

DÉVELOPPEMENT DE MODÈLES QSPR POUR LA PRÉDICTION DES PROPRIÉTÉS DE TENSIOACTIFS DÉRIVÉS DE SUCRES

Contributeurs

Théophile GAUDIN,

Patricia ROTUREAU,

Guillaume FAYET

La substitution des ressources fossiles par des matières renouvelables, notamment issues de la biomasse, est un enjeu de recherche important. Si la majorité des tensioactifs utilisés dans les détergents et en cosmétique provient aujourd'hui des ressources fossiles, les tensioactifs issus de la bioraffinerie, notamment les tensioactifs dérivés de sucres, constituent de bons candidats pour substituer les tensioactifs pétro-sourcés [1].

Aussi, l'Ineris s'est engagé dans le projet AMPHIPRED (2013-2016), financé par l'ITE PIVERT piloté par l'UTC de Compiègne et en collaboration avec l'université d'Amiens et l'ESCOM. Ce projet vise la prédiction des propriétés amphiphiles de ces tensioactifs. En effet, l'accès à ces propriétés à partir de leur structure moléculaire permet d'évaluer *a priori* les performances potentielles de nouveaux tensioactifs biosourcés dans des formulations industrielles avant même leur synthèse [2].

Dans le cadre des travaux de doctorat de Théophile Gaudin (collaboration UTC-Ineris), la méthode prédictive *Quantitative Structure Property Relationship* (QSPR), qui cherche à relier de manière quantitative une propriété expérimentale à la structure moléculaire d'une substance, a été utilisée. Si différents modèles QSPR existaient déjà pour prédire certaines propriétés des tensioactifs pétro-sourcés [3], aucun modèle validé n'avait été identifié pour les tensioactifs dérivés de sucres. L'objectif de ces travaux était de développer de tels modèles, en s'appuyant sur l'expertise de l'Ineris quant à leur développement et leur validation pour les propriétés physico-chimiques dangereuses des substances chimiques [4].

MÉTHODE

Une base de données (en cours de valorisation avec la SAS PIVERT sous le nom d'AmphInnov) contenant 2626 données issues de la littérature pour 24 propriétés amphiphiles et 659 tensioactifs dérivés de sucres a été collectée. Comme montré en **figure 1**, quatre propriétés constituant de bons indicateurs des performances des tensioactifs dans les formulations ont ensuite été étudiées: la concentration micellaire critique (CMC), la tension de surface à la CMC (γ_{CMC}), l'efficacité (pC_{20}) et le point de Krafft (T_K). Pour chacune d'elles, différents modèles QSPR ont été développés en fonction de la nature des descripteurs utilisés: d'une part, selon qu'ils sont basés sur l'intégralité de la structure du tensioactif ou sur les structures de la tête polaire et de la chaîne alkyle prises séparément (**figure 2**), et d'autre part, selon le niveau de complexité des descripteurs moléculaires, depuis les descripteurs constitutionnels. Ces derniers se calculent à partir de la simple structure semi-développée de la molécule, aux descripteurs issus de calculs de chimie quantique. Afin d'évaluer de manière fiable les performances de ces modèles, chaque modèle a fait l'objet d'une série de validations, incluant en particulier une validation externe robuste de son pouvoir prédictif en ce qui concerne des molécules non utilisées pour l'ajustement du modèle.

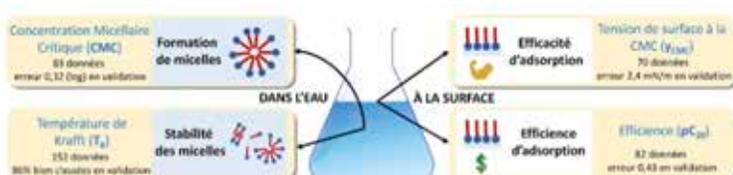
RÉSULTATS

Des modèles QSPR satisfaisants ont été obtenus pour les quatre propriétés étudiées. Pour la CMC, la tension de surface à la CMC et l'efficacité, il s'agit de modèles multilinéaires prédisant une valeur quantitative. Un exemple de modèle simple est le modèle QSPR suivant (Eq. 1), développé pour la prédiction de la CMC à l'aide de descripteurs constitutionnels calculés séparément pour la tête polaire et la chaîne alkyle:

$$\log CMC = -0,027 M_{w,c} - 20n_{rel,S,h} - 63n_{rel,Single,c} + 64,8$$

où $M_{w,c}$ et $n_{rel,Single,c}$ sont la masse molaire et le nombre relatif de liaisons simples de la chaîne alkyle et $n_{rel,S,h}$ est le nombre relatif d'atomes de soufre de la tête polaire. Comme montré en **figure 3**, ce modèle présente un bon pouvoir prédictif avec une erreur standard de 0,36 (log) pour les 27 molécules du jeu de validation.

Figure 1 / Synthèse des résultats obtenus pour la prédiction des propriétés des surfactants dérivés de sucre au cours du projet AMPHIPRED.



Si de tels modèles sont simples à mettre en œuvre, l'emploi de descripteurs quantiques permet un gain en pouvoir prédictif parfois important. Par exemple, pour γ CMC, l'erreur standard en validation du meilleur modèle dédié à γ CMC incluant les descripteurs quantiques est de 2,4 mN/m tandis que, pour le meilleur des modèles sans descripteur quantique, celle-ci est de 2,9 mN/m.

Pour le point de Krafft (T_K), des arbres de décision ont été développés classant la molécule selon que la valeur de T_K est inférieure ou supérieure à la température ambiante. Par exemple, l'arbre de décision en **figure 4** a été obtenu sur la base des mêmes types de descripteurs que ceux employés pour obtenir l'Eq. 1, et dispose aussi d'un bon pouvoir prédictif, avec 78 % des 51 molécules du jeu de validation correctement classées.

Ces travaux ont donné lieu à deux publications [5; 6] et à plusieurs communications (comme celles présentées aux congrès CESIO [7], ISGC [8] et SFGP [9]). De plus, Théophile Gaudin a reçu le prix de thèse Guy Deniérou de l'UTC en avril 2017. Les modèles développés permettent d'aiguiller les synthèses de molécules biosourcées par exemple en fonction d'un cahier des charges de propriétés recherchées comme cela est étudié actuellement dans le projet AMPHIFOAM (2016-2019).

Références

[1] Rosen, M.J.; Kunjappu, J.T.; *Surfactants and Interfacial Phenomena*, 4th ed., John Wiley & Sons, Inc., 2012.
 [2] Fayet, G.; Rotureau, P. How to use QSPR type approaches to predict properties in the context of Green Chemistry, *Biofuels, Bioproducts and Biorefining*, 10 (2016) 738-752.
 [3] Creton, B. Prediction of Surfactants' Properties using Multiscale Molecular Modeling Tools: A Review, *Oil & Gas Science and Technology - IFP Energies nouvelles*, 67 (2013) 1-14.
 [4] Dearden, J.C.; Rotureau, P.; Fayet, G. QSPR prediction of physico-chemical properties for REACH, *SAR and QSAR in Environmental Research*, 24 (2013) 279-318.
 [5] Gaudin, T.; Rotureau, P.; Pezron, I.; Fayet, G. New QSPR Models to Predict the Critical Micelle Concentration of Sugar-Based Surfactants, *Industrial & Engineering Chemistry Research*, 55 (2016) 11716-11726.
 [6] T. Gaudin, P.; Rotureau, I.; Pezron, G.; Fayet, Conformations of n-alkyl- α/β -d-glucopyranoside surfactants: Impact on molecular properties, *Computational and Theoretical Chemistry*, 1101 (2017) 20-29.
 [7] Gaudin, T.; Lu, H.; Van Hecke, E.; Drelich, A.; Dao, T.T.; Rotureau, P.; Benali, M.; Bonnet, V.; Wadouachi, A.; Pourceau, G.; Fayet, G.; Pezron, I. Data analysis of sugar-based surfactant properties: towards quantitative structure property relationships, *CESIO 2015 - 10th World Surfactant Congress and Business Convention*, 1-3 juin 2015, Istanbul (Turquie).
 [8] Gaudin, T.; Fayet, G.; Rotureau, P.; Pezron, I. New QSPR models to predict the amphiphilic properties of sugar-based surfactants, *International Symposium on Green Chemistry*, 16-19 mai 2017, La Rochelle (France).
 [9] Gaudin, T.; Rotureau, P.; Pezron, I.; Fayet, G. Modèles QSPR pour la prédiction des propriétés des tensioactifs dérivés de sucre, *16^e Congrès de la Société française de génie des procédés*, 11-13 juillet 2017, Nancy (France).

ABSTRACT /

Substitution of petroleum-based surfactants by biobased alternatives is a promising field of research and applications. Predictive models can help to identify good candidates for such substitution. But, no validated model has been identified for biobased surfactants.

In AMPHIPRED project (2013-2016), based on a large dataset of 2626 experimental values, a series of new predictive quantitative structure-property relationships (QSPR) models were developed for four amphiphilic properties of sugar-based surfactants, an important family of biobased surfactants.

A series of promising QSPR models were obtained for each property, from models with simple constitutional descriptors favouring easy applications, to quantum chemical based ones providing better predictive power. For the critical micelle concentration (CMC), the surface tension at CMC and the adsorption efficiency, multi-linear regressions were developed, to achieve quantitative predictions. For the Krafft temperature (T_K), decision trees enabled qualitative prediction to evidence if T_K is higher or lower than room temperature.

These models open new perspectives towards in silico design and screening of new bio-based surfactants with target properties as studied in AMPHIFOAM project (2016-2019).

Figure 2 / Structure moléculaire d'un tensioactif dérivé de sucre, l'octyl- β , D-glucoside, et des fragments modélisés relatifs à sa tête sucre et à sa chaîne alkyle.

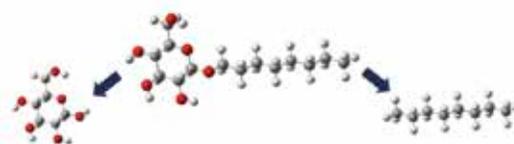


Figure 3 / Concentrations micellaires critiques (en log) expérimentales et calculées par le modèle QSPR simple présenté en Eq. 1.

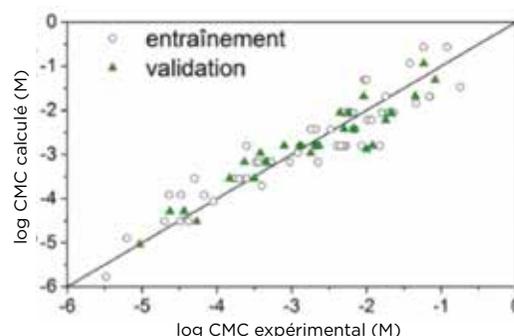


Figure 4 / Arbre de décision classant les tensioactifs dérivés de sucres selon leur température de Krafft supérieure ou inférieure à 25°C.

