

DE LA RECHERCHE À L'INDUSTRIE



PROPRIÉTÉS PHYSICO-CHIMIQUES DES COMPOSÉS ORGANIQUES: MODÈLES EXPLOITANT THÉORIES PHYSIQUES ET DONNÉES EMPIRIQUES

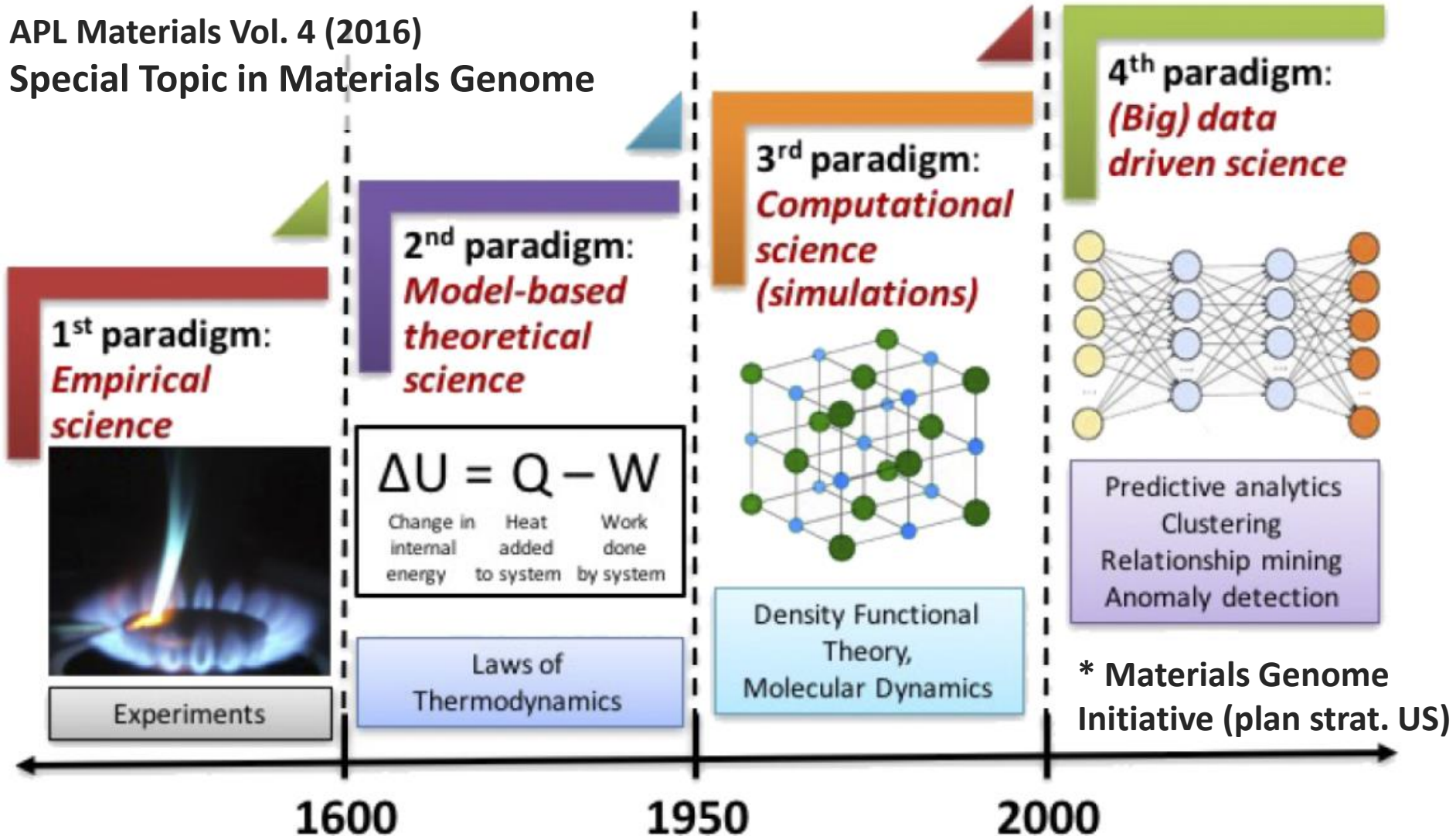
Didier MATHIEU
CEA – Le Ripault
F-37260 Monts (France)
didier.mathieu@cea.fr

26e journée du projet CASCIMODOT
Données de terrain et modèles, une interaction à double sens

Faculté de médecine, site Tonnelé, Amphi B
Tours, 22 juin 2017

Progrès en Sciences des Matériaux : le point de vue des acteurs du MGI*

APL Materials Vol. 4 (2016)
Special Topic in Materials Genome

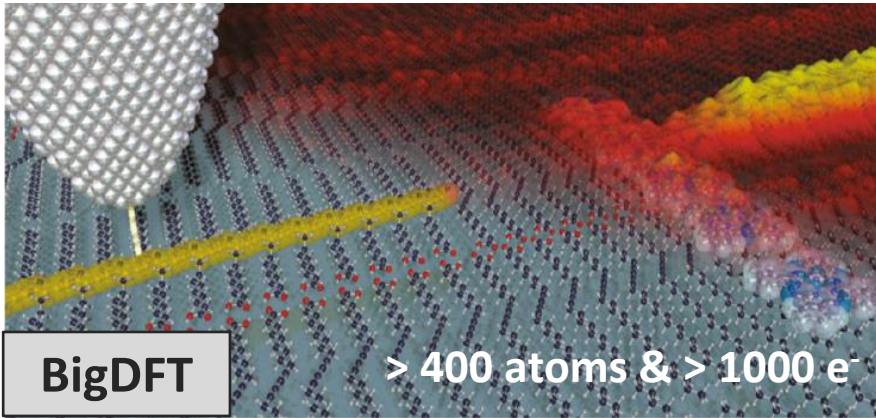


Cet exposé : paradigme #2 - modélisation des matériaux : souvent selon #3 / #4.

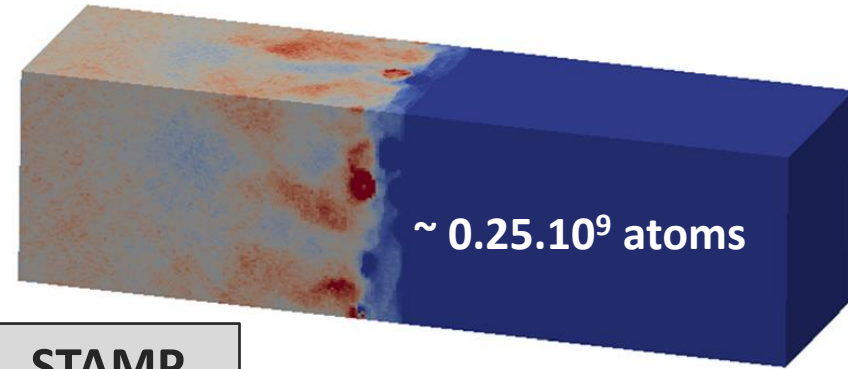
Modèles informatiques basés sur : 1) la théorie ou 2) les données

Exemples parmi les travaux menés au CEA

Theory-driven



Propriétés électroniques de nanocircuits



STAMP

Ondes de choc dans milieu poreux

PHYSICAL REVIEW X 4, 011019 (2014)

Data driven

Finding Unprecedentedly Low-Thermal-Conductivity Half-Heusler Semiconductors via High-Throughput Materials Modeling

Jesús Carrete, Wu Li, and Natalio Mingo^{*}
CEA-Grenoble, 17 Rue des Martyrs, Grenoble 38054, France

Shidong Wang, and Stefano Curtarolo[†]
*Center for Materials Genomics, Materials Science, Electrical Engineering, Physics and Chemistry,
Duke University, Durham, North Carolina 27708, USA*

(Received 26 August 2013; revised manuscript received 11 December 2013; published 19 February 2014)

QSAR / QSPR : techniques d'apprentissage automatique dans le domaine moléculaire

Apprentissage supervisé d'abord appliqué surtout aux propriétés biologiques :

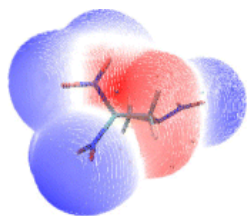
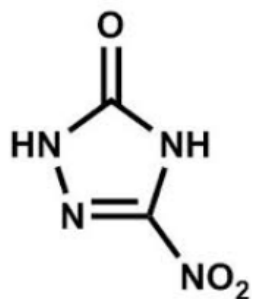
- ❑ QSAR = Quantitative Structure-Activity Relationships
- ❑ QSTR = Quantitative Structure-Toxicity Relationships

Appliqué aussi de plus en plus souvent aux propriétés physico-chimiques :

QSPR = Quantitative Structure-Property Relationships

propriété à estimer

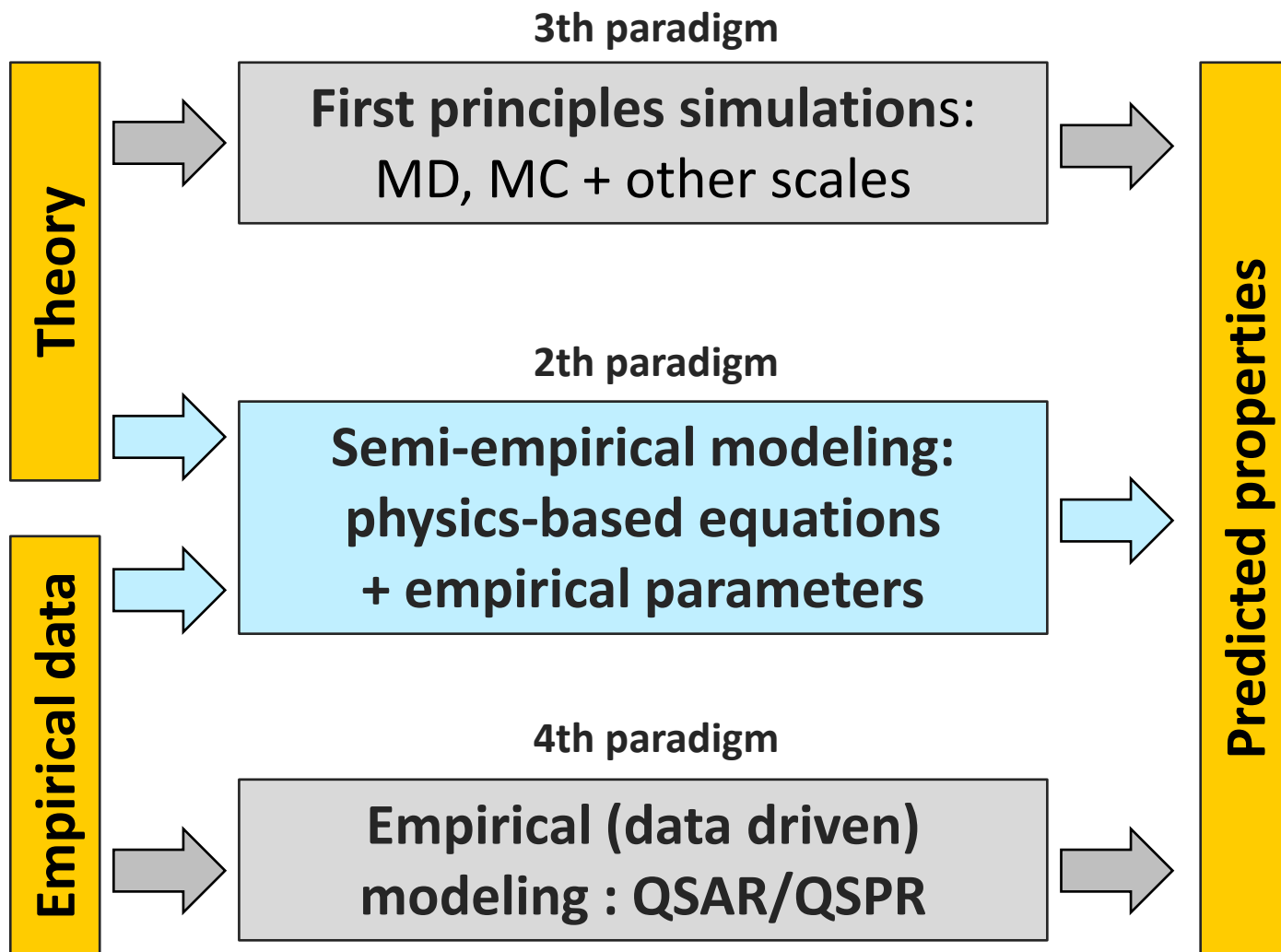
descripteurs 0D, 1D, 2D, 3D ...



	Structure	Mol. Formula	Mol. Weight	Sample Mass	Purity	Effective Mass	Effective Solides
1		C ₈ H ₁₀ N ₄ O ₂	178.1740	20g	90%	18.1g	198.0000
2		C ₁₆ H ₁₄ O ₄	294.2740	10.2g	90%	9.18g	130.0000
3		C ₁₆ H ₁₄ O ₄	294.2740	20g	90%	18.1g	130.0000
4		C ₁₆ H ₁₄ O ₄	294.2740	10.4mg	90%	9.36mg	73.0000

➔ Choix des variables + régression: MLR, PLS, Lasso, DT, RF, ANN, Cubist, GP ...

Intérêt d'une modélisation semi-empirique (approche #2) comparé aux approches #3 / #4



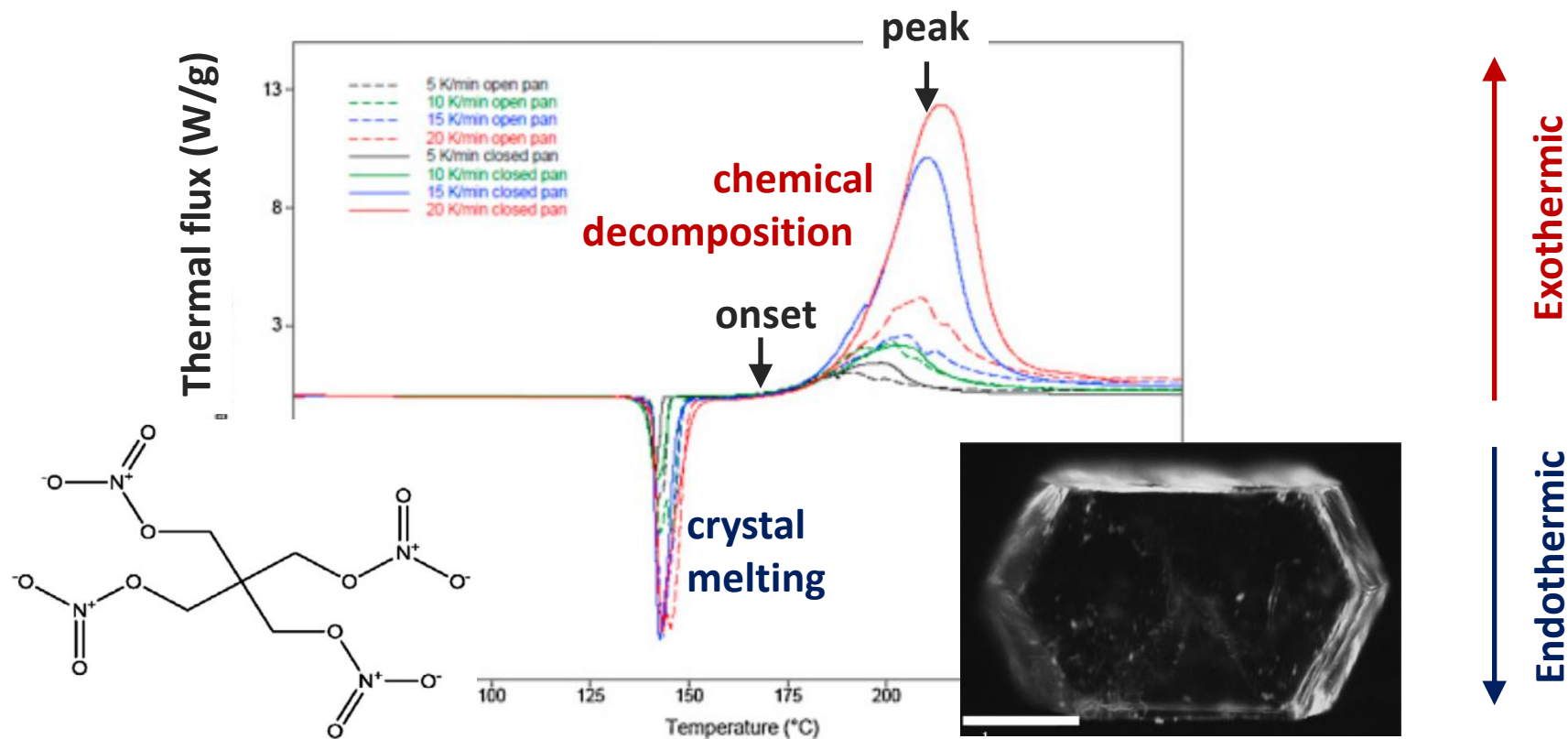
La modélisation semi-empirique exploite au mieux théorie et données

Exemples de propriétés caractérisant les substances dangereuses : 1/2

Grandeurs mesurées par analyse thermique

Exemple : DSC = Differential Scanning Calorimetry

Bhattacharia et al. *Thermochimica Acta* 617 (2015) 38–43



Grandeurs mesurées :

Chaleur de décomposition Q (hazard severity)

Température de décomposition : T_{onset} , T_{peak} ...

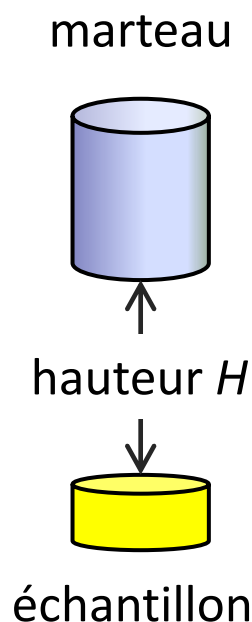
Exemples de propriétés caractérisant les substances dangereuses : 2/2

Grandeurs propres aux matériaux énergétiques (explosifs, propergols ...) :

- ❑ Performances : paramètres de détonation (énergie, célérité, pression)
paramètres de Gurney (énergie, vitesse)
- ❑ Critères de sécurité : sensibilités à l'impact, à la friction, à l'étincelle ...



Mesure de la sensibilité à l'impact



- un marteau de masse standard est lâché d'une hauteur H sur un échantillon
→ OUI/NON ?
- Ce test est répété un grand nombre de fois pour chaque hauteur H
- H_{50} = hauteur pour laquelle on détecte l'ignition (OUI) dans 50% des cas (lumière, fumée, bruit...)

Quelques difficultés pour la modélisation des propriétés de produits dangereux

- ❑ Mécanismes variés, à différentes échelles et mettant en jeu de très nombreux degrés de liberté → **simulation numérique difficile**


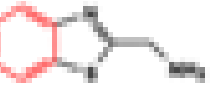
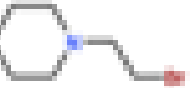
En pratique la modélisation quantitative de ces propriétés repose sur les techniques QSPR : point d'éclair, inflammabilité, sensibilités diverses ...

- ❑ Ces propriétés ne sont pas inhérentes au système étudié (plutôt à l'ensemble système/dispositif de mesure/environnement)
- ❑ Toute substance dangereuse est délicate/impossible à manipuler

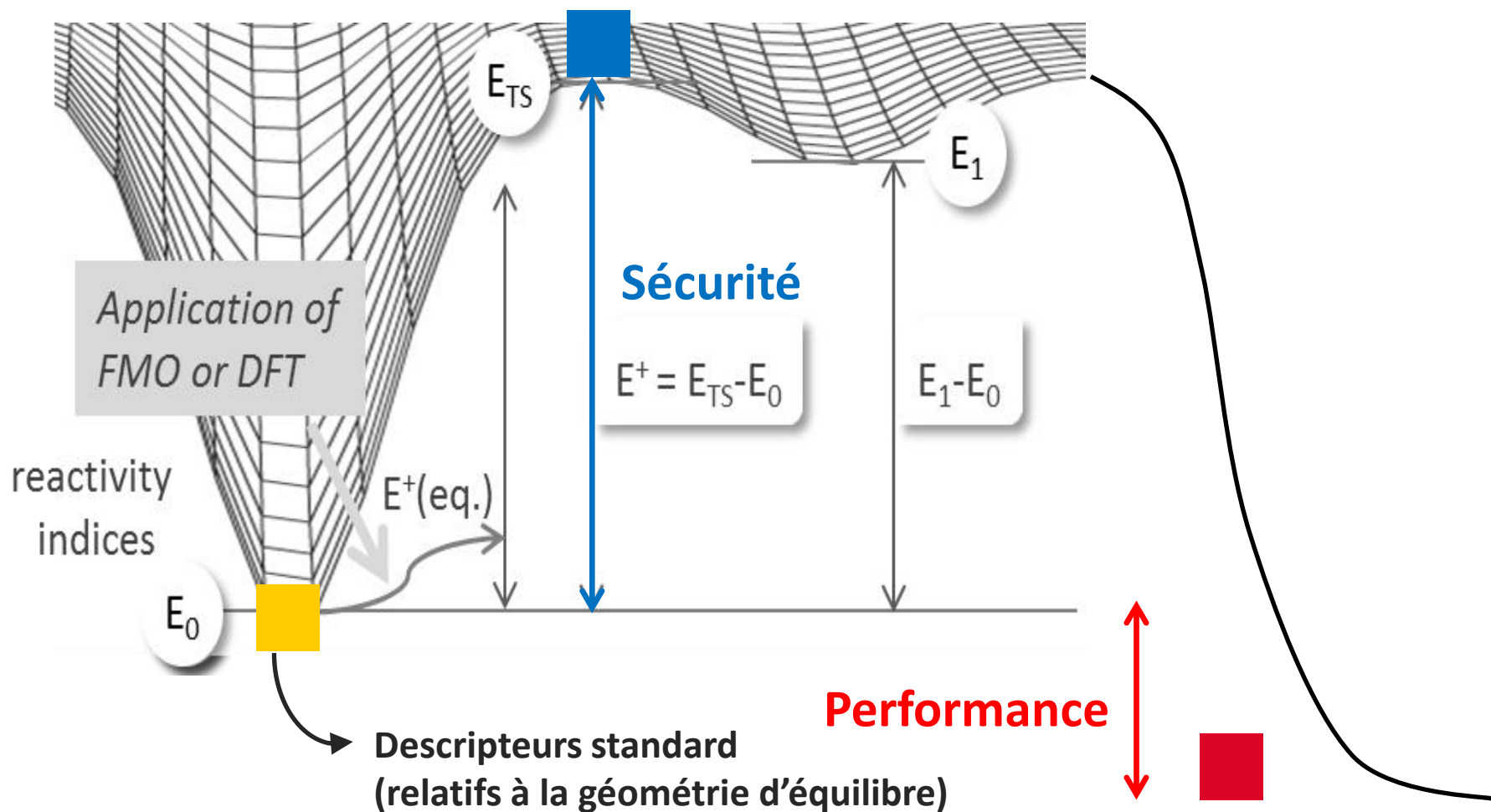
Données peu nombreuses et difficiles à acquérir



rend douteuse toute approche statistique

4		-C ₆ H ₅ CH ₂ Cl	192.6445	10.4mg	90%	14.17mg	73.5min
5		-C ₆ H ₄ Cl	144.2272				
6		-C ₆ H ₁₁ Br	192.6942				

Autre raison de l'échec des modèles QSPR : descripteurs inadaptés



Nécessité de prendre en compte énergies d'activation et de réaction : états de transition et produits de décomposition

Some guiding principles to develop semi-empirical models

- Models equations established on the basis of physical models (unlike QSPR)
- Quantities not amenable to theoretical prediction either assumed constant or estimated empirically (maybe using QSPR)
- Empirical parameters fitted against experiment or (preferably) first-principles computations (extensive data making QSPR easier)
- Extensive validation against experiment using external test sets (like QSPR models)

Considérations physiques a posteriori : « réparation » des modèles QSPR

Illustration : cas d'un modèle QSPR pour estimer l' « explosivité » de composés assimilée ici à la chaleur de décomposition molaire ΔH_{dec}

Chemical Physics Letters 467 (2009) 407–411

$$-\Delta H_{\text{dec}}^{\text{theor}} = -16.9EA + 18.1\omega - 3.6\Delta E_{\text{diss}}^{\text{theor}} + 263.8$$



$$R^2 = 0,85 \quad Q^2 = 0,73$$

Grandeur « size-consistent »

PROBLEME : aucune des variables n'est « size consistent »

Remédier à ce problème via une simple multiplication par la masse molaire ...

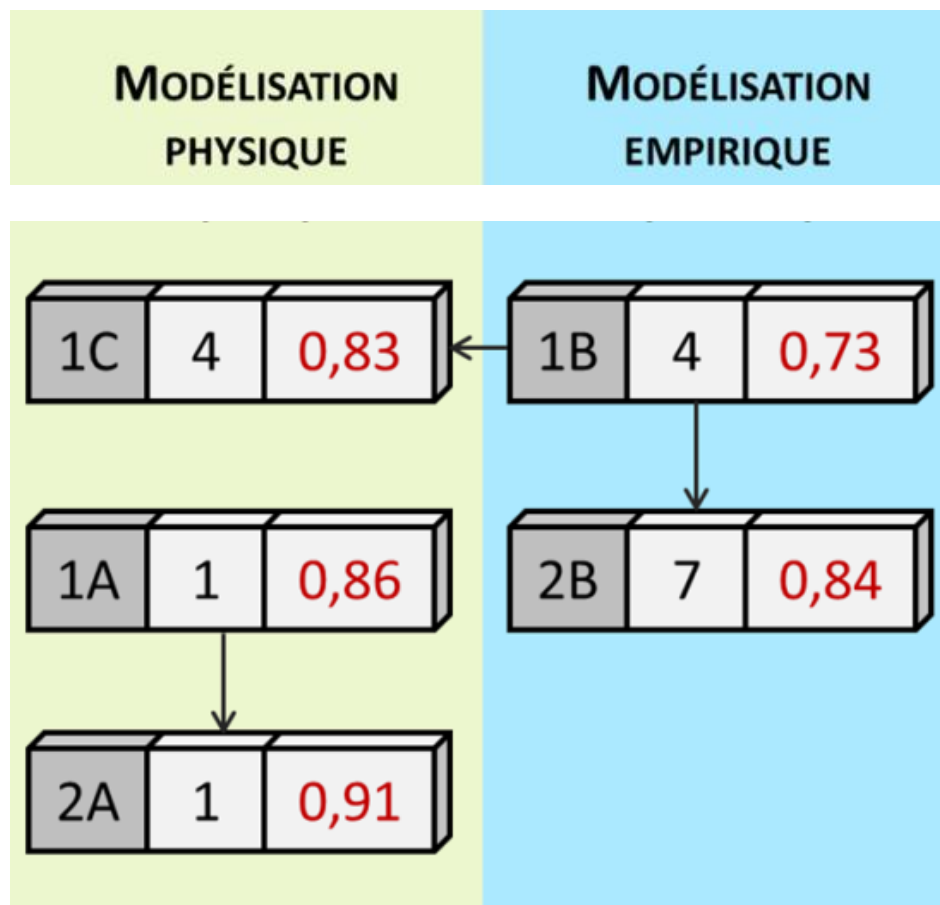
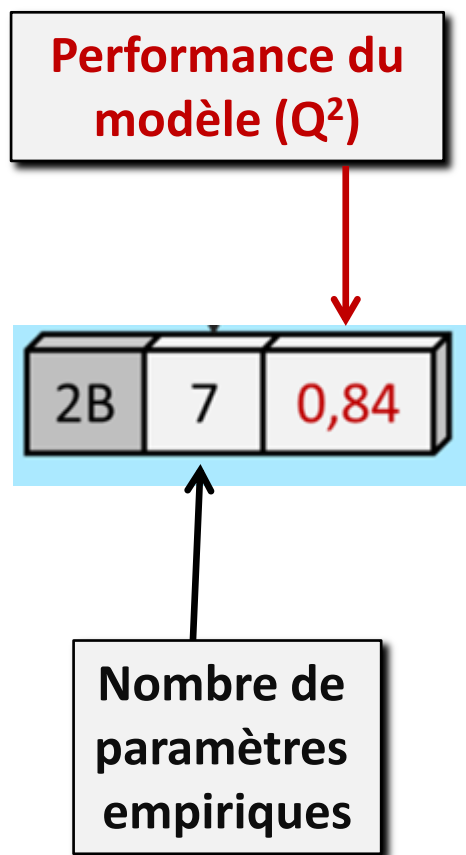
$$-\Delta H_{\text{dec}}^{\text{theor}} = (-16.9EA + 18.1\omega - 3.6\Delta E_{\text{diss}}^{\text{theor}} + 263.8) * M_w$$

... conduit immédiatement à un modèle nettement amélioré :

$$R^2 = 0,88 \quad Q^2 = 0,83$$

En principe il est cependant préférable de partir d'emblée de considérations physiques

Chaque boîte à 3 compartiments = 1 modèle

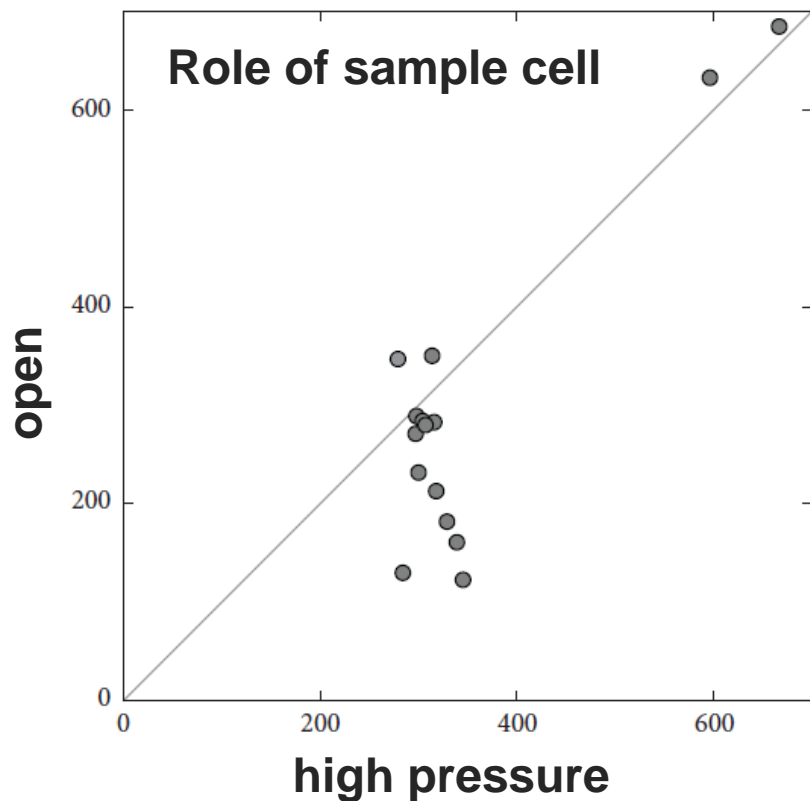


Autres aspects illustrés par la chaleur de décomposition molaire

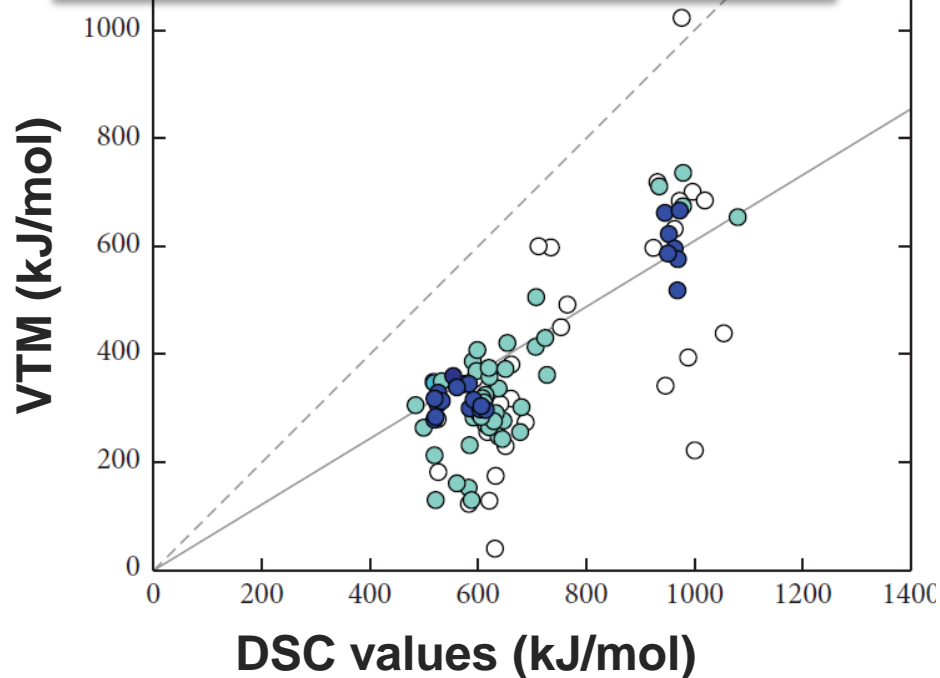
On peut estimer une **valeur théorique maximale (VTM)** définie ainsi :

$$\Delta_f H^0(\text{explosive}) - \Delta_f H^0(\text{decomp. products derived from the H}_2\text{O-CO}_2 \text{ arbitrary})$$

Au vu des incertitudes expérimentales ...



... la VTM est bien corrélée aux valeurs mesurées par DSC :
DSC = 60% (VTM)



Formulation des modèles basée d'emblée sur des considérations physiques

- Energie de Gurney E_G (performance explosif/poudre)

$$E_G/E_0 = 1 - 2 \left(\frac{1/\rho_0}{NV} \right)^{\gamma-1} \left(\frac{\gamma}{\gamma+1} \right)^\gamma.$$

- Stabilité thermique (température T_{onset})

$$T_{\text{onset}} = T_0 + c_1 dT/dt \quad T_0 = \frac{1}{\ln(Z/k_c)} \frac{E_0^\ddagger}{k_B} = \alpha \frac{E_0^\ddagger}{k_B}$$

- Sensibilité à l'impact (hauteur de chute h_{50})

$$h_{50} = \left(\frac{k_c}{k_{pr}} \right)^4 \quad k_{pr} = \frac{1}{N_A} \sum_i Z_i \exp \left(- \frac{E_i^\ddagger}{k_B T_e} \right)$$

$$k_B T_e = \eta \frac{E_c}{3N_A/2}$$

Prédiction des vitesses de Gurney : réduction des erreurs d'un facteur 2

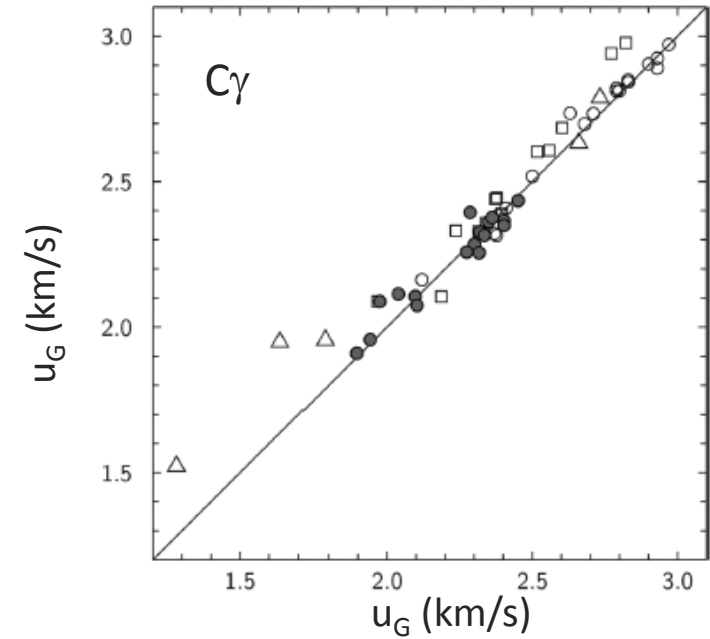
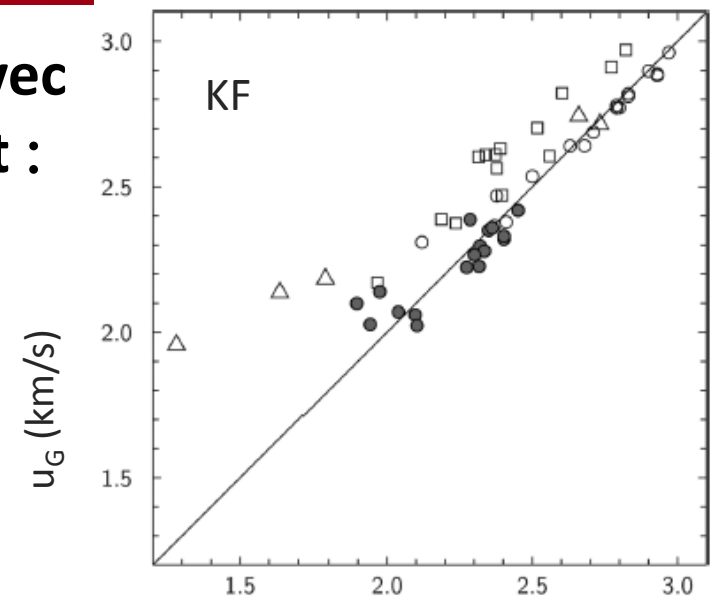
Hypothèse de base = expansion des gaz avec
exposant polytropique γ supposé constant :

$$E_G/E_0 = 1 - 2 \left(\frac{1/\rho_0}{NV} \right)^{\gamma-1} \left(\frac{\gamma}{\gamma+1} \right)^\gamma$$

Vitesse $u_G = (2E_G)^{1/2}$

Mauvaise approximation en général

Toutefois, l'ajustement de γ et
du volume V sur des données
expérimentales permet
de remédier à ses conséquences



Prédiction de températures de décomposition (T_{onset}) à partir de la th. de l'état de transition

Ingrédients du modèle :

- ❑ T_{onset} supposé correspondre à une valeur critique k_c de la constante de vitesse de la décomposition
- ❑ Théorie de l'état de transition (TST)
- ❑ Dépendance linéaire observée de T_{onset} par rapport à la vitesse de balayage en température dT/dt

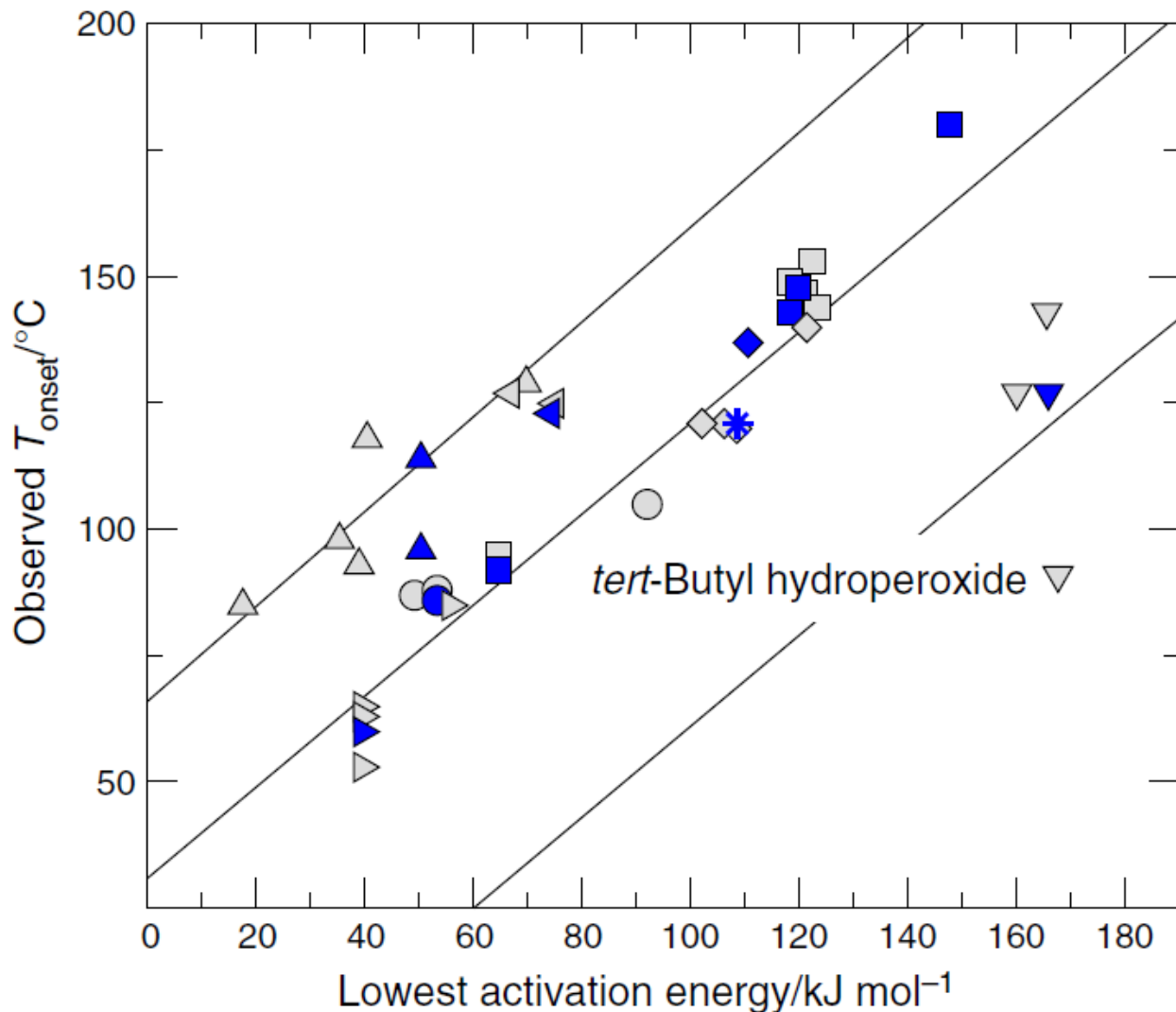
➔ Ces hypothèses conduisent à une simple relation linéaire :

$$T_{onset} = a(E^+/k_B) + b(dT/dt)$$

Deux paramètres empiriques susceptibles de dépendre de détails peu accessibles (préfacteurs cinétiques, détails expérimentaux...)

Recherche de la relation linéaire attendue entre T_{onset} et l'énergie d'activation la + faible

Résultats pour un ensemble de 38 peroxydes organiques



Corrélation observée au sein de chaque famille de peroxydes (sauf hydroperoxydes)

Nécessité d'introduire différentes constantes à l'origine (selon la famille de peroxyde considérée)

Un seul point vraiment hors-corrélation (tert-butyl hydroperoxide)

Températures de décomposition : comparison avec QSPR

	Present model	QSPR
(J. Thermal. Anal. Calorim. DOI 10.1007/s10973-017-6114-x)		(J Hazard Mater. 2014;276:216–24)
# adjustable parameters	4	4
Fitting ability (R^2)	0.93	0.84
Robustness (Q^2)	0.90	0.77
Predictive value (R_x^2)	0.90	0.80
RMS Error (training)	8	12
RMS Error (test set)	12	14
Physical consistency	YES	NO (*)

$$T_{onset} = 144 F_{00}^- + 29 n_{00} - 20 gap + 194$$

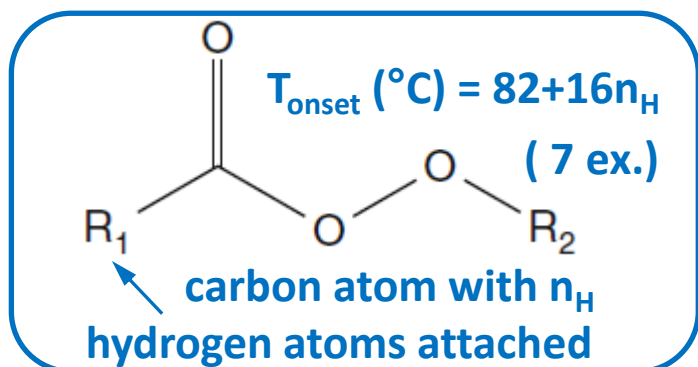
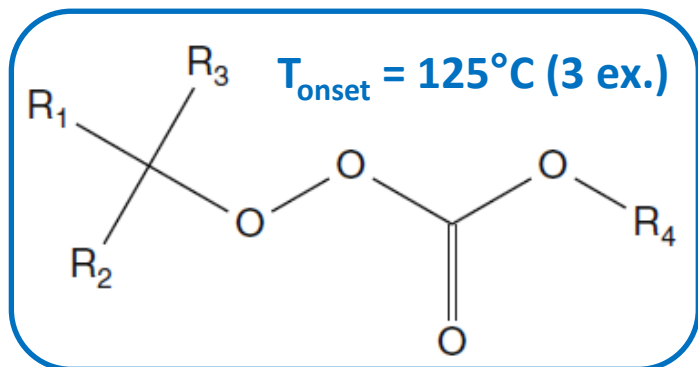
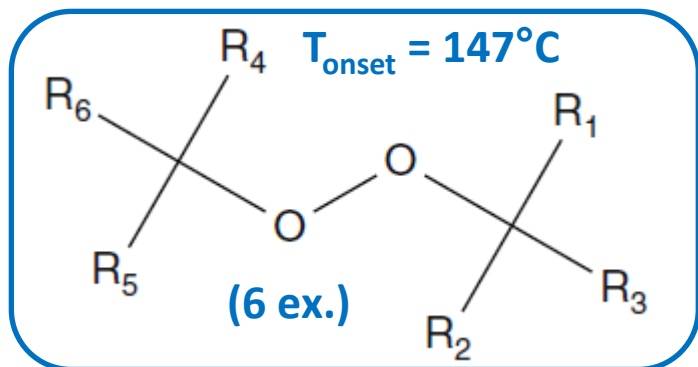
T_{onset} (QSPR): increases with # of O-O groups
decreases with band gap

$$T_{onset}(K) = 7.61 \times 10^{-3} (E_0^\ddagger / k_B) + 304 + 35I_a - 45I_h$$

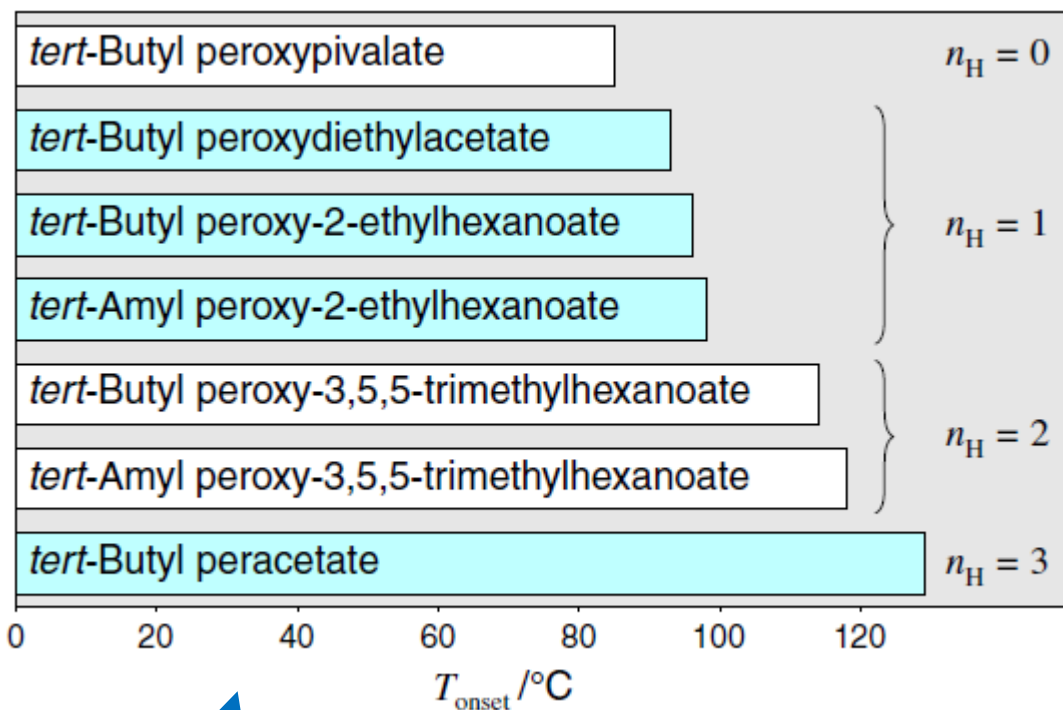
QSPR hampered by scarcity of data

S'affranchir des calculs de chimie quantique : introduction de températures de groupes

Vers une base de données de températures pour estimation rapide de T_{onset}



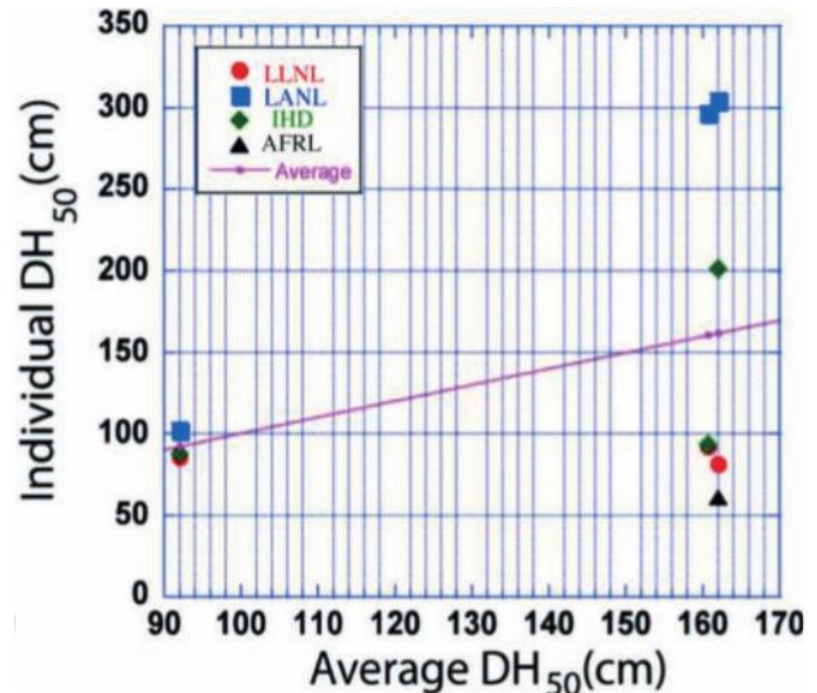
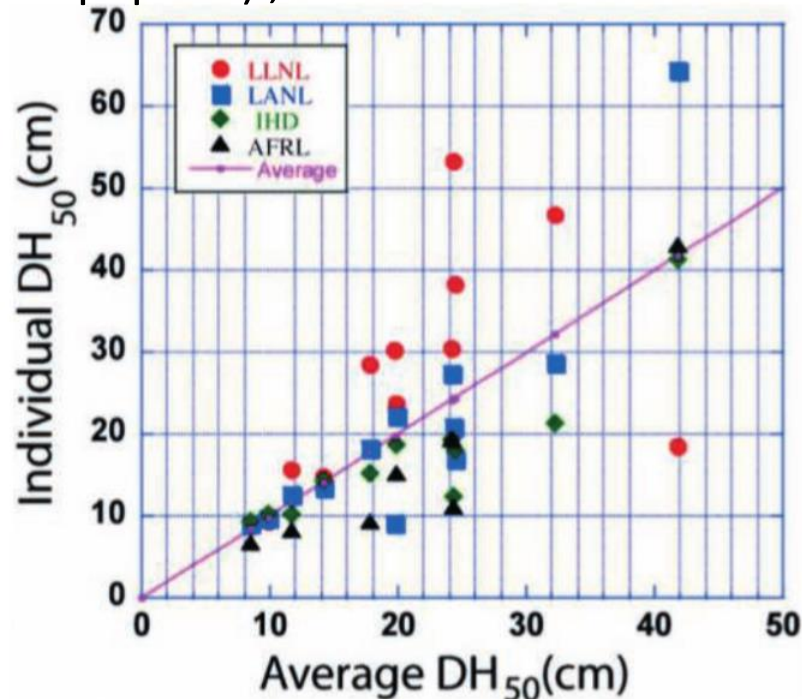
Rendue possible par la cohérence remarquable des températures mesurées avec les règles de chimie organique



Drop weight impact test: repeatability issues

Many parameters observed to play a role on measured h_{50} values:

Powder morphology, particle sizes, laboratory environment (temperature, ambient humidity...), experimental devices (weight of the hammer, use of sandpaper...) , statistical treatment of the data...



inter-laboratory comparison of h_{50} values measures for improvised explosives
M. M. Sandstrom et al. *Propellants Explos. Pyrotech.* **2015**, 40, 109 – 126

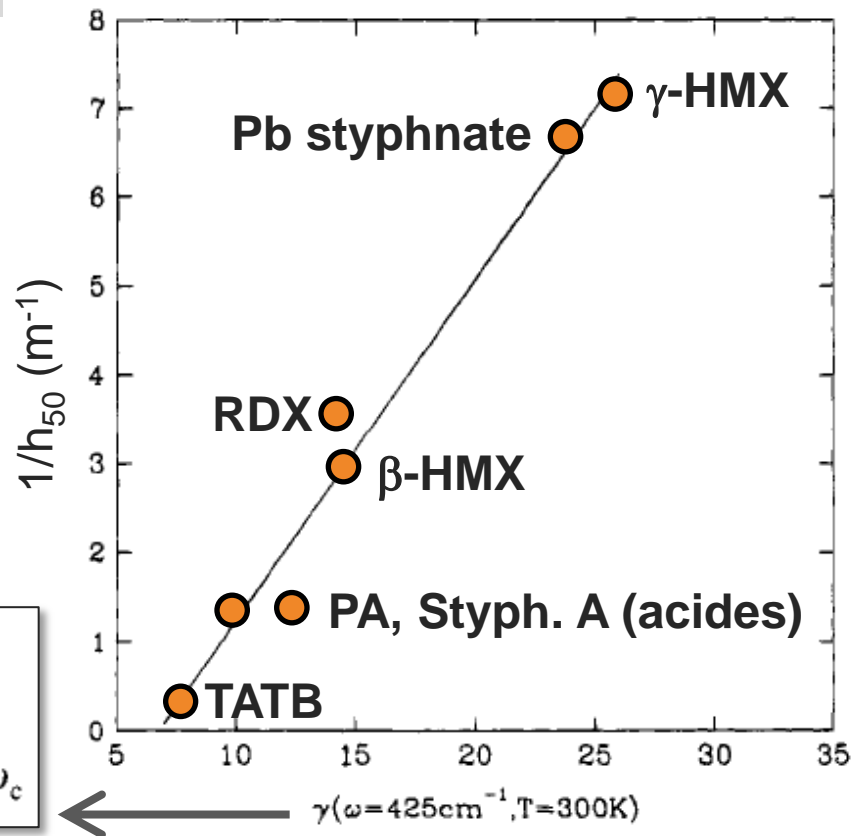
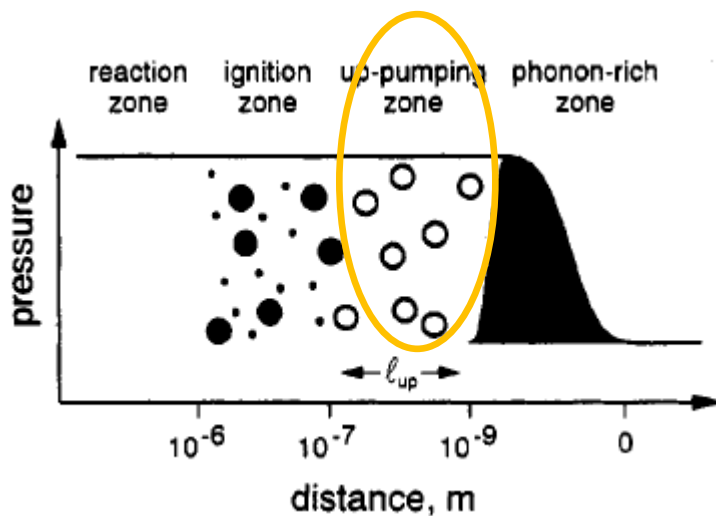


We focus on the magnitude of h_{50} : use $\log(h_{50})$ rather than h_{50}

Early approach to impact sensitivity: phonon up-pumping model

Assumes a pre-reactive rate-limiting step:
energy transfer from the impact to intramolecular vibrations

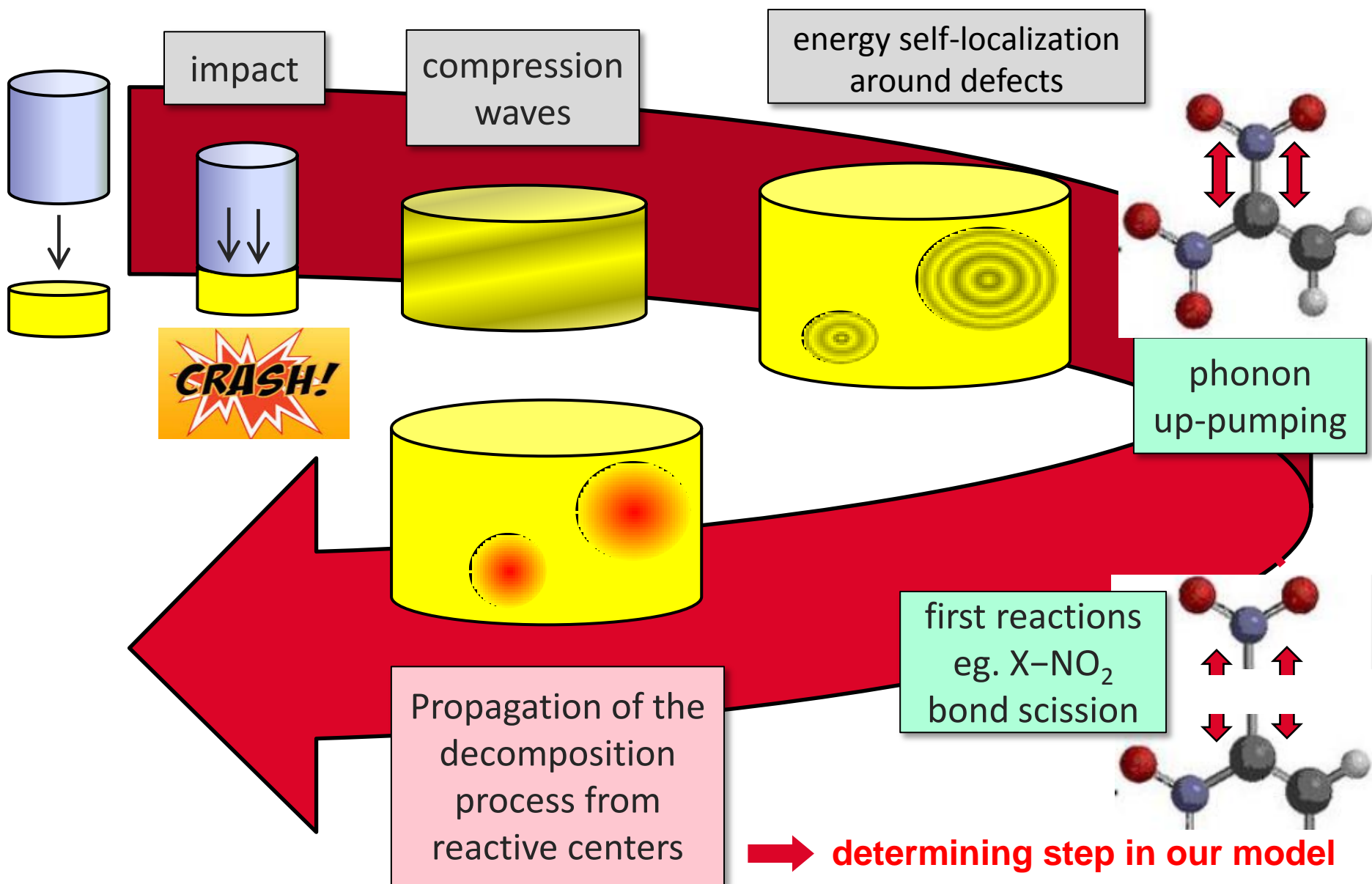
Tokmakoff et al. *J. Phys. Chem.* 97, 1901 (1993)



$$\gamma(\omega_c, T) = \frac{9|B|^2 V^2 k_B^2 T^2}{4\pi^2 \hbar} \int d\omega \frac{1}{\omega(\omega_c - \omega)} \rho(\omega) \rho(\omega_c - \omega), \quad k_B T \gg \omega_c$$

➔ Theoretically appealing but fails to provide quantitative predictions for a significant number of compounds

A number of distinct rate-limiting steps have been proposed...

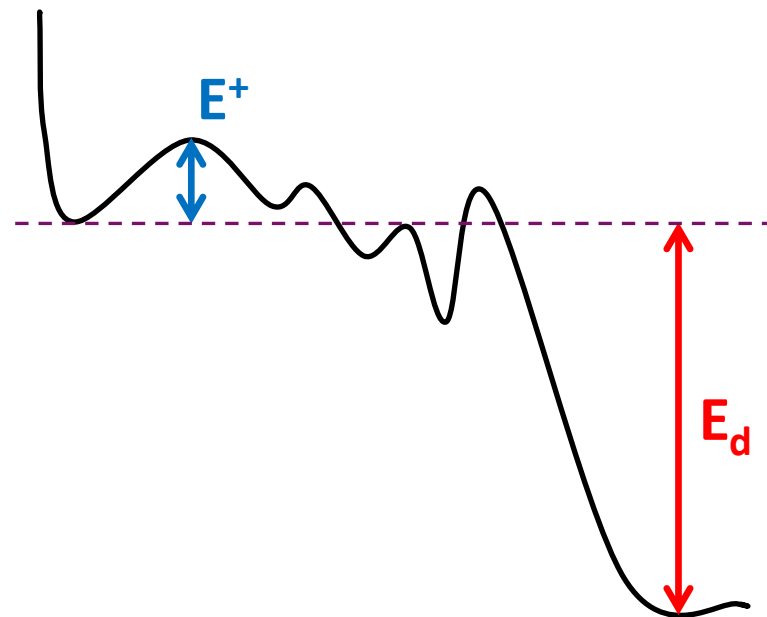


Why we believe in a rate-limiting step occurring later in the initiation process

Theoretical argument

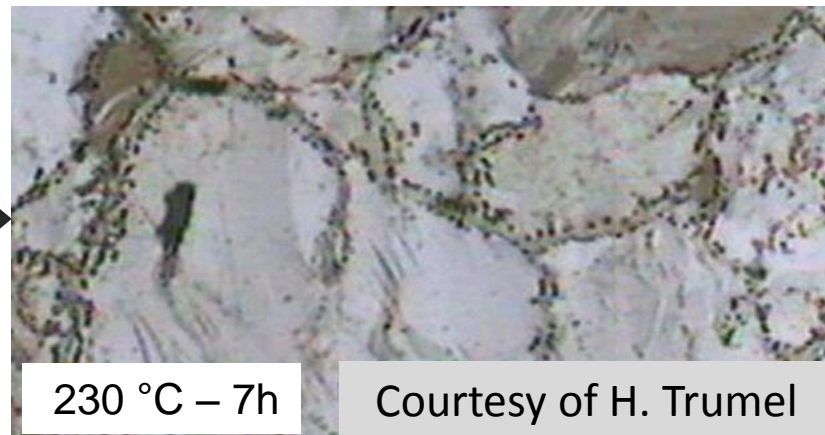
The Potential Energy Surface exhibits two main features:

- ❑ activation energies associated with early reaction steps E^+
- ❑ energy difference between unreacted molecule and decomposition products E_d



Empirical argument

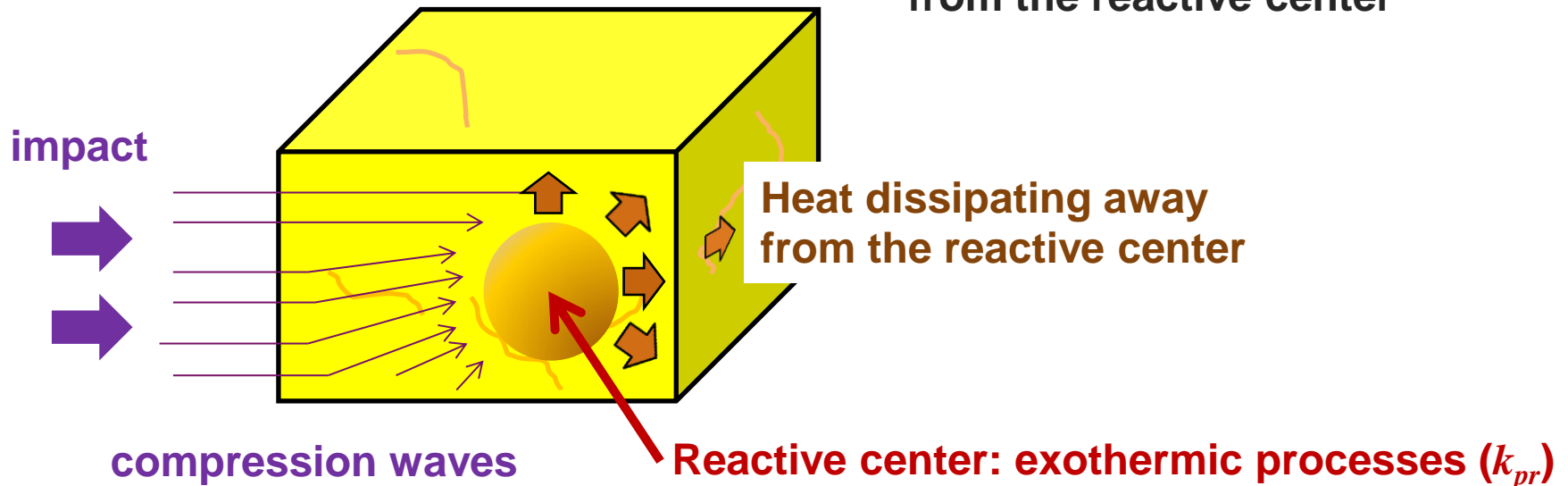
Assumption supported by micrographs of explosives subject to thermal/mechanical insults and showing holes attributed to the decomposition of molecules



Kinetic approach to impact sensitivity in a nutshell

Condition for ignition

Locally, the decomposition must be sufficiently fast (k_{pr}) compared to the energy diffusion rate away from the reactive center



➔ Sensitivity determined by an equation involving k_{pr}

Kinetic approach to impact sensitivity: working equations

Insensibility (h_{50}) increases as the rate for the decomposition process (k_{pr}) decreases compared to the energy dissipation rate

A simple power law exhibits correct asymptotic limits: $h_{50} \propto k_{pr}^{-n}$

only $N+1$ ajustables parameters

$$k_{pr} = \frac{1}{N_A} \sum_i Z_i \exp\left(-\frac{E_i^\ddagger}{k_B T}\right)$$

$$\frac{E_i^\ddagger}{k_B T} = \frac{c D_i N_A}{E_c}$$

E_c = energy content

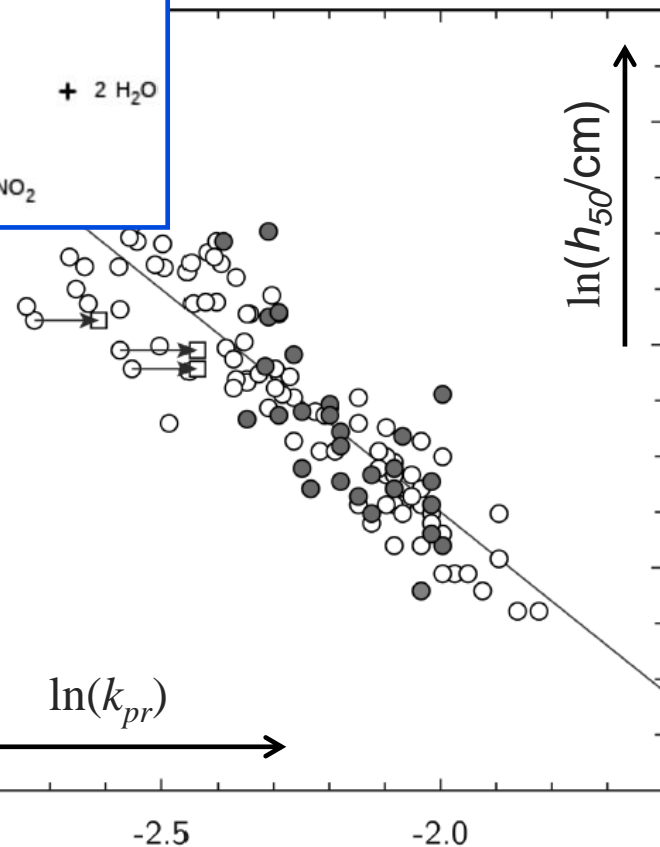
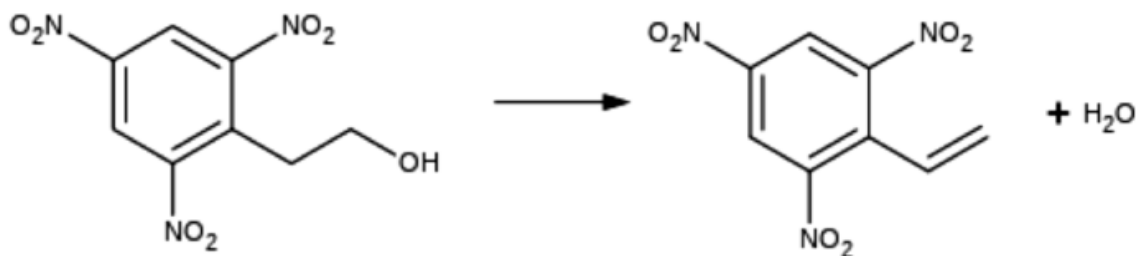
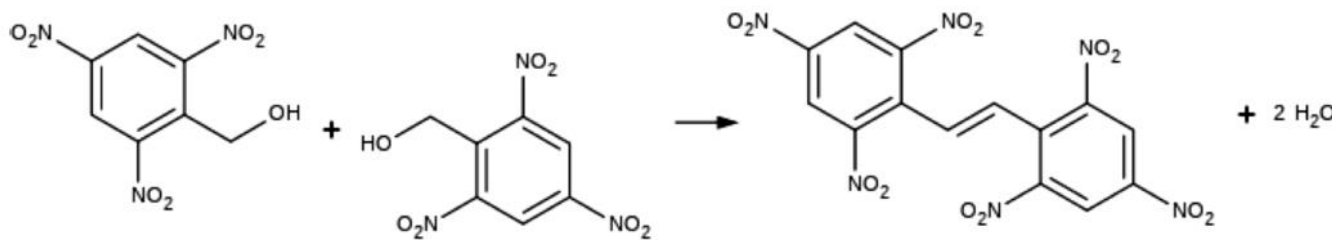
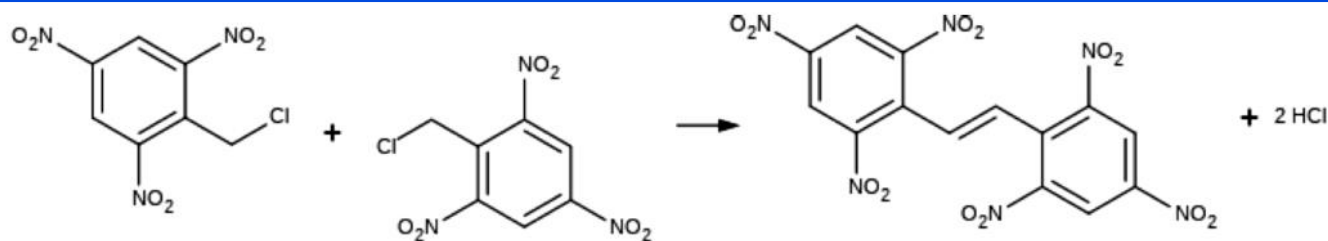
Tabulated X-NO₂ bond dissociation energies D_i

Table 1. Standard Bond Dissociation Energies (kcal/mol) Used in the Present Work^a

type of nitro	ΔG_{diss}^0
aromatic nitro group with two -NH ₂ groups in ortho positions	58.5
aromatic nitro group with -OH and -NH ₂ groups in ortho positions	60.0
aromatic nitro group with one -NH ₂ group in ortho position	56.2
aromatic nitro group with one -OH group in ortho position	56.8

→ model straightforward to develop – only 1 or 2 parameter(s) for each kind of explosophore moiety

Some elimination reactions that might take place before $X-NO_2$ bond scissions

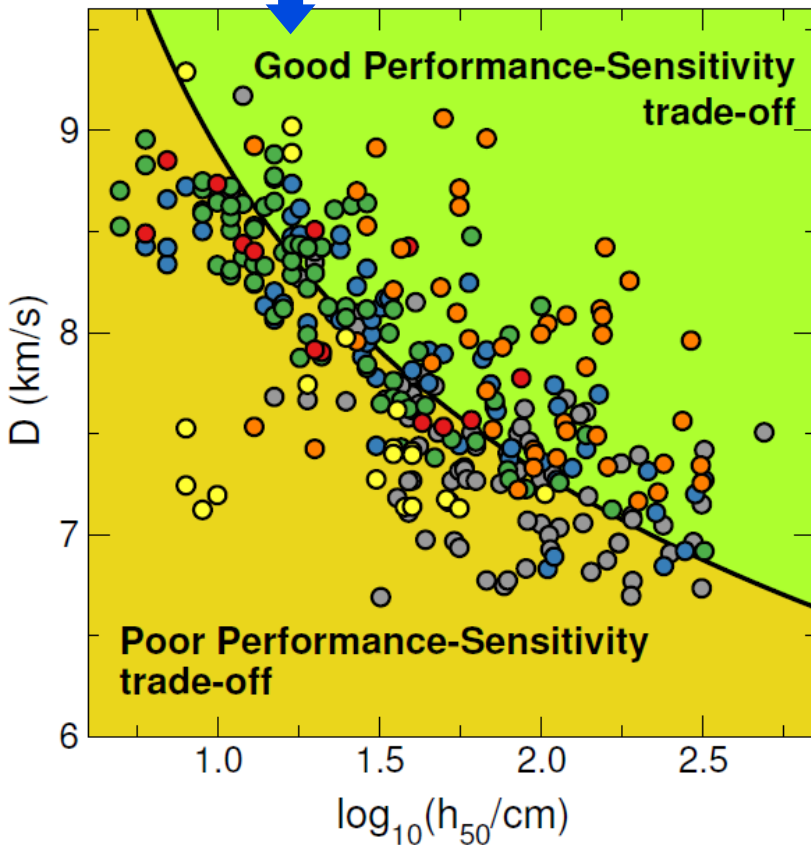


D. Mathieu, T. Alaïme,
J. Phys. Chem. A. 118, 9720-9726, 2015

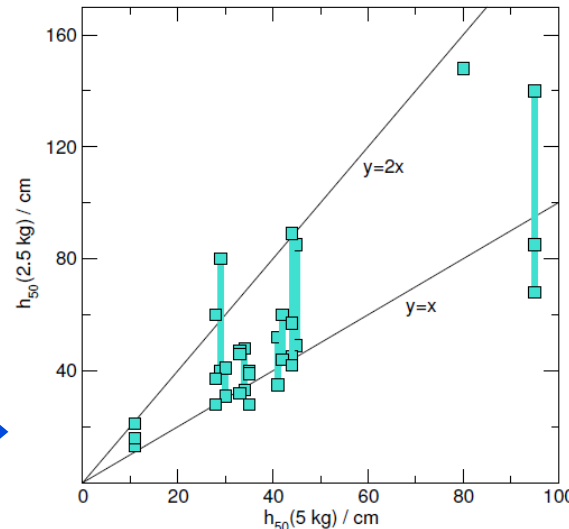
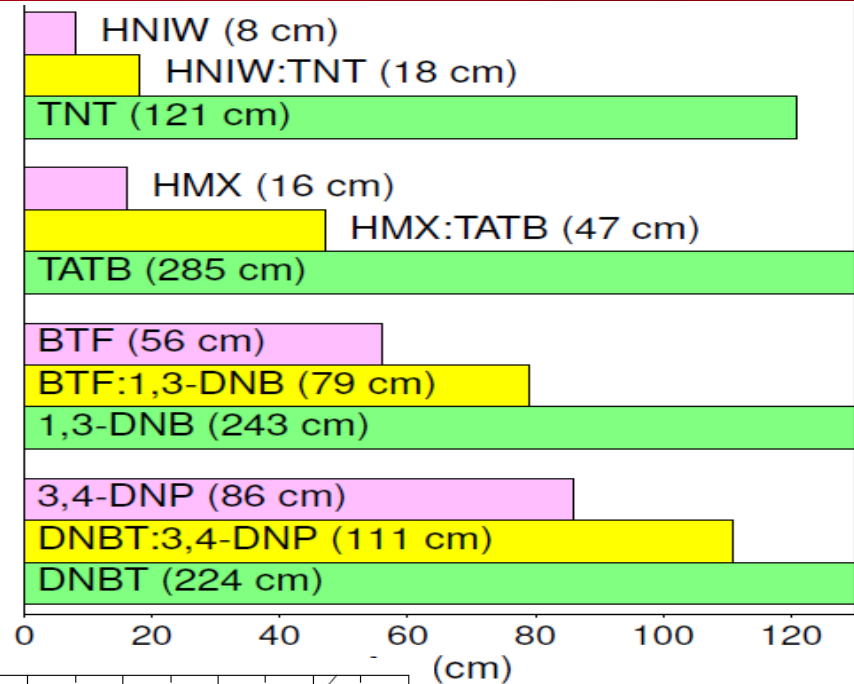
➔ small but systematic improvement of the correlation

The kinetic model of impact sensitivity yields much qualitative insight, explaining...

The origin of the performance-sensitivity trade-off



The role of experimental parameters



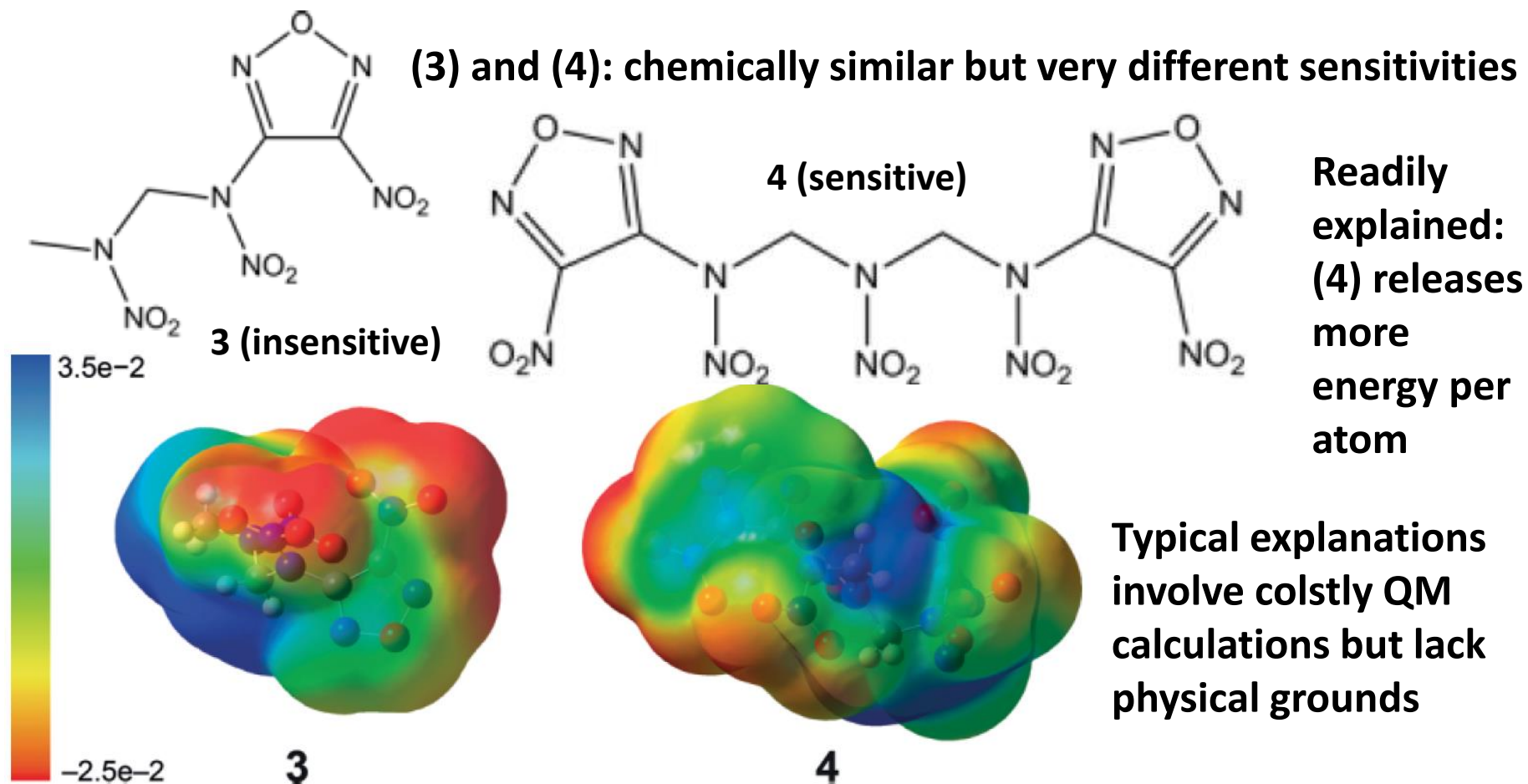
The relative sensitivity of high energy co-crystals

etc ...

Computational screening tool as well as simple back-of-the-envelope procedure

Extensive validation → kinetic model far more reliable than any alternative, either general (JMGM 62, p. 81, 2015) or highly specialized (IECR 55, p. 7569, 2016)

Usable by bench chemists for quick evaluation of the sensitivity of synthetic targets



Conclusion concernant l'évaluation des propriétés de substances dangereuses

- Les modèles établis à partir de considérations physiques sont systématiquement plus fiables que des modèles purement empiriques
- Ils conduisent à mieux connaître et comprendre les mécanismes sous-jacents
- Ils sont bien plus simples et rapides à développer, l'« apprentissage » se réduisant à caler des paramètres
- Ils permettent des calculs « à la main » une fois que les grandeurs importantes ont été calculées une fois pour toutes par chimie quantique

Conclusion concernant l'évaluation des propriétés de substances dangereuses

- Les modèles établis à partir de considérations physiques sont systématiquement plus fiables que
 - Ils ne permettent pas de rendre compte de la complexité des phénomènes
 - Ils ne permettent pas de rendre compte de la complexité des phénomènes
 - Ils ne permettent pas de rendre compte de la complexité des phénomènes
- Les données disponibles ne sont pas assez « grandes » pour rendre une approche « big data » compétitive**
- Ils ne permettent pas de rendre compte de la complexité des phénomènes
- les grandeurs importantes ont été calculées une fois pour toutes par chimie quantique

Conclusion concernant l'évaluation des propriétés de substances dangereuses

- Les modèles établis à partir de considérations physiques sont systématiquement plus fiables que
- Ils permettent de re les
- Ils permettent de per,
- Ils permettent de is que les grandeurs importantes ont été calculées une fois pour toutes par chimie quantique

Merci de votre attention