A Roadmap for Systematic Data Representations in Cyberinfrastructure for Combustion Research

Donald R. Burgess Jr. and Jeffrey A. Manion

Chemical and Biochemical Reference Data Division
National Institute of Standards and Technology
Gaithersburg, Maryland

Multiagency Coordination Committee for Combustion Research (MACCR)
4th Annual Fuels Research Review

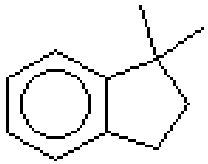
Argonne National Laboratory
Argonne, Illinois
20-22 September 2011

Information Required/Suggested in General for Chemical Kinetic Mechanisms

- Full chemical kinetic mechanisms must be made available
  - Supplementary information (does not kill trees)
- All values must be referenced
- All references ideally should be original sources
  - NIST Chemistry Webbook is NOT an acceptable reference
  - NIST Chemical Kinetics Database is NOT an acceptable reference
- All values should ideally have pedigree/datatype/method/quality
  - evaluated, experimental, ab initio, estimated, don't-know

Information Required/Suggested for Chemical Species

- Molecular formula must be provided
  - C4H5O2
- Symbol Definitions and Systematic Nomenclature should be used
  - aC3H5 allyl propen3yl
- Molecular geometries (coordinates) must be provided
  - Z-matrix, Cartesian coordinates, MDL mol files
- Molecular geometries (images) are suggested where helpful
  - 1,1-dimethyl-2,3-dihydro-1H-indene



- InChI strings ideally should be provided for molecules
  - C11H14/**c1-11(2)8-7-9-5-3-4-6-10(9)11**/h3-6H,7-8H2,1-2H3
- Chemical Abstract Registry Numbers (CASRN) are suggested
  - 4912-92-9
- Organization of molecules using Chemical Classes are suggested
  - aromatic/bicyclo/indane

Information Required / Suggested for Chemical Reactions

- All rate data ideally should be have a relevant temperature range
  - 250-425 K versus 1200-1800 K
- InChI strings ideally should be provided for transition states
  - CH3CH2Cl → C2H4 + HCl C2H5Cl/c1-2-4-3-1/h1-2H2/
- InChI strings ideally should be provided for reactions
  - CH3CH2O• + CH3• → CH3CH(=O) + CH4
  - C2H5O.CH3/c1-2(3)4;/h2H,1H3;1H3/r2-4+1'
- All reactions ideally should be reaction classified
  - Propane + C2H3 → Prop2yl + C2H4 C-H+C [CH2/CdH]
  - C2H5 + H → C2H4 + H2 C.C-H+H
  - C2H5 + H → [C2H6]* → C2H4 + H2 C-H//CC-HH

Species Symbol Definitions and Systematic Nomenclature

Trivial Symbol	Definition/Structural	Systematic Symbol
aC3H5	allyl	propen3yl
CH3CCH2	CH3-*C=CH2	propen2yl
CH3CHCH	CH3-CH=CH*	propen1yl
nC3H7	CH3-CH2-CH2*	prop1yl
iC3H7	CH3-*CH-CH3	prop2yl
nC4H3	*CH=CH-CCH	but1yn3en4yl
iC4H3	CH2C*-CCH	but1yn3en3yl
nC4H5	*CH=CH-CH=CH2	buta13dien1yl
iC4H5	CH2=C*-CH=CH2	buta13dien2yl
pC4H9	1-butyl	but1yl
sC4H9	2-butyl	but2yl
C4H81	1-butene	but1ene

Species Symbol Definitions and Systematic Nomenclature

Trivial Symbol	Definition/Structural	Systematic Symbol
C4H82	2-butene	zbut2ene, ebut2ene
iC4H8	i-butene	isobutene
C4H5-2	CH3-CC-*CH2	but2yn1yl
C4H7	*CH2-CH2-CH=CH2	but1en4yl
CH2CHCO	CH2=CH-*CO	prop1en3oyl
CH2CHCHCHO	*CH2-CH=CH-CHO	but2en4oyl
CH3CHCHCO	CH3-CH=CH-*CO	but2en1oyl
C6H3	*CH=CH-CC-CCH	hex13diyn5en6yl
C4H6	1,3-butadiene	buta13diene
C4H612	1,2-butadiene	buta12diene
C4H6-2	2-butyne	but2yne
C2H3CHO	acrolein	prop2enal

Species Symbol Definitions and Systematic Nomenclature

Trivial Symbol	Definition/Structural	Systematic Symbol
l-C6H4	CH2=CH-CC-CCH	hexa13diyn5ene
PXC12H25	1-dodecyl	dodec1yl
SXC12H25	2-dodecyl	dodec2yl
S2XC12H25	3-dodecyl	dodec3yl
C12H24	1-dodecene	dodec1ene
PXC12H23	1-dodecen-12-yl	dodec1ene12yl
C6H12	1-hexene	hex1ene
PXC6H11	1-hexen-6-yl	hex1en6yl
SXC6H11	1-hexen-5-yl	hex1en5yl
SAXC6H11	1-hexen-3-yl	hex1en3yl
SAX4-2C6H11	4-hexen-3-yl	hex4en3yl

Species Symbol Definitions and Systematic Nomenclature

Trivial Symbol	Definition/Structural	Systematic Symbol
C6H10-13	1,3-hexadiene	hexa13diene
CH3cC5H83	1-methyl-cyclopent-3-yl	cypent3yl-1me
S2XC5H11	3-pentyl	pent3yl
SAXC5H9	1-penten-3-yl	pent1en3yl
C5H8-13	1,3-pentadiene	penta13diene
cC5H9	cyclopentyl	cypentyl
cC5H8	cyclopentene	cypentene

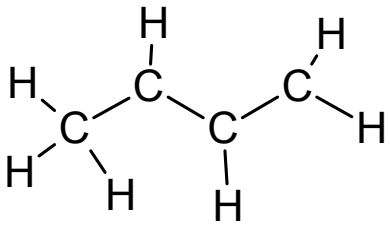
IUPAC International Chemical Identifier (InChI) is a “structural” identifier, not a “chemical” identifier

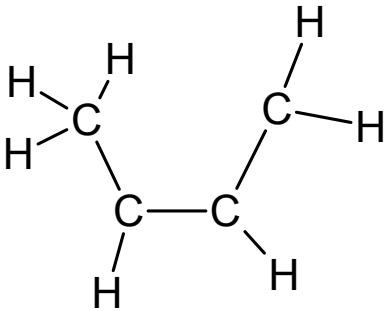
Molecular Formula: C4H7
Line Formula: CH3CHCHCH2
Trivial Name: 1-methylallyl
InChI string: C4H7/c1-3-4-2/h3-4H,1H2,2H3

<u>Two “structural” entities</u>

(*E*)-$CH_3CHCHCH_2$

(*Z*)-$CH_3CHCHCH2$





<u>Two “chemical” entities</u>

CH_3-CH=CH-CH_2(*) … terminal “group” is C*H2-Cd

CH_3-C(*)H-CH=CH_2 … terminal “group” is CdH2

IUPAC International Chemical Identifier (InChI) does not discriminate between important sp3 conformers

It is known that some chemical reactions are very dependent upon the conformation of the reactants (and of the transition states). The reactivities of a number of classes of compounds with OH have been determined to be dependent upon rotational conformations and ring conformations

- fluoroaldehydes
- fluoroformates
- chlorofluorocycloalkanes
- cyclic terpenes

(*E*)- and (*Z*)-1,2-dichloroethene are differentiated

➢ C2H2Cl2/c3-1-2-4/h1-2H/b2-1+(**b2-1-**)

(*anti*)- and (*gauche*)-1,2-dichloroethane are NOT differentiated

➢ C2H4Cl2/c3-1-2-4/h1-2H2

endo, exo, equatorial, axial ring conformations NOT differentiated

Proposed Use of InChI-ER (modified) to Uniquely Identify Reactions

(1) $CH_3CH_3 \rightarrow CH_3CH_2 + H$ — **bond fission**
C2H6/c1-2/h1-2H3 /r1-3

(2) $CH_3CH_3 \rightarrow CH_3 + CH_3$ — **bond fission**
C2H6/c1-2/h1-2H3 /r1-2

(3) $CH_3CH_3 \rightarrow CH_2{=}CH_2 + H_2$ — **1,2-elimination**
C2H6/c1-2/h1-2H3 /r1,2-3,4

(4) hexane → $CH_2{=}CH_2$ + butane — **"1,2-elimination"**
C6H14/c1-3-5-6-4-2/h3-6H2,1-2H3 /r5,6-3,4

(5) cyclohexene → $CH_2{=}CH_2$ + 1,3-butadiene — **retro Diels-Alder**
C6H10/c1-2-4-6-5-3-1/h1-2H,3-6H2 /r5-3,6-4

Proposed Use of InChI (modified) to Uniquely Identify Reactions

(6) $CH_3CH_3 + CH_3 \rightarrow CH_3CH_2 + CH_4$ **abstraction**
C2H6.CH3/c1-2;/h1-2H3;1H3 /r1-3+1'

(7) $CH_3CH_3 + OH \rightarrow CH_3CH_2 + H_2O$ **abstraction**
C2H6.HO/c1-2;/h1-2H3;1H /r1-3+1'

(8) $CH_3CH_2CH_2^* \rightarrow CH_3C^*HCH_3$ **H-atom transfer**
C3H7/c1-3-2/h1,3H2,2H3 /r2-4+1

(9) (*Z*)- CHCl=CHCl → (*E*)- CHCl=CHCl **(E),(Z) isomerization**
C2H2Cl2/c3-1-2-4/h1-2H/b2-1- /r1~2

Proposed Use of InChI (modified) to Uniquely Identify Reactions

Unimolecular Reactions: Radical Eliminations

No	Reaction Class	Reaction Form			Reaction Example	quasi-InChI / r-Layer	
1a	Bond Fission	R–X	→	R• + X•	$CH_3CH_3 \rightarrow CH_3\bullet + CH_3\bullet$	C2H6/c1-2/h1-2H3	/r1-2
					$CH_3CH_3 \rightarrow CH_3CH_2 + H\bullet$	C2H6/c1-2-3/h2H2,1H3	/r2-3
1b	Beta Fission	:RS–X	→	•R=S + X•	$CH_3CH: \rightarrow CH_2{=}CH\bullet + H\bullet$	C2H4/c1-2-3/h1H,2H2	/r2-3
2a	Beta Scission	•RS–X	→	R=S + X•	$CH_3CH_2O\bullet \rightarrow CH_2{=}O + CH_3\bullet$	C2H5O/c1-2-3/h2H2,1H3	/r1-2
					$CH_3CH_2O\bullet \rightarrow CH_3CH({=}O) + H\bullet$	C2H5O/c1-2(3)4/h2H,1H3	/r2-4
					$CH_3CH_2\bullet \rightarrow CH_2{=}CH_2 + H\bullet$	C2H5/c1-2-3/h1-2H2	/r2-3
2b	Cyclization (Internal Substitution)	•RS–X	→	-(RS)- + X•	$\bullet CH_2CH_2CH_2OOH \rightarrow -(CH_2CH_2CH_2O)- + \bullet OH$	C3H7O2/c1-2-3-5-4/h4H,1-3H2	/r1+5-4

Proposed Use of InChI (modified) to Uniquely Identify Reactions

Unimolecular Reactions: Molecular Eliminations

No	Rxn Class	Reaction Form			Reaction Example	quasi-InChI	/r-Layer
3a	**1,2-Elimination**	R(X)S(Y)	→	R=S + XY	$C_2H_5OH \rightarrow CH_2{=}CH_2 + H_2O$	C2H6O/c3-1-2-4/h3H,1-2H2	/r1,2-3,4 or /r1-3,2-4,3+4
					$CH_3CH_3 \rightarrow CH_2{=}CH_2 + H_2$	C2H4F2/c3-1-2-4/h1-2H2	/r1,2-3,4
3b	**1,1-Elimination**	R(X)(Y)	→	R: + XY	$CH_3CH_3 \rightarrow CH_3CH: + H_2$	C2H6/c1-2(3)4/h2H,1H3	/r2-3,4 or r2-3,2-4,3+4
3d	**1,2-Elimination (radical)**	R(X)S•(Y)	→	R=S• + XY	$\bullet CF_2OH \rightarrow \bullet CF(=O) + HF$	CHF2O/c2-1(3)4/h4H	/r1,2-3,5 or /r1-3,2-5,3+5
3c	**1,1-Elimination (radical)**	•RS(X)(Y)	→	•R=S + XY	$CHF_2O\bullet \rightarrow \bullet CF(=O) + HF$	CHF2O/c2-1(3,4)5	/r1-3,5
4a	**Cyclo-elimination**	R(X)S(Y)	→	R=S + X=Y	Cyclohexene → 1,3-Butadiene + C_2H_4	C6H10/c1-2-4-6-5-3-1/h1-2H,3-6H2	/r3-5,4-6
4b	**Retroene-elimination**	R(T=X)S(Y)	→	R=S + T=XY	1-Pentene → Propene + C_2H_4	C5H10/c1-2-3-4-5-6/h2H,1,3-5H2	/r3-4,5-6,1+4

Proposed Use of InChI (modified) to Uniquely Identify Reactions

No	Reaction Class	Reaction Example	quasi-InChI	/r-Layer
	ISOMERIZATIONS			
5b	**1,n Atom Transfer**	$\bullet CH_2CH_2OH \rightarrow CH_3CH_2O\bullet$	C2H5O/c1-2-3-4/h1-2H2	/r3-4+1
		$\bullet CH_2CH_2CH_3 \rightarrow CH_3C\bullet HCH_3$	C3H7/c1-3(2)4/h3H,1H2,2H3	/r3-4+1
5c	**1,2 Atom Tsf (ylidene)**	$CH_3CH: \rightarrow CH_2{=}CH_2$	C2H4/c1-2-3/h1H,2H2	/r2-3+1
6a	**Torsional Isomerization**	(*Z*)-CHCl=CHCl → (*E*)-CHCl=CHCl	C2H2Cl2/c3-1-2-4/h1-2H/b2-1+	/r2~1
		(*gauche*)-$CH_2Cl{=}CH_2Cl$ → (*anti*)-$CH_2Cl{=}CH_2Cl$	C2H4Cl2/c3-1-2-4/h1-2H2/b2-1-+	/r2~1-+
6b	**Bond Migration**	$CH_2{=}CH{-}CH_2CH_3$ → $CH_3{-}CH{=}CH{-}CH_3$	C4H8/c1-3-4(2)5/h3-4H,1H2,2H3	/r4-5+1
6c	**Structural Isomerization**	Methylenecyclopentadiene → Benzene	C6H6/c7-4-6-3-1-2-5-6/h1-5H	/r5-6+4;4-7+6
6d	**Bridged Isomerization**	$Si(H_2)Si \rightarrow :SiSiH_2$	H2Si2/c1-2(3)4	/r1+3,4
7a	**1,5-Sigmatropic Shift**	1,3-Hexadiene → 2,4-Hexadiene	C6H10/c1-3-4-5-6(2)7/h3-6H,1H2,2H3	/r6-7+1
7b	**3,3-Sigmatropic Rearrangement**	$CH_2{=}CH{-}O{-}CH_2CH{=}CH_2$ → $CH_2{=}CHCH_2CH_2CH(=O)$	C5H8O/c1-3-5-6-4-2/h3-4H,1-2,5H2	/r1+2,5-6

Proposed Use of InChI (modified) to Uniquely Identify Reactions

No	Reaction Class	Reaction Example	quasi-InChI	/r-Layer
	BIMOLECULAR			
8	**Abstraction**	$CH_3CH_3 + \bullet OH \rightarrow C_2H_5\bullet + H_2O$	C2H6.HO/c1-2-3;/h2H2,1H3;1H	/r2-3+1’
9	**Disproportionation**	$CH_3CH_2O\bullet + CH_3\bullet$ $\rightarrow CH_3CH(=O) + CH_4$	C2H5O.CH3/c1-2(3)4;/h2H,1H3;1H3	/r2-4+1’
10	**Substitution**	$CH_3OOH + CH_3\bullet \rightarrow CH_3OCH_3 + \bullet OH$	CH4O2.CH3/c1-3-2;/h2H,1H3;1H3	/r3+1’-2
		$\bullet CH_2CH_2CH_2OOH$ $\rightarrow -(CH_2CH_2CH_2O)- + \bullet OH$	C3H7O2/c1-2-3-5-4/h4H,1-3H2	/r1+5-4
	COMPOSITE			
1a/1a	**Combination/Bond Fission**	$C_2H_5O\bullet + H\bullet \rightarrow C_2H_5\bullet + \bullet OH$	C2H6O/c1-2-3-4/h2H2,1H3	/r3-4:3-2
1b/1b	**Addition/Beta Scission**	$CH_2{=}CH(OH) + CH_3\bullet$ $\rightarrow CH_2{=}CH{-}CH_3 + \bullet OH$	C3H7O/c1-3(2)4/h3-4H,1H2,2H3	/r3-4:3-2

Proposed Standard Notation Classification Scheme for Elementary Chemical Reactions

No	Class	Reactants		Products	Notation
1a	**Bond Fission**	R–X	→	R• + X•	R-X
1b	**Beta Fission**	:RS–X	→	•R=S + X•	R:S-X
2a	**Beta Scission**	•RS–X	→	R=S + X•	R.S-X
3a	**1,2-Elimination**	R(X)S(Y)	→	R=S + XY	RS-XY
3b	**1,1-Elimination**	R(X)(Y)	→	R: + XY	R-XY
3c	**1,1-Elimination (radical)**	•RS(X)(Y)	→	•R=S + XY	S-XY
3d	**1,2-Elimination (radical)**	R(X)S•(Y)	→	R=S• + XY	-XY
4a	**Cycloelimination**	R(X)S(Y)	→	R=S + X=Y	RS-XY(n)
4b	**Retroene-elimination**	R(T=X)S(Y)	→	R=S + T=XY	RS-XY\Z(n)

Proposed Standard Notation Classification Scheme for Elementary Chemical Reactions

5a	**1,n Atom Transfer**	•RS–X	→	X–RS•	S\X(n)
5b	**1,2 Atom Transfer**	•RS–X	→	X–RS•	S\X(2)
5c	**1,2 Atom Transfer (ylidene)**	:RS–X	→	X–R=S	R:S\X
6a	**Torsional Isomerization**	X/RS\Y	→	X/RS/Y	RS/X\Y
6b	**Bond Migration**	RS–X	→	X–RS	RS\X
6c	**Structural Isomerization**	X-R=S-Y	→	Y-R=S-X	RS\X/Y
6d	**Bridged Isomerization**	R(XY)S	→	RS(X)(Y)	RS\X\Y
7a	**1,5-Sigmatropic Shift**	R=TS–X	→	X–RT=S	RS\X(n)`
7b	**3,3-Sigmatropic Rearrangement**	R=TU=S	→	T=RS=U	RS\XY(n)
8	**Abstraction**	R–X + Y•	→	R• + XY	R-X+Y
9	**Disproportionation**	•RS–X + Y•	→	R=S + XY	S-X+Y
10a	**Substitution**	R–X + Y•	→	R–Y + X•	R+X-Y
10b	**Displacement**	RX + SY	→	RS + XY	R+X:-Y

Proposed Standard Notation Classification Scheme for Elementary Chemical Reactions

1/1	**Combination/Bond Fission**	R• + X•	→	S• + Y•	R-X//S-Y
2/2	**Addition/Beta Scission**	R=S–Y + X	→	R=S–X + Y	S-Y

Proposed Standard Notation Classification Scheme for Elementary Chemical Reactions

UNIMOLECULAR RADICAL ELIMINATIONS			
1a	Bond Fission	C–H	$C_2H_6 \rightarrow C_2H_5\bullet + H$
		C–O	$CH_3OCH_3 \rightarrow CH_3\bullet + CH_3O\bullet$
1b	Beta Fission	C:C–H	$CH_3CH: \rightarrow CH_2{=}CH\bullet + H$
2a	Beta Scission	C.C–H	$C_2H_5\bullet \rightarrow C_2H_4 + H$
		O.C–C	$CH_3CH_2O\bullet \rightarrow CH_2{=}O + CH_3\bullet$

Proposed Standard Notation Classification Scheme for Elementary Chemical Reactions

UNIMOLECULAR MOLECULE ELIMINATIONS			
3a	1,2-Elimination	CC–HH	$C_2H_6 \rightarrow C_2H_4 + H_2$
		CC–OH	$C_2H_5OH \rightarrow C_2H_4 + H_2O$
		CC–CH	Hexane → Propene + Ethane
		CC–CH	Cyclohexane → 1-Hexene
3b	1,1-Elimination	C–HH	$CH_3CH_3 \rightarrow CH_3CH: + H_2$
3c	1,1-Elimination (radical)	O.C–HF	CHF2O• → •CF(=O) + HF
3d	1,2-Elimination (radical)	OC.–HF	•CF2OH → •CF(=O) + HF
4a	Cycloelimination	CC–CC(4)	Cyclohexene → 1,3-Butadiene + C_2H_4
		CC–CC(2)	1,2-Dimethylcyclobutane → 2-Butene + C_2H_4
4b	Retroene-elimination	CC–CC\H(4)	1-Pentene → Propene + C_2H_4

Proposed Standard Notation Classification Scheme for Elementary Chemical Reactions

UNIMOLECULAR ISOMERIZATIONS			
5a	1,n Atom Transfer	C.C\H(3)	$CH_3CH{=}CH\bullet \rightarrow \bullet CH_2CH{=}CH_2$
5b	1,2 Atom Transfer	C.C\H(2)	$\bullet CH_2CH_2D \rightarrow CH_3CHD\bullet$
		C.O\H(2)	$\bullet CH_2CH_2OH \rightarrow CH_3CH_2O\bullet$
5c	1,2 Atom Transfer (ylidene)	C:C\H	$CH_3CH: \rightarrow CH_2{=}CH_2$
6a	Torsional Isomerization	CC/Cl\Cl	$CH_2Cl\text{-}CH_2Cl$[*gauche*] → $CH_2Cl\text{-}CH_2Cl$[*anti*]
		CC/Cl\Cl	(*Z*)-CHCl=CHCl → (*E*)-CHCl=CHCl
6b	Bond Migration	CC\H	$CH_2{=}CH{-}CH_2CH_3 \rightarrow CH_3\text{-}CH{=}CH{-}CH_3$
6c	Structural Isomerization	CC\C/H	$CH_3\text{-}CH{=}CH{-}CH_3 \rightarrow CH_2{=}C(CH_3)_2$
		CC\C/H	Benzene → Methylenecyclopentadiene
6d	Bridged Isomerization	SiSi\H\H	$Si(H_2)Si \rightarrow :SiSiH_2$
7a	1,5-Sigmatropic Shift	CC\H(5)	1,3-Hexadiene → 2,4-Hexadiene
7b	3,3-Sigmatropic Rearrangement	CC\CC(3)	1,5-Heptadiene → 3-Methyl-1,5-Hexadiene

Proposed Standard Notation Classification Scheme for Elementary Chemical Reactions

BIMOLECULAR REACTIONS			
8	Abstraction	C–H+H	$C_2H_6 + H \rightarrow C_2H_5\bullet + H_2$
		C–H+O	$C_2H_6 + OH \rightarrow C_2H_5\bullet + H_2O$
9	Disproportionation	C–H+H	$C_2H_5\bullet + H \rightarrow C_2H_4 + H_2$
10a	Substitution	O+C–O	$CH_3OOH + CH_3 \rightarrow CH_3OCH_3 + OH$
10b	Displacement	Al+O:–C	$Al(CH_3)_3 + H_2O \rightarrow Al(OH)(CH_3)_2 + CH_4$
COMPOSITE REACTIONS			
1/1	Combination/Bond Fission	C–H//C–C	$CH_3CH_2\bullet + H \rightarrow CH_3\bullet + CH_3\bullet$
		C–H//C–O	$C_2H_5O\bullet + H \rightarrow C_2H_5\bullet + OH$
2/2	Addition/Beta Scission	C.C-H//C.C-C	$CH_2{=}CH{-}CH_3 + H$ $\rightarrow CH_2{=}CH_2 + CH_3\bullet$
		C.C-C//C.C-O	$CH_2{=}CH(OH) + CH_3\bullet$ $\rightarrow CH_2{=}CH{-}CH_3 + OH$