

APPLICATION DES FRAGMENTS GÉOMÉTRIQUES AU CALCUL DE LA SOLUBILITÉ DANS L'OCTANOL

La méthode des fragments géométriques est une procédure générale pour évaluer les propriétés des liquides et des solides moléculaires à partir des formules chimiques des espèces dont ils sont constitués.

Elle s'applique plus particulièrement aux quantités déterminées par les interactions entre molécules. Elle a par exemple permis des progrès récents dans le calcul d'une grandeur d'intérêt pratique considérable : la solubilité des composés organiques dans l'octanol.

Connaître les propriétés des substances chimiques aide à se prémunir contre les risques éventuels liés à leur utilisation ou à choisir les composés les plus adaptés pour une application donnée. Les études expérimentales étant coûteuses et souvent difficiles, on manque de données pour la grande majorité des composés commercialisés, sans compter le fait qu'il serait également très utile pour les chimistes de connaître les propriétés de composés encore inconnus, afin de savoir sur quelles cibles concentrer les efforts de synthèse. Des simulations moléculaires peuvent dans certains cas pallier cette carence. Cependant, elles nécessitent des moyens inaccessibles à la majorité des ingénieurs en manque de données. De plus, la fiabilité des résultats est limitée, du fait par exemple des incertitudes sur les forces interatomiques ou de configurations atomistiques non représentatives. Il serait préférable de pouvoir calculer les propriétés d'une phase condensée sur la base d'un minimum de données, si possible uniquement à partir des formules développées des molécules.

En vue d'aboutir à de tels outils prédictifs, deux approches sont couramment utilisées. La première, dite additive, consiste à représenter la propriété considérée à l'aide de contributions associées aux différentes parties qui constituent la molécule. La seconde, désignée par le sigle anglais QSPR, signifiant relation quantitative structure-pro-

priété, consiste à la relier de manière empirique à des grandeurs appelées descripteurs, directement déduites des formules des molécules. Si l'approche QSPR est souvent considérée comme supérieure aux méthodes additives, un examen approfondi de la littérature révèle que cette idée résulte de l'utilisation erronée du principe d'additivité, certains auteurs négligeant la nécessité de modéliser des grandeurs intermédiaires ou de prendre en compte la molécule dans sa totalité.

De plus, ce principe n'est pas exploité de manière optimale par les deux principaux types de méthode additive, qui reposent sur des termes associés respectivement aux atomes et aux groupes fonctionnels. L'additivité atomique est d'un intérêt très limité, car elle ignore l'environnement des atomes. En revanche, l'additivité des groupes pose un problème inverse : le nombre de groupes fonctionnels est si élevé qu'il est difficile de déterminer les contributions de chacun d'eux. De plus, il est paradoxal de modéliser des propriétés physiques déterminées par les forces entre molécules à partir de la notion de groupe fonctionnel, destinée plutôt à rendre compte de leur réactivité.

Le CEA – Le Ripault a développé une approche additive basée sur des considérations purement géométriques. Plus précisément, cette méthode, dite des fragments géométriques, suppose que la contribution aux forces intermoléculaires d'un

atome dépend de sa nature chimique et de son accessibilité, elle-même fonction du nombre de voisins auxquels l'atome est lié et de leurs rayons **1**. Elle est conçue pour les propriétés déterminées par les forces intermoléculaires, comme la densité ou l'enthalpie de sublimation, pour lesquelles elle fournit les méthodes de calcul les plus fiables disponibles à ce jour **1,2**. Elle paraît donc prometteuse pour des propriétés d'intérêt pratique majeur, comme la solubilité des substances dans les liquides, nécessaire aussi bien pour concevoir les procédés de fabrication de l'industrie chimique que pour comprendre le devenir des molécules ingérées par l'organisme ou rejetées dans l'environnement.

La solubilité dans l'octanol à température ambiante, notée S_{oct} , est particulièrement importante pour ces dernières applications. Différentes méthodes ont été publiées pour l'évaluer, notamment une méthode de groupes ainsi qu'une large gamme de modèles QSPR. Les meilleurs résultats ont été obtenus par une technique dite Random Forest (RF). Ce modèle, qui s'apparente à une boîte noire du fait de la complexité des relations mathématiques sous-jacentes, met en jeu 86 descripteurs et un nombre encore plus élevé de paramètres empiriques **3**.

Par rapport à ces précédents modèles, la technique des fragments géométriques présente une fiabilité comparable ou supérieure. De plus, le modèle associé est plus simple et transparent, avec 19 paramètres et 18 variables représentant les contributions des différents fragments à l'enthalpie libre

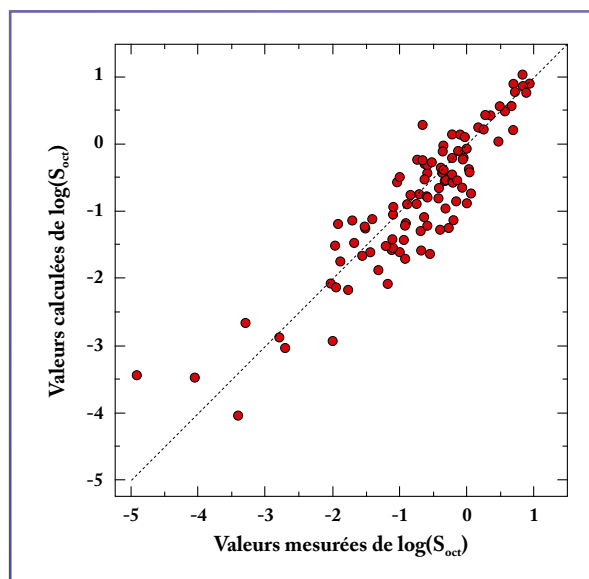


Figure 2

➔ Logarithme de la solubilité S_{oct} exprimée en moles/litre: les valeurs en ordonnée ont été calculées par la méthode des fragments géométriques, après ajustement de ses paramètres. Les valeurs en abscisse sont issues de mesures. L'accord entre les valeurs prédites et mesurées est satisfaisant comparé aux écarts observés avec les autres méthodes d'estimation.

de dissolution **4**. Un avantage notable par rapport à l'approche QSPR réside dans le fait que la cohérence des paramètres est aisément vérifiable. Par exemple, on peut vérifier (**figure 1**) que la contribution à la solubilité d'un atome de carbone augmente lorsqu'on remplace un de ses voisins par un atome d'hydrogène, comme attendu théoriquement.

Par ailleurs, pour les solides, la solubilité dans l'octanol S_{oct} dépend de l'entropie associée à la transition solide-liquide, souvent supposée constante (règle de Walden). Les fragments géométriques permettent de s'affranchir de cette hypothèse. Combinés au point de fusion expérimental, ils conduisent à d'excellents résultats (**figure 2**)

aussi précis que les données déduites de mesures indirectes **4**.

Ces résultats confirment la pertinence et la généralité de la notion de fragment géométrique pour la modélisation des propriétés des phases condensées. Diverses améliorations sont envisageables, en introduisant par exemple des charges atomiques pour prendre en compte de manière plus explicite la distribution électronique, ou une fonction d'écrantage commune à tous les atomes, en vue de réduire le nombre de paramètres empiriques.

RÉFÉRENCES

- 1** D. MATHIEU *et al.*, "Reliable and Versatile Model for the Density of Liquids Based on Additive Volume Increments", *Ind. Eng. Chem. Res.*, **55**, p. 12970-12980 (2017).
- 2** D. MATHIEU, "Simple Alternative to Neural Networks for Predicting Sublimation Enthalpies from Fragment Contributions", *Ind. Eng. Chem. Res.*, **51**, p. 2814-2819 (2012).
- 3** M. A. BUONAIUTO *et al.*, "Prediction of 1-Octanol Solubilities Using Data From the Open Notebook Science Challenge", *Chem. Cent. J.*, **9**, 50 (2015).
- 4** D. MATHIEU, "Solubility of Organic Compounds in Octanol: Improved Predictions Based on the Geometrical Fragment Approach", *Chemosphere*, **182**, p. 399-405 (2017).

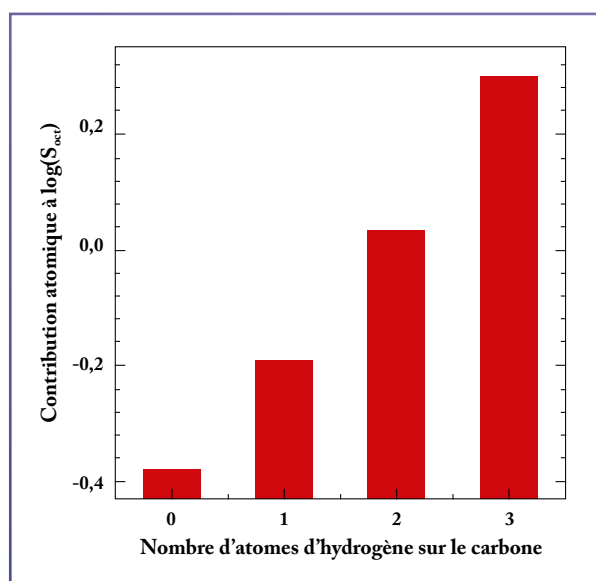


Figure 1

➔ Contributions d'un atome de carbone saturé (relié seulement par des liaisons simples) au logarithme de la solubilité dans l'octanol S_{oct} d'une molécule. Ces valeurs constituent les paramètres ajustables de la méthode des fragments géométriques. Le fait qu'elles augmentent régulièrement avec le nombre d'atomes d'hydrogène attachés à l'atome conforte leur validité ainsi que la pertinence du modèle.