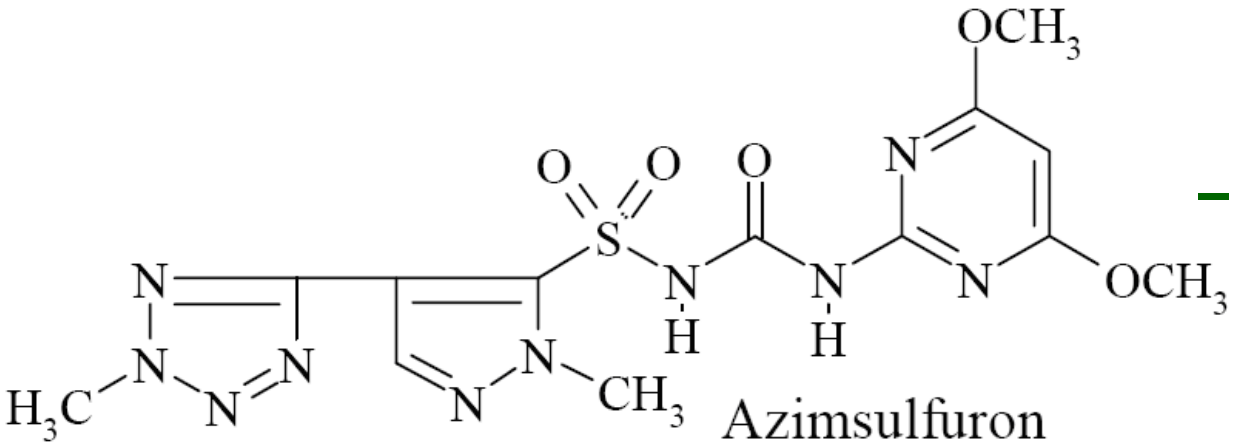
amidine → imidamide

Missing N



Due to clockwise numbering, N-atom not at the top

- PIN: (1*E*)-*N*-[(6-chloropyridin-3-yl)methyl]-*N*′-cyano-*N*-methylethanimidamide
- CA: (1*E*)-*N*-[(6-chloro-3-pyridinyl)methyl]-*N*′-cyano-*N*-methylethanimidamide (RN 160430-64-8)



Azimsulfuron

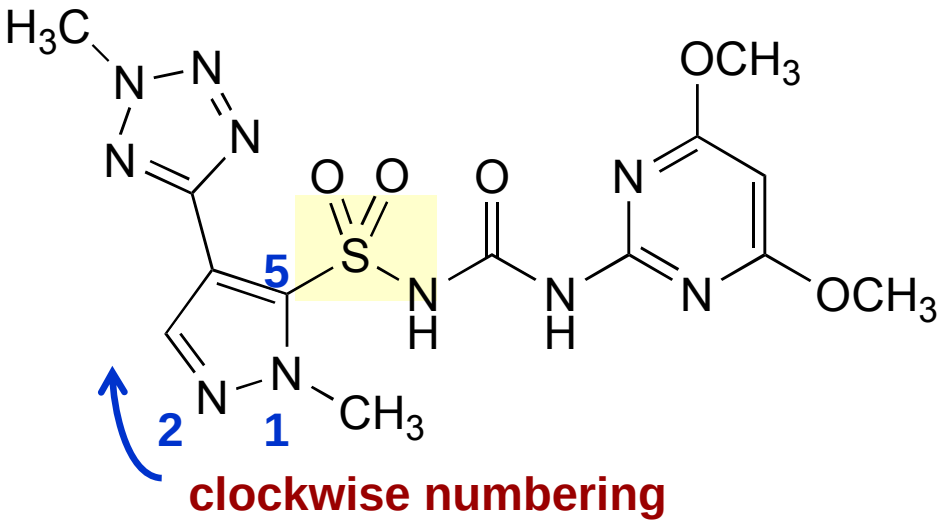
– Handbook L, p. 24 ff. -

Azimsulfuron

1-(4,6-dimethoxypyrimidin-2-yl)-3-[1-methyl-4-(2-methyl-2*H*-tetrazol-5-yl)pyrazole-5-ylsulfonyl]urea (IUPAC);
N-[[(4,6-dimethoxy-2-pyrimidinyl)amino]carbonyl]-1-methyl-4-(2-methyl-2*H*-tetrazol-5-yl)-1*H*-pyrazole-5-sulfonamide (CA; *120162-55-2*)

Resolution:

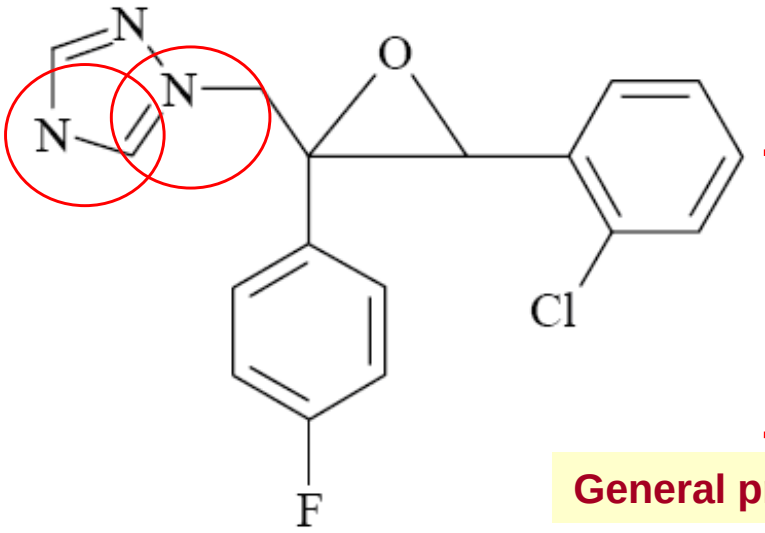
In IUPAC nomenclature, urea is not the principal group.
Principal group is the sulfonamide.



- PIN:
  N-[(4,6-dimethoxypyrimidin-2-yl)carbamoyl]-1-methyl-4-(2-methyl-2*H*-tetrazol-5-yl)-1*H*-pyrazole-5-sulfonamide
- CA:
  name ok

CIPAC Doc Checks on IUPAC/CA '2010' Accuracy

- Epoxiconazol doc. – Handbook K, page 38 ff. -

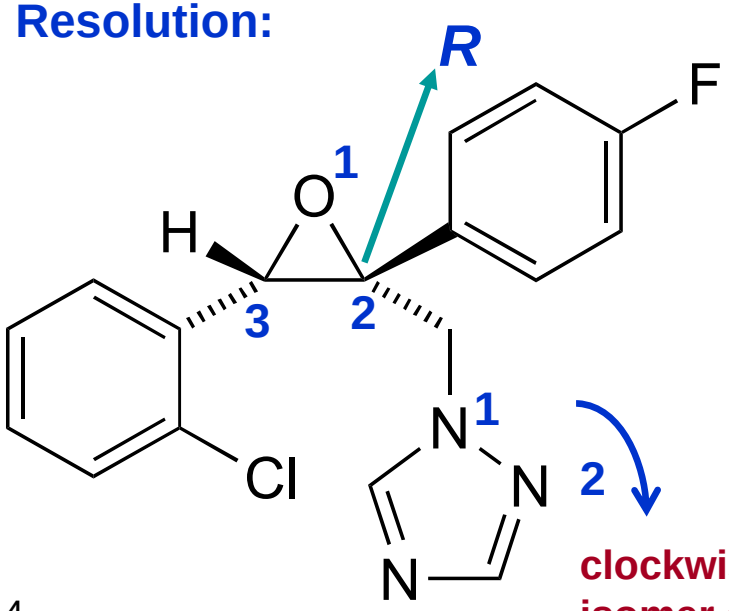


Epoxiconazole
(2*RS*, 3*SR*)-1-[3-(2-chlorophenyl)-2,3-epoxy-2-(4-fluorophenyl)propyl]-1*H*-1,2,4-triazole (IUPAC);
rel-1-[[(2*R*, 3*S*)-3-(2-chlorophenyl)-2-(4-fluorophenyl)-oxiranyl]methyl-1*H*-1,2,4-triazole (CA; *135319-73-2*)

General problems with CA RN changes

epoxy → oxiran-2-yl

Resolution:



clockwise numbering; display of the (*R*)-isomer of the lowest numbered C-atom

- PIN:
  rac-1-{[(2*R*,3*S*)-3-(2-chlorophenyl)-2-(4-fluorophenyl)oxiran-2-yl]methyl}-1*H*-1,2,4-triazole
- CA:
  rel-1-[[(2*R*,3*S*)-3-(2-chlorophenyl)-2-(4-fluorophenyl)-2-oxiranyl]methyl]-1*H*-1,2,4-triazole (RN 133855-98-8)

IUPAC Name Generation by Using Autonom 2000 and ACD/Name Vs.*11.0*

Page	Ai. code	ISO Name	Correctness by using:	
			Autonom 2000	ACD/ Name
4	649	acetamiprid	no result	
23	584	azimsulfuron	no result	
38	750	*S*-bioallethrin	no result	
45	333	deltamethrin	no result	
61	734	diflovidazin		
67	749	dinotefuran	no result	
104	174	picloram		
38	609	epoxiconazol		
77	568	kresoxim-methyl	no result	

OK | Slight discrepancies | Strong discrepancies

Software Tools for Name Generation

Autonom 2000 **- Software from Beilstein**
(e.g., embedded in ISIS Draw)

- **Output: IUPAC names (should be)**
- **Useful for smaller molecules and for training**
- **Insufficient for more complex structures and functional groups**
- **In many cases, not IUPAC-2010 state of the art**

ACD/Name - Software from ACD / Labs
(Advanced Chemistry Development, Inc.)

- **Very close to IUPAC 2010**
- **Good results even for complex molecules**
- **Quite expensive**
- **Freeware version ACD/Sketch (limited to 3 rings and up to 50 atoms)**

More References about IUPAC Names

- Selected Links -

About IUPAC Nomenclature:

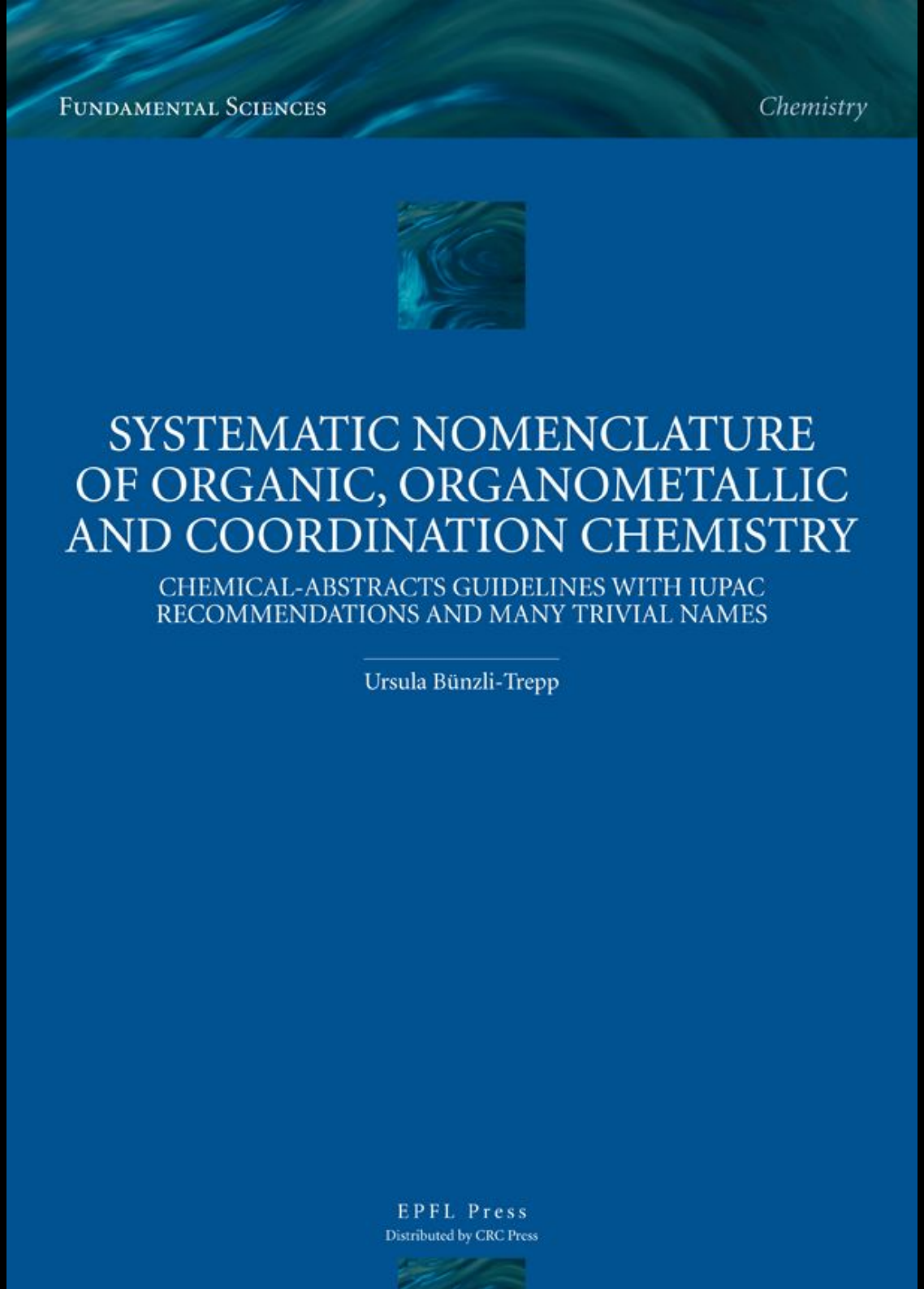
- **IUPAC Nomenclature of Organic Chemistry (Blue Book 1979), incomplete version at: http://www.acdlabs.com/iupac/nomenclature/**
- **Nomenclature of Organic Compounds, Guide 1993 (translated in German, e.g.) at: http://www.acdlabs.com/iupac/nomenclature/**
- **Nomenclature of Organic Chemistry 2004 (Preferred IUPAC Names) at: http://old.iupac.org/reports/provisional/abstract04/favre_310305.html**
- **IUPAC's Nomenclature home page, especially useful for natural products and some PINs at: http://www.chem.qmul.ac.uk/iupac/**

About CA Nomenclature:

- **No printed Chemical-Abstracts material available since 01.01.2010, see http://www.cas.org/products/print/index.html**

- **Naming and Indexing of Chemical Substances for Chemical Abstracts™, 2007 Edition, available since 2009 at: http://www.cas.org/ASSETS/58D34DD3892142D18F5C3B0A004D3A0C/indexguideapp.pdf**

- **Link to *CAS Registry Numbers*: http://www.cas.org/expertise/cascontent/registry/regsys.html**



Ref.:

Recommended book on Org. Chem. Nomenclature (2007) – 638 pages

Author:
Dr. Ursula Bünzli-Trepp, Member of IUPAC's Chemical Nomenclature Advisory Subcommittee

Visit, e.g.,

www.chemical-nomenclature.ch

IUPAC, Systematic Nomenclature for CIPAC Documentation

Summary:

- In the last 20 years, IUPAC has revised many nomenclature recommendations.
- Thereby, IUPAC names became more and more systematic, and many trivial names and names for ‘general use’ were withdrawn.
- Changes of IUPAC recommendations concern also chemical structure drawings, especially those of heterocycles.
- These changes have been only partly implemented into our CIPAC documentations, and roughly more than 80% of CIPAC docs are not compliant.

Thank you for your attention

Comments and questions for discussion are most welcome