

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE SALAH BOUBNIDER CONSTANTINE 3

N° d'ordre :
N° de série :



FACULTÉ DE GÉNIE DES PROCÉDÉS
DEPARTEMENT DE GENIE PHARMACEUTIQUE

THESE
EN VUE DE L'OBTENTION DU DIPLOME DE
DOCTORAT

**Développement d'un nouveau modèle UNIFAC
pour la prédiction de la solubilité**

Présentée par :

NOUAR Abderrahim

Devant le Jury composé de :

Présidente	BOUSHABA Rihab	Pr	Univ Salah Boubnider Constantine 3
Rapporteur	BOUHELASSA Mohamed	Pr	Univ Salah Boubnider Constantine 3
Examineurs	SAOULI Ouacil	MCA	ENPC
	MEROUANI Slimane	MCA	Univ Salah Boubnider Constantine 3
	MADANI Hakim	Pr	Université de Batna 2
	BEKKOUCHE Salim	MCA	Univ Salah Boubnider Constantine 3
Invitée	KOUTCHOUKALI Ouahiba	MAA	Univ Salah Boubnider Constantine 3

Année universitaire : 2019 / 2020

*You don't write because you want to say something
you write because you have something to say*

« *Scott Fitzgerald* »

Remerciements

Je remercie Allah le tout puissant de m'avoir ouvert les portes du savoir. Merci ALLAH, si puissant, si grand, si généreux, si judicieux lui Qui nous a facilité le chemin et nous a béni de courage pour réaliser ce travail.

En tout premier lieu, je voudrais exprimer ma gratitude à mon directeur de thèse, le professeur BOUHELASSA Mohammed, pour m'avoir donné la possibilité de travailler avec lui dans ce projet, ainsi que pour les conseils précieux et le soutien académique abondant qui m'ont été fournis à tout moment.

Je tiens à exprimer toute ma reconnaissance à Mme KOUTCHOUKALI. O, ma co-encadreur, qui tout au long de ce travail, m'a suivi et encouragé. Je tiens à la remercier pour sa patience infinie et pour ses grands efforts pendant toute la réalisation de ce modeste travail. Elle a été pour moi un guide précieux, qui a su m'éclairer, m'aider, et m'apporter tout ce dont j'avais besoin ; son expérience et ses grandes qualités d'encadrement ont été absolument indispensables à mon travail.

Je suis également reconnaissant à Mr KOUTCHOUKALI M.S, pour son soutien motivant, sa disponibilité académique, ses conseils pertinents et son intérêt pour mon travail : les nombreuses discussions que nous avons eues, m'ont été très précieuses. Je voudrais aussi le remercier pour sa présence et son altruisme, que ce soit dans le cadre professionnel ou non.

J'exprime ma reconnaissance au professeur Rafiqul Gani (Chef de Centre d'Ingénierie de Processus Assisté par Ordinateur (CAPEC), Université Technique du Danemark (DTU)) de

m'avoir accueilli au sein de leur laboratoire et pour tous les moyens qu'il a mis à ma disposition durant mon stage.

J'exprime ma profonde gratitude à toute l'équipe administrative et scientifique de la faculté de Génie des Procédés, Université Constantine 3 Salah Boubnider.

Ensuite, je remercie Mme BOUSHABA Rihab, pour avoir accepté de présider mon jury de thèse, et Messieurs SAOULI Ouacil et MADANI Hakim, membres externes de mon jury. Je remercie également Messieurs BEKKOUCHE Salim et Merouani Slimane, jurés de l'Université Constantine 3

Je tiens également à exprimer ma gratitude pour tout le soutien que j'ai reçu de mes amis, en particulier de Laaouar youcef, Telil Okba et Bourmada Mouhamed Walid. Un remerciement spécial va à l'adresse du Dr Merouani Slimane et mon ami Kacha Ameer pour le soutien moral qu'on m'a apporté.

Enfin et surtout, je souhaite remercier tous les membres de ma famille et ma belle-famille pour leur soutien et de leur compréhension tout au long mon travail de thèse. Je suis reconnaissant à mes parents, de leur soutien motivationnel continu et de leurs prières. Mes remerciements particuliers vont également

A ma femme, pour son soutien inconditionnel, sa patience et son amour, d'autant plus qu'au cours de la période de la thèse, nos deux adorables fils, Baraa et Mohamed sont nés durant les années de ce travail.

Je remercie tous ceux ou celles qui, de près ou de loin, m'ont guidés et ont contribué à la réalisation de cette thèse.

Résumé

La cristallisation de la solution d'un ingrédient pharmaceutique actif nécessite la connaissance de la solubilité dans tout l'intervalle de température étudié au cours du procédé. Cependant, lors du développement d'un nouvel ingrédient actif, cette donnée est souvent manquante. Sa détermination expérimentale est possible mais fastidieuse. La méthode de contribution du groupe UNIFAC peut être utilisée pour prédire cette propriété physique. Plusieurs modifications de ce modèle ont été proposées depuis son développement en 1977, UNIFAC modifié de Dortmund, UNIFAC Pharma modifié, KT-UNIFAC,

Dans cette étude, une nouvelle version du modèle UNIFAC est développée, en considérant la dépendance linéaire des paramètres d'interaction avec la température comme dans UNIFAC Pharma modifié, et la décomposition des groupes structurels tels que définis par le modèle KT-UNIFAC de premier ordre.

Une nouvelle méthode, adoptée pour l'estimation des paramètres d'interaction avec une grande précision, est également présentée. Cette méthode est basée sur l'utilisation de la méthode de simplex de Nelder-Mead en effectuant tous les calculs de solubilité pour les différents systèmes binaires soluté/solvant simultanément et en changeant à chaque fois la fonction objective durant la phase d'optimisation. Cette procédure est essentielle pour la réussite de ce modèle. La méthode de simplex de Nelder-Mead, codée dans le logiciel Matlab, est convertie puis utilisée dans Visual Basic (VBA). Excel est alors utilisé comme interface pour le programme d'ajustement.

Des données relatives à environ 250 systèmes binaires ont été impliquées dans l'estimation de 228 paramètres d'interaction. Ces nouveaux paramètres ont ensuite été utilisés pour calculer le coefficient d'activité et la solubilité de nombreuses molécules dans différents solvants à différentes températures. Le modèle proposé donne de meilleurs résultats que ceux données par le modèle UNIFAC original et montre un bon accord entre la solubilité expérimentale et la solubilité calculée.

Abstract

The crystallization from solution of an active pharmaceutical ingredient requires the knowledge of the solubility in the entire temperature range investigated during the process. However, during the development of a new active ingredient, this data is missing. Its experimental determination is possible, but tedious. UNIFAC Group contribution method can be used to predict this physical property. Several modifications on this model have been proposed since its development in 1977, modified UNIFAC of Dortmund, Pharma modified UNIFAC, KT-UNIFAC,

In this study, we have developed a new concept for UNIFAC model by considering the linear temperature dependence of interaction parameters as in Pharma modified UNIFAC and structural groups as defined by KT-UNIFAC first order model.

We also present a new adopted optimization method for interaction parameters with a high precision. This method is based on the use of Nelder-Mead simplex method by carrying out all the solubility calculations for the different binary systems solute / solvent simultaneously and changing each time the objective function during the optimization step. This procedure is essential for the success of the model. The Nelder-Mead simplex method has been coded in Matlab and was converted to visual basic (VBA). Then, Excel was used as a graphical interface user for the fitting program.

About 250 binary data sets were involved in the estimation of 228 interaction parameters. These new parameters were then used to calculate activity coefficient and solubility of some molecules in various solvents at different temperatures. The model gives better results than those from the original UNIFAC and shows good agreement between the experimental solubility and the calculated one.

ملخص

إن بلورة محلول لمكون صيدلاني نشط يتطلب معرفة الذوبانية في كل نطاق درجة الحرارة الذي تمت دراسته أثناء العملية. ومع ذلك ، أثناء تطوير عنصر نشط جديد، هذه البيانات تكون مفقودة. إيجادها باستعمال التجربة ممكن ولكنه مرهق جدا. يمكن استخدام طريقة مساهمة المجموعات UNIFAC للتنبؤ بهذه الخاصية الفيزيائية. قد تم اقتراح عدة تغييرات على هذا النموذج منذ تطويره في عام 1977، من بينها UNIFAC : المعدل لدورتموند، و UNIFAC فارما المعدل ، و KT-UNIFAC

في هذه الدراسة ، قمنا بتطوير مفهوم جديد لنموذج UNIFAC من خلال اعتبار العلاقة الخطية لدرجة الحرارة مع معاملات التفاعل كما هو الحال في UNIFAC فارما المعدل، وهيكل المجموعات على النحو المحدد في نموذج KT-UNIFAC للدرجة الأولى.

نقدم أيضًا طريقة جديدة معتمدة لتقدير معاملات التفاعل بدقة عالية جدًا ، وتعتمد هذه الطريقة على استخدام طريقة Nelder-Mead simplex من خلال تنفيذ جميع حسابات الذوبان للأنظمة الثنائية المختلفة المذاب / المذيب في وقت واحد، وتغيير الدالة الهدف عدة مرات أثناء عملية التقريب ، هذه الطريقة ضرورية لنجاح هذا النموذج. تم استخدام طريقة Nelder-Mead simplex في Matlab وتحويلها إلى (VBA) Visual Basic ، وتم استخدام Excel كواجهة مستخدم رسومية لبرنامج التركيب.

وقد تم استخدام حوالي 250 من مجموعات البيانات الثنائية في تقدير 228 من معاملات تفاعل. ثم استخدمت هذه المعلمات الجديدة لحساب معامل النشاط والذوبانية لجزيئات معينة في مذيبات مختلفة عند درجات حرارة مختلفة. يعطي هذا النموذج الجديد نتائج أفضل من تلك المحسوبة بـ UNIFAC الأصلي ويظهر توافق جيد بين الذوبانية التجريبية والذوبانية المحسوبة

Sommaire

Remerciements	i
Résumé	iii
Abstract.....	iv
ملخص	v
Sommaire.....	vi
Nomenclature	ix
Table des figures.....	xiv
Liste des tableaux	xviii
Introduction générale.....	1
Chapitre 1 : Partie bibliographique.....	7
1.1 Introduction	8
1.2 Généralités sur la cristallisation	9
1.3 Solubilité.....	10
1.4 Equilibre solide-liquide.....	10
1.4.1 Diagramme de phase d'équilibre solide-liquide	10
1.4.2 Équation d'équilibre	11
1.5 Mesure de la solubilité	14
Chapitre 2 : Modèles thermodynamiques de coefficient d'activité.....	16
2.1 Introduction	17
2.2 Les modèles semi-prédictifs	17
2.2.1 Le modèle de Wilson.....	17
2.2.2 Le modèle NRTL.....	19
2.2.3 Le Modèle UNIQUAC	20
2.2.4 Le modèle NRTL-SAC.....	22
2.3 Les modèles prédictifs.....	25

2.3.1	Le modèle COSMO-SAC.....	25
2.3.2	Le modèle UNIFAC et ses différentes modifications	27
2.3.2.1	Le modèle UNIFAC original	29
2.3.2.2	Le modèle UNIFAC modifié Lyngby.....	30
2.3.2.3	Le modèle UNIFAC modifié (Dortmund).....	31
2.3.2.4	Le modèle UNIFAC Pharma Modifié	32
2.3.2.5	Le modèle KT-UNIFAC.....	33
2.4	Application des modèles de coefficient d'activité aux équilibres Liquide-Solide	34
2.5	Conclusion.....	40
Chapitre 3 : Estimation des paramètres d'interaction du modèle UNIFAC Pharma Modifié ..		41
3.1	Introduction	42
3.2	Subdivision des molécules.....	42
3.3	Paramètres de volume et de surface.....	45
3.4	Estimation des paramètres d'interaction	51
3.5	Résultats et discussions	56
3.6	Conclusion.....	62
Chapitre 4 : Développement d'un nouveau modèle UNIFAC		63
4.1	Introduction	64
4.2	Structure du modèle.....	64
4.2.1	Equations du modèle	64
4.2.2	Paramètres de groupe et de surface	66
4.3	Implémentation du modèle	74
4.4	Méthode de régression	79
4.5	Estimation des paramètres d'interaction du modèle	81
4.6	Conclusion.....	87
Chapitre 5 : Prédiction de la solubilité des substances pharmaceutiques active par le nouveau modèle		88
5.1	Introduction	89

5.2	Travail réalisé.....	89
5.3	Résultats de l'estimation.....	90
5.4	Prédiction de la solubilité.....	95
5.5	Validation des paramètres d'interaction.....	97
5.6	Comparaison avec le modèle UNIFAC.....	97
5.7	Evaluation du modèle.....	103
5.8	Conclusion.....	109
	Conclusion générale	110
	Références bibliographiques	113
	Article : Solubility Prediction of Active Pharmaceutical Compounds with the UNIFAC Model	132

Nomenclature

Lettres

Romanes

<i>Lettre</i>	<i>Description</i>	<i>Unité</i>
a_{eff}	Aire de contact entre différentes molécules	m^2
A_i	Surface totale de la cavité	m^2
$A_i(\sigma)$	Surface totale de tous les segments avec une densité de charge particulière σ	m^2
A_{mn}^{XY}	Coefficient impliquant des CI et des informations stœchiométriques	
a_{nm}	Paramètre d'interaction binaire UNIFAC entre les groupements fonctionnels m et n	
a_k	Surface moléculaire de van der Waals	m^2
b_{nm}	Paramètre d'interaction pour UNIFAC modifié entre les groupes m-n	
c_{hb}	Constante pour les liaisons hydrogène	$kcal.\text{Å}^4.mol^{-1}.e^{-2}$
c_{nm}	Paramètre d'interaction pour UNIFAC modifié entre les groupes m-n	
ΔC_{P_m}	Différence de capacité calorifique molaire entre le solide et le liquide sous-refroidi	$J.mol^{-1}.K^{-1}$
d_{mn}	Distance entre les deux segments	m
F_i	Fraction de surface	
f_2^s	Fugacité du soluté dans la phase solide	Pa
f_2^l	Fugacité du soluté dans la phase liquide	Pa

f_2^0	Fugacité à l'état standard de la phase liquide	Pa
g_{ji}	L'énergie d'interaction entre une paire de molécules i-j	J
G^E	L'énergie d'excès libre	J
$G_{i,j}$	Paramètre d'interactions binaire NRTL entre i et j	
$\Delta G_{m,s \rightarrow l}$	Enthalpie libre molaire de changement d'état (solide à liquide)	$J.mol^{-1}$
$\Delta G_{i/s}^{*res}$	Enthalpie libre de restauration des charges électroniques du composé i dans la solution S	J
$\Delta H_{m,a \rightarrow d}$	Enthalpie molaire	$J.mol^{-1}$
ΔH_t^{fus}	Enthalpie de fusion au point triple	J
ΔH_m^{fus}	Enthalpie molaire de fusion	$J.mol^{-1}$
n_i	Nombre total de segments	
$n_i(\sigma_m)$	Nombre de segments avec une densité de charge de surface discrétisée σ	
$n_x^{(m)}$	Nombre d'atomes de type X dans le groupe m	
p_i	Profil sigma	
q_i	Surface de Van der Waals de la molécule i	
Q_k	Surface moléculaire de van der Waals	
R	Constante des gaz parfaits	$J.K^{-1}.mol^{-1}$
R_k	Volume moléculaire de van der Waals	
r_i	Volume de Van der Waals de la molécule i	
r_{ji}	Pondération du segment j sur la molécule i	
r_i	Pondération totale du segment sur la molécule i	
r_n	Rayon du segment de surface réel	

r_{av}	Rayon de moyenne	
ΔS_m $a \rightarrow d$	L'écart d'Entropie	$J. K^{-1}$
T	Température	K
T_t	Température au point triple	K
T_f	Température de fusion	K
Δu_{nm}	Paramètre d'interaction pour l'interaction $n - m$	
u_{ij}	Paramètre d'interaction entre les molécules i-j	
V_j	Volume molaire à température donnée de constituant j purs	
v_k	Volume moléculaire de van der Waals	
V_i'	Fractions de volume modifié	
V_i	Fractions de volume	
V_i	Volume de la cavité	
w_{R2}	Paramètre de KT-UNIFAC	
ΔW	Energie d'échange	
X	Segment hydrophobe de NRTL-SAC	
x_2	Fraction molaire (solubilité)	
x_2^{id}	Fraction molaire (solubilité) dans le cas idéal	
X_i	Fraction molaire du composant i	
X_m	Fraction molaire du sous-groupe m dans le mélange	
${}^v x_{(m)}^i$	Indice de connectivité de valence de $i^{ème}$ ordre pour le groupe m	
$Y+$	Segment polaire attracteur de NRTL-SAC	

Y-	Segment polaire répulsif de NRTL-SAC
Z	Segment hydrophile de NRTL-SAC
z	Numéro de coordination

Lettres

Grecques

Lettres	Description
α'	Constante d'énergie inadaptée
$\alpha_{i,j}$	Facteur de dispersion entre i et j de NRTL-SAC
α_{ji}	Paramètre d'interaction (NRTL, Wilson)
Γ_k	Le coefficient d'activité résiduelle du sous-groupe k dans le mélange
$\Gamma_k^{(i)}$	Le coefficient d'activité résiduelle du sous-groupe k dans la solution pure du composant i
Γ_m^{lc}	Coefficients d'activité du segment m en solution
$\Gamma_m^{lc,i}$	Coefficients d'activité du segment m du composant i
Γ_s	Coefficient d'activité pour un segment de charge σ
γ_i	Coefficient d'activité du composé i
γ_i^C	Contribution combinatoire du coefficient d'activité du composé i
γ_i^R	Contribution résiduelle du coefficient d'activité du composé i
$\gamma_{i/s}^{SG}$	Terme combinatoire de Staverman-Guggenheim
u_{ji}	Paramètre d'interaction entre les molécules i - j
$v_k^{(i)}$	Numéro du groupe de type k dans la molécule i

σ	Densité surfacique de charge	e/m^2
σ_m	Densité surfacique de charge du segment m	e/m^2
σ_n	Densité moyenne de charge en surface sur le segment m à partir de la sortie COSMO	e/m^2
σ^*	Charge induite sur la surface du conducteur sous forme vectorielle	
σ_{hb}	Sigma coupé pour la liaison hydrogène	
$\tau_{i,j}$	Interaction entre le segment i et le segment j	
τ_{ji}	Paramètre ajustable	
θ_i	Fraction surfacique de la molécule i	
θ_m	La fraction surfacique du groupe fonctionnel m dans le mélange	
ϕ_i	Fraction volumique de la molécule i	
Φ_i	Fraction molaire du segment dans le mélange	
Ψ_{nm}	Paramètre d'interaction binaire UNIFAC entre les groupements fonctionnels m et n	
λ_{ij}	L'énergie d'interaction entre une molécule du constituant i et une autre molécule du constituant j	
Λ_{ij}	Paramètre d'interaction de Wilson	

Exposants

res	Terme résiduel (entropique)
comb	Terme combinatoire (enthalpique)

Table des figures

Figure 1,1 : Diagramme de phase d'équilibre liquide-solide	11
Figure 1,2 : Cycle thermodynamique de changement de phase solide ↔ liquide.....	12
Figure 2,1 : Schéma du principe de concept de la composition locale. Cellule élémentaire d'un mélange à « n » constituants.	19
Figure 2,2 : Principe de base du modèle COSMO-SAC : a) les charges en surface de la molécule sont “éteintes”, b) La molécule est insérée dans un conducteur parfait, c) Les charges de surface de la molécule sont restituées	26
Figure 3,1 : structure chimique de l'acide benzoïque	44
Figure 3,2 : structure chimique de l'acide 4-Hydroxybenzoïque	44
Figure 3,3 : structure chimique de l'acide (méta) 3-Hydroxybenzoïque.....	44
Figure 3,4 : structure chimique de l'acide 2-Hydroxybenzoïque	44
Figure 3,5 : Représentation graphique de solubilités corrélées par le modèle UNIFAC Pharma en fonction des solubilités expérimentales pour des systèmes binaires.	56
Figure 3,6 : Représentation graphique de solubilités corrélées par le modèle UNIFAC en fonction des solubilités expérimentales pour des systèmes binaires.	57
Figure 4,1 : Capture d'écran VBA montre comment écrire des fonctions et des interfaces....	74
Figure 4,2 : Capture d'écran du menu principal	75
Figure 4,3 : Capture d'écran Excel, introduction des fractions molaires et des températures et l'obtention des valeurs calculées de γ_i et de x_i à différentes températures.	75
Figure 4,4 : Capture d'écran Excel, la matrice des paramètres d'interaction a_{mn}	75
Figure 4,5 : Capture d'écran Excel, la matrice des paramètres d'interaction b_{mn}	76

Figure 4,6 : Capture d'écran Excel, décomposition des structures chimiques des molécules avec les groupes fonctionnels correspondants.....	76
Figure 4,7 : Capture d'écran Excel, les paramètres de volume R_k et de surface Q_k	76
Figure 4,8 : Capture d'écran Excel, partie de calcul de la partie combinatoire.	77
Figure 4,9 : Capture d'écran Excel, calcul de la partie résiduelle	77
Figure 4,10 : Capture d'écran Excel, calcul de la partie résiduelle des composés purs.....	77
Figure 4,11 : Capture d'écran du logiciel d'optimisation du Simplex de Nelder-Mead.....	78
Figure 4,12 : Capture d'écran du manuel d'utilisation du logiciel d'optimisation	78
Figure 4,13 : Capture d'écran Excel : mode de décomposition des molécules.....	81
Figure 4,14 : Capture d'écran Excel : introduction des solubilités et des températures	82
Figure 4,15 : Capture d'écran : sélection du soluté et du solvant	82
Figure 4,16 : Organigramme de la méthode d'estimation des paramètres d'interaction par la méthode du Simplex de Nelder-Mead.....	84
Figure 4,17 : Schéma général de l'estimation globale : fonctionnement du Main_system avec les Sub_system(s).....	85
Figure 4,18 : Organigramme détaillant le fonctionnement entre le main_system et un sub_system	86
Figure 4,19 : Organigramme de la méthode de prédiction de la solubilité	87
Figure 5,1 : Matrice des paramètres d'interaction estimés pour 20 groupes fonctionnels	90
Figure 5,2 : Représentation graphique des solubilités corrélées en fonction des solubilités expérimentales pour des systèmes binaires utilisés dans l'estimation des groupes ; CH ₂ , ACH, OH, CH ₃ OH, ACOH, ACNH ₂ , ACNO ₂ , ACCOOH.....	96
Figure 5,3 : Représentation graphique des solubilités corrélées en fonction des solubilités expérimentales pour des systèmes binaires utilisés dans l'estimation des groupes ; CH ₂ CO, CCOO, CH ₂ O, CCl, CN.....	96

Figure 5,4 : Représentation graphique des solubilités prédites pour des systèmes non utilisés dans l'estimation en fonction des solubilités expérimentales.	97
Figure 5,5 : Solubilité corrélée en fraction molaire de l'acide picrique dans l'hexane, le cyclohexane, le dodécane, le décane, l'heptane et l'isooctane à 298,15 K	98
Figure 5,6 : Solubilité prédite en fraction molaire de l'acide salicylique dans l'hexane, le dodécane, le décane, l'heptane et l'isooctane à 298,15 K	98
Figure 5,7 : Solubilité corrélée en fraction molaire de l'acide 5-amino-salicylique dans le méthanol à différentes températures 297.15, 299.15, 301.15, 303.15, 305.15 K.....	99
Figure 5,8 : Solubilité prédite en fraction molaire de l'acide 3-nitrobenzoïque dans le 1-propanol, le 1-pentanol, le 2-butanol, le 1-hexanol et le 1-octanol à 298,15 K	99
Figure 5,9 : Solubilité prédite en fraction molaire du biphényle dans l'heptane, le 1-pentanol, le 1-heptanol, le 2-méthyl-1-pentanol et le 2-méthyl-2-butanol à 298,15 K.....	100
Figure 5,10 : Solubilité prédite en fraction molaire de l'acide benzoïque dans le 2-propanol, le 1-butanol, le 1-pentanol et le 1-heptanol à 298,15 K	100
Figure 5,11 : Solubilité prédite en fraction molaire du hexachlorobenzène dans le 2-butanol, le 2-bentanol, 2-propanol, 1-butanol, 2-methyl-2-butanol, 1-hexanol, octane à 298,15 K....	101
Figure 5,12 : Solubilité prédite en fraction molaire de l'acide salicylique dans l'acétone à différentes températures 288.15, 298.15, 308.15, 318.15 K	101
Figure 5,13 : Solubilité prédite en fraction molaire de l'acide salicylique dans l'éthyle acétate à différentes températures 288.15, 298.15, 308.15, 318.15 K.....	102
Figure 5,14 : Solubilité prédite en fraction molaire de l'hydroquinone dans l'éthyle acétate à différentes températures 290.15, 298.15, 304.15, 310.15 316.15	102
Figure 5,15 : Solubilité prédite en fraction molaire du methyl palmitate dans l'acétone à différentes températures 290,15, 298,15, 304.15, 310.15 316.15 K	103

Figure 5,16 : Solubilités prédites par NK-UNIFAC pour des systèmes non utilisés dans l'estimation en fonction des solubilités expérimentales 106

Figure 5,17 : Solubilités prédites par UNIFAC original pour les mêmes systèmes en fonction des solubilités expérimentales 107

Liste des tableaux

Tableau 2,1 : Paramètres d'interaction binaire entre les segments conceptuels NRTL-SAC .	24
Tableau 3,1 : Groupes fonctionnels du modèle UNIFAC Pharma Modifié, avec leur numéro respectif dans la matrice, pour l'estimation des nouveaux paramètres	43
Tableau 3,2 : Paramètres de volume et de surface du modèle UNIFAC Pharma Modifié	45
Tableau 3,3 : Paramètres d'interaction a_{mn} et b_{mn} estimés pour UNIFAC Pharma.....	51
Tableau 3,4 : Paramètres d'interaction b_{mn} estimés pour UNIFAC Pharma	55
Tableau 3,5 : Déviations des solubilités prédites par UNIFAC Pharma Modifié et UNIFAC original par rapport aux solubilités expérimentales.....	58
Tableau 4,1 : Groupes et paramètres de volume R_k et de surface Q_k	66
Tableau 5,1 : Groupes fonctionnels.....	91
Tableau 5,2 : Paramètres d'interaction a_{mn} et b_{mn} estimés	91
Tableau 5,3 : Déviations relatives entre solubilités prédites et expérimentales pour les modèles NK-UNIFAC et UNIFAC original.....	104

Introduction générale

La connaissance des propriétés des produits chimiques est fondamentale pour la conception et l'analyse des industries chimiques, pharmaceutiques, alimentaires, agrochimiques et connexes. Les données, relatives à ces propriétés nécessaires, peuvent être obtenues de trois manières différentes : (i) en récupérant les informations des propriétés disponibles dans les bases de données / la littérature ouverte ; (ii) en effectuant des mesures au laboratoire, et / ou (iii) en employant des méthodes de prédiction de propriétés appropriées.

Le moyen le plus courant et le plus fiable pour obtenir ces propriétés physiques consiste à expérimenter. En outre, en fonction de la propriété étudiée, un ingénieur chimiste s'appuie sur les résultats de différentes techniques expérimentales pour appliquer les connaissances acquises à l'échelle industrielle. Cependant, en raison de la complexité croissante des structures moléculaires de nombreux nouveaux composés (molécules plus grosses, groupes fonctionnels nouveaux et jamais rencontrés auparavant), le coût et le temps nécessaires à certaines expériences nécessitent la prise en compte d'autres alternatives. Cela devrait également être vrai lorsque certaines expériences ne peuvent pas être effectuées pour des raisons de sécurité (ou de danger). Des modèles mathématiques ont été développés et continuent d'être développés afin d'estimer, de corréler et de prédire les propriétés thermodynamiques des composés purs et des mélanges. La principale raison d'utiliser de tels modèles d'estimation de ces propriétés, est qu'ils représentent un moyen alternatif pour obtenir des valeurs précises de manière rapide et peu coûteuse par rapport aux expériences. Cependant, au fil du temps, le travail d'un développeur de modèle devient plus difficile en raison des exigences industrielles relatives aux propriétés des nouvelles substances découvertes ou de la prédiction des propriétés complexes qui sont devenues importantes dans une application pertinente. L'importance de ces modèles dans l'industrie chimique est cruciale, car la plupart des entreprises liées aux produits chimiques utilisent ces modèles intégrés dans leurs outils informatiques (simulateurs de processus et / ou outils de prédiction des propriétés) pour la simulation et le criblage des processus. Les simulateurs des processus tels qu'ASPEN, PROII et HYSYS reposent sur des modèles de propriétés pour la simulation

de leurs processus de séparation. D'autre part, lors de la modélisation et de la conception des équipements, ces propriétés sont essentielles pour une approche approfondie ou simplement pour prendre en compte la non-idéalité des systèmes durant l'étude.

Par conséquent, le manque de propriétés physiques et thermodynamiques nécessaires dans les bases de données restreint l'utilisation d'outils informatiques pour la synthèse-conception et la modélisation-simulation des processus chimiques. Bien que l'utilisation des valeurs des propriétés mesurées expérimentalement soit hautement souhaitable, les mesures au laboratoire peuvent prendre beaucoup de temps, être coûteuses et parfois même ne pas être réalisables. Par conséquent, il est plus pratique et plus commode d'utiliser des méthodes de prédiction pour obtenir les informations nécessaires relatives à ces propriétés, au moins aux premières étapes du processus et / ou de la conception du produit.

Plusieurs types de méthodes de prédiction des propriétés sont disponibles dans la littérature. Parmi ces méthodes, les méthodes de prédiction des propriétés basées sur la contribution de groupes (GC) sont largement utilisées dans la conception des processus et / ou des produits. Elles ont l'avantage d'être rapides, efficaces et ne nécessitent pas d'importants efforts de calcul. Bien que les applications des méthodes de contribution de groupes (pour les composés purs et pour les mélanges) dans les industries chimiques et pétrochimiques soient bien connues, il n'en est pas de même pour l'industrie pharmaceutique. Les simulateurs des processus commerciaux ne disposent pas de données et / ou de modèles de propriétés physiques et thermodynamiques nécessaires pour de nombreuses substances pharmaceutiques dans leurs bases de données, limitant ainsi la large application des méthodes et outils assistés par ordinateur pour la synthèse, la modélisation et la simulation des processus au sein de cette industrie. Les coûts associés aux processus de séparation représentent souvent une très grande partie du coût total d'une usine de traitement des substances pharmaceutiques. Par conséquent, les prédictions précises et fiables de l'équilibre des phases deviennent importantes.

L'une des principales opérations unitaires de l'industrie pharmaceutique, est la cristallisation. Au cours de cette opération, l'obtention de produits cristallisés présentant les caractéristiques souhaitables revêt une importance primordiale. Pour atteindre cet objectif, le choix d'un solvant approprié pour la cristallisation, les conditions de fonctionnement et les paramètres de conception sont des mesures très importantes pour améliorer la séparation et la purification, réduire les coûts et améliorer la qualité des produits. Les solides produits dans une unité de cristallisation affecteront les processus unitaires en aval, tels que la filtration et le séchage [1]. La cristallisation en solution est largement utilisée dans les industries chimique et

pharmaceutique au cours des phases finales et intermédiaires des processus de purification et de séparation [2].

Au cours de ces processus, la solubilité des composés organiques solides dans l'eau ou dans d'autres solvants représente une propriété thermodynamique essentielle. En raison de la contrainte temps et du coût de la réalisation des mesures dépendant de la température pour de nombreux systèmes binaires différents, la disponibilité d'une méthode fiable pour prédire cette propriété est d'une importance primordiale. Les outils, permettant d'estimer rapidement la solubilité en fonction de la température et de la composition du solvant, sont considérés comme cruciaux dans la pratique de l'ingénierie d'aujourd'hui. Une grande variété de modèles, mettant particulièrement l'accent sur la solubilité dans l'eau, ont été développés de nos jours. Mis à part les corrélations empiriques, la plupart des approches thermodynamiques reposent sur la description de ces systèmes par des modèles tels que l'équation de Wilson [3], NRTL [4] ou UNIQUAC [5]. Même s'ils fournissent des résultats satisfaisants pour certains mélanges, ces modèles n'ont aucune capacité de prédiction pour les nouveaux systèmes. Par conséquent, d'autres approches doivent être utilisées, c'est le cas des modèles de contribution des groupes tels que UNIFAC [6]–[9] et ses différentes versions ou les modèles semi-prédictifs tel que NRTL-SAC. Ce dernier, dérivé du modèle de polymère NRTL, et basé sur la conception de segments de groupe, a été proposé à des fins de conception pharmaceutique. Il montre des résultats qualitatifs mais nécessite par contre des mesures expérimentales [10,11]. Les modèles d'équation d'état d'association, pour la description des équilibres de phase, telle que l'équation d'état cubique plus association, ont donné des résultats précis, mais nécessitent un très haut degré de paramétrage [12]. En outre, une nouvelle méthodologie basée sur la procédure de modèle de filtrage de type conducteur (COSMO-RS) pour solvants réels a été proposée, démontrant une bonne capacité à prédire la solubilité, mais uniquement en ce qui concerne les mélanges binaires aqueux [13].

Pour modéliser et prédire les données manquantes, les méthodes de contribution des groupes sont souvent utilisées. Elles sont basées sur le concept que les propriétés d'une molécule se déduisent de manière additive des groupes fonctionnels qui la composent. L'avantage de ce concept est évident, puisqu'il permet de prédire les interactions entre molécules, compte tenu la connaissance des interactions entre groupes fonctionnels (relativement peu nombreux) qui les constituent. L'introduction de ce concept dans le modèle thermodynamique UNIQUAC [5], a donné naissance au modèle UNIFAC [14]. Le modèle UNIFAC original et ses extensions [15–19] ont été utilisées pour prédire, à travers le

coefficient d'activité, les équilibres liquide-vapeur mais depuis lors, il a été étendu aux équilibres liquide-liquide [20] puis, solide-liquide [21–24]. Or, dans ce dernier cas, même si le modèle UNIFAC représente convenablement les systèmes simples tels que les mélanges de solutés non électrolytes dans différents solvants [25], il n'en est pas de même pour les molécules complexes contenant plusieurs hétéroatomes [26]. Le problème avec les structures des molécules pharmaceutiques est l'absence des paramètres d'interaction représentant ces groupes dans la matrice UNIFAC existante. Et même si certains ont été déterminés, ils ont été ajustés selon une base de données de produits chimiques plus simples dont les structures sont beaucoup plus petites et échouent généralement à représenter les molécules complexes de manière adéquate [27]. De plus, un manque de données expérimentales de solubilité et de propriétés physiques pour ces molécules est réel.

Ainsi, la modélisation et la prédiction de la solubilité des nouveaux composés pharmaceutiques complexes deviennent ardues en raison de l'absence de données disponibles dans la littérature. Tous ces problèmes justifient les efforts déployés pour mettre au point des modèles de prédiction de la solubilité des ingrédients actifs pharmaceutiques et de leurs mélanges. Et ce, dans le but de réduire les délais et les coûts liés à la conception de processus pharmaceutiques.

Récemment, l'estimation de la solubilité des produits pharmaceutiques actifs, dans les alcanes, les alcools, l'eau et leurs mélanges a été réalisée en utilisant différentes versions du modèle UNIFAC [28] : UNIFAC, UNIFAC Modifié (Dortmund) et UNIFAC Pharma Modifié. Ce dernier, a été développé dans le but d'étendre l'applicabilité des méthodes de contribution de groupes aux ingrédients actifs pharmaceutiques (API). Il a été conclu que, parmi les trois méthodes de contribution des groupes, le modèle UNIFAC donne les résultats les moins précis, alors qu'UNIFAC Modifié (Dortmund) et UNIFAC Pharma Modifié fournissent des résultats similaires et satisfaisants. Cette étude montre que le modèle prédictif UNIFAC pourrait être un modèle adéquat pour estimer la solubilité de l'API si trois conditions sont remplies. D'abord, tous les groupes fonctionnels qui composent ces molécules doivent être inclus dans la matrice des paramètres. Ensuite, une importante base de données expérimentale de solubilité de ces substances doit être disponible pour effectuer une régression des paramètres d'interaction de groupes. Enfin, la dépendance de la solubilité à la température doit être également prise en compte. Plusieurs modifications du modèle UNIFAC original tel UNIFAC modifié [29], [30] et ses extensions [31], [32], KT-UNIFAC [33],

UNIFAC Pharma Modifié [34], ont déjà été proposées pour améliorer la performance du modèle.

Un modèle prédictif tel qu'UNIFAC Pharma Modifié permet de prédire la solubilité d'un soluté où aucune donnée expérimentale n'est disponible. Cependant, la matrice du modèle doit être révisée, élargie et complétée. Cela nécessite une augmentation du nombre de données de solubilité introduites dans la régression des paramètres d'interaction afin de générer davantage de paramètres d'interaction de groupe des mélanges binaires soluté-solvant sur toute la gamme de composition, de sorte que le modèle devienne également applicable dans les mélanges de solvants.

Le travail de cette thèse de doctorat porte essentiellement sur les modèles de coefficient d'activité et en particulier, les modèles UNIFAC. Il se limite aux systèmes organiques impliqués dans de nombreuses opérations unitaires des processus chimiques et biochimiques de l'industrie pharmaceutique. Les électrolytes, les polymères et les gaz non condensables ne sont pas pris en compte dans ce travail.

L'objectif de ce travail sous-tend deux accomplissements. Le premier est de développer une méthodologie pour estimer les paramètres d'interaction des groupes manquants dans le modèle UNIFAC et ses différentes versions modifiées. Le second est de développer et valider une nouvelle version du modèle UNIFAC. Cette dernière a été développée en prenant en considération la dépendance linéaire de la température des paramètres d'interaction comme dans UNIFAC Pharma modifié d'une part, et la décomposition moléculaire telle que définie par le modèle KT-UNIFAC du premier ordre, d'autre part. En outre, un processus d'optimisation de grande précision, basé sur une nouvelle méthodologie, a été utilisé pour générer les paramètres d'interaction. Le modèle proposé dans le présent travail, a été validé en comparant les solubilités prédites à des jeux de données expérimentales non utilisés pour la régression.

Cette thèse comporte cinq chapitres :

Le premier chapitre s'intéresse à l'importance de la solubilité et à sa détermination et présente une étude bibliographique générale relative aux propriétés thermodynamiques des corps purs et des mélanges, en passant par les définitions de la fugacité et du coefficient d'activité. Ce chapitre aborde, par ailleurs, les aspects thermodynamiques qui régissent les équilibres de phase liquide-solide.

Le deuxième chapitre passe en revue les différents modèles thermodynamiques prédictifs et semi-prédictifs développés pour le calcul du coefficient d'activité. Le modèle de contribution de groupe UNIFAC et les modèles qui en dérivent y sont abordés de manière plus détaillée.

Dans *le troisième chapitre*, une estimation de certains paramètres d'interaction absents de la matrice des groupes fonctionnels du modèle UNIFAC Pharma Modifié a été réalisée. Ceci dans le but d'améliorer la performance du modèle dans le calcul de la solubilité des ingrédients actifs pharmaceutiques (API's).

Le quatrième chapitre propose le modèle développé au cours de cette thèse. Il met l'accent d'une part, sur l'implémentation du modèle dans Visual Basic pour Applications (VBA) qui utilise Excel de Microsoft Office comme interface et d'autre part, sur la méthode d'optimisation utilisée pour l'estimation des paramètres d'interaction du modèle. Le travail de programmation accompli est très conséquent. La méthode d'optimisation exécutée dans VBA à travers l'interface Excel est très intéressante.

Dans *le cinquième chapitre*, le modèle développé est testé. Dans ce but, les données de solubilité, disponibles dans la littérature, de plus de 400 systèmes binaires dans des solvants purs, ont été utilisées pour estimer les paramètres d'interaction de plus de 20 groupes fonctionnels. Pour la validation du modèle, la solubilité d'environ 90 systèmes binaires soluté/solvant, non pris en considération dans l'étape d'optimisation, a été prédite. La comparaison des valeurs ainsi calculées et celles prédites par le modèle UNIFAC original avec les données expérimentales montre que le nouveau modèle donne de meilleurs résultats.

Partie bibliographique

Sommaire :

1.1	Introduction	8
1.2	Généralités sur la cristallisation	9
1.3	Solubilité.....	10
1.4	Equilibre solide-liquide.....	10
1.4.1	Diagramme de phase d'équilibre solide-liquide	10
1.4.2	Équation d'équilibre	11
1.5	Mesure de la solubilité	14

1.1 Introduction

La solubilité d'un composé solide dans un solvant est une propriété physico-chimique fondamentale, particulièrement utile dans une grande variété de phénomènes relatifs aux sciences biologiques, organiques, pharmaceutiques, environnementales et en sciences de l'ingénieur. D'un point de vue thermodynamique simple, la solubilité se traduit par l'égalité de la fugacité, de la molécule de soluté, entre l'état solide et l'état liquide. En pratique, c'est la quantité maximale de la forme cristalline du soluté qui peut être dissoute, dans un volume donné de solvant, à température et pression données. Cet état est l'expression de l'énergie libre de mélange de Gibbs, représentée par la combinaison de deux contributions. La contribution enthalpique qui reflète les interactions soluté-soluté et soluté-solvant ; et la contribution entropique qui prend en compte la taille moléculaire et le potentiel de distribution non aléatoire des molécules dans le mélange. Ensuite, la taille moléculaire et le potentiel de distribution non aléatoire sont également importants [35].

La solubilité dépend de plusieurs facteurs tels que la composition chimique, la température, le pH, la présence d'impuretés dans la solution et l'utilisation de différents solvants [36]. En particulier lorsqu'il s'agit d'espèces organiques (ou inorganiques dans des solvants non aqueux), une grande variété de solvants et de mélanges de solvants peut être utilisée. Bien qu'il existe dans la littérature un nombre important de données relatives à la solubilité de nombreux systèmes binaires et ternaires en solution aqueuse [37–40], beaucoup de combinaisons de solvants restent à étudier. La mesure de la solubilité à l'aide de méthodes expérimentales est possible mais reste une opération longue et coûteuse. En effet, elle nécessite généralement une grande quantité de soluté pur, souvent disponible en quantité limitée et très coûteux dans les premiers stades de développement d'un médicament [41].

Les données de solubilité sont essentielles pour concevoir les procédés de séparation en particulier lors de l'opération de cristallisation. Cependant, il y a un manque significatif de données de solubilité concernant les nouveaux produits industriels (tels que les produits pharmaceutiques) et les produits chimiques de spécialité [36]. En outre, les systèmes industriels forment fréquemment des polymorphes ou impliquent des impuretés et la solubilité réelle est, dans ce cas, d'une importance particulière [42]. Ainsi, de nombreux efforts sont déployés pour obtenir les données requises de solubilité en un minimum de temps et une

consommation de substance réduite [43]. Par ailleurs, la solubilité peut être obtenue soit, par mesure expérimentale, soit, à l'aide d'une méthode d'estimation ou de prédiction. Dans ce dernier cas, la méthode est limitée par la disponibilité des données thermodynamiques des composés purs et des valeurs de coefficient d'activité [44–47].

Ainsi, la connaissance de ces données est d'une importance évidente aussi bien pour la conception du procédé de cristallisation lui-même (en particulier la cristallisation par refroidissement ou la cristallisation anti-solvant) que pour le développement des modèles prédictifs de la solubilité dans un solvant pur ou un mélange binaire de solvants [48,49].

L'objectif de ce chapitre est de présenter certaines des notions clés de la cristallisation, en s'appuyant sur les recherches antérieures, en particulier celles se rapportant à solubilité d'un soluté dans une solution et aux équilibres solide-liquide. Enfin, on décrit au dernier paragraphe les différentes méthodes de mesure expérimentale de la solubilité.

1.2 Généralités sur la cristallisation

On peut définir la cristallisation comme une opération de purification et de séparation définie comme un changement de phase au cours duquel une phase solide cristalline se développe à partir d'un solide fondu ou d'une solution. Il s'agit d'un procédé classique dans l'industrie chimique. Elle a de multiples applications : procédé de purification, séparation, méthode de production, récupération de matériaux solides ... etc. De ce fait, peu d'industries chimiques, aujourd'hui, n'en fait pas appel ; il y a une étape de cristallisation à un moment ou à un autre. De plus, ce procédé permet de fabriquer des produits finaux dont les propriétés d'utilisation finale peuvent y être contrôlées. C'est le cas de l'industrie pharmaceutique pour laquelle, le contrôle du procédé se doit d'être particulièrement strict en raison de l'utilisation finale des produits. En outre, les propriétés telles que la structure cristalline (polymorphes), la morphologie, la distribution de taille ou encore la taille des cristaux doivent être parfaitement maîtrisées pour répondre aux objectifs thérapeutiques des produits. Effectivement, ces propriétés conditionnent la biodisponibilité des médicaments ou la vitesse de dissolution *in vivo* et, par conséquent, l'effet du principe actif dans l'organisme. Par ailleurs, le procédé doit être complètement reproductible. [50].

1.3 Solubilité

La solubilité est définie comme la quantité de substance solide qui passe en solution pour obtenir une solution saturée à température et pression constantes. Elle est souvent exprimée en termes de volume ou de quantité maximale du soluté qui se dissout dans un volume ou une masse de solvant donné. Les pharmacopées donnent la solubilité en termes de nombre de parties en volume de solvant requis pour dissoudre une partie en poids d'un solide, ou une partie en volume d'un liquide [51].

La solubilité est fortement influencée par de nombreux facteurs et en particulier le choix du solvant. La présence de solvants est essentielle dans toutes les étapes des processus pharmaceutiques (réaction, séparation et formulation) [52]. Pour des raisons toxicologiques, les fabricants de médicaments sont de plus en plus tenus de minimiser les quantités et le nombre de solvants utilisés dans le traitement. La sélection des solvants est donc un domaine important de la thermodynamique appliquée dans les entreprises pharmaceutiques. On estime que 30% du travail d'un groupe thermodynamique industriel peut être lié à divers aspects de la sélection de solvants [53]. Par ailleurs, les molécules traitées dans le domaine pharmaceutique sont sensiblement différentes des produits pétrochimiques communs. La forte polarisabilité, la présence d'hétéroatomes et l'existence de multiples groupes fonctionnels dans les produits pharmaceutiques rendent les molécules sensibles à une grande variété d'interactions spécifiques avec les solvants polaires. Par exemple, protonation, liaison hydrogène, solvatation spécifique....[54].

1.4 Equilibre solide-liquide

L'étude de l'équilibre solide-liquide (SLE) est d'un grand intérêt technique pour le développement et la conception de processus de séparation, tels que la cristallisation et le fractionnement. Les procédés de cristallisation sont utilisés pour la séparation des mélanges.

1.4.1 Diagramme de phase d'équilibre solide-liquide

Les données d'équilibre solide-liquide sont communément représentées sous la forme de diagrammes de phase représentant la température en fonction de la solubilité en termes de fraction molaire (diagramme T-X). La solubilité est exprimée par la fraction molaire du soluté solide dissout en solution pour les petites molécules et par la fraction massique pour les macromolécules [55].

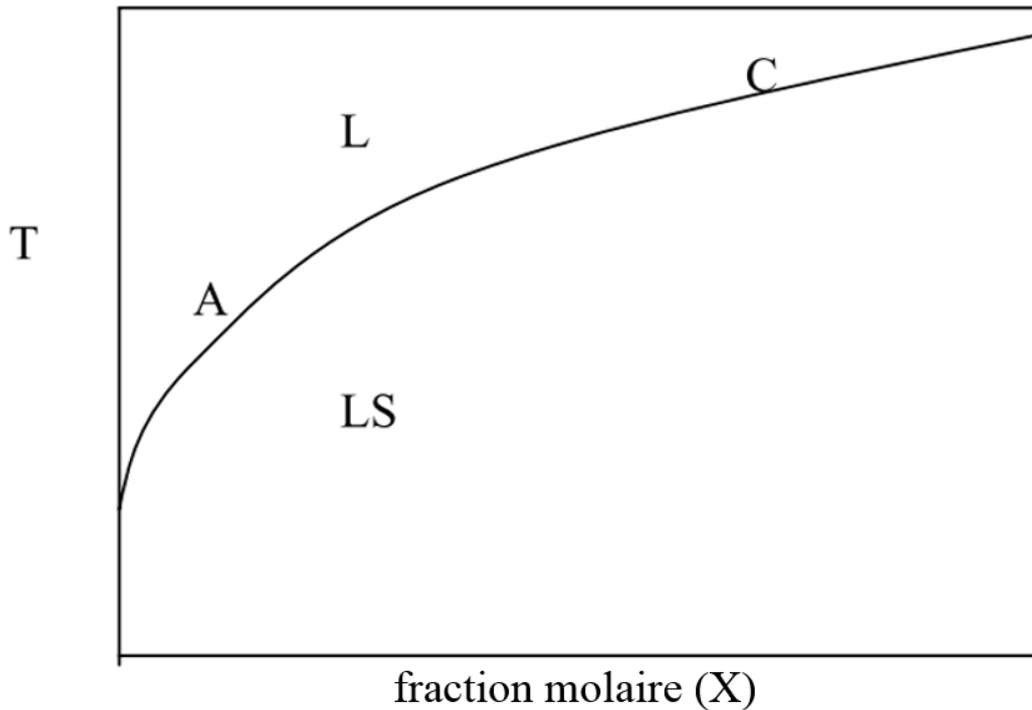


Figure 1,1 : Diagramme de phase d'équilibre liquide-solide

La courbe LS continue (AC) divise le domaine T-X en deux régions, une région liquide (L) à température plus élevée et une région solide-liquide (LS) à basse température. Dans la région SL, deux phases coexistent, le solide cristallin et la solution liquide.

1.4.2 Équation d'équilibre

Dans ce paragraphe, on désignera le solvant par l'indice 1 et le soluté par l'indice 2. L'équation d'équilibre, traduisant la solubilité d'un soluté 2 dans un solvant 1, s'exprime comme tout autre équilibre de phase, par l'égalité entre les fugacités du composé dans ses deux phases [6]. Ainsi, dans le cas de l'équilibre liquide-solide la solubilité est donnée par :

$$f_2^s = f_2^l \quad (1.1)$$

Les exposants s et l réfèrent à la phase liquide et à la phase solide respectivement.

f_2^s et f_2^l représentent les fugacités du soluté dans la phase solide et dans la phase liquide respectivement.

Si la phase solide est un composé pur, alors la fugacité du soluté solide pur f_2^s est donnée par la relation suivante [56] :

$$f_2^l = \gamma_2 x_2 f_2^0 \quad (1.2)$$

Où, x_2 est la fraction molaire, γ_2 est le coefficient d'activité de la phase liquide, et f_2^0 correspond à la fugacité du soluté pur dans un état standard hypothétique appelé liquide sous-refroidi. Les propriétés de cet état standard hypothétique peuvent être calculées avec précision, particulièrement si la température de la solution n'est pas loin du point triple du soluté [56].

En remplaçant l'équation 1.2 dans l'équation 1.1 on obtient :

$$f_2^s = \gamma_2 x_2 f_2^0 \quad (1.3)$$

En reformulant l'équation, on obtient :

$$x_2 \gamma_2 = \frac{f_2^s}{f_2^0} \quad (1.4)$$

Les fugacités f_2^s et f_2^0 sont dépendantes des propriétés du soluté et de sa phase solide, et ne dépendent pas du solvant (l'importance du solvant paraît dans le terme γ_2). Le rapport entre ces deux dernières peut être directement relié à l'écart d'enthalpie libre molaire du soluté dans les phases solide et liquide :

$$\Delta G_m = RT \ln \frac{f_2^0}{f_2^s} = RT \ln \frac{1}{x_2 \gamma_2} \quad (1.5)$$

Le schéma de la Figure I.2 représente le cycle thermodynamique du changement de phases, l'équation I.5 devient :

$$\Delta G_m = \Delta G_m = \Delta H_m - T \Delta S_m \quad (1.6)$$

Ensuite, ΔH_m se décompose en suivant les étapes b et c du cycle :

$$\Delta H_m = \Delta H_m + \Delta H_m + \Delta H_m \quad (1.7)$$

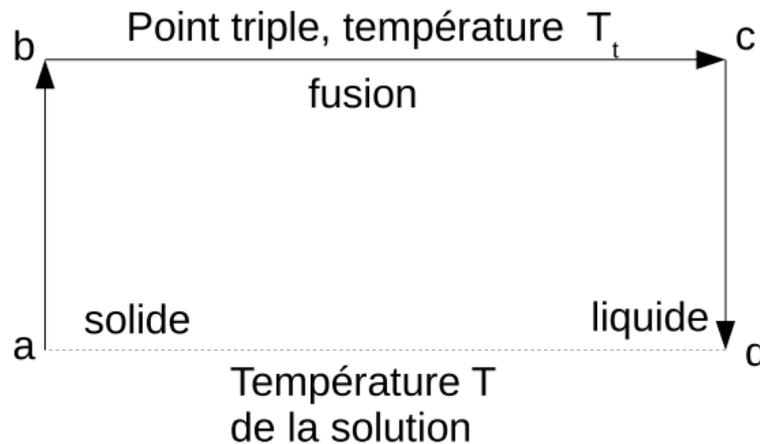


Figure 1.2 : Cycle thermodynamique du changement de phase solide ↔ liquide.

Ensuite, en considérant ΔC_{P_m} la différence de capacité calorifique entre le solide et le liquide sous-refroidi $\Delta C_{P_m} = \Delta C_{P_m} = C_{P_m}^{liquide} - C_{P_m}^{solide}$ et T_t la température au point triple on obtient :

$$\Delta H_m = \Delta H_m^{fus}(T_t) + \int_{T_t}^T \Delta C_{P_m} dT \quad (1.8)$$

De la même manière :

$$\Delta S_m = \Delta S_m^{fus}(T_t) + \int_{T_t}^T \frac{\Delta C_{P_m}}{T} dT = \frac{\Delta C_{P_m}}{T} \quad (1.9)$$

Comme au point triple, $\Delta S_m^{fus} = \frac{\Delta H_m^{fus}}{T_t}$, on obtient, en supposant que ΔC_{P_m} est constant dans le domaine de température entre T et T_t , et en introduisant dans l'équation 1.5 les équations précédentes :

$$\ln \frac{1}{x_2 \gamma_2} = \frac{\Delta H_m^{fus}}{RT_t} \left[\frac{T_t}{T} - 1 \right] + \frac{\Delta C_{P_m}}{R} \left[\ln \frac{T_t}{T} - \frac{T_t}{T} + 1 \right] \quad (1.10)$$

Sur cette dite-équation, on peut émettre quelques remarques et simplifications. De un, l'assimilation de la température de fusion T_f à la température au point triple T_t parce que pour la plupart des substances la différence entre les deux températures est minime. De deux, on assimile, habituellement, l'enthalpie de fusion au point triple ΔH_t^{fus} à l'enthalpie de fusion à la température de fusion ΔH_m^{fus} .

Les propriétés de température de fusion et d'enthalpie de fusion sont, aisément accessibles. L'équation (1.10) se réécrit alors :

$$\ln x_2 \gamma_2 = \frac{\Delta H_m^{fus}}{RT_f} \left[1 - \frac{T_f}{T} \right] - \frac{\Delta C_{P_m}}{R} \left[\ln \frac{T_f}{T} - \frac{T_f}{T} + 1 \right] \quad (1.11)$$

La deuxième simplification préoccupe l'importance que l'on attribue au terme ΔC_{P_m} . Effectivement, il est généralement pris en considération que le terme en ΔH_m^{fus} est prédominant sur les trois autres en ΔC_{P_m} . Ces derniers termes sont donc fréquemment négligés [6], dûment à cause de leur détermination expérimentale difficile due à la forte dépendance en température de capacités calorifiques au-delà de la température de fusion.

L'équation d'équilibre simplifiée la plus souvent utilisée s'exprime alors par :

$$\ln x_2 \gamma_2 = \frac{\Delta H_m^{fus}}{RT_f} \left[1 - \frac{T_f}{T} \right] \quad (1.12)$$

Dans le cas idéal pour lequel, le coefficient d'activité, qui caractérise les écarts à l'idéalité, vaut 1 ($\gamma_2 = 1$), l'expression (1.12) devient :

$$\ln x_2^{id} = \frac{\Delta H_m^{fus}}{RT_f} \left[1 - \frac{T_f}{T} \right] \quad (1.13)$$

Dans le cas d'une solution idéale, il n'y a plus de distinction entre le solvant et le soluté quant aux interactions. La solubilité idéale du soluté ne dépend que de ses propriétés thermodynamiques à l'état solide et liquide. La nature du solvant n'a subséquentement plus la moindre influence sur la solubilité.

Finalement, on peut retrouver l'équation 1.10 en décrivant l'équilibre de phase par l'égalité des potentiels chimiques des phases liquide et solide, et avec la relation de Gibbs-Helmholtz on peut la développer.

1.5 Mesure de la solubilité

De nombreuses méthodes pour mesurer la solubilité de composés organiques, telles que la chromatographie liquide à haute performance (HPLC) [57], réflexion totale atténuée (ATR) - Spectroscopie FTIR [58–60], Spectroscopie Raman [61] et mesure du mode de réflexion du faisceau focalisé (FBRM) [62]

La méthode de mesure isotherme de la solubilité, qui permet un délai prolongé pour atteindre l'équilibre, pour échantillonner et peser le résidu, et pour mesurer la concentration d'un composé dissous par une méthode analytique appropriée, est fiable [36]. Mais une telle méthode nécessite une quantité importante de matériau et un étalonnage soigné. De plus elle exige une extrapolation des données pour calculer la solubilité. Ainsi, la détermination de la solubilité en utilisant les méthodes existantes prend du temps avec des demandes importantes en main-d'œuvre. Dans l'industrie pharmaceutique, pour améliorer la prédiction de la biodisponibilité et du criblage des solvants, des méthodes expérimentales rapides de mesure précise de la solubilité ont été étudiées [63–65]. En outre, les nouveaux produits pharmaceutiques développés, qui ne sont disponibles pour analyse qu'en petites quantités, ont intensifié le besoin de méthodes utilisant des échantillons de plus en plus petits.

Les données de solubilité directe dans différents solvants ne sont pas toujours disponibles dans les industries pharmaceutiques. Les méthodes de prédiction de la solubilité sont utiles pour réduire le temps et les ressources nécessaires pour mesurer la solubilité [66].

Diverses méthodes ont été développées pour la mesure de la solubilité au cours des deux dernières décennies. Au début de la découverte, la solubilité cinétique est souvent mesurée par la turbidité [67] et systèmes de détection à base de scanners à plaques UV [67, 68]. La plupart de ces méthodes classent les dosages pour le criblage à haut débit où les solutions d'échantillon tamponné sont préparées en ajoutant des aliquotes de solutions mères de 10 mM DMSO [3]. Dans la phase de développement de la recherche sur les médicaments, la solubilité à l'équilibre des molécules de type médicament est mesurée par différentes méthodes, parmi lesquelles de nouvelles méthodes potentiométriques comme le DTT ou le CheqSol [68–72]. Cependant, la méthode de base (par rapport à laquelle les nouvelles méthodes de solubilité sont généralement validées) reste la méthode de saturation classique. Malgré les nombreuses méthodes expérimentales développées, les valeurs précises de solubilité à l'équilibre sont difficiles à obtenir, car elles sont affectées par de nombreux facteurs connus et inconnus [73, 74].

Modèles thermodynamiques de coefficient d'activité

Sommaire :

2.1	Introduction	17
2.2	Les modèles semi-prédictifs	17
2.2.1	Le modèle de Wilson.....	17
2.2.2	Le modèle NRTL.....	19
2.2.3	Le Modèle UNIQUAC	20
2.2.4	Le modèle NRTL-SAC.....	22
2.3	Les modèles prédictifs.....	25
2.3.1	Le modèle COSMO-SAC.....	25
2.3.2	Le modèle UNIFAC et ses différentes modifications	27
2.3.2.1	Le modèle UNIFAC original	29
2.3.2.2	Le modèle UNIFAC modifié Lyngby.....	30
2.3.2.3	Le modèle UNIFAC modifié (Dortmund).....	31
2.3.2.4	Le modèle UNIFAC Pharma Modifié	32
2.3.2.5	Le modèle KT-UNIFAC.....	33
2.4	Application des modèles de coefficient d'activité aux équilibres Liquide-Solide	34
2.5	Conclusion.....	40

2.1 Introduction

La prédiction de la solubilité d'un composé moléculaire solide dans un solvant, ainsi que l'estimation du coefficient d'activité de la solution à partir des données expérimentales de la solubilité, nécessitent l'estimation de l'activité du soluté dans la solution saturée. Cette dernière est souvent définie en utilisant la fusion du solide pur aux mêmes conditions thermodynamiques. Dans la littérature, l'activité du solide est généralement définie dans le même état de référence [75].

Pour estimer la solubilité, on fait souvent appel à la modélisation thermodynamique reposant sur l'enthalpie libre d'excès. Différents modèles proposent une représentation de l'enthalpie libre d'excès en fonction de la composition à température donnée pour des liquides non idéaux. Selon que ces modèles nécessitent des données expérimentales ou non ; ils se divisent en deux groupes : les modèles semi-prédictifs et les modèles prédictifs.

Les modèles semi-prédictifs les plus utilisés sont ceux de Wilson, NRTL et UNIQUAC, tandis que les modèles prédictifs sont principalement ceux dérivant de la méthode prédictive de contribution de groupe UNIFAC. Dans les paragraphes qui suivent ces différents modèles sont présentés.

2.2 Les modèles semi-prédictifs

2.2.1 Le modèle de Wilson

Développé en 1964 par Wilson [76], ce modèle, basé sur le concept de composition locale, fait appel uniquement à deux paramètres ajustables. Depuis son introduction, sous sa forme binaire, il a suscité une grande attention en raison de sa capacité à s'adapter aux systèmes fortement non idéaux. Les fractions globales du volume de la solution sont remplacées par la fraction locale du volume, liée aux ségrégations locales des molécules provoquées par les différentes énergies d'interaction entre les paires de molécules.

Le concept de composition locale diffère de la composition globale pour une solution binaire équimolaire. L'enthalpie libre d'excès pour un système binaire est donnée par :

$$\frac{G^E}{RT} = - \sum_{i=1}^n \ln \left[\sum_{j=1}^n \Lambda_{ij} x_j \right] \quad (2.1)$$

avec

$$\Lambda_{ij} = \frac{V_i}{V_j} \exp \left[\frac{-\lambda_{ij}}{RT} \right] \quad (2.2)$$

Où $i \neq j$ avec $\lambda_{ij} = a_{ij} - a_{ii}$

Dans le cas où : $i=j \rightarrow \lambda_{ij} = 1$

V_i, V_j : sont les volumes molaires à température donnée des constituants i et j purs.

λ_{ij} : est l'énergie d'interaction entre une molécule du constituant i et une autre molécule du constituant j . Donc la variation du coefficient d'activité est donnée par :

$$\ln \gamma_i = 1 - \ln \left[\sum_{j=1}^n x_j \Lambda_{ij} \right] - \sum_{k=1}^n \frac{x_k \Lambda_{ki}}{\sum_{j=1}^n (x_j \Lambda_{kj})} \quad (2.3)$$

Λ_{ij} est le paramètre d'interaction de Wilson.

Pour un mélange binaire dans lequel x_1 et x_2 représentent, respectivement les fractions molaires du solvant et du soluté, l'enthalpie libre et les coefficients d'activité sont donnés par :

$$\frac{G^E}{RT} = x_1 \ln(x_1 + x_2 \cdot \Lambda_{12}) - x_2 \ln(x_2 + x_1 \Lambda_{21}) \quad (2.4)$$

$$\ln \gamma_1 = -\ln(x_1 + x_2 \Lambda_{12}) + x_2 \left(\frac{\Lambda_{12}}{x_1 + x_2 \Lambda_{12}} - \frac{\Lambda_{21}}{x_2 + x_1 \Lambda_{21}} \right) \quad (2.5)$$

$$\ln \gamma_2 = -\ln(x_2 + x_1 \Lambda_{21}) - x_1 \left(\frac{\Lambda_{12}}{x_1 + x_2 \Lambda_{12}} - \frac{\Lambda_{21}}{x_2 + x_1 \Lambda_{21}} \right) \quad (2.6)$$

Les paramètres de Wilson Λ_{ij} peuvent être calculés à dilution infinie :

$$\ln(\gamma_1^\infty) = \ln \Lambda_{12} + 1 - \Lambda_{21} \quad (2.7)$$

et

$$\ln(\gamma_2^\infty) = -\ln \Lambda_{21} + 1 - \Lambda_{12} \quad (2.8)$$

Les paramètres d'interactions Λ_{ij} avec $i \neq j$ doivent être positifs, si $\Lambda_{12} = \Lambda_{21}$ alors la solution est idéale. Si Λ_{12} et Λ_{21} sont supérieurs à l'unité, il y a une déviation positive par rapport à l'idéalité, et dans le cas contraire c'est une déviation négative. Les paramètres Λ_{ij} sont indépendants de la composition et, à priori, aussi approximativement indépendants de la température. L'équation de Wilson, relativement simple, est, actuellement, l'une des plus recommandées pour la représentation des équilibres liquide-vapeur isobares. Cependant, son adaptation aux mélanges partiellement miscibles requiert l'introduction de paramètres supplémentaires [77].

2.2.2 Le modèle NRTL

Le modèle NRTL (Non-Random, Two-Liquids), proposé en 1968 par Renon et Prausnitz [78], est une extension de l'équation de Wilson. Il s'applique aux équilibres liquide-liquide, liquide-vapeur et liquide-liquide-vapeur pour des systèmes binaires ou à plusieurs constituants [79].

Ce modèle est basé sur la combinaison du concept de composition locale de Wilson et la théorie des deux fluides proposé par Scott en 1956 [80]. Ce dernier a proposé de considérer un mélange binaire (i, j) non comme un fluide unique mais comme un mélange de deux liquides distincts (le premier avec la molécule i au centre et le second avec la molécule j au centre). Cette théorie a permis d'exprimer l'enthalpie libre d'excès en termes de deux fluides hypothétiques par la relation : l'hypothèse que la concentration locale, autour d'une molécule, est différente de la concentration globale. La répartition non aléatoire des molécules est due à une différence entre l'énergie d'interaction de la molécule centrale avec les molécules de même espèce d'une part, et l'énergie d'interaction avec les molécules d'espèce différente, d'autre part.

Ce modèle moléculaire considère, pour un mélange de n constituants, n cellules élémentaires différentes avec la molécule occupant son centre [81].

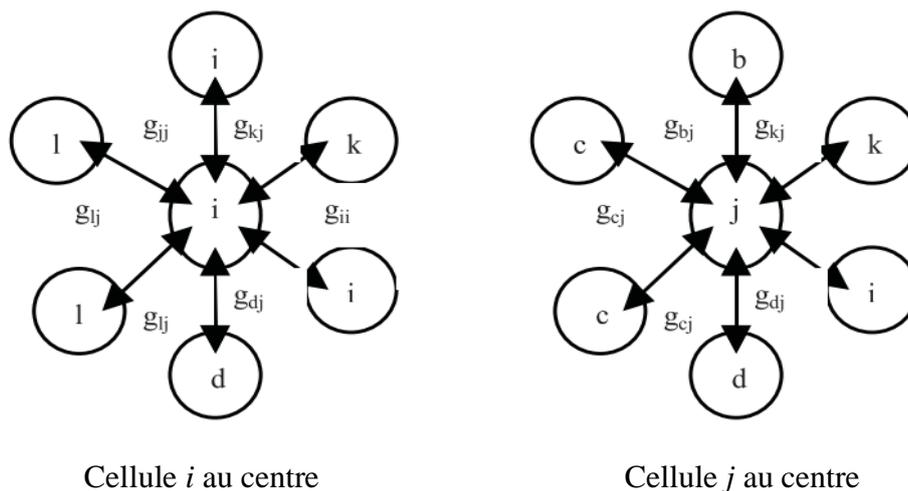


Figure 2.1 : Schéma du principe de concept de la composition locale. Cellule élémentaire d'un mélange à « n » constituants [82].

Le modèle NRTL est basé sur l'introduction de la notion de fraction molaire locale et l'expression générale pour G^E , dans le cas d'un système binaire, est :

$$\frac{G^E}{RT} = x_1 x_2 \left[\frac{\tau_{21} G_{21}}{x_1 + x_2 G_{21}} + \frac{\tau_{12} G_{12}}{x_1 G_{12} + x_2} \right] \quad (2.9)$$

La différenciation donne les expressions du coefficient d'activité suivantes :

$$\ln\gamma_1 = x_2^2 \left[\tau_{21} \left(\frac{G_{21}}{x_1 + x_2 G_{21}} \right)^2 + \frac{\tau_{12} G_{12}}{(x_1 G_{12} + x_2)^2} \right] \quad (2.10)$$

$$\ln\gamma_2 = x_1^2 \left[\tau_{12} \left(\frac{G_{12}}{x_1 G_{12} + x_2} \right)^2 + \frac{\tau_{21} G_{21}}{(x_1 + x_2 G_{21})^2} \right] \quad (2.11)$$

$$\text{Avec :} \quad \tau_{12} = \frac{g_{12} - g_{22}}{RT} \quad \tau_{21} = \frac{g_{21} - g_{11}}{RT} \quad (2.12)$$

$$G_{12} = \exp(-\alpha_{12}\tau_{12}) \quad G_{21} = \exp(-\alpha_{12}\tau_{21}) \quad (2.13)$$

g_{12} et g_{21} représentent les énergies d'interaction entre une paire de molécules 1 et 2.

Ce modèle introduit le paramètre non aléatoire, $\alpha_{12} = \alpha_{21}$. Ce dernier possède une valeur donnée, selon les espèces i et j du mélange, leur polarité et leur degré d'association.

L'équation NRTL a été appliquée pour le calcul de l'énergie libre d'excès des mélanges non électrolytes. Elle donne une bonne représentation des équilibres liquide-vapeur et liquide-liquide en ajustant trois paramètres. Elle possède l'avantage d'avoir un nombre adéquat de paramètres (trois à une température donnée) pour une bonne représentation des fortes déviations par rapport à l'idéalité. Son extension aux mélanges à plusieurs constituants ne nécessite pas de paramètres additionnels, mais uniquement les paramètres relatifs aux interactions binaires des constituants du mélange, pris deux à deux, afin de calculer la propriété de n'importe quel mélange non électrolyte [4].

2.2.3 Le Modèle UNIQUAC

Le modèle UNIQUAC [5], développé par Abrahms et Prausnitz en 1975, est un modèle semi-empirique issu de l'association entre le modèle physique quasi-réculaire de Guggenheim et le concept de composition locale. C'est le premier modèle à avoir proposé une équation en deux termes. Le premier (terme combinatoire) tient compte des différences de taille et de forme entre les molécules et le deuxième (terme résiduel) résulte des interactions entre les molécules.

L'expression générale de l'énergie d'excès de Gibbs est donnée par :

$$\frac{G^E}{RT} = \sum_i x_i \ln \frac{\phi_i}{x_i} + \frac{z}{2} \sum_i q_i x_i \ln \frac{\theta_i}{\phi_i} - \sum_i q_i x_i \ln \left[\sum_j \theta_j \tau_{ji} \right] \quad (2.14)$$

En termes de coefficient d'activité le modèle UNIQUAC s'exprime par la relation :

$$\ln\gamma_i = \ln\gamma_i^C + \ln\gamma_i^R \quad (2.15)$$

La partie combinatoire est donnée par :

$$\ln \gamma_i^c = \ln \frac{\phi_i}{x_i} + \frac{z}{2} q_i \ln \frac{\theta_i}{\phi_i} + l_i - \frac{\phi_i}{x_i} \sum_j x_j l_j \quad (2.16)$$

z : Nombre de coordination, fixé à 10,

x_i : Fraction molaire du constituant i ,

ϕ_i : Fraction de volume du constituant i ,

θ_i : Fraction de surface du constituant i ,

r_i : Volume de van der Waals de la molécule i ,

q_i : Surface de van der Waals de la molécule i ,

Avec

$$l_i = \frac{z}{2} (r_i - q_i) - (r_i - 1) \quad (2.17)$$

La partie résiduelle s'exprime par :

$$\ln \gamma_i^R = q_i \left[1 - \ln \left(\sum_j \theta_j \tau_{ij} \right) - \sum_j \left(\frac{\theta_j \tau_{ij}}{\theta_k \tau_{kj}} \right) \right] \quad (2.18)$$

Le paramètre ajustable τ_{ji} est donné par :

$$\tau_{ji} = \exp \left(\frac{u_{ii} - u_{ji}}{RT} \right) \quad (2.19)$$

Avec $\tau_{ii} = \tau_{jj} = 1$

Les fractions de volume et de surface de chaque constituant i peuvent être calculées par les relations suivantes :

$$\phi_i = \frac{r_i x_i}{\sum_{j=1}^n r_j x_j} \quad (2.20)$$

$$\theta_i = \frac{q_i x_i}{\sum_{j=1}^n q_j x_j} \quad (2.21)$$

Le modèle UNIQUAC est caractérisé par des paramètres binaires dont la valeur doit être déterminée à partir des données expérimentales, relatives pour l'essentiel, aux équilibres liquide-vapeur des mélanges binaires. L'avantage de ce type de modèle réside dans le caractère prédictif des propriétés d'équilibre pour des systèmes plus complexes (ternaire et plus). Dans le calcul des équilibres liquide-vapeur, l'équation de Wilson et les modèles NRTL et UNIQUAC conduisent à des résultats comparables [83].

2.2.4 Le modèle NRTL-SAC

Le modèle NRTL-SAC (Non-Random Two-Liquid Segment Activity Coefficient), développé par Chen et *al.* [84], est basé sur la dérivation du modèle NRTL original pour les polymères [85].

Les équations et le principe de base du calcul restent les mêmes et empruntent à UNIQUAC la division du coefficient d'activité en contribution combinatoire (entropique) et ou résiduelle (enthalpique) Equation (2.15).

L'atout de ce modèle est le principe de division des molécules considérées en segments (**le "SAC" signifie "Segment Activity Coefficient"**) qu'il offre. Inversement aux modèles de contribution de groupe, la méthode NRTL-SAC opère une division conceptuelle de chaque molécule dans la solution soit les molécules sont divisées en différentes parties conceptuelles. Ce qu'il faut comprendre du terme "conceptuelles" c'est que cette division n'est pas une représentation chimique ou physique, mais plutôt une séparation des différents comportements que les molécules sont susceptibles d'avoir. Il y a quatre comportements : polaire attracteur, hydrophobe, hydrophile et répulsif. Les segments auxquels ces comportements sont attribués sont mentionnés respectivement : Y⁻, X, Z et Y⁺.

Le fonctionnement du modèle NRTL-SAC consiste à supposer chaque molécule comme un quadruplet de valeurs correspondant à une pondération associée à chaque comportement (ou segment). Donc, l'eau sera représentée par le quadruplet [0 0 0 1] vu qu'elle a un comportement purement hydrophile. De ce fait, calculer le coefficient d'activité d'une molécule prend en considération les différentes valeurs des segments dans la solution, leurs interactions quantifiées par les paramètres d'interaction τ_{ij} . La partie combinatoire est calculée par l'équation (2.26)

$$\ln \gamma_i^c = \ln \frac{\Phi_i}{x_i} + 1 - r_i \sum_j \frac{\Phi_j}{x_j} \quad (2.22)$$

Avec les définitions :

$$r_i = \sum_j r_{ji} \quad (2.23)$$

$$\Phi_i = \frac{r_i x_i}{\sum_j r_j x_j} \quad (2.24)$$

Où :

x_i : Fraction molaire du constituant i ,

r_{ji} : Pondération du segment j sur la molécule i ,

r_i : Pondération totale du segment sur la molécule i

Φ_i : Fraction molaire du segment dans le mélange.

r_1 correspond à la valeur de X, r_2 à Y⁻, r_3 à Y⁺ et r_4 se réfère à Z.

Le terme résiduel est défini comme :

$$\ln \gamma_i^R = \ln \gamma_i^{lc} = \sum_m r_{m,i} (\ln \Gamma_m^{lc} - \ln \Gamma_m^{lc,i}) \quad (2.25)$$

Dans l'équation (2.25), les deux termes, $\ln \Gamma_m^{lc}$ et $\ln \Gamma_m^{lc,i}$ représentent les coefficients d'activité du segment m en solution et du composé i , respectivement.

Les deux termes mentionnés sont obtenus en utilisant les équations (2.26) et (2.27) :

$$\ln \Gamma_m^{lc} = \frac{\sum_j x_j G_{j,m} \tau_{j,m}}{\sum_k x_k G_{k,m}} + \sum_m \frac{x_{m'} G_{m,m'}}{\sum_k x_k G_{k,m'}} \left(\tau_{m,m'} - \frac{\sum_j x_j G_{j,m'} \tau_{j,m'}}{\sum_k x_k G_{k,m'}} \right) \quad (2.26)$$

$$\ln \Gamma_m^{lc,l} = \frac{\sum_j x_{j,l} G_{j,m} \tau_{j,m}}{\sum_k x_{k,l} G_{k,m}} + \sum_m \frac{x_{m',l} G_{m,m'}}{\sum_k x_{k,l} G_{k,m'}} \left(\tau_{m,m'} - \frac{-\sum_j x_{j,l} G_{j,m'} \tau_{j,m'}}{\sum_k x_{k,l} G_{k,m'}} \right) \quad (2.27)$$

La deuxième quantité, $\ln \Gamma_m^{lc,i}$, se calcule de la même manière que $\ln \Gamma_m^{lc}$, en ne considérant qu'une solution pure en molécule m .

Dans les deux équations (2.26) et (2.27), l se réfère au composé pur. j, k, m et m' se rapportent aux segments de chaque composé. $x_{j,l}$ est la fraction molaire du segment j dans le composé l pur uniquement. Les fractions molaires des segments dans la solution entière et dans les composants sont définies comme suit :

$$x_j = \frac{\sum_l x_l r_{j,l}}{\sum_z \sum_i x_z r_{j,z}} \quad (2.28)$$

$$x_{j,l} = \frac{r_{j,l}}{\sum_i r_{j,l}} \quad (2.29)$$

$$G_{i,j} = \exp(-\alpha_{i,j} \tau_{ij})$$

$G_{i,j}$: Paramètre d'interactions binaire NRTL entre i et j

τ_{ij} : Paramètre d'interaction entre le segment i et le segment j .

$\alpha_{i,j}$: Facteur de dispersion entre i et j de NRTL (selon l'interaction examinée, prend des valeurs différentes, en pratique 0,2 ou 0,3).

Les valeurs de ces paramètres binaires, entre deux segments conceptuels donnés, sont indiquées dans le Tableau 2.1.

Tableau 2.1 Paramètres d'interaction binaire entre les segments conceptuels NRTL-SAC

Segment (1)	X	X	Y –	Y+	X
Segment (2)	Y –	Z	Z	Z	Y+
$\tau_{1,2}$	1,643	6,547	-2,000	2,000	1,643
$\tau_{2,1}$	1,834	10,949	1,787	1,787	1,834
$\alpha_{1,2} = \alpha_{2,1}$	0,200	0,200	0,300	0,300	0,200

Les paramètres binaires entre les deux segments ont été déterminés par Chen et *al.* [84] de la manière suivante :

- 1- Pour l'interaction binaire entre les segments hydrophobe (X) et hydrophile (Z), ils ont utilisé les données expérimentales des équilibres Liquide-Liquide (LLE) pour le système hexane-eau. Les valeurs des deux paramètres sont élevées, indiquant la nature répulsive forte des deux segments.
- 2- Les valeurs d'interaction des segments binaires entre les segments hydrophobe (X) et polaire (Y) ont été déterminées à partir des données LLE du mélange hexane-acétonitrile.
- 3- Les interactions binaires entre les segments polaire (Y) et hydrophile (Z) ont été mises à jour à partir des données d'équilibre Liquide-Vapeur (VLE) disponibles dans la littérature pour le système acétonitrile-eau.
- 4- La valeur de α pour les équilibres Liquide-Liquide a été fixée à 0,2 et pour les équilibres Liquide-Vapeur à 0,3.
- 5- L'interaction entre le segment attractif polaire (Y–) et répulsif polaire (Y+) du mélange est supposée être idéale et les interactions binaires sont donc nulles.

Il est nécessaire de définir les valeurs des segments de chaque constituant dans la solution (les quadruplets [X, Y–, Y+, Z]), pour implémenter le modèle NRTL-SAC. Ceux-ci peuvent être identifiés si on connaît quatre données expérimentales de solubilité. L'intérêt principal de cette méthode est que ce quadruplet de segments moléculaires reste valable pour la même molécule dans n'importe quelle solution (une seule détermination). Ce modèle est particulièrement pratique pour la modélisation des équilibres Liquide-Solide. Il permet d'évaluer le coefficient d'activité en phase liquide de molécules complexes telles que les substances actives pharmaceutiques avec une précision "acceptable" [86].

2.3 Les modèles prédictifs

2.3.1 Le modèle COSMO-SAC

Avec les progrès de la chimie informatique, la prédiction des coefficients d'activité et des équilibres de phase sans données expérimentales est devenue possible. Les modèles basés sur COSMO (méthodes de criblage de type condensateur), tels que le modèle COSMO-SAC [87], sont maintenant des alternatives viables aux méthodes de contribution de groupe.

Le modèle COSMO-SAC (pour COnductor-like Screening MOdel for Segment Activity Coefficient) est basé sur le même principe fondateur que le modèle COSMO-RS [88]. D'abord publié par Lin et *al.* [89], ce modèle combine une approche de thermodynamique statistique et une approche de chimie quantique pour la représentation de la distribution de charge des molécules (σ -profiles). Pour déterminer le coefficient d'activité, seuls les profils σ ainsi que la surface et le volume de van der Waals sont nécessaires. Cela fait du modèle COSMO-SAC un modèle totalement prédictif.

Le modèle COSMO-SAC est basé sur le calcul de l'énergie sans solvation ΔG^{*sol} qui définit le changement d'énergie libre par rapport à une solution idéale à température et pression constantes, lorsque les molécules de soluté sont introduites dans le solvant. Dans les modèles basés sur COSMO, ce terme est divisé en deux parties. Premièrement, les charges de soluté sont désactivées et ses molécules sont insérées dans un conducteur parfait (avec une constante diélectrique infinie, ϵ). Cette première étape donne l'énergie libre de la formation de cavitation, ΔG^{*cav} . Cette dernière, dépend de la dimension spécifique de la molécule et représente la contribution entropique du coefficient d'activité. Lorsque les charges sont activées (le solvant devient un réel solvant) et l'énergie libre de restauration des charges, ΔG^{*res} , correspond alors à la contribution enthalpique [35].

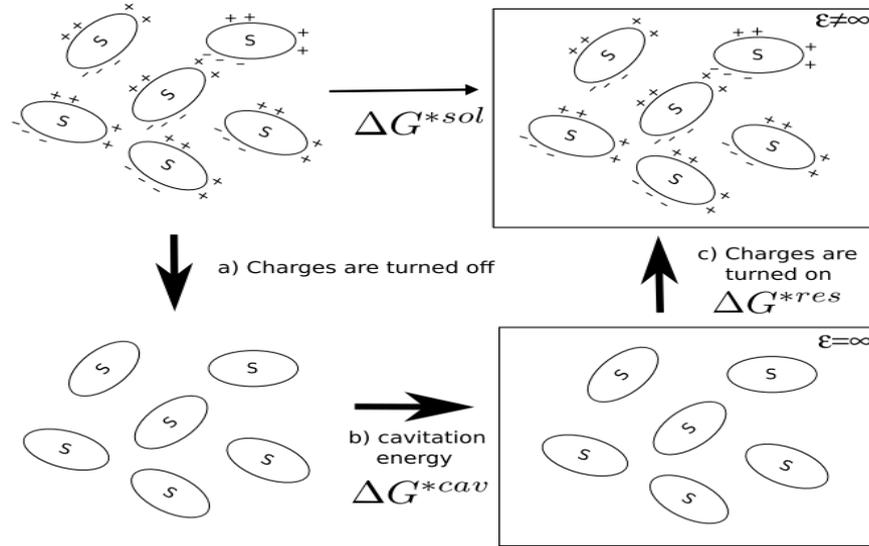


Figure 2.2 Principe de base du modèle COSMO-SAC : a) les charges en surface de la molécule sont “éteintes”, b) La molécule est insérée dans un conducteur parfait, c) Les charges de surface de la molécule sont restituées

En combinant ces deux étapes, le coefficient d'activité s'exprime comme suit :

$$\ln \gamma_{i/s} = \frac{\Delta G_{i/s}^{*res} - \Delta G_{i/i}^{*res}}{RT} + \ln \gamma_{i/s}^{SG} \quad (2.30)$$

$\gamma_{i/s}^{SG}$ est le terme entropique ou combinatoire de Staverman-Guggenheim. Il améliore, selon Lin et *al.* [90], les calculs de l'énergie libre de formation des cavités. Il est défini comme suit :

$$\ln \gamma_{i/s}^{SG} = \ln \left(\frac{\Phi_i}{x_i} \right) + \frac{Z}{2} q_i \ln \left(\frac{\theta_i}{\Phi_i} \right) + l_i - \frac{\Phi_i}{x_i} \sum_{j=1} x_j l_j \quad (2.31)$$

Φ_i , θ_i : Fractions volumique et surfacique de la molécule i , respectivement.

l_i, r_i, q_i , Φ_i et θ_i ont les mêmes expressions que dans d'autres modèles tels que UNIQUAC. Cependant, ici, la quantité q_i peut se déduire du σ -profil. En effet, la somme des surfaces de chaque segment dans le profil σ donne la surface totale q_i , autour de la molécule.

La contribution résiduelle s'écrit de la même manière que dans les méthodes de contribution de groupes mais non pas comme une somme de contributions de groupes, mais plutôt de surfaces de molécule. On attribue à chaque surface de la molécule un coefficient d'activité, ce qui se traduit par l'équation suivante :

$$\ln \gamma_s(\sigma_m) = n_i \sum_{\sigma_m} p_i(\sigma_m) [\ln \Gamma_s(\sigma_m) - \ln \Gamma_i(\sigma_m)] \quad (2.32)$$

Les valeurs utilisées pour chacun des paramètres ajustables du modèle COSMO-SAC, sont universelles et indépendantes des atomes / molécules modélisés et sont disponibles dans la littérature [88, 89].

Pour calculer l'énergie libre restituée, un profil- σ de chaque molécule est nécessaire. Ces profils, correspondant à la probabilité de trouver une surface de la molécule de densité de charge σ , sont calculés à partir de simulations chimiques quantiques. Les détails de la procédure à suivre pour obtenir les profils σ (détermination de la densité de charge surfacique et normalisation) à partir d'une molécule est disponible dans le commerce et le logiciel de simulation peut être trouvé dans la référence [91]. Une fois ces profils déterminés, les molécules sont divisées en segments. Ces segments conceptuels sont des parties de la surface de la molécule ayant la même densité de charge. L'équation du coefficient d'activité est alors écrite sous la forme :

$$\ln \gamma_{i/s} = n_i \sum_{\sigma_m} p_i(\sigma_m) [\ln \Gamma_s(\sigma_m) - \ln \Gamma_i(\sigma_m)] + \ln \gamma_{i/s}^{SG} \quad (2.33)$$

$\Gamma_s(\sigma_m)$ et $\Gamma_i(\sigma_m)$ sont les coefficients d'activité du segment σ_m en solution et dans le liquide pur, respectivement. Dans le modèle COSMO-SAC, le coefficient d'activité est obtenu en faisant la somme des contributions de chaque segment [35]. Une explication détaillée de la méthode COSMO-SAC est disponible dans le document original de Lin *et al.* [90].

2.3.2 Le modèle UNIFAC et ses différentes modifications

Le modèle UNIFAC (UNIversal Functional Activity Coefficient) est un modèle prédictif de coefficient d'activité des solutions liquides non électrolytes, basé sur le concept de contribution de groupe. Ce dernier, considère que les propriétés d'une molécule dérivent, de manière additive, de celles des groupes fonctionnels qui la composent. L'avantage de ce concept est qu'il permet de prédire les interactions entre les molécules, étant donné la connaissance des interactions des groupes fonctionnels (relativement peu nombreux) qui les constituent. La première formulation, que nous désignerons dans tout le reste de cette thèse comme le modèle UNIFAC original, a été publiée par Fredenslund *et al.* [92] en 1975. Depuis, de nombreuses versions du modèle ainsi que des extensions et révisions des paramètres de groupe UNIFAC ont été publiées [15, 16, 18, 19, 30–33, 93–99].

Gani *et al.* [33, 96, 97] ont rapporté une méthode de second ordre pour améliorer les prédictions pour les systèmes impliquant des isomères. D'autres auteurs comme Grensemann *et al.* [100] ont obtenu certains paramètres UNIFAC manquants avec les calculs COSMO-RS.

Les difficultés d'évaluation des paramètres UNIFAC sont principalement dues au manque de fiabilité de certaines données expérimentales disponibles dans la littérature. En effet, de nombreuses erreurs peuvent être commises lors de la collecte, l'analyse et la préparation des données pour leur publication [101]. Gmehling et *al.* [22] ont rapporté, que, par exemple, seulement 63 des 169 soit 37% de l'ensemble des données d'équilibre vapeur-liquide (VLE) pour les systèmes alcool-ester étaient considérés comme thermodynamiquement cohérents.

Les processus de criblage, tels que ceux utilisés par Gmehling et *al.* [22] impliquent l'application d'une variété de tests de cohérence pour l'ensemble des données VLE (Kojima et *al.* [102], Kurihara et *al.* [103], Jackson et *al.* [104], Wisniak [105], Van Ness et *al.* [106] et Herington [107]). Ces méthodes utilisent toutes des critères de réussite / échec fixes, et la décision de sélectionner ou de rejeter un ensemble de données spécifique dépend souvent du jugement personnel d'un expert, en particulier lorsque les résultats des différents tests entrent en conflit.

Récemment, une procédure plus générale et plus robuste a été développée au Centre de Recherche en Thermodynamique (TRC) de l'Institut National des Normes et de la Technologie (NIST) [108], dans lequel un seul facteur de qualité numérique QVLE est évalué sur la base d'une variété d'essais de cohérence VLE, ainsi que la prise en compte de la cohérence entre les données VLE et les pressions de vapeur des constituants purs. Les valeurs QVLE peuvent être utilisées comme facteurs de pondération pour l'analyse de régression. Cette approche élimine le besoin d'un filtrage manuel fastidieux des ensembles de données VLE. L'automatisation de ce processus présente d'autres avantages distincts : une traçabilité complète et une mise à jour simple, à mesure que de nouveaux ensembles de données deviennent disponibles et que la qualité des anciens ensembles de données soit réévaluée.

Dans toutes les évaluations précédentes des données VLE expérimentales [16, 18, 19, 30, 32, 33, 92–98, 100], les incertitudes sur les résultats expérimentaux ont été ignorées. De là, une nouvelle méthode pour l'évaluation des paramètres UNIFAC est décrite, où l'incertitude expérimentale est considérée en utilisant le facteur de qualité QVLE, décrit précédemment [101]. Les données des propriétés thermophysiques évaluées de manière critique utilisées ici (et leurs incertitudes élargies combinées [109]) sont celles évaluées avec le système expert NIST Thermo-Data-Engine (TDE) [108, 110–114]. Des évaluations préliminaires des incertitudes pour les données expérimentales ont été effectuées pendant le traitement avec le logiciel Guided Data Capture [115]. Les données expérimentales ont été

stockées dans le système d'archivage des données TRC source [116] et sont utilisées dans les évaluations critiques dynamiques effectuées par TDE, qui évaluent les estimations finales combinées des incertitudes élargies (niveau de confiance 0,95). Ces incertitudes sont basées sur l'approche de covariance, qui a été entièrement décrite dans les deux premiers articles de la série décrivant TDE [113,108]. Les incertitudes expérimentales et les écarts des courbes sont pris en compte dans le calcul des poids statistiques pour l'évaluation de la matrice de covariance. Les incertitudes évaluées sont utilisées dans l'estimation du maximum de confiance de la matrice des paramètres [108].

Deux séries de paramètres UNIFAC ont été produites; NIST-UNIFAC, où la dépendance de la température des paramètres est représentée par un seul terme, comme cela a été fait pour le modèle UNIFAC original [113], et NIST-KT-UNIFAC, où la dépendance de la température des paramètres est représentée par deux termes, comme cela a été fait pour KT-UNIFAC [33], développé par l'Université Technique du Danemark (DTU). Pour le NIST-UNIFAC et le NIST-KT-UNIFAC, la décomposition automatique des structures moléculaires a été effectuée sur la base de la liste complète des groupes et sous-groupes définis dans la référence [33].

2.3.2.1 Le modèle UNIFAC original

Le modèle UNIFAC a été développé en introduisant le concept de contribution de groupe dans le terme combinatoire du modèle UNIQUAC. Il utilise, donc, le même formalisme que ce dernier et décompose le coefficient d'activité en un terme combinatoire et un terme résiduel. Le terme combinatoire a la même forme que dans le modèle UNIQUAC (Equations 2.17, 2.18, 2.20 et 2.21). Cependant, dans ce cas, les paramètres moléculaires r_i et q_i de chaque constituant (i) sont calculés à partir de la somme des paramètres adimensionnels des groupes fonctionnels R_k et Q_k .

$$r_i = \sum_k v_k^{(i)} R_k \quad (2.34)$$

$$q_i = \sum_k v_k^{(i)} Q_k \quad (2.35)$$

$v_k^{(i)}$ est le nombre de sous-groupe k dans le composant i .

Les paramètres R_k et Q_k sont calculés depuis des volumes et des surfaces de groupe de van der Waals obtenus par Bondi [117].

La méthode optimale de décomposition est celle qui aboutit au nombre minimum de sous-groupes avec chaque sous-groupe ayant le maximum de répliquas.

Le terme résiduel dans le modèle UNIFAC est déterminé à partir de l'équation suivante :

$$\ln \gamma_i^R = \sum_k v_k^{(i)} \left[\ln \Gamma_k - \ln \Gamma_k^{(i)} \right] \quad (2.36)$$

Γ_k est le coefficient d'activité résiduel du sous-groupe k dans le mélange et $\Gamma_k^{(i)}$ est sa valeur dans une solution pure du composant i . Ce terme est ajouté, de sorte que, lorsque la fraction molaire approche l'unité, le terme $\ln \gamma_i^R$ tend vers zéro donc $\gamma_i^R \rightarrow 1$. Le coefficient d'activité résiduel du sous-groupe k dans une solution est donné par :

$$\ln \Gamma_k = Q_k \left[1 - \ln \left(\sum_m \theta_m \Psi_{mk} \right) - \sum_m \frac{\theta_m \Psi_{mk}}{\sum_n \theta_n \Psi_{nm}} \right] \quad (2.37)$$

θ_m : Fraction de surface du groupe fonctionnel m dans le mélange :

$$\theta_m = \frac{Q_m X_m}{\sum_n Q_n X_n} \quad (2.38)$$

X_m : Fraction molaire du sous-groupe m dans le mélange

$$X_m = \frac{\sum_j v_m^{(j)} x_j}{\sum_j \sum_k v_k^{(j)} x_j} \quad (2.39)$$

Ψ_{nm} : Paramètre d'interaction de groupe entre les groupes n et m et dépend de la température :

$$\Psi_{nm} = \exp \left(-\frac{\Delta u_{nm}}{RT} \right) = \exp \left(-\frac{a_{nm}}{T} \right) \quad (2.40)$$

Δu_{nm} est le paramètre d'interaction entre les groupes n et m .

Le paramètre d'interaction du groupe a_{nm} est déterminé à partir d'un large ensemble de données VLE et LLE disponibles dans la littérature et tabulés pour de nombreux sous-groupes.

Il est à noter que : $a_{mn} \neq a_{nm}$

2.3.2.2 Le modèle UNIFAC modifié Lyngby

Les modifications apportées au modèle UNIFAC original et réalisées par Larsen et *al* [23], incluent des modifications de la partie combinatoire et une expression pour les paramètres d'interaction a_{nm} en fonction de la température.

L'expression de la partie combinatoire devient alors :

$$\ln(\gamma_i^C) = \ln \left(\frac{\omega_i}{x_i} \right) + 1 - \left(\frac{\omega_i}{x_i} \right) \quad (2.41)$$

Avec :

$$\omega_i = \frac{x_i r_i^{\rho_i}}{\sum_j (x_j r_j^{\rho_i})} \quad (2.42)$$

ρ_i est relatif au paramètre de volume r_i de la molécule et varie selon le type des molécules étudiées. Kikic et *al.* [118, 119] ont fixé sa valeur à 2/3 :

$$\omega_i = \frac{x_i r_i^{2/3}}{\sum_j (x_j r_j^{2/3})} \quad (2.43)$$

L'introduction de l'exposant 2/3 permet de mieux décrire les équilibres Liquide-Vapeur des mélanges d'alcanes, spécialement lorsque la différence de taille entre les composés est importante [120].

La partie résiduelle suit la même démarche de calcul que le modèle original sauf qu'ici les auteurs utilisent trois paramètres pour décrire l'effet de la température sur les paramètres d'interaction, d'où :

$$A_{nm} = a_{nm} + b_{nm}(T - T_0) + c_{nm} \left(T \ln \left(\frac{T_0}{T} \right) + T - T_0 \right) \quad (2.44)$$

$$\Psi_{nm} = \exp \left(-\frac{A_{nm}}{T} \right) \quad (2.45)$$

$$\Psi_{nm} = \exp \left(-\frac{a_{nm} + b_{nm}(T - T_0) + c_{nm} \left(T \ln \left(\frac{T_0}{T} \right) + T - T_0 \right)}{T} \right) \quad (2.46)$$

La température de référence étant, $T_0 = 298,15\text{K}$, et a_{nm} , b_{nm} et c_{nm} ; les paramètres d'interaction.

2.3.2.3 Le modèle UNIFAC modifié (Dortmund)

Dans le modèle UNIFAC modifié (Dortmund) [22], comme dans le modèle UNIFAC original, le coefficient d'activité est la somme d'une partie combinatoire et d'une partie résiduelle.

La partie combinatoire a été modifiée de manière empirique, pour permettre de traiter des composés de tailles très différentes, et est donnée par :

$$\ln \gamma_i^c = 1 - V_i' + \ln V_i' - 5q_i \left(1 - \frac{V_i}{F_i} + \ln \left(\frac{V_i}{F_i} \right) \right) \quad (2.47)$$

Le paramètre V_i' peut être calculé en utilisant les volumes relatifs de van der Waals R_k des différents groupes

$$V_i' = \frac{r_i^{3/4}}{\sum_j(x_j r_j^{3/4})} \quad (2.48)$$

Tous les autres paramètres sont calculés de la même manière que pour le modèle UNIFAC original ; c'est-à-dire :

$$V_i = \frac{r_i}{\sum_j(x_j r_j)} \quad (2.49)$$

$$r_i = \sum_k(v_k^{(i)} R_k) \quad (2.50)$$

$$F_i = \frac{q_i}{\sum_j(x_j q_j)} \quad (2.51)$$

$$q_i = \sum_k(v_k^{(i)} Q_k) \quad (2.52)$$

La partie résiduelle peut être obtenue en utilisant les relations du modèle original (2.36), (2.37), (2.38), et (2.39).

Par rapport au modèle UNIFAC original, seules les propriétés de van der Waals relatives aux paramètres de volume et de surface ont été légèrement modifiées puisqu'elles ont été régressées en même temps que des paramètres d'interaction. Ces derniers sont considérés comme dépendants de la température pour permettre une meilleure description du comportement réel (coefficients d'activité) en fonction de la température [95].

$$\Psi_{nm} = \exp\left(-\frac{a_{nm} + (b_{nm} \times T) + (c_{nm} \times T^2)}{T}\right) \quad (2.53)$$

2.3.2.4 Le modèle UNIFAC Pharma Modifié

Le modèle UNIFAC Pharma Modifié [34] est une autre version, plus récente, du modèle UNIFAC Modifié (Dortmund). Il a été établi pour surmonter certaines limites observées lorsque ce dernier est appliqué à des solutions contenant des ingrédients actifs pharmaceutiques. Les équations de cette nouvelle version sont les mêmes que dans le modèle original, sauf en ce qui concerne l'ensemble des groupes fonctionnels disponibles, leurs paramètres de volume et de surface et leurs paramètres d'interaction.

Dans le modèle UNIFAC original, les paramètres de volume et de surface ont été calculés par la méthode de Bondi [117], tandis que les termes d'interaction ont été obtenus à partir de la régression des données expérimentales. Dans UNIFAC Pharma Modifié, les paramètres de volume et de surface ont été ajustés simultanément avec les paramètres d'interaction.

La fonction objective utilisée dans l'estimation est la suivante :

$$f(a_{nm}, a_{mn}, b_{nm}, b_{mn}, c_{nm}, c_{mn}, R_k, Q_k) = \text{Minimum} \quad (2.54)$$

Les paramètres d'interaction binaire ont trois termes a_{nm} , b_{nm} et c_{nm} . Pour réduire le nombre de paramètres, Diedrichs et *al.* [34] suggèrent de négliger certains paramètres d'interaction si leur influence sur le résultat est faible. C'est le cas des groupes fonctionnels présents dans les API's, qui interagissent à faibles concentrations. Ainsi, les paramètres d'interaction dont nous devons tenir compte, sont ceux existant entre les groupes fonctionnels de la phase solvant et ceux des principes actifs. Le terme c_{nm} est considéré comme nul.

La formule, donnant le paramètre d'interaction Ψ_{nm} de cette nouvelle version, est donnée comme suit :

$$\Psi_{nm} = \exp\left(-\frac{a_{mn} + (b_{mn} \times T)}{T}\right) \quad (2.55)$$

UNIFAC Pharma modifié n'est valable que pour une fraction molaire maximale d'API dans le mélange autour de 0,1. Cette condition devrait être suffisante pour le traitement des processus de cristallisation en solution des API's [34, 121].

2.3.2.5 Le modèle KT-UNIFAC

Dans le modèle KT-UNIFAC [33], la structure moléculaire d'un composé est considérée comme un ensemble de groupes du premier ordre et du second ordre. Les groupes du premier ordre sont capables de représenter une grande variété de composés organiques. Lorsque ces groupes deviennent insuffisants pour fournir une représentation appropriée dans le cas des molécules complexes, alors les groupes de second ordre sont impliqués pour fournir des informations plus structurales sur les fragments moléculaires des composés qui ne sont pas convenablement représentés par les groupes de premier ordre.

Il faut noter que les groupes structurels de KT-UNIFAC sont différents de ceux du modèle UNIFAC original. Selon le modèle KT-UNIFAC, les coefficients d'activité sont calculés par l'addition de trois contributions, un terme combinatoire, un terme résiduel et un terme résiduel de second ordre. Le premier prend en compte la taille et la forme moléculaire, le second prend en compte les interactions moléculaires et le troisième tient compte des effets de second ordre sur les interactions moléculaires.

$$\ln\gamma_i = \ln\gamma_i^C + \ln\gamma_i^R + w_{R2}\ln\gamma_i^{R2} \quad (2.56)$$

Le terme w_{R2} est égal à zéro pour le modèle de premier ordre et égale à 1 pour le modèle de second ordre.

La partie combinatoire est la même que dans le modèle UNIFAC original, tandis que dans la partie résiduelle, les paramètres d'interaction sont exprimés linéairement en fonction de la température.

$$\Delta u_{nm} = a_{nm,1} + a_{nm,2}(T - T_0) \quad (2.57)$$

La formule donnant Ψ_{nm} devient :

$$\Psi_{nm} = \exp\left(-\frac{a_{mk,1} + a_{mk,2}(T - T_0)}{T}\right) \quad (2.58)$$

L'équation de la partie résiduelle de second ordre et d'autres détails peuvent être trouvés dans la littérature [33].

Les trois versions du modèle UNIFAC, lorsqu'elles sont appliquées aux équilibres Liquide-Solide, nécessitent la connaissance de l'enthalpie de fusion, la température de fusion et les structures moléculaires.

2.4 Application des modèles de coefficient d'activité aux équilibres Liquide-Solide

Les modèles prédictifs, qui facilitent la détermination du solvant le plus approprié à partir d'une liste de solvants sélectionnée, nécessitent une approche modèle universelle, fournie par la thermodynamique des équilibres de phase. L'équation thermodynamique simplifiée de la solubilité, pour les systèmes eutectiques, requiert les données du point de fusion (enthalpie de fusion et température de fusion) et le coefficient d'activité en phase liquide des composés considérés. Le coefficient d'activité dans la phase liquide est une quantité auxiliaire permettant de décrire les interactions non idéales entre le soluté et le solvant.

Différents modèles de coefficients d'activité de la littérature ont été mentionnés dans le paragraphe précédent. De nombreuses recherches et améliorations ont été apportées à ces modèles afin de les rendre plus précis, ainsi que pour leur classification et la clarification de leur domaine d'application. Dans cette section, nous présenterons une analyse des modèles élaborés les plus importants ainsi que, les modifications qui y ont été apportées, la portée de leur mise en œuvre, leurs avantages et leurs limites.

Depuis le développement du premier modèle UNIFAC original [92] en 1975, le modèle UNIFAC a été étendu et décrit en détail dans une monographie. Un tableau détaillé des paramètres d'interaction des groupes a été publié dans le cadre d'un effort continu de « révision et extension ». Cela a donné lieu à une série de cinq articles [122-125] commençant à la fin des années 1970. Déjà, au tout début du développement du modèle UNIFAC original, un certain nombre de problèmes avaient été rencontrés. Parmi ces derniers, il y avait la dépendance à la température du modèle UNIFAC original. En outre, l'examen du terme combinatoire basé sur les données d'un mélange d'hydrocarbures a montré que des modifications étaient nécessaires pour améliorer les calculs de solubilité [124].

Les raisons ci-dessus ont conduit au développement des deux premières versions corrigées du modèle : UNIFAC Modifié (Lyngby et Dortmund) [126,127]. Les deux versions utilisent des équations assez similaires. Cependant, la version de Dortmund considère les paramètres de groupe comme paramètres ajustables, au même titre que les paramètres d'interaction et leur valeur ont été déterminées par ajustement des données d'équilibre de phase. C'est probablement la différence la plus importante entre les deux versions modifiées UNIFAC. La régression des paramètres pour UNIFAC Modifié "Dortmund" s'est poursuivie et a donné naissance de nombreuses mises à jour [128,129]. En revanche, une seule publication de la version "Lyngby" a été publiée. Ces dernières années, de nombreux chercheurs [129,130] ont publié des paramètres de groupes pour ce modèle.

Certaines études ont été réalisées en appliquant le modèle UNIFAC pour le calcul de la solubilité des produits pharmaceutiques dans certains solvants. Hahnenkamp et *al.* [131] ont déterminé des données d'équilibre Solide-Liquide pour le paracétamol, l'aspirine et l'ibuprofène dans plusieurs solvants. Ils ont observé que les modèles UNIFAC original UNIFAC modifié (Dortmund) ont tendance à sous-estimer systématiquement la solubilité. Le modèle UNIFAC modifié de Dortmund donne cependant de meilleurs résultats comparés à ceux d'UNIFAC original par opposition aux données expérimentales. Bouillot et *al.* [132] se sont également intéressés à la solubilité des principes actifs. Ils ont testé plusieurs modèles thermodynamiques pour la prédiction de cinq molécules pharmaceutiques tels que le paracétamol, l'ibuprofène, l'acide salicylique, l'acide 4-aminobenzoïque et l'acide benzoïque. Ils ont conclu que les modèles UNIFAC (UNIFAC et UNIFAC modifié (Dortmund)) sont généralement en accord avec les données expérimentales mais ne peuvent pas donner

suffisamment de prédictions quantitatives. En plus, ils n'ont pas réussi à décrire correctement la dépendance à la solubilité avec la température.

L'extension de la méthode de contribution des groupes UNIFAC modifiée "Dortmund" à la solubilité des substances actives a conduit à l'élaboration du modèle UNIFAC Pharma modifié [133]. La règle d'incrémentation des groupes structuraux des molécules a été révisée et le nombre des paramètres d'interaction de groupes nécessaires, a été réduit, en négligeant certaines combinaisons de paramètres d'interaction [34]. Par conséquent le champ d'application d'UNIFAC Pharma modifié est limité à une fraction molaire maximale de principe actif dans le mélange d'environ 0,1%, mais cela devrait généralement être suffisant pour le traitement des processus de cristallisation en solution des principes actifs

La large imprécision sur les résultats observés sur les mélanges d'isomères ou avec effets de proximité, a été étudiée de manière approfondie [134]. Toutes les méthodes UNIFAC existantes sont considérées comme des méthodes de premier ordre et une extension a été entreprise en utilisant l'effet des interactions de second ordre sur l'enthalpie libre d'excès. Cela a conduit à la mise au point récente d'une table de groupe combinée premier plus deuxième ordre, nommée KT-UNIFAC [33,135]. Il a été démontré que les données sur les alcools, les hydrocarbures et les aromatiques peuvent être corrélées et prédites de manière satisfaisante pour tous les isomères.

Les paramètres d'interaction de groupes du modèle UNIFAC sont indispensables à la prédiction de la solubilité, or près de 50% des paramètres sont manquants dans la matrice du modèle ; ceci en fait son principal inconvénient. Des révisions des paramètres UNIFAC ont été entreprises à plusieurs reprises, dans le passé, mais il manque toujours des paramètres d'interaction dans la matrice, en raison du manque de données expérimentales. Cela limite l'utilisation du modèle UNIFAC et son extension pour prédire les équilibres de phase pour une gamme plus large de systèmes chimiques.

Les données d'équilibre de phase Solide-Liquide ont été analysées à l'aide des modèles thermodynamiques bien connus (NRTL [136], UNIQUAC et UNIFAC original) pour différents systèmes de mélange de lipides [137–142], les performances des modèles de prédiction des données des équilibres solide-liquide sont comparées. Des écarts plus importants sont observés pour le modèle UNIFAC original, NRTL et UNIQUAC [143].

Dans le but de prédire la solubilité des médicaments Diedrichs et *al.* [34] ont évalué plusieurs modèles de coefficients d'activité (Hansen, UNIFAC original, UNIFAC modifié "Dortmund", COSMO-RS, NRTL-SAC, UNIFAC Pharma Modifié). Il ressort de cette étude que les modèles de coefficients d'activité peuvent être divisés en deux catégories :

- La première catégorie comprend les modèles corrélatifs qui reposent sur des données expérimentales de la solubilité pour déterminer les paramètres spécifiques du modèle. Ensuite, ils peuvent être utilisés pour les calculs prédictifs de la solubilité. Cette catégorie comprend les deux modèles Hansen et NRTL-SAC. Or, NRTL-SAC est nettement plus efficace que le modèle Hansen, vu que le modèle Hansen ne devrait plus être utilisé comme modèle de solubilité aujourd'hui. En mode corrélatif, NRTL-SAC peut être utilisé comme test de cohérence pour détecter des données de solubilité expérimentale de mauvaise qualité.
- La deuxième catégorie comprend les modèles qui ne nécessitent pas des données expérimentales de la solubilité du médicament ou de la molécule complexe, ce sont les purs modèles prédictifs. C'est le cas des méthodes de contribution des groupes UNIFAC, modèle UNIFAC Modifié "Dortmund" et UNIFAC Pharma Modifié, et le modèle de chimie quantique COSMO-RS. Parmi ces modèles, le modèle UNIFAC modifié "Dortmund" le meilleur qui prédit les solubilités, en général suivi par UNIFAC Pharma modifié, UNIFAC et COSMO-RS en dernier.

L'hypothèse d'un comportement idéal provoque les écarts les plus importants par rapport aux données expérimentales. Il convient de mentionner que le modèle UNIFAC Pharma modifié ne donne pas (de manière significative) de meilleurs résultats que les modèles UNIFAC original et UNIFAC modifié "Dortmund". Après tout, les paramètres du modèle UNIFAC Pharma Modifié ont été spécialement conçus pour la solubilité des médicaments dans une plage de concentration restreinte et pour les coefficients d'activité à dilution infinie. Les paramètres pour le modèle UNIFAC et UNIFAC Modifié "Dortmund" sont toutefois valables pour toute la plage de concentration. En outre, ces deux modèles n'ont même pas été développés spécifiquement pour la description des équilibres de phase solide-liquide, mais principalement sur la description des équilibres de phase Liquide-Vapeur. Néanmoins, le fait que la qualité prédictive du modèle UNIFAC Pharma Modifié ne soit pas (clairement) supérieure à celle de UNIFAC original ou encore à celle d'UNIFAC Modifié "Dortmund" doit être due au fait que la négligence de certaines combinaisons de groupes structuraux dans ce

modèle peut conduire à des conséquences graves et des écarts importants. Cela est particulièrement évident pour les molécules qui contiennent de nombreux groupes fonctionnels avec des hétéroatomes tels que les médicaments et les molécules complexes. Dans ces cas, UNIFAC Pharma modifié échoue plus rapidement que UNIFAC original et UNIFAC Modifié "Dortmund", de sorte que, UNIFAC Pharma Modifié fonctionne globalement au même niveau de prédiction que les autres méthodes de contribution de groupe. En outre, l'adaptation des paramètres pour UNIFAC Pharma Modifié a révélé que les coefficients d'activité des molécules, capables de former des liaisons hydrogène intramoléculaires, sont souvent sous-estimés. Le principe d'additivité des groupes n'est apparemment plus valable pour ces molécules car elles sont actuellement décomposées en leurs groupes structurels vu les effets de groupe voisins qui ne peuvent pas être détectés. Pour une meilleure description, des groupes structuraux spécifiques doivent être implémentés dans la règle d'incrémentation. Cette dernière, présente l'avantage, par rapport aux deux autres méthodes de contribution du groupe UNIFAC original et UNIFAC modifié "Dortmund", disant qu'un plus grand nombre de substances actives ou de molécules complexes sont incorporés dans leurs groupes structurels peut être démonté, non seulement la classification modifiée des groupes des structures, mais également le principe de réduction des paramètres utilisé par UNIFAC Pharma Modifié qui contribuent à un nombre plus élevé des systèmes informatiques. Pour le moment, seules les solubilités dans les alcanes, les alcools et l'eau peuvent être calculées. Pour une application encore plus large de Pharma UNIFAC Modifié, les paramètres d'interaction avec les groupes cétone, ester et éther sont nécessaires. Afin d'enrichir la matrice des paramètres d'interaction du modèle UNIFAC Pharma Modifié, nous avons estimé dans ce travail les paramètres d'interaction pour plus de 20 groupes fonctionnels.

La portée des modèles de coefficient d'activité est très variable. Avec NRTL-SAC basé sur la corrélation des données expérimentales de la solubilité, nous devons utiliser les données expérimentales disponibles pour estimer les paramètres du modèle de soluté, puis le modèle peut être appliqué pour prédire la solubilité de soluté pour d'autres systèmes de solvant. Pour le modèle NRTL-SAC, les paramètres spécifiques du modèle ne sont publiés que pour quelques substances actives ou molécules complexes. Par conséquent, peu de systèmes peuvent être relativement calculés avec ce modèle. Nous nous sommes appuyés sur les recherches effectuées sur ce modèle [11, 84, 144–159], pour conclure ce qui suit :

L'avantage de la méthode NRTL-SAC est de donner des prédictions quantitatives assez bonnes en raison de sa nature semi-empirique. Le modèle rend bien les ordres de grandeur, bien qu'ils soient un peu sous-estimés, et les échelles de solubilité assez fidèles. Toutefois, la méthode est très sous-tendue de ses paramètres d'entrée, y compris des quadruplets qui représentent les molécules. Cette dépendance pose deux principaux problèmes :

- La difficulté à établir ou de choisir les quadruplets,
- La pertinence de la représentation des molécules de la dite-manière.

Pour une estimation rapide de la solubilité, Le modèle NRTL-SAC semble être applicable, mais risque de ne pas être suffisamment fiable au-delà, que ce soit pour une évaluation de la sursaturation ou des résultats quantitatifs.

Par conséquent, si des estimations de la solubilité d'un médicament sont nécessaires, il est suggéré d'appliquer le modèle NRTL-SAC pour les systèmes dans lesquels des données dans un ensemble de solvants sélectionnés sont facilement disponibles. Pour les solutés où les données de solubilité expérimentale ne sont pas disponibles ou difficiles à obtenir et si tous les paramètres de groupe sont connus, l'application de la méthode UNIFAC est une option valable.

Le modèle COSMO-SAC [160–164] présente des avantages incontestables dans son fonctionnement même. Le modèle prend en considération la géométrie des molécules, les isomères, et est, totalement prédictif, excepté quelques paramètres. Toutefois, les résultats obtenus ne sont pas suffisamment précis afin de rendre le modèle pratiquement fiable pour le dimensionnement des procédés de cristallisation, ou même pour effectuer un “screening” des solvants, même si la représentation des molécules dans cette méthode semble comme une carte maitresse. Son utilisation est limitée aux composés les plus simples, même pour évaluer des équilibres d'une manière rapide . D'une façon générale et parfois de manière significative (environ deux fois plus), les solubilités prédites sont très surestimées, Néanmoins, la dépendance du modèle en température reste appropriée pour évaluer les sursaturations, et le modèle permet de mettre le doigt sur les solvants dans lesquels les solutés sont quasiment insolubles (échec de donner un ordre de grandeur valide dans ces cas-là) [50].

2.5 Conclusion

En conclusion, à l'heure actuelle, aucun des modèles présentés dans cette section ne permet de prédire avec précision la solubilité des ingrédients actifs pharmaceutiques (contenant divers groupes fonctionnels communs tels les alcools, les cétones, les amines...) dans divers solvants. Même en considérant les erreurs du modèle, la plupart du temps prises par l'imprécision sur les propriétés du solide pur, la solubilité prédite est généralement en accord avec les données expérimentales (ibuprofène avec Original UNIFAC, certains solutés avec NRTL-SAC). Les prédictions COSMO-SAC produisent généralement des valeurs plus élevées que les expérimentales, en particulier lorsque des liaisons hydrogènes peuvent se produire. Dans les différentes études mentionnées dans le paragraphe précédent, UNIFAC est plus précis que COSMO-SAC pour les APIs et donne de bons résultats pour des molécules simples telles que l'anthracène. Toutefois, même si les ordres de grandeur de solubilité sont conservés, ils ne suffisent pas pour effectuer des prévisions quantitatives. Dans le cas de NRTL-SAC, les prévisions dépendent beaucoup des paramètres utilisés (valeurs de segment). Si ces paramètres ne représentent pas le comportement correct de l'API dans le solvant choisi, les résultats ne seront pas satisfaisants. Enfin, UNIFAC et NRTL-SAC qui conservent les ordres de grandeur de solubilité peuvent constituer un bon complément des techniques expérimentales permettant de guider le choix d'un solvant de cristallisation. Mais au final, des expériences sont encore nécessaires pour obtenir des résultats quantitatifs.

Les résultats obtenus, dans les travaux mentionnés, montrent également que l'utilisation de l'équation d'équilibre la plus précise n'influence pas de manière significative la qualité des résultats. Comme les modèles ne sont pas assez précis, l'utilisation des valeurs des propriétés thermodynamiques les plus précises n'est pas nécessaire. Aussi, le terme ΔC_p peut être ignoré tant que l'erreur générée est négligeable par rapport à l'erreur du modèle. Mais, dans le cas hypothétique d'un modèle parfait, les valeurs les plus précises devraient être utilisées.

Estimation des paramètres d'interaction du modèle UNIFAC Pharma Modifié

Sommaire :

3.1	Introduction	42
3.2	Subdivision des molécules.....	42
3.3	Paramètres de volume et de surface.....	45
3.4	Estimation des paramètres d'interaction	51
3.5	Résultats et discussions	56
3.6	Conclusion.....	62

3.1 Introduction

Une étude approfondie des modèles de coefficient d'activité et leur utilisation dans le calcul de la solubilité et des équilibres Liquide-Solide a été réalisée dans le chapitre précédent. Etant donné que notre étude s'intéresse particulièrement aux composés pharmaceutiques, nous avons décidé de compléter le travail commencé par Diedrichs et *al.*, en déterminant quelques-uns des paramètres d'interaction manquants dans la matrice du modèle UNIFAC Pharma.

Dans ce but, plus de 100 systèmes binaires soluté/solvant, comportant 22 groupes fonctionnels, ont été utilisés pour estimer plus de 200 paramètres d'interaction. La méthode d'optimisation du Simplex de Nelder-Mead a permis de générer ces paramètres.

La solubilité de certains systèmes binaires a été calculée à l'aide des deux modèles de coefficients d'activité UNIFAC original et UNIFAC Pharma Modifié, puis les résultats obtenus ont été comparés.

3.2 Subdivision des molécules

Le modèle UNIFAC Pharma Modifié a mis au point une nouvelle subdivision des groupes fonctionnels, assez différente de toutes les versions modifiées précédentes du modèle classique [34].

On constate ainsi que de nouveaux groupes fonctionnels ont été ajoutés et classés en fonction du composé qui les contient. Par exemple pour le groupe CO, le modèle a considéré que le groupe fonctionnel dans l'alcane est différent du groupe fonctionnel dans la cétone. En conséquence, les deux groupes possèdent les mêmes paramètres de surface et de volume mais diffèrent dans les paramètres d'interaction.

On remarque également la présence de nouveaux groupes fonctionnels (avec une masse molaire importante) tels que CON-CONCO et N-CO-NH₂ dont les paramètres varient également en fonction du composé qui les contient. C'est le cas du groupe CO.

Pour réaliser cette partie du travail, nous avons sélectionné les groupes fonctionnels pour lesquels nous disposons des données de solubilité d'une part et des propriétés de fusion des corps purs, d'autre part. Ces groupes ainsi que leur numéro respectif dans la matrice sont regroupés dans le tableau 3.1.

Tableau 3.1 : Groupes fonctionnels du modèle UNIFAC Pharma Modifié, avec leur numéro respectif dans la matrice, pour l'estimation des nouveaux paramètres

Numéro dans la matrice	Main-group
1	CH ₂
2	AC
5 (Alcool)	CH ₂ -OH (Alcool)
6	AC-OH
6 (Ortho)	AC-OH (Ortho)
6 (Meta)	AC-OH (Meta)
10	AC-NH ₂
13 (cétone)	CO (cétone)
14	AC-CO
17	COOH
18	AC-COOH
19 (Ester)	COO- (Ester)
24	C-Cl
25	AC-Cl
33	C-Cl ₃
34 (alcane)	C-O (alcane)
34 (Ether)	C-O (Ether)
35	AC-O
44	CON
46	C≡N
48	AN
67	>S=O

Notre apport ici, concerne principalement, les paramètres d'interaction des groupes ACOH qui ont été considérés comme différents selon leur position par rapport à un autre groupe dans le cycle aromatique. Si, le composé aromatique contient la liaison AC-OH en position opposée à ce dernier, il se trouve dans la position Para. Dans ce cas, nous sommes dans le même cas que le groupe AC-OH estimé par Diedrichs et *al.* (*Exemples :* acide benzoïque (Figure 3.1) ou le 4-Hydroxybenzoïque (Figure 3.2)).

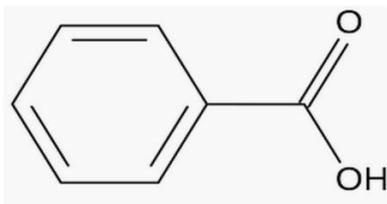


Figure 3.1 : structure chimique de l'acide benzoïque

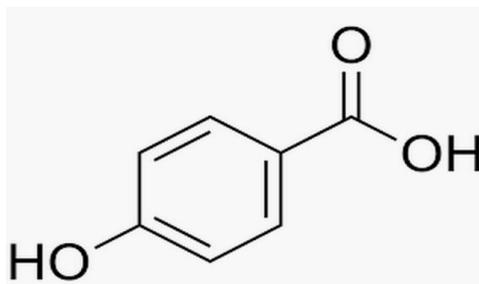


Figure 3.2 : structure chimique de l'acide 4-Hydroxybenzoïque

Si le composé aromatique contenant le groupement AC-OH est associé à un autre groupe placé en position meta, comme avec le 3-Hydroxybenzoïque (Figure 3.3), l'interaction entre ces deux groupements est différente de celle entre les mêmes groupements placés en position para.

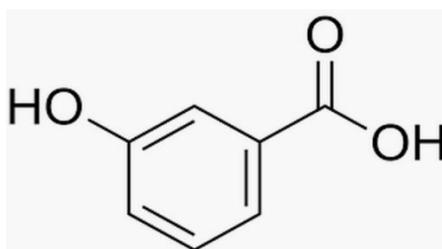


Figure 3.3 : structure chimique de l'acide (méta) 3-Hydroxybenzoïque

Le troisième cas qui peut se présenter, est celui où les mêmes groupements sont en position ortho, par exemple l'acide (ortho) 2-Hydroxybenzoïque (figure 3.4)

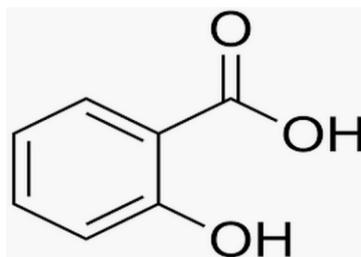


Figure 3.4 structure chimique de l'acide 2-Hydroxybenzoïque

Les paramètres d'interaction concernant ces deux derniers cas, seront certainement différents de ceux déterminés par Diedrichs considérés en position para.

3.3 Paramètres de volume et de surface

La différence essentielle entre le modèle UNIFAC Pharma Modifié et le modèle UNIFAC original concerne les paramètres de surface et de volume. Dans le premier, ces paramètres ne sont pas calculés par la méthode de Bondi [117, 165] mais régressés en même temps que les paramètres d'interaction par la méthode d'optimisation.

Le tableau 3.2 montre les paramètres de volume et de surface obtenus par le modèle UNIFAC Pharma Modifié [34].

Tableau 3.2 : Paramètres de volume et de surface du modèle UNIFAC Pharma

N°	Main-group	N°	Sub-group	R _K	Q _K
1	CH ₂	1	CH ₃	0,9000	0,9223
		2	CH ₂	0,7418	0,5940
		3	CH	0,4332	0,2393
		4	C	0,2403	0,0000
2	AC	5	ACH	0,4782	0,3600
		6	AC	0,3719	0,1099
3	C=C	7	CH ₂ =CH	1,2648	1,0584
		8	CH ₂ =C	1,1909	1,0205
		9	CH=CH	1,0207	0,9503
		10	CH=C	0,8978	0,6996
		11	C=C	0,6605	0,4080
4	C≡C	12	CH≡C	1,2920	1,0880
		13	C≡C	1,0613	0,7840
5	CH ₂ -OH	14	CH ₂ -OH	1,3248	1,0126
		15	CH ₂ -SH	1,6506	1,3680
		16	CH-OH	0,9488	0,8932

		17	CH-SH	1,4232	1,0560
		18	C-OH	0,6747	0,6424
		19	C-SH	1,1958	0,8280
		20	(=C)-OH	0,5300	0,5840
6	AC-OH	21	AC-OH	0,9847	0,7744
		22	AC-SH	1,3415	0,9480
7	C-NH ₂	23	CH ₃ -NH ₂	1,5959	1,5440
		24	CH ₂ -NH ₂	1,3691	1,2360
		25	CH-NH ₂	1,1417	0,9240
		26	C-NH ₂	0,9143	0,6960
		27	(=C)-NH ₂	0,6948	0,6960
8	C-NH	28	CH ₃ -NH	1,4338	1,2440
		29	CH ₂ -NH	1,2070	0,9360
		30	CH-NH	0,9796	0,6240
		31	C-NH	0,7521	0,3960
		32	(=C)-NH	0,5326	0,3960
9	C-N	33	CH ₃ -N	1,1866	0,9400
		34	CH ₂ -N	0,9598	0,6320
		35	CH-N	0,7324	0,3200
		36	C-N	0,5049	0,0920
		37	(=C)-N	0,2854	0,0920
10	AC-NH ₂	38	AC-NH ₂	0,9542	0,8976
11	AC-NH	39	AC-NH	0,8978	0,5160
12	AC-N	40	AC-N	0,6506	0,2120
13	CO	41	CO	0,6942	0,7040
		42	CS	1,1958	0,8280
14	AC-CO	43	AC-CO	1,0229	0,8360

		44	AC-CS	1,5610	0,9480
15	CHO	45	CHO		
16	AC-CHO	46	AC-CHO	1,3632	1,0680
17	COOH	47	COOH	1,4314	1,2140
18	AC-COOH	48	AC-COOH	1,6664	1,3440
19	COO	49	COO	0,9018	0,9680
		50	COS	1,4832	1,1600
20	HCOO	51	HCOO	1,2294	1,1880
21	AC-COO	52	AC-COO	1,4288	0,9000
		53	AC-COS	1,8484	1,2800
22	C-F	54	CH ₃ -F	1,2966	1,3080
		55	CH ₂ -F	1,0699	1,0000
		56	CH-F	0,8425	0,6880
		57	C-F	0,6150	0,4600
23	AC-F	58	AC-F	0,7475	0,5600
24	C-Cl	59	CH ₃ -Cl	1,7080	1,5760
		60	CH ₂ -Cl	1,4812	1,2680
		61	CH-Cl	1,2538	0,9560
		62	C-Cl	1,0264	0,7280
		63	Cl-(C=)	0,8069	0,7280
25	AC-Cl	64	AC-Cl	1,1562	0,8440
26	C-Br	65	CH ₃ -Br	1,8635	1,6840
		66	CH ₂ -Br	1,6368	1,3760
		67	CH-Br	1,4094	1,0640
		68	C-Br	1,1819	0,8360
27	AC-Br	69	AC-Br	1,3619	0,9720
28	C-I	70	CH ₃ -I		

		71	CH ₂ -I		
		72	CH-I		
		73	C-I		
29	AC-I	74	AC-I	1,6599	1,1240
30	C-F ₂	75	CH ₂ -F ₂		
		76	CH-F ₂		
		77	C-F ₂		
31	C-Cl ₂	78	CH ₂ -Cl ₂	2,2881	1,9960
		79	CH-Cl ₂	2,0606	1,6840
		80	C-Cl ₂	1,8332	1,4560
32	C-F ₃	81	CH-F ₃	1,6335	1,6080
		82	C-F ₃	1,4061	1,3800
33	C-Cl ₃	83	CH-Cl ₃	2,8675	2,4120
		84	C-Cl ₃	2,6401	2,1840
34	C-O	85	CH ₃ -O	1,1794	0,9793
		86	CH ₂ -O	1,0090	0,7395
		87	CH-O	0,6692	0,5148
		88	C-O	0,4634	0,2400
		89	CH ₃ -S	1,6999	1,3260
		90	CH ₂ -S	1,3594	0,9540
		91	CH-S	1,1589	0,7480
		92	C-S	0,9314	0,5200
35	AC-O	93	AC-O	0,6091	0,3600
		94	AC-S	1,0771	0,6400
36	C=N	95	CH=N		
		96	C=N		
37	ÁC-C=N	97	AC-CH=N	1,3448	0,7440

		98	ÁC-C=N	1,1173	0,5160
38	NO ₂	99	NO ₂	1,1074	1,0200
39	AC-NO ₂	100	AC-NO ₂	1,4726	1,1400
40	CONH ₂	101	CONH ₂	1,4661	1,3360
41	AC-CONH ₂	102	AC-CONH ₂	1,8312	1,4560
42	CONH	103	CONH	1,3039	1,0360
43	AC-CONH	104	AC-CONH		
44	CON	105	CON	1,0567	0,7320
45	AC-CON	106	AC-CON		
46	C≡N	107	C≡N	0,9690	0,8760
47	AC-C≡N	108	AC-C≡N	1,3342	0,9960
48	AN	109	ANH	0,5089	0,4920
		110	AN(5r)	0,3428	0,2120
		111	AN(6r)	0,3428	0,2120
49	AO	112	AO	0,2109	0,2160
		113	AS	0,7119	0,5200
50	AC-N=N	114	AC-N=N	1,4305	0,9120
51	CO-N-NH ₂	115	CO-NH-NH ₂		
		116	CO-N-NH ₂		
52	CO-N-NH	117	CO-NH-NH		
		118	CO-N-NH		
53	CO-N-N	119	CO-NH-N	1,5893	1,1280
		120	CO-N-N	1,3421	0,8240
54	SO ₂ -NH ₂	121	SO ₂ -NH ₂	2,033	1,7360
55	SO ₂ -NH	122	SO ₂ -NH	1,8708	1,4360
56	SO ₂ -N	123	SO ₂ -N	1,6236	1,1320
57	PO ₄	124	O=PO ₃	1,7007	1,6000

		125	S=PO ₃	2,3639	2,1680
		126	S=PO ₂ S	2,8319	2,4480
		127	O=PS ₃	3,1048	2,4400
58	N-CO-NH ₂	128	NH ₂ -CO-NH ₂	2,1608	2,0320
		129	NH-CO-NH ₂	1,9987	1,7320
		130	N-CO-NH ₂	1,7515	1,4280
		131	NH ₂ -CS-NH ₂	2,824	2,6000
		132	NH-CS-NH ₂	2,6618	2,3000
		133	N-CS-NH ₂	2,4146	1,9960
59	N-CO-NH	134	NH-CO-NH	1,8365	1,4320
		135	N-CO-NH	1,5893	1,1280
		136	NH-CS-NH	2,4997	2,0000
		137	N-CS-NH	2,2525	1,6960
60	N-CO-N	138	N-CO-N	1,3421	0,8240
		139	N-CS-N	2,0053	1,3920
61	O-CO-NH ₂	140	O-CO-NH ₂	1,71	1,5760
		141	S-CO-NH ₂	2,178	1,8560
62	O-CO-NH	142	O-CO-NH	1,5478	1,2760
		143	S-CO-NH	2,0158	1,5560
63	O-CO-N	144	O-CO-N	1,3006	0,9720
		145	S-CO-N	1,7686	1,2520
64	O-CO-O	146	O-CO-O		
65	CO-N-CO	147	CO-NH-CO	2,0751	1,6760
		148	CO-N-CO	1,8279	1,3720
66	CO-O-CO	149	CO-O-CO	1,7864	1,5200
67	SO	150	>S=O	1,0244	0,7760
68	SO ₂	151	SO ₂	1,3382	1,0400

69	SO ₃	152	SO ₂ -OH	1,8682	1,6240
		153	SO ₂ -O	1,5821	1,2800
70	CON-CON	154	CO-NH-CO-NH	2,6078	2,0720
		155	CO-N-CO-NH	2,3606	1,7680
		156	CO-NH-CO-N	2,3606	1,7680
		157	CO-N-CO-N	2,1134	1,4640
71	CON-CONCO	158	CO-NH-CO-NH-CO	3,379	2,7120
		159	CO-NH-CO-N-CO	3,1318	2,4080
		160	CO-N-CO-N-CO	2,8846	2,1040
72	O=C=N	161	O=C=N		
		162	=C=N		
73	H ₂ O	163	H ₂ O	0,9200	1,4000

3.4 Estimation des paramètres d'interaction

Les paramètres d'interaction des groupes fonctionnels rassemblés dans le Tableau 3.1 ont été estimés à l'aide de la méthode du Simplex de Nelder-Mead [166]. Pour cette phase d'optimisation, plus de 200 paramètres d'interaction a_{mn} et b_{mn} ont été générés pour ces groupes, en utilisant les données relatives à plus de 110 systèmes binaires soluté / solvant, tableau 3.3.

Tableau 3.3 Paramètres d'interaction a_{mn} et b_{mn} estimés pour UNIFAC Pharma

m	n	a_{mn}	b_{mn}
1	6 (Ortho)	81,7376	$5,183 \cdot 10^{-1}$
1	6 (Meta)	$2,464 \cdot 10^{-5}$	$7,467 \cdot 10^{-3}$
1	67	40,2261	$1,011 \cdot 10^{-1}$
2	6	$-1,064 \cdot 10^{-5}$	$-5,225 \cdot 10^{-3}$
2	6 (Ortho)	-34,3933	$1,162 \cdot 10^{-1}$
2	6 (Meta)	0,0000	$-3,017 \cdot 10^{-3}$
2	10	3,6070	5,7019

2	14	-28,9144	-1,3316
2	17	-15,9475	$1,632 \cdot 10^{-1}$
2	18	3,0901	-3,9437
2	19 (Ester)	36,7107	$1,014 \cdot 10^{-1}$
2	24	-23,4149	$2,053 \cdot 10^{-2}$
2	25	-117,6442	0,0000
2	33	-21,4252	0,0000
2	34 (alcane)	-28,5279	2,1637
2	34 (Ether)	-4,6050	0,0000
2	35	-5,0350	0,0000
2	44	$2,083 \cdot 10^{-3}$	$8,217 \cdot 10^{-2}$
2	46	245,4770	$-6,141 \cdot 10^{-2}$
2	48	$5,366 \cdot 10^{-3}$	1,5999
2	67	-32,8103	$1,489 \cdot 10^{-1}$
5 (Alcool)	6 (Ortho)	$-2,631 \cdot 10^{-3}$	$-7,894 \cdot 10^{-1}$
5 (Alcool)	6 (Meta)	$-4,867 \cdot 10^{-8}$	$-1,476 \cdot 10^{-5}$
6	2	$-1,699 \cdot 10^{-5}$	$-8,345 \cdot 10^{-3}$
6 (Ortho)	1	81,7736	$-7,688 \cdot 10^{-1}$
6 (Ortho)	2	-33,2940	1,1508
6 (Ortho)	5 (Alcool)	$-5,869 \cdot 10^{-3}$	-1,7657
6 (Ortho)	13 (cétone)	$-1,835 \cdot 10^{-4}$	$-5,471 \cdot 10^{-2}$
6 (Ortho)	18	$-4,573 \cdot 10^{-4}$	$-1,089 \cdot 10^{-1}$
6 (Ortho)	19 (Ester)	$-6,537 \cdot 10^{-4}$	$-1,949 \cdot 10^{-1}$
6 (Ortho)	25	-38,9626	0,0000
6 (Ortho)	33	3,1852	0,0000
6 (Ortho)	34 (alcane)	$-4,947 \cdot 10^{-4}$	$-1,475 \cdot 10^{-1}$
6 (Ortho)	35	-1,7163	0,0000
6 (Meta)	1	$2,458 \cdot 10^{-5}$	$7,450 \cdot 10^{-3}$
6 (Meta)	2	$-1,589 \cdot 10^{-5}$	$-4,817 \cdot 10^{-3}$
6 (Meta)	5 (Alcool)	$-1,788 \cdot 10^{-7}$	$-5,415 \cdot 10^{-5}$

10	2	14,6142	1,3722
10	18	$1,381 \cdot 10^{-2}$	2,5323
10	34 (alcane)	-11,6273	$1,237 \cdot 10^{-1}$
13 (cétone)	6 (Ortho)	$-2,396 \cdot 10^{-4}$	$-7,145 \cdot 10^{-2}$
14	2	84,9072	-1,0759
14	18	4,3684	$-2,345 \cdot 10^{-1}$
17	2	-47,4642	$-1,223 \cdot 10^{-1}$
17	18	7,7567	2,9218
18	2	-6,9999	4,4935
18	6 (Ortho)	$-1,583 \cdot 10^{-3}$	$1,557 \cdot 10^{-1}$
18	10	31,4192	$3,796 \cdot 10^{-1}$
18	14	142,3639	$-1,949 \cdot 10^{-1}$
18	17	-10,5145	$2,272 \cdot 10^{-1}$
18	19 (Ester)	-21,4441	$7,943 \cdot 10^{-2}$
18	24	-17,4628	$-1,492 \cdot 10^{-2}$
18	25	3,8113	0,0000
18	33	-22,9900	0,0000
18	34 (alcane)	-24,9575	1,5347
18	34 (Ether)	-3,4380	0,0000
18	44	$1,097 \cdot 10^{-3}$	$-1,803 \cdot 10^{-1}$
18	46	183,9049	$-2,123 \cdot 10^{-2}$
18	48	$-6,991 \cdot 10^{-3}$	-2,0847
18	67	-24,4620	$2,497 \cdot 10^{-2}$
19 (Ester)	2	-31,3952	$-1,833 \cdot 10^{-1}$
19 (Ester)	6 (Ortho)	$-3,064 \cdot 10^{-4}$	$-9,135 \cdot 10^{-2}$
19 (Ester)	18	7,1732	16,9777
24	2	-25,1809	$-1,365 \cdot 10^{-2}$
24	18	-1,2098	$7,324 \cdot 10^{-2}$
25	2	-109,3136	0,00

25	6 (Ortho)	-38,4030	$6,941 \cdot 10^{-3}$
25	18	$1,126 \cdot 10^{-1}$	0,0000
25	33	10,4067	0,0000
25	35	-5,6172	0,0000
33	2	-26,3754	0,0000
33	6 (Ortho)	3,1798	$4,831 \cdot 10^{-1}$
33	18	$-3,424 \cdot 10^{-1}$	0,0000
33	25	10,3744	0,0000
33	35	1,4810	0,0000
34 (alcane)	2	-19,8084	-1,5345
34 (Ether)	2	$-1,843 \cdot 10^{-4}$	$-5,495 \cdot 10^{-2}$
34 (alcane)	6 (Ortho)	-13,4990	$-1,750 \cdot 10^{-1}$
34 (alcane)	10	-26,4967	$-3,039 \cdot 10^{-1}$
34 (alcane)	18	-5,6902	0,0000
34 (Ether)	18	$1,018 \cdot 10^{-1}$	0,0000
35	2	-15,4117	0,0000
35	6 (Ortho)	-5,4144	$-5,336 \cdot 10^{-3}$
35	25	-18,0209	0,0000
35	33	1,4810	0,0000
44	2	$-8,423 \cdot 10^{-3}$	-1,8523
44	18	$1,107 \cdot 10^{-4}$	$1,107 \cdot 10^{-2}$
46	2	192,3319	$4,583 \cdot 10^{-3}$
46	18	-6,5236	1,9174
48	2	$-9,836 \cdot 10^{-4}$	$-2,936 \cdot 10^{-1}$
48	18	$-1,436 \cdot 10^{-4}$	$-4,282 \cdot 10^{-2}$
67	1	45,5220	$5,114 \cdot 10^{-1}$
67	2	-49,9322	-1,8759
67	18	-1,7154	$9,694 \cdot 10^{-3}$

Certains paramètres d'interaction b_{mn} ont été estimés pour certains groupes, considérés comme nuls, dans le modèle UNIFAC PHARMA modifié, Tableau 3.4

Tableau 3.4 Paramètres d'interaction b_{mn} estimés pour UNIFAC Pharma

m	n	b_{mn}
1	6	$1,296 \cdot 10^{-2}$
1	10	-2,8973
1	14	$1,084 \cdot 10^{-1}$
1	17	$-5,211 \cdot 10^{-2}$
1	18	-1,3851
1	44	$3,268 \cdot 10^{-1}$
1	46	3,3070
5 (Alcool)	6	$-5,823 \cdot 10^{-5}$
5 (Alcool)	10	$3,260 \cdot 10^{-1}$
5 (Alcool)	18	5,8856
6	1	$7,045 \cdot 10^{-3}$
6	5 (Alcool)	$-6,759 \cdot 10^{-5}$
10	1	$3,923 \cdot 10^{-1}$
10	5 (Alcool)	1,7918
14	1	-1,4568
17	1	$2,770 \cdot 10^{-1}$
18	1	$1,086 \cdot 10^{-1}$
18	5 (Alcool)	$5,364 \cdot 10^{-1}$

3.5 Résultats et discussions

Les paramètres d'interaction, estimés dans ce travail, ont été utilisés pour calculer la solubilité de plusieurs systèmes binaires à l'aide de deux modèles UNIFAC et UNIFAC Pharma Modifié et les résultats obtenus ont été comparés aux résultats expérimentaux. Les résultats obtenus sont représentés sur les graphes 3.5 et 3.6. Les représentations graphiques montrent la différence entre les deux modèles pour le calcul de la solubilité des composés complexes.

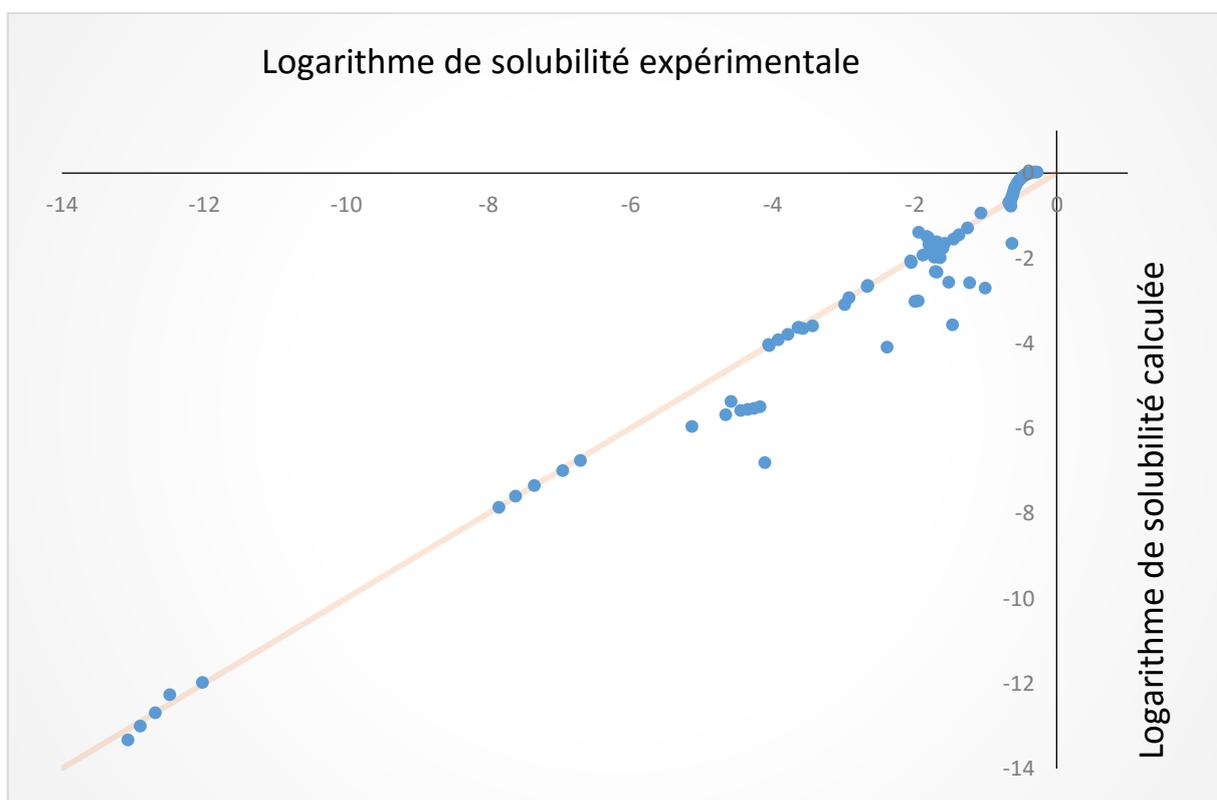


Figure 3.5 : Représentation graphique de solubilités corrélées par le modèle UNIFAC Pharma en fonction des solubilités expérimentales pour des systèmes binaires.

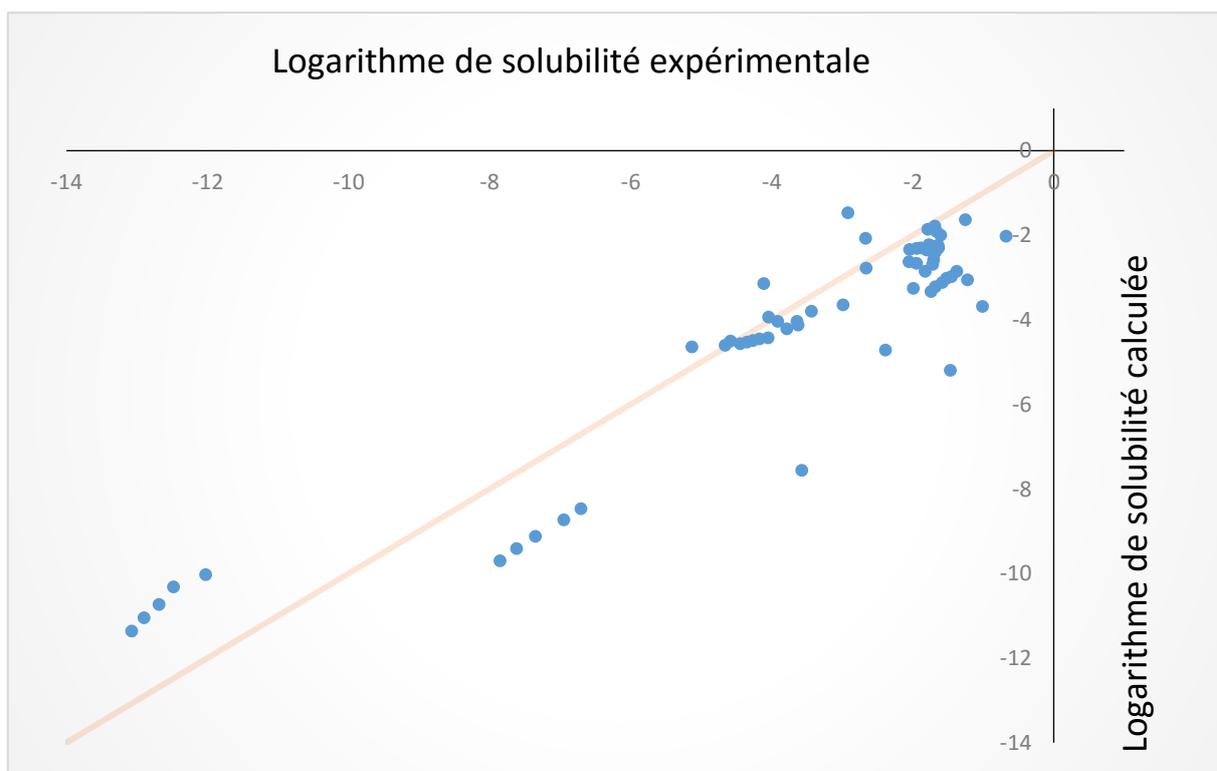


Figure 3.6 : Représentation graphique de solubilités corrélées par le modèle UNIFAC en fonction des solubilités expérimentales pour des systèmes binaires.

Afin de valider le modèle UNIFAC Pharma Modifié en utilisant les paramètres d'interaction estimés dans cette étude, on a calculé la solubilité de certains composés en utilisant les deux modèles UNIFAC original et UNIFAC Pharma Modifié. Ensuite, on a calculé les déviations des solubilités prédites (par les deux modèles) par rapport aux solubilités expérimentales pour plusieurs systèmes binaire (soluté/solvant), les déviations sont calculées par la relation :

$$ARD = \frac{1}{N} \sum \left| \frac{(X_{exp} - X_{prédite}) \times 100}{X_{exp}} \right| \quad (3.1)$$

Les résultats obtenus sont présentés dans le tableau 3.5

Dans certains cas, nous n'avons pas pu calculer la solubilité de certains composés à l'aide du modèle UNIFAC car ce dernier ne contenait pas les groupes fonctionnels représentant ces molécules, par exemple dans le cas de CON, AN et CO. Pour cette raison, la solubilité a été calculée à l'aide du modèle UNIFAC uniquement.

Tableau 3.5 Déviations des solubilités prédites par UNIFAC Pharma Modifié et UNIFAC original par rapport aux solubilités expérimentales

Systeme soluté / solvant	T (K)	UNIFAC ARD (%)	UNIFAC Pharma ARD (%)
4-Aminobenzoic acid / 1-Octanol	298,15	34,02	93,06
4-Aminobenzoic acid / 1-Butanol	298,15	17,39	86,14
4-Aminobenzoic acid / 1,4-Dioxane	298,15	9,98	1,59
4-Aminobenzoic acid / 1-Pentanol	298,15	32,22	1,33
4-Aminobenzoic acid / 1-Decanol	298,15	30,29	2,48
4-Aminobenzoic acid / 2-Butanol	298,15	98,14	7,64
4-Aminobenzoic acid / 2-Methyl-1-Propanol	298,15	12,42	1,06
4-Aminobenzoic acid / 2-Propanol	298,15	29,78	14,07
4-Aminobenzoic acid / Ethanol	298,15	47,95	10,38
Benzoic acid / 1,4-Dioxane	298,15	30,96	2,67
Benzoic acid / 1-Butanol	298,15	32,17	14,36
Benzoic acid / 1-Heptanol	298,15	47,87	14,71
Benzoic acid / 1-Hexanol	298,15	46,23	12,13
Benzoic acid / 1-Octanol	298,15	24,14	5,10
Benzoic acid / 1-Pentanol	298,15	43,84	9,49
Benzoic acid / 2-Methyl-1-Propanol	298,15	33,33	4,44
Benzoic acid / 2-Propanol	298,15	44,49	28,79
Benzoic acid / Acetone	298,15	9,31	7,81
Benzoic acid / Chlorobenzene	298,15	101,22	161,00
Benzoic acid / Cyclohexane	298,15	9,69	54,26
Benzoic acid / Decane	298,15	23,88	73,20
Benzoic acid / Diethyl Ethe	298,15	45,57	6,48
Benzoic acid / Dimethyl Sulfoxide	298,15	74,01	1,34
Benzoic acid / Ethanol	298,15	41,04	18,14
Benzoic acid / Benzyl Alcohol	298,15	16,89	106,93
Benzoic acid / Ethylene Glycol	298,15	38,98	79,61

Benzoic acid / Glycerol	298,15	166,77	93,25
Benzoic acid / Heptane	298,15	10,59	67,76
Benzoic acid / Hexane	298,15	6,29	64,15
Benzoic acid / N-Methyl Pyrrolidone	296,35	pas possible	7,08
Benzoic acid / N-Methyl Pyrrolidone	305,05	pas possible	6,47
Benzoic acid / N-Methyl Pyrrolidone	307,95	pas possible	10,91
Benzoic acid / N-Methyl Pyrrolidone	311,05	pas possible	15,63
Benzoic acid / N-Methyl Pyrrolidone	314,25	pas possible	20,38
Benzoic acid / N-Methyl Pyrrolidone	318,35	pas possible	27,37
Benzoic acid / N-Methyl Pyrrolidone	321,95	pas possible	32,64
Benzoic acid / Dimethyl acetamide	298,15	pas possible	12,14
Benzoic acid / Methyl formate	298,15	pas possible	13,99
Benzoic acid / Octane	298,15	16,07	69,92
Benzoic acid / Propionic Acid	298,15	21,24	8,83
Benzoic acid / Ethyl Acetate	298,15	pas possible	13,00
Benzoic acid / Nonane	298,15	20,19	71,69
Benzoic acid / Pentane	298,15	64,83	55,75
2-Hydroxybenzoic acid / 1,4-Dioxane	298,15	83,78	74,05
2-Hydroxybenzoic acid / 1-Butanol	298,15	41,46	34,64
2-Hydroxybenzoic acid / 1-Octanol	293,15	79,63	10,00
2-Hydroxybenzoic acid / 1-Octanol	298,15	78,18	3,70
2-Hydroxybenzoic acid / 1-Octanol	303,15	78,35	6,45
2-Hydroxybenzoic acid / 1-Octanol	310,15	78,10	9,18
2-Hydroxybenzoic acid / 1-Octanol	315,15	77,04	7,38
2-Hydroxybenzoic acid / 1-Pentanol	298,15	63,90	39,56
2-Hydroxybenzoic acid / 2-Butanol	298,15	50,92	24,72
2-Hydroxybenzoic acid / 2-Butanone	298,15	pas possible	47,06
2-Hydroxybenzoic acid / 2-Methyl-1-Propanol	298,15	50,75	73,98
2-Hydroxybenzoic acid / 2-Methyl-2-Propanol	298,15	77,41	64,68
2-Hydroxybenzoic acid / 2-Propanol	298,15	61,92	22,25

2-Hydroxybenzoic acid / Acetone	298,15	57,70	45,58
2-Hydroxybenzoic acid / Butyl Acetate	298,15	71,66	63,81
2-Hydroxybenzoic acid / Cyclohexanone	298,15	97,57	87,76
2-Hydroxybenzoic acid / Ethyl Acetate	298,15	30,09	65,25
2-Hydroxybenzoic acid / Hexane	298,15	83,13	2,47
2-Hydroxybenzoic acid / Hexane	310,15	83,05	4,47
2-Hydroxybenzoic acid / Hexane	315,15	82,72	5,13
2-Hydroxybenzoic acid / Tetrahydrofurane	298,15	93,09	81,52
2-Hydroxybenzoic acid / Dibutyl ether	298,15	90,17	81,67
4-Hydroxybenzoic acid / Hexane	293,15	457,33	22,01
4-Hydroxybenzoic acid / Hexane	298,15	545,21	9,40
4-Hydroxybenzoic acid / Hexane	310,15	778,83	24,47
4-Hydroxybenzoic acid / Hexane	315,15	641,61	5,42

Dans certains cas on trouve des déviations supérieures à 100 %. Ceci est très logique, car dans certains cas, les solubilités dans un solvant donné sont très faibles et le moindre écart peut nous donner des résultats en dehors de l'intervalle de grandeur dans laquelle la solubilité expérimentale a été mesurée. Par exemple, dans le cas de l'acide 4-hydroxybenzoïque, la solubilité expérimentale dans l'hexane est de $2,1 \cdot 10^{-6}$, et la solubilité prédite par le modèle UNIFAC est de $11,704 \cdot 10^{-6}$. Donc, d'après l'équation 3.1, l'ARD est calculé comme suit :

$$ARD = \left| \frac{(X_{exp} - X_{prédite}) \times 100}{X_{exp}} \right| = \left| \frac{(2,1 \cdot 10^{-6} - 11,704 \cdot 10^{-6}) \times 100}{2,1 \cdot 10^{-6}} \right| = 457,33$$

On peut constater que les résultats, obtenus par le modèle UNIFAC Pharma Modifié, sont généralement meilleurs que ceux issus du modèle UNIFAC original. Même si, dans certains cas, le modèle UNIFAC original donne de meilleurs résultats, cela doit être dû au fait que ces composés ont déjà été inclus dans le même modèle afin d'estimer les paramètres d'interaction.

Il est également clair, à la lumière des résultats, que le modèle UNIFAC Pharma Modifié a permis de calculer la solubilité des isomères tels que l'acide 2-hydroxybenzoïque et l'acide 4-hydroxybenzoïque, alors que le modèle UNIFAC a échoué dans ce cas. Cela signifie

qu'on considère correctement la différence entre les paramètres d'interaction en fonction de la position du groupe fonctionnel sur le cycle aromatique.

Toutefois, il convient de noter que même si l'estimation des paramètres d'interaction par le modèle UNIFAC Pharma Modifié a donné des résultats très satisfaisants, comme le montre la Figure 3.5, il a montré également son incapacité à calculer la solubilité de certains systèmes binaires. Cela peut être dû à plusieurs raisons, notamment :

- En premier lieu, le calcul des paramètres de volume et de surface dans ce modèle a été effectué par une estimation, ce qui a créé une différence significative entre ceux-ci et ceux calculés selon la méthode de Bondi utilisée dans le modèle UNIFAC. Par conséquent, il est préférable d'éviter les estimations dans ce cas, car il existe des méthodes pour calculer le volume et la surface de ces groupes fonctionnels par le calcul de volume de van der Waals, en utilisant la méthode Bondi [167] ou d'autres méthodes [168].
- En second lieu, le modèle UNIFAC Pharma Modifié a considéré une nouvelle subdivision de molécules pour les grands groupes, le modèle n'a pas pris en considération les interactions entre les groupes formés pour ce grand groupe fonctionnel. Cela conduit à une estimation inefficace car plusieurs interactions ne sont pas prises en compte, c'est pourquoi on constate de bons résultats avec les composés qui ont été inclus dans le processus d'estimation contrairement à certains composés qui n'ont pas été inclus.
- Pour les groupes qui ont plusieurs paramètres d'interaction, par exemple $\text{CH}_2\text{-OH}$, on trouve que le modèle considère que leurs interactions dans le composé alcoolique sont différentes de leurs interactions dans les alcanes. Dans le Tableau 3.5, le modèle a permis de prédire la solubilité de certains composés alcooliques et a échoué avec d'autres composés. Cela indique que cette considération n'est pas précise dans le processus d'estimation.

D'après les études réalisées et les estimations effectuées, on pense qu'il est préférable de ne calculer qu'un paramètre d'interaction pour chaque groupe après l'inclusion d'une combinaison de composés alcooliques et alcalins dans le processus d'estimation. Parce que quelle que soit la composition, le groupe reste le même et la différence réside dans les interactions avec les autres groupes.

3.6 Conclusion

Dans ce chapitre, notre objectif consistait à déterminer des paramètres d'interaction manquants dans la matrice des groupes fonctionnels du modèle UNIFAC Pharma Modifié. Et ce, pour améliorer les performances du modèle dans le calcul de la solubilité des substances actives pharmaceutiques. Nous avons, par ailleurs, fait en sorte que le modèle prenne en compte la différence entre les isomères comme c'est le cas des isomères de l'acide Hydroxybenzoïque.

Les nouveaux paramètres d'interaction estimés ont permis de calculer la solubilité pour plusieurs composés dont les données n'ont pas été utilisés lors de l'estimation. Les résultats ont montré que le modèle UNIFAC Pharma Modifié est meilleur que le modèle original pour prédire la solubilité des composés pharmaceutiques dans la plupart des cas. En revanche, le modèle UNIFAC était préférable dans certains cas, ce qui nous a amenés à déduire que :

- Il n'est pas convenable de considérer que les paramètres d'interaction de certains groupes varient en fonction du composé qu'ils contiennent, comme dans le cas de CO (cétone), CO (alcane), C-O (Ether), C-O (alcane), COO- (Ester), COO- (alcane), CO (cétone), CO (alcane), CH₂-OH (Alcool), CH₂-OH (alcane).
- Il est préférable d'inclure tous les composés chimiques ensemble dans le processus d'estimation afin de trouver des paramètres d'interaction communs entre ces groupes.
- Il n'est pas approprié de considérer les grands groupes avec une longue chaîne, chose qui conduit à la non-prise en compte des interactions entre les petits groupes qui les constituent. De ce fait il s'avère plus utile d'opter pour une subdivision des molécules en groupes plus petits où les interactions entre ces derniers sont mieux prises en considération.
- Il est nécessaire de considérer la différence entre les paramètres d'interaction en fonction de la position du groupe fonctionnel sur le cycle aromatique.

Toutes ces conclusions nous ont amenés à réfléchir sur l'élaboration d'un nouveau modèle UNIFAC exempt des failles identifiées

Développement d'un nouveau modèle

UNIFAC

Sommaire :

4.1	Introduction	64
4.2	Structure du modèle.....	64
4.2.1	Equations du modèle	64
4.2.2	Paramètres de groupe et de surface	66
4.3	Implémentation du modèle	74
4.4	Méthode de régression	79
4.5	Estimation des paramètres d'interaction du modèle	81
4.6	Conclusion.....	87

4.1 Introduction

Dans ce chapitre, un nouveau modèle UNIFAC est développé à partir de KT-UNIFAC et UNIFAC Pharma modifié, présentés dans le chapitre précédent. Ce nouveau modèle, nommé NK-UNIFAC est basé sur les mêmes équations utilisées dans UNIFAC Pharma modifié [34] mais en considérant une dépendance linéaire de la température des paramètres d'interaction.

La décomposition de la structure chimique des molécules ainsi que les paramètres de volume et de surface des groupes fonctionnels sont les mêmes que celles adoptées par KT-UNIFAC [33].

Par ailleurs, une nouvelle méthode adoptée pour l'estimation des paramètres d'interaction avec une très grande précision a été utilisée et est également présentée dans ce chapitre. Cette méthode basée sur l'utilisation de la méthode du Simplex de Nelder-Mead effectue tous les calculs de solubilité pour les différents systèmes binaires soluté/solvant simultanément. Et ce, en changeant à chaque fois la fonction objective pendant l'optimisation des paramètres. Cette approche est essentielle à la réussite de ce modèle. La méthode du Simplex de Nelder-Mead intégrée dans Matlab, a été convertie sous Visual Basic (VBA), puis utilisée sous Excel avec une interface graphique pour le programme d'ajustement.

4.2 Structure du modèle

4.2.1 Equations du modèle

Comme le modèle UNIFAC Pharma modifié et les autres versions du modèle UNIFAC, NK-UNIFAC se décompose en deux termes : un terme combinatoire qui tient compte des différences de taille et de volume et un terme résiduel associé aux interactions énergétiques entre les groupes [169].

$$\ln \gamma_i = \underbrace{\ln \gamma_i^C}_{\text{Partie combinatoire}} + \underbrace{\ln \gamma_i^R}_{\text{partie résiduelle}} \quad (4.1)$$

✓ La partie combinatoire

$$\ln \gamma_i^{comb} = 1 - V_i' + \ln V_i' - 5q_i \left(1 - \frac{V_i}{F_i} + \ln \frac{V_i}{F_i} \right) \quad (4.2)$$

Avec :

$$V_i = \frac{r_i}{\sum_j(x_j r_j)} \quad (4.3)$$

$$F_i = \frac{q_i}{\sum_j(x_j q_j)} \quad (4.4)$$

$$V'_i = \frac{r_i^{3/4}}{\sum_j(x_j r_j^{3/4})} \quad (4.5)$$

V_i, V'_i sont des fractions de volume, F_i est la fraction de surface.

Les paramètres du constituant pur r_i et q_i correspondent respectivement au volume moléculaire et à la surface moléculaire de Van Der Waals. Ils sont calculés par sommation des paramètres adimensionnels R_k et Q_k

$$r_i = \sum_k(v_k^{(i)} R_k) \quad (4.6)$$

$$q_i = \sum_k(v_k^{(i)} Q_k) \quad (4.7)$$

$v_k^{(i)}$ est le nombre des groupes de type k dans la molécule i .

R_k et Q_k appelés paramètres de volume de surface, respectivement, sont recueillis à partir des volumes et des surfaces des groupes de Van Der Waals, V_k et A_k donnés par Bondi [165]. Ces paramètres s'expriment par :

$$R_k = \frac{V_k}{15,17} \quad (4.8)$$

$$Q_k = \frac{A_k}{(2,5 \times 10^9)} \quad (4.9)$$

✓ Partie résiduelle

La partie résiduelle est la somme des différentes contributions de chaque groupe du corps dissous dans la solution, moins la somme des différentes contributions dans l'environnement du composé pur.

$$\ln \gamma_i^{Res} = \sum_k [v_k^{(i)} (\ln \Gamma_k - \ln \Gamma_k^{(i)})] \quad (4.10)$$

Γ_k : coefficient résiduel d'activité du groupe

$\Gamma_k^{(i)}$: coefficient résiduel d'activité du groupe k dans une solution de référence contenant seulement des molécules de type i .

$v_k^{(i)}$: Nombre des groupes de type k dans une molécule de type i .

Les différentes contributions du groupe, dans tout environnement contenant des groupes de type 1, 2...n, sont supposés être, seulement fonction des concentrations et de la température.

$$\ln \Gamma_k = Q_k \left[1 - \ln(\sum_m (\theta_m \Psi_{mk})) - \sum_m \left(\frac{\theta_m \Psi_{km}}{\sum_n (\theta_n \Psi_{nm})} \right) \right] \quad (4.11)$$

$$\theta_m = \frac{Q_m X_m}{\sum_n (Q_n X_n)} \quad (4.12)$$

$$X_m = \frac{\sum_j (v_m^{(j)} x_j)}{\sum_j \sum_n (v_n^{(j)} x_j)} \quad (4.13)$$

X_m : Fraction du groupe m dans le mélange

θ_m : Fraction de surface du groupe m

$$\Psi_{nm} = \exp \left(- \frac{a_{nm} + (b_{nm} \times T)}{T} \right) \quad (4.14)$$

Ψ_{nm} : facteur de Boltzmann correspondant au paramètre d'interaction a_{nm} et b_{nm} .

a_{nm} , b_{nm} caractérisent les interactions entre les groupes n et m , pour chaque paire de groupes. Il y a quatre valeurs de paramètres d'interaction : $a_{mn} \neq a_{nm} \neq b_{mn} \neq b_{nm}$.

4.2.2 Paramètres de groupe et de surface

Les groupes, les sous-groupes et les paramètres de volume et de surface (R_k et Q_k) dans le modèle NK-UNIFAC, proposé dans cette étude, sont les mêmes que ceux de KT-UNIFAC [33].

Tableau 4.1 : Groupes et paramètres de volume R_k et de surface Q_k

N1	Groupe principal	N2	Sous-groupe	R_k	Q_k
1	CH ₂	1	CH ₃	0,9011	0,8480
		2	CH ₂	0,6744	0,5400
		3	CH	0,4469	0,2280
		4	C	0,2195	0,0000
2	C=C	5	CH ₂ =CH	1,3454	1,1760
		6	CH=CH	1,1167	0,8670

		7	CH ₂ =C	1,1173	0,9880
		8	CH=C	0,8886	0,6760
		9	C=C	0,6605	0,4850
3	ACH	10	ACH	0,5313	0,4000
		11	AC (link)	0,3652	0,1200
		12	AC (cond)	0,3125	0,0840
4	ACCH ₂	13	ACCH ₃	1,2663	0,9680
		14	ACCH ₂	1,0396	0,6600
		15	ACCH	0,8121	0,3480
		16	ACC	0,5847	0,0840
5	OH	17	OH	1,0000	1,2000
6	CH ₃ OH	18	CH ₃ OH	1,4311	1,4320
7	H ₂ O	19	H ₂ O	0,9200	1,4000
8	ACOH	20	AC-OH	0,8952	0,6800
9	CH ₂ CO	21	CH ₃ CO	1,6724	1,4880
		22	CH ₂ CO	1,4457	1,1800
		23	CHCO	1,2182	0,8680
		24	CCO	0,9908	0,6400
10	CHO	25	CHO	0,9980	0,9480
11	CCOO	26	CH ₃ COO	1,9031	1,7280
		27	CH ₂ COO	1,6764	1,4200
		28	CHCOO	1,4489	1,1080
		29	CCOO	1,2215	0,8800
12	HCOO	30	HCOO	1,2420	1,1880
13	CH ₂ O	31	CH ₃ O	1,1450	1,0880
		32	CH ₂ O	0,9183	0,7800
		33	CH-O	0,6908	0,4680
		34	CO	0,9183	0,2400
		35	CH ₂ O, f.	0,9183	1,1000

14	CNH ₂	36	CH ₃ NH ₂	1,5959	1,5440
		37	CH ₂ NH ₂	1,3692	1,2360
		38	CHNH ₂	1,1417	0,9240
		39	CHNH ₂	1,1417	0,9240
		40	CNH ₂	0,9275	0,6960
15	(C) ₂ NH	41	CH ₃ NH	1,4337	1,2440
		42	CH ₂ NH	1,2070	0,9360
		43	CHNH	0,9795	0,6240
16	(C) ₃ N	44	CH ₃ N	1,1865	0,9400
		45	CH ₂ N	0,9597	0,6320
17	ACNH ₂	46	ACNH ₂	1,0600	0,8160
		47	ACNH	0,8978	0,5160
		48	ACN	0,6506	0,2120
18	PYRIDINE	49	C ₅ H ₅ N	2,9993	2,1130
		50	C ₅ H ₄ N	2,8332	1,8330
		51	C ₅ H ₃ N	2,6670	1,5530
19	CCN	52	CH ₃ CN	1,8701	1,7240
		53	CH ₂ CN	1,6434	1,4160
		54	CHCN	1,4160	1,1040
		55	CCN	1,1885	0,8760
20	COOH	56	COOH	1,3013	1,2240
		57	HCOOH	1,5280	1,5320
21	CCl	58	CH ₂ Cl	1,4654	1,2640
		59	CHCl	1,2380	0,9520
		60	CCl	1,0060	0,7240
22	CCl ₂	61	CH ₂ Cl ₂	2,2564	1,9880
		62	CHCl ₂	2,0606	1,6840
		63	CCl ₂	1,8016	1,4480
23	CCl ₃	64	CHCl ₃	2,8700	2,4100

		65	CCl ₃	2,6401	2,1840
24	CCl ₄	66	CCl ₄	3,3900	2,9100
25	ACCl	67	ACCl	1,1562	0,8440
26	CNO ₂	68	CH ₃ NO ₂	2,0086	1,8680
		69	CH ₂ NO ₂	1,7818	1,5600
		70	CHNO ₂	1,5544	1,2480
		71	CNO ₂	1,3270	1,0200
27	ACNO ₂	72	ACNO ₂	1,4199	1,1040
28	CS ₂	73	CS ₂	2,0570	1,6500
29	CH ₃ SH	74	CH ₃ SH	1,8770	1,6760
		75	CH ₂ SH	1,6510	1,3680
		76	CHSH	1,4232	0,2280
		77	CSH	1,1958	0,0000
30	FURFURAL	78	FURFURAL	3,1680	2,4810
31	DOH	79	(CH ₂ OH) ₂	2,4088	2,2480
32	I	80	I	1,2640	0,9920
33	BR	81	BR	0,9492	0,8320
34	C#C	82	CH#C	1,2920	1,0880
		83	C#C	1,0613	0,7840
35	DMSO	84	CH ₃ SOCH ₃	2,8266	2,4720
36	ACRY	85	CH ₂ =CHCN	2,3144	2,0520
37	CICC	86	Cl-C=C	0,7910	0,7240
38	ACF	87	ACF	0,6948	0,5240
39	DMF	88	DMF	3,0856	2,7360
		89	CON(Me) ₂	2,8589	2,4280
		90	CONMeCH ₂	2,6322	2,1200
		91	HCON ₂ CH ₂	2,6322	2,1200
		92	CON ₂ CH ₂	2,4054	1,8120
40	CF ₂	93	CHF ₃	1,5781	1,5480

		94	CF ₃	1,4060	1,3800
		95	CHF ₂	1,2011	1,1080
		96	CF ₂	1,0105	0,9200
		97	CH ₂ F	1,0514	0,9800
		98	CHF	0,8240	0,6680
		99	CF	0,6150	0,4600
41	COO	100	COO	1,3800	1,2000
42	SiH ₂	101	SiH ₃	1,6035	1,2632
		102	SiH ₂	1,4443	1,0063
		103	SiH	1,2853	0,7494
		104	Si	1,0470	0,4099
43	SIO	105	SiH ₂ O	1,4838	1,0621
		106	SiHO	1,3030	0,7639
		107	SIO	1,1044	0,4657
44	NMP	108	NMP	3,9810	3,2000
45	CClF	109	CCl ₃ F	3,0356	2,6440
		110	CCl ₂ F	2,2287	1,9160
		111	HCCl ₂ F	2,4060	2,1160
		112	HCClF	1,6493	1,4160
		113	CClF ₂	1,8174	1,6480
		114	HCClF ₂	1,9670	1,8280
		115	CClF ₃	2,1721	2,1000
		116	CCl ₂ F ₂	2,6243	2,3760
46	CONCH ₂	117	CONHCH ₃	2,2050	1,8640
		118	HCONHCH ₂	2,2050	1,8840
		119	CONHCH ₂	1,9782	1,5760
		120	CONH ₂	1,4661	1,3360
47	OCCOH	121	C ₂ H ₅ O ₂	2,1226	1,9040
		122	C ₂ H ₄ O ₂ -1	1,8952	1,5920

		123	C ₂ H ₄ O ₂ -2	1,8952	1,5920
48	CH ₂ S	124	CH ₃ S	1,6130	1,3680
		125	CH ₂ S	1,3863	1,0600
		126	CHS	1,1589	0,7480
		127	CS	0,9314	0,5200
49	MORFOLIN	128	MORFOLIN	3,4740	2,7960
50	THIOPHEN	129	C ₄ H ₄ S	2,8569	2,1400
		130	C ₄ H ₃ S	2,6908	1,8600
		131	C ₄ H ₂ S	2,5247	1,5800
51	C=C=C	132	CH ₂ =C=CH ₂	1,5405	1,4200
		133	CH ₂ =C=C	1,4838	2,7480
		134	CH=C=CH	1,7732	0,8640
52	ACC=C	135	ACCH=CH ₂	1,7106	1,2960
		136	ACCH=CH	1,9394	1,6080
		137	ACC=CH ₂	1,4825	1,1080
53	ACC#C	138	ACC#CH	1,6572	1,2080
		139	ACC#C	1,4265	0,9040
54	CH ₂ (cyclique)	140	CH ₂ , cyclique	0,6744	0,5400
		141	CH, cyclique	0,4469	0,2280
		142	C, cyclique	0,2195	0,0000
55	C=C(cyclique)	143	CH=CH, cyclique	1,1167	0,8670
		144	CH ₂ =C, cyclique	1,1173	0,9880
		145	CH=C, cyclique	0,8886	0,6760
		146	C=C, cyclique	0,6605	0,4850
56	ACCO	147	ACCO	1,1365	0,7240
57	ACCHO...	148	ACCHO	1,3632	1,0680
58	ACCOO	149	ACCOO	1,3672	1,0000
59	ACOOC	150	ACOOC	1,3672	1,0000
60	ACO	151	ACO	0,6091	0,3600

Chapitre 4. Développement d'un nouveau modèle UNIFAC

61	NH ₂	152	NH ₂	0,6948	0,6960
62	AN	153	AN	0,5060	0,2120
63	ACCN	154	AC-CN	1,3342	0,9960
64	CN	155	CN	0,9690	0,8760
65	ACCOOH	156	ACCOOH	1,6664	1,3440
66	Cl	157	Cl	0,7910	0,7240
67	F	158	F	0,3771	0,4400
68	ACI	159	ACI	1,3619	0,9720
69	ACBR	160	ACBr	1,6599	1,1240
70	NO ₂	161	NO ₂	1,1074	1,4200
71	ACSH	162	ACSH	1,3415	0,1200
72	SH	163	SH	1,0771	1,0400
73	SO	164	SO	1,0244	0,7760
74	SO ₂	165	SO ₂	1,3382	1,0400
		166	O(SO)O	1,5122	1,2560
75	SO ₃	167	O(SO ₂)	1,5821	1,2800
76	SO ₄	168	SO ₄	1,8260	1,5200
77	ACSO	169	ACSO	1,3896	0,8960
78	ACSO ₂	170	ACSO ₂	1,7034	1,1600
79	CONCO	171	CONHCO	1,5307	1,3440
		172	CONCO	2,0751	1,6760
80	ACCONH	173	ACCONH ₂	1,8279	1,3720
		174	ACCONH	1,8312	1,4560
81	ACNHCO	175	ACNHCO	1,6691	1,1560
		176	ACNHCOH	1,8958	1,4640
		177	ACN(CO)H	1,6486	1,1600
		178	AC-NCO	1,4219	0,8520
82	UREA	179	UREA	2,1608	2,0320
		180	NH ₂ CONH	1,9987	1,7320

		181	NH ₂ CON	1,7515	1,4280
		182	NHCONH	1,8365	1,4320
		183	NHCON	1,5893	1,1280
		184	NCON	1,3421	0,8240
83	ACNCON	185	ACNHCON ₂	2,3639	1,8520
		186	ACNHCON	2,2017	1,5520
84	ONO	187	ONO	1,1074	1,4200
85	ONO ₂	188	ONO ₂	1,3514	1,6600
86	CN(cyclique)	191	CH=N, cyclique	0,8438	0,5240
		192	C=N, cyclique	0,6157	0,3360
87	N(cyclique)	193	NH, cyclique	0,5326	0,3960
		194	N, cyclique	0,2854	0,0920
88	O(cyclique)	195	O, cyclique	0,2439	0,2400
89	CO(cyclique)	196	C=O, cyclique	0,3652	0,0840
90	ACCONH	200	ACCHNOH	1,7390	1,2280
91	ACSH	201	ACS	1,0771	1,0400
92	CONCH ₂	202	HCONH	172,0000	1,5307
93	S(cyclique)	203	S, cyclique	0,7119	0,5200
94	SO ₂ (cyclique)	204	SO ₂ , cyclique	1,3382	1,0400
95	O=COC=O	208	O=COC=O	1,7864	1,5200
96	OCOO	209	OCOO	1,2591	1,1200
97	CH ₂ OCH ₂	211	CH ₂ OCH ₂	1,5926	1,3200
		212	CH ₂ OCH	1,3652	1,0080
		213	CH ₂ OC	1,1378	0,7800
		214	CHOCH	1,1378	0,6960
		215	CHOC	0,9103	0,4680
		216	COC	0,6829	0,2400
98	NCO	217	NCO	1,0567	0,7320

4.3 Implémentation du modèle

Tous les programmes de calcul ont été codés en Visual Basic (VBA), y compris, notamment, les formules du modèle UNIFAC, les fonctions utilisées dans l'exploration des données et l'appel à d'autres programmes. Le schéma de la Figure 4.1, en illustre un aperçu.

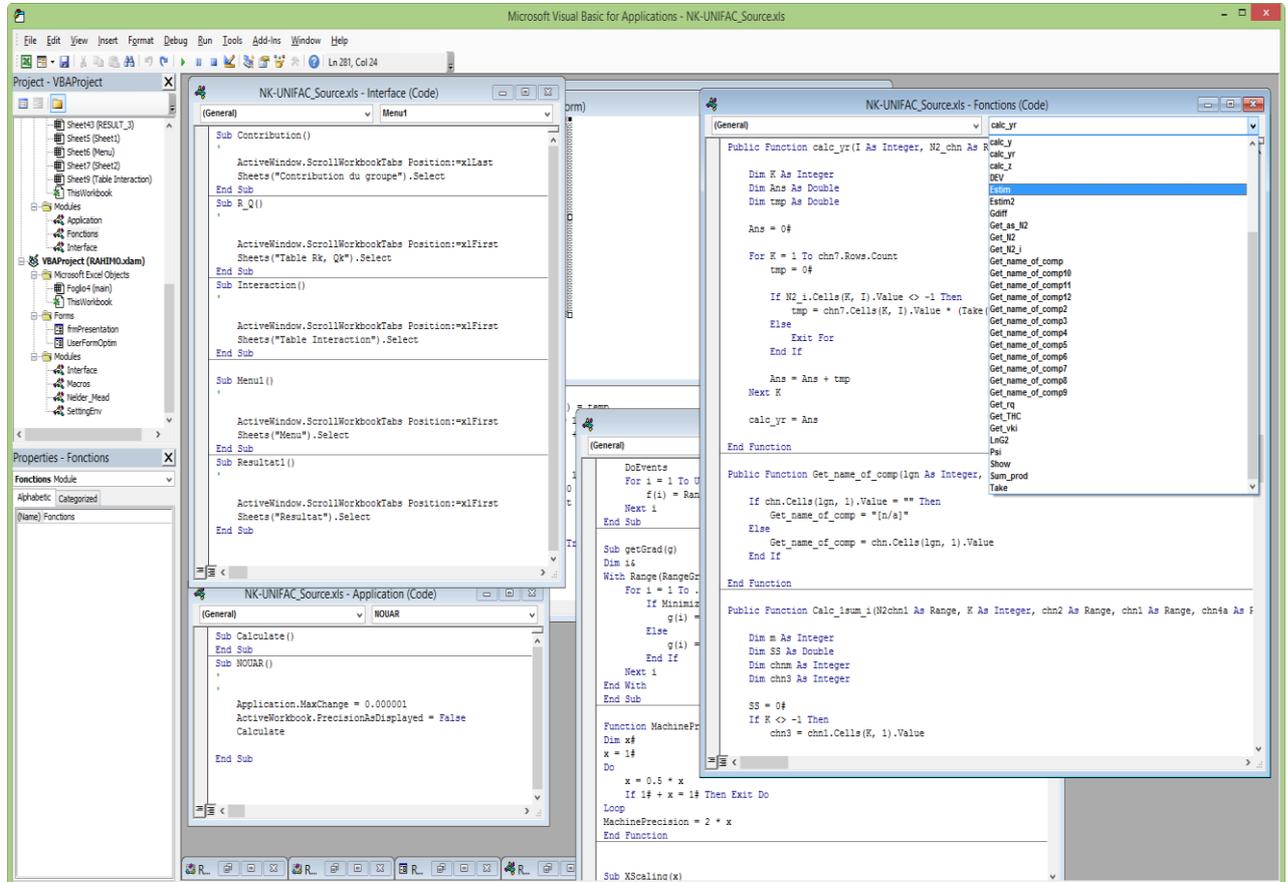


Figure 4.1 Capture d'écran VBA montre comment écrire des fonctions et des interfaces

Excel a été utilisé comme une interface d'entrée des données, le suivi des calculs et l'affichage des résultats comme indiqué dans les Figures [4.2 – 4.10]

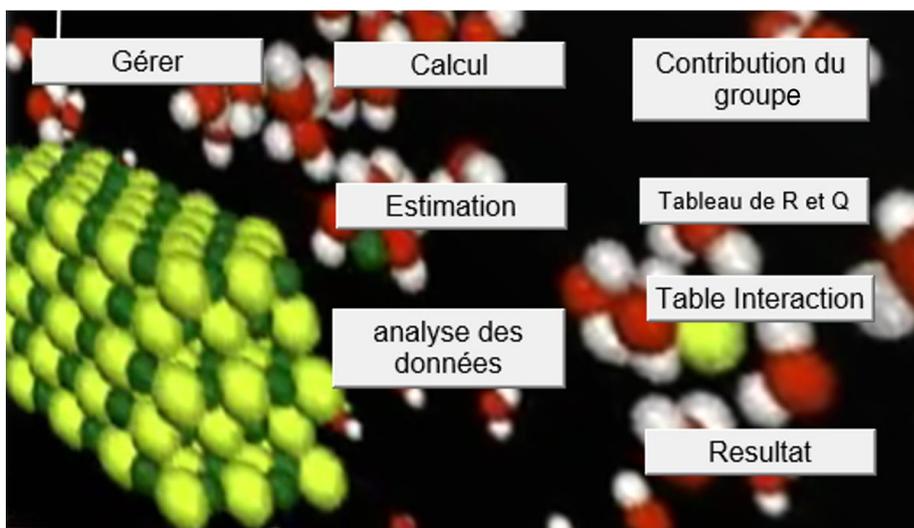


Figure 4.2 Capture d'écran du menu principal.

Temperature (K) T =		297,15	299,15	301,15	303,15	305,15
Composants						
La fraction molaire dans différentes températures						
No	ne choisissez pas "[n/a]"					
1	5-AMINOSALICYLIC ACID	0,1705	0,1876	0,2018	0,2154	0,2401
2	METHANOL	0,8295	0,8124	0,7982	0,7846	0,7599
		Coefficient d'activité, γ_i exp sans ΔC_p			Coefficient d'activité, γ_i exp avec ΔC_p	
No						
1						
2						
		ln γ_i calculé			γ_i calculé	
No						
1						
2						
		La solubilité X_i sans ΔC_p			La solubilité X_i avec ΔC_p	
No						
1						
2						

Figure 4.3 Capture d'écran Excel, introduction des fractions molaires et des températures et l'obtention des valeurs calculées de γ_i et de x_i à différentes températures.

Menu										
a_{mn}	n	1	2	3	4	5	6	7	8	9
m	Name	ACH	ACOH	ACNH2	ACCOOH	ACCI	ACNO2	AN	H2O	CH
1	ACH	0	-614,708232	-1433,789354	97,0114118	187	796,8113854	0	752,329954	-231,50
2	ACOH	-1,537945449	0	-305,5	-22,3275062	6915	-42,83245534	0	-1259,03318	421,48
3	ACNH2	-599,3054877	-253,1	0	-381,232028	617,5	-314,0500315	0	-98,241522	9195,2
4	ACCOOH	-198,7468693	-244,7429	197,1920438	0	-79,08	-160,9968754	0	-55,8417642	-551,2
5	ACCI	-97,27	622,3	334,9	993,4	0	2213	0	-274,5	106
6	ACNO2	1981,27558	-84,0123739	-469,5327709	-359,1374	-123,1	0	0	-56,122457	-455,2
7	AN	0	0	0	0	0	0	0	0	0

Figure 4.4 Capture d'écran Excel, la matrice des paramètres d'interaction a_{mn} .

b_{mn}	n	1	2	3	4	5	6	7	8	9
m	Name	ACH	ACOH	ACNH2	ACCOOH	ACCI	ACNO2	AN	H2O	CH3
1	ACH	0	-1,49075294	-1,888379066	2,93838154	0	-11,22610091	0	-0,03350098	-0,13068
2	ACOH	-2,162616252	0	0	-1,48076829	0	0,330334678	0	1,77333935	-3,54924
3	ACNH2	0,108830618	0	0	-5,9847795	0	0,118539322	0	-2,41518657	0
4	ACCOOH	0,630854824	-10,1326115	1,500443357	0	0	4,54436857	0	-0,63656872	6,60509
5	ACCI	0	0	0	0	0	0	0	0	0
6	ACNO2	-0,124674468	-16,1548814	-14,62400873	0,05711169	0	0	0	-1,50482111	0,54763
7	AN	0	0	0	0	0	0	0	0	0

Figure 4.5 Capture d'écran Excel, la matrice des paramètres d'interaction b_{mn} .

Tableau des composants											Menu		
Name:											$T_r(K)$	ΔH_i	ΔC_p
Name: water											0,92	1,4	
Sub group:	H2O												
No of subgr:	1												
	19	-1	-1	-1	-1	-1	-1	-1	-1	-1			
Name: benzene											3,1878	2,4	
Sub group:	ACH												
No of subgr:	6												
	10	-1	-1	-1	-1	-1	-1	-1	-1	-1			
Name: ethanol											2,5755	2,588	
Sub group:	CH3	CH2	OH										
No of subgr:	1	1	1										

Figure 4.6 Capture d'écran Excel, décomposition des structures chimiques des molécules avec les groupes fonctionnels correspondants.

Les paramètres R and Q pour chaque subgroup					Menu
N1	N2				
Main group	Subgroup	nom de Groupe	R_k	Q_k	
1	1	CH3	0,9011	0,8480	
1	2	CH2	0,6744	0,5400	
1	3	CH	0,4469	0,2280	
1	4	C	0,2195	0,0000	
2	5	CH2=CH	1,3454	1,1760	
2	6	CH=CH	1,1167	0,8670	
2	7	CH2=C	1,1173	0,9880	
2	8	CH=C	0,8886	0,6760	
2	9	C=C	0,6605	0,4850	
3	10	ACH	0,5313	0,4000	
3	11	AC(link)	0,3652	0,1200	
3	12	AC(cond)	0,3125	0,0840	
4	13	ACCH3	1,2663	0,9680	
4	14	ACCH2	1,0396	0,6600	
4	15	ACCH	0,8121	0,3480	
4	16	ACC	0,5847	0,0840	
5	17	OH	1,0000	1,2000	
6	18	CH3OH	1,4311	1,4320	
7	19	H2O	0,9200	1,4000	
8	20	AC-OH	0,8952	0,6800	
9	21	CH3CO	1,6724	1,4880	

Figure 4.7 Capture d'écran Excel, les paramètres de volume R_k et de surface Q_k .

La méthode Simplex dans Matlab est utilisée et convertie en format VBA, puis une interface d'estimation est créée et intégrée dans Excel, ceci est montré dans les figures 3.11 et 3.12.

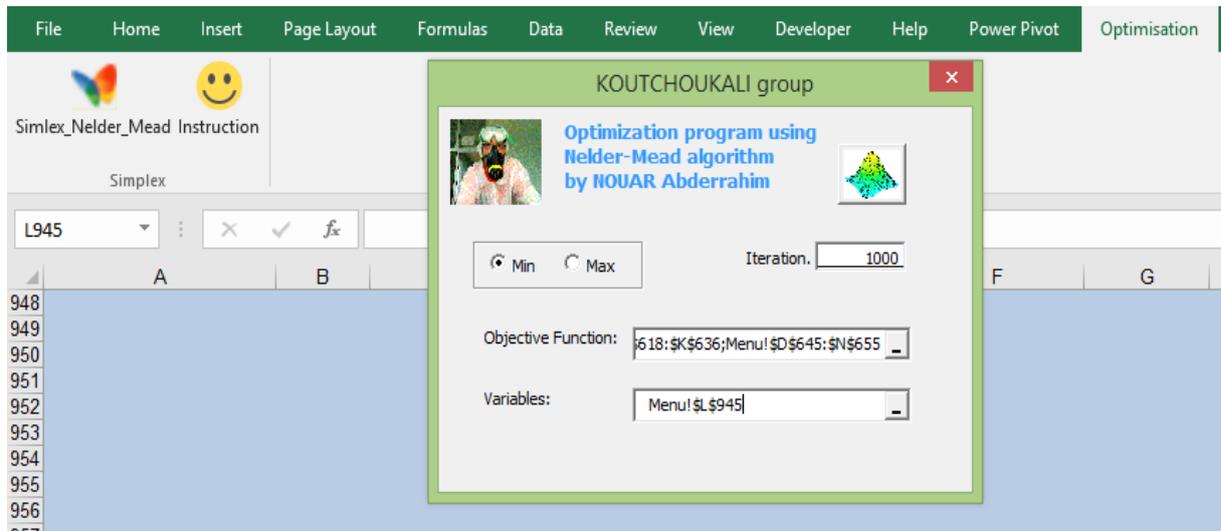


Figure 4.11 Capture d'écran du logiciel d'optimisation du Simplex de Nelder-Mead.

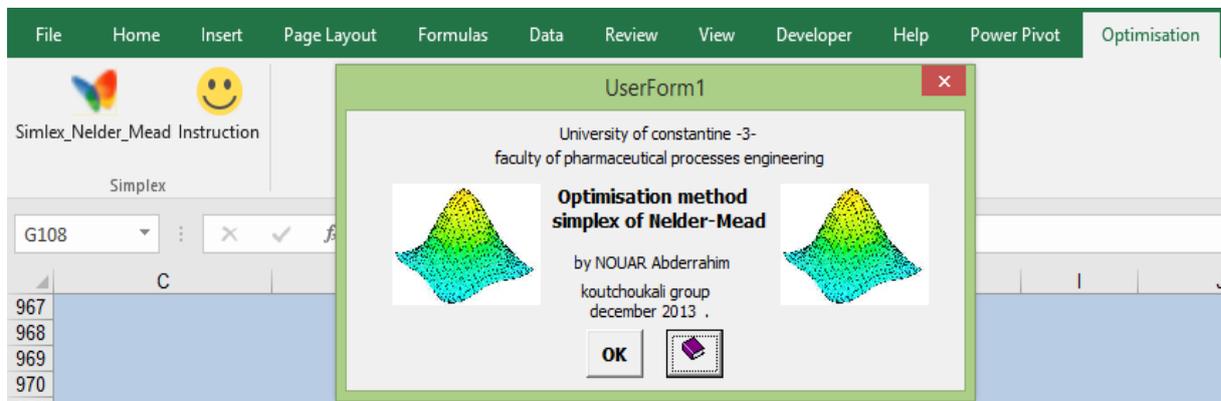


Figure 4.12 Capture d'écran du manuel d'utilisation du logiciel d'optimisation.

L'utilisation de VBA avec l'interface Excel a facilité la saisie des données et a permis le suivi des calculs. La gestion des diverses tâches et l'analyse des résultats et des données de manière systématique et rapide.

4.4 Méthode de régression

Avant de régresser les paramètres d'interaction des groupes fonctionnels (GIPs), il était nécessaire de décider si la régression devait être effectuée en série ou en parallèle. Les équivalences lors de la régression des paramètres d'interaction des groupes (GIPs) dans tous les modèles UNIFAC sont évidentes. Historiquement, les GIPs des modèles UNIFAC ont été régressés en série. Fredenslund et *al.* [14] se sont abstenus de régresser en parallèle l'ensemble (ou une partie importante du total) des GIPs, principalement parce que les ressources informatiques à l'époque n'étaient pas assez puissantes.

En outre, la surveillance de la procédure d'optimisation n'a pas été jugée pratique (et cette déclaration est toujours valide de nos jours). Dans son travail pour UNIFAC-LLE, Magnusson [170] a régressé les GIPs en parallèle pour éviter les problèmes de convergence et les problèmes de pondération "insolubles" en raison de la présence de différents types de mélanges et de composés dans les ensembles des données expérimentales. Suivant le même raisonnement, Abildskov [24] a régressé les GIPs du modèle KT-UNIFAC en utilisant une procédure parallèle en termes des solutions de groupe allant de 2 solutions de groupe (par exemple, interactions $\text{CH}_2 - \text{C}=\text{C}$, $\text{CH}_2 - \text{ACH}$, $\text{CH}_2 - \text{ACCH}_2$) et au-delà. Cependant, Abildskov indique clairement que la procédure de construction utilisée dans son travail (les paramètres sont régressés de manière séquentielle, les résultats des régressions de paramètres précédentes étant utilisés comme des suppositions initiales pour l'ajout de nouveaux paramètres) peut être considérée comme une procédure mixte "garder les avantages du montage en série tout en montage en parallèle".

Dans les travaux de González et *al.* [171], une approche similaire à celle utilisée par Abildskov [24] a été adoptée. Comme cela a été fait pour les GIPs dans l'UNIFAC d'origine, dans leur travail, les paramètres d'interaction d'atome (AIPs) du modèle UNIFAC-CI ont d'abord été régressés à l'aide de données ne contenant que des atomes de C. Deuxièmement, toutes les interactions C-C étant fixes, les AIPs des interactions C-O sont régressés. Ensuite, les interactions O-O, C-N, O-N, N-N, C-Cl, O-Cl, N-Cl, Cl-Cl, C-S, O-S, N-S, Cl-S et S-S ont été ajoutées. À la fin (ou entre toute la procédure si nécessaire), tous les AIPs ont été régressés pour une optimisation simultanée.

La signification de cette expression "optimisation simultanée" vise à inclure tous les paramètres en même temps dans la régression des GIPs et cela ne signifie pas que l'ordinateur effectue tous les calculs en même temps pour tous les systèmes. C'est-à-dire qu'ils choisissent d'inclure tous les systèmes dans la régression ou de sélectionner un ensemble de systèmes à

chaque fois. Ce que Abildskov a fait c'est la régression pour chaque ensemble de systèmes, puis de les collecter tous au cours d'une procédure de régression complète.

D'un point de vue mathématique, la complexité et la non-linéarité des expressions UNIFAC conduisent à un problème de régression non linéaire qui, le plus souvent, se traduit par un problème d'optimisation multimodal. Kratch et *al.* [172] ont étudié l'optimisation des GIPs de l'UNIFAC à l'aide des algorithmes évolutionnaires lors de la régression simultanée et séquentielle des paramètres des GIPs. Il a été confirmé que l'un des problèmes les plus courants lors de la régression en série est le risque d'incompatibilité entre les GIPs. La raison en est qu'il peut arriver qu'aucun jeu de paramètres ne puisse être trouvé afin de fournir des résultats satisfaisants, du fait que les paramètres précédemment ajustés restent constants.

D'autre part, il a été constaté que la régression simultanée de tous les paramètres conduisait à un espace de recherche plus complexe avec des exigences de calcul importantes pour les algorithmes d'optimisation employés. Leur travail, lors de la régression d'un ensemble de quatre GIPs dans UNIFAC modifié (Dortmund), n'a pas permis de déterminer avec précision l'approche la plus pratique, peut-être (encore) en raison de la complexité du scénario d'optimisation des paramètres lors de la régression des GIPs de UNIFAC. Lors de l'optimisation poussée des paramètres pour le modèle UNIFAC, la plupart (sinon la totalité) des travaux publiés utilisent des algorithmes d'optimisation locaux (Nelder-Mead [166], Levenberg [173], Marquadt [174]). Aucun travail publié n'a été trouvé en utilisant des algorithmes d'optimisation globaux pour régresser les GIPs de modèle UNIFAC.

On a sélectionné autant de systèmes binaires que possible pour régresser les paramètres d'interaction pour le moins de groupes fonctionnels possibles. À chaque fois, on fait la régression pour un ensemble différent de systèmes binaires contenant de nouveaux groupes fonctionnels, où les GIPs estimés sont fixés. Après avoir régressé les paramètres d'interaction pour le plus grand nombre de groupes fonctionnels, on effectue une procédure de régression qui inclut l'ensemble de ces derniers, y compris des systèmes binaires pouvant contenir un grand nombre de ces groupes.

Ce qui est nouveau, c'est que les optimisations sont effectuées en même temps, les fonctions objectives de tous les systèmes binaires sont calculées en même temps de manière à garantir la plus grande précision possible dans le processus d'estimation. En effet, chaque système binaire est indépendant dans son cas, comme indiqué dans la section 4.5.

Lors de la régression d'un très grand nombre de paramètres, il y a toujours des problèmes à résoudre. Un schéma de régularisation de la fonction objective a été utilisé, le cas échéant, pour contrôler la variance des GIPs et pour accélérer les calculs. Lorsque de bons résultats ne pouvaient pas être obtenus, les valeurs de régularisation ont été modifiées et l'optimisation a été effectuée à nouveau.

Les fonctions objectives sont très importantes, puisqu'une bonne sélection significative peut entraîner des temps de calcul plus courts et un meilleur ensemble de paramètres pour les phénomènes à corrélérer. Le choix correct est particulièrement important. Différentes fonctions objectives ont été appliquées à chaque fois durant le processus d'estimation.

Afin de pouvoir comparer équitablement le modèle NK-UNIFAC avec le modèle UNIFAC de référence, les fonctions objectives initiales ont été préservées dans la mesure du possible.

4.5 Estimation des paramètres d'interaction du modèle

Pour atteindre les paramètres d'interaction de groupe optimaux, le processus d'estimation passe par plusieurs étapes. Celles-ci sont détaillées ci-dessous :

1. Les composés (soluté et solvant) sont décomposés en groupes fonctionnels comme indiqué sur la figure 4.13.

Nom:	3-NITROBENZOIC		
Sub group:	ACH	AC-COOH	AC-NO2
Nbr de subgr	4	1	1
	1	8	10
Nom:	BENZOIC ACID		
Sub group:	ACH	AC-COOH	
Nbr de subgr	5	1	
	1	8	-1
Nom:	SALICYLIC ACID		
Sub group:	ACH	AC-OH	AC-COOH
Nbr de subgr	4	1	1
	1	4	8

Figure 4.13 Capture d'écran Excel : mode de décomposition des molécules.

2. Les solubilités expérimentales des différents systèmes binaires avec les températures, sont introduites figure 4.14 et 4.15.

Temperature (K) T =		297,15	299,15	301,15	303,15	305,15
No	Composants	La fraction molaire dans différentes températures				
	ne choisissez pas "[n/a]"	↗	↖	↘	↙	↕
1	5-AMINOSALICYLIC ACID	0,1705	0,1876	0,2018	0,2154	0,2401
2	METHANOL	0,8295	0,8124	0,7982	0,7846	0,7599

Figure 4.14 Capture d'écran Excel : introduction des solubilités et des températures.

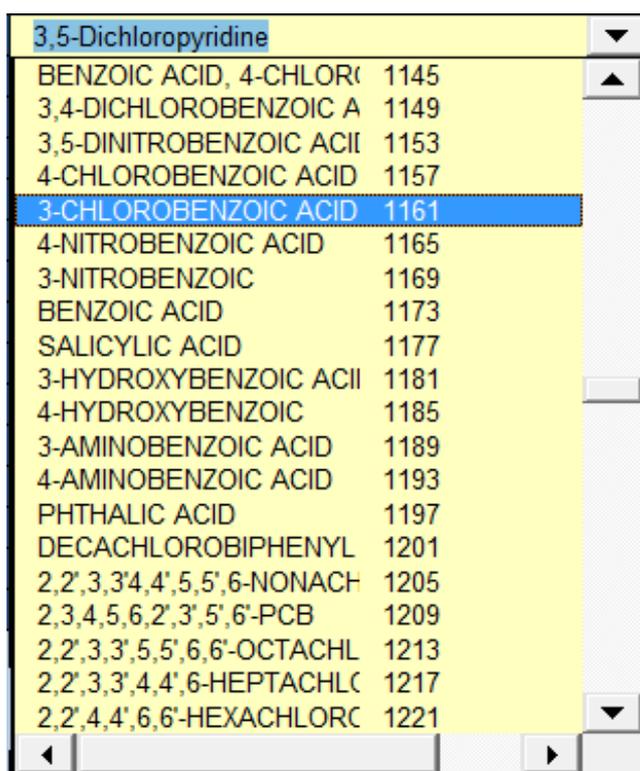


Figure 4.15 Capture d'écran : sélection du soluté et du solvant.

3. Dans le but d'estimer les paramètres d'interaction, on commence par une première estimation pour laquelle on attribue la valeur initiale de zéro à tous les paramètres d'interactions. Au cours de cette première étape, on utilise la fonction objective suivante dans l'itération k :

$$OF(x_{i,calc}^k, \gamma_{i,calc}^k) = \sum_{i=1}^n \sum_{j=1}^m [x_{i,exp} - x_{i,calc}^k]_j^2 + \sum_{i=1}^n \sum_{j=1}^m [\gamma_{i,exp} - \gamma_{i,calc}^k]_j^2 \quad (4.15)$$

Le but de cette estimation est de trouver les paramètres d'interaction initiaux pour la prochaine estimation.

4. Après l'arrêt de l'estimation à la valeur minimale de la fonction objective, les paramètres d'interaction obtenus sont utilisés comme paramètres initiaux dans la deuxième phase d'estimation. On utilise, au cours de cette dernière la méthode du Simplex de Nelder-Mead [175–177] où la fonction objective est la suivante :

$$OF(x_{i,calc}^k) = \sum_{i=1}^n \sum_{j=1}^m [x_{i,exp} - x_{i,calc}^k]_j^2 \quad (4.16)$$

5. Lorsque la fonction objective atteint la valeur la plus basse possible, elle est remplacée par la formule suivante :

$$OF(\gamma_{i,calc}^k) = \sum_{i=1}^n \sum_{j=1}^m [\gamma_{i,exp} - \gamma_{i,calc}^k]_j^2 \quad (4.17)$$

6. L'estimation est poursuivie, en changeant à chaque fois la fonction objective OF entre les trois formules mentionnées (4.15), (4.16) et (4.17), jusqu'à ce que la fonction se maintienne à la valeur la plus basse possible.
7. Au final, les paramètres d'interaction obtenus sont utilisés comme valeurs initiales dans la dernière estimation dans laquelle la somme des différences carrées de la solubilité est utilisée comme fonction objective.

Cette méthode est plus précise parce qu'on estime les valeurs initiales nécessaires dans l'estimation globale. Lorsqu'on se focalise, dans l'estimation, sur la somme des différences carrées de la solubilité, on va avoir une divergence significative sur les différences carrées du coefficient d'activité. Et si on fait l'estimation en utilisant la somme des différences carrées du coefficient d'activité, les différences carrées de la solubilité seront augmentées. On n'aura pas assez de précision.

La méthode d'estimation des paramètres d'interaction est très importante, car une fois obtenus, ces derniers permettent un calcul précis de la solubilité sans la nécessité de recourir à des résultats expérimentaux.

Une bonne initialisation est également très importante dans les méthodes d'optimisation car elle garantit l'obtention de bons résultats. L'organigramme suivant montre un aperçu général de la méthode d'estimation :

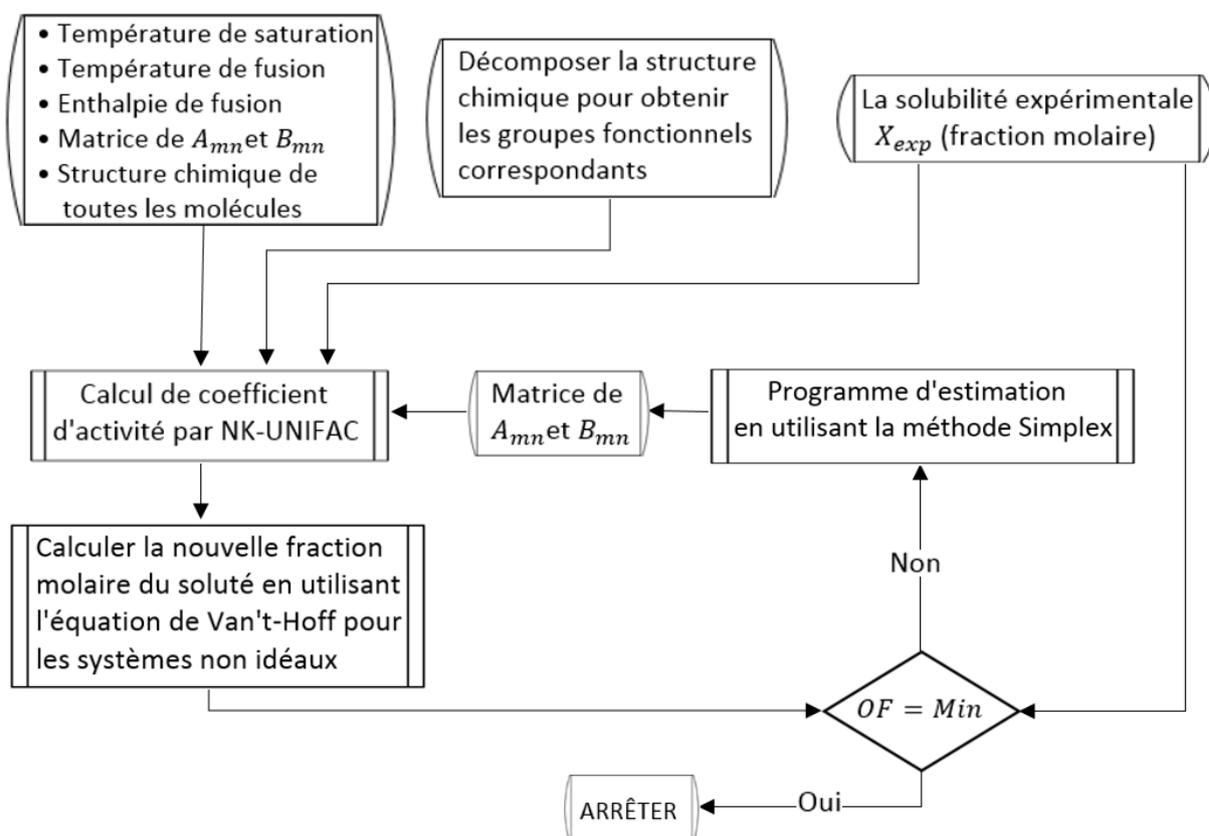


Figure 4.16 Organigramme de la méthode d'estimation des paramètres d'interaction par la méthode du Simplex de Nelder-Mead

OF est la fonction objective, elle peut prendre trois valeurs $OF(x)$, $OF(y)$ ou bien $OF(xy)$ comme indiqué dans les équations (4.15 – 4.17).

Les coefficients d'activité et les solubilités pour les systèmes binaires (soluté-solvant) sont calculés en même temps dans la même itération, les fonctions objectives des mélanges, la fonction objective globale et la somme de toutes les fonctions individuelles.

Excel et VBA sont utilisés pour créer plusieurs ensembles de classeurs et de feuilles. Chaque dossier contient un certain nombre de feuilles et chaque feuille renferme les solubilités à une ou plusieurs températures d'un système binaire (soluté-solvant) donné. Chaque feuille, appelée Sub_system possède son propre système de calcul associé aux fonctions dans VBA.

Tous les Sub_system(s) sont liés à un système fondamental, considéré comme classeur parent pour tous les autres classeurs et appelé Main_System. Ce dernier contient la décomposition de toutes les molécules, les paramètres de volume et de surface, la matrice des

paramètres d'interaction et le programme d'optimisation. La relation du dossier parent Main_System avec les Sub_system(s) est schématisée sur la Figure 4.17.

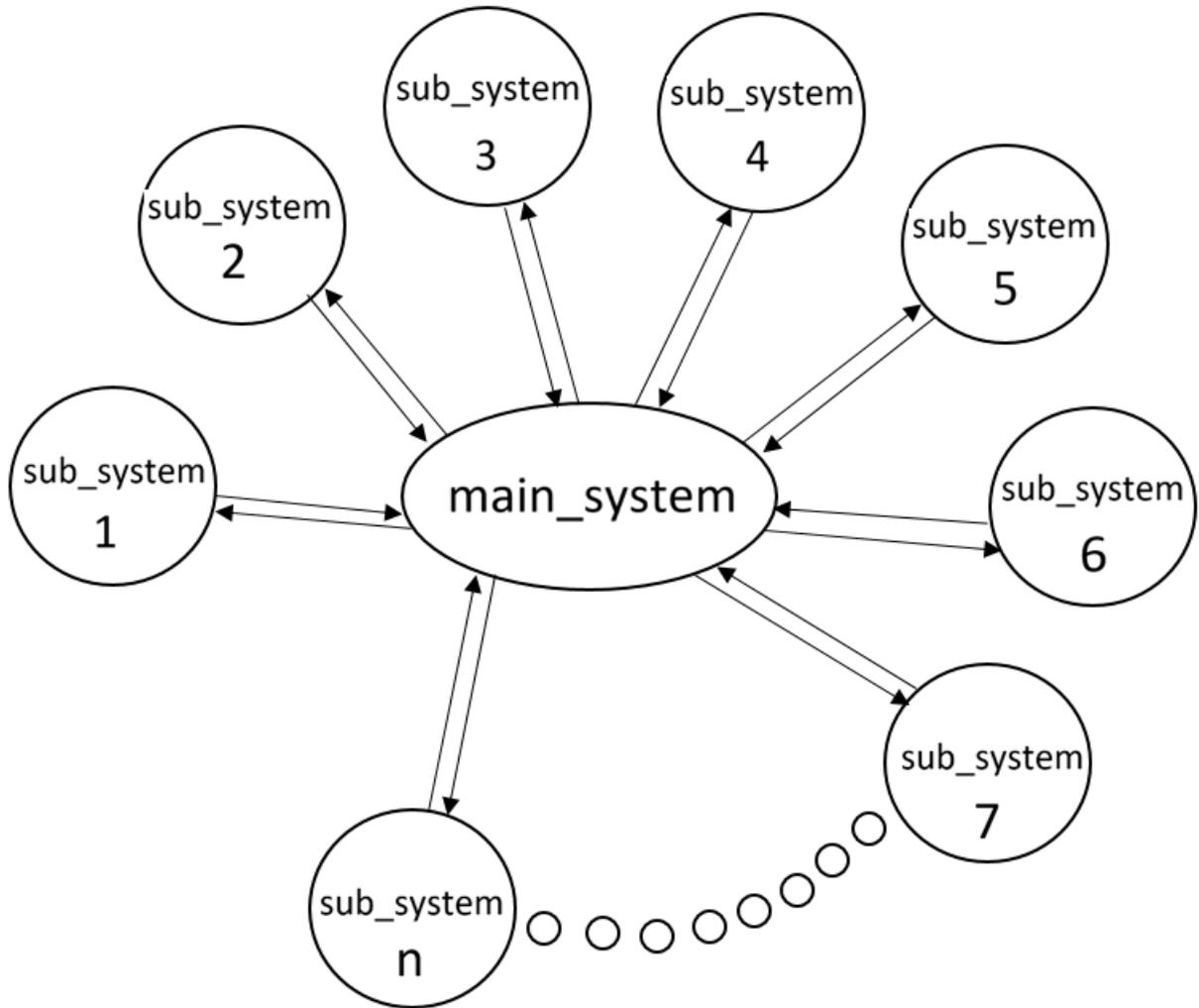


Figure 4.17 Schéma général de l'estimation globale : fonctionnement du Main_system avec les Sub_system(s)

Grâce au Main_system, il est possible de gérer, calculer, analyser les données et de voir les changements qui se produisent dans tous les sub_system(s) pendant l'estimation.

Les résultats finaux peuvent également être vus, la figure 4.18 illustre l'organigramme de calcul et la relation entre le Main_System et le Sub_system.

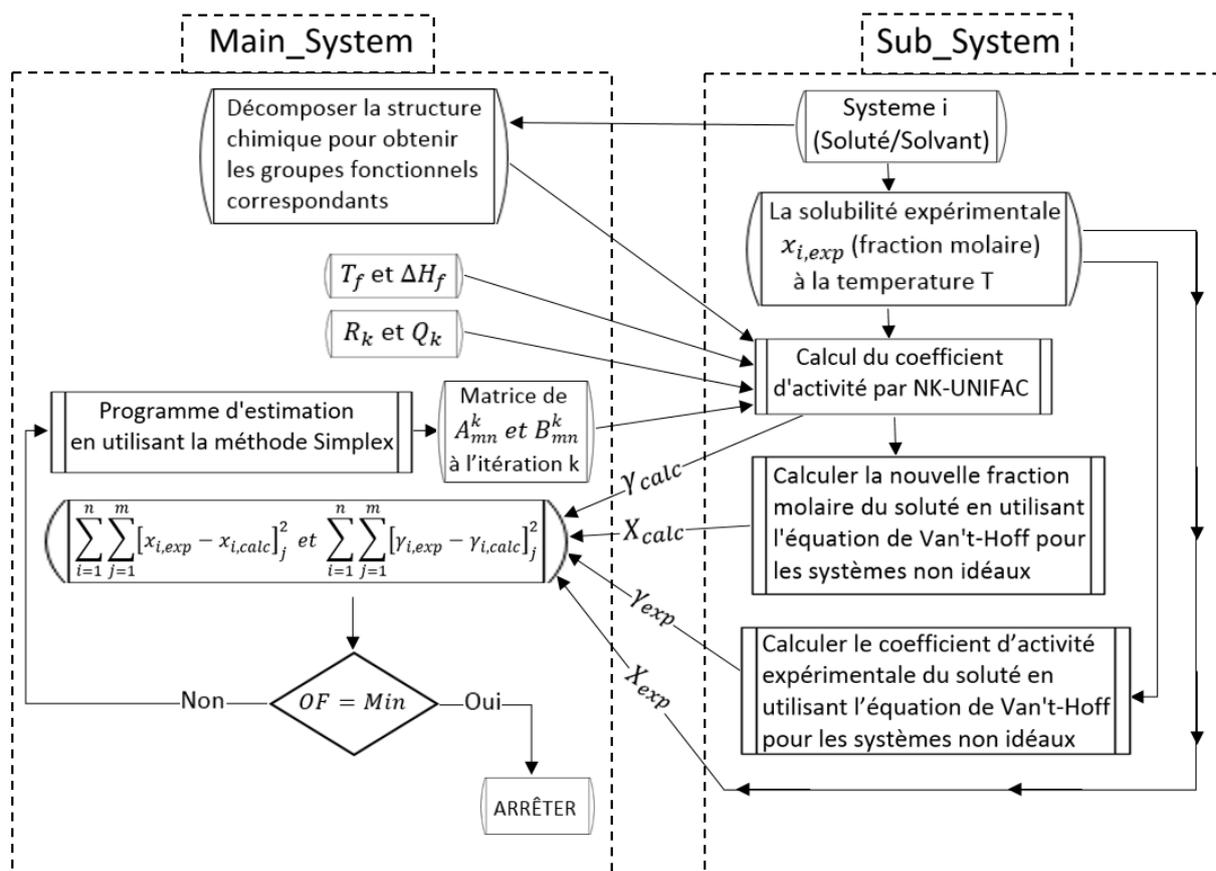


Figure 4.18 Organigramme détaillant fonctionnement entre le main_system et un sub_system.

Après l'achèvement de l'estimation et l'obtention de résultats satisfaisants pour les paramètres d'interaction requis, le programme NK-UNIFAC est prêt à calculer la solubilité des autres composés sans qu'il soit nécessaire de faire appel à des données expérimentales.

Afin de calculer la solubilité d'une molécule, la structure chimique du soluté et du solvant est décomposée pour obtenir les groupes fonctionnels correspondants.

Les fractions molaires X_0 sont initialisées, ensuite les fractions molaires X_1 sont calculées en utilisant les formules de NK-UNIFAC. La formule générale est :

$$X_i^{n+1} = f(X_i^n) \quad (4.18)$$

Une méthode d'optimisation est utilisée pour estimer la solubilité facilement. Le schéma de figure 4.19 montre d'une manière générale l'organigramme de prédiction de la solubilité.

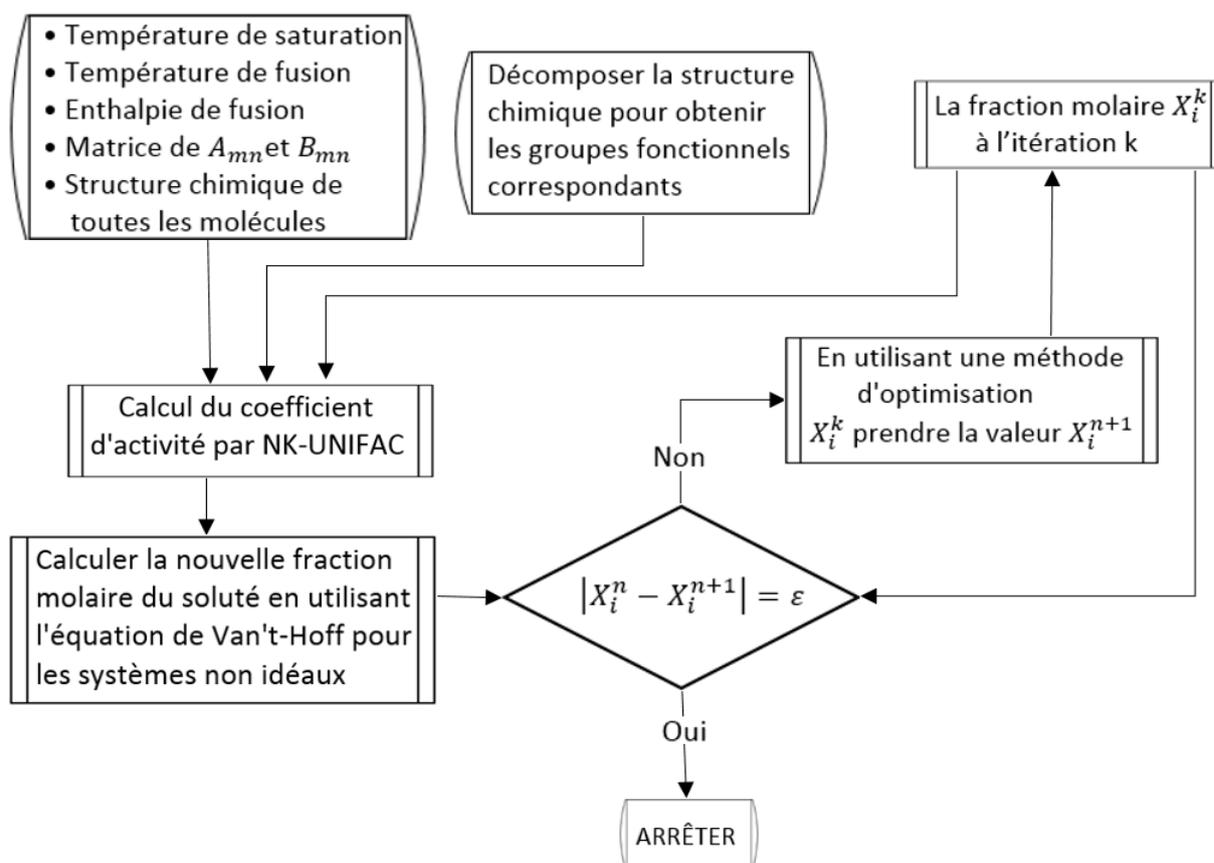


Figure 4.19 : Organigramme de la méthode de prédiction de la solubilité

4.6 Conclusion

Dans cette étude, une nouvelle méthode d'évaluation des paramètres UNIFAC est présentée. Cette nouvelle méthode a été déployée pour produire deux ensembles de paramètres d'interaction groupe-groupe a_{mn} et b_{mn} basés sur les définitions de KT-UNIFAC [33] et UNIFAC Pharma modifié [34]. La dépendance de la température des paramètres est représentée par deux termes (comme dans UNIFAC Pharma Modifié), et la décomposition de la structure chimique des molécules et les paramètres de volume et de surface des groupes fonctionnels sont les mêmes que ceux utilisés dans KT-UNIFAC [135].

Tous les programmes de calcul ont été écrits en Visual Basic (VBA). L'utilisation de VBA avec l'interface Excel a facilité la saisie des données. Elle a permis, en outre, de suivre l'estimation et les calculs, de gérer les diverses tâches et d'analyser les résultats et les données de manière systématique et rapide.

La validation du modèle est probablement l'étape la plus importante dans le développement des outils de prédiction fiables. Sans une validation approfondie, l'exactitude et la robustesse de la prédiction des nouveaux composés ne peuvent être jugées [178].

Prédiction de la solubilité des substances pharmaceutique active par le nouveau modèle

Sommaire :

5.1	Introduction	89
5.2	Travail réalisé.....	89
5.3	Résultats de l'estimation	90
5.4	Prédiction de la solubilité.....	95
5.5	Validation des paramètres d'interaction.....	97
5.6	Comparaison avec le modèle UNIFAC.....	97
5.7	Evaluation du modèle.....	103
5.8	Conclusion.....	109

5.1 Introduction

Dans ce chapitre, on présente les résultats obtenus dans cette étude. On donne, dans un premier temps, les résultats de l'estimation incluant 250 systèmes binaires (soluté/ solvant). Cette estimation a permis de générer 228 paramètres d'interaction pour 13 groupes fonctionnels. On donne la matrice correspondante.

Ces paramètres ont ensuite été validés en utilisant 88 systèmes binaires non inclus dans l'étape de régression des paramètres. Les solubilités prédites par NK-UNIFAC puis comparées aux solubilités expérimentales montrent que les résultats sont très satisfaisants [179].

Par ailleurs, le modèle UNIFAC original a été utilisé pour prédire les solubilités de certains systèmes binaires qui contiennent des groupes définis dans ce modèle. Les résultats obtenus ont été comparés à ceux du nouveau modèle NK-UNIFAC et aux données expérimentales.

Il ressort de cette comparaison, à travers les représentations graphiques de la solubilité, que le nouveau modèle donne des résultats plus proches aux résultats expérimentaux que ceux du modèle UNIFAC original.

5.2 Travail réalisé

Les données de solubilité de plus de 400 systèmes binaires (soluté-solvant), dans un intervalle de température [297,15 - 315,15 K], disponibles dans la littérature, sont considérées pour estimer les paramètres d'interaction de 20 groupes fonctionnels. Parmi ces groupes, les résultats obtenus pour 13 groupes fonctionnels ont été validés.

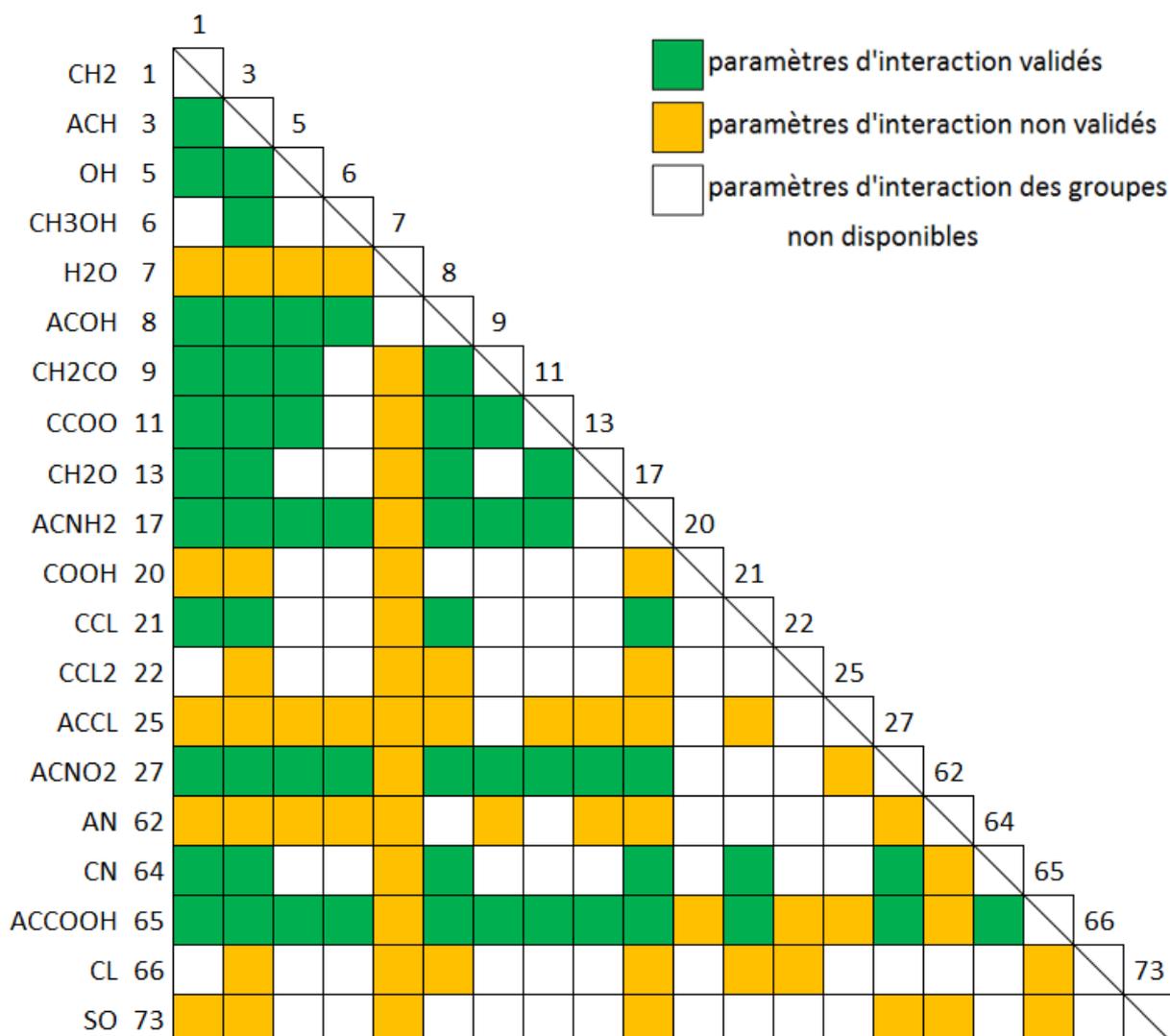


Figure 5.1 Matrice des paramètres d'interaction estimés pour 20 groupes fonctionnels

5.3 Résultats de l'estimation

L'estimation a été réalisée en plusieurs étapes. A chaque étape, un certain nombre de systèmes binaires est inclus et les paramètres d'interaction de groupes non disponibles dans la matrice sont estimés. Le tableau suivant montre la liste des groupes, munis de leurs numéros respectifs dans la matrice, dont les paramètres ont été estimés.

Tableau 5.1 Groupes fonctionnels

1	CH ₂
3	ACH
5	OH
6	CH ₃ OH
8	ACOH
9	CH ₂ CO
11	CCOO
13	CH ₂ O
17	ACNH ₂
21	CCI
27	ACNO ₂
64	CN
65	ACCOOH

Les paramètres d'interaction a_{mn} et b_{mn} estimés dans ce travail sont montrés dans le Tableau 5.2

Tableau 5.2 Paramètres d'interaction a_{mn} et b_{mn} estimés

m	n	a_{mn}	b_{mn}
1	3	457,5662	1,6567
1	5	1041,8026	-1,1869
1	8	-174,1714	1,8452
1	9	1032,8741	$-5,667 \cdot 10^{-1}$
1	11	353,9085	-1,1276
1	13	222,5837	2,9078
1	17	453,0707	-14,1871
1	21	663,6319	$1,191 \cdot 10^{-1}$
1	27	-1672,2371	1,7946
1	64	-1229,4124	$5,591 \cdot 10^{-1}$
1	65	-1942,5303	6,3423

3	1	-231,5086	$-1,307 \cdot 10^{-1}$
3	5	-3794,5739	12,7646
3	6	18,7680	$2,838 \cdot 10^{-1}$
3	8	-614,7082	-1,4908
3	9	-288,1830	$5,100 \cdot 10^{-1}$
3	11	191,8475	$-2,665 \cdot 10^{-2}$
3	13	-261,7626	$6,407 \cdot 10^{-2}$
3	17	-1433,7894	-1,8884
3	21	2545,0580	$2,506 \cdot 10^{-4}$
3	27	796,8114	-11,2261
3	64	2898,6508	$2,963 \cdot 10^{-2}$
3	65	97,0114	2,9384
5	1	1908,1081	-6,6490
5	3	322,7674	$-6,894 \cdot 10^{-2}$
5	8	45,8229	-11,5628
5	9	-36,9993	$3,301 \cdot 10^{-1}$
5	11	1,8820	$-3,769 \cdot 10^{-1}$
5	17	20180,8799	-83,6405
5	27	465,3309	6,6972
5	65	-1352,1542	4,4841
6	3	125,0602	$-4,946 \cdot 10^{-2}$
6	8	-13,7722	-13,9303
6	17	-8211,7757	9,2592
6	27	1047,5594	$6,411 \cdot 10^{-2}$
6	65	-3,0344	$2,765 \cdot 10^{-1}$
8	1	421,4888	-3,5493
8	3	-1,5379	-2,1626
8	5	-191,0139	3,5574

8	6	-20,0612	4,8933
8	9	-569,7766	$-1,532 \cdot 10^{-1}$
8	11	-288,5179	$-3,217 \cdot 10^{-1}$
8	13	-119,9539	$1,026 \cdot 10^{-1}$
8	21	-3,0187	-1,7304
8	27	-42,8325	$3,303 \cdot 10^{-1}$
8	64	44,1762	$8,193 \cdot 10^{-1}$
8	65	-22,3275	-1,4808
9	1	21,1064	$4,969 \cdot 10^{-2}$
9	3	691,7017	$-8,671 \cdot 10^{-1}$
9	5	521,2705	$-7,058 \cdot 10^{-1}$
9	8	$-5,066 \cdot 10^{-8}$	$7,347 \cdot 10^{-8}$
9	11	-369,8708	$-3,821 \cdot 10^{-1}$
9	17	$1,755 \cdot 10^{-5}$	$1,850 \cdot 10^{-5}$
9	27	$-1,688 \cdot 10^{-5}$	$-3,292 \cdot 10^{-5}$
9	65	182,6055	$-5,445 \cdot 10^{-1}$
11	1	146,1972	$-4,557 \cdot 10^{-2}$
11	3	42,1931	1,1088
11	5	217,7224	$1,177 \cdot 10^{-1}$
11	8	-2,5680	$3,257 \cdot 10^{-1}$
11	9	-14,0754	$-2,242 \cdot 10^{-1}$
11	13	$-6,751 \cdot 10^{-1}$	$7,071 \cdot 10^{-1}$
11	17	$2,177 \cdot 10^{-6}$	$1,350 \cdot 10^{-6}$
11	27	$-1,728 \cdot 10^{-4}$	$-3,145 \cdot 10^{-4}$
11	65	-10,8871	$-1,829 \cdot 10^{-1}$
13	1	36,7950	-1,6664
13	3	4208,3839	$1,504 \cdot 10^{-2}$
13	8	$-3,275 \cdot 10^{-1}$	$-1,716 \cdot 10^{-1}$

13	11	$-7,003 \cdot 10^{-1}$	$3,193 \cdot 10^{-1}$
13	27	$-7,058 \cdot 10^{-3}$	$-3,099 \cdot 10^{-3}$
13	65	-31,0566	$3,643 \cdot 10^{-1}$
17	1	9195,2604	0,0000
17	3	-599,3055	$1,088 \cdot 10^{-1}$
17	5	3405,8585	1,0115
17	6	-8,0286	8,1204
17	9	-1,4168	$-1,889 \cdot 10^{-1}$
17	11	1,4876	$-1,783 \cdot 10^{-1}$
17	21	-591,6884	$-3,839 \cdot 10^{-1}$
17	27	-314,0500	$1,185 \cdot 10^{-1}$
17	64	-1144,4302	$-4,957 \cdot 10^{-1}$
17	65	-381,2320	-5,9848
21	1	-277,6494	$-1,027 \cdot 10^{-1}$
21	3	399,7085	$-3,695 \cdot 10^{-1}$
21	8	$-3,403 \cdot 10^{-2}$	$-2,111 \cdot 10^{-1}$
21	17	-41,5941	$-1,383 \cdot 10^{-2}$
21	65	516,1970	$-9,946 \cdot 10^{-2}$
27	1	-455,2281	$5,476 \cdot 10^{-1}$
27	3	1981,2756	$-1,247 \cdot 10^{-1}$
27	5	284,7005	-2,4388
27	6	934,1809	-4,2080
27	8	-84,0124	-16,1549
27	9	76,5343	-1,8805
27	11	-12,9711	-1,3555
27	13	-3,9727	-1,3455
27	17	-469,5328	-14,6240
27	64	-1330,9924	$5,199 \cdot 10^{-1}$

27	65	-359,1374	$5,711 \cdot 10^{-2}$
64	1	777,6103	$-1,837 \cdot 10^{-1}$
64	3	6475,9249	$-3,294 \cdot 10^{-7}$
64	8	$-1,606 \cdot 10^{-7}$	$-9,494 \cdot 10^{-7}$
64	17	$1,388 \cdot 10^{-6}$	$8,497 \cdot 10^{-6}$
64	27	-3,0716	$1,171 \cdot 10^{-1}$
64	65	29,1353	$-6,596 \cdot 10^{-2}$
65	1	-551,2107	6,6051
65	3	-198,7469	$6,309 \cdot 10^{-1}$
65	5	-142,2851	15,4378
65	6	10,4286	$-1,846 \cdot 10^{-1}$
65	8	-244,7429	-10,1326
65	9	1283,5688	$2,006 \cdot 10^{-2}$
65	11	36,8238	$-7,078 \cdot 10^{-1}$
65	13	116,2035	$6,053 \cdot 10^{-2}$
65	17	197,1920	1,5004
65	21	-546,6243	$4,930 \cdot 10^{-1}$
65	27	-160,9969	4,5444
65	64	2646,7420	$3,092 \cdot 10^{-2}$

5.4 Prédiction de la solubilité

Les solubilités corrélées pour les composés inclus dans l'estimation sont comparées aux solubilités expérimentales. Les deux représentations graphiques, données en Figures 5.2 et 5.3, montrent deux types de systèmes binaires contenant différents groupes.

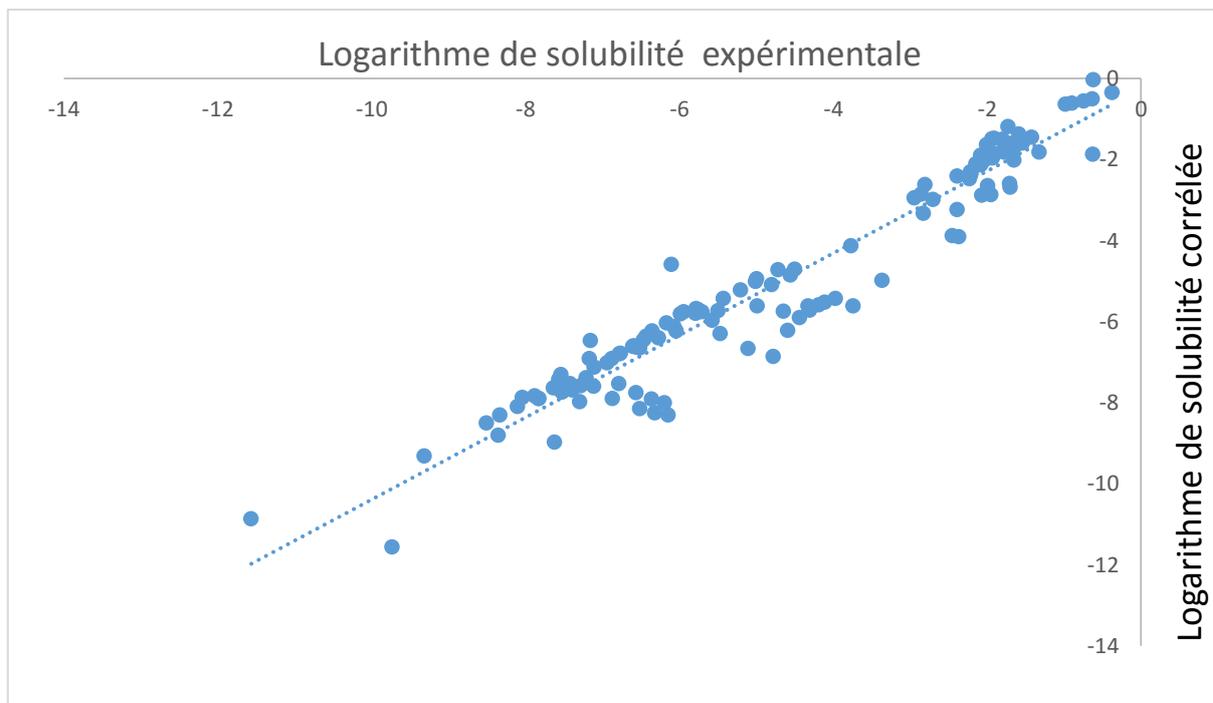


Figure 5.2 Représentation graphique des solubilités corrélées en fonction des solubilités expérimentales pour des systèmes binaires utilisés dans l'estimation des groupes ; CH₂, ACH, OH, CH₃OH, ACOH, ACNH₂, ACNO₂, ACCOOH

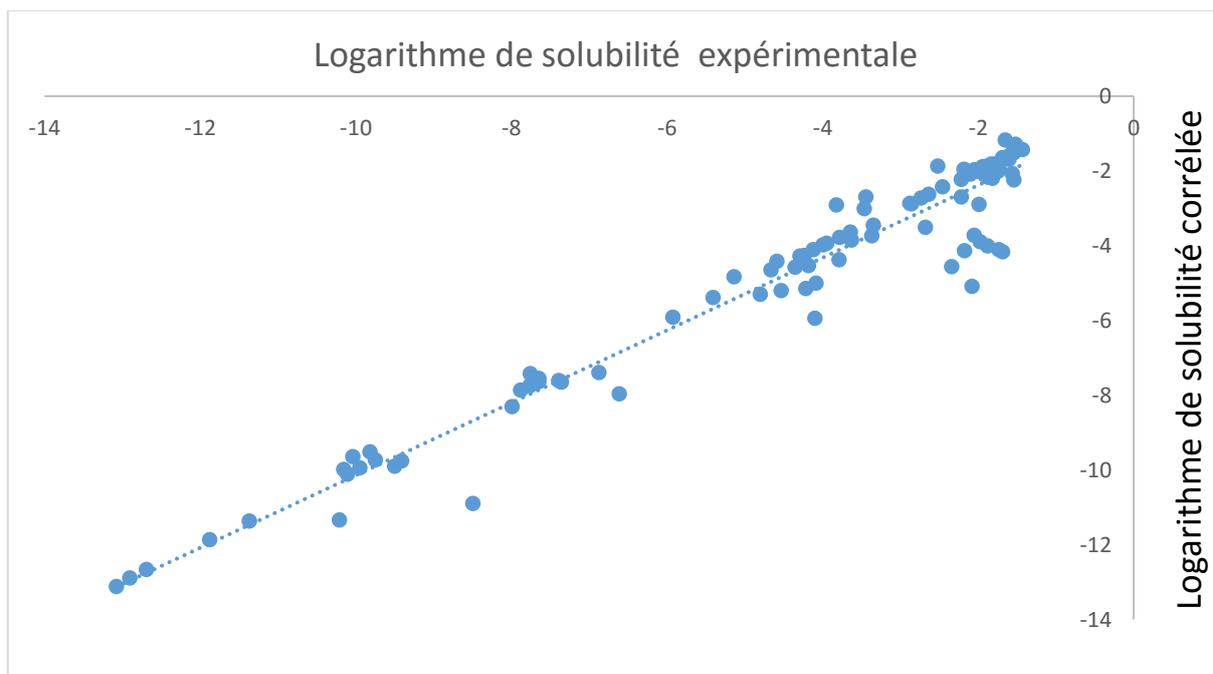


Figure 5.3 Représentation graphique des solubilités corrélées en fonction des solubilités expérimentales pour des systèmes binaires utilisés dans l'estimation des groupes ; CH₂CO, CCOO, CH₂O, CCl, CN

5.5 Validation des paramètres d'interaction

Pour la validation de ces paramètres d'interaction, environ 90 systèmes binaires soluté/solvant ont été considérés dans la prédiction de la solubilité et en les comparant avec des données expérimentales, on a trouvé les résultats illustrés dans la Figure 4.3

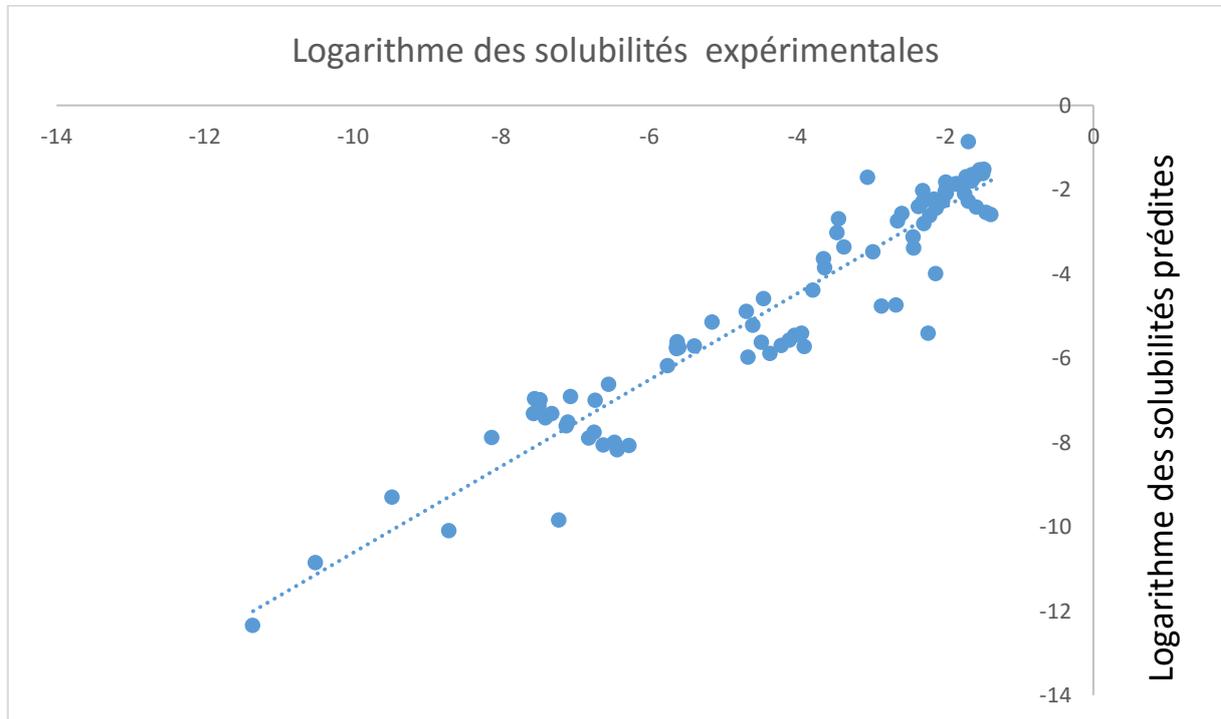


Figure 5.4 : représentation graphique des solubilités prédites pour des systèmes non utilisés dans l'estimation en fonction des solubilités expérimentales.

5.6 Comparaison avec le modèle UNIFAC

Dans cette étude, nous avons limité la comparaison entre le nouveau modèle NK-UNIFAC et UNIFAC original aux systèmes pour lesquels des paramètres d'interaction sont disponibles dans ce dernier.

Les trois premières figures 5.4, 5.5 et 5.6 montrent que la solubilité corrélée avec la nouvelle approche du modèle UNIFAC donne de meilleurs résultats que le modèle UNIFAC original. Les figures suivantes de 5.7 jusqu'à 5.13 montrent la solubilité des systèmes non utilisés dans l'estimation des paramètres d'interaction, confirment le rapprochement entre les données de solubilité prédites et expérimentales.

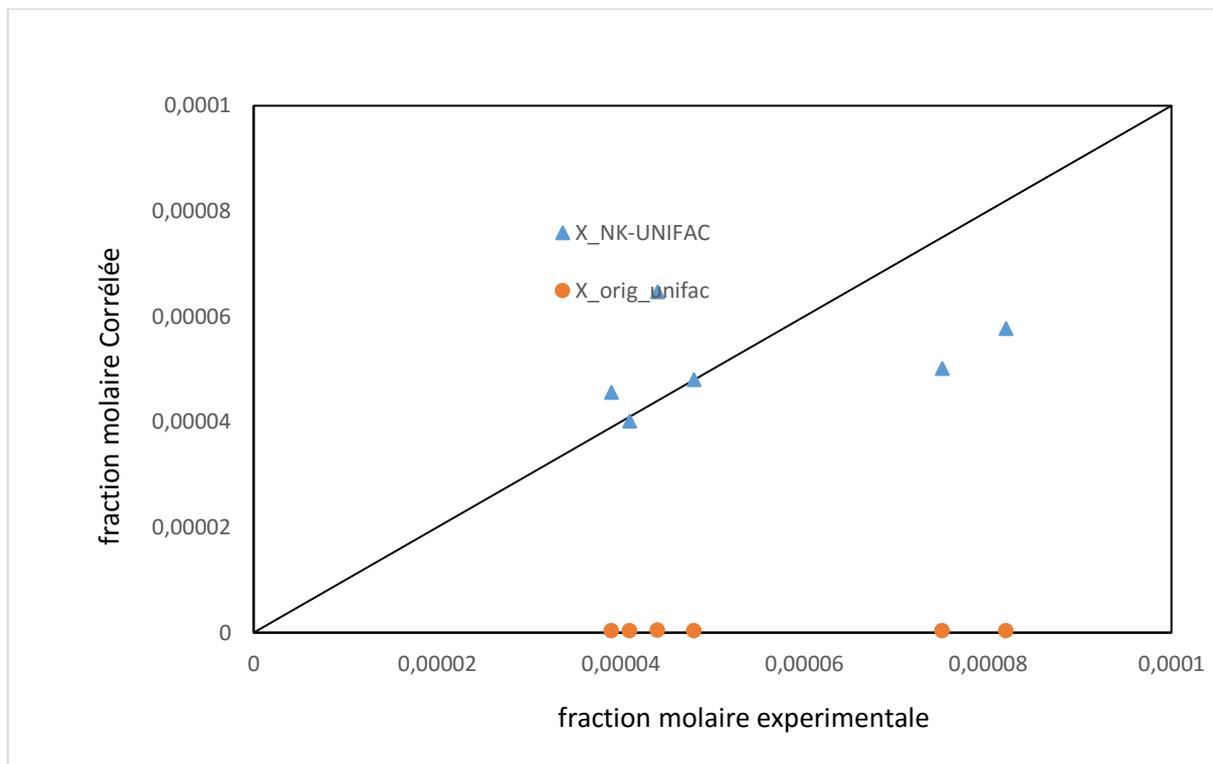


Figure 5.5 Solubilité corrélée en fraction molaire de l'acide picrique dans l'hexane, le cyclohexane, le dodécane, le décane, l'heptane et l'isooctane à 298,15 K

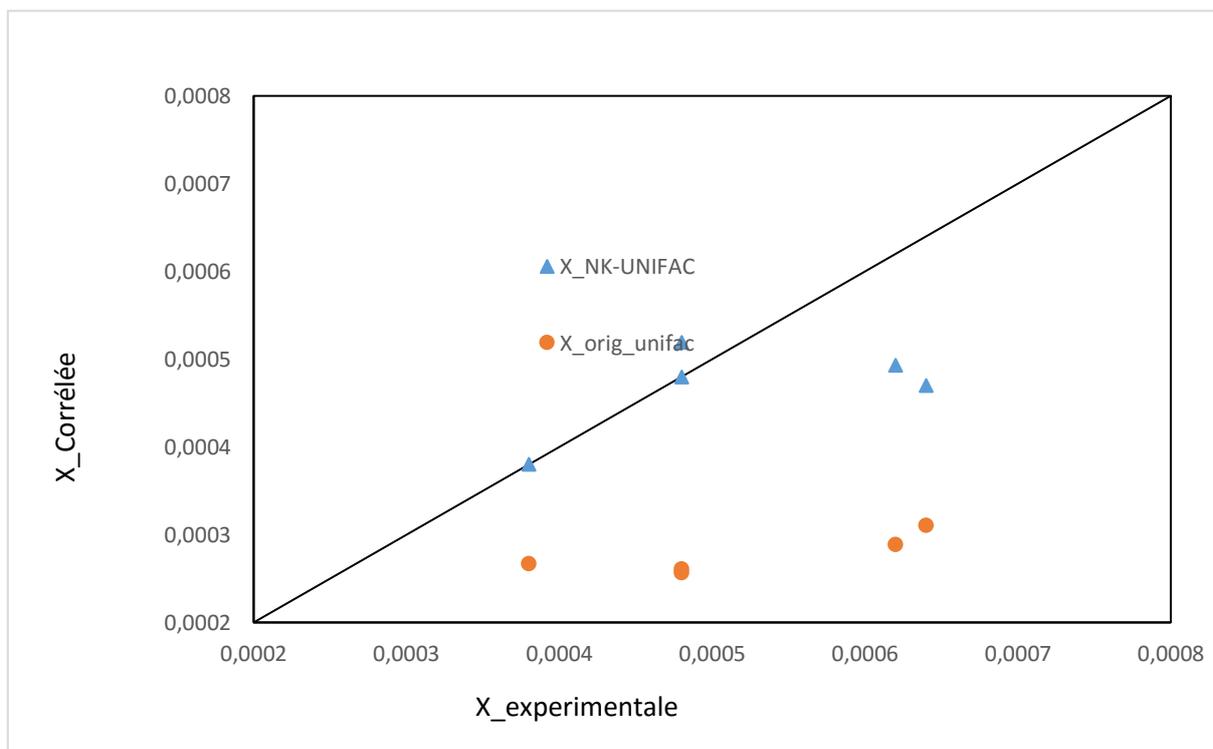


Figure 5.6 Solubilité prédite en fraction molaire de l'acide salicylique dans l'hexane, le dodécane, le décane, l'heptane et l'isooctane à 298,15 K

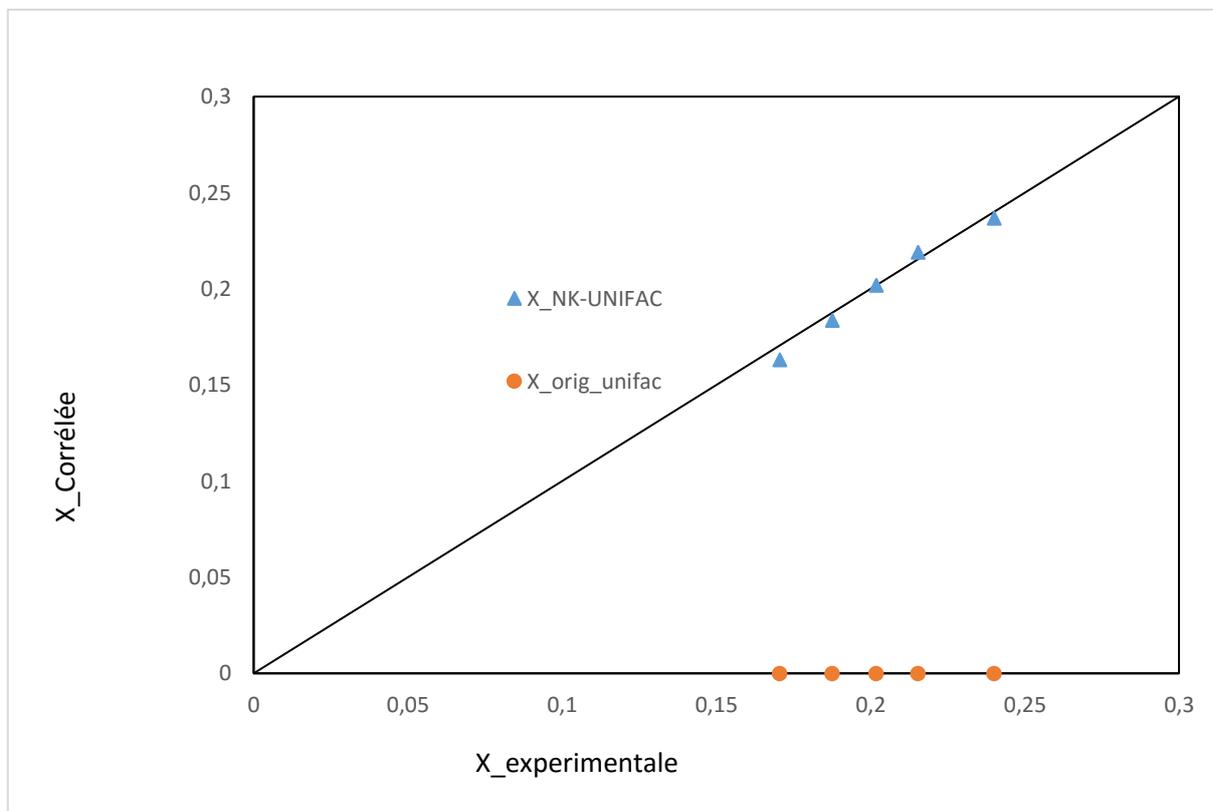


Figure 5.7 Solubilité corrélée en fraction molaire de l'acide 5-amino-salicylique dans le méthanol à différentes températures 297,15, 299,15, 301,15, 303,15, 305,15 K

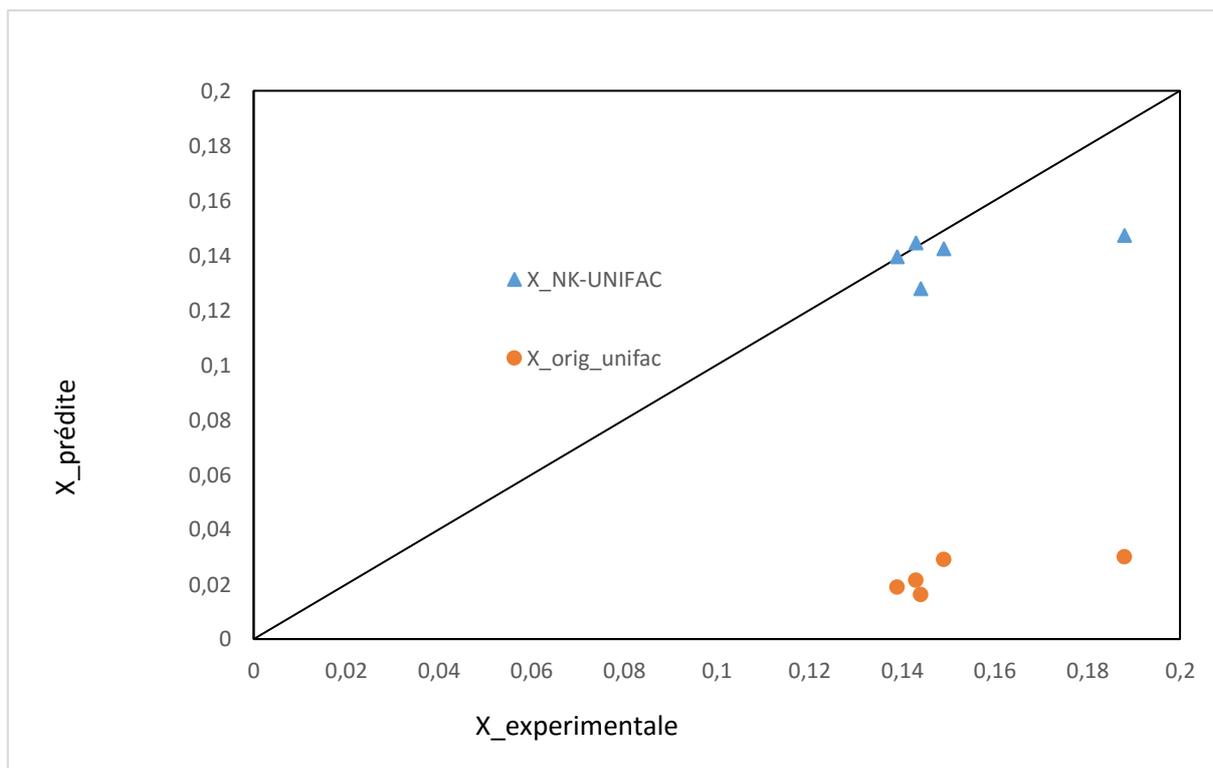


Figure 5.8 Solubilité prédite en fraction molaire de l'acide 3-nitrobenzoïque dans le 1-propanol, le 1-pentanol, le 2-butanol, le 1-hexanol et le 1-octanol à 298,15 K

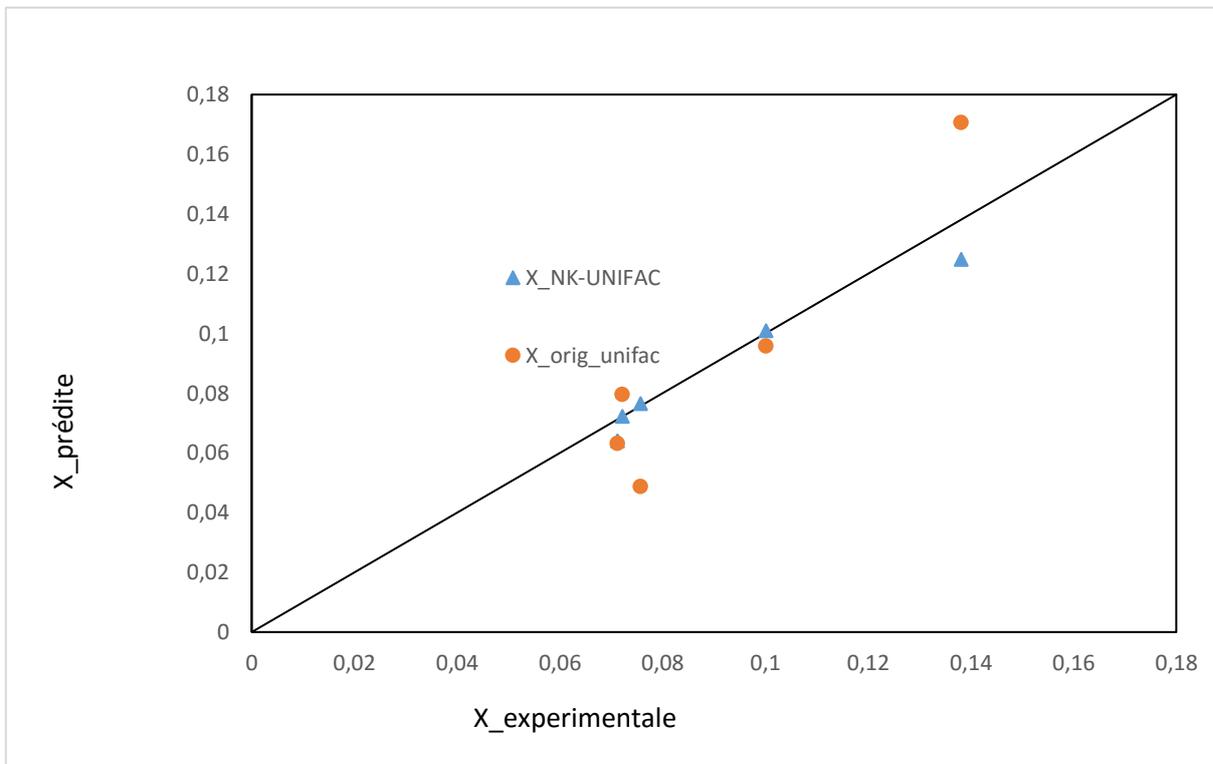


Figure 5.9 Solubilité prédite en fraction molaire du biphenyle dans l'heptane, le 1-pentanol, le 1-heptanol, le 2-méthyl-1-pentanol et le 2-méthyl-2-butanol à 298,15 K

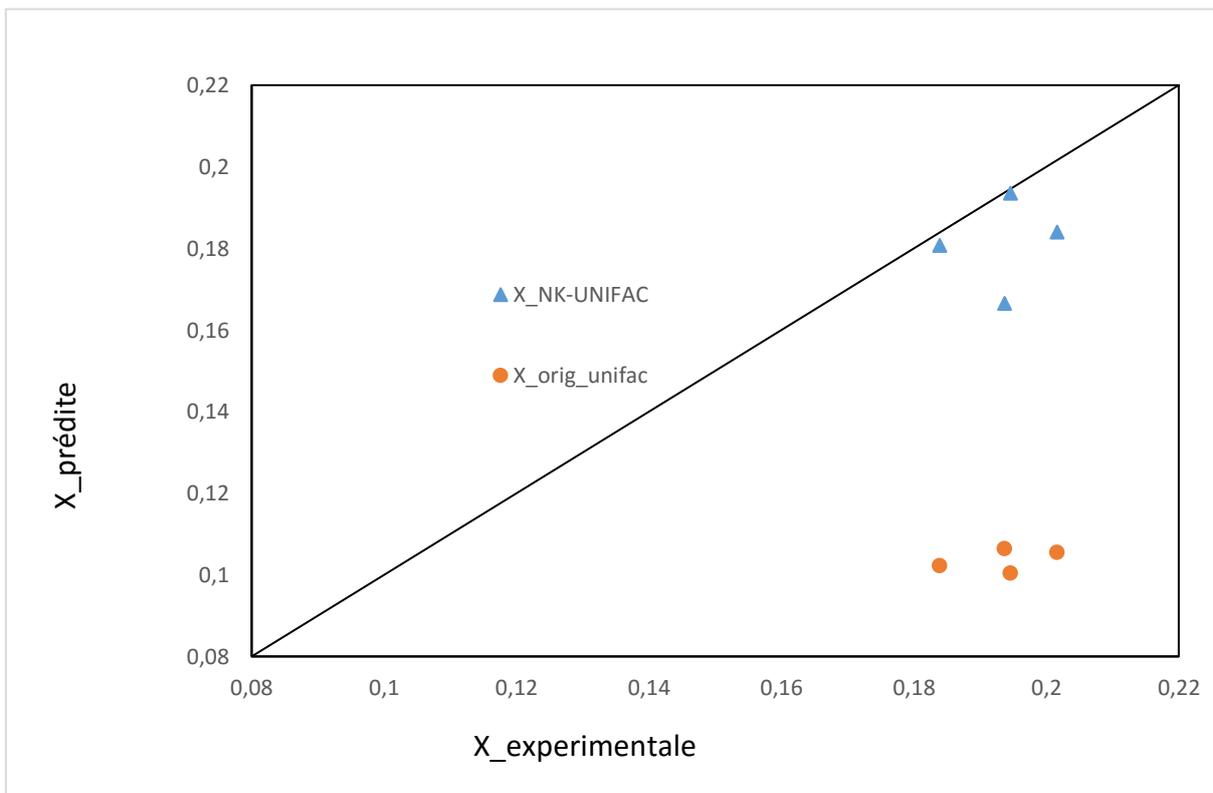


Figure 5.10 Solubilité prédite en fraction molaire de l'acide benzoïque dans le 2-propanol, le 1-butanol, le 1-pentanol et le 1-heptanol à 298,15 K

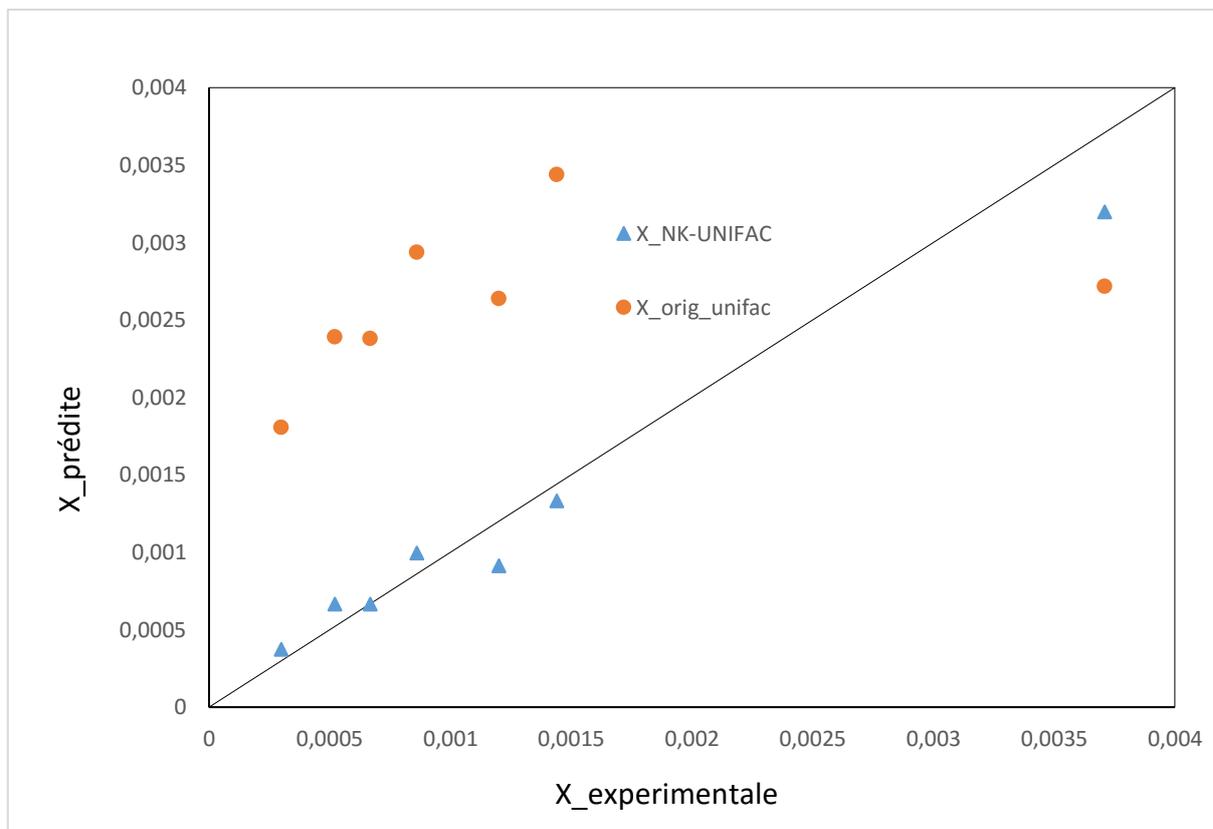


Figure 5.11 Solubilité prédite en fraction molaire du hexachlorobenzène dans le 2-butanol, le 2-bentanol, 2-propanol, 1-butanol, 2-methyl-2-butanol, 1-hexanol, octane à 298,15 K

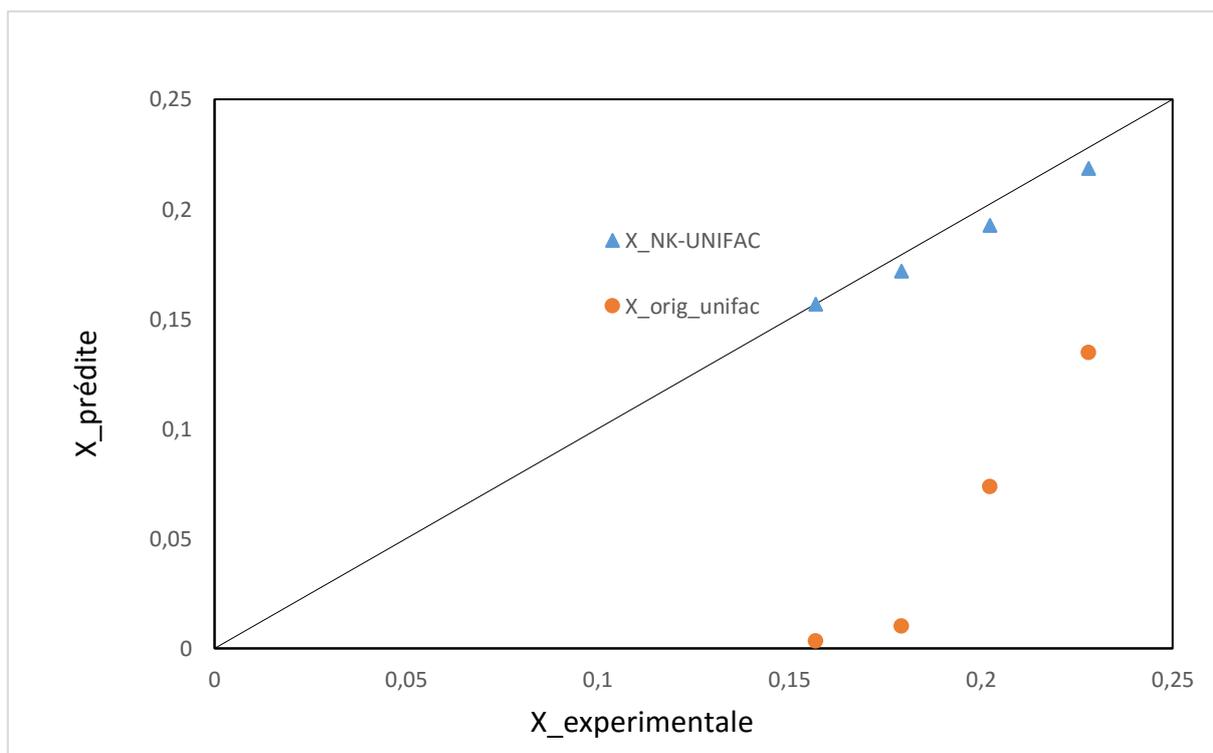


Figure 5.12 Solubilité prédite en fraction molaire de l'acide salicylique dans l'acétone à différentes températures 288,15, 298,15, 308,15, 318,15 K

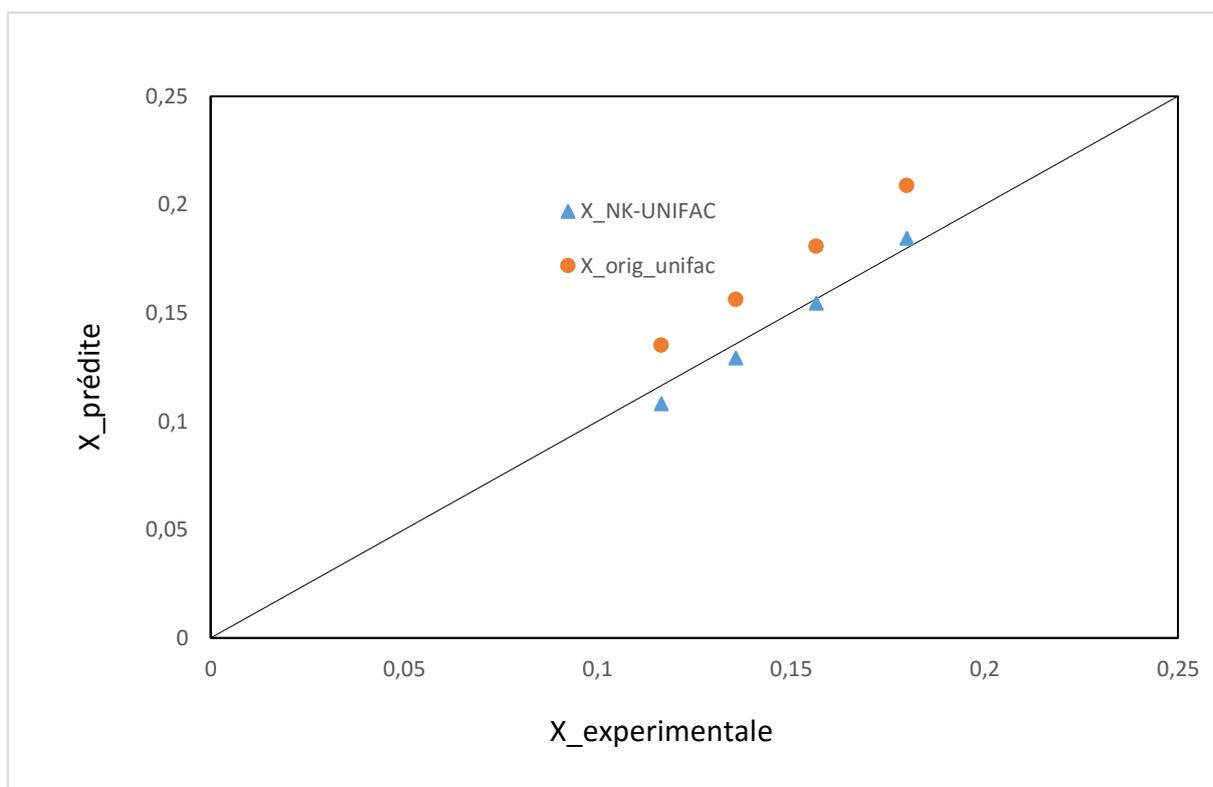


Figure 5.13 Solubilité prédite en fraction molaire de l'acide salicylique dans l'éthyle acétate à différentes températures 288,15, 298,15, 308,15, 318,15 K

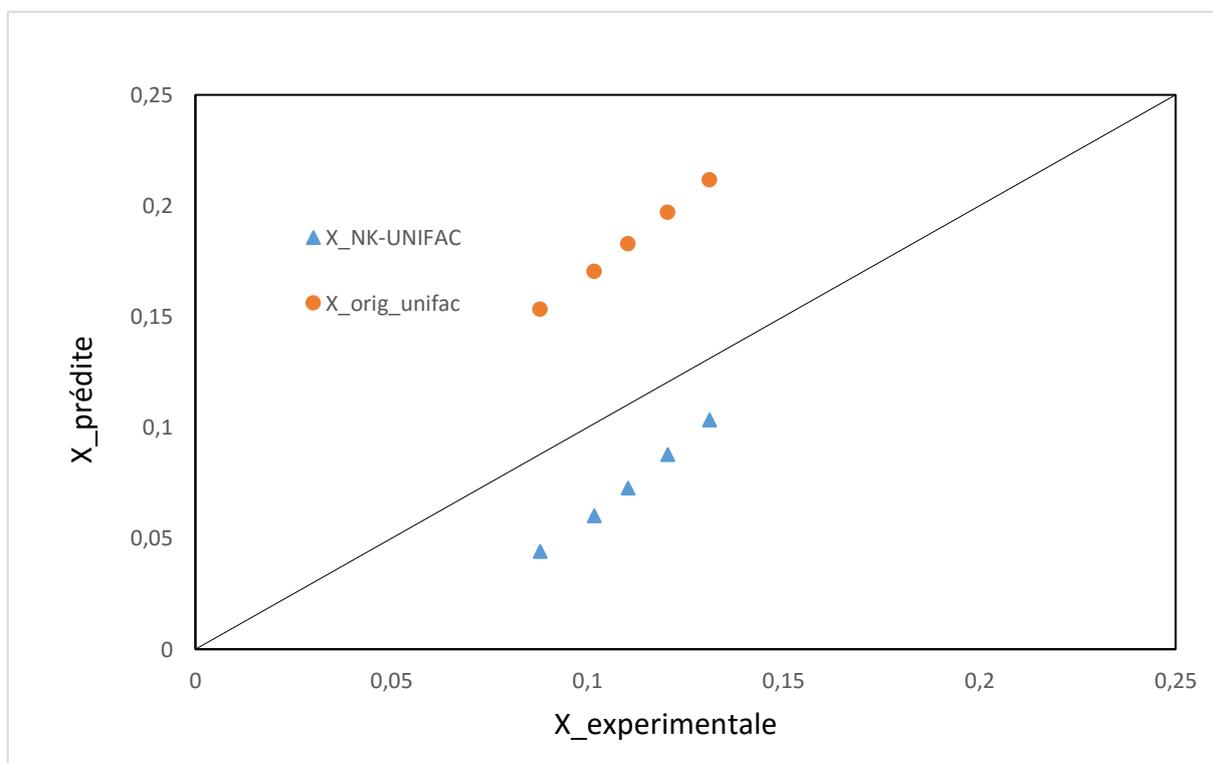


Figure 5.14 Solubilité prédite en fraction molaire de l'hydroquinone dans l'éthyle acétate à différentes températures 290,15, 298,15, 304,15, 310,15 316,15 K

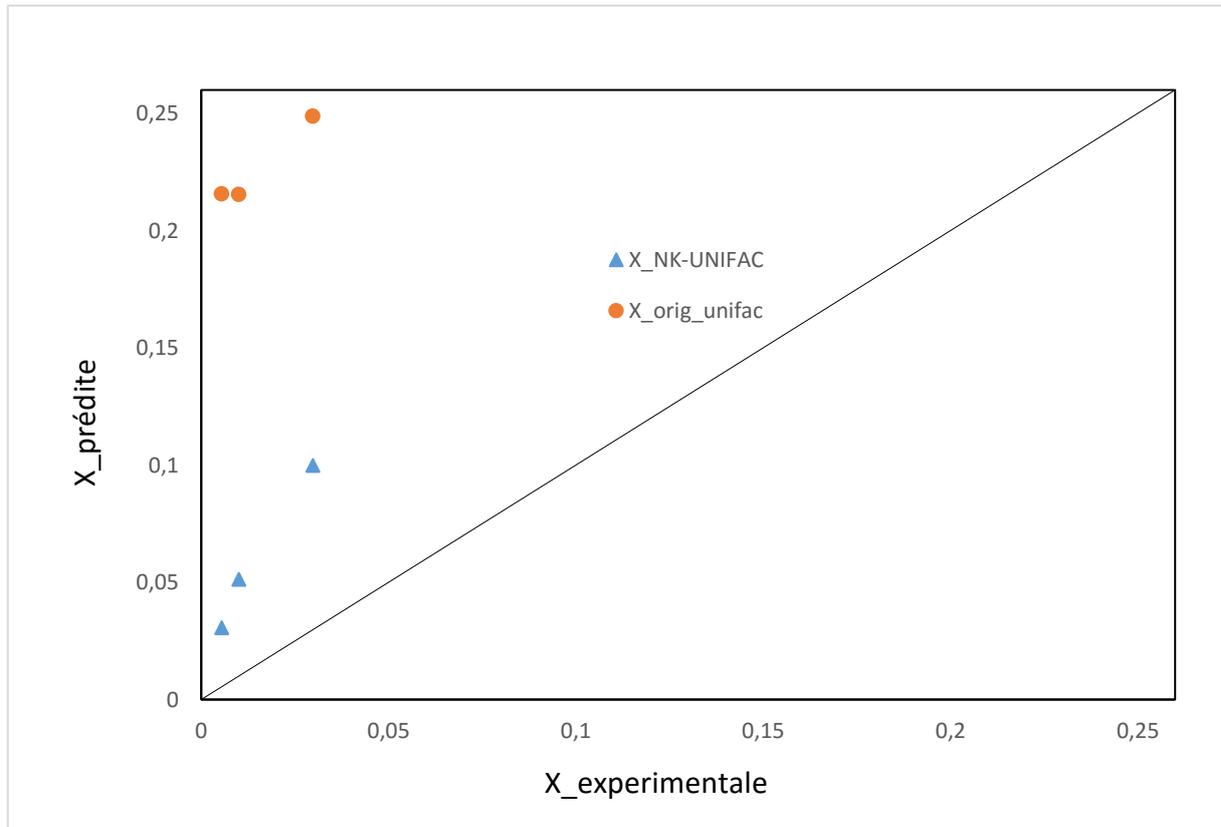


Figure 5.15 Solubilité prédite en fraction molaire du methyl palmitate dans l'acétone à différentes températures 290,15, 298,15, 304,15, 310,15 316,15 K

5.7 Evaluation du modèle

Les déviations entre les solubilités prédites et les solubilités expérimentales ont été calculées pour le nouveau modèle NK-UNIFAC et le modèle UNIFAC original en utilisant la relation suivante :

$$ARD = \frac{1}{N} \sum \left| \frac{(X_{exp} - X_{prédite}) \times 100}{X_{exp}} \right| \quad (5.1)$$

Le tableau 5.3 montre quelques déviations calculées pour les deux modèles.

Tableau 5.3 Déviations relatives entre solubilités prédites et expérimentales pour les modèles NK-UNIFAC et UNIFAC original

Systeme (Soluté / Solvant)	T (K)	NK-UNIFAC ARD (%)	UNIFAC ARD (%)
4-Chlorobenzoic acid / Dibutyl ether	298,15	0,03	85,06
4-Hydroxybenzoic acid / 4-Methyl-2-Pentanone	298,15	16,69	79,46
4-Chlorobenzoic acid / Butyl acetate	298,15	19,75	77,74
Salicylic acid / Butyl acetate	298,15	24,33	23,11
Benzoic acid / Triacetin	298,15	8,32	1,86
Benzoic acid / N-Butyl lactate	298,15	24,36	53,67
Benzoic acid / MTBE	298,15	14,80	58,57
3-Nitrobenzoic / Dibutyl ether	298,2	0,02	97,87
3-Nitrobenzoic / Ethyl acetate	298,2	0,01	86,04
Hexachlorobenzene / 2-Propanol	298,15	25,92	506,58
Hexachlorobenzene / 1-Butanol	298,15	0,03	256,88
Hexachlorobenzene / 2-Methyl-2-Butanol	298,15	24,04	119,87
Hexachlorobenzene / 2-Butanol	298,15	28,45	359,19
Hexachlorobenzene / 1-Hexanol	298,15	7,52	138,79
Hexachlorobenzene / Octane	298,15	13,87	26,71
Hexachlorobenzene / Butyl acetate	298,15	0,03	548,13
Hexachlorobenzene / Decane	298,15	28,62	39,68
Hexachlorobenzene / MTBE	298,15	35,24	188,85
Hexachlorobenzene / 2-Pentanol	298,15	15,98	241,77
Salicylic acid / Acetone	288,15	0,03	97,97
Salicylic acid / Acetone	298,15	2,82	94,43
Salicylic acid / Acetone	308,15	3,41	63,68
Salicylic acid / Acetone	318,15	3,03	41,00
Salicylic acid / Ethyl acetate	288,15	10,37	16,11
Salicylic acid / Ethyl acetate	298,15	6,10	15,25
Salicylic acid / Ethyl acetate	308,15	1,48	15,48
Salicylic acid / Ethyl acetate	318,15	2,79	16,10

Hydroquinone / Ethyl acetate	298,8	50,40	72,71
Hydroquinone / Ethyl acetate	304,6	36,00	69,01
Hydroquinone / Ethyl acetate	310,8	25,42	64,87
Hydroquinone / Ethyl acetate	316,7	18,08	60,76
Methyl palmitate / Acetone	273,85	54,23	2698,83
Methyl palmitate / Acetone	284,05	73,87	611,96
Methyl palmitate / Acetone	290,65	51,24	120,75
Picric acid / Hexane	298,15	16,77	99,11
Picric acid / Cyclohexane	298,15	46,90	99,06
Picric acid / Dodecane	298,15	33,16	99,51
Picric acid / Decane	298,15	29,59	99,57
Picric acid / Heptane	298,15	0,01	99,29
Picric acid / Isooctane	298,15	2,07	99,17
Salicylic acid / Hexane	298,15	8,15	46,55
Salicylic acid / Dodecane	298,15	26,55	51,47
Salicylic acid / Decane	298,15	20,49	53,52
Salicylic acid / Heptane	298,15	0,11	45,69
Salicylic acid / Isooctane	298,15	0,00	29,86
5-Aminosalicylic acid / Methanol	297,15	4,22	100,00
5-Aminosalicylic acid / Methanol	299,15	2,04	100,00
5-Aminosalicylic acid / Methanol	301,15	0,04	100,00
5-Aminosalicylic acid / Methanol	303,15	1,71	99,99
5-Aminosalicylic acid / Methanol	305,15	1,36	99,99
3-Nitrobenzoic / 1-Propanol	298,15	4,41	80,52
3-Nitrobenzoic / 1-Pentanol	298,15	1,11	84,89
3-Nitrobenzoic / 2-Butanol	298,15	21,61	83,98
3-Nitrobenzoic / 1-Hexanol	298,15	0,42	86,30
3-Nitrobenzoic / 1-Octanol	298,15	11,23	88,67
Biphenyl / 1-Heptanol	298,15	0,87	4,30
Biphenyl / Heptane	298,15	9,59	23,56

Biphenyl / 2-Methyl-1-Pentanol	298,15	0,17	10,43
Biphenyl / 2-Methyl-2-Butanol	298,15	9,96	11,10
Biphenyl / 1-Pentanol	298,15	1,10	35,59
Benzoic / 1-Butanol	298,15	8,73	47,64
Benzoic acid / 1-Pentanol	298,15	1,70	44,38
Benzoic acid / 1-Heptanol	298,15	0,55	48,37
Benzoic acid / 2-Propanol	298,15	13,97	45,02

Les solubilités de plusieurs systèmes binaires non incluses dans le processus d'estimation à différentes températures ont été prédites en utilisant le nouveau modèle NK-UNIFAC. Les solubilités de ces mêmes systèmes binaires ont également été estimées par le modèle UNIFAC original. Ensuite, les résultats obtenus par les deux modèles ont été comparés aux solubilités expérimentales. Cette comparaison est illustrée dans les Figures 5.16 et 5.17.

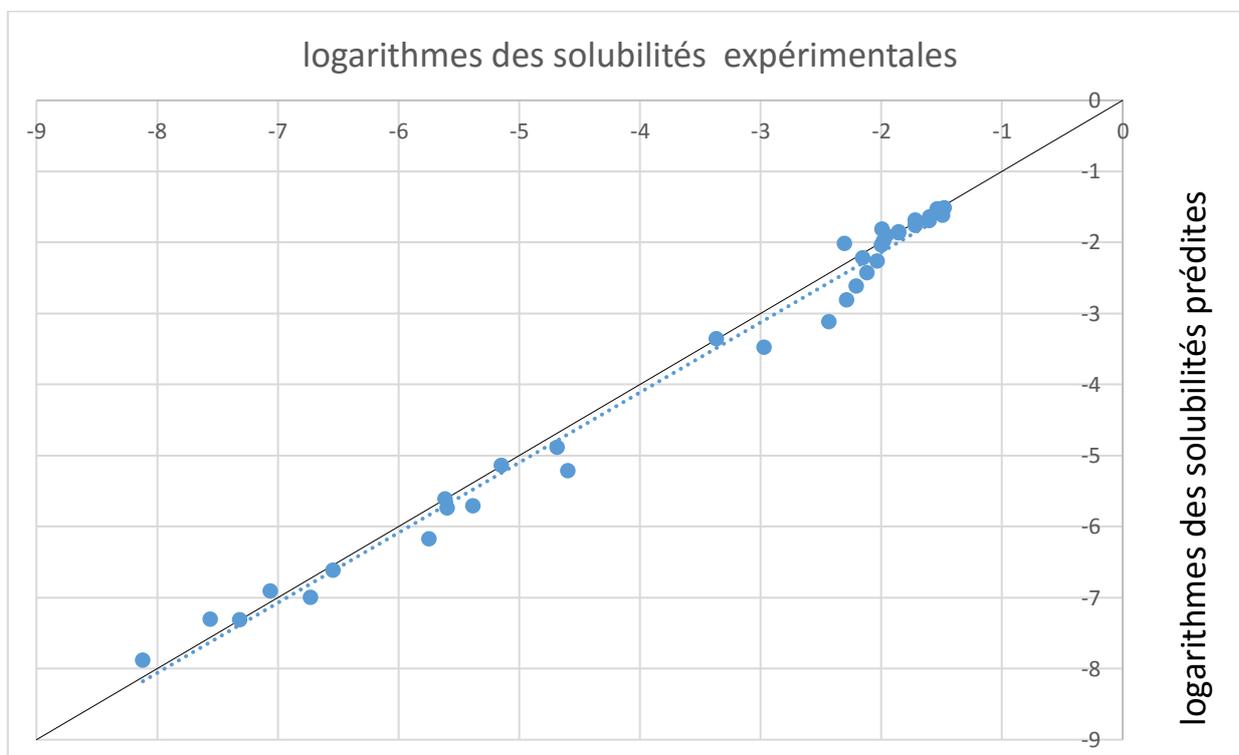


Figure 5.16 Solubilités prédites par NK-UNIFAC pour des systèmes non utilisés dans l'estimation en fonction des solubilités expérimentales

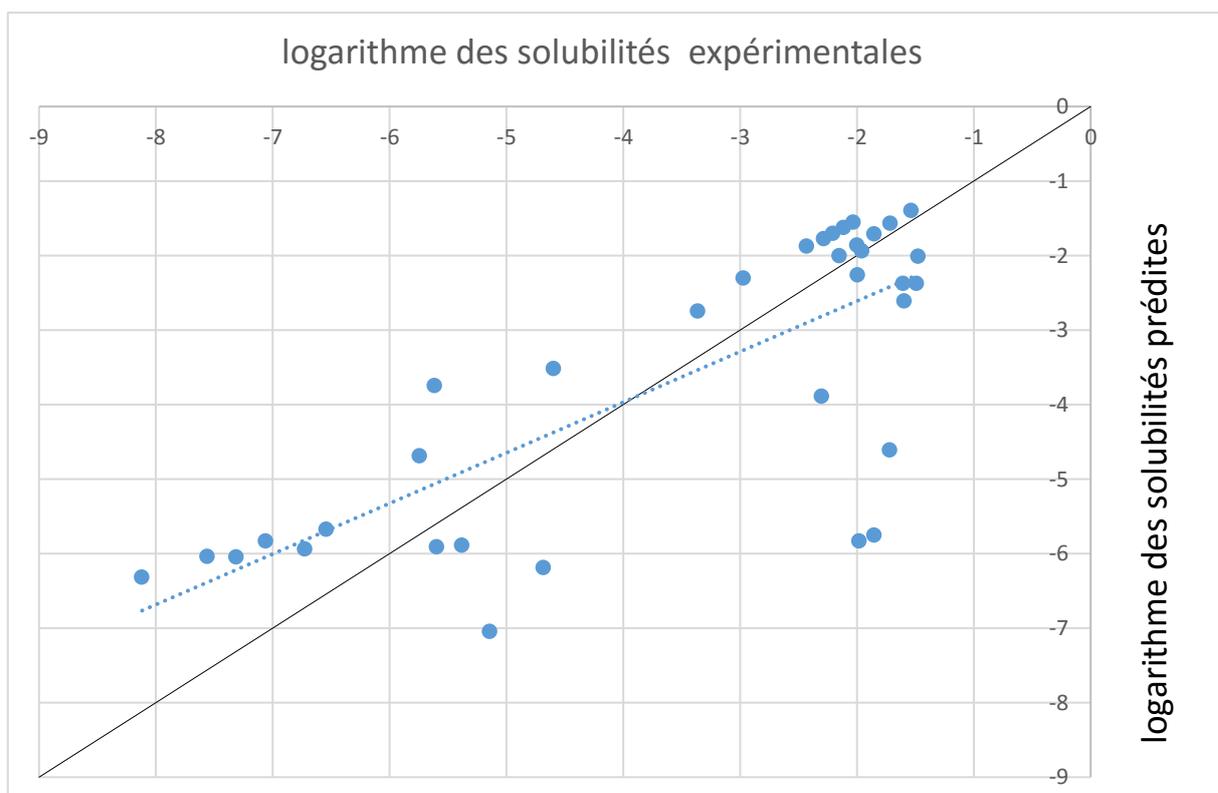


Figure 5.17 Solubilités prédites par UNIFAC original pour les mêmes systèmes en fonction des solubilités expérimentales

Dans le processus d'estimation, on commence par sélectionner des composés ne contenant pas un grand nombre de groupes fonctionnels (figure 5.1 et 5.2). On recueille le plus grand nombre de systèmes binaires contenant ces groupes, ensuite, on effectue le processus d'estimation comme expliqué au chapitre 4.

Lorsque les paramètres d'interaction souhaités sont obtenus, ils sont utilisés pour calculer la solubilité des composés non inclus dans le processus d'estimation, afin de valider le modèle. Les résultats obtenus sont ensuite comparés aux résultats du modèle UNIFAC d'origine.

On peut conclure à partir de ces résultats que le nouveau modèle NK-UNIFAC donne une prédiction plus précise que l'original.

Les résultats sont globalement satisfaisants et ce, pour plusieurs raisons parmi lesquelles on mentionne : La méthode de subdivision des molécules en petits groupes fonctionnels, comme dans KT-UNIFAC, rend l'évaluation plus efficace. En effet, prendre une longue

chaîne ne permet pas de prendre en compte les interactions entre les petits groupes qui composent ces chaînes.

Le processus d'estimation est très efficace et précis, il dépend de la méthode d'optimisation Nelder-mead et estime les paramètres d'interaction en même temps. Il estime également les paramètres d'interaction initiaux et modifie la fonction objective à chaque fois. Dans certains cas, on observe quelques écarts mineurs de solubilité, dus principalement aux isomères. Le modèle KT-UNIFAC ne prend pas en compte la géométrie des molécules. La différence entre la solubilité des isomères est parfois très grande, ce qui conduit à une estimation inefficace avec ces groupes fonctionnels

Autre chose est que les résultats empiriques disponibles ne sont pas tous mesurés avec une grande précision. Dans certains cas, on a trouvé pour certains composés dans les mêmes solvants et à la même température, que les résultats de la solubilité sont très différents. Cela peut être dû à la technique utilisée pour mesurer la solubilité ou aux conditions de mesure. Cette différence dans la mesure de la solubilité des mêmes composés fait que le processus d'estimation se comporte de manière erronée.

Peu importe la précision et l'efficacité du processus d'optimisation, il est impossible d'obtenir de bons résultats après avoir commencé avec des données erronées ou incomplètes.

Il en va de même pour les propriétés du corps pur (enthalpie de fusion, température de fusion). L'absence de mesure précise de ces caractéristiques peut entraîner une dérivation du processus d'estimation et des résultats médiocres.

Il est très important de bien mesurer la solubilité pour améliorer les performances du modèle KT-UNIFAC.

On voit ici la relation étroite entre l'expérimental et la modélisation. Pour que la modélisation soit efficace, elle doit être basée sur des mesures précises afin de former des équations valides et d'estimer efficacement les paramètres du modèle.

Cependant, on est toujours à la recherche d'une solubilité expérimentale pour de nouveaux systèmes, le processus de mise à jour de la matrice des paramètres d'interaction se poursuivra jusqu'à ce que les paramètres optimaux soient trouvés pour tous les systèmes.

5.8 Conclusion

Les données de solubilité de plus de 400 systèmes binaires soluté/solvant (pur) sont considérées pour estimer les paramètres d'interaction pour 20 groupes fonctionnels. L'intervalle de température va de 297,15 à 315,15 K. Toutes les données de solubilité ont été recueillies dans la littérature. Parmi ces groupes, les résultats obtenus pour 13 groupes fonctionnels ont été validés. Pour calculer les paramètres d'interaction de ces derniers, nous avons utilisé environ 250 données expérimentales de la solubilité dans différentes températures. 228 paramètres d'interaction a_{mn} et b_{mn} ont été estimés.

Selon les résultats présentés, la méthode proposée a donné des résultats qui, comparés à un modèle UNIFAC original, sont plus précis pour prédire les données de solubilité des composés complexes.

Certains de ces nouveaux paramètres d'interaction de groupes peuvent devoir être révisés en utilisant une base de données de solubilité plus étendue et la matrice des paramètres du modèle peut être complétée et étendue.

Conclusion générale

Plusieurs enjeux et exigences en thermodynamique liés à la conception des produits chimiques complexes notamment les produits pharmaceutiques, ainsi quelques exemples de nos activités de recherche dans ce domaine ont été décrits. L'impact économique du développement des produits pharmaceutiques et les défis posés par l'établissement d'une compréhension basée sur un modèle des produits chimiques continueront à inspirer les programmes de recherche thermodynamique dans les années à venir et indiqueront des lacunes dans les capacités actuelles de modélisation.

Dans ce travail, on a étudié en profondeur les différents modèles qui calculent le coefficient d'activité et de ce fait le calcul de la solubilité. Après avoir évalué ces modèles, étudié leurs avantages et limites, les avoir comparés et connaître le champ de leur application, nous avons conclu que le modèle UNIFAC était le plus largement utilisé et le plus crédible. Plusieurs versions de ce dernier ont été modifiées pour améliorer sa performance, par exemple le modèle UNIFAC modifié (Dortmund, Lyngby), UNIFAC Pharma modifié, CI-UNIFAC et KT-UNIFAC.

L'un de ces modèles est UNIFAC Pharma modifié, qui a été développé pour calculer la solubilité des composés pharmaceutiques. Au début de notre travail, on y a apporté quelques modifications et estimé les paramètres d'interaction qui constituent sa matrice.

Après avoir comparé les résultats du calcul de la solubilité à l'aide des modèles UNIFAC et UNIFAC Pharma modifié, il était clair pour nous que ce dernier était plus fiable. Mais ce modèle est limité par plusieurs facteurs, ce qui nous a incité à penser à en développer un autre.

Dans notre recherche en vue du développement d'un nouveau modèle de calcul de la solubilité, nous avons développé un nouveau concept pour le modèle UNIFAC en considérant la dépendance linéaire de la température des paramètres d'interaction comme dans UNIFAC Pharma modifié, et les groupes structurels tels que définis par le modèle KT-UNIFAC du

premier ordre. Plus de 220 paramètres d'interaction Amn et Bmn pour 13 groupes fonctionnels ont été estimés et validés pour le nouveau modèle NK-UNIFAC.

Plusieurs enquêtes et processus d'optimisation qui ne sont pas rapportés ici ont été effectués en coulisse, pour s'assurer qu'un ensemble optimal de paramètres d'interaction pouvait être obtenu. Certains processus d'optimisation prennent des mois pour trouver quelques paramètres d'interaction optimaux, c'est pourquoi une approche systématique a été développée afin de régresser les paramètres d'interaction des groupes fonctionnels avec une consommation minimale du temps et avec la plus grande fiabilité possible. Par conséquent, l'ensemble des paramètres d'interaction pour les groupes fonctionnels estimés est suffisamment bon pour prédire la solubilité de molécules complexes contenant ces dernières.

Tous les programmes de calcul ont été écrits en Visual basic (VBA). L'utilisation de VBA avec l'interface Excel nous facilite la saisie des données, nous permet de suivre l'estimation et les calculs, gérer les diverses tâches et d'analyser les résultats et les données de manière systématique et rapide.

Selon les résultats présentés, la méthode proposée a donné des résultats qui, comparés à un modèle UNIFAC original, sont plus précis pour prédire les données de solubilité des composés complexes. Certains de ces nouveaux paramètres d'interaction de groupe peuvent devoir être révisés en utilisant une base de données de solubilité plus étendue et la matrice des paramètres du modèle peut être complétée et étendue.

À partir des travaux présentés aux chapitres 2-5, les réalisations et les contributions de ce projet de thèse peuvent être résumées comme suit :

1. Etude et analyse de différents modèles prédictifs de la solubilité, indiquant les avantages et les limites ainsi que les améliorations apportées
2. Ajouter quelques modifications au modèle UNIFAC PHARMA modifié afin d'améliorer sa performance, d'estimer les paramètres d'interaction de certains groupes fonctionnels manquants, et de montrer ses limites dans la prédiction de la solubilité en raison des chevauchements entre groupes fonctionnels qui empêchent une bonne estimation des paramètres d'interaction
3. Développer une nouvelle version du modèle UNIFAC, que nous avons appelée KT-UNIFAC, où un programme a été développé pour ce dernier et une interface a été créée pour faciliter la saisie des données et obtenir les résultats

4. Développer une nouvelle méthode d'estimation des paramètres d'interaction jamais vue auparavant. Un article détaillé sur cette méthode sera bientôt publié, on rappelle que cette dernière dépend de la méthode d'optimisation de Nelder-Mead. La nouveauté est que tous les paramètres d'interaction sont évalués simultanément.

5. La nouvelle méthode d'estimation a été utilisée pour estimer les paramètres d'interaction du nouveau modèle NK-UNIFAC où un ordinateur spécial (poste de travail) a été utilisé. Ce dernier est caractérisé par de nombreux processeurs logiques et une énorme mémoire vive RAM.

6. Application du modèle KT-UNIFAC à des études des cas sélectionnées concernant l'investigation de la solubilité de systèmes complexes.

Par conséquent, on aimerait donner quelques recommandations et suggestions pour les travaux futurs. Les recommandations sont les suivantes :

- Recueillir le plus grand nombre de données expérimentales sur la solubilité des composés complexes, en particulier des composés pharmaceutiques, ceci afin d'estimer les paramètres d'interaction des groupes fonctionnels manquants dans la matrice.
- Ajouter un nouveau terme dans le modèle NK-UNIFAC afin de prendre en compte la géométrie des molécules, Ainsi, la précision du processus d'estimation des paramètres d'interaction. La prédiction de la solubilité est donc accrue
- Développer le modèle NK-UNIFAC pour calculer la solubilité dans les mélanges de solvants. Le programme nécessaire a été mis au point en attendant la collecte de données expérimentales de la solubilité dans les mélanges de solvants.

Références bibliographiques

- [1] Z. Q. Yu, J. W. Chew, P. S. Chow, and R. B. H. Tan, “Recent advances in crystallization control: An industrial perspective,” *Chemical Engineering Research and Design*, vol. 85, no. 7 A, pp. 893–905, 2007, doi: 10.1205/cherd06234.
- [2] B. Y. Shekunov and P. York, “Crystallization processes in pharmaceutical technology and drug delivery design,” *J. Cryst. Growth*, vol. 211, no. 1, pp. 122–136, 2000, doi: 10.1016/S0022-0248(99)00819-2.
- [3] R. H. Perry, D. W. Green, and J. O. Maloney, *Perry’s chemical engineers’ handbook*. New York, 1999.
- [4] H. Renon and J. M. Prausnitz, “Local compositions in thermodynamic excess functions for liquid mixtures,” *AIChE J.*, vol. 14, no. 1, pp. 135–144, 1968, doi: 10.1002/aic.690140124.
- [5] D. S. Abrams and J. M. Prausnitz, “Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems,” *AIChE J.*, vol. 21, no. 1, pp. 116–128, Jan. 1975, doi: 10.1002/aic.690210115.
- [6] J. M. Prausnitz, R. N. Lichtenthaler, and E. G. de Azevedo, *Molecular thermodynamics of fluid-phase equilibria*, 3rd ed. Upper Saddle River, N.J: Prentice Hall PTR, 1999.
- [7] A. Fredenslund, R. L. Jones, and J. M. Prausnitz, “Group-contribution estimation of activity coefficients in nonideal liquid mixtures,” *AIChE J.*, vol. 21, no. 6, pp. 1086–1099, 1975, doi: 10.1002/aic.690210607.
- [8] S. Banerjee, “Calculation of Water Solubility of Organic Compounds with UNIFAC-Derived Parameters,” *Environ. Sci. Technol.*, vol. 19, no. 4, pp. 369–370, 1985, doi: 10.1021/es00134a012.

- [9] S. O. Derawi, G. M. Kontogeorgis, and E. H. Stenby, “Application of group contribution models to the calculation of the octanol - water partition coefficient,” *Ind. Eng. Chem. Res.*, vol. 40, no. 1, pp. 434–443, 2001, doi: 10.1021/ie000361n.
- [10] C.-C. Chen and Y. Song, “Solubility Modeling with a Nonrandom Two-Liquid Segment Activity Coefficient Model,” *Ind. Eng. Chem. Res.*, vol. 43, no. 26, pp. 8354–8362, 2004, doi: 10.1021/ie049463u.
- [11] C. Chen, “Correlation and Prediction of Drug Molecule Solubility in Mixed Solvent Systems with the Nonrandom Two-Liquid Segment Activity Coefficient (NRTL - SAC) Model,” pp. 4816–4824, 2006.
- [12] M. B. Oliveira, M. J. Pratas, I. M. Marrucho, A. J. Queimada, and J. A. P. Coutinho, “Description of the mutual solubilities of fatty acids and water with the CPA EoS,” *AIChE J.*, vol. 55, no. 6, pp. 1604–1613, 2009, doi: 10.1002/aic.11766.
- [13] B. Schröder, L. M. N. B. F. Santos, I. M. Marrucho, and J. A. P. Coutinho, “Prediction of aqueous solubilities of solid carboxylic acids with COSMO-RS,” *Fluid Phase Equilib.*, vol. 289, no. 2, pp. 140–147, 2010, doi: 10.1016/j.fluid.2009.11.018.
- [14] A. Fredenslund, J. Gmehling, and P. Rasmussen, *Vapor-liquid equilibria using UNIFAC: a group contribution method*. Amsterdam; New York; New York: Elsevier Scientific Pub. Co. ; Distributors for the U.S. and Canada, Elsevier North-Holland, 1977.
- [15] S. Skjold-Jorgensen, B. Kolbe, J. Gmehling, and P. Rasmussen, “Vapor-liquid equilibria by UNIFAC group contribution. Revision and extension,” *Ind. Eng. Chem. Process Des. Dev.*, vol. 18, no. 4, pp. 714–722, 1979.
- [16] J. Gmehling, P. Rasmussen, and A. Fredenslund, “Vapor-liquid equilibria by UNIFAC group contribution. Revision and extension. 2,” *Ind. Eng. Chem. Process Des. Dev.*, vol. 21, no. 1, pp. 118–127, 1982, doi: 10.1021/i200016a021.
- [17] D. Tiegs, P. Rasmussen, J. Gmehling, and A. Fredenslund, “Vapor-liquid equilibria by UNIFAC group contribution. 4. Revision and extension,” *Ind. Eng. Chem. Res.*, vol. 26, no. 1, pp. 159–161, 1987.

- [18] H. K. Hansen, P. Rasmussen, A. Fredenslund, M. Schiller, and J. Gmehling, "Vapor-Liquid Equilibria by UNIFAC Group Contribution. 5. Revision and Extension," *Ind. Eng. Chem. Res.*, vol. 30, no. 10, pp. 2352–2355, 1991, doi: 10.1021/ie00058a017.
- [19] R. Wittig, J. Lohmann, and J. Gmehling, "Vapor-liquid equilibria by UNIFAC group contribution. 6. Revision and extension," *Ind. Eng. Chem. Res.*, vol. 42, no. 1, pp. 183–188, 2003, doi: 10.1021/ie020506l.
- [20] T. Magnussen, P. Rasmussen, and A. Fredenslund, "UNIFAC parameter table for prediction of liquid-liquid equilibria," *Ind. Eng. Chem. Process Des. Dev.*, vol. 20, no. 2, pp. 331–339, Apr. 1981, doi: 10.1021/i200013a024.
- [21] S. Gracin, T. Brinck, and Å. C. Rasmuson, "Prediction of Solubility of Solid Organic Compounds in Solvents by UNIFAC," *Am. Chem. Soc.*, pp. 5114–5124, 2002.
- [22] J. Gmehling, J. Li, and M. Schiller, "A modified UNIFAC model. 2. Present parameter matrix and results for different thermodynamic properties," *Ind. Eng. Chem. Res.*, vol. 32, no. 1, pp. 178–193, 1993, doi: 10.1021/ie00013a024.
- [23] B. L. Larsen, P. Rasmussen, and A. Fredenslund, "A Modified UNIFAC Group-Contribution Model for Prediction of Phase Equilibria and Heats of Mixing," *Ind. Eng. Chem. Res.*, vol. 26, no. 11, pp. 2274–2286, 1987, doi: 10.1021/ie00071a018.
- [24] J. Abildskov, *Extended Group Contribution Model for Polyfunctional Phase Equilibria*. 1998.
- [25] J. G. Gmehling, T. F. Anderson, and J. M. Prausnitz, "Solid-liquid equilibria using UNIFAC," *Ind. Eng. Chem. Fundam.*, vol. 17, no. 4, pp. 269–273, 1978.
- [26] B. Moller, J. Rarey, and D. Ramjugernath, "Activity of complex multifunctional organic compounds in common solvents," University of KwaZulu-Natal, Durban, 2009.
- [27] K. M. Ng, R. Gani, and K. Dam-Johansen, *Chemical Product Design: Towards a Perspective through Case Studies: Towards a Perspective through Case Studies*, vol. 23. Elsevier, 2006.
- [28] A. Diedrichs and J. Gmehling, "Solubility calculation of active pharmaceutical

- ingredients in alkanes, alcohols, water and their mixtures using various activity coefficient models,” *Ind. Eng. Chem. Res.*, vol. 50, no. 3, pp. 1757–1769, 2010.
- [29] U. Weidlich and J. Gmehling, “A modified UNIFAC model. 1. Prediction of VLE, hE, and γ_{∞} ,” *Ind. Eng. Chem. Res.*, vol. 26, no. 7, pp. 1372–1381, 1987.
- [30] J. Gmehling, J. Li, and M. Schiller, “A modified UNIFAC model. 2. Present parameter matrix and results for different thermodynamic properties,” *Ind. Eng. Chem. Res.*, vol. 32, no. 1, pp. 178–193, 1993, doi: 10.1021/ie00013a024.
- [31] J. Gmehling, J. Lohmann, A. Jakob, J. Li, and R. Joh, “A modified UNIFAC (Dortmund) model. 3. Revision and extension,” *Ind. Eng. Chem. Res.*, vol. 37, no. 12, pp. 4876–4882, 1998.
- [32] J. Gmehling, R. Wittig, J. Lohmann, and R. Joh, “A modified UNIFAC (Dortmund) model. 4. Revision and extension,” *Ind. Eng. Chem. Res.*, vol. 41, no. 6, pp. 1678–1688, 2002, doi: 10.1021/ie0108043.
- [33] J. W. Kang, J. Abildskov, R. Gani, and J. Cobas, “Estimation of mixture properties from first- and second-order group contributions with the UNIFAC model,” *Ind. Eng. Chem. Res.*, vol. 41, no. 13, pp. 3260–3273, 2002, doi: 10.1021/ie010861w.
- [34] A. Diedrichs, “Evaluation und Erweiterung thermodynamischer Modelle zur Vorhersage von Wirkstofflöslichkeiten,” pp. 1–253, 2010.
- [35] B. Bouillot, S. Teychené, and B. Biscans, “An evaluation of thermodynamic models for the prediction of drug and drug-like molecule solubility in organic solvents,” *Fluid Phase Equilib.*, vol. 309, no. 1, pp. 36–52, 2011, doi: 10.1016/j.fluid.2011.06.032.
- [36] A. S. Myerson and R. Ginde, *Handbook of Industrial Crystallization (Crystals, Crystal Growth, and Nucleation)*, no. October. 2002.
- [37] P. Novotný and O. Söhnel, “Densities of Binary Aqueous Solutions of 306 Inorganic Substances,” *J. Chem. Eng. Data*, vol. 33, no. 1, pp. 49–55, 1988, doi: 10.1021/je00051a018.
- [38] M. Broul, J. Nývlt, and O. Söhnel, “Solubility in inorganic two-component systems,” *Elsevier Sci. Amsterdam*, 1981.

- [39] H. Stephen and T. Stephen, "Solubilities of inorganic and organic compounds. Volume 2, ternary and multicomponent systems, part 2," *J. Chem. Educ.*, vol. 42, no. 8, p. 460, 1965, doi: 10.1021/ed042p460.2.
- [40] W. F. Linke, "Compounds, Solubilities : inorganic and metal-organic," *Washington, D.C. Am. Chem. Soc.*, 1958.
- [41] X. Zhang, G. Qian, X. Yang, C. Hu, and X. Zhou, "Solid-liquid equilibrium of dicyandiamide in different solvents," *Fluid Phase Equilib.*, vol. 363, pp. 228–232, 2014, doi: 10.1016/j.fluid.2013.12.001.
- [42] W. Beckmann, "Crystallization of Pharmaceutical Compounds Polymorphs, Pseudo-Polymorphs and Particle Formation," *Eng. Life Sci.*, vol. 3, no. 3, pp. 113–120, 2003, doi: 10.1002/elsc.200390013.
- [43] D. Sapoundjiev, H. Lorenz, and A. Seidel-Morgenstern, "Determination of solubility data by means of calorimetry," *Thermochim. Acta*, vol. 436, no. 1–2, pp. 1–9, 2005, doi: 10.1016/j.tca.2005.06.031.
- [44] S. M. Walas, *Phase Equilibria In Chemical Engineering*. 1985.
- [45] J. M. Prausnitz, R. N. Lichtenthaler, and G. Edmundo, "Molecular thermodynamics of fluid-phase equilibria," *The Journal of Chemical Thermodynamics*, vol. 2, no. 1, pp. 158–159, 1970, doi: 10.1016/0021-9614(70)90078-9.
- [46] T. C. Frank, J. R. Downey, and S. K. Gupta, "Quickly screen solvents for organic solids," *Chem. Eng. Prog.*, vol. 95, no. 12, pp. 41–61, 1999.
- [47] J. Gmehling and Bärbel Kolbe, "Thermodynamik, zweite, überarbeitete Auflage, VCH, Weinheim," *Berichte der Bunsengesellschaft für Phys. Chemie*, vol. 97, no. 6, pp. 842–842, 1992, doi: 10.1002/bbpc.19930970623.
- [48] G. Zhao and W. Yan, "Solubilities of betulin in chloroform + methanol mixed solvents at T = (278.2, 288.2, 293.2, 298.2, 308.2 and 313.2) K," *Fluid Phase Equilib.*, vol. 267, no. 1, pp. 79–82, 2008, doi: 10.1016/j.fluid.2008.02.013.
- [49] I. V. Prikhod'ko, F. Tumakaka, and G. Sadowski, "Application of the PC-SAFT equation of state to modeling of solid-liquid equilibria in systems with organic

- components forming chemical compounds,” *Russ. J. Appl. Chem.*, vol. 80, no. 4, pp. 542–548, 2007, doi: 10.1134/S1070427207040040.
- [50] B. Bouillot, “Approches thermodynamiques pour la prédiction de la solubilité de molécules d’intérêt pharmaceutique,” Institut National Polytechnique de Toulouse (INP Toulouse), 2011.
- [51] J. Larsson, “Methods for Measurement of Solubility and Dissolution Rate of Sparingly Soluble Drugs,” Lund University, 2009.
- [52] A. D. dos Santos, A. R. C. Morais, C. Melo, R. Bogel-Łukasik, and E. Bogel-Łukasik, “Solubility of pharmaceutical compounds in ionic liquids,” *Fluid Phase Equilib.*, vol. 356, pp. 18–29, 2013, doi: 10.1016/j.fluid.2013.07.020.
- [53] P. Kolář, J. W. Shen, A. Tsuboi, and T. Ishikawa, “Solvent selection for pharmaceuticals,” *Fluid Phase Equilib.*, vol. 194–197, pp. 771–782, 2002, doi: 10.1016/S0378-3812(01)00716-6.
- [54] J. Abildskov and G. M. G. M. Kontogeorgis, “A New Challenge of Applied Thermodynamics,” *Chem. Eng. Res. Des.*, vol. 82, no. November, pp. 1505–1510, 2004, doi: 10.1205/cerd.82.11.1505.52036.
- [55] S. K. Maity, “Modeling and Simulation of Solid-Liquid Equilibrium Perturbed-Chain Statistical Associating Fluid Theory,” *Chem. Eng.*, no. January, 2003.
- [56] T. Fornari, E. Ibañez, G. Reglero, and R. P. Stateva, “Analysis of Predictive Thermodynamic Models for Estimation of Polycyclic Aromatic Solid Solubility in Hot Pressurized Water,” *Open Thermodyn. J.*, vol. 5, pp. 40–47, 2011.
- [57] J. C. Givand, A. S. Teja, and R. W. Rousseau, “Effect of relative solubility on amino acid crystal purity,” *AIChE J.*, vol. 47, no. 12, pp. 2705–2712, 2001, doi: 10.1002/aic.690471210.
- [58] D. D. Dunuwila and K. a. Berglund, “ATR FTIR spectroscopy for in situ measurement of supersaturation,” *J. Cryst. Growth*, vol. 179, no. 1–2, pp. 185–193, 1997, doi: 10.1016/S0022-0248(97)00119-X.
- [59] T. Togkalidou, M. Fujiwara, S. Patel, and R. D. Braatz, “Solute concentration

- prediction using chemometrics and ATR-FTIR spectroscopy,” *J. Cryst. Growth*, vol. 231, no. 4, pp. 534–543, 2001, doi: 10.1016/S0022-0248(01)01518-4.
- [60] M. Fujiwara, P. S. Chow, D. L. Ma, and R. D. Braatz, “Paracetamol Crystallization Using Laser Backscattering and ATR-FTIR Spectroscopy: Metastability, Agglomeration, and Control,” *Cryst. Growth Des.*, vol. 2, no. 5, pp. 363–370, 2002, doi: 10.1021/cg0200098.
- [61] J. A. Falcon and K. A. Berglund, “Monitoring of Antisolvent Addition Crystallization with Raman Spectroscopy,” *Cryst. Growth Des.*, vol. 3, no. 6, pp. 947–952, 2003, doi: 10.1021/cg0340446.
- [62] P. Barrett and B. Glennon, “Characterizing the metastable zone width and solubility curve using lasentec FBRM and PVM,” *Chem. Eng. Res. Des.*, vol. 80, no. 7, pp. 799–805, 2002, doi: 10.1205/026387602320776876.
- [63] C. D. Bevan and R. S. Lloyd, “A high-throughput screening method for the determination of aqueous drug solubility using laser nephelometry in microtiter plates,” *Anal. Chem.*, vol. 72, no. 8, pp. 1781–1787, 2000, doi: 10.1021/ac9912247.
- [64] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, “Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings,” *Advanced Drug Delivery Reviews*, vol. 64, no. SUPPL. pp. 4–17, 2012, doi: 10.1016/j.addr.2012.09.019.
- [65] H. G. Brittain, “A Method for the Determination of Solubility of Metastable Crystal Phases Based on Total Light Scattering,” *Langmuir*, vol. 12, no. 2, pp. 601–604, 1996, doi: 10.1021/la950977v.
- [66] Y. Yi, D. Hatzivramidis, A. S. Myerson, M. Waldo, V. G. Beylin, and J. Mustakis, “Development of a small-scale automated solubility measurement apparatus,” *Ind. Eng. Chem. Res.*, vol. 44, no. 15, pp. 5427–5433, 2005, doi: 10.1021/ie049215y.
- [67] L. Pan, Q. Ho, K. Tsutsui, and L. Takahashi, “Comparison of chromatographic and spectroscopic methods used to rank compounds for aqueous solubility,” *J. Pharm. Sci.*, vol. 90, no. 4, pp. 521–529, 2001, doi: 10.1002/1520-6017(200104)90:4<521::AID-JPS1009>3.0.CO;2-B.

- [68] A. Avdeef, *Absorption and Drug Development: Solubility, Permeability, and Charge State*. 2012.
- [69] A. Avdeef, “pH-metric Solubility. 1. Solubility-pH Profiles from Bjerrum Plots. Gibbs Buffer and pKa in the Solid State.pdf,” *Pharm. Pharmacol. Commun.*, vol. 4, pp. 165–178, 1998, doi: 10.1111/j.2042-7158.1998.tb00328.x.
- [70] A. Avdeef and C. M. Berger, “pH-metric solubility.: 3. Dissolution titration template method for solubility determination,” *Eur. J. Pharm. Sci.*, vol. 14, no. 4, pp. 281–291, 2001, doi: 10.1016/S0928-0987(01)00190-7.
- [71] A. Avdeef, C. M. Berger, and C. Brownell, “pH-metric solubility. 2: Correlation between the acid-base titration and the saturation shake-flask solubility-pH methods,” *Pharm. Res.*, vol. 17, no. 1, pp. 85–89, 2000, doi: 10.1023/A:1007526826979.
- [72] M. Stuart and K. Box, “Chasing equilibrium: Measuring the intrinsic solubility of weak acids and bases,” *Anal. Chem.*, vol. 77, no. 4, pp. 983–990, 2005, doi: 10.1021/ac048767n.
- [73] D.J.W. Grant and H. G. Brittain, “Physical Characterization of Pharmaceutical Solids,” *Marcel Dekker, New York*, pp. 321–386, 1995.
- [74] E. Baka, J. E. A. Comer, and K. Takács-Novák, “Study of equilibrium solubility measurement by saturation shake-flask method using hydrochlorothiazide as model compound,” *J. Pharm. Biomed. Anal.*, vol. 46, no. 2, pp. 335–341, 2008, doi: 10.1016/j.jpba.2007.10.030.
- [75] F. L. Nordström and Å. C. Rasmuson, “Determination of the activity of a molecular solute in saturated solution,” *J. Chem. Thermodyn.*, vol. 40, no. 12, pp. 1684–1692, 2008, doi: 10.1016/j.jct.2008.06.016.
- [76] M. O. A. S. Ferreira, “Modelling of association effects by group-contribution: application to natural products,” *Faculdade de Enderharia Universidade do Porto*, 2003.
- [77] G. M. Wilson, “Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing,” *J. Am. Chem. Soc.*, vol. 86, no. 2, pp. 127–130, 1964, doi: 10.1021/ja01056a002.

- [78] H. Renon and J. M. Prausnitz, "Local compositions in thermodynamic excess functions for liquid mixtures," *AIChE J.*, vol. 14, no. 1, pp. 135–144, 1968, doi: 10.1002/aic.690140124.
- [79] A. Belabbaci, "Etude des équilibres liquide – vapeur des mélanges { amines – eau } et { alcools – hydrocarbures } : mesure et modélisation," Université Abou-Bekr Belkaid de Tlemcen, 2012.
- [80] R. L. Scott, "Corresponding states treatment of nonelectrolyte solutions," *J. Chem. Phys.*, vol. 25, no. 2, pp. 193–205, 1956, doi: 10.1063/1.1742853.
- [81] P. Atkins and de paula Julio, "Elements of Physical Chemistry," *Science (80-.)*, vol. 26, no. 670, pp. 588–588, 2009, doi: 10.1126/science.26.670.588-a.
- [82] K. Bitchikh, W. Louaer, M. A, and G. J, "Experimental and Modelling of liquid –solid equilibria," *EDP Sci.*, vol. 00011, pp. 1–6, 2009.
- [83] M. SIMOND, "Modélisation et simulation d 'interactions moléculaires Systèmes Eau + Amine," Université Blaise Pascal, 2010.
- [84] C. Chen and Y. Song, "Solubility Modeling with a Nonrandom Two-Liquid Segment Activity Coefficient Model," pp. 8354–8362, 2004.
- [85] C. C. Chen, "A segment-based local composition model for the gibbs energy of polymer solutions," *Fluid Phase Equilib.*, vol. 83, no. C, pp. 301–312, 1993, doi: 10.1016/0378-3812(93)87033-W.
- [86] B. Bouillot, *Introduction aux méthodes thermodynamiques pour le génie des procédés*. Mines Saint-Etienne, 2016.
- [87] A. Klamt and G. Schüürmann, "COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient," *J. Chem. Soc. Perkin Trans. 2*, no. 5, pp. 799–805, 1993, doi: 10.1039/P29930000799.
- [88] A. Klamt, V. Jonas, T. Bürger, and J. C. W. Lohrenz, "Refinement and parametrization of COSMO-RS," *J. Phys. Chem. A*, vol. 102, no. 26, pp. 5074–5085, 1998, doi: 10.1021/jp980017s.

- [89] S. T. Lin and S. I. Sandler, “A priori phase equilibrium prediction from a segment contribution solvation model,” *Ind. Eng. Chem. Res.*, vol. 41, no. 5, pp. 899–913, 2002, doi: 10.1021/ie001047w.
- [90] S. T. Lin and S. I. Sandler, “Prediction of octanol-water partition coefficients using a group contribution solvation model,” *Ind. Eng. Chem. Res.*, vol. 38, no. 10, pp. 4081–4091, 1999, doi: 10.1021/ie990391u.
- [91] E. Mullins, Y. A. Liu, A. Ghaderi, and S. D. Fast, “Sigma profile database for predicting solid solubility in pure and mixed solvent mixtures for organic pharmacological compounds with COSMO-based thermodynamic methods,” *Ind. Eng. Chem. Res.*, vol. 47, no. 5, pp. 1707–1725, 2008, doi: 10.1021/ie0711022.
- [92] A. Fredenslund, R. L. Jones, and J. M. Prausnitz, “Group-contribution estimation of activity coefficients in nonideal liquid mixtures,” *AIChE J.*, vol. 21, no. 6, pp. 1086–1099, 1975, doi: 10.1002/aic.690210607.
- [93] E. A. Macedo, U. Weidlich, J. Gmehling, and P. Rasmussen, “Communications Vapor-Liquid Equilibria by UNIFAC Group Contribution. Revision and Extension. 3,” *Ind. Eng. Chem. Process Des. Dev.*, vol. 22, no. 4, pp. 676–678, 1983, doi: 10.1021/i200023a023.
- [94] D. Tiegs, J. Gmehling, P. Rasmussen, and A. Fredenslund, “Vapor—Liquid Equilibria by UNIFAC Group Contribution. 4. Revision and Extension,” *Ind. Eng. Chem. Res.*, vol. 26, no. 1, pp. 159–161, 1987, doi: 10.1021/ie00061a030.
- [95] U. Weidlich and J. Gmehling, “A Modified UNIFAC Model. 1. Prediction of VLE, hE, and $^3\infty$,” *Ind. Eng. Chem. Res.*, vol. 26, no. 7, pp. 1372–1381, 1987, doi: 10.1021/ie00067a018.
- [96] J. Abildskov, R. Gani, P. Rasmussen, and J. P. O’Connell, “Beyond basic UNIFAC,” *Fluid Phase Equilib.*, vol. 158–160, pp. 349–356, 1999, doi: 10.1016/S0378-3812(99)00091-6.
- [97] J. Abildskov, L. Constantinou, and R. Gani, “Towards the development of a second-order approximation in activity coefficient models based on group contributions,” *Fluid Phase Equilib.*, vol. 118, no. 1, pp. 1–12, 1996, doi: 10.1016/0378-

3812(95)02846-3.

- [98] B. L. Larsen, P. Rasmussen, and A. Fredenslund, "A Modified UNIFAC Group-Contribution Model for Prediction of Phase Equilibria and Heats of Mixing," *Ind. Eng. Chem. Res.*, vol. 26, no. 11, pp. 2274–2286, 1987, doi: 10.1021/ie00071a018.
- [99] A. Jakob, H. Grensemann, J. Lohmann, and J. Gmehling, "Further Development of Modified UNIFAC (Dortmund) : Revision and Extension 5," *Ind. Eng. Chem. Res.*, vol. 45, no. 23, pp. 7924–7933, 2006.
- [100] H. Grensemann and J. Gmehling, "Performance of a conductor-like screening model for real solvents model in comparison to classical group contribution methods," *Ind. Eng. Chem. Res.*, vol. 44, no. 5, pp. 1610–1624, 2005, doi: 10.1021/ie049139z.
- [101] J. W. Kang *et al.*, "Quality assessment algorithm for vapor-liquid equilibrium data," *J. Chem. Eng. Data*, vol. 55, no. 9, pp. 3631–3640, 2010, doi: 10.1021/je1002169.
- [102] K. Kojima, H. Man Moon, and K. Ochi, "Thermodynamic consistency test of vapor-liquid equilibrium data. Methanol ~ water, benzene ~ cyclohexane and ethyl methyl ketone ~ water," *Fluid Phase Equilib.*, vol. 56, no. C, pp. 269–284, 1990, doi: 10.1016/0378-3812(90)85108-M.
- [103] K. Kurihara, Y. Egawa, K. Ochi, and K. Kojima, "Evaluation of thermodynamic consistency of isobaric and isothermal binary vapor-liquid equilibrium data using the PAI test," in *Fluid Phase Equilibria*, 2004, vol. 219, no. 1, pp. 75–85, doi: 10.1016/j.fluid.2004.01.018.
- [104] P. L. Jackson and R. A. Wilsak, "Thermodynamic consistency tests based on the Gibbs-Duhem equation applied to isothermal, binary vapor-liquid equilibrium data: data evaluation and model testing," *Fluid Phase Equilib.*, vol. 103, no. 2, pp. 155–197, 1995, doi: 10.1016/0378-3812(94)02581-K.
- [105] J. Wisniak, "A new test for the thermodynamic consistency of vapor-liquid equilibrium," *Ind. Eng. Chem. Res.*, vol. 32, no. 7, pp. 1531–1533, 1993, doi: 10.1021/ie00019a030.
- [106] H. C. Van Ness, S. M. Byer, and R. E. Gibbs, "Vapor-Liquid equilibrium: Part I. An

- appraisal of data reduction methods,” *AIChE J.*, vol. 19, no. 2, pp. 238–244, 1973, doi: 10.1002/aic.690190206.
- [107] E. F. G. Herington, “Tests for the consistency of experimental isobaric vaporliquid equilibrium data,” *J. Inst. Pet.*, vol. 37, pp. 457–470, 1951.
- [108] V. Diky, C. D. Muzny, E. W. Lemmon, R. D. Chirico, and M. Frenkel, “ThermoData Engine (TDE): Software implementation of the dynamic data evaluation concept. 2. Equations of state on demand and dynamic updates over the web,” *J. Chem. Inf. Model.*, vol. 47, no. 4, pp. 1713–1725, 2007, doi: 10.1021/ci700071t.
- [109] R. D. Chirico, M. Frenkel, V. V. Diky, K. N. Marsh, and R. C. Wilhoit, “ThermoML- An XML-based approach for storage and exchange of experimental and critically evaluated thermophysical and thermochemical property data. 2. Uncertainties,” *J. Chem. Eng. Data*, vol. 48, no. 5, pp. 1344–1359, 2003, doi: 10.1021/jc034088i.
- [110] V. Diky, R. D. Chirico, A. F. Kazakov, C. D. Muzny, and M. Frenkel, “ThermoData Engine (TDE): Software implementation of the dynamic data evaluation concept. 3. Binary mixtures,” *J. Chem. Inf. Model.*, vol. 49, no. 2, pp. 503–517, 2009, doi: 10.1021/ci800345e.
- [111] V. Diky, R. D. Chirico, A. F. Kazakov, C. D. Muzny, and M. Frenkel, “ThermoData engine (TDE): Software implementation of the dynamic data evaluation concept. 4. Chemical reactions,” *J. Chem. Inf. Model.*, vol. 49, no. 12, pp. 2883–2896, 2009, doi: 10.1021/ci900340k.
- [112] V. Diky *et al.*, “ThermoData Engine (TDE): Software implementation of the dynamic data evaluation concept. 5. Experiment planning and product design,” *J. Chem. Inf. Model.*, vol. 51, no. 1, pp. 181–194, 2011, doi: 10.1021/ci100373t.
- [113] M. Frenkel, R. D. Chirico, V. Diky, X. Yan, Q. Dong, and C. Muzny, “ThermoData Engine (TDE): Software implementation of the dynamic data evaluation concept,” *J. Chem. Inf. Model.*, vol. 45, no. 4, pp. 816–838, 2005, doi: 10.1021/ci050067b.
- [114] M. D. Frenkel *et al.*, “ThermoData Engine (TDE) Version 5.0 (Pure compounds, Equations of State, Binary mixtures, and Chemical Reactions),” *ThermoData Engine, NIST Stand. Ref. Database*, 2010.

- [115] V. V. Diky, R. D. Chirico, R. C. Wilhoit, Q. Dong, and M. Frenkel, “Windows-based guided data capture software for mass-scale thermophysical and thermochemical property data collection,” *J. Chem. Inf. Comput. Sci.*, vol. 43, no. 1, pp. 15–24, 2003, doi: 10.1021/ci025534t.
- [116] M. Frenkel, Q. Dong, R. C. Wilhoit, and K. R. Hall, “TRC SOURCE Database: A Unique Tool for Automatic Production of Data Compilations,” *Int. J. Thermophys.*, vol. 22, no. 1, pp. 215–226, 2001, doi: 10.1023/A:1006720022161.
- [117] A. Bondi, “van der Waals Volumes and Radii,” *J. Phys. Chem.*, vol. 68, no. 3, pp. 441–451, 1964, doi: 10.1021/j100785a001.
- [118] I. K. D. Tiegs, J. Gmehling, A. Medina, M. Soares, J. Bastos, P. Alessi, “Activity Coefficients at Infinite Dilution , Vol. IX, aus der Reihe: Chemistry Data Series, Part 1 + 2, Dechema, Frankfurt,” *Berichte der Bunsengesellschaft für Phys. Chemie*, vol. 92, no. 6, pp. 761–761, 2014, doi: 10.1002/bbpc.198800186.
- [119] I. Kikic and P. Alessi, “Liquid-liquid equilibrium for the activity coefficient determination: Effect of mutual solubility between binary systems and auxiliary solvents,” *Can. J. Chem. Eng.*, vol. 55, no. 1, pp. 78–81, 1977, doi: 10.1002/cjce.5450550114.
- [120] G. J. Maximo, A. J. A. Meirelles, and E. A. C. Batista, “Boiling point of aqueous d-glucose and d-fructose solutions: Experimental determination and modeling with group-contribution method,” *Fluid Phase Equilib.*, vol. 299, no. 1, pp. 32–41, 2010, doi: 10.1016/j.fluid.2010.08.018.
- [121] S. A. Febra and E. Committee, “Implementation and evaluation of two models for solubility of active pharmaceutical ingredients (APIs) Application on solvent selection in process design,” 2013.
- [122] G. J. Gloor, F. J. Blas, E. M. Del Río, E. De Miguel, and G. Jackson, “A SAFT-DFT approach for the vapour-liquid interface of associating fluids,” in *Fluid Phase Equilibria*, 2002, vol. 194–197, pp. 521–530, doi: 10.1016/S0378-3812(01)00774-9.
- [123] S. Gupta and J. D. Olsont, “Industrial Needs in Physical Properties,” *Ind. Eng. Chem. Res.*, vol. 42, no. 25, pp. 6359–6374, 2003, doi: 10.1021/ie030170v.

- [124] R. Gani, S. Christensen, and C. Gonzalez-Jimenez, “A new solvent selection method for organic synthesis,” *AIChE Annu. Meet.*, 2003.
- [125] E. C. Carlson, “Don’t gamble with physical properties for simulations,” *Chem. Eng. Prog.*, vol. 92, no. 10, pp. 35–46, 1996.
- [126] R. Agarwal, Y. K. Li, O. Santollani, M. A. Satyro, and A. Vieler, “Uncovering the realities of simulation - part 2,” *Chem. Eng. Prog.*, vol. 97, no. 6, pp. 64–72, 2001.
- [127] S. Karaborni, “Computer Simulations of Nanostructures in the Petrochemical and Pharmaceutical Industries,” *Plenary Pap. Present. FOMMS, USA*, 2003.
- [128] U. Von Stockar and L. A. M. Van Der Wielen, “Thermodynamics in biochemical engineering,” *J. Biotechnol.*, vol. 59, no. 1–2, pp. 25–37, 1997, doi: 10.1016/S0168-1656(97)00167-3.
- [129] S. I. Sandler and H. Orbey, “The thermodynamics of long-lived organic pollutants,” *Fluid Phase Equilib.*, vol. 82, no. C, pp. 63–69, 1993, doi: 10.1016/0378-3812(93)87129-O.
- [130] T. Lindvig, M. L. Michelsen, and G. M. Kontogeorgis, “Liquid–Liquid Equilibria for Binary and Ternary Polymer Solutions with PC-SAFT,” *Ind. Eng. Chem. Res.*, vol. 43, no. 4, pp. 1125–1132, 2004, doi: 10.1021/ie030443m.
- [131] I. Hahnenkamp, G. Graubner, and J. Gmehling, “Measurement and prediction of solubilities of active pharmaceutical ingredients,” vol. 388, pp. 73–81, 2010, doi: 10.1016/j.ijpharm.2009.12.036.
- [132] B. Bouillot, S. Teychené, and B. Biscans, “An evaluation of thermodynamic models for the prediction of drug and drug-like molecule solubility in organic solvents,” *Fluid Phase Equilib.*, vol. 309, no. 1, pp. 36–52, 2011, doi: 10.1016/j.fluid.2011.06.032.
- [133] A. Diedrichs and J. Gmehling, “Solubility calculation of active pharmaceutical ingredients in alkanes, alcohols, water and their mixtures using various activity coefficient models,” *Ind. Eng. Chem. Res.*, vol. 50, no. 3, pp. 1757–1769, 2011, doi: 10.1021/ie101373k.
- [134] I. A. Kouskoumvekaki, G. J. P. Krooshof, M. L. Michelsen, and G. M. Kontogeorgis,

- “Application of the Simplified PC-SAFT Equation of State to the Vapor–Liquid Equilibria of Binary and Ternary Mixtures of Polyamide 6 with Several Solvents,” *Ind. Eng. Chem. Res.*, vol. 43, no. 3, pp. 826–834, 2004, doi: 10.1021/ie034082m.
- [135] J. W. Kang *et al.*, “A new method for evaluation of UNIFAC interaction parameters,” *Fluid Phase Equilib.*, vol. 309, no. 1, pp. 68–75, 2011, doi: 10.1016/j.fluid.2011.07.001.
- [136] J. Nti-gyabaah, “Solubility and activity coefficient of pharmaceutical compounds in liquid organic solvents,” The State University of New Jersey, 2009.
- [137] M. C. Costa, M. P. Rolemberg, L. A. D. Boros, M. A. Krähenbühl, M. G. De Oliveira, and A. J. A. Meirelles, “Solid-liquid equilibrium of binary fatty acid mixtures,” *J. Chem. Eng. Data*, vol. 52, no. 1, pp. 30–36, 2007, doi: 10.1021/je060146z.
- [138] R. Coelho, P. G. Dos Santos, M. R. Mafra, L. Cardozo-Filho, and M. L. Corazza, “(Vapor + liquid) equilibrium for the binary systems {water + glycerol} and {ethanol + glycerol, ethyl stearate, and ethyl palmitate} at low pressures,” *J. Chem. Thermodyn.*, vol. 43, no. 12, pp. 1870–1876, 2011, doi: 10.1016/j.jct.2011.06.016.
- [139] M. C. Costa, L. A. D. Boros, J. A. P. Coutinho, M. A. Krähenbühl, and A. J. A. Meirelles, “Low-temperature behavior of biodiesel: Solid-liquid phase diagrams of binary mixtures composed of fatty acid methyl esters,” in *Energy and Fuels*, 2011, vol. 25, no. 7, pp. 3244–3250, doi: 10.1021/ef2004199.
- [140] M. P. Rolemberg, “Equilíbrio sólido-líquid de ácidos graxos e triglicerídeos: determinação experimental e modelagem,” *State Univ. Campinas*, 2002.
- [141] L. A. D. Boros, “Mathematical thermodynamics Modeling and of solid-liquid equilibrium of fatty systems,” *Univ. Campinas*, 2005.
- [142] M. C. COSTA, “Experimental determination of solid-liquid equilibrium for binary systems of saturated fatty acids: a study detailed of the solid phase,” *Univ. Campinas*, 2008.
- [143] L. P. Cunico, “Modelling of phase equilibria and related properties of mixtures involving lipids,” Technical University of Denmark, 2015.

- [144] C.-C. Chen, "Toward development of activity coefficient models for process and product design of complex chemical systems," *Fluid Phase Equilib.*, vol. 241, no. 1–2, pp. 103–112, 2006, doi: 10.1016/j.fluid.2006.01.006.
- [145] C. Chen and P. A. Crafts, "Correlation and Prediction of Drug Molecule Solubility with the NRTL-SAC Model," *16th Eur. Symp. Comput. Aided Process Eng. 9th Int. Symp. Process Syst. Eng. Elsevier B.V.*, no. 2004, pp. 859–864, 2006.
- [146] C.-C. Chen, "Computer method and system for predicting physical properties using a conceptual segment model," *United States, Pat. Appl. Publ.*, pp. 1–70, 2011.
- [147] C.-C. Chen, "Computer method and system for predicting physical properties using a conceptual segment-based ionic activity coefficient," *United States Pat. Appl. Publ.*, pp. 1–38, 2011.
- [148] C.-C. Chen, "methods of modeling physical properties of chemical mixtures and articles of use," *United States Pat.*, pp. 1–36, 2010.
- [149] C.-C. Chen, "Computer method and system for predicting physical properties using a conceptual segment model," *United States Pat. Appl. Publ.*, pp. 1–79, 2008.
- [150] T. Spyriouni, B. Bouillot, S. Teychené, X. Krokidis, and B. Biscans, "Solubility of pharmaceuticals : a comparison of a PC-SAFT-BASED approach and NRTL-SAC," no. March, pp. 19–20, 2012.
- [151] F. L. Mota, A. J. Queimada, A. E. Andreatta, S. P. Pinho, and E. A. Macedo, "Calculation of drug-like molecules solubility using predictive activity coefficient models," *Fluid Phase Equilib.*, vol. 322–323, pp. 48–55, 2012, doi: 10.1016/j.fluid.2012.02.003.
- [152] E. Sheikholeslamzadeh and S. Rohani, "Prediction of Solubility of Active Pharmaceutical Ingredients in Single Solvents and Their Mixtures — Solvent Screening," pp. 1–24, 2015.
- [153] D. J. Widenski, A. Abbas, and J. A. Romagnoli, "Effect of the Solubility Model on Antisolvent Crystallization Predicted Volume Mean Size," *IFAC Proc. Vol.*, vol. 7, no. PART 1, pp. 207–212, 2009, doi: 10.3303/CET0917107.

- [154] E. Conte, Azizul A. Mustafa, Peter A. Crafts, and R. Gani, “Verification and Prediction of Solubilities of Active (Pharmaceutical) Ingredients In Solvents and Solvent Mixtures,” *AIChE J.*, pp. 28–29, 2011.
- [155] F. L. Mota, A. Queimada, and S. P. Pinho, “Solubilities of Some Pharmaceutical Compounds in Pure Solvents,” *18th Eur. Conf. Thermophys. Prop.*, 2008.
- [156] F. Ruether and G. Sadowski, “Modeling the Solubility of Pharmaceuticals in Pure Solvents and Solvent Mixtures for Drug Process Design,” vol. 98, no. 11, pp. 4205–4215, 2009, doi: 10.1002/jps.
- [157] D. M. Sevillano, “On the Prediction of Polyphenol Properties,” Technische Universiteit Delft, 2015.
- [158] F. L. Mota, A. P. Carneiro, A. J. Queimada, S. P. Pinho, and E. A. Macedo, “Temperature and solvent effects in the solubility of some pharmaceutical compounds: Measurements and modeling,” vol. 37, pp. 499–507, 2009, doi: 10.1016/j.ejps.2009.04.009.
- [159] M. Salimi, B. Zarenezhad, H. Fakhraian, and E. Choobdari, “Thermodynamic Modeling of Chiral Compounds Solubility Using Correlative and Predictive Models,” *J. Appl. Solut. Chem. Model.*, p. 5030, 2015.
- [160] R. J. Oldland, R. M. Davis, E. Marand, and Y. A. Liu, “Predicting Phase Equilibria Using COSMO-Based Thermodynamic Models and the VT-2004 Sigma-Profile Database Predicting Phase Equilibria Using COSMO-Based Thermodynamic Models and the VT-2004 Sigma-Profile Database,” no. November, 2004.
- [161] E. Mullins, Y. A. Liu, A. Ghaderi, and S. D. Fast, “Sigma profile database for predicting solid solubility in pure and mixed solvent mixtures for organic pharmacological compounds with COSMO-based thermodynamic methods,” *Ind. Eng. Chem. Res.*, vol. 47, no. 5, pp. 1707–1725, 2008, doi: 10.1021/ie0711022.
- [162] R. P. Gerber and R. D. P. Soares, “Assessing the reliability of predictive activity coefficient models for molecules consisting of several functional groups,” *Braz. J. Chem. Eng.*, vol. 30, no. 1, pp. 1–11, 2013, doi: 10.1590/S0104-66322013000100002.

- [163] L. S. Ferreira and J. O. Trierweiler, "Modeling and simulation of the polymeric nanocapsule formation process," *IFAC Proc. Vol.*, vol. 7, no. PART 1, pp. 405–410, 2009, doi: 10.1002/aic.
- [164] H. Yang, H. Ye, S. Zhai, and G. Wang, "First-principles prediction of phase equilibria using the PR+COSMOSAC equation of state," *2011 Int. Symp. Adv. Control Ind. Process.*, no. May, pp. 258–261, 2011, doi: 10.1002/apj.
- [165] A. B. Shell, "van der Waals Volumes and Radi," *Phys. Chem.*, pp. 441–451, 1965.
- [166] J. a. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, pp. 308–313, 1964, doi: 10.1093/comjnl/7.4.308.
- [167] A. Bondi, "Physical Properties of Molecular Crystals, Liquids, and Glasses," *John Wiley Sons, New York, London, Sydney*, 1968.
- [168] Y. H. Zhao, M. H. Abraham, and A. M. Zissimos, "Fast calculation of van der Waals volume as a sum of atomic and bond contributions and its application to drug compounds.," *J. Org. Chem.*, vol. 68, no. 19, pp. 7368–7373, 2003, doi: 10.1021/jo034808o.
- [169] A. Nouar and K. Lekmine, "Prédiction de la solubilité De substances d'intérêt pharmaceutique Par le modèle UNIFAC," Université Mentouri Constantine, 2011.
- [170] T. Magnusson, "Prediction of Liquid-Liquid Equilibria using UNIFAC," Danmarks Tekniske Højskole, Lyngby Denmark, 1980.
- [171] H. E. González Villalba, "Development of Group Contribution Plus Models for Properties of Organic Chemical Systems," Technical University of Denmark 1, 2009.
- [172] C. Kracht, H. Geyer, P. Ulbig, and S. Schulz, "Parameter Optimization of Group Contribution Methods in High Dimensional Solution Spaces," 2007, pp. 647–657.
- [173] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Q. Appl. Math.*, vol. 2, no. 2, pp. 164–168, 1944, doi: 10.1090/qam/10666.
- [174] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Source J. Soc. Ind. Appl. Math. J. Soc. INDUST. APPL. MATH*, vol. 11,

no. 2, pp. 431–441, 1963.

- [175] T. Z. BOULMEZAOUD, “Programmation linéaire et Méthode du simplexe (en bref),” Université de Versailles Saint-Quentin-en-Yvelines.
- [176] S. Reveliotis, “The Simplex Method,” *Georgia Institute of Technology*, pp. 87–96, 1997.
- [177] S. Teukolsky, “Multidimensional Optimization : The Simplex Method,” *Biostatistics*. [Online]. Available: <http://www.numerical-recipes.com/>.
- [178] C. Wassvik, “Computational Analysis of Aqueous Drug Solubility – Influence of the Solid State,” Uppsala University, 2006.
- [179] A. Nouar, I. Benmessaoud, O. Koutchoukali, and M. S. Koutchoukali, “Solubility Prediction of Active Pharmaceutical Compounds with the UNIFAC Model,” *Int. J. Thermophys.*, vol. 37, no. 3, 2016, doi: 10.1007/s10765-015-2033-x.