STRUCTURE-ACTIVITY RELATIONSHIPS (SARS)

SARs attempt to describe the reactivity trends in a reaction class
- Rate coefficients / photolysis frequencies / equilibrium constants
- Product distributions / quantum yields / population fractions

Goal: accurate and easy prediction of kinetic parameters where no direct data is available

SARs are often used to help derive (semi)explicit kinetic mechanism
  MCM, SAPRC, GeckoA, ... (see also Webinar on mechanisms)

Other uses:
- Rationalization of theoretical and experimental results
- Reduction of detail to the main reactivity trends and its underlying chemical reasons
Problem domains have become more complex and more subtle

E.g. atmospheric chemistry: highly oxygenated molecules, aerosols, low-yield products
Trend towards more detailed mechanisms to capture the details and lower-yield products
Requires more broadly applicable SARs, with improved accuracy

(See also: Vereecken et al., Int. J. Chem. Kin. 50, 435-469, 2018, "Perspective on ...mechanisms ..and SARs...")

Deriving an "accurate and broadly applicable" SAR implies:

- Wide range of training data
- Often many SAR parameters
- More rigorous validation

Example: Peroxy radical H-migration SAR:
How to formulate SARs for human/software use?

- **Molecular structure**: backbone, functionalities, hetero-atoms, ...
  This includes group-additivity schemes, graph-based prediction schemes

- **Condensed structural indices**: degree of unsaturation, C:O:H, Randić index, ...

- **Molecular properties**: Ionisation potential, dipole moment, electron affinity,...

- **Quantum chemical properties**: HOMO/LUMO, Fukui indices, hardness/softness,...

Easiest to use are SARs based solely on:

- Molecular structure / graph (draw and count)
- Easy mathematics and clear applicability
- No "chemists' judgment" needed

Example: alkoxy radical decomposition:

\[ E_b = 17.8 \text{ kcal mol}^{-1} + \sum F_{\text{substituents}} \]

\[ k(298 K) = A \times \exp \left( -\frac{E_b}{kT} \right) \]

Vereecken and Peeters, PCCP 11, 9062-9074 (2009)
Where to get training and validation data sets

**Experimental data:**
- Extensive data sets available for e.g. VOC + OH / NO$_3$ / Cl / O$_3$ /...
- Less direct data available for reactive intermediates: R•, RO•, RO$_2$•, SCI, ...
- Less information on temperature and pressure dependence, bulk vrs. elementary,

**Theoretical data:**
- Easier access to data on reactive intermediates that are experimentally hard
- Ability to do systematic series, e.g. number and type of substitutions, reaction spans,
- Access to T-dependence, P-dependence harder but possible

**Ideal world: Combining experimental and theoretical data**
- Theo&Exp are highly complementary, with different strengths and weaknesses
- Synergetic: Experimental reference data allows cheaper theoretical calculations

Next slides: recent SARs based on (mostly) theoretical data
THEORY-BASED SAR FOR RO$_2$ ISOMERISATION

H-migration in peroxy radicals (autoxidation):

ROO$^\cdot$ → HOOQ$^\cdot$ → HOOQOO$^\cdot$ (can be repeated)

Important for: OH regeneration, formation of highly oxidized molecules (HOM), ...


Wide range of spans and substitutions (see drawing), easy to use (lookup tables)

Using near-exclusively theoretical data on 100s of reactions (original and literature)

Good agreement with scarce experiments (typically factor 2)

Improvement and validation needed
Ring closure in unsaturated peroxy radicals (also auto-oxidation)

Important in biogenic RO$_2$ (e.g. multi-unsaturated VOC such as isoprene, terpenoids,...)

\[
\begin{align*}
\text{\textbullet}\text{OO} & \rightarrow \text{\textbullet O} - \text{O} + \text{\textbullet O} - \text{O} \\
& \text{ (k(298K) up to } 10^3 \text{ s}^{-1})
\end{align*}
\]

SAR: Vereecken, Vu, and Nguyen (in preparation, 2020)

non-cyclic unsaturated RO$_2$, endo/exo-cyclic product radical, alkyl substituents only

- OO$^\cdot$: 0-2 alkyl substituents
- 4- to 8-membered rings
- Spectator alkyl substituents
- C=C : 0-3 alkyl substituents

No direct experimental validation data available, to our knowledge

Indirect validation e.g. for β-pinene + OH (Vereecken et al. 2004, 2012; Kaminski et al. 2017)
REACTIONS OF CRIEGEE INTERMEDIATES

Unimolecular isomerisation / decomposition of SCI

Important intermediates in ozonolysis reaction
Yields oxygenates, OH, ring structures,..
Reactions are often faster than any bimolecular reaction

SAR : Vereecken et al. (PCCP 19, 31599-31612, 2017)
Wide range of aliphatic, unsaturated, oxygenated substituents
Easy to use: stereo- and site-specific lookup tables
Good agreement with recent theoretical work
most recent experiments (factor ~2)
SCI + H₂O / (H₂O)₂ : SAR Vereecken et al. (PCCP 19, 31599-31612, 2017)

Water reaction important mostly for $E$-RCHO SCI

Good agreement with recent experimental and theoretical data (~1 order of magnitude)

**SAR update:** Vereecken (in preparation, 2020)

Focusing on multi-unsaturated compounds (terpenoids) and their 1$^{\text{st}}$ generation products

Reactions of unsaturated SCI: refinement of SAR

Reaction of oxygenated SCI: –OH, –OOH, –C(=O)OH, –OR, –ROR, epoxy, –ONO₂

Extremely fast ring closure reactions (OH, OOH, COOH, small to large rings, see below)

H₂O: Water-catalyzed unimolecular reaction **AND** oxy-substituent-catalysed water reaction
REATIONS OF ALKOXY RADICALS

Unimolecular decomposition and isomerisation

Important after $\text{RO}_2 + \text{NO} / \text{RO}_2 / \text{HO}_2$

Isomerisation (H-shift) can lead to autoxidation; fragmentation leads to smaller fragments.

Relevant for $\text{RO}_2$ measurement with LIF (see talk Novelli et al.)

SAR: Vereecken and Peeters (PCCP 2009, 2010)

- Original SAR based on lower-level calculations (B3LYP)
- Validation against higher-level work
- CCSD(T)//M06-2X: good agreement
- Deeper validation against literature study:
  → ongoing work
SAR update I: define 2-substituent parameters

Decomposition / isomerisation leading to secondary fragmentation:

$\text{ONO}_2^\bullet \rightarrow \text{H}_2\text{C}=\text{O} + \text{ONO}_2 + \text{NO}_2$

→ Update for nitrate-substituted RO

Interaction between substituents

Simple additive model does not work

→ update with second set of terms

$E_b = 17.8 \text{ kcal mol}^{-1} + \sum F_{\text{substituents}} + \sum F^{\text{NO}_2}_{\text{substituents}}$

(Experimental validation: see talk Novelli et al.)

see Novelli et al. (PCCP, submitted)

Vereecken et al. (in preparation, 2020)
SAR update II: epoxidation important in unsaturated RO:

$\text{Update for epoxy-substituted RO: decomposition, H-shift isomerisation}$

$\text{fast reactions: } k(298K) \approx 10^8 \text{ s}^{-1}$

$\text{reversible at similar speeds:}$

$\text{competes with } O_2 \text{ addition on radical site}$

$\text{epoxy group leads to slow reactions when in active site}$

$\text{(see Novelli et al., PCCP, submitted)}$

Application example: isoprene + NO$_3$ chemistry: epoxidation affect fate of alkoxy radicals

$\text{see Vereecken et al. (in preparation, 2020)}$
CONCLUSIONS

Theoretical kinetic studies
- Allow characterization of systematic series of reaction classes
- Can have strong predictive capabilities for rate coefficients across a range of reactions

Higher-level calculations on reactions already covered by SARs
- Enable validation of existing SARs outside training data set

Recent new and updated theory-based SARs:
- H-migration in RO₂ radicals
- Ring closure reactions in unsaturated RO₂ radicals
- Unimolecular and water reactions of unsaturated and oxygenated Criegee intermediates
- Nitrated alkoxy radicals
- Epoxidized alkoxy radicals