

INCOMPLETE DRAFT

DOCUMENTATION OF THE SAPRC-16 MECHANISM GENERATION SYSTEM

Interim Report to California Air Resources Board Contract No. 11-761

William P. L. Carter

May 25, 2019

College of Engineering Center for Environmental Research and Technology (CE-CERT)
University of California, Riverside, California 92521

SUMMARY

This document gives a description of the automated chemical mechanism generation system that is used to derive gas-phase atmospheric reaction mechanisms for reactive organic compounds for use in atmospheric chemical models. In particular, this system is used to derive the SAPRC-16 that is being developed. The system, called MechGen uses a computer program to derive fully explicit mechanisms for the reactions of many types of emitted organic compounds and their oxidation products when they react in the atmosphere in the presence of oxides of nitrogen and other pollutants, and then uses the results to derive lumped mechanisms suitable for use in atmospheric models. This report gives an overview of system, describes the procedures used to generate detailed explicit mechanisms, and the procedures used to derive lumped reactions for the SAPRC-16 chemical mechanism for atmospheric modeling. The scientific basis for the generation of the detailed mechanisms, including the assumptions, approximations, and assignments or estimations of the rate constants or branching ratios involved, are described in a separate document (Carter, 2019a). The system runs on a programming platform originally developed for text-based and is accessible by anyone over the internet using a web-based interface, with a telnet login being used to program and manage the system and input assignments. The web-based interface can be used to generate complete mechanisms for selected compounds or to examine individual reactions to see documentation output concerning how individual reactions and rate constants are derived. It is currently available at <http://mechgen.cert.ucr.edu>.

ACKNOWLEDGEMENTS

This work was supported in part by the California Air Resources Board primarily through contract no 11-761 and in part by the University of California Retirement system. The author wishes to thank Dr. Ajith Kaduwela, the CARB project officer, for his support and helpful discussions. The opinions and conclusions in this report are entirely those of the author.

CONTENTS

1.	Introduction.....	1
1.1.	Background.....	1
1.2.	Overview.....	2
2.	Programming Platform and Major Components.....	5
2.1.	User Interfaces.....	7
2.1.1.	Telnet Access.....	7
2.1.2.	Online access.....	7
3.	Operation of System.....	9
3.1.	Specification of Reactants.....	9
3.2.	Reaction Generation.....	13
3.2.1.	Single Step Reactions.....	13
3.2.2.	Full Mechanism Generation.....	14
3.3.	Processing Mechanisms for SAPRC-16.....	18
3.3.1.	Process Lumping for Compounds.....	18
3.3.2.	Mechanism Condensation.....	21
3.3.3.	Deriving Mechanisms for Mixtures.....	24
3.4.	(Additional discussion to be added).....	26
4.	Programming and Operational Details.....	27
4.1.	(Discussion of Installation and Initial Setup to be added).....	27
4.2.	(Discussion of programming and entering assignments using Telnet to be added).....	27
4.3.	Operations using the Telnet Interface.....	27
4.3.1.	Creating Reactants.....	27
4.3.2.	(discussion of Additional Telnet Commands to be added).....	27
4.4.	Operations Using the Web Interface.....	27
4.4.1.	Creating Reactants.....	27
4.4.2.	Running Single Step Reactions.....	28
4.4.3.	(discussion of Additional Web Commands to be added).....	28
5.	References.....	29
6.	Appendix A. List of MOO objects.....	31

LIST OF TABLES

Table 1.	Types of stable compounds whose reactions are supported by MechGen.....	3
Table 2.	Summary of types of reactions supported by the mechanism generation system. "VOC" indicates any stable compound supported by the system (see Table 1).....	4
Table 3.	List of groups and group designations used to specify organic reactants.....	10
Table 4.	List of elementary or single-group species names recognized by the system.....	11
Table 5.	Examples of designations of selected representative compounds and radicals.....	12
Table 6.	Classification of products formed in explicit mechanism generations.....	19
Table 7.	List of MOO objects used in the MechGen system, indicating numbers of programs and properties associated with each and the total numbers of lines of codes in the programs.....	31
Table 8.	List of MOO objects used to specify individual types of radical or non-radical groups.....	33
Table 9.	List of MOO objects used to specify individual types of non radical product compounds or counter species.....	40

LIST OF FIGURES

Figure 1.	Flow diagram of the process of generating mechanisms for emitted and selected organic product compounds.....	15
Figure 2.	Flow diagram of process lumping for generated mechanisms.....	20
Figure 3.	Flow diagram of the process of estimating upper limit intermediate and product yields.....	23
Figure 4.	Flow diagram showing how a processed mechanism is used to derive lumped reactions can be included in the SAPRC-16 mechanism.....	24
Figure 5.	Flow diagram showing how mechanisms for mixtures are initially processed.....	26

1. INTRODUCTION

1.1. Background

When most volatile organic compounds (VOCs) are introduced into the atmosphere in the presence of sunlight, they can react in various ways to form various radicals that react further to form oxidized organic products, including gas-phase toxics and secondary organic aerosol (SOA). In the presence of oxides of nitrogen (NO_x) emitted from combustion sources these radicals also promote the formation of ozone and oxidized nitrogen compounds that affect air quality. In order to assess control strategies to determine how best to reduce ozone, SOA, and other toxic secondary pollutants it is necessary to use air quality models that include chemical mechanisms that predict how the formation of these secondary pollutants depend on emissions of VOCs and NO_x , and how they may change if the compositions of the emitted VOCs change due to regulations or other economic or policy changes. For these predictions to be reliable the chemical mechanisms used need to be based on the best available science, and appropriately represent the actual reactions that are occurring.

Many hundreds of types of organic compounds are emitted into the atmosphere, from both anthropogenic and biogenic sources. The atmospheric reaction mechanisms for most these compounds are complex, and for larger molecules can involve a large number of reactive intermediates and form a large number of oxidized organic products. Most of these organic products also react in the atmosphere, forming even more intermediates and products. In addition, in almost all cases these mechanisms involve reactions whose rate constants are unknown and have to be estimated. Because of the complexity, for practical reasons it is necessary either to greatly simplify the mechanisms for most VOCs, use extensive lumping or condensations in VOC representations, or use an automated chemical mechanism generation system to generate the mechanisms. In the case of atmospheric reaction mechanisms of VOCs, the existing chemical mechanism systems include the GECKO-A system developed by Aumont and co-workers (Aumont et al, 2005) and the SAPRC mechanism generation system that is the subject of this document.

The SAPRC mechanism generation system was developed by the author for deriving portions of the SAPRC-90 through SAPRC-07 atmospheric chemical mechanisms (Carter, 1990, 2000, 2010a,b). The first version, used for SAPRC-90 (Carter, 1990) was written in FORTRAN and generated mechanisms only for alkanes, using the procedures and estimates documented by Carter and Atkinson (1985). When SAPRC-99 was being developed the mechanism generation system was moved to another programming platform that was more suitable for this application (see below), was completely re-written, and was extended to cover a much wider range of acyclic and monocyclic compounds, including monoalkenes, alcohols, ethers, esters, aldehydes, ketones and organic nitrates in addition to alkanes (Carter, 2000). This version was used to determine the net effects of these compounds in the presence of NO_x that was incorporated into SAPRC-99, though reactions in the absence of NO_x were not generated. A number of updates to the system were made when SAPRC-99 was updated to SAPRC-07, including the ability to generate mechanisms for dialkenes and bi- and polycyclic compounds, reactions of chlorine atoms and chlorine-containing compounds, and reactions of alkynes and amines (Carter, 2010). It was still used only to generate mechanisms in the presence of NO_x , and aromatics were still not supported. The system was further updated for use in the current version of SAPRC, designated SAPRC-16 (Carter, 2019b), and the methods used to implement its output into the mechanisms was revised. It now includes the ability to generate mechanisms for benzene and alkylbenzenes, and reactions and product formation that occurs in the absence of NO_x is now supported. This version of the system, which is referred to as MechGen in the remainder of this document, is documented in this report.

Note that this report focuses on the use and capabilities of the mechanism generation system itself, specifically how to use it to derive explicit mechanisms, how the results can be output, and how the results can be used to lumped mechanisms for modeling. The scientific basis of the detailed mechanism generation process, including estimation procedures employed for the various types of reactions, assignments of known reactions and rate constants, and assumptions and approximations employed, is described in a separate document (Carter, 2019a). This report is sufficient for using the capabilities of the system and understanding how condensed mechanisms are derived from the detailed generated reactions, but the companion report is necessary for documenting the scientific basis of the detailed generated mechanisms.

1.2. Overview

The SAPRC-16 mechanism generation system, or MechGen, is capable of generating fully explicit mechanisms for the atmospheric reactions of many types of organic compounds and the intermediate radicals they form. Although in principle it could operate like GECKO-A (Aumont et al, 2005) and generate all the reactions of a selected compound and its oxidation products leading either to nonvolatile compounds or CO or CO₂, in practice this system is designed to be used to generate reactions leading to stable products, with the subsequent reactions of only radical intermediates being automatically generated. Reactions of stable product compounds can be generated as a separate step, in the same manner as used for the starting compound.

Note that MechGen is designed to estimate and generate mechanisms for lower tropospheric modeling only, so the estimates are applicable primarily for temperatures at or near 300 K and pressures at or near 1 atmosphere. Although temperature-dependent rate constants are estimated for some reactions, for others the rate constants or branching ratios are only applicable for 300 K and 1 atmosphere. In particular, MechGen is not currently designed for estimating mechanisms for combustion modeling or very low pressure systems. This is discussed further in the companion report giving the details of how the mechanisms are derived.

Table 1 lists the types of stable compounds whose reactions can be generated and shows the types of initial atmospheric reactions that can be generated for them. Table 2 lists the types of reactions can be generated, including reactions of intermediate radicals as well as reactions of stable compounds. These include reactions of stable compounds by H-atom abstractions by OH, NO₃, and Cl radicals, additions to double bonds by these radicals and by O₃ and O³P, and photolyses at various groups. The types of radicals that can be generated include carbon-centered radicals that in most cases react primarily with O₂, peroxy radicals that in most cases react with NO, NO₂, NO₃, HO₂, or other peroxy radicals and in some cases also have unimolecular reactions, alkoxy radicals that can react with O₂ or by various types of unimolecular reactions, and excited and stabilized Crigeer biradicals. The details of their reactions and how they are generated are described in various sections of this document as indicated on Table 2.

The following section gives a brief summary of the programming platform, an overview of programming concepts useful for understanding the system, and user interface options. The major operations of the system, including how reactants are specified, how explicit mechanisms are generated, and how explicit mechanisms are used to derive lumped mechanisms for compounds and mixtures, are described in Section 3. This includes how reactants (whether stable or radicals) are specified, how explicit mechanisms are generated, and how explicit mechanisms. Section 4 additional discussion about programming aspects, software components and maintaining the system, including components and options used specifically when developing the SAPRC-16 atmospheric chemical mechanism (Carter, 2016).

Table 1. Types of stable compounds whose reactions are supported by MechGen.

Type of Compound	Reactions	Exceptions not supported
Alkanes	OH, NO ₃ , Cl	Some estimates for polycyclic compounds are not always reliable
Alkenes (including multiple double bonds)	OH, O ₃ , NO ₃ , O ³ P, Cl	See above. Not all reactions of unsaturated Cl-addition radicals can be generated.
Alkynes	OH, NO ₃ , Cl	See above
Aromatic hydrocarbons	OH, Cl	Naphthalenes and other PANs
Aldehydes, ketones, hydroperoxides, organic nitrates [a]	OH, NO ₃ , Cl, hv	
Alcohols, ethers, esters, other oxygenates [a]	OH, NO ₃ , Cl	
Unsaturated aldehydes, ketones, hydroperoxides, and nitrates	OH, O ₃ , NO ₃ , O ³ P, Cl, hv	Aromatic oxygenates
Other unsaturated oxygenates	OH, O ₃ , NO ₃ , O ³ P, Cl	Aromatic oxygenates
Amines [a]	OH, NO ₃ , Cl	Amines with no α or β hydrogens
Peroxonitrates	Unimolecular, OH, NO ₃ , hv	
Bi- and polyfunctional compounds	OH, O ₃ , NO ₃ , O ³ P, Cl, hv, as applicable	Estimates are generally much more uncertain and less reliable

[a] No double or triple bonds except for carbonyls or nitrates, if applicable.

Table 2. Summary of types of reactions supported by the mechanism generation system. "VOC" indicates any stable compound supported by the system (see Table 1).

Reactant(s)	Type of reactions (* indicates a significant change for this version)	Section
VOC + OH	H-atom abstraction Addition to double bonds Addition to aromatic rings (OH only)	
VOC + O ₃	Addition to double bonds followed by Criegee biradical formation. Excited adduct addition to amines, followed by decomposition forming OH	
VOC + NO ₃	H-atom abstraction Addition to double bonds	
VOC + O ³ P	Addition to double bonds	
VOC + hv	Breaking the weakest bond in aldehydes, ketones, hydroperoxides, α -dicarbonyls, PAN compounds, and organic nitrates More rapid photolysis of dinitrates and carbonyl nitrates and carbonyl hydroperoxides	
VOC Uni.	Unimolecular decompositions of peroxy nitrates	
Carbon-centered Radicals	Unimolecular decompositions of radicals with α -nitro, α -nitrate or α -peroxy groups Reaction with O ₂ with H-abstraction from α -OH groups Addition of O ₂ to radicals with allylic resonance Reactions of O ₂ with aromatic - OH adducts Addition of O ₂ to other alkyl radicals	
Peroxy or Acyl peroxy Radicals	Reactions with NO forming the corresponding alkoxy radical or organic nitrate Cyclization of aromatic OH-O ₂ adducts Unimolecular H-shift reactions forming hydroperoxides Reactions with NO ₂ forming the corresponding peroxy nitrate or PAN Reaction with HO ₂ forming the corresponding hydroperoxide Reaction with NO ₃ forming NO ₂ and the corresponding alkoxy radical Reaction with the total of other peroxy or acyl peroxy radicals forming the corresponding alkoxy radical, carbonyl compound, or alcohol	
Alkoxy Radicals	α -H abstraction by O ₂ forming the corresponding carbonyl compound β -scission decompositions H-shift isomerizations Ester rearrangement	
Excited Criegee biradicals	Decompositions, stabilization, or rearrangements of saturated biradicals Internal addition to the double bond of unsaturated biradicals, followed by O-O scission and epoxide formation.	

2. PROGRAMMING PLATFORM AND MAJOR COMPONENTS

The current version of MechGen is incorporated into an online MOO system, which was originally developed as a programmable text-base virtual reality system (MOO, 2016, 2014, 1997). This type of text-based system is no longer widely used for online virtual reality experiences and the programming system is no longer being developed or supported, but features of the MOO object-oriented programming language, which is very similar to Python (Python, 2016), made it much better suited for mechanism generation applications than Fortran or other programming languages that the author happens to be familiar with, so that is why it was used for its initial development. In theory this system could be converted to Python or another platform whose underlying programming system is still being supported. This would allow it to continued to be maintained into the future as a collaborative effort and with more people being available who can program for it. However, this would be a major effort that is well beyond the scope of the current project, and there is no urgent incentive to do so since the current system performs satisfactorily. Its online access capabilities, discussed below, provide advantages that may be more difficult to implement using other platforms.

The MOO programming system, like Python, uses various software "objects" to organize procedures, operations, and data. Since the MOO system was developed to represent virtual realities, these objects persist in memory and are available whenever the program is running. Many of these can be thought of as analogues to physical objects, while others are used mainly to organize data or software. Normally the MOO system that hosts the mechanism generation system is running all the time, to allow users to access it via telnet or a web interface (see below) whenever desired. When the program is shut down these objects are saved, and they are restored when the program is resumed and re-loaded. These include "generic" or "parent" objects that have programs or commands, subroutines, or functions (called "verbs") and assignments of data (called "properties") associated with them, and "descendent" or "child" objects that "inherit" the verbs and properties of their parent object, and have property values that are associated with their particular instance. Although descendant objects can also be persistent, in many cases the normal operation of MechGen involves creating and deleting ("recycling") many descendant objects.

MechGen uses a number of different types of software objects that are used for various purposes. The most important of these in terms of understanding how the system works and is organized are summarized below.

- A "reactant" object is used to define and store information for a specific reactant compound or radical. The generic reactant object includes commands and subroutines involved in carrying out operations on reactants such as building the structure or initiating reactions, and properties with assignments or information shared by all reactants. Descendant reactant objects define a specific reactant and have properties indicating the structural groups it contains and its associated reaction chamber object (see below), and results of its reactions after they have been carried out. An important property of non-generic reactant objects is its "identity", which is a character string uniquely defining its structure, as discussed in Section 3.1. Mechanism generation starts with creating a persistent reactant object defining the initial reactant, and involves creating temporary reactant objects that are intermediates and products formed in the various reactions, which are deleted once the reactions involving them are listed and saved on properties of the initial reactant object.
- A "group" object is used to define a chemical group within a molecule, such as $\text{-CH}_2\text{-}$, -CO- , $\text{-CH(OO}\cdot\text{)-}$, etc (see Section 3.1 for a complete list). Note that these can have more than atom other than perhaps hydrogen or oxygen, of which there may be up to three. Each type generic

group object contains the subroutines and functions used to generate or estimate reactions involving the type of group, and their properties contain data used in making estimates. Each descendant group object has properties indicating what is its reactant object, and which groups are bonded to it with it double, triple, or aromatic bonds. Descendent group objects for specific reactants are temporary, being created when needed to carry out a reaction on the object using the information in the reactant's "identity" string, and then deleted when the reaction has been processed. Properties and verbs on these group objects is where many of the chemistry assignments are made, and are discussed elsewhere (Carter, 2019a).

- A "reaction chamber" object is treated like a "room" in the MOO virtual reality system where the logged-in user must be "located" to create reactants and generate mechanisms. Its commands and subroutines control the user interface (whether telnet or web access -- see following section), and its properties specify the environment where the reactions occur (e.g., temperature and pressure) and the database objects that are used to determine which set of assignments and mechanism generation and lumping procedures to employ (of which there are several options), and give other mechanism generation options. Different "chambers" can be used to generate mechanisms for different temperatures and pressures and with different mechanism generation options or lumping procedures, though most mechanisms are generated using default chambers with options appropriate for SAPRC-16 (Carter, 2016).
- Various database objects are used to contain assignment data that are used to generate mechanisms. This includes the database of manually assigned (e.g., measured) rate constants or branching ratios that are to be used instead of estimates, the database of thermochemical group values for estimating heats of reaction (needed for many rate constant estimates), parameters used for rate constant estimates that are not part of group properties, the database of SAPRC detailed VOC model species names and structures and default types of reactions to be generated for them, and several others. This allows for the possibility of generating mechanisms using different assignment data.
- The "lumping" objects are used to control the mechanism generation operations and the processing of mechanism generation output into various types of lumped mechanism. The lumping object used is specified as a property of the reaction chamber object for this purpose. Therefore, there is a separate such object for each type of lumped mechanism, and separate chambers are used to generate mechanisms for each. Although in principle mechanism generation is independent of the lumping method, the types of reactions that need to be generated depend on the lumping procedure employed. For example, SAPRC-99 through SAPRC-11 only need results of peroxy + NO reactions because the results of other bimolecular peroxy reactions are not used, while SAPRC-16 and explicit mechanisms need the results of all the bimolecular reactions.
- Various "player" objects are associated with each user that can log onto the system using the telnet interface (see Section 2.1.1). These control the user operations using the telnet interface. The main player object in MechGen has the name "MechGen" and "owns" all the programs and objects associated with mechanism generation (as opposed to those that are part of the core MOO system). There also has to be an administrative user (called a "wizard" in MOO systems) to do necessary systems maintenance or programming of the core MOO and also parts of the MechGen system that can impact sensitive parts of the system, so a "wizard" player object (named "MGadmin") is included for this purpose. (Passwords will be provided to those who obtain independent copies of the MechGen system, which is available on request.) There are also "guest" player objects that allow others to access the MOO system, though their capabilities when using telnet access is limited. Most general users would access the system using the web interface, discussed below in Section 2.1.2.

2.1. User Interfaces

2.1.1. Telnet Access

MechGen can be accessed using a Telnet client to log in with administrative access to program the system, input its assignment data, generate reactions, process results for mechanism implementation, and download the results in text files for incorporation into the mechanism. This method of access is similar to entering a text-based virtual reality system, with the system requiring a player object name and password for the user to log into the system. Currently available login names are "MechGen" (for normal use of the system) "MGadmin" (for system maintenance or programming), and "Guest" (no password required, but minimal capabilities), as indicated above. If the name and password is accepted, then the user's telnet connection is connected to the player object that is assigned that name, which in turn is associated with a "room" where it was "located" when it used the system last. The room description is then output to the user's Telnet screen, and the user can type in commands recognized by the room or the player object itself. Normally the room associated with the "MechGen" player object would be a "reaction chamber" object, with the current default being that with assignments appropriate for SAPRC-16.

MechGen is built upon an underlying LamdaMoo-based virtual reality system (MOO, 2014), and the underlying system has other capabilities and uses many other objects and utilities. In order to maintain such a system, it is necessary to log in as "MGadmin", which has the needed system-level permissions for this purpose from time to time. A discussion of what's needed to administer MOO systems in general is beyond the scope of this document (see MOO, 2014, and references therein).

Although the MOO system is capable of allowing access via Telnet clients for others to work with the system, this capability is not currently implemented in a secure enough fashion to be practical. Therefore, only users with "MechGen" or "MDadmin" passwords can usefully access the system via telnet. However, different instances of the system can be installed and run separately, under control of different users who know the administrative passwords, which have to be provided with copies of the system. Examples of commands and operations available when logged in as mechanism generation system owner is given in Section 4.

2.1.2. Online access

Although telnet access is needed to exploit the full capability of the system, MechGen can also be accessed using a web interface to generate reactions of selected systems, and obtain associated documentation. This is a good way to explore the predictions of the system, and obtain explanations of how or why it generated the reactions and estimated the rate constants involved, and to get lists of generated reactions and products formed when compounds are reacted completely. The system can also be used to obtain information about most of the various types of estimation methods and lists of assigned reactions used when generating complete mechanisms for compounds that are on the list of SAPRC detailed model species. However, the web system cannot be used to program the system or make or change any of its assignments or capabilities.

It is necessary to log in to access the web system, but only so the system can assign a unique "reaction chamber" for each user, and to prevent simultaneous users from interfering with each other, and to prevent "web spiders" from affecting the system. Once logged in, users can create radical or stable reactants (there is help on how to designate structures using the standard MechGen designation, or Smiles notation or SAPRC detailed model species names can also be used -- see also Section 3.1), or can select compounds to react from a list of compounds on the current SAPRC species list. The system can

be used to generate reactions of stable compound with atmospheric species such as OH, O₃, NO₃, O³P, photolysis or by unimolecular reaction where applicable, or atmospheric reactions of radical species, and will output documentation information indicating estimated or assigned rate constants or branching ratios used and the products formed. The system can then be used to react the products, seeing how they are estimated to react and also has the capability of generating full mechanisms online. Examples of use of the online system are discussed in the following section in conjunction with the discussions of the various operations of the system.

3. OPERATION OF SYSTEM

3.1. Specification of Reactants

The structure of an organic reactant or radical is specified by giving the "groups" in the molecule or radical, and indicating the groups each are bonded to, and the type of bond. Groups are parts of molecules that are treated as units in the system, and contain no more than one carbon, nitrogen, or halogen atom, and also can contain zero to three hydrogen atoms or zero to three oxygen atoms. The list of groups that can be used for compounds with more than one group are given in Table 3, elementary species or compounds with only one group are listed in Table 4, and illustrative examples are shown on Table 5. Note also the following:

- Neighboring groups in non-branched and non-cyclic structures are separated by either "-", "=", or "#" symbols, depending on whether their bonds are single, double, or triple.
- Branched structures are indicated using "("'s to show third or fourth groups bonded to groups with more than two neighbors. If the bond between the center group and the 3rd or 4th group is a double or triple bond, it is indicated using a "=" or "#" after the "(", as shown for isoprene in Table 5. This is similar to the treatment of branched structure using the SMILES notation (SMILES, 2016). Other examples are also shown on Table 5.
- Monocyclic structures are indicated using "*" symbols to indicate groups bonded to other groups using single bonds. The "*" can be before or after the group name, but the standard format is after then name. Bi- and polycyclic structures are indicated using "*1", "*2", etc. symbols to indicate groups that are bonded together, and should always be given after the group name but before any bond designation. Examples are shown on Table 5.
- Aromatic structures can be created using "CH" and "C" groups with alternating single or double bonds as with other cyclic polyalkenes. However, the standard designation, which reflects the fact that separate groups are used for aromatic carbons, is to designate carbons in aromatic rings using "aCH" or "aC", depending on whether it has a hydrogen or a third group bonded to it, with the aromatic bonds designated as if they were single bonds. See the example for toluene on Table 5.
- Allylic radicals can be created using carbon-centered radical groups and alkene groups to designate one of the resonance structures, and it does not matter which structure is used for this purpose. However, the standard designation, which reflects the fact that separate groups are used for such radicals and the radical center is at more than one location, is to designate carbons where the radical center may be located as ".aCH2", "aCH[.]" or "aC[.]" and the other carbons in the allylic structure as if they are aromatic (i.e, alternating double and single bonds). This is only applicable for carbon-centered radicals with conjugated C=C double bonds. See, for example, the structures shown for methyl allyl radicals and the OH+benzene adduct on Table 5.
- Cis and trans isomerization about a double bond are indicated using "^" and "v" symbols, analogous to the use of "\" and "/" symbols in the SMILES notation. See the examples on Table 5. Note that specifications of cis-trans isomerization for cyclic compounds is not currently supported, nor is specification of optical isomers.

Although one can often enter structures in more than one way (see examples on Table 5), the system uses an algorithm to generate a unique structure definition string for each structure. These are the MechGen structures shown as the first example of the MechGen structures on Table 5. This is done so that the structure definition string can be used to determine if two products or intermediate species

Table 3. List of groups and group designations used to specify organic reactants

Non-Radical Groups			Radical Groups [a]		
Types, names, bonds [b]		Note [c]	Group types, names, bonds [b]	Note [c]	
Alkane			Carbon-centered radical groups		
-CH3	CH3-		-CH2.	.CH2-	
-CH2-			-CH[.]		
-CH()			-C[.]()		
-C()			=CH.	.CH=	
Alkene			=C[.]	-C[.]=	
=CH2	CH2=		-aCH2.	.aCH2-	[e]
=CH-	-CH=		-aCH[.]		[e]
=C()	-C()		-aC[.]()		[e]
=C=			-pC[.]		[f]
Alkyne			-CO.	CO[.]	
#CH	HC#		Peroxy radical groups		
#C-	-C#		-CH2OO.	.OOCH2-	
Aromatic or Allylic			-CH[OO.]		
-aCH-		[d]	-C[OO.]()		
-aC()		[d]	-CO[OO.]	CO[OO.]	
Oxygenate			=CHOO.	.OOCH=	
-CHO	HCO-		=C[OO.]	-C[OO.]	[f]
-CO-			-pC[OO.]		
-OH	HO-		Alkoxy radical groups		
-O-			-CH2O.	.OCH2-	
Nitrate, nitro, nitroso			-CH[O.]		
-ONO2	O2NO-		-C[O.]()		
-NO2	NO2-		=CH[O.]	CH[O.]	
-NO	NO-	1	=C[O.]	-C[O.]	
Amino			-CO2.	CO[O.]	
-NH2	NH2-		-pC[O.]		[f], 5
-NH-			Crige biradicals		
-N()			-CHOO[excited]	-CHOO[stab]	4
Imine			-COO[excited]	-COO[stab]	4
=NH	NH=	6	CHOO[excited]	CHOO[stab]	4
=N-	-N=	6	N-Containing radicals and excited adducts		
Halogen			-NH.	.NH-	
-F	F-	2, 6	-N[.]		
-Cl	Cl-	3	-NH2[Oexcited]	NH2[Oexcited]	[g]
-Br	Br-	2, 6	-NH[Oexcited]		
-I	I-	2, 6	-N[Oexcited]()		
			Carbenes		
			-CH[.]	CH[.]	7
			-C[.]		7

[a] Reactants are referred to as "radicals" if they have a radical group and as "stable" if they do not. Reactants cannot have more than one radical group.

Table 3 (continued)

- [b] "-", "=", "#" indicate that the group is bonded to another group with a single, double, or triple bond, respectively. "-" is also used to designate aromatic bonds or allylic bonds between groups with "a" or "p" prefixes. "()" indicates that the group is bonded to a third or fourth group with a single bond, e.g., "CH₃-C(CH₃)(CH₃)-CH₃ for the "-C()()-" group in neopentane. If two names are given then either form can be used to designate reactants with this type of group.
- [c] Codes for level of support for reactions of group are as follows:
- 1 Reactants can be built with this group but generating their reactions is not currently supported.
 - 2 Reactions of alkoxy radicals containing this group cannot be generated because some parameters needed for rate constant estimates have not been derived.
 - 3 Reactions of at least some radicals containing this group cannot be generated because of lack of necessary thermochemical group additivity values.
 - 4 Currently, the only reactions generated for stabilized Crigiee biradicals are reactions with water forming the corresponding acid.
 - 5 Reactions are not generated with these radicals, and they are treated as final products. Compounds with these groups should be lumped with appropriate phenoxy radical model species when deriving lumped mechanisms where these radicals are formed.
 - 6 The system generates reactions for these compounds but the estimates may not be appropriate.
 - 7 Currently the system does not generate reactions that form carbene radicals, but they are available for future use as photolysis products. They are assumed to react exclusively with O₂ to form the corresponding excited Crigiee biradical.
- [d] The "aC" codes are used to indicate carbon centers in aromatic rings and also non-radical unsaturated in allylic groups. See discussion in text.
- [e] The "aC[.]" codes are used to indicate radical centers in allylic groups. See discussion in text.
- [f] These are phenyl, phenyl peroxy, or phenoxy radical centers.
- [g] These are used as intermediates in reactions of O₃ with amines.

Table 4. List of elementary or single-group species names recognized by the system.

a) Elementary Products		b) Reacting Organic Radicals		c) Atmospheric Species [a]		d) Other Elementary Species	
				Name	Notes	Name	Notes
CH ₄	O ₂	HCO.	CH ₂ OO[excited]	OH		HO ₂	[b]
HCHO	H ₂ O	CH ₃ .	CH ₂ OO[stab]	NO ₃	[b]	NO	[b]
CO	HF	CH ₃ OO.	CH ₂ [..]	O ₃		NO ₂	[b]
CO ₂	HCl	CH ₃ O.		O ₃ P	[c]	H.	
	HBr	HCO ₂ .		Cl.		F.	
	HI			HV	[c]	Br.	
				uni	[c, d]	I.	

- [a] These are species that can react with stable compounds, depending on the compound. See Table 1.
- [b] Reactions of these species with peroxy radicals can be generated.
- [c] These are currently not formed as products in any reaction.
- [d] This is a virtual species that can be used to designate unimolecular reactions of non-radical compounds in contexts where the type of reaction needs to be specified.

Table 5. Examples of designations of selected representative compounds and radicals.

Compound	Smiles [a]	MechGen Structures [a]
propane	CCC	CH3-CH-CH3
propene	C=CC	CH2=CH-CH3; CH3-CH=CH2
methyl acetylene	C#CC	CH#C-CH3; CH3-C#CH
2-methyl propene	CC(C)C	CH3-CH(CH3)-CH3
2-methyl-2-ethyl butane	CCC(C)(C)CC	CH3-CH2-C(CH3)(CH3)-CH2-CH3; CH3-C(CH3)(CH2-CH3)-CH2-CH3
isoprene	C=CC(=C)C	CH2=CH-C(=CH2)-CH3; CH2=CH-C(-CH3)=CH2
2-butenes (mixed cis & trans isomers)	CC=CC	CH3-CH=CH-CH3
cis-2-butene	C/C=C\C	CH3-^CH=CH-vCH3; CH3-vCH=CH-^CH3
trans-2-butene	C/C=C/C	CH3-^CH=CH-^CH3; CH3-vCH=CH-vCH3
trans-3-methyl-2-pentene	C/C=C(\C)CC	CH3-^CH=C(vCH3)-CH2-CH3; CH3-CH2-C(^CH3)=CH-^CH3
cyclopropane	C1CC1	CH2*-CH2-CH2*; *CH2-CH2-*CH2
bicyclo [1.1.1] heptane	C1C2CC1C2	CH2*1-CH*2-CH2-CH*1-CH2*2; CH*12-CH2-CH(CH2*1)-CH2*2
spiropentane	C1CC12CC2	CH2*1-CH2-C*12-CH2-CH2*2; C*12(CH2-CH2*1)-CH2-CH2*2
toluene	Cc1ccccc1	CH3-aC*-aCH-aCH-aCH-aCH-aCH*; CH3-C*=CH-CH=CH-CH=CH*
naphthalene	c12ccccc1ccc2	aC*12-aCH-aCH-aCH-aCH-aC*1-aCH-aCH-aCH-aCH*2
2-propyl nitrate	CC(C)ON(=O)=O	CH3-CH(CH3)-ONO2; CH3-CH(ONO2)-CH3
ethylene glycol ethyl ether acetate	CCOCCOC(C)=O	CH3-CH2-O-CH2-CH2-O-CO-CH3
2-propyl radicals	C[CH]C	CH3-CH[.]-CH3
2-propyl peroxy radicals	CC(C)O[O]	CH3-CH[OO.]-CH3
methyl allyl radicals	C=C[CH]C	CH3-aCH[.]-aCH-aCH2; CH3-CH[.]-CH=CH2; or CH3-CH=CH-CH2.
OH+benzene adduct	OC1C=CC=C[CH]1	HO-CH*-aCH[.]-aCH-aCH[.]-aCH-aCH[.]*

[a] The SMILES and first MechGen structure code given are those generated by the system. Subsequent MechGen structures, if given, are alternatives that can be used to create the same compound. Other SMILES or MechGen structures may also be acceptable for creating some of these compounds.

generated by the system are the same compound. Therefore, the structure specification generated by the system when a new molecule is specified may be slightly different than the one input by the user, though they would refer to the same compound. This is the case when the structures is entered using SMILES as well.

Reactants can also be created using the SMILES notation rather than the using the MechGen notation discussed above. However, the system saves the structural information using the MechGen codes, which has more direct relationship with the structural group objects used by the system, and MechGen codes are used when displaying reactions, intermediates, and products created when mechanisms are generated. However, the system can output the SMILES notation for any reactant if requested.

Reactants can also be created using pre-defined SAPRC "detailed model species" (DMS) names that are used for organic compounds whose mechanisms have been derived for SAPRC-16 and previous versions (Carter, 2010, 2010a, 2016). These are the names that identify the compound in the SAPRC mechanisms. A listing of SAPRC detailed model species that are defined and supported by the current version of MechGen are given in Table ?? in Appendix ?. For example, the first six compounds listed on Table 5 have DMS names of "PROPANE", "PROPENE", "ME-ACTYL", "2-ME-C3", "33-DM-C5" and "ISOPRENE".

3.2. Reaction Generation

Once the reactant is created, its reactions can be generated either in "single step" mode or "react completely" mode. Single step mode consists of generating only the initial reaction(s) of the specified compound, but not the reactions of the products formed. React completely mode not only generates the initial reactions of the compound, but also generates the reactions of the radical products formed, as described in Section 3.2.2. Normally the react completely mode is used to derive mechanisms for modeling, and the specifics of the process, and the post-processing employed to derive lumped depends on the mechanism and lumping approach. However, running reactions in single step mode is useful to obtain information on how the system estimates a compound's reactions, since the results show not only the products formed and the rate constants or branching ratios involved if there are competing processes, but also the system outputs documentation concerning the estimates or assignments used when the reaction(s) of the species was generated. Because potentially a large number of species and reactions are involved, documentation output is not produced when running the reactions in react completely mode.

3.2.1. Single Step Reactions

How single step reactions are initiated depends on whether the reactant is a stable compound or a radical. If the compound is a radical, i.e., has a radical group as indicated on the right side of Table 3, then all the reactions of the radical that are estimated to be non-negligible under tropospheric conditions are generated. The specific reactions that can be generated for each type of radical, and how they are estimated, are discussed in detail elsewhere (Carter, 2019a). These could include unimolecular reactions or reactions with O₂, and for peroxy radicals they can also include reactions with NO, NO₂, NO₃, HO₂, generic RO₂, and generic RCO₃ (acyl peroxy) radicals, and for stabilized Criegee biradicals they include reaction with H₂O.

In the case of stable compounds (i.e., compounds without radical groups as indicated on Table 3), it is necessary to specify the atmospheric species with the compound is reacting, which could be OH, O₃, NO₃, O³P, hv, or "uni" (for unimolecular decomposition, such as for peroxy nitrates). If the compound does not react with the species (or have a unimolecular reaction), or reacts only at a negligible rate under

atmospheric conditions, then no reaction is generated. If the compound reacts in more than one way with the species (e.g., the species adding or abstracting from different parts of the molecule, or having more than one photolysis or unimolecular reaction), then all the possible reactions are generated.

The types of single-step reactions generated for compounds with radical groups are summarized on Table 2. In the case of alkoxy and carbon-centered radicals, it is assumed that unimolecular reaction and/or reaction with O₂ are the major loss processes and that there are no other bimolecular reactions. In the case of peroxy and acyl peroxy radicals, unimolecular reactions (where applicable) and reactions with NO, NO₂, NO₃, HO₂, and generic RO₂ and RCO₃ are generated. Because of the many possible peroxy and acyl peroxy radicals that can be formed, it is not practical to generate all the possible peroxy+peroxy cross-reactions, so instead reactions with a only a single generic or lumped RO₂ and a single lumped RCO₃ species are generated. These are converted to reactions with "SumRO2" and "SumRCO3" model species when these reactions are processed for the SAPRC-16 mechanism. The generic RO₂ species is assumed to have an abstractable α -hydrogen that can be transferred to the peroxy radical reacting with it, while the generic RCO₃ species does not.

In all cases, the system outputs documentation information about the generated reaction to the user, and the products formed in all the reactions are generated. The web-interface user can then select one of the products and generate its reactions, form its products, and see the associated documentation, if desired. The product objects formed are temporary, and are deleted when another reactant is created or reacted.

3.2.2. Full Mechanism Generation

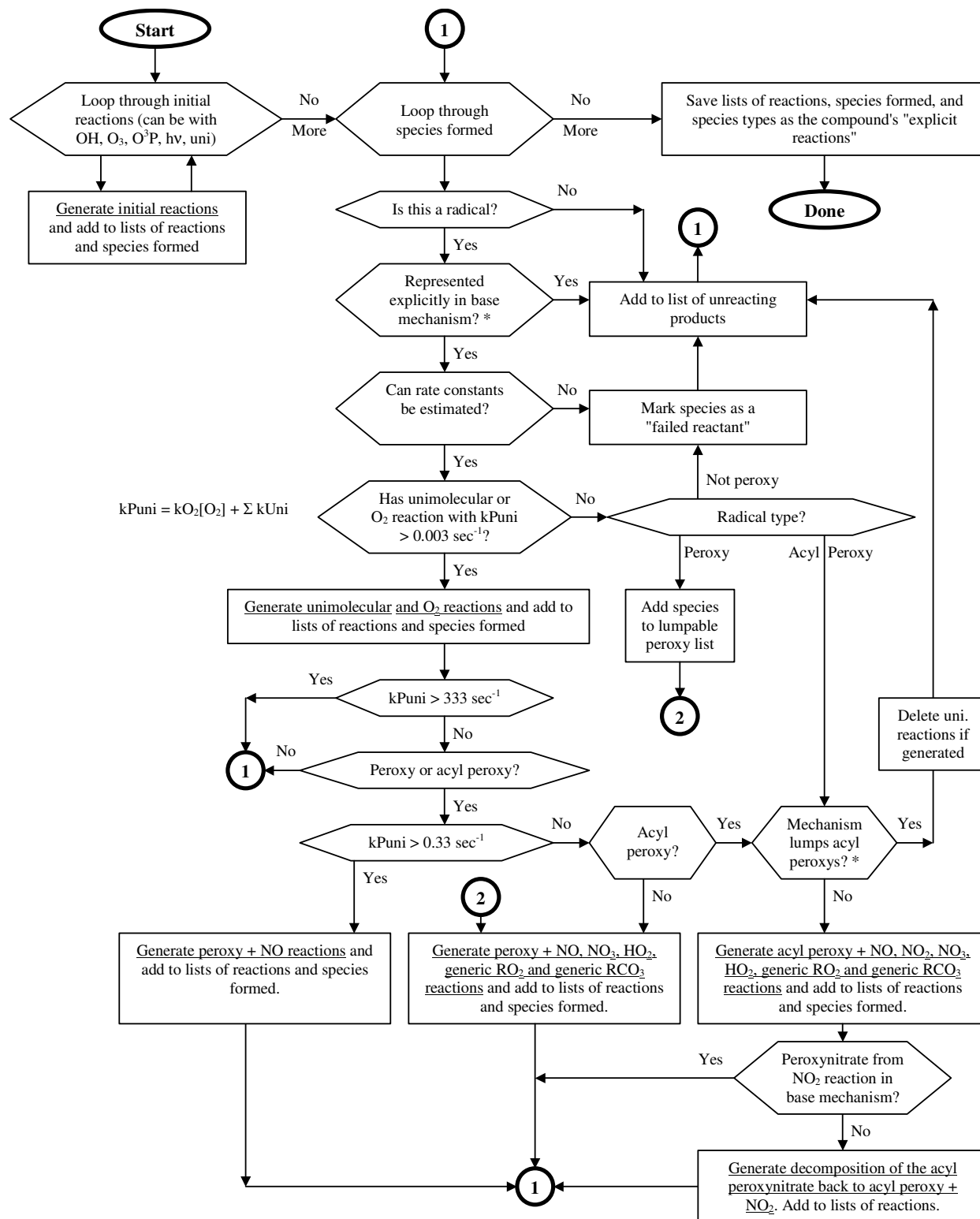
Full mechanism generation for stable compounds involves first reacting the compounds with all or a designated subset of its possible atmospheric reactions, and then reacting most of the radicals formed, using procedures that depend in some cases on the lumping method employed. Reactions of products or radical species that are already part of the base mechanism¹, or that are represented by lumped model species whose mechanisms do not depend on the compounds they represent², are not generated. The net result is a sequence of reactions of the starting compound and the radicals whose reactions are generated, and lists of the various types of products and reacting intermediates involved. These results are stored as properties on the starting reactant, where they can be processed further and can be output to the user as discussed below.

The algorithm used for explicit mechanism generation is shown schematically in Table 1. The steps involved are summarized below.

1. The structure of the organic compound whose mechanism is to be estimated is provided as an input to the system as described in Section 3.1. The types of initial reactions that the compound can undergo can optionally also be assigned, based on the type of compound and the needs of the

¹ The base mechanism consists of the reactions of the inorganics and common organic products or explicitly represented compounds or intermediates that are included with the mechanism along with processed mechanism generation output. Reactions of species in the base mechanism do not have to be generated because they are already part of the mechanism. Examples of such species for SAPRC-16 include methane, methyl and ethyl peroxy, formaldehyde, acetyl peroxy radicals, PAN, and others (see Carter (2019b) for the list of SAPRC-16 base mechanism).

² For example, SAPRC-16 represents the various acyl peroxy radicals using either explicit or lumped model species whose mechanisms are assigned based on representative compounds. Since the SAPRC-16 representation does not depend on reactions generated for these specific radicals, their reactions do not need to be generated.



* E.g., Methyl peroxy, ethyl peroxy, phenoxy, t-butoxy, are explicit, and acetyl peroxys are either explicit or lumped in SAPRC-16

Figure 1. Flow diagram of the process of generating mechanisms for emitted and selected organic product compounds.

model where the mechanism will be used. For example, almost all compounds are assigned as reacting with OH radicals, alkenes are designated as reacting with OH, O₃, NO₃, and O³P, aldehydes as reacting with OH, NO₃ or by photolysis, etc. Reactions of organics with Cl radicals can also be generated, but since currently SAPRC-16 does not currently include halogen chemistry, these reactions are not currently included when generating mechanisms for SAPRC-16. If the types of initial reactions are not assigned, then all possible initial reactions are attempted, though those that are not supported or estimated to be unimportant are not included.

2. All possible modes of initial reactions believed to be potentially important under atmospheric conditions are generated and the rate constant for each route is estimated or an assigned branching ratio is used if data are available. The types of reactions considered are summarized on Table 2, and the details of how the reactions are estimated or assigned are described by Carter (2019a). The resulting reactions are added to the list of explicit reactions, along with their assigned or estimated relative or absolute rate constants. Each explicit reaction refers to an elementary process, with no lumping or combining of consecutive processes. Attempts to react compounds with species whose reactions are not supported, such as photolysis or ozone reactions for alkanes, result in no reactions being generated. Each explicit reaction is assigned a "yield", which is the rate constant for this reaction divided by the total rate constant for of the compound with the atmospheric species with which it reacts. If the yield is less than 0.5%³ relative to the total overall reaction then the reaction and products are deleted, and the yields of all competing reactions are increased as needed to make the total yield 100% for reaction with each type of atmospheric species.
3. The products of the generated reactions are examined to determine how they are to be processed. No processing is done for products that have been previously generated and examined,, other than retaining them as products of the reactions. If the new product is a stable compound or a type of radical that is represented explicitly by a model species in the mechanism for which the reactions are generated then they are treated as an end product in the system and their subsequent reactions are not generated. The latter include explicitly represented radicals such as OH, HO₂, and, for SAPRC-16, methyl and ethyl peroxy, t-butoxy, phenoxy, acetyl peroxy, and several other peroxy radicals (see Carter, 2017). New radical products that are not represented explicitly in the mechanism are added to the bottom of the list of intermediate products whose reactions are to be generated.
4. The intermediate radicals to be reacted are processed on a first-in, first out basis. Note that the intermediate will not have been reacted previously because only new intermediates are added to the list of radicals to be processed in step 3. The intermediate is removed from the list before its reactions are generated.
5. The unimolecular reactions and (if applicable) reactions with O₂ are processed first. The yields of each such reaction are determined based on the ratio of the rate constant for the reaction to the total of rate constants for all these reactions, where the rate constant for reaction with O₂ is converted into a pseudo-unimolecular rate constant by multiplying it by the atmospheric O₂ concentration. Routes whose yields are less than 0.5%³ relative to the total are deleted, and the yields of competing reactions are increased so the total remains 100%. The relative yields are calculated using their estimated or assigned rate constants, with the pseudo-unimolecular rate constant for reaction with O₂ calculated using the atmospheric O₂ concentration.

³ This "de-minimis yield" parameter is an option that can be varied if desired. See Section ?? (not yet added)

6. If the radical is a peroxy or acyl peroxy radical and the sum of the unimolecular and/or pseudo-unimolecular rate constants is greater than 133 s^{-1} (see note⁴), or if the radical has unimolecular or O_2 reactions but is not a peroxy radical, then the mechanism generation process returns to step 3 to process products of the reactions of this radical and continue the reaction generation process.
7. If no unimolecular or O_2 reactions are generated and the radical is not a peroxy or acyl peroxy radical, then something went wrong because all radicals other than perhaps peroxy or acyl peroxy are expected to have unimolecular or O_2 reactions. This might occur if reactions cannot be generated because a needed thermochemical or kinetic parameter assignment is missing. In this case, the radical is added to the list of unreacting products and also marked as a "failed reactant". The process then goes to step 4 to continue processing the list of radicals (and in so doing find out if there are other such "failed reactants" generated from the starting compound). However, mechanism generation is considered unsuccessful if there are failed reactants and subsequent processing of the reactions of the compound for the mechanism cannot be done.
8. If the radical is a peroxy or acyl peroxy radical that has unimolecular reactions with total rate constants of greater than 0.33 s^{-1} (see note⁴), then it is assumed that the only significant loss processes for the radical are unimolecular reactions or reaction with NO . Although reactions of acyl peroxy radicals with NO_2 forming a PAN analogue can also occur at significant rates, it is assumed that the unimolecular reactions with this minimum rate constant are sufficiently fast that the PAN analogue does not build up and become an important net sink for the radical, so its formation can be ignored. Therefore the reaction with NO is generated, and the yields associated with these reactions are relative to the total rate of reaction with NO . The procedure then returns to step 3 to process the products of these unimolecular and NO reactions, and then proceed with the reaction generation process.
9. If the radical is an acyl peroxy radical with no unimolecular reactions⁵ or a total unimolecular rate constant less than 0.33 s^{-1} (see note⁴), and if the mechanism for which the reaction is being generated uses lumped species to represent acyl peroxy radicals, then it is treated as an unreacting product and its reactions are not generated, and the system goes back to step 4 to react the next radical on the list and continue the generation process. Its reactions are not generated because they are not used in the model if it is represented by a lumped species. If the mechanism does not lump this type of acyl peroxy radical, then its reactions with NO , NO_2 , NO_3 , HO_2 , and generic RO_2 and RCO_3 are generated and the process returns to step 3 to process the products of these reactions and then continue with the reaction generation process.
10. If the radical is a peroxy radical or an is acyl peroxy radical that is not represented by lumped model species, and it has either no estimated unimolecular reactions or a total unimolecular rate constant less than 0.33 s^{-1} (see note⁴), then its reactions with NO , NO_2 (for acyl peroxy only), NO_3 , HO_2 , and generic RO_2 and RCO_3 are generated and the process returns to step 3 to process the products of these reactions and then continue with the reaction generation process. The reactions of non-acyl peroxy radicals with NO_2 are not generated because it is expected that the peroxy nitrates formed will decompose sufficiently rapidly to reform the radicals and NO_2 that their formation do not constitute net sinks for the radicals.

⁴ The parameters "kSlow", "kSlowPAN", and "kFast", with default values of 3.3×10^{-3} , 0.33, and 133 s^{-1} , control the processing of bimolecular peroxy radical reactions based on their total pseudo-unimolecular reaction rates. They can be modified if desired for specific mechanisms. The values used for SAPRC-16 were derived based on test calculations but the possibility that different values may be more appropriate for some model applications has not been ruled out. See Section ?? (not yet added).

⁵ If the total of unimolecular rate constants for a peroxy or acyl peroxy radical is estimated to be less than kSlow (see note 4) then its unimolecular reactions are ignored.

11. If the reaction of an acyl peroxy radical with NO_2 , forming an acyl peroxy nitrate (PAN analogue), is generated, then the thermal decomposition reaction re-forming the acyl peroxy radical and NO_2 is also generated. Note that this reaction is not counted as a source of the acyl peroxy radical for yield estimation purposes because there is a near-equilibrium between the PAN analogue and its acyl peroxy + NO_2 precursors.
12. This process is complete once the list of radicals to be reacted has been completely processed.

The result of this process is a list of explicit reactions and their relative or absolute rate constants, and lists of final products and intermediate reactant radicals that were generated. The various products and intermediates involved are also listed and are classified as indicated on Table 6 – these classifications affect the subsequent processing of the mechanism as discussed in the following sections. This is referred to as the "explicit mechanism" for first generation reactions of the subject compound. The explicit reactions of the products, if needed for the complete mechanism, can be generated separately using the same procedures. The results are stored on the "reactant" object, and can be used as input for subsequent processing for use with SAPRC-16 or other lumped or reduced mechanisms.

3.3. Processing Mechanisms for SAPRC-16

Although in principle the explicit mechanism could be used to represent the atmospheric reactions of the compound, the number of reactions and intermediate species are generally much larger than needed for most model applications. Even if complete chemical detail is desired, the number of species and reactions that need to be simulated could be reduced significantly by using "process lumping" as described below in Section 3.3.1. This results in a "processed mechanism" can then be used if minimal chemical approximations are desired. This can be further reduced and adapted to the condensed mechanism it is to be used with by using "species lumping" as described elsewhere (Carter, 2019b). The resulting "lumped mechanism" can then be used with SAPRC-16 (Carter, 2019b) or a similar lumped or reduced mechanism.

3.3.1. Process Lumping for Compounds

Process lumping involves combining reactions or processes to reduce the number of species and reactions in the explicit mechanism without introducing significant chemical approximations. The overall process as used for SAPRC-16 is shown schematically on Figure 2. Although developed for SAPRC-16, it could be used for other mechanisms of similar structure. Process lumping for SAPRC-16 involves three steps, the first two of which introduces no chemical approximations except for assuming steady state for some rapidly reacting species, and the third that introduces only minor impacts for representing organic peroxy + organic peroxy reactions, which have to be represented approximately to avoid huge numbers of reactions to represent processes that are generally relatively unimportant under atmospheric conditions. These are as follows:

1. Merge Reactions. This involves combining all elementary reactions involving the same reactants into a single, combined reaction forming the products of all the competing reactions. Each product of the reaction is given a yield, which is the ratio of the rate constant for all the reactions forming this product, relative to the total rate constant for the reactant. This reduces the number of reactions in the mechanism, which may or may not make mechanism implementation in models more efficient, depending on the mechanism compiler software used. The main advantage of this step is that the algorithms for subsequent process and product lumping are generally much simpler. It introduces no chemical approximation unless the rate constants for the competing reactions have different temperature dependences, and the model is applied at different temperatures than used when the product yields are

Table 6. Classification of products formed in explicit mechanism generations

Description	How Processed for SAPRC-16
Non-radical product	Reactions are not generated. If needed, they can be generated for the overall mechanism as a separate step
Radical whose reactions are already included in the lumped mechanism	Reactions not generated because they are already in the model where the processed reactions will be used.
Non-peroxy radicals and peroxy and acyl peroxy radicals with $k_{\text{Puni}} [a] > 333 \text{ s}^{-1}$.	Only unimolecular or reactions with O_2 (if any) are generated. These reactants are replaced by the set of products they form during the "reaction lumping" stage. This is equivalent to using the steady state approximation for this species when running the model simulation.
Peroxy radical with $k_{\text{Puni}} < 3 \times 10^{-3} \text{ s}^{-1}$ at 300K.	Can be lumped with other such peroxy radicals formed in the same reaction(s) during the "reaction lumping" stage with minimal chemical approximation. Reactions with NO, NO_3 , and generic peroxy and acyl peroxy radicals are generated and processed. Reactions forming these radicals are indicated as also forming "SumRO2" when processed for the lumped mechanism.
Peroxy radicals with k_{Puni} between 3×10^{-3} and 0.33 s^{-1} at 300K.	Not lumped with other peroxy radicals during the "reaction lumping" stage. Unimolecular reactions and reactions with NO, NO_3 , HO_2 and generic peroxy and acyl peroxy radicals are generated and processed. Reactions forming these radicals are indicated as also forming "SumRO2" when processed for the lumped mechanism.
Peroxy and acyl peroxy radicals with k_{Puni} between 0.33 and 333 s^{-1} at 300K	Not lumped with other peroxy radicals during the "reaction lumping" stage. Unimolecular reactions and reactions with NO are generated and processed. Other bimolecular reactions are assumed to be negligible because the unimolecular reaction should dominate whenever NO is sufficiently low.
Acyl peroxy radicals with $k_{\text{Puni}} < 0.33 \text{ s}^{-1}$ at 300K	Products in this category are treated as non-reacting when generating for SAPRC-16 because all acyl peroxy radicals that do not have unimolecular reactions faster than 0.33 s^{-1} are represented by either explicit or lumped model species. Their unimolecular reactions, if any, are ignored. If needed for other mechanisms, their unimolecular reactions (if faster than $3 \times 10^{-3} \text{ s}^{-1}$) and their reactions with NO, NO_2 , NO_3 , HO_2 , and generic peroxy and acyl peroxy radicals can be generated and processed, and reactions forming these radicals are also indicated as forming "SumRCO3" when processed for the lumped mechanism.
Radicals whose reactions cannot be generated	If radicals are formed whose reactions cannot be generated because of lack of thermochemical or other needed assignment data, then they are treated as unreacting products. The mechanism generation process is treated as unsuccessful and not processed further if such products are generated with non-negligible yields, though an upper limit can optionally be set to allow further process if their yields are low enough. In that case, these species are treated as if they were unreactive and their formation is ignored.

[a] k_{Puni} is the sum of all unimolecular rate constants + (the rate constant for reaction with O_2 x the atmospheric O_2 concentration).

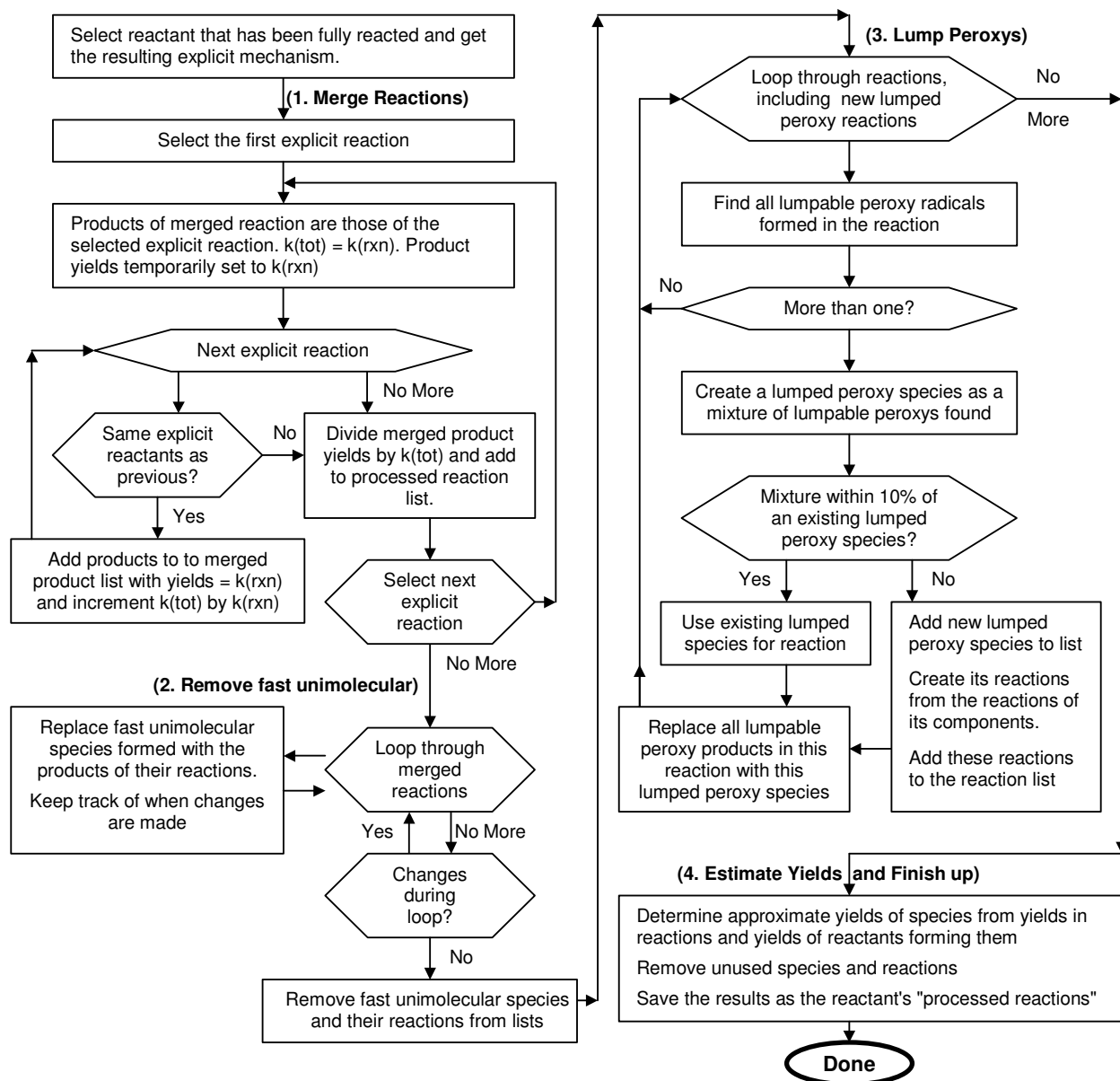


Figure 2. Flow diagram of process lumping for generated mechanisms.

derived when the reactions are merged. If the reactant reacts only one way, the reaction is retained with the same products, but is reformatted so that each product is given a yield of 1.0.

2. Remove Fast Unimolecular. This involves replacing rapidly reacting intermediate species that do not have bimolecular reactions other than with O_2 with the products that they form. This is done recursively as shown on Figure 2. This introduces no chemical approximation other than using the steady state approximation for the intermediate, which is appropriate for most of the radical intermediates in atmospheric mechanisms other than perhaps peroxy and acyl peroxy radicals under some conditions. This is also appropriate for those peroxy and acyl peroxy radicals that have sufficiently rapid unimolecular reactions that their bimolecular reactions are not generated (See Figure 1 and Table 6). Note that peroxy

radical intermediates that have bimolecular reactions cannot be removed from the mechanism and replaced by the products they form without introducing chemical approximations because the products they form depend on the relative levels of the NO_x and peroxy species with which they react. However, this step removes a significant number of reactive intermediates in most mechanisms, potentially making its implementation much more efficient.

3. **Lump Peroxys.** Competing reactions of larger reactants can result in the formation of many peroxy radicals that are formed from the same or very similar processes, but that cannot be replaced by their products because the products they form depend on NO_x and peroxy levels. However, if the peroxy radicals do not have unimolecular reactions and have the same total rate constants for their bimolecular reactions, then all peroxy radicals formed by a single process can be lumped as one peroxy radical model species, forming products determined by the relative formation rates of each of the component radicals from the common process. SAPRC-16, like most current mechanisms of varying detail, assumed that all the higher peroxy radicals have the same total rate constants for their reactions with NO_x species, HO_2 and acyl peroxy radicals. Although the rate constants for reactions of non-acyl peroxy radicals with other non-acyl peroxy radicals can be highly variable (ref?), SAPRC-16, like most lumped mechanisms, use the same rate constant for these reactions because it allows for much simpler representations of these very minor processes. Peroxy + peroxy reactions are negligible when NO_x is present, and test calculations show they are generally minor compared to peroxy + HO_2 or peroxy + acyl peroxy under most realistic atmospheric conditions when NO_x is low. Therefore, peroxy lumping has essentially no effect on model predictions using lumped mechanisms such as SAPRC-16, and only minor effects on mechanisms, such as MCM, that use different rate constants for different peroxy + peroxy reactions.

3.3.2. Mechanism Condensation

Although the process lumping procedures discussed in the previous section can significantly reduce the number of reactions and reactive intermediates needed to represent the atmospheric mechanisms without significant chemical approximation, the numbers of model species representing peroxy radicals can still be quite large for larger compounds. Often many of these model species are needed to represent only relatively minor processes and their reactions can be simplified without significant effects on model predictions, though this is not always the case. Condensed mechanisms such as the SAPRC mechanisms through SAPRC-99 (Carter, 1990, 2000) and the Carbon Bond mechanisms (e.g., Yarwood et al, 2005) remove these peroxy model species by the non-radical products they ultimately form when they react with NO , and use various "chemical operator" model species to represent effects of the additional NO to NO_2 conversion and NO consumption caused by multi-step mechanisms and radical termination and hydroperoxide formation caused by peroxy + HO_2 or peroxy + peroxy reactions. This requires the assumption that the organic products formed from peroxy + HO_2 or peroxy + peroxy have similar reactivity effects as those formed when peroxy + NO reactions dominate. This is generally considered acceptable if the primary objective of the model application is ozone prediction, but is problematical if prediction of secondary organic aerosol (SOA) is a priority, because the peroxy + HO_2 and peroxy + peroxy generally form higher yields of lower volatility compounds than peroxy + NO reactions. Although such highly condensed mechanisms are still widely used in models to predict SOA, it is necessary to either ignore known effects of NO_x levels on SOA formation, or to use separate SOA models in the models that include representations of the NO_x effects on SOA precursor formation that are missing from the gas-phase mechanism.

The SAPRC-07 and SAPRC-11 mechanisms (Carter, 2010a,b; Carter and Heo, 2012) deal with this problem by using separate chemical operator model species to represent precursors for each type of organic model species formed from peroxy + NO and peroxy + HO_2 reactions so that the dependences of their yields on NO_x and radical levels can be more accurately simulated (Carter, 2010a,b). This gives

similar results as the RADM through RACM2 mechanism (Stockwell et al, 1997; Goliff et al, 2013), which retains peroxy radical model species but lumps all of those formed from the same initial reactions, and still has to use chemical operators to represent the additional NO to NO₂ or NO to nitrate conversions of multi-step mechanisms. However, neither of these methods can adequately represent mechanisms that can form peroxy radicals that can undergo unimolecular reactions at rates competitive with their bimolecular reactions with NO or other radicals. Such unimolecular reactions of peroxy radicals were not previously considered in atmospheric mechanisms, but are now considered to be non-negligible. In addition, the use of operators to represent additional NO to NO₂ conversions or NO consumptions in multi-step accurately represent the dependence of product yields on NO_x levels under moderately low NO_x conditions where peroxy + NO and peroxy + HO₂ reactions occur at similar rates.

The approach used for SAPRC-16 is to use the more approximate SAPRC-99/Carbon Bond method to remove the peroxy radical model species formed in relatively low yields in the initial reactions, while retaining those that are formed in relatively high yields. The yield cutoff point is determined by an input parameter in the MechGen system, with SAPRC-16 using a cutoff value of 10%. However, the actual yields of peroxy radicals formed depend on reaction conditions, depending on relative importances of the initial atmospheric reactions of the starting compound, and the competitions in the reactions of peroxy radicals formed that form other peroxy radicals. Therefore, we use approximate upper limit yield estimates, derived as discussed above, for the purpose of determining which peroxy species can be treated more approximately.

For this purpose, we weigh each of the initial reactions of the starting VOC (with OH, O₃, etc.) as being equally important, and assume that the only significant bimolecular reactions of peroxy radicals are with NO, which is the dominant process under conditions where ozone is formed. Note that the reactions of peroxy radicals with NO₂ are reversed so the impact of this on peroxy yields is not counted, the reactions of peroxy radicals with HO₂ do not form radical intermediates, and their reactions with NO₃ and other peroxy radicals form a subset of the intermediates formed when they react with NO, so not counting these reactions does not result in the relative importances of reactive intermediates being underestimated. If these intermediates are important in the presence of NO_x, most if not all of them will also be formed when NO_x is absent, though generally their yields would be somewhat lower. The specific procedure employed is shown schematically on Figure 3.

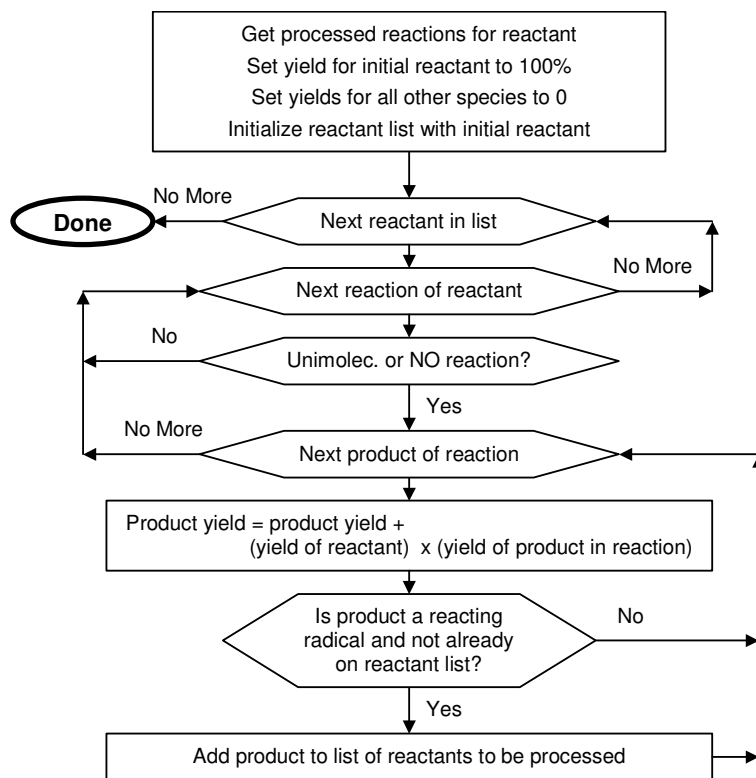


Figure 3. Flow diagram of the process of estimating upper limit intermediate and product yields.

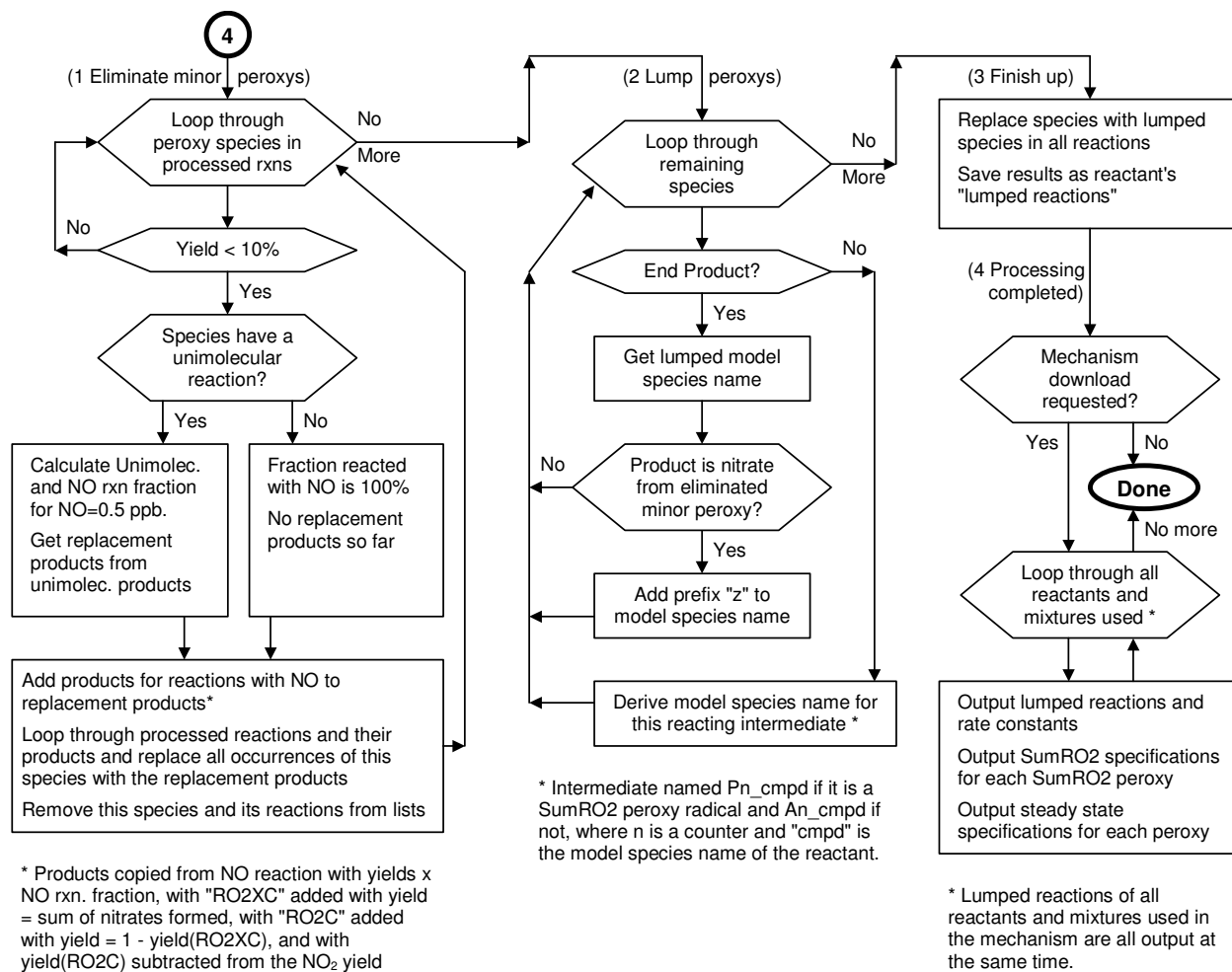


Figure 4. Flow diagram showing how a processed mechanism is used to derive lumped reactions can be included in the SAPRC-16 mechanism.

3.3.3. Deriving Mechanisms for Mixtures

The SAPRC-16 mechanism (Carter, 2019b) includes a number of lumped model species whose mechanisms are derived from explicit mechanisms generated for a specified mixture of compound that these model species are intended to represent. In previous versions of SAPRC, the approach used was to derive condensed mechanisms for the mixtures as weighted averages of the various mechanistic parameters for the components, where the parameters consisted of rate constants for the initial reactions, plus the yields for the various product and radical operator model species (Carter, 2000, 2010a,b). However, this approach can only be used if the lumped reactions of the compounds with atmospheric species such as OH and O₃ can be represented by single overall reaction forming only model species in the base mechanism. This approach cannot be used for SAPRC-16 because, as discussed in the following sections, the lumped mechanisms for most compounds require use of compound-specific peroxy radicals that can undergo competing reactions, so in general a single overall reaction cannot be used to represent the lumped mechanism.

Instead, for SAPRC-16, the explicit mechanisms for each of the components are used to derive explicit mechanisms for the mixture as if it were a single compound. The flow diagram for this process is shown on Figure 5. The result is a detailed mechanism for the lumped model species that can be treated by the software as if it is a mechanism for a single compound, where the various lumping approaches can be applied.

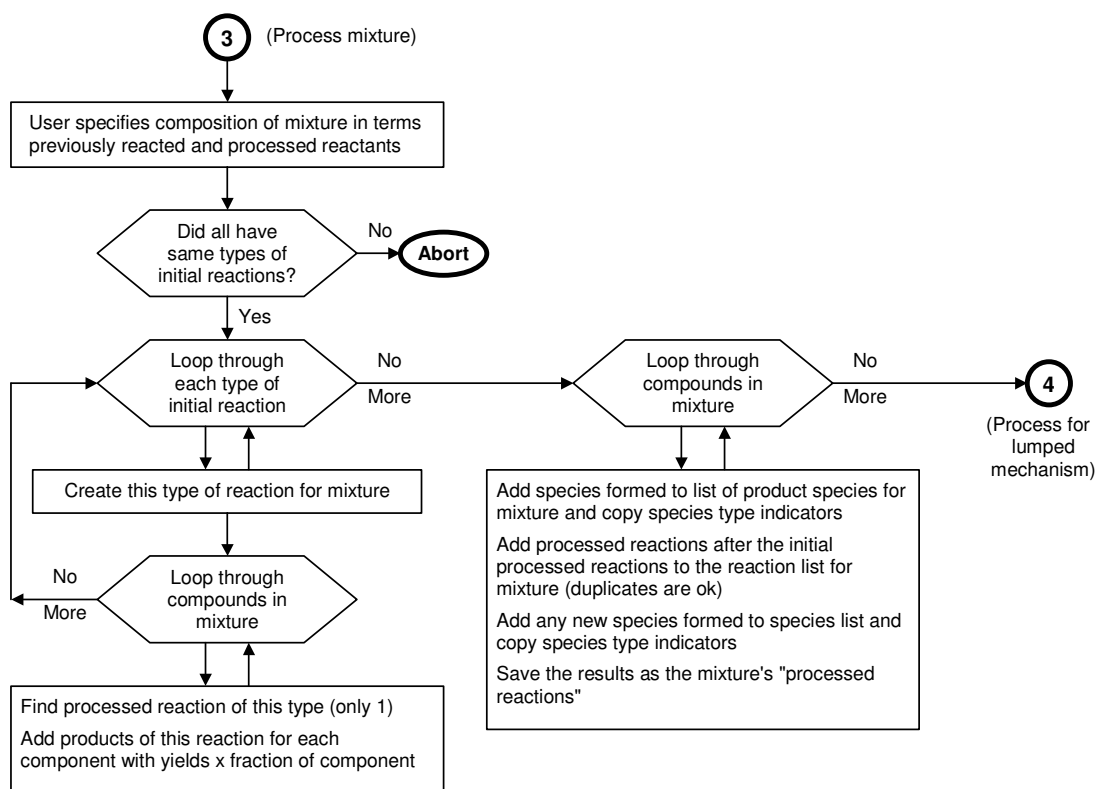


Figure 5. Flow diagram showing how mechanisms for mixtures are initially processed.

3.4. (Additional discussion to be added)

4. PROGRAMMING AND OPERATIONAL DETAILS

- 4.1. (Discussion of Installation and Initial Setup to be added)
- 4.2. (Discussion of programming and entering assignments using Telnet to be added)
- 4.3. Operations using the Telnet Interface
 - 4.3.1. Creating Reactants

Administrative users logged in using the Telnet interface (see Section 2.1.1) can create new reactants using the "build", "dms", or "build-smiles" commands, as follows:

```
build name as structure
build-smiles name as smilesstructure
dms dmsname
```

where *name* is the name the user wants to give the reactant (which must not be used elsewhere on the user or in the reaction chamber room), *structure* is the MechGen structure, *smilesstructure* is the SMILES structure, and *dmsname*. Note that MechGen commands using the Telnet interface are case insensitive, as are MechGen structures and DMS names, though SMILES structures are case-sensitive, since lower case means aromatic. A reactant object with the given name and structure is created in the reaction chamber room, where commands to generate reactions with it and separate commands to output the results can be given, as described later. The command fails and an error message is given if (1) an invalid *structure* or *smilesstructure* is given; (2) the name *dmsname* is not a valid pre-defined SAPRC detailed model speciesname; or (3) an object named *name* already exists in the reaction chamber room or held by the user.

4.3.2. (discussion of Additional Telnet Commands to be added)

4.4. Operations Using the Web Interface

4.4.1. Creating Reactants

When first logged into the system, or after selecting the "Reload Main Menu" link on the bottom of most pages, the user is presented with a web page that has a space for the user to enter the MechGen or Smiles structure of the reactant. The user can enter the structure in MechGen or Smiles format (MechGen is tried first), or a detailed model species name. If the structure or DMS name is valid, the user is displayed a page containing information about the compound (including its MechGen and Smiles codes, atomic formula, molecular weight, heat of formation (if it can be estimated), and the groups in the structure. Additional information on this page depends on whether the reactant specified is a stable compound or a radical.

If the compound specified is a stable compound, the page also displays links that the user can use to generate its reactions, as discussed in the following sections. These include links to generate single-step or multi-step reactions of the compounds with OH with other atmospheric reactions or photolyses if applicable, depending on the compound (see Table 1). It also gives the SAPRC-16 lumped model species used to represent this compound (Carter, 2016), and, if possible, estimates its vapor pressure using the Evaporation method of Compernelle et al (2011).

If the compound specified is a radical, the page automatically shows all the single-step reactions it can generate with the radical, as shown on Table 2. If the radical is assigned a lumped model species in

the SAPRC mechanism then this is shown as well, though most are not because they are replaced by the products that they ultimately form in the mechanism (Carter, 2016). The display of products formed in these reactions (and also in single-step reactions of stable compounds if the link to generate these are chosen) are active hyperlinks, and clicking on them generates the reactions (for radicals) or reaction menu (for stable compounds) for the product formed. This way, the user can go through a sequence of reactions without having to enter the names or structures of the intermediates involved. However, this is for information purposes only; it is not the normal way that mechanisms are generated for implementation in models.

The main menu also has a link that can be used to obtain links for the types of pre-defined SAPRC detailed model species, which includes normal, branched, and cyclic alkanes, monoalkenes, etc. (currently there are a total of 35 types in all). Clicking on these links will list all compounds of these types, and the user can select the compound to get the same page as would result if the compound structure or DMS name were entered at the main menu. This method can only be used for creating the (currently) 687 stable organic compounds that have been represented in this and previous versions of SAPRC.

4.4.2. Running Single Step Reactions.

For web users can create the reactant in the main menu or from the list of SAPRC detailed model species as described above in Section 4.4.1. The reactions are automatically generated for radical species, while menus are given for reactions of stable compounds with atmospheric species. One can generate the single step reactions of the compound with the various atmospheric species (or unimolecular reactions if estimated to be non-negligible for atmospheric conditions) by selecting the appropriate link in the menu.

In either case, the web page output gives the reactions, products, an associated documentation. The organic products formed are active links, and clicking on them either generates the reactions of the product if it is a radical, or goes to the compound information page and reaction menu if it is a stable product. This way the user can go through a sequence of reactions to generate the documentation and products of potential interest.

In the previous versions of the SAPRC mechanism, an automated procedure was used to derive mechanisms for the alkanes, but molecule-by-molecule assignments or various lumping or condensation approaches were used for all the other VOCs. In this version, an automated procedure is now used to derive the mechanisms for a much wider variety of compounds, which includes almost all compounds for which mechanistic assignments have been made except for the aromatics and terpenes. This procedure, and the estimation methods and assignments that it employs, are discussed in this section.

4.4.3. (discussion of Additional Web Commands to be added)

5. REFERENCES

- Aumont B., Szopa S. and Madronich S., Modelling the evolution of organic carbon during its gas-phase tropospheric oxidation: development of an explicit model based on a self generating approach, *Atmospheric Chemistry Physics*, 5, 2497-2517, 2005.
- Carter, W. P. L. (1990): "A Detailed Mechanism for the Gas-Phase Atmospheric Reactions of Organic Compounds," *Atmos. Environ.*, **24A**, 481-518.
- Carter, W. P. L. (2000): "Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment," Report to the California Air Resources Board, Contracts 92-329 and 95-308, May 8. Available at <http://www.cert.ucr.edu/~carter/absts.htm#saprc99>.
- Carter, W. P. L. (2010a): "Development of the SAPRC-07 Chemical Mechanism and Updated Ozone Reactivity Scales," Final report to the California Air Resources Board Contract No. 03-318. January 27. Available at www.cert.ucr.edu/~carter/SAPRC.
- Carter, W. P. L. (2010b): "Development of the SAPRC-07 Chemical Mechanism," *Atmospheric Environment*, 44, 5324-5335.
- Carter, W. P. L. (2019a): "Estimates and Assignments used in the SAPRC-16 Mechanism Generation System," Draft Interim Report to California Air Resources Board Contract No. 11-761, March 4, 2019. Available on request.
- Carter, W. P. L. (2019b): "Preliminary Documentation of the SAPRC-16 Mechanism," Draft Interim Report to California Air Resources Board Contract No. 11-761, March 4, 2019. Available on request.
- Carter, W. P. L., and R. Atkinson (1985): "Atmospheric Chemistry of Alkanes", *J. Atmos. Chem.*, 3, 377-405, 1985.
- Carter, W. P. L. and G. Heo (2012): "Development of Revised SAPRC Aromatics Mechanisms," Report to the California Air Resources Board Contracts No. 07-730 and 08-326, April 12, 2012. Available at <http://www.cert.ucr.edu/~carter/absts.htm#saprc11>.
- Goliff, W. S., Stockwell, W. R. and Lawson, C. V. (2013): "The regional atmospheric chemistry mechanism, version 2," *Atmos. Environ.* 68, 174-185.
- MOO (1997): "LambdaMOO Programmer's Manual For LambdaMOO Version 1.8.0p6," currently available at <http://www.hayseed.net/MOO/manuals/ProgrammersManual.html>, dated March, 1997.
- MOO (2014): "MOO-Cows FAQ," by Ken Fox, available at <http://www.moo.mud.org/moo-faq/moo-faq.html#toc1>, last updated October 8, 2014.
- MOO (2016): Wikipedia entry for LambdaMoo, available at <https://en.wikipedia.org/wiki/LambdaMOO>, dated May, 2016.
- Stockwell, W.R., F. Kirchner, M. Kuhn, and S. Seefeld (1997): "A new mechanism for regional atmospheric chemistry modeling," *J. Geophys. Res.*, 102, 25847-25880.

Yarwood, G., S. Rao, M. Yocke, and G. Whitten (2005): "Updates to the carbon bond chemical mechanism: CB05. Final report to the U.S. EPA, RT-0400675. Available at http://www.camx.com/publ/pdfs/CB05_Final_Report_120805.pdf.

6. APPENDIX A. LIST OF MOO OBJECTS

Table 7. List of MOO objects used in the MechGen system, indicating numbers of programs and properties associated with each and the total numbers of lines of codes in the programs.

Object internal ID and level [a]	Verbs and Props [b]			Description
	nV	nVC	nP	
0 mech_master	43	899	48	Master Chemical Mechanism Generic and Database Object
1 mech_Atoms	8	153	12	Parent of all objects that have atoms associated with them
2 mech_group	45	1119	107	Generic Group
3 radical	5	63	8	Generic Radical Center
4 alkoxy (see Table 8 for types)	32	2559	49	Generic Alkoxy Radical
4 peroxy (see Table 8 for types)	24	1081	12	Generic Peroxy Radical
4 alkyl (see Table 8 for types) (See Table 8 for other types)	11	1274	12	Generic Carbon Centered Radical
3 nonrad (see Table 8 for types)	-	-	-	Generic non-radical group
3 undefined_group	1	14	-	Undefined Group
3 null_g	-	-	-	null group
3 syn_ind	-	-	-	Group used to indicate the syn/anti position relative to the neighboring groups
2 reactprod	11	93	13	Generic Reactant or Product
3 Mech_product (see Table 9 for types)	10	152	11	Generic Unreacting Product
3 reactant	149	6872	80	Generic Reactant Molecule and/or Radical
1 mech_pool	6	173	-	Object Pool for Chemical Mechanism Generation Chambers
1 fullmech	10	289	7	Generic Full Mechanism Generator
1 reactant_container	76	2955	17	Generic Reactant Storage Container (for bulk processing)
2 fullmech_reactant_container	-	-	-	Generic Reactant Container used by Fullmech
1 Mechgen_Util	29	1653	4	General utility verbs for mechanism generation system.
1 fullmech_cont	12	142	1	Generic Full Mechanism Generator Container
1 MGdb	11	254	6	Generic MechGen Database
2 AtomDB	2	51	84	This database object contains information on all types of atoms.
2 thermo	20	678	6	Thermochemical Group Values for Heats of Formation (generic)
3 thermo15	4	193	1115	SAPRC 15 Thermochemical Group Values for Heats of Formation

Table 7 (continued)

Object internal ID and level [a]	Verbs and Props [b]			Description
	nV	nVC	nP	
2 Mechgen_help	6	85	3	Mechgen Help Database
2 kest	31	2487	27	VOC + abstracting radical rate constant estimation database (generic)
3 S15est	-	-	-	SAPRC-15 VOC + abstracting radical rate constant estimation database
2 gen_mech_asn_set	36	1334	20	Generic Set of Mechanistic Assignments or Assumptions
3 SAPRC16_mech_asn	-	-	482	SAPRC-16 Mechanism Assignments
2 raddb	11	349	3	Generic organic radical rate constant estimation database
3 S15rad	-	-	-	SAPRC-15 organic radical rate constant estimation database
2 orgdb	9	155	3	Organic Compound name Database
3 DMS_db	29	538	849	Detailed Model Species Database
1 GenLump	80	6748	33	Generic Lumping Rules
2 SAPRC16	-	-	-	SAPRC-16 Lumping
1 MechWeb	32	949	18	Mechanism Generation System Web Interface
1 mixdb	13	1108	19	Generic defined mixture
0 Mech_Est	-	-	363	Core Option Set Mech_Est Database (lists core MechGen objects)
0 mech_chamber	164	4583	74	Generic Chemical Mechanism Reaction Chamber
1 Mech_Web_Chamber	47	1098	18	Generic Web-Enabled Chemical Mechanism Reaction Chamber
0 Mech_owner	75	2214	17	Non-wizard MechGen programmer object (named "MechGen")
0 gen_pooled	4	25	2	Generic Pooled Object (unused but available for use).

[a] The level number in the first column and the indentation of the object ID in the second indicate which objects are descendants of of which, where an object inherits the programs and properties associated with its parent. Level 0 means the parent of the object is not part of the core MOO system, and level i means that the parent object is level i-1. The parent of an object with level >0 is the first object above it with a lower level or a lower amount of indentation.

[b] nV = number of verbs (programs) associated with the object; nVC = total number of lines of verb code in all the programs; and nP = number of properties defined on the object. (Note that the object also inherits the properties of its parents, which may be given different values than on the parent.). The totals for objects listed on this table are 1046 programs, 42,340 lines of program code, and 3523 properties defined on these objects. Note that each object inherits programs and properties of its parents, and though either can be over-written with different programs or values.

Table 8. List of MOO objects used to specify individual types of radical or non-radical groups.

Object internal ID	and level [a]	Verbs and Props [b]			Description
		nV	nVC	nP	
<u>Alkoxy Radical Groups</u>					
5	CH2O_rad	-	-	-	-CH2O.
5	CH3O_rad	-	-	-	CH3O.
5	acyl_oxy_rad	1	20	-	-CO2.
5	CH[O_rad]	-	-	-	-CH[O.]
5	C[O_rad]	-	-	-	>C[O.]
5	HCO2_rad	1	9	-	HCO2.
5	NH_Orad	-	-	-	NH[O.]
5	N_Orad	-	-	-	N[O.]
<u>Peroxy Radical Groups</u>					
5	acyl_peroxy_rad	1	4	-	-CO[OO.]
5	CH3OO_rad	-	-	-	CH3OO.
5	CH2[OO_rad]	-	-	-	-CH2OO.
5	CH[OO_rad]	-	-	-	-CH[OO.]
5	C[OO_rad]	-	-	-	>C[OO.]
5	CH_db_OO_rad	1	35	-	=CHOO.
6	C_db_OO_rad	-	-	-	=COO.
5	PhenPeroxyRad	-	-	-	Phen[OO.]
<u>Alkyl Radical Groups</u>					
5	CH3_rad	-	-	-	CH3.
5	CH2[rad]	-	-	-	-CH2.
6	CH2[aRad]	-	-	-	-aCH2.
5	C[rad]	-	-	-	>C[.]
6	C[aRad]	-	-	-	>aC[.]
5	Acyl_rad	2	69	1	-CO.
5	CH[rad]	-	-	-	-CH[.]
6	CH[aRad]	-	-	-	-aCH[.]
5	vinyl	1	74	4	vinyl
6	CH_dbond_rad	-	-	-	=CH.
7	CH_dbond_arad	-	-	-	=aCH.
6	C_dbond_rad	-	-	-	=C.
7	C_dbond_arad	-	-	-	=aC[.]
5	PhenRad	-	-	-	Phen[.]
5	Formyl_rad	1	8	-	HCO.
<u>Other Radical Types and Groups</u>					

Table 8 (continued)

	Object internal ID and level [a]	Verbs and Props [b]			Description
		nV	nVC	nP	
4	Oxy_rad	-	-	-	Oxy Radical Center
4	H_atom	1	8	-	H.
4	peroxy_rad	-	-	-	-OO.
4	dbC[O_rad]	1	12	-	=C[O(rad)]-
5	dbCH[O_rad]	-	-	-	=CH[O(rad)]
5	PhenoxyRad	1	3	-	Phen[O.]
4	amino	2	80	4	Generic Amino Radical
5	NH_rad	-	-	-	NH.
5	N_rad	-	-	-	N.
4	Nx_O_g	1	21	-	Nx[O]
5	NH2_O_g	-	-	-	NH2[O]
5	NH_O_g	-	-	-	NH[O]
4	HO2_g	-	-	-	[HO2]
4	carbene	1	10	-	Generic Carbene
5	methylene_carbene	-	-	-	CH2:
5	monosub_carbene	-	-	-	C2:
5	disub_carbene	-	-	-	C:
4	Criegee	2	415	8	Generic Criegee Biradical
5	CH2OO_cr	2	139	-	CH2OO
5	COO_cr	2	304	-	-C[OO]-
5	CHOO_cr	2	69	-	-CHOO
6	CHOOsyn	1	157	-	-CHOO[syn]
6	CHOOanti	1	15	-	-CHOO[anti]
<u>Nonradical Groups</u>					
4	F_Group	-	-	-	-F
4	Br_Group	-	-	-	-Br
4	I_Group	-	-	-	-I
4	Cl_group	-	-	-	-Cl
4	Nitro_group	-	-	-	-NO2
4	phot_carb_g	3	309	1	Photolyzable Carbonyl
5	aldehyde_g	5	175	-	-CHO
6	formate	-	-	-	HCO-O-
6	dcaldHCO	-	-	-	HCO-CO-
5	carbonyl_g	3	264	-	-CO-
6	EsterCO	-	-	-	-CO-O-
7	carbonateCO	-	-	-	carbonate-CO
6	acidCO	-	-	-	-CO-OH

Table 8 (continued)

	Object internal ID and level [a]	Verbs and Props [b]			Description
		nV	nVC	nP	
7	peroxyacidCO	-	-	-	-CO-O-OH
6	dcbketCO	-	-	-	-CO-CO-
7	dcbacidCO	-	-	-	CO(CO)-OH
6	ketone_g	-	-	-	-CO-R
4	Nitroso_group	-	-	-	-NO
4	Carbonyl_O	-	-	-	=O
4	nitrite_group	-	-	-	-ONO
4	alkane_g	-	-	-	-CHx
5	CH3_g	-	-	-	-CH3
6	CH3_on_aro	-	-	-	CH3(aro)
6	CH3_NX_g	-	-	-	CH3(Nx)
7	CH3_NH2_g	-	-	-	CH3(NH2)
7	CH3_NH_g	-	-	-	CH3(NH)
7	CH3_N_g	-	-	-	CH3(N)
6	CH3_OX_g	-	-	-	CH3(Ox)
6	CH3_OH_g	-	-	-	CH3(OH)
6	CH3_ONO2_g	-	-	-	-CH3-ONO2
6	CH3_OOH_g	-	-	-	CH3(OOH)
5	CH2_g	-	-	-	-CH2-
6	CH2_on_aro	-	-	-	CH2(aro)
6	CH2_ONO2_g	-	-	-	-CH2-ONO2
6	CH2_NX_g	-	-	-	CH2(Nx)
7	CH2_NH2_g	-	-	-	CH2(NH2)
7	CH2_NH_g	-	-	-	CH2(NH)
7	CH2_N_g	-	-	-	CH2(N)
6	CH2_OX_g	-	-	-	CH2(Ox)
6	CH2_OH_g	-	-	-	CH2(OH)
6	CH2_OOH_g	-	-	-	CH2(OOH)
5	CH_g	-	-	-	>CH-
6	CH_on_aro	-	-	-	CH(aro)
6	CH_ONO2_g	-	-	-	>CH-ONO2
6	CH_NX_g	-	-	-	CH(Nx)
7	CH_NH2_g	-	-	-	CH(NH2)
7	CH_NH_g	-	-	-	CH(NH)
7	CH_N_g	-	-	-	CH(N)
6	CH_OX_g	-	-	-	CH(Ox)
6	CH_OH_g	-	-	-	CH(OH)
6	CH_OOH_g	-	-	-	CH(OOH)

Table 8 (continued)

Object internal ID and level [a]	Verbs and Props [b]			Description	
	nV	nVC	nP		
5	tertiary_C_g	-	-	-	>C<
6	C_ONO2_g	-	-	-	>C-ONO2-
6	C_OX_g	-	-	-	C(Ox)
6	C_OH_g	-	-	-	C(OH)
6	C_NX_g	-	-	-	C(Nx)
6	C_OOH_g	-	-	-	C(OOH)
4	O_g	-	-	-	O-g
5	alcohol_g	2	59	-	-OH
6	hydroperoxy_g	1	19	-	HO-O-
6	acidOH	-	-	-	HO-CO-
7	peroxyacidOH	1	19	-	HO-O-CO-
6	OH_on_amine	-	-	-	OH[Nx]
6	Phenol_grp	-	-	-	HO-aC<
5	O_notH_grp	-	-	-	-O-X
6	ether_g	-	-	-	-O-
7	peroxy_g	1	43	-	-O-O-
8	peroxyacidO	-	-	-	-O-OH
8	hydroperoxy_ether	-	-	-	-O(OH)
8	O2NO2_O	-	-	-	O[O2NO2s]
9	PAN_O	1	52	-	-O(CO)(ONO2)
10	dcPAN_O	-	-	-	O[dcPANs]
6	EsterO	-	-	-	-O(CO)
7	carbonateO	-	-	-	-O-CO-O-
7	formateO	-	-	-	-O-CHO
4	amino_g	-	-	-	-NHx
5	t_amino_group	-	-	-	-N<
5	s_amino_grp	-	-	-	-NH-
5	p_amino_grp	-	-	-	-NH2
4	NO3_group	1	15	2	-ONO2
4	HCHO_g	-	-	-	[HCHO]
4	CO2_g	-	-	-	[CO2]
4	O2_g	-	-	-	[O2]
4	NO3rad_g	-	-	-	[NO3]
4	NO2_g	-	-	-	[NO2]
4	arogrp	4	110	2	-Aro<
5	AroC	-	-	-	aC
5	AroCH	-	-	-	aCH
5	PAHC	-	-	-	pahC

Table 8 (continued)

	Object internal ID and level [a]	Verbs and Props [b]			Description
		nV	nVC	nP	
4	N_O_g	-	-	-	-N[O]-
4	S_grp	-	-	-	-S-
4	SH_grp	-	-	-	-SH
4	SO_grp	-	-	-	-SO-
4	S_dbond_grp	-	-	-	=S
4	Si_grp	-	-	-	>Si<
4	SiH_grp	-	-	-	-SiH<
4	SiH2_grp	-	-	-	-SiH2-
4	SiH3_grp	-	-	-	-SiH3
4	SiO_grp	-	-	-	-SiO-
4	Si_dbond_grp	-	-	-	=Si<
4	N_tbond_grp	-	-	-	#N
4	HaloAlk_g	-	-	-	HaloAlk
5	CF3_g	-	-	-	-CF3
5	CF2_g	-	-	-	-CF2-
5	CCl3_g	-	-	-	-CCl3
5	CBr3_g	-	-	-	-CBr3
5	CF2Cl_g	-	-	-	-CF2Cl
5	CF2Br_g	-	-	-	-CF2Br
5	CFCl2_g	-	-	-	-CFCl2
5	CCl2Br_g	-	-	-	-CCl2Br
5	CFBr2_g	-	-	-	-CFBr2
5	CClBr2_g	-	-	-	-CClBr2
5	CFClBr_g	-	-	-	-CFClBr
5	CCl2_g	-	-	-	-CCl2-
5	CBr2_g	-	-	-	-CBr2-
5	CFCl_g	-	-	-	-CFCl-
5	CFBr_g	-	-	-	-CFBr-
5	CClBr_g	-	-	-	-CClBr-
4	HaloAlkH_g	-	-	-	HaloAlkH
5	CHF2_g	-	-	-	-CHF2
5	CHCl2_g	-	-	-	-CHCl2
5	CHBr2_g	-	-	-	-CHBr2
5	CHFC1_g	-	-	-	-CHFC1
5	CHFBr_g	-	-	-	-CHFBr
5	CHClBr_g	-	-	-	-CHClBr
4	N_D2	-	-	-	-N(=)= (Nitrogen group with two dative bonds)

Table 8 (continued)

Object internal ID and level [a]	Verbs and Props [b]			Description
	nV	nVC	nP	
4 SO2_grp	-	-	-	-SO2-
3 gen_unsat_group	4	75	15	Generic Unsaturated Carbon Center
4 allene_grp	-	-	-	=C=
5 Allene_mid_C_s1	-	-	-	C(=CH2)(=CH)
5 Allene_mid_C_s2u	-	-	-	C(=CH2)(=C)
5 Allene_mid_C	-	-	-	C(=CH2)2
5 Allene_mid_C_S2s	-	-	-	C(=CH)(=CH)
5 Allene_mid_C_s3	-	-	-	C(=CH)(=C)
5 Allene_mid_C_s4	-	-	-	C(=C)2
4 gen_alkyne_grp	2	41	-	#CHx
5 C_tbond_grp	-	-	-	#C
6 C_tbond_CH_grp	-	-	-	C#(CH)
7 Cme_tbond_CH_grp	-	-	-	C(CH3)#(CH)
6 C_tbond_C_grp	-	-	-	C#(C)
7 Cme_tbond_C_grp	-	-	-	C(CH3)#(C-)
8 Cme_tbond_Cme_grp	-	-	-	C(CH3)#(C-CH3)
7 C_tbond_Cme_grp	-	-	-	C#(C-CH3)
6 Cme_tbond_grp	-	-	-	#C(CH3)
5 CH_tbond_grp	-	-	-	#CH
6 CH_tbond_CH_grp	-	-	-	CH#(CH)
6 CH_tbond_C_grp	-	-	-	CH#(C)
7 CH_tbond_Cme_grp	-	-	-	CH#(C-CH3)
4 N_dbond_grp	-	-	-	-N=
4 NH_dbond_grp	-	-	-	=NH
4 alkene_g	-	-	1	=CHx
5 CH2_dbond_grp	-	-	-	=CH2
6 CH2_dbond_CH_grp	-	-	-	CH2=(CH)
7 CH2_dbs_CH_Cx_CH2_grp	-	-	-	CH2=(CH-Cx=CH2)
7 CH2_dbs_CH_Cx_CH_grp	-	-	-	CH2=(CH-Cx=CH)
7 CH2_dbs_CH_Cx_C_grp	-	-	-	CH2=(CH-Cx=C)
6 CH2_dbond_C_grp	-	-	-	CH2=(C)
7 CH2_dbs_C_Cx_CH2_grp	-	-	-	CH2=(C-Cx=CH2)
7 CH2_dbs_C_Cx_CH_grp	-	-	-	CH2=(C-Cx=CH)
7 CH2_dbs_C_Cx_C_grp	-	-	-	CH2=(C-Cx=C)
6 CH2_dbond_CH2_grp	-	-	-	CH2=(CH2)

Table 8 (continued)

Object internal ID and level [a]	Verbs and Props [b]			Description	
	nV	nVC	nP		
6	CH2_dbond_allene_grp	-	-	-	CH2=(C=)
5	CH_dbond_grp	-	-	-	=CH-
6	CH_dbond_CH2_grp	-	-	-	CH=(CH2)
6	CH_dbond_CH_grp	-	-	-	CH=(CH)
7	CH_dbs_CH_Cx_CH2_grp	-	-	-	CH=(CH-Cx=CH2)
7	CH_dbs_CH_Cx_CH_grp	-	-	-	CH=(CH-Cx=CH)
7	CH_dbs_CH_Cx_C_grp	-	-	-	CH=(CH-Cx=C)
6	CH_dbond_C_grp	-	-	-	CH=(C)
7	CH_dbs_C_Cx_CH2_grp	-	-	-	CH=(C-Cx=CH2)
7	CH_dbs_C_Cx_CH_grp	-	-	-	CH=(C-Cx=CH)
7	CH_dbs_C_Cx_C_grp	-	-	-	CH=(C-Cx=C)
6	CH_dbond_allene_grp	-	-	-	CH=(C=)
5	C_dbond_grp	-	-	-	=C<
6	C_dbond_CH2_grp	-	-	-	C=(CH2)
6	C_dbond_CH_grp	-	-	-	C=(CH)
7	C_dbs_CH_Cx_CH2_grp	-	-	-	C=(CH-Cx=CH2)
7	C_dbs_CH_Cx_CH_grp	-	-	-	C=(CH-Cx=CH)
7	C_dbs_CH_Cx_C_grp	-	-	-	C=(CH-Cx=C)
6	C_dbond_C_grp	-	-	-	C=(C)
7	C_dbs_C_Cx_CH2_grp	-	-	-	C=(C-Cx=CH2)
7	C_dbs_C_Cx_CH_grp	-	-	-	C=(C-Cx=CH)
7	C_dbs_C_Cx_C_grp	-	-	-	C=(C-Cx=C)
6	C_dbond_allene_grp	-	-	-	C=(C=)

[a] See footnote [a] in Table 8.

[b] See footnote [b] in Table 8.

Table 9. List of MOO objects used to specify individual types of non radical product compounds or counter species.

Object internal ID and level [a]	Verbs and Props [b]			Description	
	nV	nVC	nP		
4	R2O2	-	-	-	NO to NO2 conversion
4	CO	-	-	-	Carbon Monoxide
4	HCHO	-	-	-	HCHO
4	CH4	-	-	-	Methane
4	CO2	-	-	-	CO2
4	null_p	-	-	-	No product at all
4	H2O	-	-	-	H2O
4	rxnspec	6	108	3	Primary reacting species
5	HO2	1	10	-	HO2
5	absrad	3	279	4	Abstracting radical
6	I_Atom	-	-	-	Iodine Atom
6	Br_Atom	-	-	-	Br Atom
6	F_Atom	-	-	-	F Atom
6	Cl_Atom	-	-	-	Cl Atom
6	hydroxyl	-	-	-	Hydroxyl
6	NO3_rad	1	10	-	NO3
5	Ozone	2	141	1	O3
5	O3P	5	291	2	O3P
5	HV	6	357	4	HV
5	NO	1	10	-	NO
5	O2	-	-	-	O2
5	NO2	1	10	-	Nitrogen Dioxide
5	LumpRO2S	1	10	-	Lumped organic peroxy radicals
5	unirect	1	10	-	Unimolecular reactor
5	LumpRCO3s	1	10	-	Lumped acyl peroxy radicals
4	HNO3	-	-	-	HNO3
4	HCl	-	-	-	HCl
4	HBr	-	-	-	HBr
4	HF	-	-	-	HF
4	hi	-	-	-	HI
4	R2O3	-	-	-	NO2 to NO3 conversion operation
4	RO2RO	-	-	-	RO2. to RO. conversion
4	RO2ROH	-	-	-	RO2. to ROH conversion
4	RO2RCO	-	-	-	RO2. to RCHO or RCOR' conversion
4	RCO3RO	-	-	-	RCO3. to RCO2. conversion
4	RCO3ROH	-	-	-	RCO3. to RCO-OH conversion

Table 9 (continued)

Object internal ID and level [a]		Verbs and Props [b]			Description
		nV	nVC	nP	
4	HNO	-	-	-	HNO
4	NH3	-	-	-	NH3
4	H2	-	-	-	H2

[a] See footnote [a] in Table 8.

[b] See footnote [b] in Table 8.