

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

VERS L'UTILISATION DE LA SPECTROSCOPIE RAMAN COUPLÉE À LA CHIMIE
QUANTIQUE POUR LA CARACTÉRISATION ROUTINIÈRE DE MOLÉCULES ORGANIQUES

MÉMOIRE PAR ARTICLE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
À LA MAÎTRISE EN CHIMIE

PAR

JASON MALENFANT

AVRIL 2023

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.04-2020). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Pour débiter, je remercie chaleureusement mon directeur de maîtrise Mathieu Frenette, pour son appui et son écoute, pour des discussions fort stimulantes, et pour m'avoir donné la latitude nécessaire ainsi que la poussée dont j'avais besoin pour trouver ma voie. J'en retire une des expériences collaboratives les plus appréciables qui soit. Merci également de m'avoir confié des charges d'assistant d'enseignement en séances d'exercices d'un cours magistral, et en laboratoire. Ce furent d'excellentes occasions d'affiner mes compétences de vulgarisation scientifique, et je l'espère, de transmettre un certain amour de la chimie aux personnes qui ont suivi les cours.

Il va de soi de mentionner mes collègues et stagiaires qui ont vaillamment apporté leurs compétences à ces travaux - en ordre chronologique : Hugo Racine-Béchar, Chloé Samard, Lucille Kuster, Maxime Denis, Kouassi Signo, Yohann Gagné et Hugo Frappé. Particulièrement, à Lucille, pour son écoute sans faille, ses conseils, son attitude joviale, son soutien constant, et surtout pour sa patience à mon égard dans mes erreurs de parcours. Merci également à Antoine Juneau et Taylor Hope pour m'avoir assisté dans mes débuts en chimie computationnelle et pour m'avoir motivé à donner le meilleur de moi-même en mettant la barre aussi haute. Par extension, je remercie également le reste du groupe de recherche dans lequel j'ai pu œuvrer ces dernières années; pour vos suggestions et vos retours toujours constructifs. Ce fut un réel plaisir de travailler avec chacune et chacun d'entre vous.

Dans le département de chimie-biochimie de l'Université, merci au professeur Ali Nazemi, pour m'avoir pris sous son aile pour mon tout premier stage en recherche, durant mes études au baccalauréat. Je suis reconnaissant pour les bons conseils que j'ai reçus à certains moments clé de mon parcours dans les études supérieures, et pour nos discussions toujours cordiales qui m'ont souvent redonné du cran. Également je mentionne avec beaucoup d'affection Jacqueline Tieu, technicienne hors pairs, avec qui j'ai pu œuvrer à la fois comme étudiant et comme démonstrateur, pour avoir rendu mes études et mon travail agréables.

Enfin, un merci spécial pour ma famille immédiate; Annie, Daniel et Marie-Eve, pour m'avoir soutenu dans l'atteinte de mes buts depuis le début et jusqu'à la fin.

“ “ “

N'oubliez pas, les enfants, la différence entre la science et déconner, c'est de noter ce que vous faites.

“ “ “

- Adam Savage, 'Les Stupéfiants'

TABLE DES MATIÈRES

REMERCIEMENTS	ii
LISTE DES FIGURES	vii
LISTE DES TABLEAUX	ix
LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES	x
LISTE DES SYMBOLES ET DES UNITÉS.....	xi
RÉSUMÉ.....	xii
ABSTRACT	xiii
CHAPITRE 1 INTRODUCTION.....	14
1.1 Contexte.....	14
1.2 De la spectroscopie Raman	16
1.2.1 Théorie fondamentale	16
1.2.2 Applications typiques	18
1.2.3 Microscope Raman confocal.....	21
1.3 Éléments de Théorie de la Fonctionnelle de la Densité	22
1.3.1 De l'approche Kohn–Sham	23
1.3.2 Les familles de fonctionnelles	25
1.3.2.1 Les fonctionnelles à approximation de la densité locale/de spin (LDA et LSDA).....	25
1.3.2.2 Les fonctionnelles à approximation par gradient généralisé (GGA).....	25
1.3.2.3 Les fonctionnelles à approximation par gradient méta-généralisé (mGGA).....	26
1.3.2.4 Les fonctionnelles hybrides.....	26
1.3.2.5 Les fonctionnelles double-hybrides	27
1.3.3 Les bases	28
1.3.4 Corrections choisies à l'énergie DFT	32
1.3.4.1 Correction d'interactions longue distance	32
1.3.4.2 Correction sur les bases	33
1.3.5 Choisir un niveau de théorie	34
1.3.6 Le calcul d'optimisation géométrique des molécules	35
1.3.7 Le calcul de spectres Raman théoriques.....	36
1.3.7.1 Méthode dynamique.....	36
1.3.7.2 Méthode statique.....	37
1.3.7.3 Modélisation de l'environnement chimique	38
Solvatation implicite	39
Solvatation explicite.....	40
Modèles multiéchelle.....	40
1.4 Conclusion du chapitre	41
CHAPITRE 2 MÉTHODES DE CORRÉLATION QUANTITATIVE DE SPECTRES RAMAN.....	42

2.1	Éléments de traitement des signaux spectroscopiques vibrationnels.....	42
2.1.1	Élargissement des pics ponctuels	42
2.1.2	Correction harmonique.....	43
2.1.3	Rééchantillonnage	44
2.1.4	Troncature	44
2.1.5	Normalisation.....	44
2.2	Corrélation quantitative de signaux spectraux – méthodes connues	44
2.3	iROMA : Approche manuelle	45
2.3.1	Principes	45
2.3.2	Résultats et discussion	48
2.4	Logiciel SARA : Approche automatisée.....	57
2.4.1	Principes	57
2.4.2	Résultats et discussion	61
2.5	Conclusion du chapitre	63
	CHAPITRE 3 ARTICLE: TOWARDS ROUTINE ORGANIC STRUCTURE DETERMINATION USING RAMAN MICROSCOPY AND SPECTRAL MATCHING FROM r^2 SCAN-3c DFT CALCULATIONS	64
	CONCLUSION	73
	ANNEXE A DONNÉES DFT	74
	ANNEXE B SPECTRES RAMAN THÉORIQUES ET EXPÉRIMENTAUX	93
	ANNEXE C PARAMÈTRES D'ACQUISITION INSTRUMENTALE DES SPECTRES	108
	ANNEXE D INFORMATION EN SUPPORT DE L'ARTICLE DU CHAPITRE 3	110
1.	Simple computational protocol for modelling the Raman spectrum of organic solids as dimers	
	112	
2.	Molecular modelling strategy comparison.....	113
3.	DFT Data	114
3.1	ORCA Input file example for pre-optimization.....	114
3.2	ORCA Input file example for final frequency calculation	114
3.3	Optimized atomic coordinates	115
4.	Raman spectra.....	169
5.	Materials	181
6.	Synthesis of Borneol/Isoborneol.....	181
7.	Measurement parameters	182
8.	SARA Compression function	182

9. References 183

APPENDICE A iROMA (Code du programme pour Python 3) 184

RÉFÉRENCES 201

LISTE DES FIGURES

Figure 1-1 Diagramme d'une transition vibrationnelle en absorption infrarouge et par diffusion Raman de type Stokes	17
Figure 1-2 Schéma d'un microscope confocal Raman classique.	21
Figure 1-3 Schématisation du processus de détermination de la géométrie la plus basse en énergie.....	35
Figure 1-4 Caricature de la métaphore de l'objet modélisé appliqué aux solides moléculaires réels.	39
Figure 2-1 Données Raman DFT avant (a) et après (b) élargissement par profil de Voigt.	43
Figure 2-2 Profil de la courbe de pénalité pour la distance entre pics homologues	46
Figure 2-3 Interface utilisateur du logiciel de traitement spectral iROMA.	48
Figure 2-4 Matrice de corrélation fictive pour le scénario ayant une prédiction parfaite des spectres Raman, et une note qui représenterait la corrélation avec une justesse parfaite.	49
Figure 2-5 Matrice de notes de corrélation iROMA (en pourcentages) pour 20 molécules test. Chaque colonne représente un spectre calculé à B3LYP/6-311+G(2d,2p), et chaque ligne, un spectre expérimental acquis dans les conditions décrites à l'Annexe C. Les notes dans la région des mauvaises correspondances qui sont plus élevées qu'une des bonnes correspondances (dans la diagonale) sont indiqués en rouge.	49
Figure 2-6 Données spectrales expérimentales (noir, trait plein) et théoriques (trait bleu pointillé) sélectionnées représentant (a) le camphre, (b) la phénacétine, et (c) l'acide 1,3, 5-benzènetricarboxylique.....	51
Figure 2-7 Données spectrales d'une mauvaise correspondance sélectionnée. Presqu'aucun pic n'est communs aux spectres, mais par proximité de pics dans les régions denses du spectre, une note de 60% est accordée à la paire.....	52
Figure 2-8 Matrice de corrélation pour un jeu de molécules modifié. Les spectres théoriques ont été calculés avec solvatation implicite du THF, et les échantillons expérimentaux étaient solubilisés dans le THF. Figure reproduite du rapport de stage de Lucille Kuster, avec permission, en modifiant les identifiants des composés pour les harmoniser avec ce document. Les valeurs en rouge représentent les notes dans la région des mauvaises correspondances qui sont plus élevées qu'une des bonnes correspondances.	54
Figure 2-9 Comparaison spectrale pour les deux conditions expérimentales et les deux systèmes de modélisation de la phénacétine (M). Les spectres en bleu sont théoriques, et ceux en noir, expérimentaux.....	55
Figure 2-10 Comparaison spectrale pour les deux conditions expérimentales et les deux systèmes de modélisation de la coumarine (C). Les spectres en bleu sont théoriques, et ceux en noir, expérimentaux.....	56

Figure 2-11 Comparaison d'un spectre (a) avant et (b) après la compression des intensités. On remarque l'intensification des pics semblables à celui à 1200 cm^{-1} et la disparition de certains pics près de la ligne de base..... 59

Figure 2-12 Matrice de corrélation SARA des vingt spectres-tests de la Fig. 2-5 (modélisation en phase gazeuse, échantillons solides). Les valeurs en rouge représentent les notes dans la région des mauvaises correspondances qui sont plus élevées qu'une des bonnes correspondances. 61

Figure 2-13 Paire de spectres donnant une corrélation surévaluée à 71% avec SARA. 62

LISTE DES TABLEAUX

Tableau 1.1 - Résumé de quelques techniques d'analyse courantes en chimie organique	14
--	----

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES

B3LYP	Fonctionnelle d'échange Becke à trois paramètres avec corrélation Lee/Yang/Parr
BSSE	Erreur de superposition de la base
C-PCM	<i>Conductor-like Polarizable Continuum Model</i>
DFT	Théorie de la Fonctionnelle de la Densité de Kohn-Sham
DZ(V)	(Valence) Double zêta
ECP	Potentiel Effectif de Cœur
gCP	Correction par rééquilibrage à partir de la géométrie
GGA	Approximation par gradient généralisé
GP	Phase Gazeuse
HF	Hartree-Fock
HK	Hohenberg-Kohn
IR	Infrarouge
iROMA	<i>interactive Raman Organic Match Assessment</i>
KS	Kohn-Sham
LDA	Approximation de la densité locale
LSDA	Approximation de la densité locale de spin
mGGA	Approximation par gradient méta-généralisé
MM	Mécanique moléculaire
MP2	Théorie de la perturbation Møller–Plesset de second ordre
MS	Spectrométrie de Masse
QM	Mécanique quantique
QZ(V)	(Valence) Quadruple zêta
RMN	Résonance Magnétique Nucléaire
RR	Raman en condition de Résonance
SARA	<i>Similarity Assessment of Raman Arrays</i>
SCAN	Fonctionnelle <i>Strongly Constrained and Appropriately Normed</i>
SCF	<i>Self-Consistent Field</i> , champ auto-cohérent
SERS	Spectroscopie Raman Exaltée par la Surface
TZ(V)	(Valence) Triple zêta
UV-Vis	Spectroscopie ultra-violet et visible
WCCA	Corrélation Croisée Moyenne Pondérée, « <i>Weighted Cross-Correlation Average</i> »

LISTE DES SYMBOLES ET DES UNITÉS

cm^{-1}	Centimètres réciproques (1/cm)
E_x	Énergie d'échange électronique
E_c	Énergie de corrélation électronique
E_{xc}	Énergie d'échange-corrélation
v_{xc}	Potentiel d'échange-corrélation
pm	picomètre
$\tilde{\nu}$	Nombre d'onde
$\Delta\tilde{\nu}$	Différence de nombre d'onde

RÉSUMÉ

L'entreprise de créer des nouveaux composés organiques a été ponctuée d'innovations en méthodes de caractérisation des molécules. Le défi de l'identification d'une substance avec certitude est parfois considéré comme résolu, puisque des méthodes comme la résonance magnétique nucléaire et la spectrométrie de masse sont très puissantes. Néanmoins, certaines méthodes présentent des inconvénients majeurs pour les chimistes qui font des synthèses à multiples étapes puisque le rendement global de ces synthèses peut être très bas. La microscopie Raman est capable de donner un signal unique à une molécule avec seulement une dizaine de microgrammes d'un solide, sans préparation d'échantillon, et pourrait être utilisée comme méthode de caractérisation de routine non-destructrice. Les spectres riches en information générés par cette technique analytique sont difficiles à assigner pour un humain. Ce projet de maîtrise propose de coupler la microscopie Raman à une méthode de modélisation des spectres Raman par théorie de la fonctionnelle de la densité pour confirmer l'obtention de produits de synthèse organique. Deux méthodes de comparaisons spectrales ont été développées pour obtenir une note de corrélation entre un spectre théorique et un spectre expérimental. Finalement, une publication soumise démontre le potentiel du couplage DFT-Raman, ainsi qu'une méthode pour générer efficacement des spectres Raman théoriques de solides organiques par modélisation bimoléculaire.

Mots clés : Microscopie Raman, DFT, corrélation spectrale, caractérisation organique.

ABSTRACT

The enterprise of creating new organic compounds has been often enhanced by innovations in methods of characterizing molecules. The old challenge of identifying a substance with certainty is sometimes considered as solved, since methods such as nuclear magnetic resonance and mass spectrometry are quite powerful. However, a few well-known methods show major inconveniences for chemists who perform multiple steps syntheses where global yield can be low. Raman microscopy can give a unique signal to a molecule with ten micrograms of solid product, and could be used as a routine, non-destructive characterization method. The information-rich spectra thus generated can be a puzzle to assign for humans, therefore this master's project proposes to couple Raman microscopy with a Raman spectrum modelling method using density functional theory to confirm the creation of an organic product. Two methods for quantitatively comparing a theoretical Raman spectrum and an experimental spectrum were developed. Finally, a submitted article shows the potential of the DFT-Raman coupling, and a method to adequately generate theoretical Raman spectra of organic solids with bimolecular modelling.

Keywords : Raman Microscopy, DFT, spectral correlation, organic characterization.

CHAPITRE 1

INTRODUCTION

Ce premier chapitre est consacré à la mise en place du contexte scientifique, à la description de l'objectif principal du projet de recherche derrière ce mémoire, et à l'explication détaillée des bases théoriques nécessaires à la compréhension des fondements du projet.

1.1 Contexte

En chimie comme dans toute discipline scientifique, le principe du « fardeau de la preuve » dicte qu'une déclaration devrait être supportée par des preuves fournies par la personne ou le groupe qui la proclame. En chimie organique, la création de nouvelles molécules implique donc la démonstration de la réussite d'une réaction, ou d'une suite de réactions chimiques. La caractérisation de composés organiques nécessite l'emploi de techniques variées qui apportent chacune une information différente quant à la nature de la substance générée. À cet effet, une large sélection d'instruments peut être utilisée par l'organicien.ne pour démontrer la réussite d'une synthèse chimique. La spectrométrie de masse est une technique sensible pouvant fournir la masse exacte d'une molécule, donnant un compte du total massique des atomes présents, mais peu d'information sur l'arrangement de ces atomes. Pour cela, c'est la résonance magnétique nucléaire, ou RMN, qui amène souvent le plus d'information pertinente pour la chimie organique, n'en témoigne la révolution qu'elle a engendrée. La RMN du proton renseigne sur le niveau de blindage électromagnétique des atomes d'hydrogène présents sur une molécule, permettant d'obtenir des indices par inférence sur le type et l'emplacement relatif des carbones et de certains hétéroatomes qui composent le squelette de la molécule. (Pavia *et al.*, 2009) La RMN du carbone donne le niveau de blindage électromagnétique pour l'isotope ¹³ du carbone, ce qui permet une analyse structurale plus profonde, surtout si on utilise une expérience avec transfert de polarisation, homo- ou hétéronucléaire.

Tableau 1.1 - Résumé de quelques techniques d'analyse courantes en chimie organique

Technique	Information	Inconvénient
MS	Total des atomes	Destructif
UV-Vis	Transitions électroniques	Peu d'information
RMN ¹ H	Emplacement des atomes hydrogénés	Nécessite la solubilisation
RMN ¹³ C	Structure de la chaîne carbonée	Peu sensible
XRD	Structure exacte	Nécessite un monocristal
IR	Vibration moléculaire	Lecture difficile

Des méthodes basées sur la lumière existent aussi, comme la spectroscopie dans le domaine ultraviolet et visible (UV-Vis), permettant d'observer l'absorption lumineuse résultant d'une transition électronique d'un niveau d'énergie à un autre. Cette méthode fournit assez peu d'information en elle-même, et n'est pas applicable à un large éventail de substances puisque les transitions électroniques ne donnent pas toujours une absorption dans la bande du spectre électromagnétique observée, et fournit des pics larges et peu nombreux. À l'opposé des fréquences lumineuses de l'ultraviolet par rapport à la bande visible du spectre, on trouve les infrarouges. Ces longueurs d'onde auront en pratique toujours des pics d'absorption lors de l'observation d'échantillons organiques, puisque ce sont les mouvements de vibration de groupes d'atomes dans la molécule qui capturent les photons de ces énergies. La spectroscopie infrarouge (IR) est donc régulièrement utilisée pour valider la présence ou l'absence d'un groupe fonctionnel sur le composé d'intérêt, comme un acide carboxylique ou un alcyne. Chaque type de vibration a une énergie lui correspondant, et la transition vibrationnelle sera excitée par la longueur d'onde lumineuse lui étant associée. Les spectres vibrationnels contiennent beaucoup d'information, mais ceux obtenus par absorption directe présentent souvent des pics assez larges, rendant leur interprétation ardue. La technique « cousine » de l'absorption infrarouge, la spectroscopie Raman, génère des spectres ayant la sensiblement la même information sur les fréquences vibrationnelles, mais qui sont plus lisibles puisque les pics sont plus fins.

Là où, avec la RMN, il est possible et prescrit d'assigner chaque signal à un type de noyau atomique ayant un environnement chimique spécifique, avec la spectroscopie vibrationnelle cet exercice devient rapidement futile, pour deux raisons : 1) la quantité de pics dans un spectre est faramineuse, et 2) il est difficile d'assigner un mouvement spécifique effectué par des atomes spécifiques avec uniquement l'énergie du mode vibrationnel et l'intensité du signal. Si l'on souhaite utiliser plus d'un ou deux pics, on traitera plutôt l'information spectrale comme une empreinte digitale permettant d'identifier une substance organique. Ainsi, il devient nécessaire d'avoir un point de comparaison externe à l'échantillon afin de savoir si le signal obtenu est le bon. Pour des molécules connues et bien caractérisées, il suffit de trouver le spectre de référence existant dans une banque de données, et de le comparer au spectre acquis après la synthèse du produit. Pour pallier l'inexistence de spectre de référence pour une molécule fabriquée pour la toute première fois, on peut utiliser un modèle informatique basé sur la mécanique quantique (QM) peut être utilisé pour générer un spectre théorique pouvant être comparé au spectre expérimental obtenu par l'analyste. Dans le cas de cet ouvrage, la théorie de la fonctionnelle de la densité (DFT) est utilisée pour générer un spectre Raman calculé permettant d'évaluer la validité du spectre obtenu en laboratoire. L'information relative à cet aspect se trouve dans la section 1.3.7.

Dans la perspective de générer une preuve de la synthèse d'un composé, le dernier enjeu considérable concerne l'évaluation elle-même de la correspondance entre le spectre Raman expérimental, et la référence théorique. La vision humaine est certes une faculté remarquable, mais il est évident qu'une méthode robuste, répétable et quantitative est de mise. À cette fin, le chapitre 2 sera consacré à l'élaboration d'une méthode de corrélation spectrale faite sur mesure pour comparer un spectre Raman instrumental avec un spectre théorique calculé par DFT.

Globalement, l'idée centrale du projet est de tenter d'élever la spectroscopie Raman à une méthode de caractérisation de routine en chimie de synthèse organique en la couplant à une prédiction DFT du spectre évalué quantitativement.

Les prochaines sections, 1.2 et 1.3, sont consacrées à l'explication de la théorie de la spectroscopie Raman et de la DFT respectivement, en partant des bases communes aux chimistes de tous les horizons.

1.2 De la spectroscopie Raman

Commençons par expliquer la technique analytique choisie pour ce projet de maîtrise. La spectroscopie Raman est somme toute assez peu utilisée en synthèse organique de nos jours, et il sera pertinent de la connaître un minimum pour comprendre les enjeux de modélisation de ses spectres par chimie quantique.

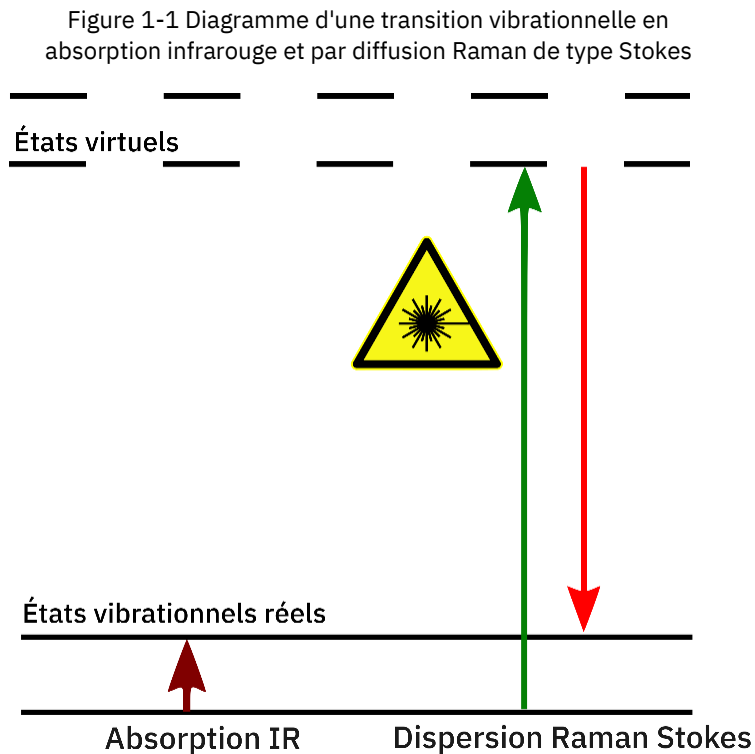
1.2.1 Théorie fondamentale

Le processus de diffusion Raman a été co-découvert par Sir Chandrasekhara Venkata Raman et son étudiant Sir Kariamanikkam Srinivasa Krishnan au 20^e siècle, (Raman, 1953) découverte qui valut à Sir Raman le prix Nobel de physique de 1930. On donne le nom de diffusion Raman à tout évènement de diffusion inélastique d'un rayonnement, lorsque l'énergie incidente vient d'un photon. Ce type de collision est dit « inélastique » parce que l'énergie du photon n'est pas la même avant et après l'évènement. Les deux sous-catégories de diffusion Raman sont la diffusion Stokes, et Anti-Stokes, qui sont sans surprise l'inverse l'une de l'autre. La première est favorisée dans les conditions normales de laboratoire est la diffusion Stokes, et la seconde augmente avec la température de l'échantillon. En diffusion Stokes, la molécule observée gagne de l'énergie qui est « volée » au photon du laser, induisant un décalage vers le rouge de la lumière incidente. (Larkin, 2011) Pour une molécule, ce décalage est égal à la différence entre l'énergie de l'état vibrationnel excité et l'état initial de la molécule. Ceci nous permet d'introduire les règles de sélection pour le phénomène Raman. La principale dicte que pour qu'une transition vibrationnelle soit active en spectroscopie Raman, elle doit induire un changement de polarisabilité α du nuage électronique

de la molécule (Équation 1.1). Ce changement de polarisabilité dans le temps (t) dépend de la valeur d'équilibre de polarisabilité à l'équilibre pour la molécule donnée (α_0), la distance internucléaire (r), la séparation internucléaire maximale (r_{max}), et la fréquence de la vibration (ν_{vib}).

$$\alpha = \alpha_0 + (r_{max} \cos(2\pi\nu_{vib}t)) \left(\frac{\partial\alpha}{\partial r}\right)_e \quad (1.1)$$

Une conséquence – qui s'applique pour les molécules centrosymétriques - est celle d'exclusion mutuelle avec l'absorption infrarouge; qui dicte si une transition est permise en IR (lorsque le changement vibrationnel induit une différence de moment dipolaire), elle sera défavorable en Raman, et inversement.(Smith et Dent, 2005) En pratique, pour les molécules de toutes symétries, il arrive souvent que les pics soient visibles avec les deux instruments, mais l'intensité d'un pic dans le spectre d'une des techniques sera complémentaire avec l'intensité de l'autre. Donc en Raman, si ces conditions sont réunies, un photon incident interagit avec la matière, et la molécule passe par un état quantique virtuel (inobservable puisqu'extrêmement bref) avant de redescendre à un niveau vibrationnel réel supérieur au niveau initial (Figure 1-1). La longueur d'onde du photon sortant est plus grande, suffisamment pour être séparée des photons incidents, qui sont généralement d'une source monochromatique lorsqu'on souhaite détecter un signal à des fins analytiques.



Sachant que la diffusion élastique de la lumière est déjà un phénomène peu probable (on parle d'environ 0,05% des photons qui font de la diffusion Rayleigh), la diffusion Raman est un processus dit fortement inefficace, puisqu'environ un photon sur dix millions (0.00001%) sera diffusé de façon inélastique. Pour cette raison, la présence de fluorescence – qui est un processus plus favorable que la diffusion Raman - dans l'échantillon aux longueurs d'ondes à observer peut gravement interférer avec l'analyse. On doit aussi envoyer une grande quantité de lumière pour avoir suffisamment de photons Raman pour en faire la détection avec un tube photomultiplicateur, ou encore un capteur matriciel de type CCD (*charged coupled device*). À cette fin, un laser est une très bonne source lumineuse puisqu'intrinsèquement monochromatique, et potentiellement très intense. Pour compléter l'assemblage et en faire un instrument analytique, il n'y a qu'à ajouter un filtre pour séparer la lumière du laser de celle issue de la diffusion Raman, un détecteur polychromatique, et un système de traitement informatique, comme la majorité des spectromètres modernes.

1.2.2 Applications typiques

La spectroscopie Raman connaît déjà un certain succès en chimie des matériaux où elle permet de caractériser de nombreux composés comme le graphène et ses dérivés, dont les propriétés mécaniques et conductrices sont convoitées. Ces réseaux de carbones sp^2 bi ou tridimensionnels sont de bons exemples permettant d'illustrer le mécanisme de la Raman en analyse de matériaux, parce que leur structure très carbonée est particulièrement polarisable. La position et la largeur des pics Raman présents sur un spectre de matériau basé sur le graphène peut renseigner sur le nombre de feuillets superposés, la présence et la quantité de défauts dans la structure, (Orlando *et al.*, 2021) et la macrostructure - s'il s'agit de nanotubes, de feuillets, ou autre. (Wu *et al.*, 2018) La spectroscopie Raman y est un outil précieux qui catalyse le développement de nombreux composants comme des transistors nouveau genre.

L'étude des minéraux peut aussi être assistée par Raman. On peut utiliser la spectroscopie Raman pour valider l'authenticité de pierres précieuses, comme le diamant et le saphir. (Batista *et al.*, 2019) Une utilisation possiblement plus inspirante du point de vue du grand public se trouve dans l'exploration spatiale. Une mention spéciale sera donnée à l'astromobile *Perseverance*, qui a touché le sol martien en février 2021, et est équipé de deux instruments utilisant la Raman. Le principal, SHERLOC, est situé sur la tourelle du bras robotisé du véhicule, et permet d'amener l'appareil directement à 5 cm de la surface à observer. (Bhartia *et al.*, 2021) L'instrument couplé à un laser UV est capable d'analyser les minéraux de la Planète Rouge, mais a comme principal objectif d'essayer de trouver des composés organiques pouvant représenter des traces de vie passée ou présente sur le sol martien. L'analyse minérale plus approfondie

– littéralement – est plutôt effectuée avec SuperCam, l'appareil multifonctions couronnant le mat de *Perseverance*, et constituant la « tête » de l'astromobile.(Wiens *et al.*, 2021) SuperCam est une suite d'instruments comprenant un spectromètre atomique de plasma induit par laser servant à pulvériser et analyser le sol martien en dessous de la surface, mais est aussi équipé d'un spectromètre Raman à longue portée (jusqu'à environ sept mètres), permettant d'analyser toute sorte de cibles minérales que SHERLOC ne pourrait atteindre. Le laser du spectromètre Raman de SuperCam est moins énergétique, à 532 nm, permettant l'obtention de spectres avec une source d'excitation différente de SHERLOC. En effet, si un échantillon fait de la fluorescence, l'intensité lumineuse due à ce phénomène peut facilement étouffer le signal Raman, qui vient d'un phénomène peu efficace tel qu'expliqué à la section 1.2.1. Ainsi, plus l'énergie du laser est élevée, ou plus sa longueur d'onde est courte, plus les chances d'obtenir un signal de fluorescence sont élevées, jusqu'à l'atteinte des UV profonds (250 nm ou moins) où les deux phénomènes sont séparés en longueur d'onde. Ainsi, les chromophores donnent souvent des signaux intenses en spectroscopie Raman, ce qui en fait une technique de choix avec un laser peu énergétique (par exemple 1064 nm) pour l'étude de pigments anciens depuis la fin du siècle dernier.(Edwards *et al.*, 1997) Il y a des techniques permettant de diminuer ou supprimer ce signal,(Wei *et al.*, 2015) mais typiquement s'il est impossible de diminuer l'énergie du laser utilisé, on tente l'utilisation d'une méthode permettant au signal Raman lui-même de surpasser la fluorescence.

Il y a premièrement la spectroscopie Raman en condition de résonance (*Resonance Raman* ou *RR*). Pour avoir une multiplication de l'intensité du signal diffusé de plusieurs ordres de grandeur, il suffit d'avoir une source de lumière incidente dans une zone d'absorption de la molécule à l'étude, souvent dans la partie visible du spectre lumineux.(Smith et Dent, 2005) La nature non-destructrice de la lumière - utilisée en quantité raisonnable - permet une utilisation archéologique de la Raman, par exemple pour l'identification de pigments utilisés dans des peintures, des encres, et des teintures de vêtements avec la RR.(M. Edwards et Chalmers, 2005) Il est à noter que la fluorescence peut demeurer un problème dans ces applications. Sachant qu'on peut simuler les bandes de combinaison qui apparaissent dans les spectres RR avec la DFT,(de Souza *et al.*, 2019) il serait possible de l'utiliser pour ces travaux, même si elle ne serait pas la meilleure technique puisque le choix d'analytes s'en retrouve limité, et l'amplification du signal est sélective.

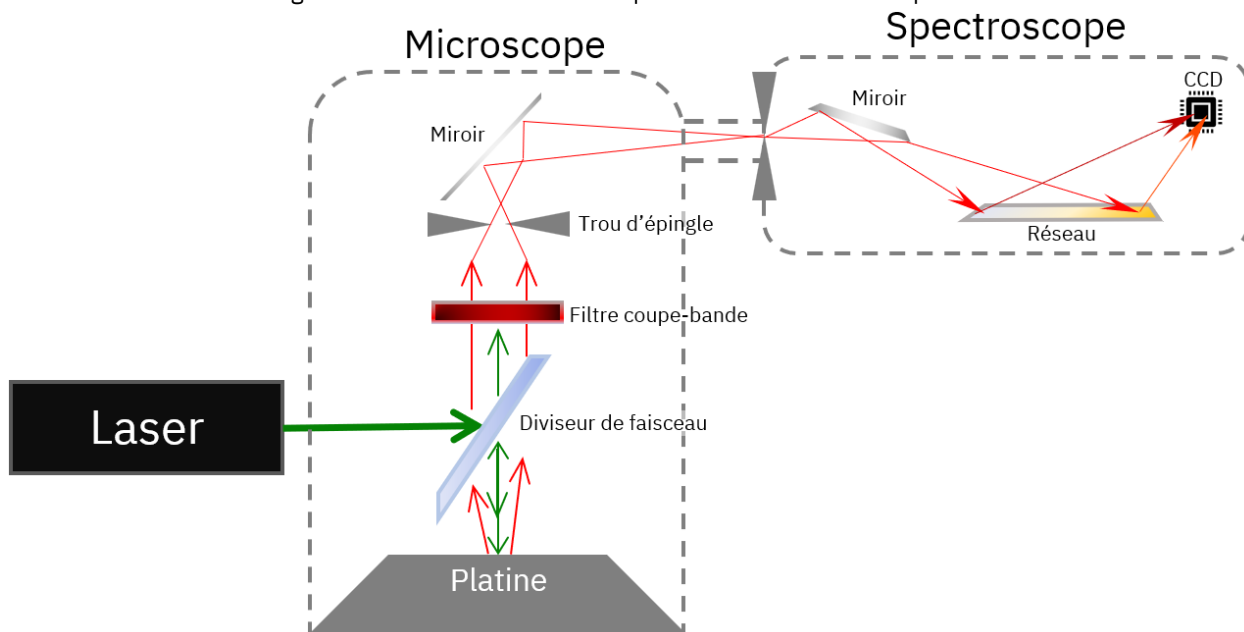
Finalement, l'analyse de biomolécules est un champ d'applications en plein essor pour la spectroscopie Raman.(Kuhar *et al.*, 2018) Elle est utilisée par exemple pour étudier les changements biochimiques résultats de maladies comme la tuberculose ou le cancer.(Kopec et Abramczyk, 2018) Ce champ

d'applications est idéal pour introduire la deuxième technique permettant d'avoir des signaux Raman très forts : la Spectroscopie Raman Exaltée par la Surface (SERS). Principale méthode Raman permettant une intensification très forte du signal, la SERS est au cœur d'un nombre grandissant de protocoles d'analyse de molécules, surtout organiques, de diverses natures. Ici, l'amplification du signal provient de l'excitation du plasmon de surface (l'oscillation du plasma électronique) d'un résonateur métallique comme une nanostructure d'or ou d'argent. Pour une explication plus en profondeur du mécanisme derrière la SERS, la revue de 2008 parue dans *Annual Review of Analytical Chemistry* est conseillée. (Stiles *et al.*, 2008) Il est alors possible d'obtenir une détection spectrale à une concentration de 10^{-18} M, dans des conditions favorables de résonance. (Kudelski, 2008) Ce type d'exaltation est utile pour toutes les applications où les concentrations en analytes sont très faibles, incluant ces dernières années de nombreuses utilisations en chimie bioanalytique et chimie biologique médicale, aussi diversifiées que la détection du glucose en temps réel en concentrations physiologiques *ex-vivo*, (Shah *et al.*, 2007) et *in-vivo*, (Huang *et al.*, 2020) ou encore la quantification de marqueurs de spores bactériennes. (Vidic *et al.*, 2020) Puisque la réponse instrumentale en SERS fournit une signature qui dépend de l'orientation de la molécule analyte par rapport à la surface excitée, la technique de traitement spectrale de pointe pour les données obtenues par SERS est l'utilisation de modèles d'intelligence artificielle comme l'apprentissage automatique pour la détection ou la classification des substances détectées à travers la réponse instrumentale. À cet effet, le lectorat est dirigé vers une revue parue dans *Trends in Analytical Chemistry* chapeauté par le professeur Jean-François Masson et co-écrite avec divers collègues. (Lussier *et al.*, 2020) On y recense de nombreuses approches chimométriques utilisant la SERS appliquées en alimentation, chimie criminalistique, microbiologie chimique, et diagnostics médicaux. Les modèles informatiques d'intelligence artificielle et d'apprentissage profond y sont entraînés à reconnaître les formes spectrales que peuvent prendre les différents analytes, et sont souvent déployés pour une reconnaissance en temps réel, en multiplex ou non. Les modèles basés sur l'intelligence artificielle permettent notamment de discriminer les signaux venant d'une matrice complexe comme un milieu biologique, mais sont plus complexes à déployer, parce qu'ils nécessitent un jeu de données d'entraînement très étendu et complet. Malgré son extraordinaire sensibilité, la variabilité des signatures spectrales SERS générées par l'analyse d'une même substance par rapport à l'angle d'incidence avec le plasmon (qui donnera une composante aléatoire dans le temps aux spectres acquis) signifie que les données expérimentales sont plus difficiles à modéliser précisément avec les méthodes de chimie computationnelle bien documentées. (Aroca, 2006) Cela signifie que la SERS n'est pas une candidate idéale pour le projet ci-présenté, et ne sera pas utilisée même si un couplage avec la DFT a déjà été réalisé. (Ciubuc *et al.*, 2017)

1.2.3 Microscope Raman confocal

Pour faire de la spectroscopie Raman une méthode acceptable par les chimistes organiciens, qui ont de hautes attentes face à une potentielle nouvelle méthode analytique de routine, l'applicabilité de l'appareillage doit être aussi large que possible. Une des idées centrales proposées dans ce projet est la mise en avant de la microscopie Raman confocale pour l'obtention d'information sur la nature de la substance à l'étude, qui est très versatile.

Figure 1-2 Schéma d'un microscope confocal Raman classique.



Le principe de fonctionnement d'un tel instrument est à peu de choses près ce qu'on peut s'imaginer d'un hybride entre un microscope confocal et un spectromètre Raman (Figure 1-2). L'échantillon est posé sur une platine généralement motorisée dans les deux axes du plan formé par le plateau principal (x et y), et parfois aussi dans l'axe vertical (z). Cela permet de faire des cartographies hyperspectrales riches en informations, et si la conception de l'instrument le permet, un balayage en profondeur de l'échantillon pour lesquelles les sciences de la vie ont trouvé diverses utilités. (Gomes da Costa *et al.*, 2019) Lors d'une analyse, la lumière incidente monochromatique générée par un laser d'une longueur d'onde choisie est amenée à l'échantillon à une certaine distance de la source. Ce faisceau sera amené à converger sur une petite région de l'échantillon. Le phénomène Raman se produit, et la lumière diffusée est captée par l'objectif. Le faisceau diffusé passe par le système optique du microscope, dans lequel un « trou d'épingle » se trouve, placé sur un plan dit confocal avec le plan focal de l'échantillon. Cela signifie que seule la lumière ayant pu être concentrée sur la région spécifique de l'échantillon sera en mesure de progresser vers le détecteur, et toute lumière parasite est éliminée, incluant la lumière venant de les

tranches de l'échantillon qui ne sont pas en focus.(Giridhar *et al.*, 2017) Après le « trou d'épingle », la lumière polychromatique contenant l'information spectrale est passée par un monochromateur, et est acheminée au détecteur. Dans le cas de l'instrument utilisé pour l'acquisition des données présentées dans cet ouvrage, le Renishaw in-Via, il s'agit de l'arrangement assez commun d'un réseau de diffraction suivi d'un capteur CCD permettant l'acquisition de différentes fréquences lumineuses en simultanément, évitant de devoir faire un long balayage du monochromateur – si la fenêtre spectrale désirée est à l'intérieur d'un certain intervalle. Avec cette technique, une cartographie serait réalisée en faisant une matrice de spectres point par point à différentes coordonnées de la topographie de l'échantillon.

Bien sûr, ce type d'instrument n'est pas conçu pour l'analyse de produits en vrac dans un réacteur industriel, et c'est pour cette spécificité que le microscope confocal Raman est la catégorie de spectromètres Raman choisie. Puisque le faisceau incident est focalisé sur une très petite région de l'espace (approximativement $8 \mu\text{m}^3$ dans le cas de l'instrument utilisé avec son objectif 50X), une très faible quantité de matière est requise pour obtenir un spectre Raman haute résolution, et avec un excellent rapport signal-bruit. L'aspect de confocalité signifie également qu'un spectre peut être obtenu à travers un flacon étanche ayant une atmosphère inerte, permettant l'analyse de composés sensibles à l'air qui seraient fabriqués dans une boîte à gants.

Le microscope Raman confocal représente un instrument produisant des spectres riches en information, et nécessitant une très faible quantité de matière pour obtenir des données avec un bon rapport signal-bruit. Cette quantité importante de pics est difficile à complètement analyser pour un humain. Il devient donc pertinent de compléter les données expérimentales avec de l'information externe pour valider l'empreinte digitale générée par l'échantillon. Bien qu'il existe des banques de spectres Raman de composés connus, la synthèse organique implique fondamentalement la caractérisation de nouvelles substances. De là vient la nécessité d'utiliser la chimie computationnelle : un spectre Raman théorique peut alors être comparé au spectre acquis afin de confirmer l'identité de l'échantillon.

1.3 Éléments de Théorie de la Fonctionnelle de la Densité

La DFT est un modèle en chimie computationnelle permettant la simulation d'atomes et de molécules en prenant en compte le comportement des orbitales électronique de chaque atome simulé. Ainsi, cette technique de chimie quantique est intrinsèquement beaucoup plus précise que les modèles de mécanique moléculaire (MM) où les atomes sont représentés par des charges ponctuelles. Certes, l'approche MM permet de calculer des systèmes immenses, comme des protéines entières, qu'il est impensable de

modéliser par DFT. Cependant, il est impossible d'obtenir une description juste des propriétés physiques des molécules sans utiliser un modèle chimique basé sur la mécanique quantique (QM). Le choix de la DFT s'impose parmi les méthodes en QM parce qu'elle représente un bon équilibre entre la charge de calcul et la précision des résultats obtenus. Les propriétés physiques peuvent être calculées avec significativement plus de précision avec la DFT qu'avec la théorie Hartree-Fock (HF) ou les méthodes semi-empiriques, à un coût de calcul considérable mais néanmoins réalisable sur une grappe de calcul moderne. Comme la chimie numérique est un monde vaste – on parle de « soupe alphabétique » tellement il y a d'acronymes à découvrir, et on qualifie régulièrement la DFT de « zoo » dans la littérature – les principes fondamentaux en seront étayés pour permettre une compréhension élémentaire du sujet.

1.3.1 De l'approche Kohn–Sham

La description mathématique d'un système chimique passe en grande partie par la description de ses électrons. Même en considérant le noyau atomique comme stationnaire (approximation de Born-Oppenheimer), la réalité du comportement électronique tel que comprise par la physique quantique contemporaine décrirait la structure électronique d'une molécule comme une grande équation de Schrödinger « à plusieurs corps » (à n électrons), décrivant toutes les interactions entre tous les électrons avec la théorie des fonctions d'ondes. Cependant, cette équation serait à toute fin pratique irrésolvable par un ordinateur classique. Sur le précepte de la simplifier en un modèle calculable, Walter Kohn et Lu Jeu Sham ont publié en 1965 ce qui est devenu la théorie de la fonctionnelle de la densité Kohn-Sham (KS DFT) en raffinant des théorèmes d'abord élaborés par Kohn et Pierre Hohenberg¹, dits « théorèmes HK ». (Hohenberg et Kohn, 1964) L'idée centrale amenée par les théorèmes HK est la preuve que l'énergie électronique totale d'un système à n électrons, dont la fonction d'ondes est basée sur x^n variables, peut en principe être déterminée à partir d'une fonctionnelle (une fonction mathématique qui est appliquée sur d'autres fonctions) de la densité électronique dans l'espace qui dépend donc de seulement trois variables; une par dimension spatiale. Chaque terme de l'énergie électronique totale pourrait être calculé seulement à partir de la densité électronique en fonction de l'espace, ce qui nécessitera l'application d'un potentiel effectif – calculé avec plus ou moins de précision selon la méthode finale - pour incorporer les interactions avec les autres électrons du système. Puisque la fonctionnelle « universelle » de l'énergie électron-

¹ Pour des raisons de concision un malheureux raccourci historique est ici fait. Le modèle réellement précurseur à la DFT est celui de Thomas et Fermi datant de 1927, dont l'évaluation de l'énergie cinétique électronique est considérée comme trop approximative. De ce modèle découle ensuite la théorie Hartree-Fock dans laquelle les électrons se déplacent dans un seul potentiel, contre deux pour la KS. Aussi, une heureuse conséquence des équations de KS est que leur processus de résolution permet d'obtenir « automatiquement » l'énergie cinétique avec toute la précision de la théorie des fonctions d'onde.

électron en fonction de la densité électronique, nécessaire pour obtenir le potentiel effectif exact à appliquer sur chaque électron, demeure à ce jour inconnue, le développement réalisé par Kohn et Sham est plutôt axé sur la résolution de n équations de type Schrödinger à un seul électron. Ces équations plus faciles à calculer constituent des « orbitales » hypothétiques : elles n'ont pas de réalité physique et ne sont ni des orbitales atomiques ni des orbitales moléculaires, mais ont comme caractéristique de générer la densité exacte des électrons dans le système à plusieurs corps. (Parr et Weitao, 1995)

En physique quantique, l'interaction d'échange décrit l'échange entre deux particules indiscernables comme l'électron. L'énergie qui lui est associée est l'énergie d'échange E_x . Le phénomène de corrélation électronique est le résultat de l'influence imposée sur le mouvement d'un électron par la présence de tous les autres électrons, et son énergie en chimie quantique est notée E_c . En DFT ces deux énergies sont souvent additionnées pour donner le terme d'échange-corrélation E_{xc} , mais il est bon de se rappeler qu'il vient de deux composantes distinctes.

L'électron de chaque orbitale (que nous appellerons ici orbitale KS) est affecté par un potentiel externe, qui dépend de la structure et de la composition élémentaire du système, et par le potentiel effectif interélectronique. Une innovation des équations Kohn-Sham a été d'appliquer un potentiel effectif en extrayant l'énergie d'échange-corrélation (E_{xc}) pour la rendre directement calculable en fonction de la densité électronique, tout comme les autres termes décrivant l'énergie totale du système. Sans E_{xc} et le potentiel v_{xc} qui en découle, le processus mathématique est réduit en une équation simplifiée (dite « de Hartree » dans ce cas), qui fait partie du modèle Hartree-Fock antérieur à la KS DFT. Ainsi cette énergie E_{xc} , qui englobe les phénomènes quantiques d'échange et de corrélation, donne un résultat plus juste avec la KS DFT qu'avec les précédents modèles. (Kohn et Sham, 1965) La façon d'approximer E_{xc} est le principal élément qui différencie les familles de fonctionnelles en DFT, qui sont le sujet de la prochaine section. Puisque E_{xc} dépend de la densité électronique, qui elle-même dépend de l'orbitale KS non-interagissante à un électron, qui à son tour dépend du potentiel effectif appliqué dans la fonctionnelle, il est nécessaire de résoudre les équations KS de façon autocohérente, c'est-à-dire de façon itérative jusqu'à convergence asymptotique des orbitales KS. Le processus est appelé SCF (*self-consistent field*, ou champ auto-cohérent) par emprunt à la théorie Hartree-Fock dont il est utile d'utiliser les rouages, malgré la manière différente de résoudre l'équation de Schrödinger.

Si ces travaux ont valu à Walter Kohn 50% du prix Nobel de chimie de 1998, de nombreuses avancées se sont greffées à cette idée centrale pour transformer la KS DFT en la DFT moderne. On recense par exemple

la redéfinition de certains termes mathématiques pour étendre l'applicabilité de la méthode,(Parr et Weitao, 1995) ou encore l'utilisation de jeux de bases, autre emprunt à la Hartree-Fock (HF). Leur application sera couverte plus en détails à la section 1.3.3 sur les bases.

1.3.2 Les familles de fonctionnelles

Les diverses fonctionnelles d'échange-corrélation en fonction de la densité électronique en fonction de l'espace, permettant d'obtenir l'énergie d'échange-corrélation de chaque orbitale KS, sont réparties en groupes selon le type d'approximation de l'énergie d'échange et de corrélation utilisée. On les place parfois sur une échelle « de Jacob » (métaphore de John P. Perdew, physicien américain de renom en DFT),(Perdew et Schmidt, 2001) dont les échelons partent du monde Hartree (« l'enfer », parfois appelé « la terre » de Hartree) et montent jusqu'à la précision dite chimique (le « paradis »), qui serait hypothétiquement une représentation si fidèle de la réalité qu'à toute fin pratique la fonctionnelle pour E_{xc} serait assez précise pour toute situation. L'atteinte d'une telle fonctionnelle reste à être démontrée, mais le problème demeure que les niveaux de théorie les plus élevés en DFT ont un coût de calcul rédhibitoire pour les systèmes qui nous intéressent.

1.3.2.1 Les fonctionnelles à approximation de la densité locale/de spin (LDA et LSDA)

Le premier échelon permettant tout juste de distinguer la DFT du monde Hartree est représentée par la simple idée d'approximer l'énergie d'échange-corrélation en la calculant sur un gaz d'électrons uniforme – modèle hypothétique où des électrons interagissent entre eux dans un solide dont les charges positives sont distribuées uniformément dans l'espace - ayant une densité comparable à la densité locale, plutôt que pour l'environnement électronique avoisinant l'orbitale KS.(Sousa *et al.*, 2007) Or, la densité électronique, même locale, d'un système moléculaire est rarement homogène. Cette approximation quoique très rapide à calculer – car simple et connue analytiquement pour l'échange - est si grossière qu'on lui applique une correction. La fonctionnelle LSDA SPW92 contient une composante prenant le spin en compte, et est paramétrisée de façon empirique pour la corrélation électronique. De nos jours il est facile et recommandé pour la grande majorité des applications d'utiliser une approximation plus précise.

1.3.2.2 Les fonctionnelles à approximation par gradient généralisé (GGA)

Ici, l'idée de calculer E_x et E_c à partir du gaz d'électrons uniforme est préservée, mais on lui impose un dégradé de densité dans l'espace, afin de palier à l'homogénéité du gaz uniforme. À partir de cet échelon, les fonctionnelles d'échange et de corrélation sont fréquemment différentes, permettant des combinaisons de fonctionnelles modulaires ayant différents auteurs. Les premières fonctionnelles

d'échange GGA reconnues ont été la B88 (Axel Becke), PB91 (John Perdew et collègues) et la PBE (Perdew-Burke-Ernzerhof). Les fonctionnelles de corrélation ayant eu du succès rapidement furent la P86 (Perdew), LYP (Lee, Yang et Parr) et aussi la PBE dans sa composante de corrélation. À titre d'exemple on peut faire les combinaisons : B88 + LYP = BLYP, ou B88 + P86 = BP86 pour avoir une fonctionnelle d'échange-corrélation complète.(Mardirossian et Head-Gordon, 2017) Ces fonctionnelles améliorent drastiquement la performance de la LSDA, à un coût toujours relativement faible sur un ordinateur moderne, mais leur précision est peut être légèrement insuffisante pour le calcul de certaines propriétés physiques comme les hauteurs absolues de barrières énergétiques des réactions, malgré l'addition de paramètres semi-empiriques.

1.3.2.3 Les fonctionnelles à approximation par gradient méta-généralisé (mGGA)

Le troisième échelon de l'échelle de Jacob des fonctionnelles représente l'ajout d'un autre ingrédient au calcul des énergies de corrélation et d'échange par le gaz d'électrons uniforme. Là où les GGA incluent la densité et sa première dérivée dans le calcul du potentiel d'échange-corrélation, un ingrédient pouvant être ajouté à une GGA est la dérivée seconde de la densité sous forme d'une Laplacienne. Une autre possibilité plus commune est d'y ajouter l'énergie cinétique. Ces fonctionnelles montrent une amélioration de la performance lors de certains calculs, mais peuvent avoir des problèmes avec les systèmes ayant beaucoup d'interactions faibles.(Johnson *et al.*, 2009) On retrouve les fonctionnelles SCAN et la récente r^2 SCAN (d'intérêt dans le cadre de ce document), TPSS et certaines fonctionnelles Minnesota dans cette famille. Le goulot d'étranglement de la précision de ces fonctionnelles est la description de l'énergie d'échange E_x , que le prochain échelon sera responsable d'améliorer.

1.3.2.4 Les fonctionnelles hybrides

Les fonctionnelles hybrides conservent le calcul de E_c l'énergie de corrélation tel que décrit pour les fonctionnelles GGA ou mGGA en utilisant le système du gaz d'électrons uniforme, avec un gradient de densité électronique. En revanche, pour le calcul de l'énergie d'échange E_x , c'est l'échange exact qui est utilisé, et mélangé avec une autre méthode, typiquement GGA ou mGGA.(Becke, 1993b) De ce mélange est tiré le caractère hybride de la fonctionnelle. L'échange exact est celui tiré de la théorie Hartree-Fock, où il est calculé d'une façon qui dépasse largement la portée de cet ouvrage. Il devrait être suffisant de dire que le calcul Hartree-Fock est de complexité et de précision se situant entre la MM et la DFT, et que l'énergie d'échange s'en retrouve donc mieux prédite par rapport aux échelons inférieurs de l'échelle de Jacob puisque l'échange Hartree-Fock est non-local, c'est-à-dire que les interactions d'échange électroniques sont modélisées au-delà de l'environnement immédiat de la coordonnée spatiale centrale

à l'électron. Il est important de noter que l'échange Hartree-Fock est exact du point de vue de la théorie HF, et non exact par rapport à la KS DFT, donc des approximations sont incluses dans ce traitement de E_x . Il a été observé que de mélanger l'échange HF à l'échange de type LDA par combinaison linéaire des énergies permet d'obtenir les meilleurs résultats.(Perdew *et al.*, 1996) Cela implique d'ajuster la pondération de chaque contribution, ce qui ajoute un aspect semi-empirique à la fonctionnelle hybride. Malgré cela, la performance de ces fonctionnelles est excellente, tellement que l'hybride la plus populaire, B3LYP, a constitué pratiquement une hégémonie de la DFT durant plusieurs années : 80% des publications en DFT de 1990 à 2006 l'ont utilisé.(Sousa *et al.*, 2007)

À titre d'exemple, la fonctionnelle B3LYP est dite à trois paramètres, qui sont trois nombres desquels sont dérivés les poids pour la combinaison linéaire des énergies. Les trois paramètres sont déterminés par ajustement (« *fitting* ») à des quantités physiques calculées comme des énergies d'atomisations, conférant un second caractère semi-empirique aux fonctionnelles hybrides. Les termes mathématiques décrivant E_{xc} pour la B3LYP comprend une E_x LDA, une E_x HF, l' E_x type GGA provenant de la fonctionnelle d'échange B88, une E_c LDA, et l' E_c de type GGA de la fonctionnelle de corrélation LYP.(Becke, 1993a) La PBE0 est également une fonctionnelle de choix parmi les hybrides.

Une sous-catégorie des hybrides contient les fonctionnelles hybrides séparées par la distance (*range-separated hybrid density functionals*) et ajoutent l'idée d'avoir différents pourcentages d'échange Hartree-Fock selon la distance depuis la coordonnée centrale de l'électron intéressé, permettant d'affiner le calcul davantage.

1.3.2.5 Les fonctionnelles double-hybrides

L'évolution logique du concept de la fonctionnelle hybride est d'également ajouter un terme à l'évaluation de l'énergie de corrélation E_c issu d'une autre méthode que la LDA, en plus du terme déjà ajouté à l'énergie d'échange E_x pour générer les fonctionnelles hybrides.(Zhao *et al.*, 2004) Ici la méthode de choix est une méthode souvent plus complète que la DFT elle-même, comme la théorie de la perturbation Møller-Plesset d'ordre x (si $x \geq 2$) notée MP x .(Zhao *et al.*, 2004) Considérée comme méthode « post-Hartree-Fock » de par sa précision supérieure à la DFT, cette méthode est prohibitive dans son coût. Par conséquent, on ne l'utilise pour une modélisation complète des électrons que sur des systèmes moléculaires assez réduits, et plus couramment pour améliorer un modèle plus simple, comme ici dans les double-hybrides pour calculer E_c seulement. Néanmoins dans son incarnation de second ordre (MP2), elle demeure une bonne façon d'améliorer la précision des énergies de corrélation, et par extension des

énergies d'échange-corrélation. On parle souvent d'un doublement du temps de calcul en utilisant une fonctionnelle double-hybride par rapport à une hybride, même en utilisant certaines approximations. La B2PLYP de Stefan Grimme, version double-hybride de B3LYP, utilise justement la théorie de la perturbation de second ordre couplée à une GGA pour le calcul de l'énergie de corrélation. (Grimme, 2006)

Les fonctionnelles double-hybrides séparées par la distance sont prisées pour les calculs nécessitant la plus grande rigueur possible, comme les calculs d'énergies de réactions et de hauteur de barrières d'activation. (Santra *et al.*, 2022)

Pour résumer, la fonctionnelle de la densité permet de déléguer le calcul de l'équation de Schrödinger à plusieurs corps en résolvant n équations de « type Schrödinger » appelées orbitales Kohn-Sham à un seul électron plutôt qu'une équation de Schrödinger à n électrons. Le calcul de l'énergie d'échange-corrélation électronique déterminé par la fonctionnelle est effectué avec différentes approximations, qui définissent les familles de fonctionnelles de la densité.

1.3.3 Les bases

Pour réduire davantage la charge de calcul, une autre astuce importante est déployée constamment en DFT. En effet, les ordinateurs classiques ne sont pas particulièrement efficaces pour résoudre des équations différentielles partielles, qui constituent une certaine partie d'une fonction d'onde quantique. Pour simplifier le calcul, on peut approximer la fonction d'onde mono-électronique en une combinaison linéaire de fonctions de base. Dans le contexte de la DFT, elles représentent des orbitales atomiques centrées sur le noyau atomique. On appelle un ensemble de ces fonctions une « base » (*basis set*). On utilise souvent plusieurs fonctions de base pour représenter une orbitale, permettant une représentation plus fidèle de l'orbitale atomique.

Idéalement les bases étaient composées d'orbitales de Slater (choisies pour leur réalité physique comme solutions à l'équation de Schrödinger pour l'atome d'hydrogène), mais une réalisation de Samuel Francis « Frank » Boys a eu un impact considérable sur la chimie numérique : l'orbitale de Slater peut être évaluée avec une certaine précision comme une somme de courbes gaussiennes dites « primitives ». (Boys, 1950) Cette idée a été attrapée au vol par John A. Pople avec une telle vigueur que le sol de la chimie en tremble toujours. Il s'avère que les gaussiennes sont significativement plus simples à calculer par ordinateur que les fonctions de Slater. Ainsi, Pople implémenta l'utilisation de gaussiennes primitives pour générer les fonctions de base représentant les orbitales dans un logiciel d'importance historique, *Gaussian*. Le travail

de Pople sur *Gaussian* et d'autres aspects de la chimie quantique comme la création de la famille de bases Pople parfois utilisées de nos jours lui ont valu la deuxième moitié du prix Nobel de chimie de 1998 avec Walter Kohn.

Une base, donc, est un ensemble de fonctions décrivant des orbitales atomiques. Il est commun de séparer le nombre de fonctions de base assignées à un atome en ayant une quantité constante de primitives par orbitales pour les électrons de cœur, et une quantité différente pour les électrons de valence. Dans ces bases « à valence séparée », on peut moduler le niveau de détail attribué aux deux types d'électrons. On attribuera généralement une seule fonction de base par orbitale pour les électrons de cœur, et au moins deux fonctions de base par orbitale de valence avec chacune un nombre variable de gaussiennes primitives, puisque ce sont généralement les orbitales de valence qui participent le plus activement à la chimie. La qualité zêta d'une base indique le nombre de fonctions de base attribuées à une orbitale électronique de valence : double zêta (DZ), triple zêta (TZ), quadruple zêta (QZ), etc. Une QZ sera donc plus longue à calculer qu'une DZ à cause du nombre de fonctions de base, mais permettra une meilleure description de la structure électronique de valence de l'atome. Aussi, une orbitale comprenant une seule primitive est dite « contractée », raccourci qui peut s'avérer utile pour les atomes lourds. (Jensen, 2013) La forme et la quantité de ces orbitales est une des raisons pour laquelle il existe une quantité faramineuse de bases, on parle alors de différence de patrons de contraction. Le niveau de contraction d'une base est donc une considération lors du choix de base pour le lancement d'un calcul DFT tout comme le zêta.

Une base est « minimale » lorsqu'elle ne contient que les fonctions nécessaires à l'endiguement des électrons de l'atome. Il est souvent préférable d'utiliser une base ayant un peu plus d'orbitales, pour que les électrons puissent y occuper une fraction de leur densité, ce qui représente une meilleure description d'un atome réel.

Lorsqu'on ajoute une ou des fonctions de base pour améliorer la polarisation de la densité électronique de l'atome lorsqu'il est dans une molécule, on parle d'un ajout de « fonction de polarisation ». Par exemple, pour l'atome d'hydrogène, on ajouterait une deuxième voire une troisième orbitale s, et une orbitale p à 3 fonctions de base (p_x , p_y , p_z). Il est également très commun d'ajouter dans une orbitale une ou des primitives ayant un profil plat et large. Ces fonctions appelées « diffuses » étendent les fonctions de base pour qu'elles atteignent une région plus éloignée du noyau, ce qui est important pour la modélisation des moments dipolaires et des interactions intra-et intermoléculaires.

Il existe une immense variété de bases. Elles sont souvent publiées en familles, ayant des propriétés ou des applications communes. Par exemple, les bases Cohérentes en Corrélation (cc-) ont comme propriété de proposer une augmentation régulière de l'énergie de corrélation avec l'ajout de primitives non-contractées de haut moment angulaire. Ici quelques familles de bases à orbitales gaussiennes pertinentes seront abordées.

- Les bases Pople sont les premières bases à fonctions gaussiennes à avoir connu un grand succès, et on pourrait spéculer qu'elles ont été une composante du premier niveau de théorie que plusieurs ont testé lors de leur premier contact avec la chimie numérique. Elles sont notées x-xyzG où x est le nombre de gaussiennes primitives additionnées pour représenter l'orbitale des électrons de cœur, y, le nombre de primitives de la première orbitale d'électrons de valence, et z, le nombre de primitives pour la seconde orbitale de valence. Le nombre de chiffres entre le trait d'union et la lettre 'G' (pour 'gaussienne') indique le zêta. Ainsi, 3-21G est une DZ ayant un total de six primitives par atome : trois pour les électrons de cœur, et un total de trois pour les électrons de valence.

La notation G* dénote l'ajout de fonctions de polarisation, et la notation +G, de fonctions diffuses. 6-311+G* est donc une base triple zêta avec un total de 11 primitives, plus des primitives diffuses et de polarisation. Pour le calcul de spectres Raman, la base Pople utilisée par notre groupe de recherche constituant un équilibre acceptable entre coût et précision est la 6-311+G(2d,2p) où deux orbitales d et deux orbitales p sont ajoutées pour la polarisation, de même que des fonctions diffuses. Un désavantage avec ces bases est qu'elles ne sont pas définies pour tous les atomes du tableau périodique – nous verrons pourquoi très bientôt. Ainsi, pour une molécule organique contenant un atome de brome, par exemple, il faut spécifier à la majorité des programmes de DFT l'utilisation d'une base différente pour cet atome.

- La famille Def2, souvent appelée Karlsruhe d'après l'endroit où elle a été créée en 2005, a été élaborée par Reinhart Ahlrichs et collègues pour le programme TURBOMOLE. (Weigend et Ahlrichs, 2005) Disponibles en plusieurs qualités, elles prennent la notation Def2-nZVP pour « deuxième base par défaut [pour TURBOMOLE] à valence n zêta avec fonctions de polarisation ». Pour cette famille, n vaut S, T, ou Q pour double, triple, et quadruple. L'ajout d'un deuxième 'P' à la fin du nom de la base indique l'ajout de fonctions de polarisation supplémentaires, et il n'y a pas de base ayant une fonction diffuse pour strictement la famille créée par Ahlrichs. Les bases Def2 sont considérées comme un bon choix moderne et contemporain pour beaucoup d'applications grâce à leur précision très constante d'un atome à un autre à travers tout le tableau périodique, pour un zêta donné. (Hill, 2013) Pour un niveau de complexité donnant un temps de calcul semblable, les

bases Karlsruhe ont généralement une performance supérieure aux bases Pople, et sont définies pour un grand nombre d'atomes – tout le tableau périodique à l'exception des actinides.

- Pour obtenir des bases d'une grande précision, orientée pour les calculs de propriétés physiques, l'équipe de Rappoport a généré une famille de bases cousine aux Def2. (Rappoport et Furche, 2010) L'ajout de fonctions de base sur une fonctionnelle existante (processus appelé « augmentation ») est ici représenté par l'ajout de la lettre 'D', pour 'diffuses', même si les fonctions ajoutées ne sont que légèrement diffuses. Donc, la base la plus complète pour cette famille est appelée Def2-QZVPPD, et est très lourde à calculer. En général on utilise les QZ pour les calculs plus exigeants comme les hauteurs de barrières réactionnelles; pour les calculs qui nous intéressent, une base triple zêta correctement augmentée suffit.

Maintenant que le vocabulaire central à propos des bases est défini, il est temps de répondre à une question mise en place précédemment, soit la raison pour laquelle une base peut ne pas être définie pour tous les atomes du tableau périodique. Quant à la définition d'une base pour un nombre d'atomes donné, il faut considérer qu'avec le numéro atomique, le nombre d'orbitales, et donc de fonctions primitives à calculer, augmente. Or, il y a un point pour chaque système numérique où le calcul est considéré en surcharge. Une technique qui « doit » être employée si on souhaite calculer un certain nombre d'atomes est l'utilisation de pseudopotentiels. Pour chaque atome, les électrons de cœur (ou un sous-ensemble de ceux-ci) sont simplifiés en un potentiel unique « ressenti » par les électrons de valence, permettant de réduire davantage la charge de calcul. L'acronyme ECP (*effective core potential*, potentiel de cœur effectif) est utilisé lorsque les pseudopotentiels électroniques sont complétés d'une fonction améliorant son profil asymptotique en fonction de la distance du noyau. Les bases modernes utilisent des ECP régulièrement, et presque assurément pour les atomes lourds. La création de pseudopotentiels de même que de patrons de contraction demande cependant des calculs qui auraient été assez lourds à l'époque où Pople développait ses bases à fonctions gaussiennes, et chaque atome supplémentaire défini dans une base représentait un certain travail de développement et de nombreux tests, le tout à effectuer sur des ordinateurs des années 1970.

En bref, les bases sont des ensembles de fonctions représentant des orbitales atomiques constituées de courbes gaussiennes primitives utilisées pour accélérer le calcul des équations d'orbitales de Kohn-Sham à un électron. Différentes bases conviennent pour différentes applications, et différentes tailles de systèmes à modéliser.

1.3.4 Corrections choisies à l'énergie DFT

1.3.4.1 Correction d'interactions longue distance

En DFT, la sélection d'une fonctionnelle combinée à une base, principales composantes du « niveau de théorie », représente un des grands choix à effectuer pour se lancer dans la modélisation d'un système. Certains ajouts peuvent être faits au niveau de théorie pour le rendre plus juste. Par exemple, il est connu qu'intrinsèquement, la KS DFT ne peut pas tenir compte des interactions de dispersion de longue distance, environ plus de 300 pm. (Parr et Weitao, 1995) Au fil des années d'activité de la DFT, divers théoriciens ont développé des méthodes permettant une correction pour ces interactions. Une approche est d'intégrer une composante à la fonctionnelle de la densité, une autre est de calculer une énergie de correction par paires d'atomes, souvent additionnée à l'énergie totale du système. Une des plus populaires est la série de corrections 'D', (Grimme *et al.*, 2010) dont la D4 est la plus récente représentante. (Caldeweyher *et al.*, 2017)

Le principe est de calculer pour chaque paire d'atomes dans le système une énergie qui dépend d'un coefficient assigné selon la nature des deux atomes impliqués, d'une polarisabilité précalculée, et de la distance interatomique. Pour les D et D2, les coefficients de dispersion étaient tabulés, et les valeurs venaient de calculs DFT. Pour la version D3, la flexibilité de l'idée est améliorée en ayant un coefficient basé sur des nombres de coordination d'atomes. (Goerigk, 2017) Enfin, la D4 complète la D3 en calculant à la volée la polarisabilité dipolaire en fonction de la charge partielle atomique, ce qui réduit le caractère empirique de la correction. Les D à D3 n'avaient comme seule donnée d'entrée la géométrie moléculaire (les coordonnées atomiques), et la D4 a comme entrées la géométrie et les charges. Cela confère à ces méthodes de correction une bonne flexibilité, et donne un résultat en quelques secondes de calcul, au maximum, pour un système de grande taille. Si les D et D2 sont considérées comme dépassées, D3 et D4 sont fortement recommandées pour compléter un niveau de théorie, sauf dans de très rares exceptions.

Les corrections de dispersions sont utilisées pour modéliser les interactions de London à longue distance : il faut donc diminuer la contribution des atomes s'ils sont proches de l'atome étudié pour éviter le comportement asymptotique de la correction. Une considération technique de l'utilisation de la D3 est donc le choix du type d'amortissement (« *dampening* ») de la correction. Généralement l'amortissement BJ (Johnson et Becke, 2005) est recommandée pour la D3, mais la correction D4 est préférable pour sa précision améliorée et comprend un seul type d'amortissement basé sur la Becke-Johnson – choisie pour sa meilleure performance avec la D3. (Caldeweyher *et al.*, 2019) Au meilleur de la connaissance de l'auteur,

il n'existe aucune raison de ne pas utiliser de correction D3 ou D4, sachant qu'il est plus rapide que leur calcul est extrêmement rapide. C'est également vrai pour la prédiction de spectres Raman, puisque la modélisation des interactions London modifiera la géométrie optimisée, de laquelle le calcul des fréquences de vibration dépend comme nous le verrons bientôt.

1.3.4.2 Correction sur les bases

Puisque les bases sont une approximation, leur utilisation n'est pas exempte de problèmes. L'un d'eux a tendance à apparaître fortement lorsqu'une base est de qualité zêta ou double zêta, et dans une moindre mesure avec les bases triple zêta, est employée. L'erreur de superposition des bases (BSSE, *Basis Set Superposition Error*) se produit lorsque deux atomes sont rapprochés de telle sorte qu'au moins une orbitale atomique de chaque atome se chevauche. Chaque atome peut alors utiliser une ou des fonctions de base de l'autre atome, agrandissant artificiellement la base d'une façon qui n'est pas réaliste puisque l'énergie est artificiellement minimisée lors du rapprochement de certains atomes. C'est un problème qui nous concerne puisque la distance interatomique est un des facteurs influençant les fréquences vibrationnelles. Différentes solutions à ce problème existent. Seule celle qui est utilisée dans cet ouvrage sera décrite.

La correction par rééquilibrage à partir de la géométrie (gCP, *geometrical CounterPoise*) consiste à calculer la valeur de la BSSE, puis de la soustraire à l'énergie KS DFT finale, comme le fait la correction D4. La BSSE est calculée uniquement à partir des coordonnées atomiques. En amont, pour chaque combinaison fonctionnelle-base pour laquelle les auteurs ont souhaité que la correction fonctionne, ils ont effectué un calcul mesurant sur une banque de dimères à diverses distances la différence d'énergie entre la petite version de la base (zêta ou double zêta) et la plus grande version possible avec en plus une augmentation par extrapolation. Ces différences d'énergie précalculées, approximativement proportionnelles à la BSSE, servent à construire une courbe en fonction de l'intégrale de superposition des orbitales (déterminés à partir de la même banque de dimères) avec quatre paramètres ajustés de façon empirique. Au moment d'appliquer la correction, l'intégrale de superposition est déterminée pour chaque paire d'atomes dans un certain rayon, et grâce aux quatre paramètres empiriques incrustés dans la courbe, une énergie de correction pour la paire d'atomes est rapidement extraite – un peu plus d'une minute pour un système à 2000 atomes. (Kruse et Grimme, 2012) La correction gCP à appliquer à l'énergie DFT est donc la somme des corrections individuelles pour chaque paire d'atomes.

1.3.5 Choisir un niveau de théorie

Les choix à faire pour avoir un niveau de théorie convenable pour la modélisation d'un système donné fait l'objet de nombreuses publications. On résumera ici les quelques considérations jugées pertinentes pour l'application présente, qui est la génération d'un spectre Raman de molécules organiques solides.

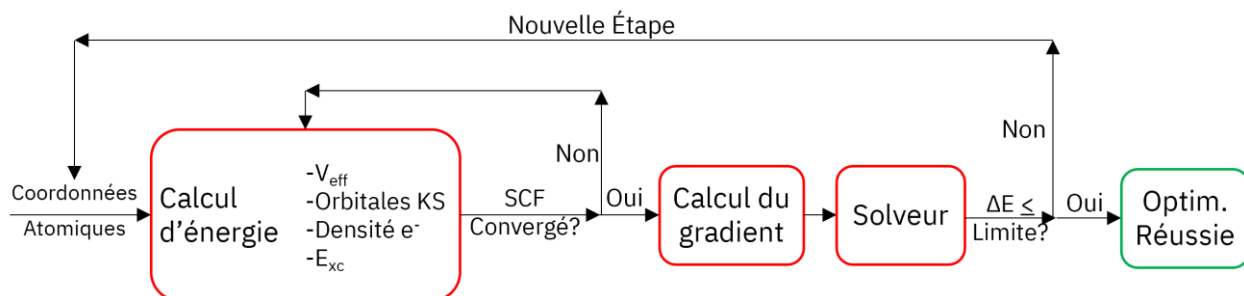
- Pour le choix d'une fonctionnelle,
 - ✓ Les LDA sont à proscrire;(Wong, 1996)
 - ✓ Pour une molécule de petite et moyenne taille (jusqu'à environ 70 atomes) une fonctionnelle hybride est préférée - la séparation de la portion d'échange HF par la distance (*Range-separation*) semble avoir peu d'impact sur les spectres Raman;(Jimenez-Hoyos *et al.*, 2008)
 - ✓ Utiliser une fonctionnelle mGGA pour les grandes molécules ou une GGA comme la PBE dans les cas extrêmes (120 atomes et plus, nécessite tout de même une puissance de calcul considérable);
 - ✓ B3LYP est une bonne hybride pour la génération de spectres Raman théoriques, mais devient difficile à calculer pour les molécules de taille d'intérêt aux organiciens.(Merrick *et al.*, 2007)
- Pour le choix d'une base,
 - ✓ Ne pas utiliser une base minimale pour générer des spectres Raman;
 - ✓ Une base quadruple zêta est généralement de taille superflue, les bases triple zêta augmentées donnent une qualité acceptable;
 - ✓ Si on doit absolument utiliser une base double zêta comme la Def2-SVP, s'assurer d'avoir une correction de la BSSE;
 - ✓ Pour bien modéliser les spectres Raman, les fonctions diffuses sont de mise, à moins d'avoir une fonctionnelle particulièrement robuste;(Ferrari et Bennett, 2019)
 - ✓ En cas de doute, Def2-TZVPD est une excellente base pour cette application pour les systèmes petits et moyens, et est définie pour un grand nombre d'éléments.
- Pour le choix d'une correction d'interactions de dispersion,
 - ✓ Si la correction D4 n'est pas disponible pour la combinaison fonctionnelle-base choisie, s'assurer d'activer la D3BJ, ou une autre méthode le cas échéant. Une correction de dispersion devrait en principe toujours être appliquée.

Avant de passer à la suite, une mention sera faite pour les personnes souhaitant se documenter pour trouver un niveau de théorie à utiliser. Une étude citée dans quelques articles récents donne un aperçu des résultats générés en Raman avec différentes bases.(Cheeseman et Frisch, 2011) Cependant, les quelques articles récents qui citent cette étude ne semblent pas être au fait que depuis sa publication en 2011, il y a plus d'une décennie, les bases recommandées comme la aug-cc-PVDZ ont été supplantées en termes de qualité, et la puissance des grappes de calcul haute performance a considérablement augmentée. Ainsi, cela peut servir de « récit à morale » indiquant de bien regarder l'année de publication de recommandations en chimie numérique, puisque c'est un domaine en plein essor. À titre d'exemple, les fonctionnelles de la densité assistées par apprentissage automatique sont en émergence, dont une qui a été produite pour la prédiction de spectres RMN.(Novitskiy et Kutateladze, 2022)

1.3.6 Le calcul d'optimisation géométrique des molécules

Une fois le niveau de théorie choisi, un.e chimiste peut générer un ensemble de coordonnées atomiques représentant la géométrie de la molécule. Pour obtenir un spectre Raman, il faut d'abord obtenir une structure qui soit minimisée en énergie. Pour y arriver, un calcul itératif d'optimisation géométrique est réalisé sur les coordonnées atomiques de départ. Chaque logiciel a des subtilités qui lui est propre, mais en général, pour la géométrie de départ, le champ auto-cohérent (SCF, défini à la section 1.3.1) est exécuté pour obtenir un ensemble d'orbitales Kohn-Sham convergées, et donc une énergie du système. Ensuite, le logiciel calcule un gradient d'énergie dans l'espace, et passe cette matrice à un algorithme solveur chargé de déterminer la prochaine géométrie à calculer.(Schlegel, 2011) Divers algorithmes ayant des fonctionnements internes fascinants existent, mais seront souvent traités en « boîte noire » par beaucoup de théoriciens, chose que nous ferons ici. Un des plus répandus est le DIIS (inversion directe [de l'hessienne] dans l'espace itératif).(Császár et Pulay, 1984) On retiendra que le solveur fournit un déplacement des atomes appelé « étape » permettant la construction d'une nouvelle géométrie, et le cycle recommence par la résolution du SCF. Chaque boucle SCF donne l'énergie du système à une géométrie donnée, et la boucle d'optimisation s'arrête quand la géométrie donne un gradient et un ΔE par

Figure 1-3 Schématisation du processus de détermination de la géométrie la plus basse en énergie.



rapport à la structure précédente ayant une valeur inférieure à un seuil déterminé par le logiciel, ou par l'utilisateur (voir Figure 1-3). Ce seuil devrait être assez bas pour assurer un minimum énergétique et prendre en considération le niveau de « bruit numérique » de la méthode employée, un artefact de la résolution numérique de certaines équations si une solution analytique est irréalisable ou non-implémentée.

La géométrie de départ donnée au logiciel devrait être aussi près de la réalité que possible même si la géométrie sera optimisée, pour deux raisons. La première est liée au gain de temps de calcul : plus la géométrie se rapproche du minimum énergétique de la DFT, moins d'étapes d'optimisation seront nécessaires à l'obtention de la bonne structure et un gain de temps sera réalisé. La seconde raison concerne la géométrie finale obtenue. En effet, même le meilleur algorithme solveur ne peut que trouver un minimum local d'énergie électronique par rapport aux coordonnées atomiques. Comme il est souhaitable de déterminer le minimum global en énergie pour les atomes donnés, la géométrie de départ doit être de la meilleure conjecture possible.

Dans la majorité des cas, et surtout si la molécule comporte de nombreux degrés de liberté, on effectue d'abord une étude conformationnelle pour choisir une géométrie de départ. Pour ce faire, on crée des géométries de départ différentes, sur lesquelles on effectue un premier calcul d'optimisation, à un niveau de théorie légèrement inférieur à celui visé par le calcul final. Selon le cas on choisit la ou les structures les plus basses en énergie, qu'on relance avec le niveau de théorie souhaité pour l'obtention des données désirées.

1.3.7 Le calcul de spectres Raman théoriques

Il est connu que les molécules vibrent, mais comment (bien) modéliser les modes vibrationnels avec leur fréquence, et l'intensité spectrale de ces modes vibrationnels?

1.3.7.1 Méthode dynamique

L'approche qui semble la plus intuitive sur le plan théorique pour calculer les vibrations moléculaires en est actuellement à ses débuts, parce qu'elle demande des simulations très longues à calculer. Elle sera brièvement exposée pour la postérité, puisque l'idée semble prometteuse.

On peut se dire que la solution logique au problème de représenter un mouvement vibrationnel qui se produit de lui-même dans la nature est de simplement laisser les molécules vibrer d'elles-mêmes, et

observer leur comportement. C'est le postulat proposé par la méthode dynamique.(Thomas *et al.*, 2013) On place plusieurs copies de la molécule optimisée dans une « boîte », et on fait une simulation de dynamique moléculaire, idéalement par premiers principes (*ab initio molecular dynamics*, AIMD). Il s'agit de la version QM de la dynamique moléculaire, elle-même basée sur la MM, ou mécanique moléculaire newtonienne. La simulation doit d'abord être thermostatée - puisque les calculs HF et KS DFT sont généralement faits à 0K – et on calcule ensuite le système à divers moments subséquents dans le temps, à une résolution temporelle donnée.(Luber *et al.*, 2014) On extrait enfin la trajectoire des atomes à partir des coordonnées de chaque étape dans le temps, et on effectue une analyse statistique du mouvement de tous les atomes en même temps, pour extraire les fréquences vibrationnelles obtenues de façon explicite. Des logiciels comme TRAVIS existent déjà pour faire le traitement post-simulation et générer un spectre Raman.(Brehm *et al.*, 2020) Pour obtenir les intensités des pics, on calcule le changement de polarisation dans le temps (et donc à travers le mouvement vibrationnel). Globalement, cette idée continue à être raffinée par quelques groupes pour faciliter le calcul avec les grappes informatiques contemporaines.(Aprà *et al.*, 2020)

La simulation est gourmande, puisqu'elle est effectuée à un niveau de théorie permettant une certaine précision sur la polarisabilité, et sur un système souvent grand. En effet, puisqu'une analyse statistique du mouvement est à la base des nombres d'ondes générés, il faut une certaine quantité de copies de la molécule qui évoluent ensemble pour avoir une représentation correcte des vibrations de l'échantillon, ce qui limite actuellement - aucune grande amélioration ni approximation n'a encore été développée – cette méthode à des molécules comportant une quinzaine à une vingtaine d'atomes chacune. Un des principaux avantages de cette méthode est l'inclusion intrinsèque des interactions intermoléculaires qui se produisent pour des solides organiques.(Kirchner *et al.*, 2021) On pourrait aussi spéculer comme application future la modélisation d'un échantillon organique amorphe contenant un faible pourcentage de solvant utilisé pour la purification, et les vibrations spécifiques en résultant.

1.3.7.2 Méthode statique

La méthode bien établie pour l'obtention de spectres Raman par DFT est basée en partie sur un modèle classique. En premier lieu, pour une géométrie optimisée où les distances interatomiques sont toutes à leur position la plus stable, on calcule une matrice hessienne – en DFT elle dépend de la double dérivée partielle de l'énergie par rapport aux coordonnées atomiques – avec approximation. En connaissant la masse de chaque atome, on peut utiliser les équations classiques de force et de l'oscillateur harmonique, pour transformer la matrice hessienne de sorte que ses valeurs propres représentent les fréquences de

déplacement sinusoïdales des atomes. Pour une description plus complète du processus mathématique, le cinquième chapitre du livre *Density Functional Theory: A Practical Introduction* traite du calcul de fréquence en plus grands détails qu'il ne serait pertinent d'amener ici. (Sholl et Steckel, 2009) Il devrait être mentionné que puisque le modèle de l'oscillateur harmonique est utilisé, une certaine déviation des fréquences est attendue, puisque la réalité démontre que la courbe d'énergie versus la distance interatomique est asymétrique, ce qui n'est pas le cas dans l'approximation harmonique couramment employée en DFT. Il est possible de réaliser les calculs avec un modèle anharmonique, (Lin *et al.*, 2021) mais la charge de calcul est plus grande. Souvent, on se contente d'appliquer un facteur de correction harmonique sur les fréquences calculées, permettant une modélisation acceptable des positions des pics. (Merrick *et al.*, 2007)

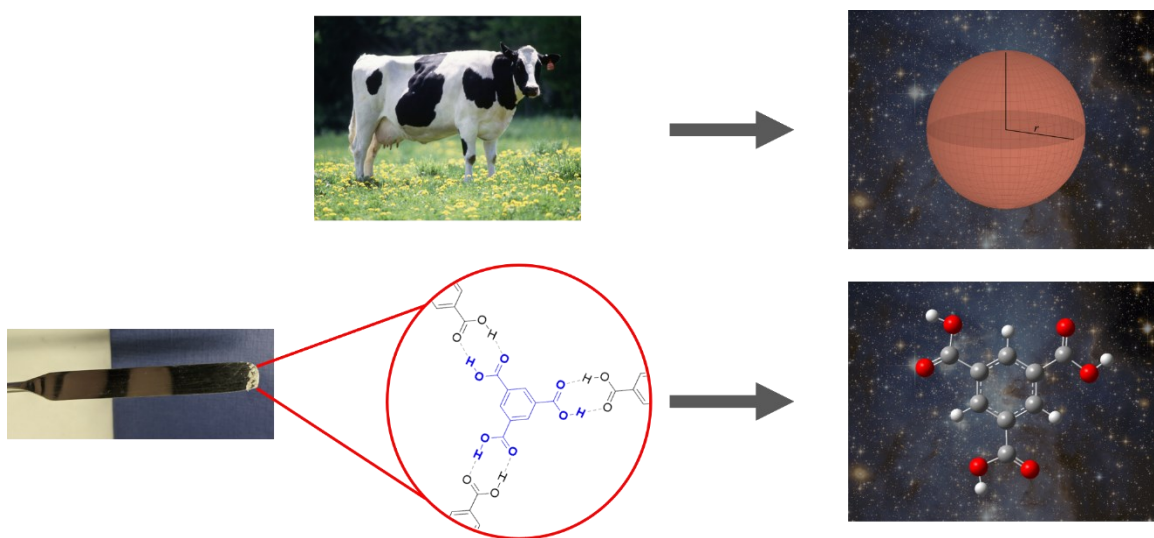
Une fois les fréquences vibrationnelles obtenues, il faut déterminer l'activité Raman de ces modes vibrationnels pour avoir l'intensité des pics prédits. Pour ce faire, il « suffit » d'évaluer la dérivée seconde de la polarisabilité par rapport aux coordonnées nucléaires à différents moments des mouvements vibrationnels. Sur ORCA, le logiciel utilisé dans ce projet, le calcul doit souvent être fait de façon numérique, et demande une certaine quantité de ressources de calcul – incluant le temps. Certains logiciels incorporent des approches indirectes comme l'idée d'évaluer le gradient d'énergie potentielle en présence d'un champ électrique externe. (Komornicki et McIver, 1979) En général, il semble que la prédiction des intensités Raman donne des résultats qui ne sont pas aussi précis que les fréquences vibrationnelles. (Zvereva *et al.*, 2011) D'un côté, il est fortuit que les fréquences soient bien prédites, parce que l'énergie de transition vibrationnelle est plus caractéristique que l'intensité relative des pics. Par contre, si l'intensité prédite est trop différente, il devient difficile de comparer un spectre théorique au spectre expérimental lui correspondant. Cette constatation est à l'origine des travaux présentés au Chapitre 2 sur la création de méthodes de corrélation spectrale entre un spectre Raman expérimental et un spectre calculé par DFT. La principale raison à l'origine de la qualité inférieure des prédictions d'intensités (par rapport aux nombres d'ondes qui sont bien prédits) est qu'elles viennent d'une propriété de troisième ordre, la double dérivée du tenseur de polarisabilité, ce qui signifie que l'exigence sur la justesse de la structure électronique est accrue, et les approximations de la DFT font que parfois ces intensités prédites dérogent un peu de la réalité.

1.3.7.3 Modélisation de l'environnement chimique

La DFT est en général assez bonne pour prédire les spectres Raman, lorsque le système est bien décrit. Une bonne description d'un système moléculaire comprend certes le choix des approximations incluses

pour le calcul (le niveau de théorie), mais attardons-nous sur le choix de ce qui est modélisé. Beaucoup de calculs en chimie numérique sont effectués sur une seule copie d'une molécule en phase gazeuse, dans un vide absolu. Une référence humoristique pour illustrer une faille de cette approche est le début d'hypothèse posé par certains physiciens : « postulons une vache sphérique de masse uniforme, dans le vide parfait de l'espace intersidéral » (Figure 1-4). La leçon à tirer de cette métaphore est que le choix de l'objet que l'on passe dans un modèle est tout aussi important que le choix du modèle lui-même. À ce sujet, il existe quelques technologies permettant d'avoir un « objet chimique » différent, pour permettre de mieux modéliser l'environnement de la molécule d'intérêt.

Figure 1-4 Caricature de la métaphore de l'objet modélisé appliqué aux solides moléculaires réels.



La prédiction des fréquences vibrationnelles est assez sensible à la géométrie de départ. Par conséquent, l'environnement chimique de la molécule est un enjeu dans le calcul DFT des spectres Raman, tel qu'il est expliqué précédemment dans cette section. Or dans le cadre de ce projet, l'échantillon-type ciblé par l'analyse Raman est un solide, soit amorphe ou cristallin. Il va de soi que – et c'est ce qui a été observé durant les essais préliminaires antérieurs à mes travaux - lorsque l'objet chimique modélisé ne comporte qu'une seule molécule en phase gazeuse, il ne peut y avoir d'interactions intermoléculaires. Puisque la DFT est une méthode ayant un certain coût computationnel, des méthodes de compensation existent.

Solvatation implicite

Des calculs en chimie numérique qui ne sont pas forcément basés sur l'obtention de spectres vibrationnels ont rapidement eu besoin d'une façon de représenter les interactions entre le système atomique et un

solvant, souvent pour être plus fidèle à une expérience réalisée en laboratoire, parfois pour faciliter un processus comme le transfert de charge. La littérature sur le sujet est assez bien cimentée, et des revues accessibles couvriront les modèles de solvation communs pour les intéressés.e.s.(Tomasi *et al.*, 2005) Un des modèles implicite les plus populaires, C-PCM (*Conductor-like Polarizable Continuum Model*),(Barone et Cossi, 1998) consiste à placer une surface autour de la molécule (le continuum) qui peut se polariser, et affecter les électrons de la molécule. Ainsi, en modulant la constante diélectrique du continuum, on peut simuler l'effet de polarisation de divers solvants. Cette méthode est une des plus rapides puisque seulement les parties de la molécule faisant face au continuum auront leur interaction modélisée (« *solvent-accessible surface* »). Cependant, son impact sur la modélisation de spectres Raman est plutôt faible : puisque seules les interactions électrostatiques avec le solvant sont modélisées,(Mennucci *et al.*, 2002) l'effet sur la polarisabilité, qui détermine les intensités, est très faible, et son effet sur la géométrie varie selon le système étudié.

Solvatation explicite

Comme on pourrait s'y attendre, le fait de placer des molécules de solvant directement dans le modèle et de les simuler au même niveau de théorie que le sujet principal donne de très bons résultats pour les spectres Raman,(Hopmann *et al.*, 2011) mais demande un temps de calcul considérable, et rend la convergence du SCF plus difficile. L'approche appelée *cluster-in-a-liquid* propose de modéliser une seule molécule d'intérêt, avec quelques molécules de solvant explicites.(Perera *et al.*, 2016)

Modèles multiéchelle

Un compromis raisonnable est d'incruster des particules pouvant se déplacer, et ainsi « participer » plus activement à la géométrie de la molécule à l'étude, sans les modéliser en pleine KS DFT. Les modèles multiéchelle représentent précisément cette idée : tout le système incluant les molécules de solvant est modélisé en mécanique moléculaire (MM), mais en plus, le sujet principal est modélisé en mécanique quantique (QM). Par arithmétique on arrive à une énergie QM/MM.(Groenhof, 2013) Les molécules de solvant sont donc présentes, et peuvent se déplacer, mais sont modélisées comme des charges ponctuelles. La QM/MM supplémentée d'une couche polarisable le C-PCM a été démontrée à quelques reprises,(Curutchet *et al.*, 2009) avec des résultats prometteurs, qui en ferait une méthode intéressante si elle ne nécessitait pas l'intervention d'un théoricien formé spécialement.(Steindal *et al.*, 2011) Il est connu que les modèles multiniveaux demandent un minimum d'attention dans leur préparation, et ne sont pas encore des méthodes dites « boîte noire ».(Clemente *et al.*, 2010) Les modèles de solvation implicite et explicite sont les seuls qui sont communément implémentés avec un calcul de spectre vibrationnel par

méthode statique; les modèles multiéchelle peuvent en théorie être calculés de façon dynamique, mais ne le sont que rarement. Pour une comparaison entre certaines méthodes basées sur les calcul dynamique et statique, une revue parue dans *Advanced Theory and Simulation* pourra assouvir la curiosité du lectorat.(Kirchner *et al.*, 2021)

1.4 Conclusion du chapitre

La microscopie Raman est une technique de haute sensibilité, et non-destructrice. Pour obtenir une prédiction DFT du spectre Raman d'une molécule, un minimum de compréhension de la DFT est requis pour choisir un niveau de théorie qui fait un bon compromis entre le coût et la qualité. L'utilisation d'une méthode DFT un peu plus clé-en-main sera couverte dans l'article soumis et réimprimé au Chapitre 3.

Lorsqu'un.e chimiste a une paire de spectres (un expérimental et un théorique), il conviendra de déterminer leur niveau de correspondance. C'est justement le sujet du prochain chapitre.

CHAPITRE 2

MÉTHODES DE CORRÉLATION QUANTITATIVE DE SPECTRES RAMAN

Ce chapitre couvre le fonctionnement de deux logiciels développés que j'ai créés en amont des travaux présentés dans l'article, et permettant d'effectuer des comparaisons quantitatives de spectres. Après quelques éléments de théorie à connaître et un survol littéraire traitant spécifiquement de corrélation entre spectres Raman, chaque logiciel sera présenté, puis une conclusion commune aux deux logiciels fermera le chapitre.

Puisque les méthodes de DFT actuelles modélisent bien la position des pics sur le spectre (leur nombre d'ondes), mais rencontrent certaines difficultés intrinsèques à bien prédire l'intensité relative de ces pics, un certain biais humain est compris dans l'évaluation qualitative de la correspondance entre un spectre expérimental et un spectre généré par la DFT. Pour compenser ce biais et aussi favoriser une analyse répétable, deux méthodes d'obtention d'un indice de corrélation spectrale chiffré sont ici testées. J'ai été auteur de la première méthode - présentée à la section 2.3 -, et du logiciel servant d'implémentation du concept, iROMA. La deuxième méthode décrite en 2.4 a été adaptée d'un algorithme existant, avec une étape de prétraitement que j'ai créée sur mesure pour les données Raman. Le logiciel que j'ai écrit pour tester cette méthode s'appelle SARA.

2.1 Éléments de traitement des signaux spectroscopiques vibrationnels

Afin de familiariser le lectorat avec certaines techniques de transformation des données spectrales employées par les deux logiciels, cette section regroupe les explications et les mathématiques liées à ces techniques de prétraitement.

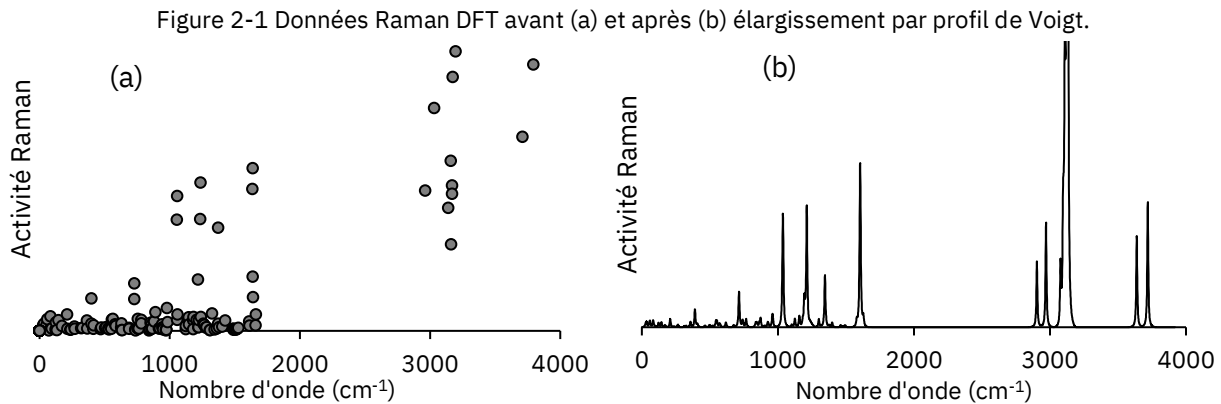
2.1.1 Élargissement des pics ponctuels

La chimie numérique fournit des spectres théoriques sous la forme d'un ensemble de coordonnées $(\tilde{\nu}_{DFT}, i_{DFT})$ où $\tilde{\nu}_{DFT}$ est le nombre d'onde en cm^{-1} , et i_{DFT} , l'intensité du pic associé (sans unités). Or, un spectre expérimental est constitué de pics ayant une certaine largeur, et souvent doté « d'épaules ». Pour les deux logiciels, il est pertinent de transformer le spectre théorique pour lui donner un aspect semblable à un spectre obtenu sur un instrument.

Lors de mes tests, la méthode ayant donné qualitativement le résultat le plus proche d'un spectre réel est l'utilisation d'un profil de Voigt avec certains paramètres ajustés pour élargir chaque paire de coordonnées spectrales. Le profil de Voigt est la combinaison d'une distribution Gaussienne et d'une distribution Lorentzienne, et sera définie dans mes logiciels comme :

$$V = \int_{-60}^{60} \left(\frac{e^{-\tilde{\nu}^2/2\sigma^2}}{\sigma\sqrt{2\pi}} \frac{\gamma}{\pi(\tilde{\nu}^2+\gamma^2)} d\tilde{\nu} \right) \quad (2.1)$$

avec $\sigma = 1$ et $\gamma = 4$, soit les coefficients déterminant la proportion de chaque type de distribution dans la fonction finale, déterminés de façon empirique par simple étude visuelle. L'intégration est faite sur une plage de 121 cm^{-1} , puisqu'avec les paramètres de σ et γ utilisés, à l'extérieur de l'intervalle de -60 à 60 cm^{-1} la courbe est considérée largement assez asymptotique pour être coupée à cet endroit. À noter que si σ vaut 0, on obtient un profil purement gaussien, et si $\gamma = 0$, un profil purement lorentzien. Pour obtenir un spectre, une courbe de Voigt est générée pour chaque pic avec l'intensité i_{DFT} . L'ensemble des courbes de Voigt individuelles placées au bon endroit sur le spectre est additionné, et la résultante est le spectre théorique aux axes $(\tilde{\nu}_{Théo}, i_{Théo})$ prêt à être pré-traité tout comme un spectre expérimental. À la Figure 2-1 se trouve le résultat de cette transformation sur des données Raman obtenues par DFT.



2.1.2 Correction harmonique

Tel que mentionné à la section 1.3.7.2, l'approximation harmonique du calcul des fréquences des modes vibrationnels implique des résultats légèrement divergents de la réalité. L'application d'un simple facteur de correction produit systématiquement un spectre calculé plus fidèle à l'expérimental.

$$\tilde{\nu}_{Corrigé} = k_{corr} \tilde{\nu}_{Initial} \quad (2.2)$$

Ici, le facteur k_{corr} prend une valeur qui est généralement près de l'unité, souvent entre 0,96 et 1,02. Par exemple, pour le niveau de théorie B3LYP/6-311G* l'Institut National des Standards et Technologie (NIST) recense sur sa base de données CCCBDB (cccbdb.nist.gov) un facteur de 0,9663, arrondi dans le cadre de mes travaux à 0,970.

2.1.3 Rééchantillonnage

Cette étape appliquée de la même façon pour tous les types de spectres analysés dans ce travail permet de prendre des spectres ayant une résolution différente, ou une résolution qui change à travers un spectre, et de générer un spectre « identique » ayant une résolution constante, et fixée dans le cadre de ce projet à 1 cm^{-1} . Le rééchantillonnage permet d'avoir des spectres directement comparables point-par-point, ce qui est nécessaire à la deuxième méthode de corrélation utilisée (section 2.4). Dans le cas de ces travaux, les nouveaux points sont calculés si nécessaire par interpolation linéaire entre les deux points les plus proches.

2.1.4 Troncature

La troncature consiste à couper les spectres de façon à les obtenir dans une fenêtre spectrale ayant les mêmes dimensions. En d'autres mots, pour chaque point en nombre d'onde dans un des deux spectres, il doit y avoir un point d'intensité connue au même nombre d'onde pour l'autre spectre. La fenêtre spectrale conservée pour une paire de spectres donnée s'étend donc du nombre d'onde minimum le plus élevé des deux, au maximum le plus bas (on garde l'intersection des spectres $A \cap B$).

2.1.5 Normalisation

Chaque spectre Raman traité a de grandes chances d'avoir une échelle d'intensités différente. Il est de mise de rétablir les données sur une échelle commune aux spectres à mettre en corrélation. Quelques méthodes sont employées pour la spectroscopie vibrationnelle, (Lasch, 2012) mais celle qui a été choisie et qui sera implémentée est une des plus simple à calculer et à interpréter : la normalisation min-max.

$$i' = \frac{i - i_{min}}{i_{max} - i_{min}} \quad (2.3)$$

2.2 Corrélation quantitative de signaux spectraux – méthodes connues

Les méthodes communes de corrélation quantitative de spectres vibrationnels sont souvent conçues pour la recherche spectrale dans une base de données où un spectre-requête est comparé à des dizaines de

milliers de spectres possibles avec un algorithme simple, permettant de générer un classement des composés les plus probables.(ATSM, 2015) Cela implique que les spectres comparés sont de même nature (expérimentale) et auront une allure très similaire.(Li *et al.*, 2006) Certaines méthodes communes comme la distance euclidienne, (Chu *et al.*, 2014) l'indice de Jaccard,(Garg *et al.*, 2015) l'indice de Tanimoto,(Bajusz *et al.*, 2015) et les méthodes basées sur un indice de qualité - HQI, « *Hit Quality Index* », (Galeev *et al.*, 2019; Gryniewicz-Ruzicka *et al.*, 2012; Park *et al.*, 2017; Rodriguez *et al.*, 2011; Zhang *et al.*, 2021) ont un problème en commun pour cette application. Afin de permettre une identification rapide (en quelques secondes) du spectre parmi toutes les possibilités que contient la base de données, les systèmes de pointage doivent donner un résultat rapidement, pour une paire de spectres donnés. Ainsi, tous les indices de similarité mentionnés dans la liste ci-haut sont des méthodes calculées point par point, c'est-à-dire que pour chaque nombre d'onde, l'intensité entre les deux spectres est comparée. La performance des meilleures de ces méthodes est très bonne pour la recherche spectrale en librairie, mais pour comparer un spectre calculé par DFT avec un spectre expérimental, un algorithme de notation doit être en mesure de « regarder sur les côtés » de chaque pic pour tenir compte de la prédiction DFT qui peut donner un décalage de part et d'autre d'un pic expérimental.

À cette fin, une approche a été développée (section 2.3), et une autre adaptée (section 2.4) pour le calcul de scores de corrélation.

2.3 iROMA : Approche manuelle

iROMA, pour “*interactive Raman Organic Match Assessment*” (évaluation interactive de correspondance organique par Raman) est la méthode que j'ai créée en premier, basé sur un calcul assez simple permettant de quantifier la correspondance entre deux spectres Raman, peu importe leur provenance. Afin d'alléger cette section, le code du programme utilisé comme implémentation d'iROMA se trouve à l'Appendice A.

2.3.1 Principes

Pour une paire de spectres, il est aisé pour un humain de voir directement si un pic donné dans un spectre possède un pic homologue dans l'autre spectre, et si les dimensions des pics homologues se ressemblent. C'est ce qu'iROMA tente d'émuler, en partant de connaissances spécifiques sur le spectre, comme le nombre de paires de pics (n), qui est évaluée durant le processus. La note donnée à une paire de spectres ayant n paires de pics homologues (p) dépend de trois éléments :

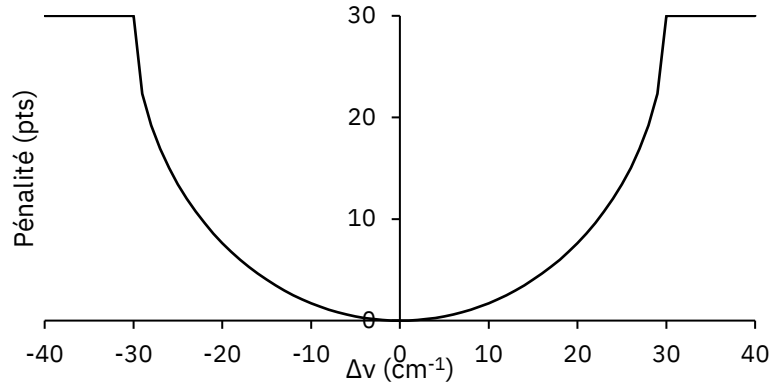
1. Le nombre total – pour les deux spectres - de pics n'ayant pas d'homologue (qu'on appellera dorénavant le nombre de pics « seuls », ou m , qui est ajusté par 0.5 pour diminuer la pénalité si un seul pic est seul parmi les deux spectres). La sous-note pour les pics seuls est donnée par l'équation :

$$s_{seuls} = \left(1 - \frac{(m-0.5)}{n}\right) \quad (2.4)$$

C'est donc la quantité de pics seuls relative au nombre de paires de pics n dans la paire de spectres qui est prise en compte.

2. La distance entre les pics d'une paire d'homologues se trouvant chacun dans un des spectres, $\Delta\tilde{\nu}_p$ (en cm^{-1}), est évaluée par rapport à une limite maximale $\Delta\tilde{\nu}_{lim}$, fixée à 30 cm^{-1} . La distance entre les pics d'une paire $\Delta\tilde{\nu}_p$ est placée dans une fonction mathématique qui renvoie une pénalité selon la grandeur de la distance. La fonction d'hypoténuse selon Pythagore a été choisie pour son profil représentant une pénalité douce à de faibles distances, et d'intensité grandissante jusqu'à devenir drastique, puis totale à une distance de $\Delta\tilde{\nu}_{lim}$.

Figure 2-2 Profil de la courbe de pénalité pour la distance entre pics homologues



Pour obtenir le sous-score de déplacement, on soustrait la somme des pénalités à un bassin de « points » obtenu en multipliant $\Delta\tilde{\nu}_{lim}$ au nombre de paires de pics n . L'équation de sous-score résultante est :

$$s_D = \frac{n\Delta\tilde{\nu}_{lim} - \sum_{p=1}^n \left(\Delta\tilde{\nu}_{lim} - \sqrt{\Delta\tilde{\nu}_{lim}^2 - \Delta\tilde{\nu}_p^2} \right)}{n\Delta\tilde{\nu}_{lim}} \quad (2.5)$$

3. La différence d'intensité entre les pics d'une paire d'homologues Δi_p est évaluée selon le même principe que le déplacement, mais avec une pénalité augmentant de façon linéaire. Ici le seuil d'acceptabilité Δi_{lim} est fixé à 0,65 (sur une possibilité de 1) puisque tel qu'il a été expliqué, la DFT est moins précise dans sa prédiction des intensités Raman.

$$S_I = \left(\frac{n\Delta i_{lim} - \sum_{p=1}^n |\Delta i_p|}{n\Delta i_{lim}} \right) \quad (2.6)$$

On donne donc une sous-note pour chacun des éléments pris en compte, et on fait une moyenne pondérée des trois sous-notes :

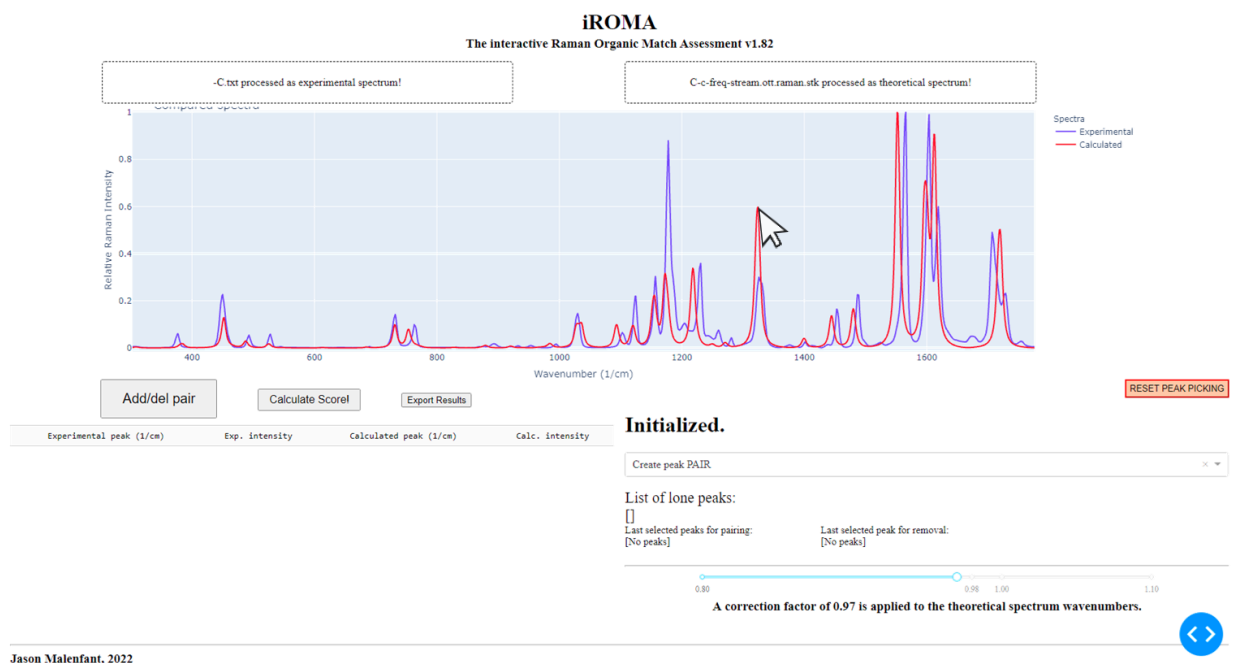
$$S_{iROMA} = w_D(S_D) + w_{seuls}(S_{seuls}) + w_I(S_I) \quad (2.7)$$

En mettant les poids optimisés de façon empirique pour chaque composante, et en insérant les équations (2.4), (2.5) et (2.6) dans l'équation (2.7) on obtient l'équation de score pour une paire de spectres:

$$S_{iROMA} = 0.40 \left(\frac{30n - \sum_{p=1}^n \left(30 - \sqrt{900 - \Delta \tilde{\nu}_p^2} \right)}{30n} \right) + 0.55 \left(1 - \frac{(m-0.5)}{n} \right) + 0.05 \left(\frac{0.65n - \sum_{p=1}^n |\Delta i_p|}{0.65n} \right) * 100\% \quad (2.8)$$

Une personne attentive aura donc compris qu'un certain nombre d'informations complexes à extraire de façon automatisée sont à connaître à propos des données qui sont analysées. Pour tester l'approche, un logiciel de preuve de concept avec une interface permettant l'utilisation d'un être humain comme système de sélection des pics (« *peak picking* ») a été créé avec le langage Python. Son utilisation est maintenue simple pour en permettre l'accessibilité : l'application fonctionne dans un navigateur web, et on peut y glisser un fichier de spectre expérimental, et un fichier de spectre théorique. Les spectres sont traités selon les étapes décrites à la section 2.1, avec le facteur de correction harmonique pouvant être modifié avec le déplacement d'une jauge à glissière (« *slider* »). L'humain clique sur un pic du spectre expérimental, puis sur le pic homologue du spectre théorique, suivi d'un bouton permettant de valider la création de la paire de pics. Un outil permet de changer le type de sélection, pour cliquer sur l'ensemble des pics seuls contenus dans les spectres. Une fois la sélection des pics complétée, la pression d'un dernier bouton permet de faire le calcul et générer la note accordée à la correspondance des deux spectres.

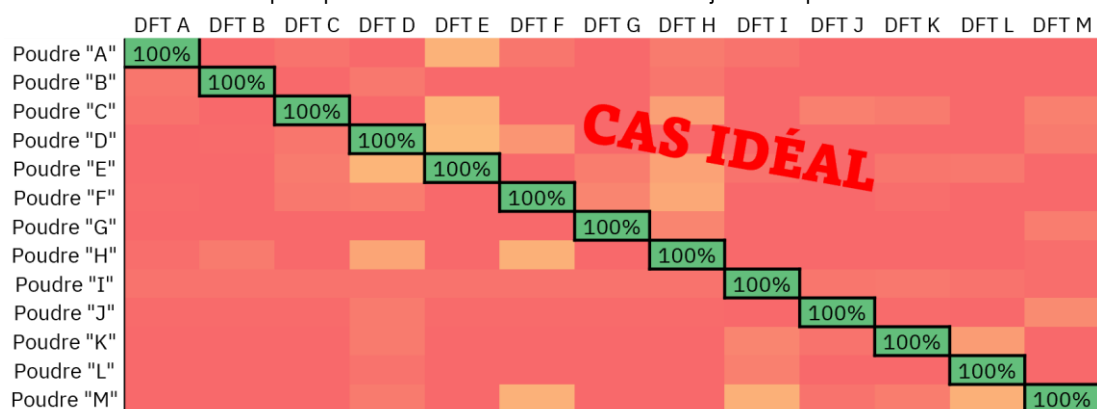
Figure 2-3 Interface utilisateur du logiciel de traitement spectral iROMA.



2.3.2 Résultats et discussion

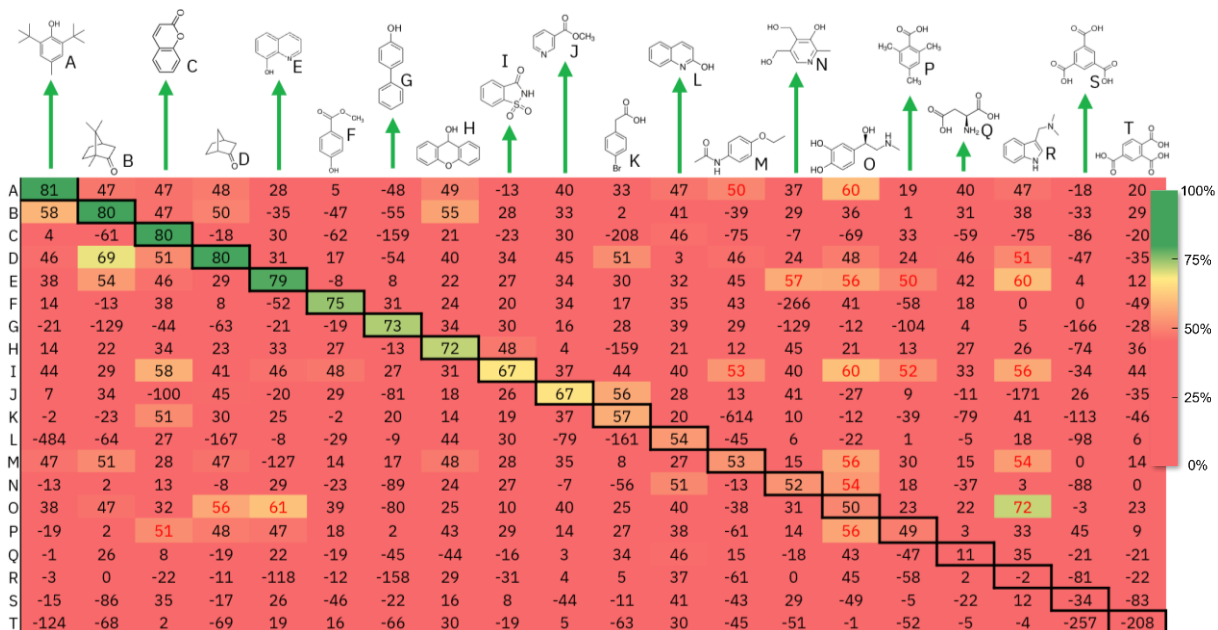
Pour évaluer la performance du logiciel iROMA, un jeu de données a été généré en double aveugle en prenant un ensemble de composés (initialement identifiés avec seulement un symbole) représentant une certaine diversité de groupes fonctionnels, qui seraient des modèles réduits de molécules d'intérêt en chimie organique. Un spectre Raman a été acquis sur les poudres avec les paramètres donnés à l'Annexe C, et un spectre théorique a été calculé pour chacune des substances analysées (données en Annexe A). Le logiciel iROMA a été utilisé pour évaluer la correspondance de chacun des spectres expérimentaux avec chacun des spectres théoriques. Cela signifie que pour 20 molécules, 40 spectres ont été traités, deux à la fois, et 400 scores de corrélation ont été générés. Pour afficher ces résultats de façon compacte, une matrice de corrélation peut être construite, avec chaque ligne représentant un spectre expérimental, et chaque colonne, un spectre théorique. Ainsi, une donnée dans la matrice est le score iROMA entre le spectre expérimental de la ligne, et le spectre théorique de la colonne. Idéalement, la ligne diagonale parcourant la matrice de la cellule (0,0) à la cellule (n,n), qui représente les scores des « bons *matches* » (les correspondances réelles) devrait avoir des notes approchant 100%, et toutes les valeurs à l'extérieur de cette diagonale, des notes très basses. Voici une représentation visuelle de ce scénario idéal :

Figure 2-4 Matrice de corrélation fictive pour le scénario ayant une prédiction parfaite des spectres Raman, et une note qui représenterait la corrélation avec une justesse parfaite.



En parallèle de l'acquisition des spectres Raman expérimentaux avec l'instrument Renishaw inVia, des calculs DFT ont été effectués avec le logiciel Gaussian au niveau de théorie B3LYP/6-311+G(2d,2p), en phase gazeuse (GP). Ce choix était informé par une contrainte : la puissance de calcul accessible pour ce projet durant la première phase de recherche et de développement était limitée à deux ordinateurs octocœurs, ce qui est très peu. La fonctionnelle B3LYP demeure un bon choix de qualité, mais la base utilisée était le résultat d'un compromis entre la puissance de calcul utilisée, et les choix disponibles dans la version 9 de Gaussian (d'où la base type Pople), seul logiciel de calcul DFT installé sur l'équipement de calcul disponible pour ce projet à l'époque.

Figure 2-5 Matrice de notes de corrélation iROMA (en pourcentages) pour 20 molécules test. Chaque colonne représente un spectre calculé à B3LYP/6-311+G(2d,2p), et chaque ligne, un spectre expérimental acquis dans les conditions décrites à l'Annexe C. Les notes dans la région des mauvaises correspondances qui sont plus élevées qu'une des bonnes correspondances (dans la diagonale) sont indiqués en rouge.

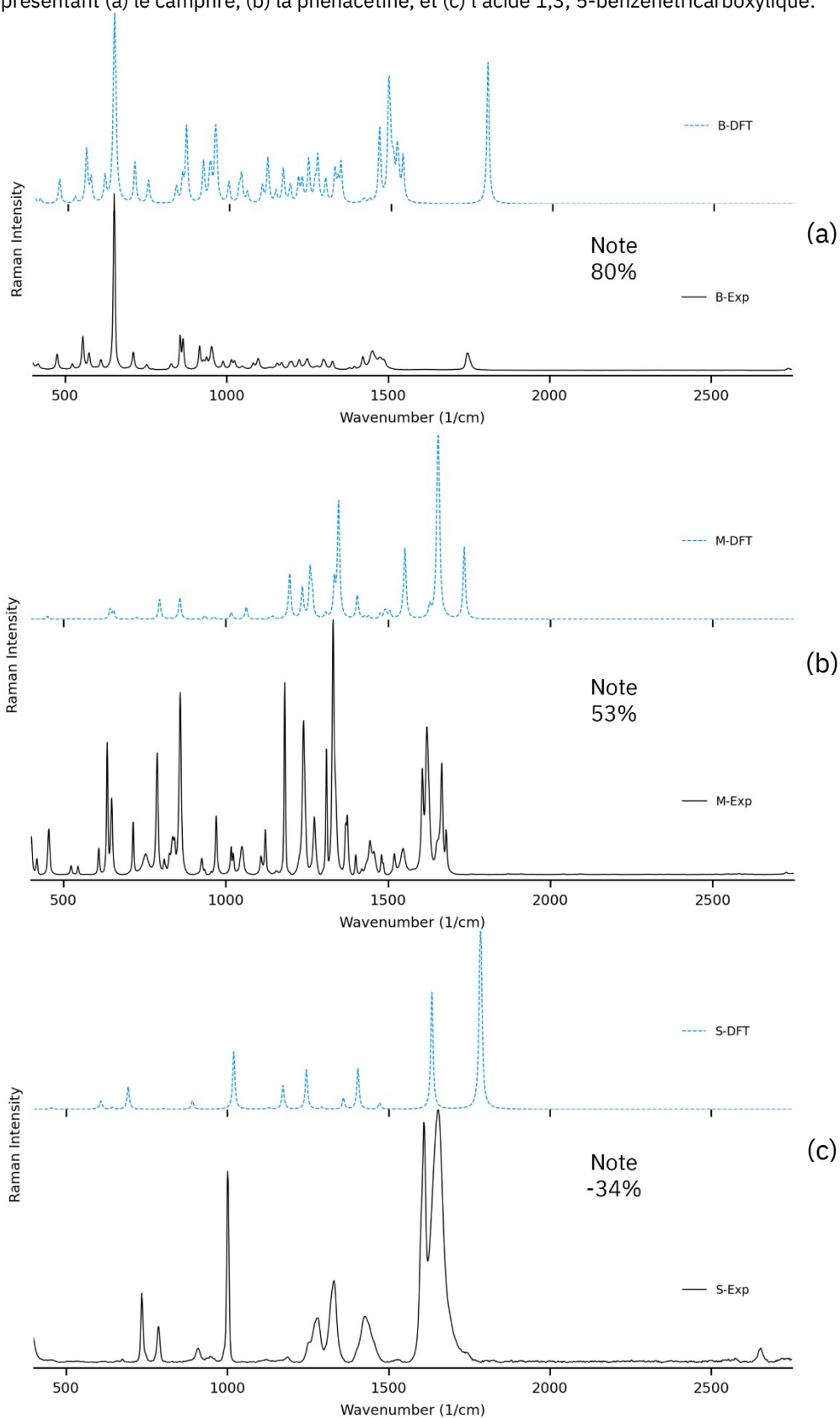


Comme le montre la Figure 2-5, certaines des molécules choisies sont bien prédites. Néanmoins, les personnes ayant fait une lecture attentive de l'introduction auront remarqué que certaines des molécules utilisées devraient manifester des interactions intermoléculaires très fortes, que la DFT sur une molécule en phase gazeuse aura de faibles chances de reproduire. C'est ce que l'on voit particulièrement dans le tiers à droite de la matrice. Ce choix a été fait pour tester les limites d'une méthode « classique » d'un certain âge avec les contraintes matérielles et logicielles existantes. Le résultat peut sembler insatisfaisant à première vue, mais ces données permettent certains apprentissages fort pertinents pour la suite.

De plusieurs observations à faire concernant ces données, la plus frappante est la diagonale de correspondances réelles qui comprend de nombreux faux négatifs. On constate également plusieurs faux positifs à l'extérieur de la diagonale. Pour comprendre d'où vient le problème, il convient de regarder quelques paires de spectres analysées avec le logiciel. À noter que tous les spectres acquis et calculés sont accessibles à l'Annexe B. La Figure 2-6 en montre trois qui donnent des scores à différents endroits de la diagonale; soit le camphre (molécule B) à 80%, la phénacétine (molécule M) à 53%, et l'acide 1,3,5-benzènetricarboxylique (molécule S) à -34%. Ces trois paires de spectres ont été choisies parce qu'elles montrent bien les éléments diagnostiques à souligner pour discuter de la méthode. L'ensemble des paires de spectres de la diagonale se retrouvent à l'Annexe B, avec les conditions d'acquisition des spectres expérimentaux. Pour le camphre, on peut constater d'un seul coup d'œil un spectre DFT qui donne une bonne représentation du spectre expérimental de la poudre. Hormis dans la zone allant de 1000 à 1500 cm^{-1} , où quelques rares pics sont absents du spectre DFT si on scrute à la loupe, on a une correspondance satisfaisante. Tel qu'attendu, les intensités relatives de certains pics ne sont pas parfaitement modélisées. Une note de 80% paraît qualitativement un bon jugement de la corrélation spectrale, quoique légèrement sévère. Cette paire de spectres montre la puissance de la DFT dans de bonnes conditions : on peut obtenir des résultats d'une grande qualité si les conditions sont réunies.

Pour la phénacétine, la situation se complique. Il manque une certaine quantité de pics, ou leur intensité est telle qu'on ne les aperçoit qu'à peine. Quelques pics sont présents mais apparaissent avec un certain décalage. Il y a tout de même certains pics qui sont au bon endroit. Encore une fois, une note de 53% semble être un bon jugement de la corrélation spectrale. Globalement cette paire ne devrait pas avoir une note lui conférant le statut de « positif ».

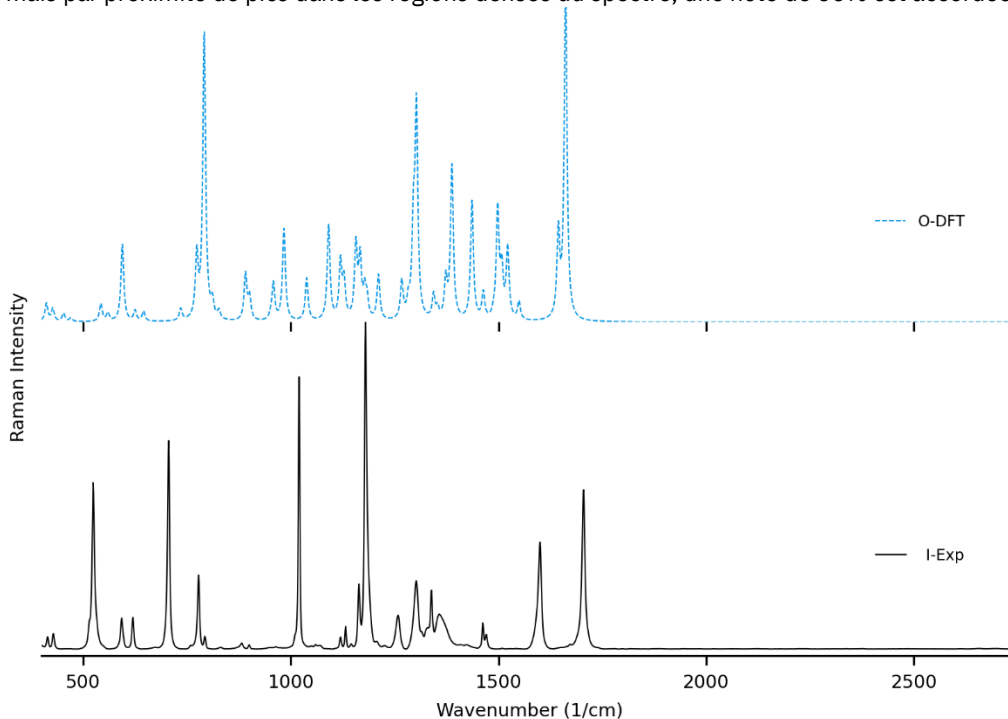
Figure 2-6 Données spectrales expérimentales (noir, trait plein) et théoriques (trait bleu pointillé) sélectionnées représentant (a) le camphre, (b) la phénacétine, et (c) l'acide 1,3, 5-benzènetricarboxylique.



Pour l'acide carboxylique choisi, on ne compte que trois pics qui sont approximativement au bon endroit : ceux à environ 1000 cm^{-1} , 1420 cm^{-1} , et 1600 cm^{-1} . Le reste est clairement mal modélisé, ce qui est facile à expliquer. Dans un système réel, la géométrie d'une molécule individuelle dépendra fort probablement d'un réseau étendu d'acides faisant des ponts hydrogènes très forts. Or, la position la plus stable pour les protons et les oxygènes impliqués dans le système unimoléculaire sera forcément différente si aucune copie du triacide ne peut interagir avec la molécule. Pour cette paire, une note de -34% (qu'on peut poser comme synonyme de 0%) paraît un jugement correct de la corrélation spectrale – ou l'absence de celle-ci en l'occurrence. Puisqu'il y a très peu de pics modélisés, ceux qui sont au bon endroit pourraient dans d'autres cas l'être par chance, et il est adéquat de dire que la corrélation est approximativement nulle.

Abordons maintenant la quantité importante de nombres en rouge présents dans la matrice de la Figure 2-5. Ces valeurs sont des scores pour lesquels une valeur sur la diagonale de correspondances réelles est supérieure, constituant un quasi-faux positif. Ici, en comparant une nouvelle paire de spectres faisant partie de ces cas, en l'occurrence le spectre expérimental de la saccharine (I) avec le spectre théorique de l'adrénaline (O), on peut constater que les « empreintes digitales » ne sont pas les mêmes (Figure 2-7).

Figure 2-7 Données spectrales d'une mauvaise correspondance sélectionnée. Presqu'aucun pic n'est communs aux spectres, mais par proximité de pics dans les régions denses du spectre, une note de 60% est accordée à la paire.



Il se trouve que par coïncidence, certains pics homologues sont suffisamment près pour permettre une corrélation notée à 60%. Cette note se situe entre celles pour la correspondance réelle de la saccharine (à 67%) et la correspondance réelle de l'adrénaline (de 50%). Pour plusieurs indices de similarité, une valeur en deçà de 80% est considérée comme un négatif, notion que l'on peut vaguement tenter d'appliquer ici. Une grande part du problème semble être à nouveau la modélisation partiellement incorrecte des deux molécules.

Ainsi, les exemples montrés et l'ensemble des spectres de la matrice indiquent un certain niveau de succès de la mesure de similarité entre les spectres par iROMA, mais pointent vers un problème de modélisation. Dans un premier temps, le niveau de théorie employé ne comportait aucune correction des interactions de dispersion de longue distance, rendant un spectre moins fidèle de certains composés ayant de l'encombrement stérique intramoléculaire. Cela est une hypothèse qui n'explique que partiellement le fait que la note maximale attribuée à une corrélation n'était que de 81%, même si souvent les meilleures paires de spectres étaient très similaires. Une autre composante est l'intensité des pics qui semble dépendre plus fortement du processus de normalisation qu'initialement prévu, surtout lorsqu'un pic très intense « écrase » le reste du spectre – situation que le logiciel SARA tentera de corriger. Puisque la prédiction des intensités n'a qu'un poids de 5% dans la note iROMA, cette piste ne semble à nouveau pas la bonne. Il est possible que l'absence de calibration des nombres d'onde sur l'instrument lors de l'acquisition des spectres soit responsable d'une partie de l'erreur observée, ceci dit la déviation devrait être plus modérée s'il s'agissait de la seule cause. Une hypothèse aussi plausible serait de considérer le facteur de correction harmonique utilisé sur les spectres théoriques. En effet, la littérature sur le sujet préconise l'utilisation d'un seul facteur de correction par niveau de théorie, par souci de répétabilité. En observant plus attentivement les spectres de la Figure 2-6(a), on peut voir que la région à droite du spectre théorique possède un décalage relativement constant avec le spectre expérimental, même après l'application d'un facteur de correction. Cela est symptomatique d'un problème de fluctuation de l'erreur du modèle harmonique. Pour corriger cette difficulté, trois options sont possibles :

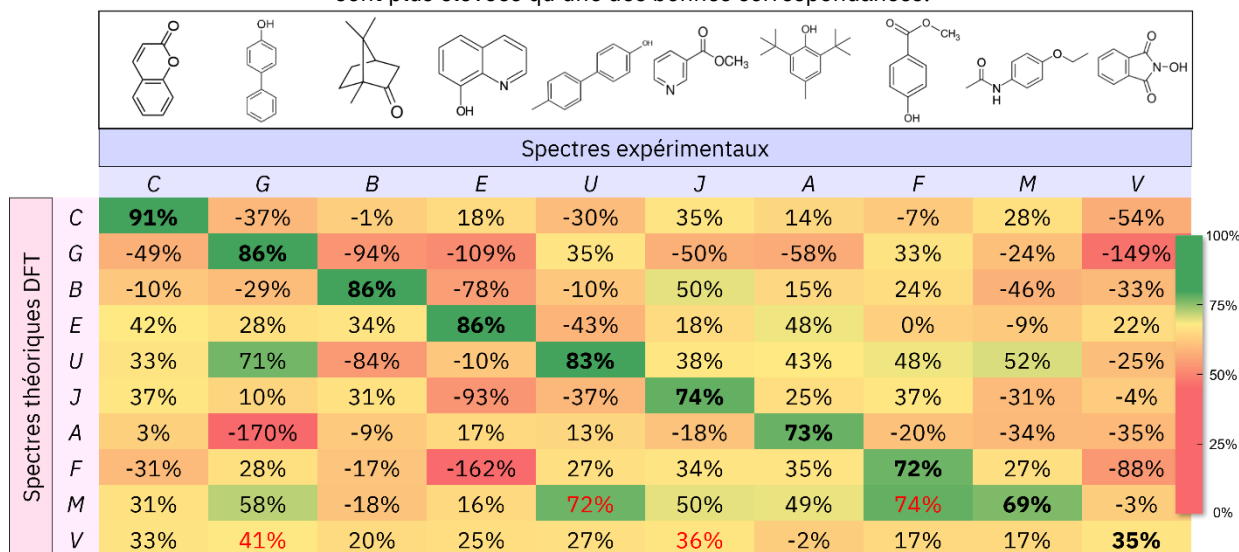
1. Exécuter le calcul avec un oscillateur anharmonique, complexité qu'il serait préférable de ne pas ajouter aux organiciens;
2. Ajuster le facteur de correction de façon individuelle pour chaque paire de spectre, ce qui cause l'ajout indésirable d'une variable ajustable permettant plus de souplesse vis-à-vis des scores;
3. Utiliser un niveau de théorie pour lequel le facteur de correction est plus stable.

Pour la méthode de calcul utilisée dans la publication faisant l'objet du Chapitre 3, la solution 3 est employée pour aider à améliorer la situation. Cependant, le niveau de théorie utilisé pour cette partie du projet n'existait pas au moment où iROMA était en utilisation.

Durant les dernières étapes de confection du logiciel SARA, un deuxième et dernier ensemble de données a été généré avec iROMA. Cette fois, un jeu de molécules modifié a été caractérisé, et de façon différente.

Afin de diminuer la présence d'interactions intermoléculaires, et donc de s'approcher du modèle computationnel, les spectres ont été acquis dans le solvant tétrahydrofurane (THF). Le choix de ce solvant a deux composantes : d'une part, il ne présente que deux signaux importants en Raman, ce qui facilite l'analyse spectrale, et réduit les interférences. D'autre part, c'est un solvant accessible, plutôt versatile, et ayant une certaine utilisation en chimie de synthèse. Concernant le choix des molécules, les triacides et le trialcool ont été abandonnés, puisqu'il sera difficile de bien modéliser ces systèmes avec les contraintes préalablement exposées. Également, les molécules trop peu solubles dans le THF ont été effacées, puisque la Raman en solution nécessite la concentration la plus élevée possible pour obtenir un signal adéquat. Pour arriver à une dizaine de composés, des ajouts ont été faits. La molécule U est une version méthylée de la G, et la V est un analogue de la I qui est plus soluble dans le solvant utilisé. Idéalement, U et G devraient être discriminées par leur spectre Raman.

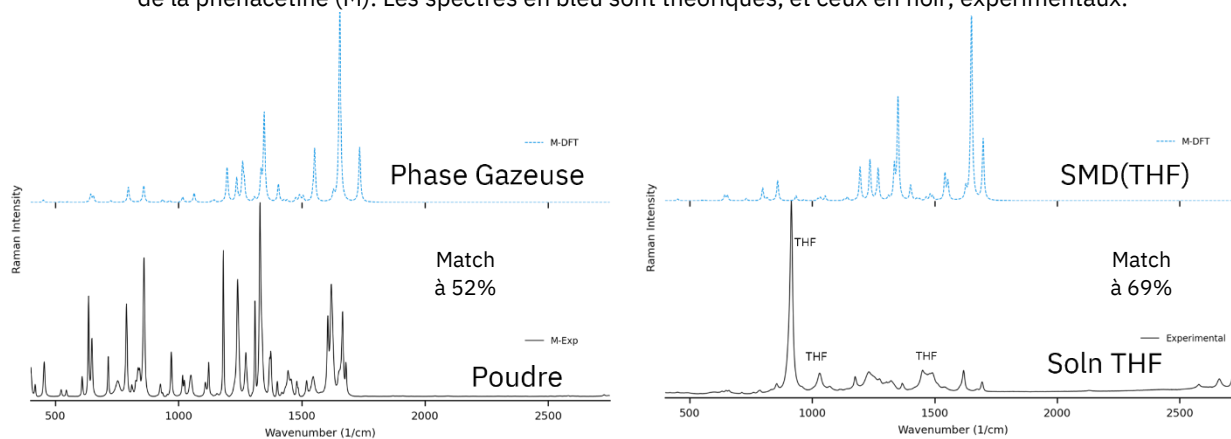
Figure 2-8 Matrice de corrélation pour un jeu de molécules modifié. Les spectres théoriques ont été calculés avec solvation implicite du THF, et les échantillons expérimentaux étaient solubilisés dans le THF. Figure reproduite du rapport de stage de Lucille Kuster, avec permission, en modifiant les identifiants des composés pour les harmoniser avec ce document. Les valeurs en rouge représentent les notes dans la région des mauvaises correspondances qui sont plus élevées qu'une des bonnes correspondances.



Concernant la chimie numérique utilisée, un modèle de solvation implicite a été utilisé afin d'en tester l'applicabilité. Le modèle populaire SMD, (Marenich *et al.*, 2009) une évolution de la CPCM qui prend en compte la densité électronique, a été employé. Le niveau de théorie B3LYP/6-311+G(2d,2p):SMD(THF) tel qu'implémenté dans Gaussian 9 a été utilisé. La Figure 2-8 détaille les corrélations spectrales obtenues de cette façon.

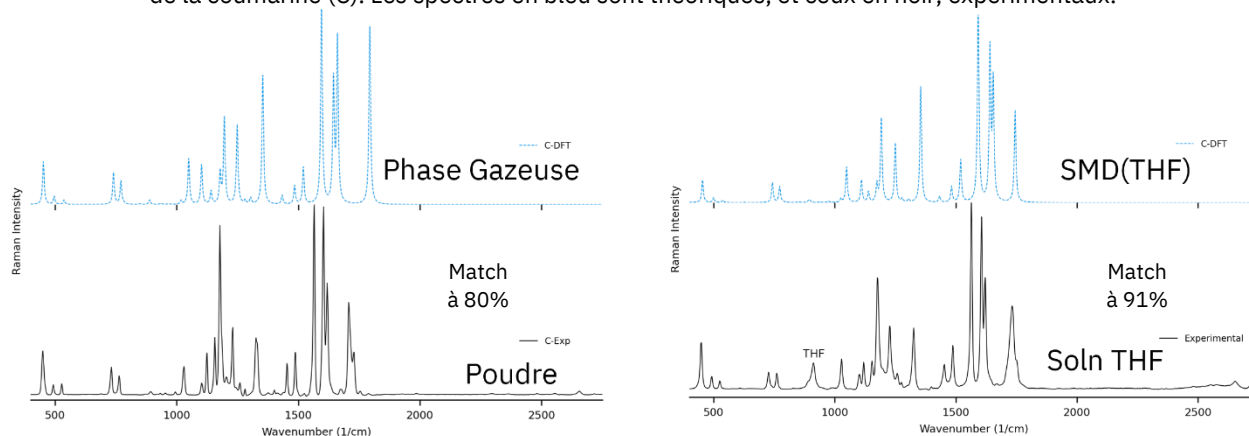
On remarque globalement de bonnes corrélations sur la diagonale des bonnes correspondances, avec quelques améliorations par rapport à la première étude en phase solide. Comme exemple, la phénacétine (molécule M) passe de 52% dans la première matrice de corrélation, à 69% dans la seconde. Dans ce cas, on observe une simplification du spectre expérimental lors du passage en solution, qui peut être expliquée par la diminution réussie des interactions intermoléculaires par rapport à la phase cristalline. Cette simplification du spectre diminue le nombre de pics seuls du spectre expérimental, réduisant la pénalité qui leur est associée. Les deux spectres théoriques (en bleu sur la Figure 2-9) sont qualitativement plutôt similaires, puisqu'une seule molécule est modélisée, et la solvation implicite SMD modifie peu la géométrie. La meilleure correspondance qualitative et quantitative semble dans ce cas être attribuable au fait qu'en solution, les interactions entre deux molécules de phénacétine sont plus rares, et forcément le spectre calculé à partir d'une seule molécule sera plus proche de cette réalité.

Figure 2-9 Comparaison spectrale pour les deux conditions expérimentales et les deux systèmes de modélisation de la phénacétine (M). Les spectres en bleu sont théoriques, et ceux en noir, expérimentaux.



Un avantage de l'analyse iROMA avec échantillons en solution est que les signaux Raman du solvant utilisé peuvent simplement être ignorés.

Figure 2-10 Comparaison spectrale pour les deux conditions expérimentales et les deux systèmes de modélisation de la coumarine (C). Les spectres en bleu sont théoriques, et ceux en noir, expérimentaux.



On remarque la même chose à la Figure 2-10 avec le meilleur match de la matrice réalisée en solution : le spectre expérimental est plus proche du spectre DFT. Une seconde fois, on constate que le modèle de solvation SMD ne change qu'assez peu l'apparence du spectre Raman.

Dernier aspect concernant la Figure 2-8, les « quasi-faux-positifs » et faux-positifs, identifiés en rouge, dont la quantité relative semble avoir légèrement diminué. Une hypothèse valable serait d'attribuer ce changement à la même raison pour laquelle les notes ont monté : une simplification des spectres expérimentaux permet moins de « surcorrélations » (« *overfitting* ») avec des pics qui sont absents dans le spectre calculé.

Malgré la performance acceptable d'iROMA, deux facteurs ont précipité l'usage du logiciel SARA, et favorisé l'arrêt de l'emploi d'iROMA. Le premier facteur, et le plus important, est la durée de l'analyse spectrale nécessaire pour l'obtention d'une grille de scores. La première matrice présentée, comportant vingt molécules (soit 400 corrélations) a nécessité plus de deux journées et demie de sélection de pics ininterrompue à l'auteur de ce document et du logiciel, qui avait considérablement d'expérience à son utilisation. La deuxième matrice présentée, composée de dix molécules (pour 100 corrélations), a nécessité plus d'une journée pour une tierce personne, qui était néophyte à l'utilisation d'iROMA. Clairement, la sélection de pics manuelle devrait être remplacée par un système automatisé. Il est possible que ce processus serait un problème qu'une intelligence artificielle pourrait être entraînée à résoudre, chose qui n'a pas été tentée, faute d'expertise technique. Les quelques algorithmes de « *peak picking* » testés nécessitaient tous l'intervention d'un utilisateur entraîné pour avoir les bons paramètres de sélection, et la sélection des pics seuls devait tout de même être faite par un humain. Globalement, ce

temps d'analyse est beaucoup trop élevé pour une méthode que l'on souhaite déployée pour faire de l'analyse de routine.

Le deuxième facteur ayant promu l'utilisation de SARA est que l'assignation manuelle des pics avec iROMA, qui devait être un processus assez aisé, s'est révélé contenir des ambiguïtés inattendues. Dans un certain nombre de régions spectrales, des pics coalescés présentaient une difficulté d'assignation par rapport au spectre homologue. Cette situation peut engendrer la création de paires de pics qui ne sont pas reproductibles d'un utilisateur à un autre, ce qui est problématique et peut influencer la note.

2.4 Logiciel SARA : Approche automatisée

Cette section servira de présentation rapide du logiciel SARA (*Similarity Assessment of Raman Arrays*) en Français. SARA fait l'objet de la publication constituant le Chapitre 3 de ce mémoire par article, donc la méthode sera présentée en Anglais à cet endroit. Ici, il sera possible d'aller dans de plus amples détails sur le logiciel lui-même que ne peut le faire l'article, pour des raisons de concision, mais l'article contient l'ensemble des résultats plus intéressants générés avec ce logiciel.

2.4.1 Principes

SARA est un programme Python, avec comme objectif un certain niveau d'automatisation et de parallélisation. Il exploite la puissante librairie de code numérique NumPy pour les opérations matricielles. Le programme prend comme entrées un répertoire contenant un fichier textuel pour chaque spectre, et un facteur de correction harmonique (qui peut être fixé à l'unité dans le cas où la correction est déjà appliquée ou n'est pas nécessaire). Les spectres théoriques et expérimentaux sont distingués soit par extension de fichier (qui sont données à SARA au préalable), ou par une suite de caractère dans le nom des fichiers (comme « exp », « theo », et autres). Les formats de données textuelles reconnus par SARA comprennent les CSV (*comma-separated values*, ou valeurs séparées par des virgules), les fichiers venant du logiciel Renishaw WiRE, et les spectres théoriques exportés par GaussView et ORCA. De ce point, le logiciel commence le pré-traitement et le calcul des notes à raison d'une paire de spectres par cœur du processeur. Les notes attribuées aux paires de spectres sont compilées dans un bloc de texte type CSV, qui peut être enregistré dans un fichier, ou importé directement dans un logiciel tableur.

La mesure de correspondance qui est centrale à ce logiciel est la corrélation croisée moyenne pondérée ou WCCA, pour « *weighted cross-correlation average* », (de Gelder *et al.*, 2001) d'abord développée pour l'analyse de signaux en diffraction des poudres par rayons X. Il a été déterminé que l'algorithme peut aussi

servir pour les spectres Raman, mais nécessite de l'adaptation pour l'analyse d'un spectre théorique avec un spectre expérimental. (Fairchild *et al.*, 2009) La WCCA consiste à créer une carte de la corrélation croisée CC entre l'intensité spectrale d'un point du spectre expérimental $i_{\tilde{\nu}}$ avec l'intensité spectral du point du spectre expérimental à une position décalée $i_{\tilde{\nu}+\Delta\tilde{\nu}}$, en fonction du décalage $\Delta\tilde{\nu}$ allant à gauche du point observé ($\Delta\tilde{\nu} < 0$), et à sa droite ($\Delta\tilde{\nu} > 0$) :

$$CC(\Delta\tilde{\nu}) = \int i_{\tilde{\nu}} j_{\tilde{\nu}+\Delta\tilde{\nu}} d\tilde{\nu} \quad (2.9)$$

Cette carte est pondérée en intensité par une fonction symétrique centrée à zéro $w(\Delta\tilde{\nu})$ qui peut ressembler à une gaussienne, ou plus simplement, une fonction linéaire « triangulaire » dont l'intensité maximale est au centre à $\Delta\tilde{\nu} = 0$. La pondération permet de diminuer l'apport des régions éloignées du point d'intérêt à l'intégrale d'intensité spectrale. En d'autres mots, elle agit comme un « filtre » qui pénalise progressivement le déplacement de densité spectrale à partir de la position du point d'intérêt. Dans le cas de SARA, la fonction de pondération prend la forme de l'Équation 2.10, qui a fait ses preuves dans l'article publié par Fairchild et collègues, avec $\alpha = 0.00001$ et $\beta = 4$ optimisés empiriquement par l'auteur de ce travail. Cette fonction a une intensité de 0 à un déplacement de $\pm 65 \text{ cm}^{-1}$.

$$w(\Delta\tilde{\nu}) = \frac{2}{\alpha |\Delta\tilde{\nu}|^{\beta+2}} \quad (2.10)$$

Au final, la note SARA attribuée à une paire de spectres découle de l'équation 2.11 :

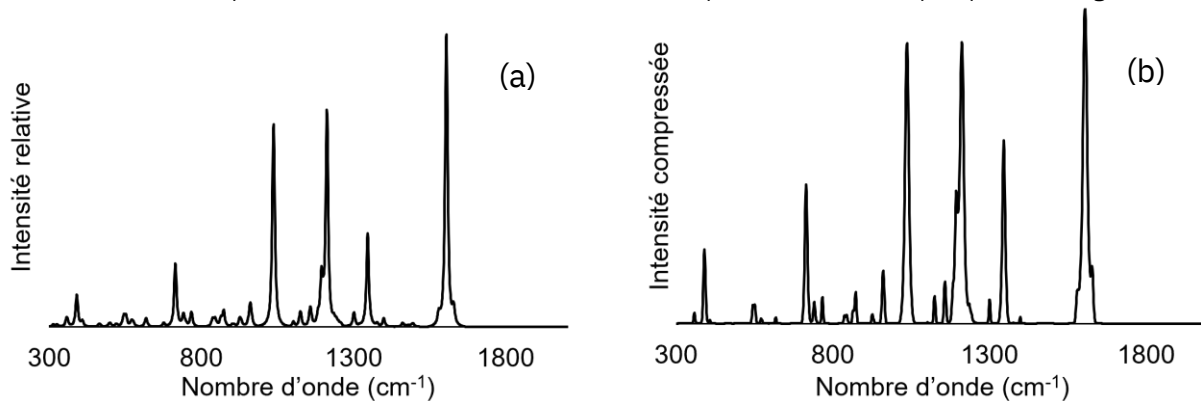
$$SARA \text{ Score} = \frac{\int_{-65}^{65} w(\Delta\tilde{\nu}) CC(\Delta\tilde{\nu}) d(\Delta\tilde{\nu}) * 100\%}{\sqrt{\int_{-65}^{65} w(\Delta\tilde{\nu}) AC_{Exp}(\Delta\tilde{\nu}) d(\Delta\tilde{\nu}) \int_{-65}^{65} w(\Delta\tilde{\nu}) AC_{DFT}(\Delta\tilde{\nu}) d(\Delta\tilde{\nu})}} \quad (2.11)$$

Puisque cette méthode est basée sur l'intégrale de l'intensité des pics, il est apparu lors de tests préliminaires qu'il est primordial de transformer les spectres pour leur permettre de se ressembler davantage, et mettre de façon effective une plus grande importance sur la position des pics dans les spectres. iROMA permettait de contourner ce problème par assignation manuelle des paires de pics, mais toute méthode basée sur la densité spectrale sera plus affectée par l'inexactitude de modélisation des spectres Raman par DFT. L'idée apportée par l'auteur de ce document pour confronter ce problème et rendre possible l'adaptation de la WCCA au problème qui nous intéresse est d'ajouter une étape de pré-traitement consistant à faire passer tous les spectres dans un filtre de compression qui change l'intensité de chaque point du spectre par *mapping* à une fonction ressemblant légèrement à une logarithmique

optimisée de façon empirique pour la comparaison d'un spectre Raman théorique à un qui est expérimental. Sa forme finale est donnée par l'équation 2.12. Pour plus de détails sur cet aspect, voir l'article du chapitre suivant et l'information en support de l'article à l'Annexe D. Le résultat de cette transformation est un genre de « compression verticale » des spectres, qui diminue « l'information » sur l'intensité sans aller jusqu'à l'approche plus drastique des spectres « code barre ». Ainsi, les pics de hauteur intermédiaire vont tendre vers des valeurs similaires, les pics de très faible intensité seront diminués davantage, et les pics très intenses vont tendre vers l'intensité maximale. Cette astuce force les intensités des spectres à plus se ressembler, sans induire des faux positifs supplémentaires en comparaison à iROMA.

$$y = 1 - 1.16 (5x + 1.1)^{-0.7(x+1)} + 0.11(46^{-350x^2 - 9.8x - 0.0686}) \quad (2.11)$$

Figure 2-11 Comparaison d'un spectre (a) avant et (b) après la compression des intensités. On remarque l'intensification des pics semblables à celui à 1200 cm⁻¹ et la disparition de certains pics près de la ligne de base.



Les points de très faible intensité sont réduits, permettant l'élimination du bruit et des très petits pics moins caractéristiques, tout en mettant les petits pics à des intensités plus similaires les uns aux autres. Un effet secondaire désirable de ce traitement est la réduction des bordures en bas des pics, produisant globalement un spectre épuré. En bref, les pics plus intenses sont légèrement augmentés, tandis que les pics moyens sont plus fortement augmentés.

Après la compression, le calcul de la note par WCCA est amorcé. Ce calcul dépend d'un signal appelé corrélation croisée, qui est donné en fonction d'un décalage horizontal $\Delta\tilde{\nu}$. La corrélation croisée est le produit de l'intensité au point $\tilde{\nu}$ sur un des deux spectres et de l'intensité au point $\tilde{\nu} + \Delta\tilde{\nu}$ sur l'autre spectre. Chaque point du signal de corrélation croisée en fonction de $\Delta\tilde{\nu}$ est la somme des produits pour ce décalage. Ainsi, on obtient une mesure du déplacement moyen de densité spectrale entre les deux

spectres. Cette méthode est attirante puisqu'elle « regarde » intrinsèquement à côté de chaque point du spectre. C'est aussi pour cela que la transformation des intensités des spectres est importante avant le calcul de la note de corrélation.

En somme, les étapes qui résument brièvement le « pipeline » SARA pour une paire de spectres donnée sont les suivantes :

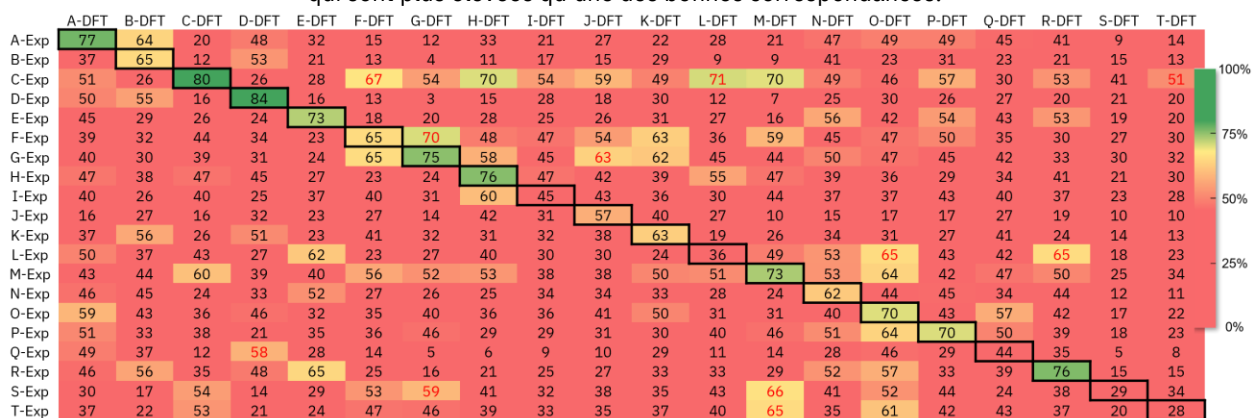
1. Ingestion des deux fichiers et de leurs données. Si nécessaire, les pics théoriques sont élargis par profil de Voigt.
2. Les données spectrales sont triées par nombre d'ondes croissants.
3. Le facteur de correction harmonique défini par l'utilisateur est appliqué au spectre théorique.
4. Les spectres sont rééchantillonnés à une résolution de 1 cm^{-1} par interpolation linéaire.
5. Les intensités spectrales des deux spectres sont normalisées individuellement par méthode min-max afin de tout mettre sur un pied d'égalité (spectres allant de 0 à 1 d'intensité).
6. La fonction de compression est appliquée séparément sur les intensités des deux spectres.
7. Les fonctions d'autocorrélation des deux spectres sont générées (pour servir de normalisation de la densité spectrale par rapport à la corrélation croisée, et minimisant l'impact de la fonction de compression sur la note), et le signal de corrélation croisée entre les deux spectres est généré.
8. Une fenêtre de pondération identique est appliquée sur les trois signaux pour discriminer les pics homologues lointains et favoriser les pics homologues rapprochés.
9. Les trois signaux de corrélation pondérés sont intégrés, et la note est calculée pour la paire de spectres.

Le Chapitre 3 (l'article) et son document d'informations complémentaires contiennent plus en détail les équations exactes du calcul de WCCA, de la fenêtre de pondération, et de la fonction de compression.

2.4.2 Résultats et discussion

En passant les mêmes spectres utilisés pour démontrer iROMA par SARA, on peut avoir une matrice 20x20 avec des scores calculés automatiquement, en une trentaine de secondes. Cela permet de comparer les résultats directement entre les deux logiciels. Une démonstration plus ciblée (c'est-à-dire pour la discrimination de produits de synthèse) est faite dans l'article présenté au prochain chapitre, avec des résultats plus intéressants qu'ici, puisque l'application réelle du logiciel est démontrée, et donne des « *matches* » plus satisfaisants.

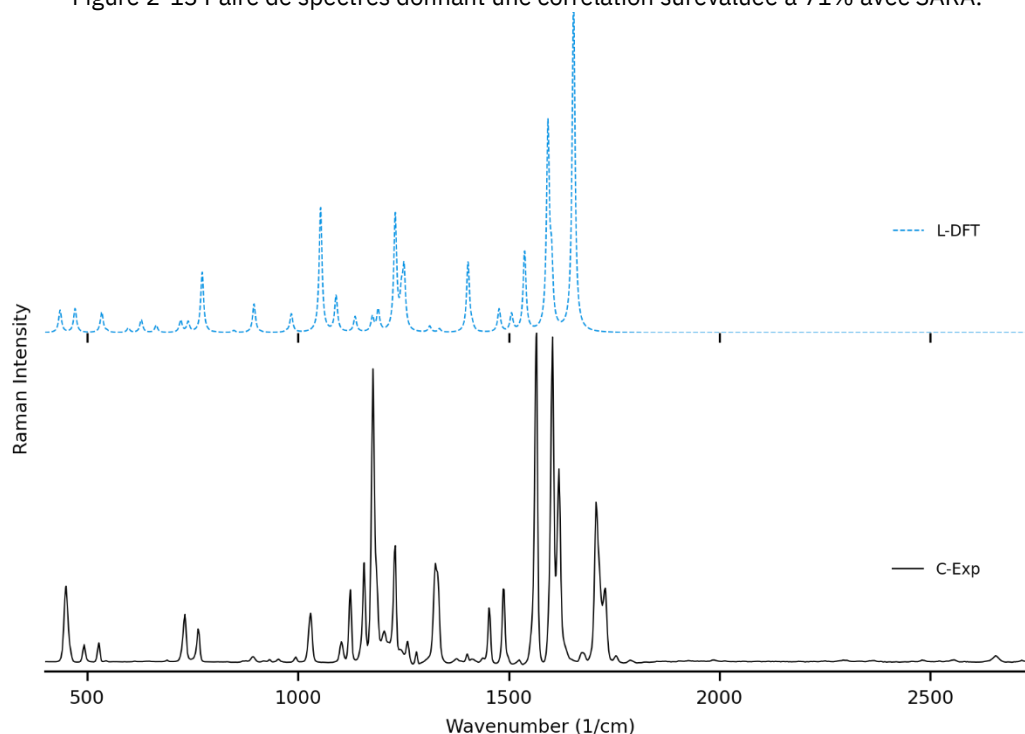
Figure 2-12 Matrice de corrélation SARA des vingt spectres-tests de la Fig. 2-5 (modélisation en phase gazeuse, échantillons solides). Les valeurs en rouge représentent les notes dans la région des mauvaises correspondances qui sont plus élevées qu'une des bonnes correspondances.



Les spectres étant les mêmes que pour la première matrice de corrélation présentée, le lectorat est référé à la Figure 2-6 et à l'Annexe B pour observer les paires de spectres de la diagonale.

Les cas difficiles ne sont pas les mêmes que pour iROMA. La correspondance exagérée évaluée dans la section précédente entre la molécule O et la I (Figure 2-7) donne maintenant un vrai négatif. Cependant, le biais de cette méthode est axé sur les pics à forte intensité. En effet, la principale difficulté avec SARA est la surévaluation des pics très intenses. L'effet est mitigé par l'application de l'étape de compression décrite précédemment, mais le problème peut demeurer, comme c'est le cas pour la valeur en rouge représentant la note accordée à la paire L-DFT (spectre théorique de la 2-hydroxyquinoline) et C-Exp (spectre expérimental de la coumarine) où, entre autres, les pics à 1600 cm⁻¹ se chevauchent pour contribuer à une corrélation croisée artificiellement haute (Figure 2-13).

Figure 2-13 Paire de spectres donnant une corrélation surévaluée à 71% avec SARA.



On relève dans la matrice pour SARA 11 valeurs en rouge, contre 19 pour iROMA. Cela donne une indication que moins de faux positifs sont attendus avec cette méthode pour une application réelle. Comme la méthode est automatique et dépend de la densité spectrale, il n'est pas possible de forcer l'appariement des pics légèrement distants comme on peut le faire avec iROMA. D'une certaine façon cela en fait une méthode plus sévère, et les notes accordées à plusieurs bonnes correspondances reflètent les problèmes de modélisation de la position des pics déjà connus avec les paramètres DFT employés jusqu'à présent. Pour cette raison, et pour sa rapidité par rapport à iROMA, SARA est une méthode plus adéquate pour l'analyse de routine en chimie de synthèse organique. Globalement, le taux de succès, qui sera pour l'instant défini par le nombre de molécules dont la meilleure note sur l'axe horizontal se trouve dans la diagonale, sans égard à la valeur absolue de la note, augmente légèrement avec SARA, pour les mêmes paires de spectres.

SARA est dans l'incapacité d'évaluer des spectres acquis en solution, parce que les pics du solvant ne peuvent être ignorés, et les notes s'en retrouveraient gravement dégradées. Une soustraction spectrale serait certes possible, mais des tests effectués en amont de l'écriture de ces mots révèlent que cette méthode est fortement incommode à automatiser de façon répétable.

2.5 Conclusion du chapitre

Deux méthodes de comparaison quantitative de la similarité de spectres Raman ont été développées spécialement pour le cas où un spectre Raman expérimental d'un composé pur est comparé à un spectre prédit par chimie numérique. La première est basée sur des distances simples des pics pairés manuellement et de pics seuls sélectionnés à la main, et la seconde, sur la corrélation croisée calculée de façon automatique. Un ensemble de molécules diverses a été analysé par ces méthodes pour tester leur comportement lorsque des données différentes les unes des autres sont traitées. Le logiciel SARA donne une détection plus fiable des vrais négatifs, mais légèrement moins bonne pour les vrais positifs. Il nécessite substantiellement moins de temps d'analyse que par iROMA. SARA est donc choisi pour les mesures dans des conditions réelles de synthèse, inclus dans le prochain chapitre.

Pour compléter SARA, il ne manque qu'une méthode de chimie computationnelle permettant de prédire plus justement le spectre Raman de solides organiques. À cet effet, un niveau de théorie récemment publié a permis de changer la façon dont les spectres théoriques sont générés depuis la phase test de SARA. Le prochain chapitre est consacré à la publication contenant les meilleurs résultats de ce projet, qui tirent intégralement parti de ce niveau de théorie. Le protocole DFT contenu dans le document « *Supporting Information* » de l'article rapporté à l'Annexe D a le potentiel de réaliser l'objectif fixé par mes travaux.

CHAPITRE 3

ARTICLE: TOWARDS ROUTINE ORGANIC STRUCTURE DETERMINATION USING RAMAN MICROSCOPY AND SPECTRAL MATCHING FROM r²SCAN-3c DFT CALCULATIONS

Dans ce chapitre se trouve l'article qui regroupe la continuation et la culmination des travaux présentés jusqu'à présent. Le niveau de théorie mis en vedette est le r²SCAN-3c,(Grimme *et al.*, 2021) qui combine la fonctionnelle de la densité mGGA contemporaine r²SCAN,(Furness *et al.*, 2020) avec une base triple zêta modifiée pour optimiser la balance rapidité-précision spécifiquement avec la fonctionnelle r²SCAN, ainsi que les corrections D4 pour la dispersion et gCP pour l'erreur de superposition de la base.

La rapidité d'exécution des calculs r²SCAN-3c n'est pas un obstacle à la qualité de leurs résultats,(Grimme *et al.*, 2021) ce qui permet une approche nouvelle à la prédiction de spectres Raman de solides organiques : un modèle bimoléculaire (ou modélisation par dimères) permet de calculer intrinsèquement assez d'interactions intermoléculaires pour améliorer le spectre Raman théorique qui est comparé à l'expérimental de façon notable. Il est néanmoins toujours difficile de modéliser un triacide avec cette technique, donc les molécules-test choisies pour l'article ont été rajustées pour contenir quelques composés qui présentent une amélioration entre une modélisation uni- et bimoléculaire.

Le niveau de théorie utilisé est implémenté dans le robuste logiciel ORCA, qui a l'avantage d'être gratuit pour les groupes de recherche académiques, ce qui le rend plus accessible pour les chimistes de synthèse.

Le logiciel SARA est employé pour discriminer des produits de synthèse d'intérêt pour les organiciens issus de différentes étapes d'une suite de réactions.

Towards Routine Organic Structure Determination using Raman Microscopy

Jason Malenfant, Lucille Kuster,[‡] Yohann Gagné,[‡] Kouassi Signo, Maxime Denis, Sylvain Canesi*, Mathieu Frenette*

Department of Chemistry, NanoQAM and Centre Québécois de Matériaux Fonctionnels (CQMF), Université du Québec à Montréal, Montreal, Quebec H3C 3P8, Canada

Supporting Information Placeholder

ABSTRACT: Raman microscopy can reveal a compound-specific vibrational “fingerprint” from micrograms of material with no sample preparation. We expect this increasingly available instrumentation to routinely assist synthetic chemists in structure determination; however, interpreting the information-dense spectra can be challenging for unreported compounds. Appropriate theoretical calculations using the highly efficient r²SCAN-3c method can accurately predict peak positions but are less precise in matching peak heights. To limit incorrect biases while comparing experimental and theoretical spectra, we introduce a user-friendly software that gives a match score to assist with structure determination. The capabilities and limitations of this approach are demonstrated for several proof-of-concept examples including the characterization of intermediates in the total synthesis of deoxyaspidospermine.

Introduction

Following synthesis, chemists rely on a range of analytical techniques to verify molecular structures (Table 1). As a characterization method, visual inspection is clearly insufficient—most synthetic organic molecules are white to off-white powders. Still, a sample’s interaction with visible light can reveal surprisingly precise information about its composition. Visible light and molecules can undergo inelastic collisions, first described by Sir Chandrasekhara Venkata Raman,¹ establishing the basis for a spectroscopic technique that now bears his name. Raman spectroscopy is a relatively underused structure elucidation method by synthetic organic chemists, but this characterization method is prevalent in material science.² As with infrared spectroscopy, Raman spectroscopy reports a vibrational spectrum that is specific to a molecular structure and geometry. The peaks in Raman spectra are often narrower than those in infrared spectra, which better isolates and distinguishes the numerous peaks found in the fingerprint region (~200-1800 cm⁻¹).

Raman microscopy is an increasingly available analytical tool with several advantages that foreshadow its use in routine organic structure determination. Samples are non-destructively measured directly in their powder form, without sample preparation or the need for expensive deuterated solvents. Air-, moisture- or temperature-sensitive samples can be measured in an inert atmosphere (through a quartz vial) or directly under liquid nitrogen.^{2d} Confocal Raman microscopy can measure vibrational spectra with a spatial resolution of less than ~5 μm³ or about 10 μg of solid sample. With such a low sample size needed to record a vibrational spectrum and the often-remarked fingerprint-quality of these spectra, Raman microscopy is primed to assist chemists in routine structure determination.

Table 1. Comparison between common structure elucidation techniques and Raman microscopy.

Technique	Information	Sub-mg?	Non-destructive?	No sample preparation?
Infrared	Fingerprint	No ✗	Yes ✓	Yes ✓
Mass Spec.	Mass	Yes ✓	No ✗	Yes ✓
¹ H NMR	Structural	Yes ✓	Yes ✓	No ✗
¹³ C NMR	Structural	No ✗	Yes ✓	No ✗
Conventional Raman	Fingerprint	No ✗	Yes ✓	Yes ✓
Raman Microscopy	Fingerprint	Yes ✓	Yes ✓	Yes ✓

While Raman spectra databases can confirm a match for reported compounds,³ the same exercise is more difficult for unreported spectra. Fortunately, vibrational modes are amongst the most precisely predicted properties by Density Functional Theory (DFT) calculations. Peak positions in predicted spectra rely mostly on molecular bonding and geometry; both these properties are accurately modeled by relatively inexpensive DFT calculations. While it is challenging to interpret a medium-sized molecule’s Raman spectrum without a valid comparison, matching an experimental spectrum to a DFT-predicted spectrum is more straightforward.

DFT-correlated Raman microscopy can be integrated in an organic synthetic process as illustrated in Figure 1. Following purification of synthetic products, via preparatory thin layer chromatography (TLC) for example, as little as 10 μg of material are needed to be analyzed with an appropriate confocal Raman microscope setup. The correlation between experimental and calculated peak positions is often quite good, as indicated qualitatively with dashed lines in Figure 1.

While DFT calculations can offer satisfactory matches in the number and position of Raman peaks, predicted Raman intensities require the third derivative of electronic densities—a calculation that is far from precise.⁴ Despite poorly predicted peak heights, matching experimental Raman spectra with DFT calculations has confirmed many molecular structures,⁵ including elusive reactive intermediates.⁶ Because it favorably predicts molecular geometries, the ubiquitous B3LYP has been the hegemonic functional in DFT-correlated Raman spectroscopy. Recently, Grimme and co-workers introduced the r²SCAN-3c method which is implemented in the free-for-academics ORCA software.⁷ The r²SCAN-3c method will reproduce, in a fraction of the time, structural accuracy of triple- ξ B3LYP calculations (e.g., B3LYP/6-311+g(2d,2p)).⁸ Less

expensive calculations allow researchers to quickly predict Raman spectra, which opens the door for routine structure elucidation by DFT-correlated Raman spectroscopy. Herein, we describe the application of r^2 SCAN-3C calculations in routine organic structure prediction from molecular solids (i.e., powder samples).

As with most spectra comparisons, deciding if a predicted spectrum corresponds to the correct structure is a subjective and a potentially biased exercise. To help with this decision, we implement a pipeline for spectral processing and quantitative match assessment. This approach yields a user-friendly percent match between predicted and experimental spectra for user-generated structures. We demonstrate DFT-correlated spectral matching in several proof-of-concept examples, including the match of synthetic intermediates from the total synthesis of deoxyaspidodispermine.⁹

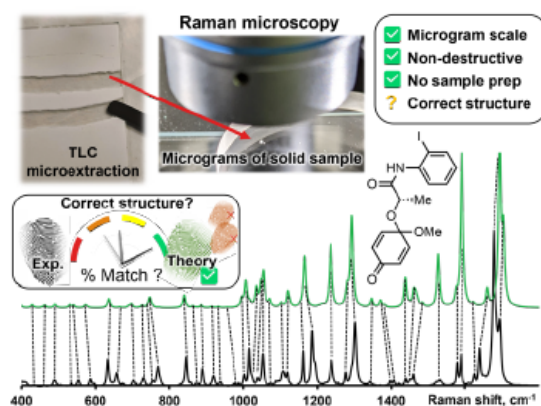


Figure 1. Proposed workflow to incorporate Raman microscopy as a tool for routine structure elucidation following synthesis. Theoretical calculations can help match an experimental vibrational fingerprint with proposed structures, however, assessing a match's quality without bias is challenging.

Method

SARA, or "Similarity Assessment of Raman Arrays", is the name given to the software written by author J.M. to process Raman spectra. It was programmed in numerical Python with a certain level of automation in mind. The script takes as input a directory path for a folder containing Raman data files to be compared. Theoretical and baseline-corrected experimental spectrum can be differentiated to SARA by either file extension, or an identifier string ("exp" or "theo") in the filename. Compatible file formats include comma-separated values (CSV), Renishaw WiRE files, as well as Gaussian- and ORCA-calculated spectra. Once spectra-containing files are submitted, the software starts pre-processing and calculating the match score for every pair of submitted experimental and theoretical spectra according to the algorithm described in Figure 2.

The core correlation metric employs the weighted cross-correlation average (WCCA) as first described by De Gelder *et al.*¹⁰ and originally used for powder X-ray diffraction. This method was shown to outperform the root mean square error metric with Raman spectra for small molecules.¹¹ In order to adapt this algorithm in a way that limits the errors in modeled peak intensities, spectra are passed through a compression filter which maps the intensities according to a pre-defined logarithm-like function (Figure S4). This compression stretches the intensity of peaks towards the spectrum's maximum intensity, rendering peak intensities closer to one another without taking a drastic "barcode spectrum" approach.

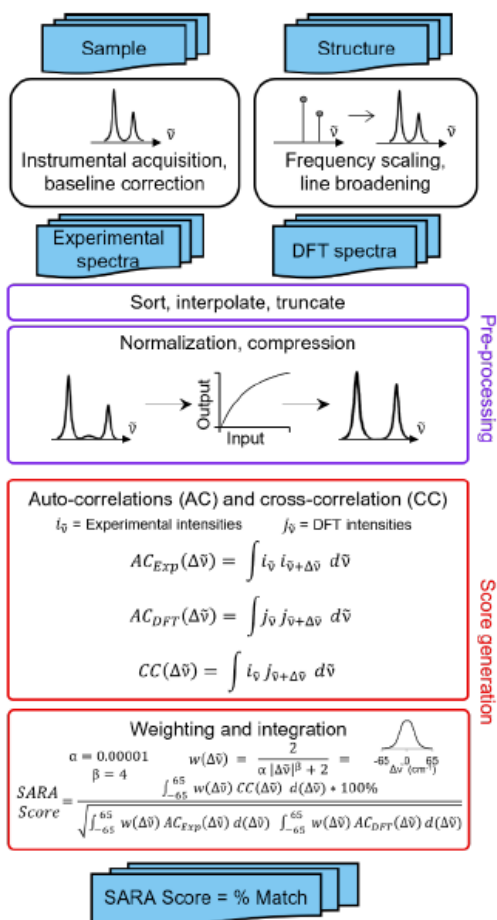


Figure 2. Proposed algorithm to generate match scores between experimental spectra and theoretical calculations.

For each pair of spectra, the score calculation starts with pre-processing, defined in six steps: (1) The files are parsed. If necessary, peak broadening using a specifically parametrized Voigt profile¹² is applied to files consisting of only theoretical wavenumbers to obtain a realistic spectrum. (2) Both spectra are sorted by ascending wavenumber. (3) A user-defined theoretical frequency correction factor is applied to the theoretical spectrum—a 0.98 correction factor was used in this study. (4) Both spectra are resampled to obtain a directly comparable resolution of 1 cm^{-1} . (5) The spectral intensities are individually normalized using the min-max method. (6) The compression function (equation S1) is applied to the intensities. With the pre-processing complete, the match score can be calculated, as described by De Gelder and colleagues, with a Karfunkel window using parameters shown in Figure 2. One autocorrelation signal is integrated per spectrum, and their cross-correlation is computed, allowing for the SARA score to be assigned to the pair of spectra. When all the scores are calculated, a text-based matrix of scores is generated in CSV format for exporting to a file or spreadsheet software. The match score closest to 100 for a specific DFT-calculated spectrum will indicate the most likely structure for each experimental spectrum according to our algorithm. Since Raman spectra are more accurately predicted in position than in intensity, SARA was designed to penalize shifts in peak position more severely than mismatches in peak height (Figure 3).

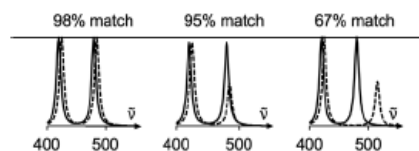


Figure 3. Spectra comparisons and their match scores as calculated by the SARA algorithm. A significant change in peak height will not be significantly penalized (middle), whereas a shift in peak position will result in a significantly lowered score (right).

Computational details

All DFT calculations were carried out using parallel¹³ ORCA^{14,15} version 5¹⁶ at the r²SCAN-3c level of theory⁵ with the resolution of identity approximation.¹⁷ The recently introduced r²SCAN-3c level of theory uses the re-regularized meta-GGA functional r²SCAN¹⁸ with the basis set Def2-mTZVPP. It automatically applies the geometrical counterpoise¹⁹, and D4 corrections,^{20,21} which compensate for the basis set superposition error and dispersion interactions,

respectively. Geometrical conformer search, dimer preparation protocols, example inputs and data processing are included in the Supporting Information. Figure S5 shows the ~15x reduction in calculation time for r²SCAN-3c versus B3LYP/Def2TZVPD.

Results and discussion

The advantages of Raman microscopy in routine organic structure predictions are clear: the method is non-destructive and requires sub-milligram quantities of material. Additionally, the information-rich experimental data are molecule-specific and, in several cases, geometry-specific. Raman spectra have been used to distinguish between crystalline polymorphs^{22,23} and polymer orientation,²⁴ however, the differences between these spectra are relatively minor. Raman microscopy also contributed prominently to the identification of pigments used in historical artifacts due to the non-destructive and spatially resolved nature of Raman microscopy.²⁵ The SARA matching algorithm was applied to identify the solid-state structure of two pigments, Pigment Red 1 (PR1) and Pigment Red 3 (PR3) (see Figure 4).

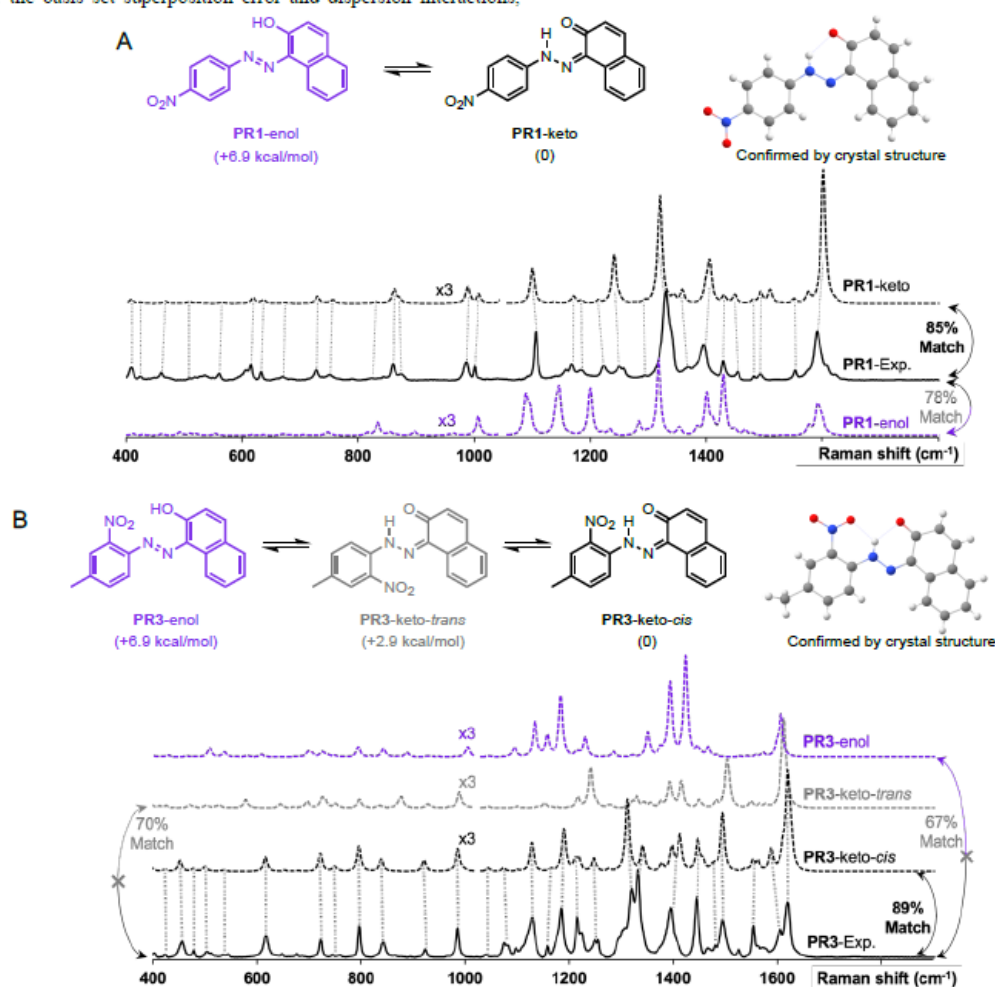


Figure 4. Both pigments PR1 (A) and PR3 (B) are more stable in their keto vs enol form according to r²SCAN-3C calculated energies. The more stable keto forms are confirmed by SARA match scores against other plausible structures, and by the reported crystal structures for PR1²⁷ and PR3²⁸.

These molecules can exist as several conformers and crystal structures are reported for both molecules.²⁶⁻²⁸ Both compounds are clearly identified as the keto forms according to bond lengths, i.e., **PR1**-keto and **PR3**-keto-*cis* versus other possible enantiomers/geometries **PR1**-enol and **PR3**-keto-*trans* or **PR3**-enol. The structures were calculated with the level of theory r²SCAN-3c and free energy calculations confirm the most stable form. Visual inspection of DFT-predicted spectra with the reported experimental spectra does offer a qualitatively similar conclusion. Satisfyingly, treatment of Raman data with SARA gives the highest match score for the correct structure. We observe for the pigment **PR1** that the **PR1**-keto isomer gives the best correlation with a match score of 85%. For molecule **PR3**, the **PR3**-keto-*cis* isomer gives the best correlation at 89%.

Inspired by the promise of Raman microscopy in the routine structure elucidation of synthetic products, we applied the SARA algorithm to verify its performance against a series of 12 small molecules (Figure 5A). An ideal outcome would be perfect match scores when comparing experimental and theoretical spectra for the same molecule. Additionally, no higher scores should be found in the comparison of dissimilar structures (i.e., false positives). Each of the 12 powder samples were measured using Raman microscopy (see SI for experimental details). Predicting their Raman spectra from monomers in the gas phase gave decent results but presented several problems. Several molecules gave good match scores when comparing experimental and theoretical spectra (ranging from 54-94%), as seen in the diagonal of Figure 4B, however, there were several notable false positives for dissimilar structure.

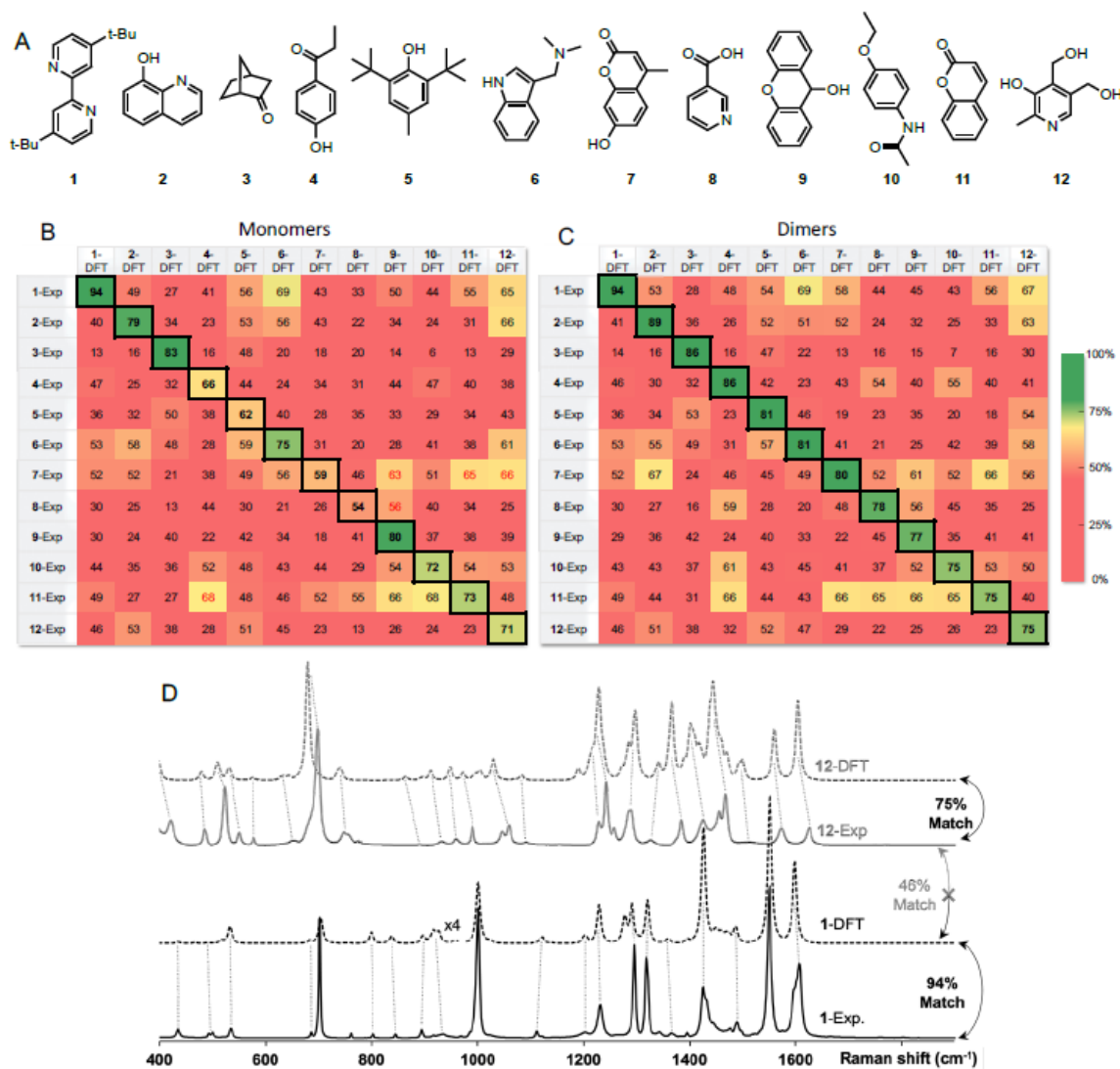
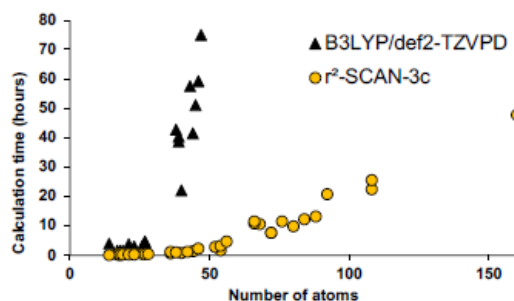


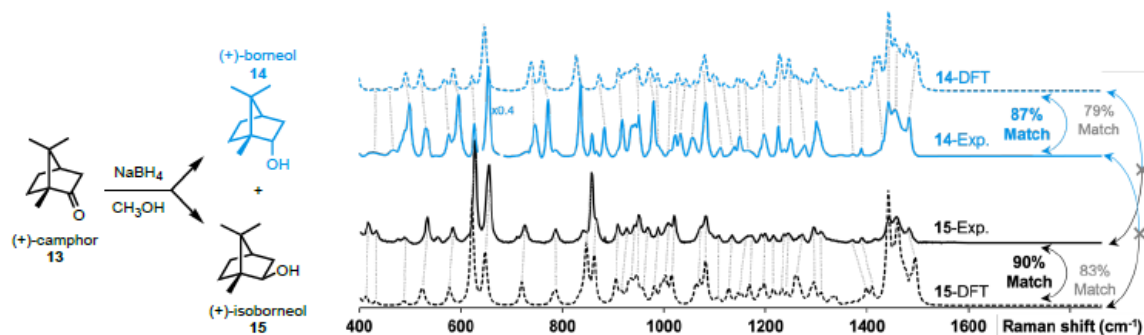
Figure 5. (A) Structure of molecules **1** to **12**. (B) Match scores for experimental and DFT-predicted Raman spectra using solid samples of molecules **1** to **12**. The ideal result would be a perfect diagonal with no false positives. Modeling experimental spectra of solid powders as single molecules (monomers) gives several false positives and some unsatisfactory scores in the diagonal. (C) Optimized dimers gave satisfying match scores in the diagonal and no false positives for these 12 small molecules. (D) Selected examples of match scores for pairs of Experimental and DFT-predicted Raman spectra using dimers.

1 Instead of predicting Raman spectra from single molecules in the
 2 gas phase, an ideal approach would be to use precise crystalline
 3 structures as input. Most molecules do not have reported crystal
 4 structures, of course, and crystal structure prediction is at the fore-
 5 front of computational chemistry research.²⁰ In addition, precise
 6 calculations for repeating unit cells of medium-sized molecules are
 7 currently prohibitively time-consuming. As a compromise between
 8 gas-phase monomers and full crystalline structures, we explored di-
 9 mers of molecules. Both B3LYP/def2-TZVP and the recent com-
 10 posite method r^2 SCAN-3c offered accurate spectral matches, how-
 11 ever, r^2 SCAN-3c calculations showed dramatically lower compu-
 12 tational cost as the system size increases (Figure 6).



23 **Figure 6.** Computational time required for a geometry optimization
 24 and Raman spectrum calculation (using 16-cores) as a function of the
 25 total number of atoms using B3LYP/def2-TZVP (black triangles)
 26 and r^2 SCAN-3c (yellow circles).

27 After standard geometry optimization, several dimer geometries
 28 were generated from chemical intuition, with pairs selected to max-
 29 imize H-bonding or pi-stacking. The lowest electronic energy geo-
 30 metry was then chosen to perform a frequency calculation and to
 31 predict a Raman spectrum.



32 **Figure 7.** Application of the proposed workflow introduced in Figure 1. Milligram-scale reduction of **13** yields a mixture of **14** and **15** which
 33 were separated via preparative thin layer chromatography and solvent extraction. Analysis of ~15 μ g of each sample by Raman microscopy
 34 gave good matches with DFT-predicted spectra using dimers. Calculated match scores using the SARA algorithm confirmed the assignments.

35 The proposed methodology is expected to be particularly beneficial
 36 in multi-step syntheses, where milligram-scale synthetic attempts
 37 can yield microgram quantities of isolatable products. In such sys-
 38 tems, mass spectrometry can give a good idea of possible products,
 39 and Raman microscopy would further confirm or deny proposed
 40 structures. With this goal in mind, we followed a recently reported
 41 total synthesis of deoxyaspodispermine using Raman microscopy
 42 as a structure elucidation tool.⁹ A total of 6 structures, molecules
 43 **16** to **21** (Figure 8A), were isolated and analyzed by Raman mi-
 44 croscopy (see SI for experimental and theoretical spectra).

The match scores between experimental and predicted spectra from
 dimers of the same 12-molecule set are shown in Figure 5C. The
 diagonal correlation showed better match scores ranging from 75
 to 94%. Using dimers, importantly, no false positives were found,
 showing promise for the use of this method in routine organic struc-
 ture determination. Figure 5D shows the correlation scores between
 the theoretical spectra (modeled with dimers) and the experimental
 spectra between molecule **1** and **12** with a 94% and 75% match,
 respectively. Spectra comparison between the DFT predicted spec-
 tra of **1** and experimental spectrum of **12** show expected differences
 that explain a lower match score of 46%.

The effectiveness of our protocol was evaluated in the identifica-
 tion of two products resulting from a reaction (see Figure 7). The
 reduction of (+)-camphor (**13**) with NaBH₄ yields a mixture of two
 isomers, (+)-borneol (**14**) and (+)-isoborneol (**15**) (see SI for a syn-
 thetic protocol). Following their separation via preparative thin
 layer chromatography and microextraction, the Raman spectra of
 the extracted solids were acquired. The experimental data presented
 in Figure 6 were obtained with ~15 μ g of each sample, i.e., the
 smallest spec of sample visible by the naked eye.

While they are quite similar, some differences are visually notice-
 able between the experimental spectra of the two isomers. Inter-
 preting these spectra, unaided by DFT-calculations, would be chal-
 lenging. The DFT-calculated spectra for dimers of **13** and **14** are,
 unsurprisingly, similar as well. Subjectively, the position of peak
 predictions satisfyingly overlay with experimental peak positions
 as indicated by the dotted lines that link pairs of spectra. Quantita-
 tively, the SARA algorithm can also help pair experimental spectra
 to isomers **13** and **14**. The correlation scores obtained with the
 SARA algorithm are higher for the experimental and theoretical
 spectra of the same molecule (87% for **14** and 90% for **15**) than for
 different molecules (83% and 79%). This proof-of-concept work-
 flow shows how Raman microscopy, highly efficient r^2 SCAN-3c
 DFT calculations and the SARA software can differentiate and
 identify two isomeric products using very little isolated product.

Comparing experimental and DFT-predicted spectra for similar
 molecules gave satisfyingly high match scores, varying between
 78% and 95% using monomers or dimers (Figure 8B and 8C, re-
 spectively). The match score matrix was able to reliably identify
 the correct molecule, except for the final transformation from **20** to
21 where relatively small structural changes are noted. A near false-
 positive is observed in the final result (Figure 8C), however, the
 correct assignment becomes obvious with the 95% match score for
 molecule **20**.

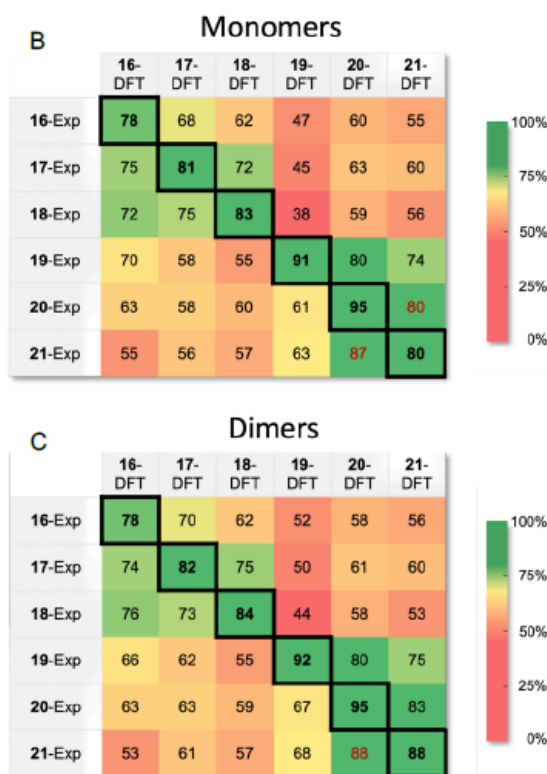
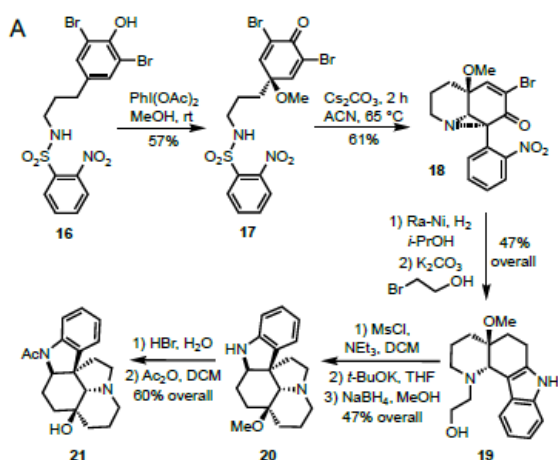


Figure 7. (A) Structure of intermediates analyzed by Raman microscopy during the racemic synthesis of deoxyaspidospermine.⁹ Match scores between experimental Raman spectra for compounds **16** to **21** and calculated Raman spectra using the r^2 SCAN-3c level of theory for monomers (B) and dimers (C).

Special Considerations

As with any analytical technique, special considerations must be kept in mind to acquire adequate Raman data.³⁰ The following pitfalls should be avoided. As mentioned throughout the text, samples were measured on quartz since laboratory glass has a significant Raman signal. Another common enemy of Raman spectroscopy is

fluorescence—a very weak fluorescent signal can easily dominate Raman signals. If the fluorescence signal is not completely overpowering, proper baseline correction can extract Raman spectra. While Raman signals are more intense with short wavelength excitation, longer wavelength excitations are often used to avoid 2-photon excitation and subsequent fluorescence. Another common pitfall arises from coloured impurities in samples. The Raman signal of impurities that absorb the laser excitation can be greatly enhanced due to the resonance Raman effect. In these cases, the Raman spectrum could be disproportionately contaminated by that of the coloured impurity.

Conclusion

As it becomes more widely available, Raman microscopy will play a growing role in routine structure determination. From sub-milligram quantities of material, without sample preparation, information-rich Raman spectra can be generated. While a well-resolved spectrum is specific to a molecule's bonding and geometry, unambiguous interpretation of a molecule's Raman spectrum can be challenging. Fortunately, DFT calculations are well-suited to predict the number and position of Raman peaks, particularly in the fingerprint region (~ 200 to 1800 cm^{-1}). We find the r^2 SCAN-3c method, as implemented in the ORCA software, to give adequately predicted Raman spectra with remarkable speed compared to similarly accurate methods.

While peak positions are well-predicted by DFT calculations, peak heights remain poorly described. To limit bias when comparing experimental and DFT-calculated Raman spectra, an algorithm that limits the impact of peak height was developed to give a score, out of 100, for a proposed structure. This simple output confirmed the correct structure between several proposed structures, including isomers and conformers.

Some molecules required a better description of their intermolecular forces to properly predict the Raman spectra of their powder forms. In these cases, an appropriate dimer afforded a significant improvement over single molecules in predictive power. This approach was validated with several proof-of-concept examples, including the characterization of synthetic intermediates in the total synthesis of deoxyaspidospermine. Computationally-aided spectral matching will continue to improve, and, as it does, we expect Raman microscopy to find growing applicability in routine organic structure determination.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at <https://pubs.acs.org/doi/>. ORCA input file examples and computational protocol; match scores for monomer, dimer and tetramer modeling for molecules **1** through **12**; information on optimized structures from which Raman spectra were calculated, including optimized atomic coordinates and energies; experimental and theoretical Raman spectra of all samples; commercial sources for all chemicals; synthesis results for the reduction of molecule **13**; equation for SARA's pre-processing compression function.

Python Code

The open-source code for the SARA program can be found at author J.M.'s GitHub repository (<https://github.com/jasonMa142/SARA>).

AUTHOR INFORMATION

Corresponding Author

Mathieu Frenette (frenette.mathieu@uqam.ca)
Sylvain Canesi (canesi.sylvain@uqam.ca)

Author Contributions

The initial idea was conceived by M.F., L.K., and J.M. J.M. developed the algorithm, wrote the Python code, measured the Raman spectra and generated the theoretical spectra. Y.G. extracted/calculated spectra found in Figure 4 and assisted with conformer searches. K.S., M.D., and S.C. provided synthetic samples and helped design/guide the studies. M.F., L.K., and J.M. wrote the initial draft which was edited with the other authors. †These authors contributed equally.

Notes

The authors declare no competing financial interests.

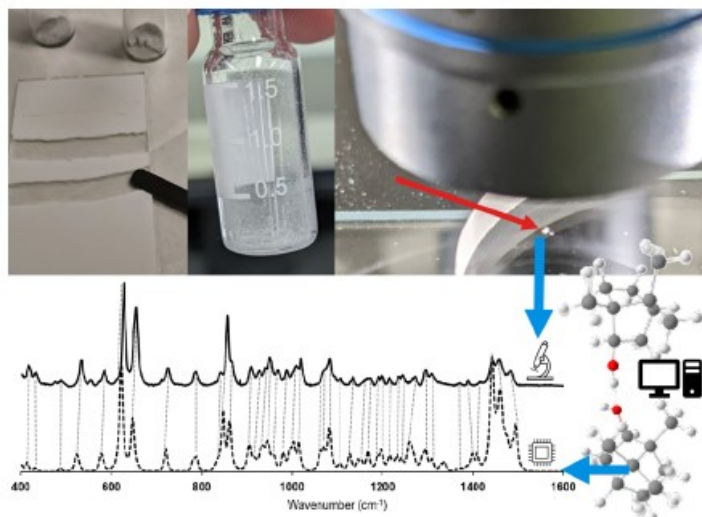
ACKNOWLEDGMENT

We gratefully acknowledge NSERC for funding in the form of a Discovery Grant to M.F. and S.C. as well as an equipment grant to M.F. that allowed the purchase of a Raman Microscope, which is maintained by Gwénaél Chamoulaud at the NanoQAM and CQMF research platforms. Calculations were conducted using the open-sourced software ORCA using computational resources provided by Calcul Québec (calculquebec.ca) and the Digital Research Alliance of Canada (alliancecan.ca).

REFERENCES

- (1) Raman, C. V. A new radiation. *Proceedings of the Indian Academy of Sciences - Section A* **1953**, *37*, 333-341.
- (2) (a) Wu, J.B.; Lin, M.L.; Cong, X.; Liu, H.N.; Tan, P.H. Raman spectroscopy of graphene-based materials and its applications in related devices. *Chemical Society Reviews* **2018**, *47*(5), 1822-1873. (b) Kuhar, N., Sil, S., Verma, T.; Umapathy, S. Challenges in application of Raman spectroscopy to biology and materials. *RSC advances*, **2018**, *8*(46), 25888-25908; (c) Orlando, A., Franceschini, F., Muscas, C., Pidkova, S., Bartoli, M., Rovere, M. and Tagliaferro, A., 2021. A comprehensive review on Raman spectroscopy applications. *Chemosensors*, *9*(9), 262. (d) Compton, R. N.; Hammer, N. I. Raman Under Liquid Nitrogen (RUN). *Journal of Physics: Conference Series* **2014**, *548* (1), 012017. DOI: 10.1088/1742-6596/548/1/012017.
- (3) Robertson, J. L.; Senger, R. S.; Talty, J.; Du, P.; Sayed-Issa, A.; Avelar, M. L.; Ngo, L. T.; Gomez De La Espriella, M.; Fazili, T. N.; Jackson-Akers, J. Y.; et al. Alterations in the molecular composition of COVID-19 patient urine, detected using Raman spectroscopic/computational analysis. *PLoS ONE* **2022**, *17* (7), e0270914. DOI: 10.1371/journal.pone.0270914.
- (4) Zvereva, E. E.; Shagidullin, A. R.; Katsyuba, S. A. Ab initio and DFT predictions of infrared intensities and Raman activities. *Journal of Physical Chemistry A* **2011**, *115* (1), 63-69. DOI: 10.1021/jp108057p.
- (5) (a) Juneau, A.; Frenette, M. Raman Spectra of Persistent Radical Anions from Benzophenone, Fluorenone, 2, 2'-Bipyridyl, 4, 4'-Di-tert-butyl-2, 2'-dipyridyl, and Anthracene: Excellent Agreement between DFT and Experiment for Highly Delocalized Radical Systems. *The Journal of Physical Chemistry B*, **2022** *125*(6), 1595-1603. DOI: 10.1021/acs.jpcc.0c04742. (b) Kiefer, J.; Fries, J.; Leipertz, A. Experimental Vibrational Study of Imidazolium-Based Ionic Liquids: Raman and Infrared Spectra of 1-Ethyl-3-methylimidazolium Bis(Trifluoromethylsulfonate)imide and 1-Ethyl-3-methylimidazolium Ethylsulfate. *Applied Spectroscopy* **2007**, *61* (12), 1306-1311. DOI: 10.1366/000370207783292000. (c) Käppler, A.; Fischer, D.; Oberbeckmann, S.; Schemewski, G.; Labrenz, M.; Eichhorn, K.-J.; Voit, B.. Analysis of Environmental Microplastics by Vibrational Microspectroscopy: FTIR, Raman or Both? *Analytical and Bioanalytical Chemistry* **2016**, *408* (29), 8377-8391. DOI: 10.1007/s00216-016-9956-3.
- (6) (a) Juneau, A.; Lepage, I.; Sabbah, S. G.; Winter, A. H.; Frenette, M. Mechanistic Insight into Phenol Dearomatization by Hypervalent Iodine: Direct Detection of a Phenoxenium Cation. *The Journal of Organic Chemistry* **2022**, *87* (21), 14274-14283. DOI: 10.1021/acs.joc.2c01765. (b) (1) Čirić-Marjanović, G.; Trchová, M.; Stejskal, J. The chemical oxidative polymerization of aniline in water: Raman spectroscopy. *Journal of Raman Spectroscopy* **2008**, *39* (10), 1375-1387. DOI: https://doi.org/10.1002/jrs.2007.
- (7) Neese, F. The ORCA program system. *WIREs Computational Molecular Science* **2012**, *2* (1), 73-78. DOI: https://doi.org/10.1002/wcms.81.
- (8) Grimme, S.; Hansen, A.; Ehlert, S.; Mewes, J.-M. r²SCAN-3c: A "Swiss army knife" composite electronic-structure method. *The Journal of Chemical Physics* **2021**, *154* (6), 064103. DOI: 10.1063/5.0040021.
- (9) Signo, K.; Canesi, S. Synthesis of Deoxyaspidodispermine Based on a Functional Protecting Group Strategy. *Organic Letters* **2022**, *24* (27), 4939-4942. DOI: 10.1021/acs.orglett.2c01878.
- (10) de Gelder, R.; Wehrens, R.; Hageman, J. A. A generalized expression for the similarity of spectra: application to powder diffraction pattern classification. *Journal of Computational Chemistry* **2001**, *22* (3), 273-289. DOI: 10.1002/1096-987x(200102)22:3<273::Aid-jcc1001>3.0.Co;2-0.
- (11) Fairchild, S. Z.; Bradshaw, C. F.; Su, W.; Guharay, S. K. Predicting Raman spectra using density functional theory. *Applied Spectroscopy* **2009**, *63* (7), 733-741. DOI: 10.1366/000370209788700991.
- (12) Sundius, T. Computer fitting of Voigt profiles to Raman lines. *Journal of Raman Spectroscopy* **1973**, *1* (5), 471-488. DOI: https://doi.org/10.1002/jrs.1250010506.
- (13) Neese, F.; Wennmo, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *The Journal of Chemical Physics* **2020**, *152* (22), 224108. DOI: 10.1063/5.0004608.
- (14) Neese, F. The SHARK integral generation and digestion system. *Journal of Computational Chemistry* **2022** *44*, (3). DOI: 10.1002/jcc.26942.
- (15) Lehtola, S.; Steigemann, C.; Oliveira, M. J. T.; Marques, M. A. L. Recent developments in libxc — A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1-5. DOI: 10.1016/j.softx.2017.11.002.
- (16) Neese, F. Software update: The ORCA program system—Version 5.0. *WIREs Computational Molecular Science* **2022**, *12* (5), e1606. DOI: 10.1002/wcms.1606.
- (17) Neese, F. An improvement of the resolution of the identity approximation for the formation of the Coulomb matrix. *Journal of Computational Chemistry* **2003**, *24* (14), 1740-1747. DOI: 10.1002/jcc.10318.
- (18) Furness, J. W.; Kaplan, A. D.; Ning, J.; Perdew, J. P.; Sun, J. Accurate and Numerically Efficient r²SCAN Meta-Generalized Gradient Approximation. *The Journal of Physical Chemistry Letters* **2020**, *11* (19), 8208-8215. DOI: 10.1021/acs.jpclett.0c02405.
- (19) Kruse, H.; Grimme, S. A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. *The Journal of Chemical Physics* **2012**, *136* (15), 154101. DOI: 10.1063/1.3700154.
- (20) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A generally applicable atomic-charge dependent London dispersion correction. *The Journal of Chemical Physics* **2019**, *150* (15), 154122. DOI: 10.1063/1.5090222.
- (21) Caldeweyher, E.; Mewes, J.-M.; Ehlert, S.; Grimme, S. Extension and evaluation of the D4 London-dispersion model for periodic systems. *Physical Chemistry Chemical Physics* **2020**, *22* (16), 8499-8512. DOI: 10.1039/D0CP00502A. DOI: 10.1039/D0CP00502A.
- (22) Paudel, A.; Rajada, D.; Rantanen, J. Raman spectroscopy in pharmaceutical product design. *Advanced Drug Delivery Reviews* **2015**, *89*, 3-20. DOI: 10.1016/j.addr.2015.04.003.
- (23) Roy, S.; Chamberlin, B.; Matzger, A. J. Polymorph Discrimination using Low Wavenumber Raman Spectroscopy. *Organic Process Research & Development* **2013**, *17* (7), 976-980. DOI: 10.1021/op400102e.
- (24) Richard-Lacroix, M.; Pellerin, C. Orientation and Structure of Single Electrospun Nanofibers of Poly(ethylene terephthalate) by Confocal Raman Spectroscopy. *Macromolecules* **2012**, *45* (4), 1946-1953. DOI: 10.1021/ma202749d.
- (25) Edwards, H. G. M.; Brooke, C. J.; Tait, J. K. F. Fourier Transform Raman Spectroscopic Study of Pigments from English Mediaeval Wall Paintings. *Journal of Raman Spectroscopy* **1997**, *28* (2-3), 95-98. DOI: 10.1002/(sici)1097-4555(199702)28:2/3<95::Aid-jrs74>3.0.Co;2-0.
- (26) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database. *Acta Crystallographica Section B* **2016**, *72* (2), 171-179. DOI: 10.1107/S2052520616003954.
- (27) A. Whitaker, *Zeitschrift für Kristallographie, Kristallgeometrie, Kristallographie, Kristallchemie* **1980**, *152* (227).
- (28) A. Whitaker, *Zeitschrift für Kristallographie, Kristallgeometrie, Kristallographie, Kristallchemie* **1978**, *147* (99).
- (29) Weatherby, J. A.; Rumson, A. F.; Price, A. J. A.; Roza, A. O. d. I.; Johnson, E. R. A density-functional benchmark of vibrational free-energy corrections for molecular crystal polymorphism. *The Journal of Chemical Physics* **2022**, *156* (11), 114108. DOI: 10.1063/5.0083082.
- (30) (1) Vandenabeele, P. Raman Spectroscopy in Daily Lab-life. In *Practical Raman Spectroscopy – An Introduction*, John Wiley & Sons, Ltd, **2013**; 101-148.

Table of Contents artwork



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

CONCLUSION

La cible de ce projet de maîtrise était de rapprocher la spectroscopie Raman d'une méthode de caractérisation de routine pour la chimie de synthèse organique. Pour ce faire, la microscopie Raman confocale a été couplée aux calculs de théorie de la fonctionnelle de la densité afin de permettre la validation des empreintes digitales vibrationnelles de molécules jamais caractérisées par Raman.

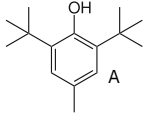
Deux logiciels de corrélation spectrale utilisant des approches différentes ont été développés pour l'obtention d'une note de corrélation quantitative entre un spectre Raman théorique et un spectre expérimental. Le logiciel iROMA permet d'évaluer la distance entre des pics sélectionnés manuellement de même que la quantité relative de pics n'ayant aucun homologue, et le logiciel SARA permet de comparer automatiquement la densité spectrale compressée entre les deux spectres par corrélation croisée.

Malgré la performance encourageante du système combiné SARA-r²SCAN-3c démontrée dans un article, il y a toujours des avenues à explorer. Dans un futur où la puissance de calcul n'est plus contraignante, il pourrait être envisageable de coupler le calcul de spectres Raman avec un système de prédiction de la structure cristalline en amont, qui déterminerait la géométrie moléculaire la plus probable sur laquelle calculer le spectre. Une pseudo-cellule cristalline pourrait ensuite être calculée par méthode multiniveaux comme la QM/MM. Une méthode de chimie quantique future pourrait également mieux prédire les intensités des pics Raman, donnant une meilleure corrélation avec les spectres expérimentaux. La performance de SARA avec r²SCAN-3c pourrait également être évaluée pour des systèmes inorganiques, ouvrant un autre champ d'applications de routine pour la spectroscopie Raman confocale. Un développement en cours dans le groupe de recherche du Pr. Frenette est d'étendre la méthode iROMA pour l'analyse de spectres déconvolués. Cette idée permettrait de sauter l'étape d'élargissement des pics théoriques, et de plutôt analyser les spectres expérimentaux comme des coordonnées uniques situées au sommet des pics. Ce processus devrait être plus facilement automatisable. Finalement, l'idée démontrée au chapitre 3 pourrait être testée avec des produits secondaires de réactions n'ayant pas un taux de conversion de 100% pour une seule espèce.

Dans l'ensemble, les résultats contenus dans l'article soumis représentent une atteinte partielle de l'objectif du projet, et l'implémentation de la méthode de corrélation spectrale en langage Python permet une discrimination généralement satisfaisante des paires de spectres Raman.

ANNEXE A
DONNÉES DFT

Modélisation en phase gazeuse, B3LYP/6-311+G(2d,2p) d'une seule molécule avec Gaussian 9

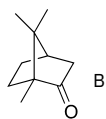


Nom : BHT (Hydroxytoluène butylé)

Énergie électronique finale: -661.492385 E_h

Géométrie optimisée :

```
C -1.24523800 0.16367700 -0.00001100
C -1.20234100 1.55543600 -0.00058900
C -0.01017000 2.27502900 -0.00066600
C 1.18052100 1.56344700 -0.00054100
C 1.22655600 0.16554400 0.00001200
C -0.00694100 -0.51492700 0.00038500
H -2.12717400 2.11146800 -0.00111500
H 2.10187000 2.12401400 -0.00101200
C -2.59402000 -0.59064200 -0.00006600
C -3.79181900 0.37849300 -0.00052600
H -3.80544600 1.01585200 -0.88524400
H -4.71542600 -0.20093600 -0.00046900
H -3.80566800 1.01640600 0.88378900
C -2.72377600 -1.46604700 1.26833200
H -3.69061900 -1.97323600 1.26970000
H -1.94436400 -2.21954200 1.32423300
H -2.66952600 -0.84847900 2.16632600
C -2.72332400 -1.46673000 -1.26803200
H -1.94390500 -2.22027400 -1.32323100
H -3.69017500 -1.97390700 -1.26949800
H -2.66872100 -0.84965400 -2.16634200
C 2.58729600 -0.57417200 -0.00007500
C 2.74960500 -1.43803900 -1.27708300
H 3.71761700 -1.94146900 -1.26649400
H 1.98762300 -2.20683200 -1.40519700
H 2.70639400 -0.80636700 -2.16433500
C 2.75014700 -1.43733300 1.27735200
H 3.71821700 -1.94064600 1.26667900
H 2.70716600 -0.80519300 2.16428100
H 1.98834100 -2.20618400 1.40617800
C 3.77593700 0.40757400 -0.00057100
H 3.77880500 1.04437200 -0.88470400
H 3.77913100 1.04487200 0.88320200
H 4.70704800 -0.15980700 -0.00059200
C -0.02215200 3.78471000 0.00080100
H -0.51820200 4.17960900 0.88987300
H 0.98970000 4.18791200 -0.01904100
H -0.55370900 4.18104700 -0.86666400
O -0.07060900 -1.89144100 0.00078000
H 0.81420800 -2.25763000 0.00109200
```

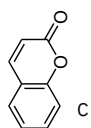


Nom : (+)-Camphre

Énergie électronique finale : -466.068451 E_h

Géométrie optimisée :

```
C 0.86708700 1.24691200 1.07585700
C -0.59644900 1.26429000 0.59751700
C 0.29404300 -0.50444000 -0.55036600
H 1.41804800 2.16064100 0.85301300
H 0.97578600 1.05880200 2.14370600
C -0.91748600 -0.25564500 0.41873100
C -0.57839900 1.78292200 -0.85705000
H -1.58151500 1.99583900 -1.22158800
H -0.00512700 2.70516200 -0.94446400
C 0.07351700 0.60217500 -1.63440800
H -0.58121600 0.21802800 -2.41558400
H 1.01090500 0.87765600 -2.11648500
H -1.27632100 1.80078400 1.25765100
C 1.44337600 0.06386000 0.28839300
O 2.58362600 -0.32854800 0.31631800
C -2.29220700 -0.52455200 -0.20688300
H -2.43677400 -1.59307900 -0.37188400
H -3.08083900 -0.19257600 0.47151500
H -2.44695500 -0.02214000 -1.15858700
C -0.83644600 -1.07203100 1.71924600
H -1.59030200 -0.72887100 2.43025800
H -1.03607900 -2.12546700 1.51931200
H 0.13236700 -1.01846500 2.21272400
C 0.54147000 -1.89517400 -1.09652400
H 1.45107800 -1.91158100 -1.69685200
H 0.67379300 -2.62214800 -0.29513800
H -0.28689800 -2.22410300 -1.72522900
```



Nom : Coumarine

Énergie électronique finale : -497.175023 E_h

Géométrie optimisée :

```
C 2.59635900 -1.00273700 -0.00001900
C 1.29531700 -1.48084000 -0.00001600
C 0.24157100 -0.57276700 0.00000000
C 0.47364900 0.81056500 0.00001400
C 1.79999600 1.26780700 0.00000000
C 2.85351300 0.37198700 -0.00001300
H 3.41912800 -1.70405600 -0.00002900
H 1.07802200 -2.53894100 -0.00001900
C -0.67545900 1.67508000 -0.00000600
H 1.98610000 2.33375400 -0.00000400
H 3.87227400 0.73205600 -0.00002200
C -1.92123700 1.16538900 -0.00000600
```

H -0.52043600 2.74684300 -0.00006000
 H -2.80531800 1.78372900 -0.00012200
 C -2.15871400 -0.27174900 0.00020300
 O -3.23310300 -0.81080100 -0.00012900
 O -1.02436500 -1.08042400 0.00004300

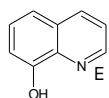


Nom : Norcamphre

Énergie électronique finale : -348.090929 E_h

Géométrie optimisée :

C 0.52175100 1.29874900 -0.39626100
 C -0.92397700 1.03435200 0.06820700
 C 0.17143300 -0.87002600 0.66480000
 H 0.63443500 1.52170300 -1.45703300
 H 0.98827300 2.11209600 0.16451900
 C -0.67891700 0.21671200 1.35730200
 H -1.59451900 -0.18030300 1.79567000
 H -0.13337600 0.77072100 2.12206100
 C -1.54905500 -0.02929300 -0.86196000
 H -2.61862800 -0.12507300 -0.67709600
 H -1.42267000 0.22346700 -1.91421500
 C -0.79180800 -1.33657000 -0.46626700
 H -1.47533800 -2.08771100 -0.07190900
 H -0.25734500 -1.78857500 -1.30007800
 H -1.52512300 1.93294500 0.18490100
 H 0.59134500 -1.66931200 1.26728600
 C 1.22497700 -0.01753700 -0.03130400
 O 2.37081400 -0.31103500 -0.26515100

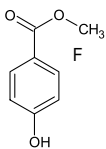


Nom : 8-Hydroxyquinoline

Énergie électronique finale : -477.304101 E_h

Géométrie optimisée :

C -1.99779200 -1.49037000 0.00000100
 C -0.69527100 -1.92662500 0.00000200
 C 0.35849800 -0.98143900 -0.00000100
 C 0.04670000 0.40582500 -0.00000200
 C -1.31672500 0.82224300 -0.00000100
 C -2.32032300 -0.11699300 -0.00000100
 H 2.02592400 -2.35825200 -0.00000200
 H -2.80394500 -2.21122900 -0.00000200
 H -0.46385100 -2.98287000 0.00000400
 C 1.73230300 -1.31644500 -0.00000100
 H -3.35050900 0.20869800 -0.00000400
 C 2.25257400 1.02702600 0.00000100
 C 2.67514500 -0.32037000 0.00000000
 H 2.98487900 1.82594900 0.00000200
 H 3.73135100 -0.54863700 0.00000000
 O -1.58947100 2.14522800 0.00000100
 H -0.72771300 2.59649400 0.00000200
 N 0.98699700 1.38441600 -0.00000100



Nom : 4-Hydroxybenzoate de méthyle

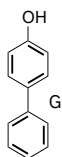
Énergie électronique finale : -535.526046 E_n

Géométrie optimisée :

```

C -1.72997600 -1.27205200 -0.00015600
C -0.35444200 -1.11655200 -0.00018000
C 0.22092100 0.15901100 -0.00007400
C -0.61391500 1.27845400 0.00020300
C -1.99113100 1.13259100 0.00013000
C -2.55203900 -0.14548500 0.00000600
H -2.18321900 -2.25270000 -0.00026900
H 0.28310800 -1.98731500 -0.00027200
H -0.16568800 2.26125800 0.00028800
H -2.63245000 2.00552200 0.00026600
O -3.89948100 -0.35332000 0.00014700
H -4.36134500 0.49063300 -0.00067700
C 1.68744200 0.37818900 -0.00006800
O 2.21972100 1.46542800 -0.00034000
O 2.38700300 -0.77944500 0.00025200
C 3.81666700 -0.63659600 0.00011700
H 4.20933800 -1.64831300 -0.00132400
H 4.14541000 -0.09662700 -0.88528600
H 4.14573600 -0.09912600 0.88694200

```



Nom : *p*-phénylphénol

Énergie électronique finale : -538.690504 E_n

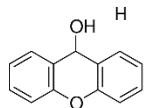
Géométrie optimisée :

```

C -0.29489000 0.00238500 0.00219000
C -1.02381100 -1.12356700 -0.39399500
C -2.41292200 -1.12419200 -0.39595300
C -3.10943500 0.01303000 0.00428900
C -2.40617300 1.14609300 0.40448300
C -1.01991900 1.13382700 0.39977700
C 1.18778500 -0.00074900 0.00058900
C 1.90838700 -1.13198500 0.40470200
C 3.29807400 -1.13584700 0.40118800
C 4.00211000 -0.00623400 -0.00380800
C 3.30115200 1.12618700 -0.40631300
C 1.91146900 1.12787200 -0.40536900
H -0.50065100 -2.00810900 -0.73018300
H -2.95345400 -2.00593300 -0.71891200
H -2.95489300 2.02061000 0.72381800
H -0.48963200 2.01458900 0.73485200
H 1.37513900 -2.00806500 0.74791800
H 3.83142700 -2.01937900 0.72531400

```

H 5.08338800 -0.00825100 -0.00535500
 H 3.83685900 2.00787600 -0.73153800
 H 1.38056700 2.00640500 -0.74581400
 O -4.47693300 0.07543800 0.02448100
 H -4.84424600 -0.76416100 -0.26663600

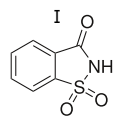


Nom : Xanthen-9-ol

Énergie électronique finale : -652.046872 E_h

Géométrie optimisée :

C -1.00485300 0.01071900 -3.57850800
 C -1.60906700 -0.12389800 -2.34097700
 C -0.83313100 -0.04300700 -1.18405700
 C 0.54316200 0.16305600 -1.24733600
 C 1.12848900 0.30092200 -2.51041100
 C 0.37176900 0.22781000 -3.66803700
 C 1.39114900 0.20759300 0.00000000
 C 0.54316200 0.16305600 1.24733600
 C -0.83313100 -0.04300700 1.18405700
 C -1.60906700 -0.12389800 2.34097700
 H -2.67275300 -0.28778800 2.24430600
 C -1.00485300 0.01071900 3.57850800
 C 0.37176900 0.22781000 3.66803700
 C 1.12848900 0.30092200 2.51041100
 H -1.60538100 -0.04986000 -4.47557400
 H -2.67275300 -0.28778800 -2.24430600
 H 2.19803400 0.45320200 -2.56795800
 H 0.84432200 0.33748200 -4.63380300
 H -1.60538100 -0.04986000 4.47557400
 H 0.84432200 0.33748200 4.63380300
 H 2.19803400 0.45320200 2.56795800
 O -1.51392500 -0.17214800 0.00000000
 O 2.39311400 -0.82981500 0.00000000
 H 1.94268300 -1.68158900 0.00000000
 H 1.99204200 1.11844500 0.00000000



Nom : Saccharine

Énergie électronique finale : -948.533587 E_h

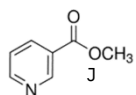
Géométrie optimisée :

C 0.81663000 0.70091800 -0.01310700
 C 0.31169700 -0.59178900 0.00279100
 C 1.12683900 -1.70796900 0.01455700
 C 2.50341600 -1.49161300 0.01608000
 C 3.02749400 -0.19782300 0.00286800
 C 2.18796300 0.91190400 -0.01261900
 H 0.71776000 -2.70740700 0.02362300
 H 3.17307500 -2.33983800 0.02785700
 H 4.09917800 -0.05833200 0.00386200

```

H 2.57656100 1.91990000 -0.02440200
C -0.23938500 1.75273200 -0.02267300
O -0.07578200 2.94863600 0.02231300
N -1.47740900 1.12803700 -0.09568500
H -2.33250700 1.64108400 0.05550100
S -1.47032700 -0.57367500 0.00142300
O -2.00263800 -1.13817500 -1.21462700
O -2.01844200 -0.98933800 1.27146500

```



Nom : Nicotinate de méthyle

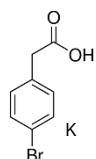
Énergie électronique finale : -476.313115 E_n

Géométrie optimisée :

```

C -2.91904100 -0.36123600 0.00000800
C -2.48730600 0.96226400 0.00003700
C -1.12498600 1.21312800 0.00003200
C -0.24494200 0.13190600 -0.00000300
C -0.78136200 -1.15852600 -0.00004300
H -3.97788800 -0.59203400 0.00000200
H -3.20514600 1.77001100 0.00006500
H -0.72865800 2.21836200 0.00005300
H -0.12548500 -2.01810100 -0.00006200
C 1.21969000 0.39955500 -0.00001400
O 1.70527800 1.50646400 -0.00007600
O 1.95110700 -0.73161600 0.00008000
C 3.37865800 -0.54833000 0.00001000
H 3.68979800 -0.00117600 0.88701300
H 3.79799600 -1.54882700 -0.00052100
H 3.68962500 -0.00027200 -0.88648900
N -2.09079800 -1.40847400 -0.00003600

```



Nom : Acide para-bromophénoxyacétique

Énergie électronique finale : -3109.051775 E_n

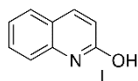
Géométrie optimisée :

```

C 1.12681300 -1.09185100 -0.00011300
C 1.92574400 0.03680300 0.00003400
C 1.35857200 1.30732600 0.00015500
C -0.01935200 1.44075300 0.00015800
C -0.83704700 0.30821200 0.00006100
C -0.26182800 -0.96089000 -0.00010700
H 1.57044300 -2.07592300 -0.00023700
H 1.98643000 2.18548900 0.00025700
H -0.48212100 2.41694400 0.00025600
H -0.86668300 -1.85419200 -0.00023900
Br 3.83460500 -0.14923600 -0.00001300
O -2.18138100 0.55542700 -0.00002800

```


C -3.05879200 -0.54138000 -0.00001700
H -2.92268700 -1.17247600 -0.88393600
H -2.92265000 -1.17250700 0.88386200
C -4.48397600 -0.02554900 -0.00005000
O -4.83454700 1.11866600 -0.00036100
O -5.33881500 -1.07759400 0.00032600
H -6.23675900 -0.71660400 0.00027300

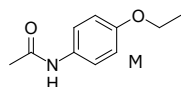


Nom : 2-Hydroxyquinoline

Énergie électronique finale : -477.312194 E_h

Géométrie optimisée :

C -2.88740100 0.33487600 0.00000300
C -1.85667500 1.24294300 -0.00002000
C -0.51351000 0.80435900 -0.00001300
C -0.23463200 -0.59293300 0.00001500
C -1.31312800 -1.50499200 0.00003600
C -2.60923100 -1.04786600 0.00003100
H 0.42333200 2.75812700 -0.00005700
H -3.91314400 0.67623500 -0.00000100
H -2.06080300 2.30606000 -0.00004200
C 0.59600400 1.68962700 -0.00003500
H -1.08693100 -2.56174400 0.00005800
H -3.42765700 -1.75496700 0.00004700
C 2.02506900 -0.21603200 0.00000300
C 1.86628400 1.19472500 -0.00003000
H 2.73938700 1.82928500 -0.00005100
N 1.04375500 -1.07346100 0.00002300
O 3.29407100 -0.68917000 -0.00000800
H 3.23029000 -1.65364800 0.00000800



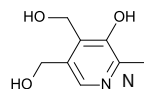
Nom : Phénacétine

Énergie électronique finale : -594.297658 E_h

Géométrie optimisée :

C 1.67204600 0.36792700 0.00038000
C 0.99695900 1.59165800 0.00026600
C -0.38393400 1.62413700 0.00017100
C -1.13510900 0.44144300 0.00009700
C -0.46274700 -0.77846600 0.00026000
C 0.93149300 -0.81152800 0.00042300
H 1.57245900 2.50617700 0.00028200
H -0.88967200 2.58230500 0.00007500
H -1.02571600 -1.69581800 0.00024400
H 1.42057300 -1.77302700 0.00064800
N -2.54383800 0.56558100 -0.00012100
H -2.88386700 1.51155300 -0.00068000
C -3.50132200 -0.41641600 -0.00011000
O -3.25269400 -1.60979700 0.00044200
C -4.93282900 0.09047900 -0.00075000
H -5.44039100 -0.30655400 0.87724800
H -5.02060300 1.17622800 -0.00084000

H -5.43970100 -0.30666300 -0.87909600
 O 3.03572800 0.43787200 0.00069100
 C 3.78243100 -0.77950700 -0.00001400
 H 3.52447200 -1.36786700 -0.88547600
 H 3.52584600 -1.36823400 0.88561500
 C 5.25370200 -0.41996000 -0.00108500
 H 5.85771800 -1.32764800 -0.00192200
 H 5.50794900 0.16397700 -0.88489700
 H 5.50939000 0.16329500 0.88275500

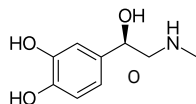


Nom : Pyridoxine

Énergie électronique finale : -592.081582 E_h

Géométrie optimisée :

C 1.83846200 -0.70421400 0.02787700
 C -0.04370700 -1.97837000 -0.27742900
 C -0.89681500 -0.88343800 -0.23310400
 C -0.33273800 0.38643700 -0.03387500
 C 1.05617900 0.46399800 0.09227900
 H -0.45096600 -2.97238800 -0.42354500
 N 1.28212400 -1.89903900 -0.15278200
 C -1.19477000 1.61378400 0.11594200
 H -2.18165500 1.44645700 -0.30941200
 H -1.32605700 1.83860800 1.17857800
 C -2.37536200 -1.09403300 -0.39489700
 H -2.75659300 -0.50613600 -1.23784000
 H -2.56131900 -2.14762700 -0.61385400
 C 3.33035400 -0.62259500 0.17458600
 H 3.76345600 0.02890200 -0.58607200
 H 3.60803400 -0.19895600 1.14136400
 H 3.75427600 -1.61874600 0.08364500
 O 1.71533300 1.63755600 0.30028600
 H 1.11580500 2.36475900 0.06197600
 O -0.56528200 2.73488800 -0.54485700
 H -1.04987100 3.53494100 -0.32427500
 O -3.04918700 -0.70599200 0.81369400
 H -3.98650200 -0.89757400 0.71765400



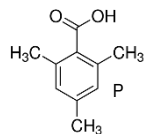
Nom : Adrénaline

Énergie électronique finale : -631.406588 E_h

Géométrie optimisée :

C -0.59337500 -1.42369400 -0.12187900
 C -0.04590700 -0.26010300 0.41101300
 C -0.85909700 0.86317500 0.57047600
 C -2.19265100 0.83550600 0.19304500
 C -2.73146400 -0.33655500 -0.34480400
 C -1.93441500 -1.45920400 -0.49677000
 H 0.02320700 -2.30332900 -0.22388200

H -0.47106700 1.77677400 1.00195900
H -2.36170200 -2.36776500 -0.90456100
O -2.96711300 1.94789600 0.36085700
H -3.86189300 1.74207100 0.06585600
O -4.06733700 -0.27667100 -0.68384800
H -4.36338500 -1.13077500 -1.00939100
C 1.41784500 -0.18010400 0.78215100
H 1.51359300 0.47742000 1.65801300
C 2.26196200 0.41381400 -0.35569400
H 1.90232400 1.42161400 -0.60191000
H 2.12798800 -0.20812600 -1.24295000
O 1.91369600 -1.47117700 1.10840100
H 2.87236800 -1.41485300 0.98457700
N 3.67688800 0.37421200 0.01453300
H 3.84734400 1.03147300 0.76676000
C 4.59040400 0.65198300 -1.09165700
H 4.40437500 1.61460200 -1.58681400
H 5.61651800 0.64590900 -0.72625400
H 4.49834600 -0.13379200 -1.84169900

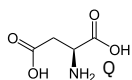


Nom : Acide 2,4,6-triméthylbenzoïque

Énergie électronique finale : -538.941889 E_n

Géométrie optimisée :

C 1.55489700 -1.17926800 0.07322000
C 0.16412500 -1.23233100 0.04941000
C -0.55457800 -0.02069100 -0.01538600
C 0.12382900 1.21168100 -0.08553000
C 1.51751500 1.20268500 -0.08616000
C 2.25213400 0.02438900 0.00530400
H 2.10817000 -2.10781300 0.13989600
H 2.04207600 2.14678100 -0.16428100
C -0.51107400 -2.58116000 0.10793900
H -0.96525200 -2.83548700 -0.84873500
H 0.21458300 -3.35133700 0.36338400
H -1.31070900 -2.60760100 0.84756200
C -0.58292400 2.54242400 -0.19960600
H -1.38644100 2.51837300 -0.93474700
H -1.02870700 2.83857000 0.74930800
H 0.12389900 3.31407000 -0.49919000
C 3.75830500 0.04872200 0.04976000
H 4.15566100 0.95031100 -0.41465300
H 4.11584100 0.02624200 1.08199700
H 4.18500000 -0.81467400 -0.46022300
C -2.04306000 -0.08615100 -0.05259600
O -2.69436400 -0.86388900 -0.70909900
O -2.64682900 0.81983200 0.76128000
H -3.59960000 0.69323500 0.64409300

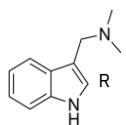


Nom : Acide L-aspartique

Énergie électronique finale : -512.504647 E_n

Géométrie optimisée :

```
C 0.69623000 0.45531500 0.79332500
H 1.20370200 0.47241700 1.76083900
C -0.76263100 0.06465700 1.06001800
H -1.22157600 0.76940500 1.75761100
H -0.80416400 -0.91681700 1.52900700
O -2.80757700 -0.55706800 0.04026100
H -3.30314600 -0.52953600 -0.79143300
N 0.93443100 1.75365900 0.15645900
H 0.25018300 1.92164300 -0.57197100
H 0.91174600 2.51594900 0.81880700
O 2.39752700 -0.20388200 -0.79259300
H 2.34232100 0.77415900 -0.74927700
C 1.42680400 -0.65519300 0.00680600
O 1.17788100 -1.82342300 0.14475100
C -1.61773100 0.03493700 -0.18287600
O -1.31484600 0.49923300 -1.25647300
```

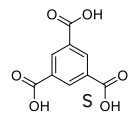


Nom : Gramine

Énergie électronique finale : -537.257696 E_h

Géométrie optimisée :

```
C 0.70158900 -0.08472900 -0.23895800
C 1.83524200 0.71845500 0.04317600
C 3.09854500 0.17259400 0.26836000
C 3.21949400 -1.20624100 0.20713300
C 2.10850700 -2.02284300 -0.06952300
C 0.85502900 -1.47746400 -0.28943300
C -0.41822900 0.80553600 -0.41673900
C 0.06411000 2.07272800 -0.23754400
H 3.95418000 0.79908800 0.48300000
H 4.18498600 -1.66246200 0.37746600
H 2.23892200 -3.09576500 -0.10478600
H 0.00103200 -2.11071100 -0.48236300
H -0.45626200 3.01410400 -0.29009400
N 1.41675300 2.03218500 0.03585400
H 1.99712300 2.83129900 0.20929700
C -1.82197400 0.44042000 -0.78814500
H -2.41803900 1.36541800 -0.88790900
H -1.82534600 -0.03547400 -1.77247200
N -2.45761300 -0.50284400 0.13740100
C -2.61886000 0.05055900 1.47412600
H -3.29031200 0.92686500 1.49253800
H -3.03602200 -0.70944500 2.13435400
H -1.65253000 0.35465900 1.87073200
C -3.72673400 -0.97678200 -0.39263800
H -4.14831000 -1.72646000 0.27672500
H -4.47080900 -0.16990000 -0.51075900
H -3.57291400 -1.43999800 -1.36740700
```

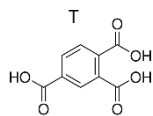


Nom : Acide 1,3,5-Benzènetricarboxylique

Énergie électronique finale : -798.241723 E_h

Géométrie optimisée :

```
C -1.33734500 -0.39084500 0.00001200
C -0.33217100 -1.35384400 -0.00000400
C 1.00753600 -0.96274200 0.00000600
C 1.33889200 0.38925100 0.00003300
C 0.33036900 1.35401200 0.00005200
C -1.00622000 0.96487500 0.00004500
H -2.37760000 -0.67678600 0.00000600
H 1.77545000 -1.72047700 -0.00001400
H 0.60279200 2.39788100 0.00007100
C -2.11982000 1.95405500 0.00005700
O -3.29090900 1.66422700 -0.00000800
C -0.63269900 -2.81264300 -0.00005100
O 0.20322000 -3.68251500 0.00004900
C 2.75254000 0.85846800 0.00003900
O 3.08781200 2.01737200 0.00005400
O 3.64573000 -0.15775000 -0.00010300
H 4.52566900 0.24443900 -0.00014000
O -1.95941400 -3.07779300 -0.00003000
H -2.05146800 -4.04090400 0.00004000
O -1.68713900 3.23594400 -0.00008000
H -2.47573300 3.79644700 -0.00015700
```



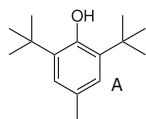
Nom : Acide 1,2,4-Benzènetricarboxylique

Énergie électronique finale : -798.229762 E_h

Géométrie optimisée :

```
C -1.33702000 -1.56499700 -0.00455700
C -1.79452800 -0.24777100 0.00259800
C -0.87736700 0.80093100 -0.01060600
C 0.49133600 0.54778000 -0.02391900
C 0.94752400 -0.78049300 -0.00907500
C 0.02656400 -1.82385900 -0.01345900
H -2.04359600 -2.38064400 -0.00296300
H -1.24234000 1.81585200 -0.02593900
H 0.38600400 -2.84258200 -0.02004500
C -3.24385900 0.09105600 0.01197200
O -3.68693600 1.21286900 0.00496700
O -4.04150700 -1.00426300 0.03037700
H -4.95382300 -0.68226600 0.03481900
C 2.39963700 -1.16349500 0.02537200
O 2.92454400 -1.86611700 -0.79734500
C 1.45926900 1.67250800 -0.16153600
O 2.58281500 1.56000900 -0.58430500
O 0.93556500 2.86349000 0.20776300
H 1.61702000 3.53044400 0.04323300
O 3.01803000 -0.72154800 1.13823200
H 3.94730200 -0.98628800 1.07263900
```

Modélisation avec solvataion implicite (SMD, THF) à B3LYP/6-311+G(2d,2p) d'une
seule molécule avec Gaussian 9



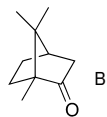
Nom : BHT

Énergie électronique finale : -8162864 kJ/mol

Géométrie optimisée :

C	1.24649700	0.16213700	0.00018900
C	1.20483300	1.55535100	-0.00383600
C	0.01146900	2.27460300	-0.00450700
C	-1.18041200	1.56172200	-0.00331000
C	-1.22796800	0.16290100	0.00048400
C	0.00725400	-0.51795000	0.00290800
H	2.12930600	2.11241200	-0.00729500
H	-2.10157200	2.12285600	-0.00617400
C	2.59800100	-0.59016700	-0.00045700
C	3.79414500	0.38033900	-0.00353900
H	3.80906200	1.02053900	0.87945500
H	4.71836600	-0.19908700	-0.00347100
H	3.80721800	1.01725800	-0.88892500
C	2.72970700	-1.46797500	-1.26608400
H	3.69860900	-1.97198500	-1.26716800
H	1.95372400	-2.22615900	-1.32178300
H	2.67339400	-0.85442200	-2.16745200
C	2.73312200	-1.46395300	1.26759800
H	1.95741700	-2.22206900	1.32774400
H	3.70211000	-1.96778800	1.26769100
H	2.67910000	-0.84757800	2.16718300
C	-2.59260200	-0.57356500	-0.00054000
C	-2.76176300	-1.43663000	1.27498700
H	-3.74435300	-1.91122900	1.27046000
H	-2.02677600	-2.23441900	1.38878400
H	-2.69259100	-0.81265600	2.16686000
C	-2.75799500	-1.43873400	-1.27507700
H	-3.73956100	-1.91548700	-1.27160500
H	-2.68869900	-0.81574400	-2.16763100
H	-2.02069900	-2.23471500	-1.38628600
C	-3.77530500	0.41455700	-0.00282400
H	-3.77855000	1.05282800	0.88084100
H	-3.77675600	1.05095100	-0.88782400
H	-4.70754800	-0.15124800	-0.00322800
C	0.02182800	3.78399600	0.00533400
H	0.65766300	4.18498600	-0.78653600
H	-0.98093800	4.18683100	-0.13556300

H	0.40465600	4.17476500	0.95157000
O	0.07198200	-1.89195100	0.00566200
H	-0.81329700	-2.26301000	0.00710100

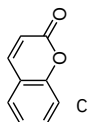


Nom : (+)-Camphre

Énergie électronique finale : -12223696 kJ/mol

Géométrie optimisée :

C	0.87830500	1.24019300	1.06964600
C	-0.58482700	1.27219800	0.59039500
C	0.29290300	-0.51120700	-0.54515000
H	1.43819100	2.14869100	0.84459000
H	0.98043100	1.05047700	2.13842000
C	-0.92067000	-0.24499700	0.42017100
C	-0.55846200	1.78378100	-0.86603700
H	-1.55911900	2.00247700	-1.23442900
H	0.02625500	2.69864000	-0.95566000
C	0.08245200	0.59337000	-1.63534500
H	-0.57837900	0.20642800	-2.40986000
H	1.02144600	0.85810500	-2.12116700
H	-1.25736700	1.81845800	1.24930300
C	1.44163200	0.05324000	0.28932300
O	2.58611200	-0.34497100	0.31828500
C	-2.29384600	-0.50653800	-0.21039900
H	-2.45058900	-1.57624800	-0.35886400
H	-3.08151400	-0.15448400	0.45985600
H	-2.43793600	-0.01678300	-1.17068700
C	-0.85277000	-1.05193500	1.72599300
H	-1.60270600	-0.69081200	2.43295200
H	-1.06733400	-2.10450500	1.53492200
H	0.11520300	-1.00617800	2.22334400
C	0.51880100	-1.90680800	-1.08639800
H	1.40131800	-1.93488800	-1.72721400
H	0.67606000	-2.62925600	-0.28433600
H	-0.33396300	-2.23813500	-1.68063900

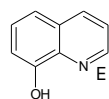


Nom : Coumarine

Énergie électronique finale : -1305375 kJ/mol

Géométrie optimisée :

C	-2.59484600	-1.00417400	-0.00000900
C	-1.29344600	-1.48355300	0.00000300
C	-0.24331200	-0.57242200	0.00000000
C	-0.47265900	0.81045700	-0.00001100
C	-1.79974100	1.26918900	-0.00002200
C	-2.85171600	0.37149000	-0.00002400
H	-3.41782300	-1.70559300	-0.00000900
H	-1.08204700	-2.54344500	0.00001200
C	0.67419400	1.67503500	-0.00001000
H	-1.98364400	2.33535500	-0.00003200
H	-3.87127700	0.73020000	-0.00003400
C	1.92207900	1.16466900	0.00000500
H	0.51812000	2.74619900	-0.00001500
H	2.80257500	1.78874500	0.00001000
C	2.15197000	-0.26682200	0.00001000
O	3.23097100	-0.81578900	0.00003700
O	1.02889800	-1.07604400	0.00001300



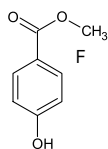
Nom : 8-Hydroxyquinoline

Énergie électronique finale : -1253194 kJ/mol

Géométrie optimisée :

C	1.99861500	-1.49183100	0.00000000
C	0.69526800	-1.92642900	0.00000000
C	-0.35869000	-0.97988700	0.00000000
C	-0.04612800	0.40657300	0.00000000
C	1.31737100	0.82073800	0.00000000
C	2.32154900	-0.11704800	0.00000000
H	-2.02499600	-2.35935300	0.00000000
H	2.80555700	-2.21227800	0.00000000
H	0.46110100	-2.98233900	0.00000000
C	-1.73281300	-1.31722400	0.00000000
H	3.35399200	0.20373100	0.00000000
C	-2.25483200	1.02650600	0.00000000
C	-2.67578600	-0.32095400	0.00000000
H	-2.98919300	1.82351000	0.00000000

H	-3.73222800	-0.54858500	0.00000000
O	1.59065400	2.14708800	0.00000000
H	0.72830700	2.59952700	0.00000000
N	-0.98787100	1.38520300	0.00000000

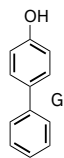


Nom : 4-Hydroxybenzoate de méthyle

Énergie électronique finale : -1736779 kJ/mol

Géométrie optimisée :

C	-1.72643200	-1.27415900	-0.00007400
C	-0.35126900	-1.11418100	-0.00003500
C	0.22151700	0.16383500	0.00004300
C	-0.62012900	1.28013600	0.00008200
C	-1.99696500	1.13128400	0.00004700
C	-2.55337000	-0.14974300	-0.00003500
H	-2.17059600	-2.25963300	-0.00013700
H	0.28290100	-1.98769400	-0.00006500
H	-0.18562500	2.26924200	0.00014200
H	-2.64153000	2.00146700	0.00008300
O	-3.89707500	-0.35870800	-0.00008000
H	-4.36126200	0.48825900	-0.00004400
C	1.68827400	0.38118500	0.00007300
O	2.22066600	1.47296300	0.00001500
O	2.38488700	-0.77076700	0.00003500
C	3.82024300	-0.64535700	-0.00003200
H	4.19982200	-1.66259900	-0.00006600
H	4.15858500	-0.11746100	-0.88940600
H	4.15867000	-0.11749000	0.88932700



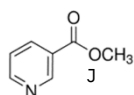
Nom : *p*-phénylphénol

Énergie électronique finale : -1414374 kJ/mol

Géométrie optimisée :

C	-0.29596700	-0.00397000	0.00196400
C	-1.02643600	1.13344200	-0.36317300
C	-2.41566900	1.13457300	-0.36530500
C	-3.11151400	-0.01463800	0.00401100
C	-2.40782900	-1.15904600	0.37302100

C	-1.02053900	-1.14703200	0.36850900
C	1.18810900	0.00015200	0.00065600
C	1.91395300	-1.13955300	-0.37371800
C	3.30445900	-1.13616300	-0.37425700
C	4.00499700	0.00822700	-0.00308000
C	3.29893100	1.14854800	0.37018700
C	1.90845000	1.14386900	0.37342900
H	-0.50648400	2.03007000	-0.67152800
H	-2.95805900	2.02424700	-0.66178500
H	-2.95168000	-2.04583800	0.66841500
H	-0.49304100	-2.03955200	0.67604700
H	1.38657400	-2.03074400	-0.68589100
H	3.84088400	-2.02692100	-0.67380400
H	5.08670500	0.01129300	-0.00447100
H	3.83110100	2.04223300	0.66862800
H	1.37670700	2.03182000	0.68748900
O	-4.47712800	-0.07327800	0.02129300
H	-4.84136100	0.77915700	-0.24690800

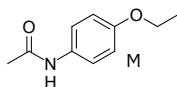


Nom : Nicotinate de méthyle

Énergie électronique finale : -1250591 kJ/mol

Géométrie optimisée :

C	-2.92134000	-0.36721100	0.00008400
C	-2.49324800	0.95622900	0.00008300
C	-1.13159600	1.21176200	-0.00001300
C	-0.24484500	0.13558900	-0.00010800
C	-0.77461800	-1.15724900	-0.00009300
H	-3.97919500	-0.60060800	0.00016000
H	-3.21349200	1.76174800	0.00015500
H	-0.75025400	2.22270800	-0.00002000
H	-0.11766300	-2.01598200	-0.00016600
C	1.22102500	0.40279900	-0.00022900
O	1.70373300	1.51456600	-0.00004100
O	1.94928100	-0.72238900	-0.00000200
C	3.38332400	-0.55457500	0.00019400
H	3.79113300	-1.56037700	0.00036100
H	3.70261400	-0.01680700	-0.88991600
H	3.70235100	-0.01661900	0.89028500
N	-2.08454600	-1.41079200	-0.00000200

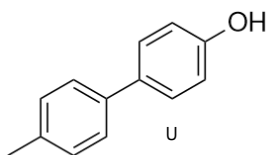


Nom : Phénacétine

Énergie électronique finale : -1223696 kJ/mol

Géométrie optimisée :

C	1.67411300	0.37103600	-0.00072600
C	0.99044900	1.59207100	-0.00016900
C	-0.39106900	1.62046800	0.00027600
C	-1.13854500	0.43363700	0.00026400
C	-0.45710900	-0.78251500	-0.00020500
C	0.93840000	-0.81248000	-0.00072700
H	1.55682300	2.51315700	-0.00013700
H	-0.90272200	2.57491900	0.00068700
H	-1.01130800	-1.70517400	-0.00015400
H	1.43091300	-1.77235500	-0.00119100
N	-2.54591300	0.55754500	0.00059300
H	-2.88396500	1.50696800	0.00110400
C	-3.50575000	-0.41369400	0.00027100
O	-3.25979900	-1.61551900	0.00001600
C	-4.93005900	0.09867600	-0.00005800
H	-5.44237000	-0.29437900	0.87830500
H	-5.00621000	1.18485700	0.00169700
H	-5.44085100	-0.29137400	-0.88066300
O	3.03528100	0.44546900	-0.00143900
C	3.78884900	-0.77723100	0.00040600
H	3.53184500	-1.36399300	-0.88512500
H	3.53046200	-1.36206200	0.88681000
C	5.25764400	-0.41788600	0.00105000
H	5.85612500	-1.32980400	0.00226600
H	5.51919700	0.16133200	-0.88496100
H	5.51806200	0.16299300	0.88630400

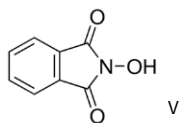


Nom : 4-(4-méthylphényl) phénol

Énergie électronique finale : -1517634 kJ/mol

Géométrie optimisée :

C	2.83228900	-1.13847400	0.36396600
C	1.44404800	-1.13742600	0.36794000
C	0.71660500	0.00402200	0.00025600
C	1.44597900	1.14003900	-0.36921100
C	2.83738000	1.13473300	-0.36974700
C	3.55961200	-0.00203200	-0.00432600
H	3.36103100	-2.03625800	0.66074800
H	0.91805600	-2.02998300	0.67932800
H	0.92386600	2.03504200	-0.68002700
H	3.36923300	2.02934900	-0.66902100
C	-0.76710900	0.00573000	0.00202400
C	-1.49671000	-1.13426200	-0.35736700
C	-1.49422800	1.14910100	0.36312500
C	-2.88614600	-1.13774800	-0.35880700
H	-0.97571500	-2.03180100	-0.66136700
C	-2.88171900	1.15936600	0.36683600
H	-0.96820900	2.04391400	0.66661200
C	-3.58392100	0.01218400	0.00407900
H	-3.42707400	-2.03001300	-0.65029900
H	-3.42672200	2.04707200	0.65746700
O	-4.95018300	0.06881600	0.02096900
H	-5.31271800	-0.78577600	-0.24242900
C	5.06621300	-0.01145700	0.00069200
H	5.46930400	0.93275100	-0.36349300
H	5.45676900	-0.17595900	1.00755500
H	5.45987400	-0.81153200	-0.62958800



Nom : n-Hydroxyphthalimide

Énergie électronique finale : -1544984 kJ/mol

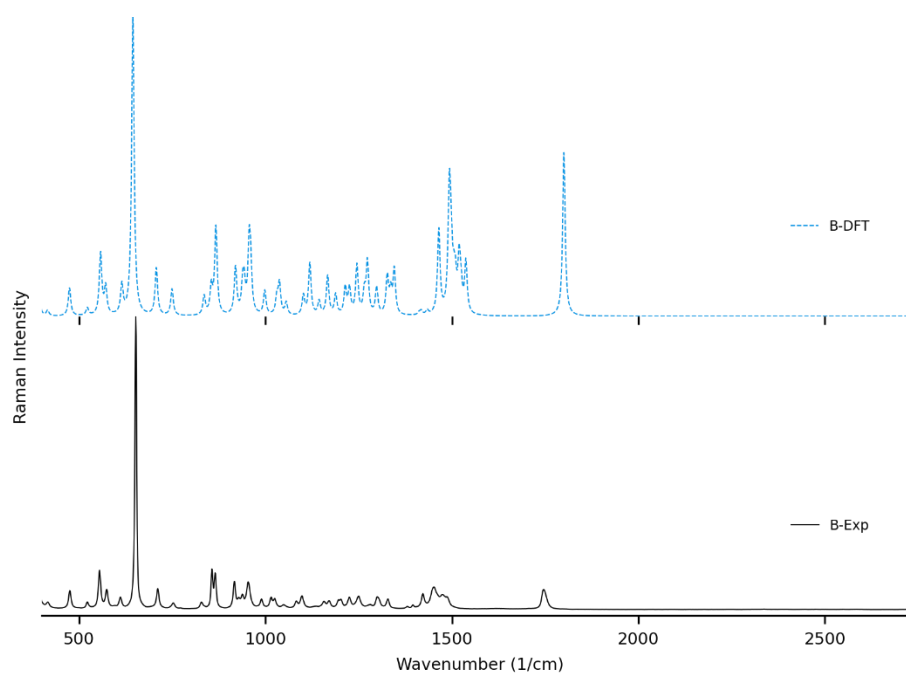
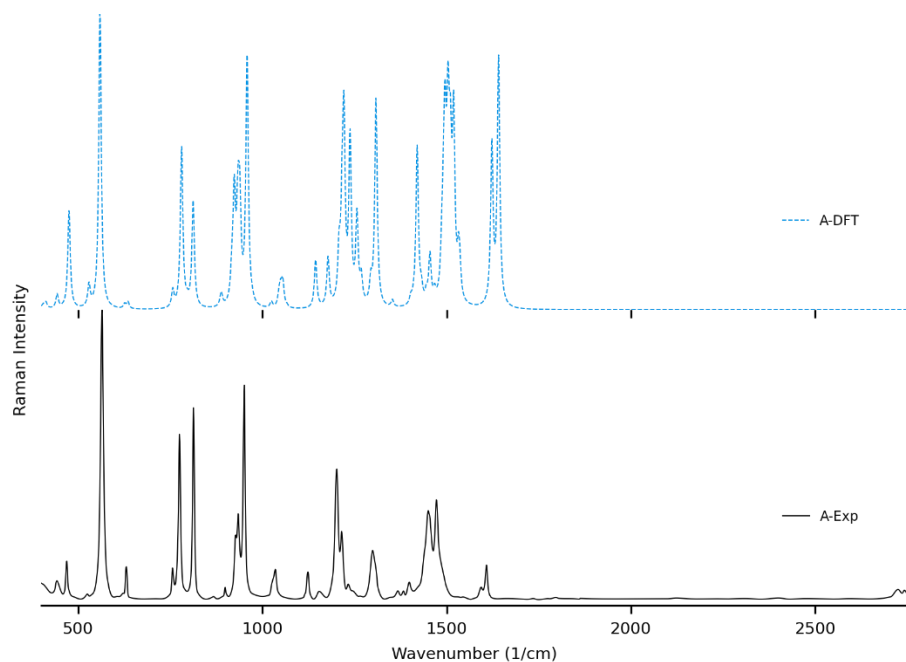
Géométrie optimisée :

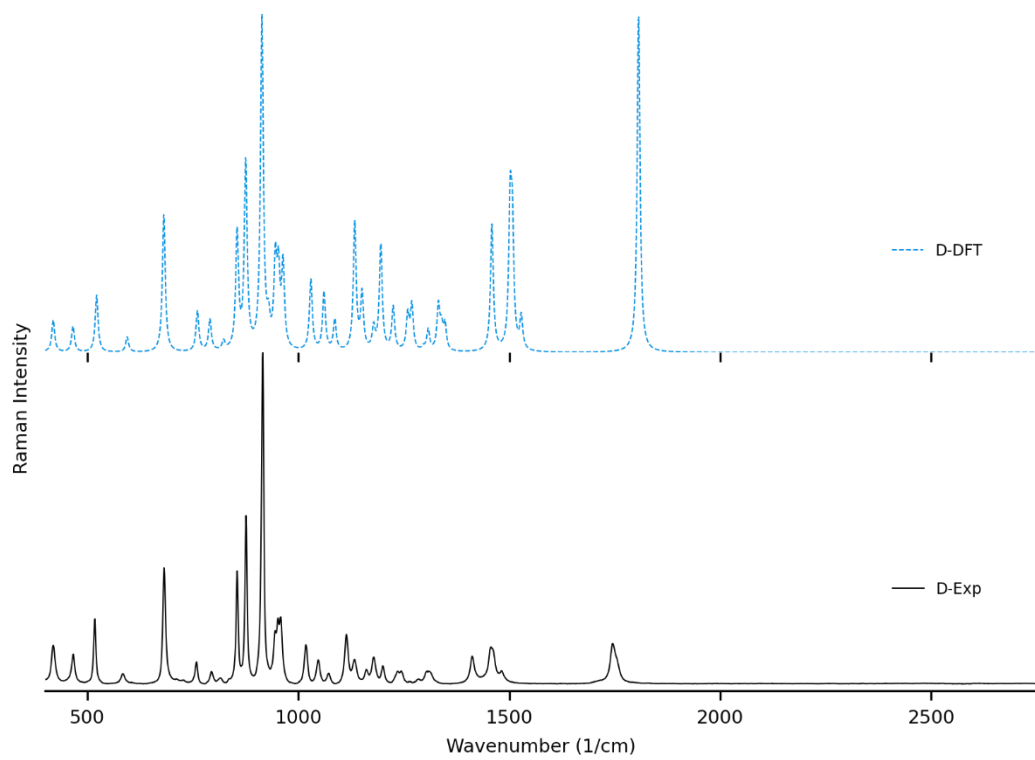
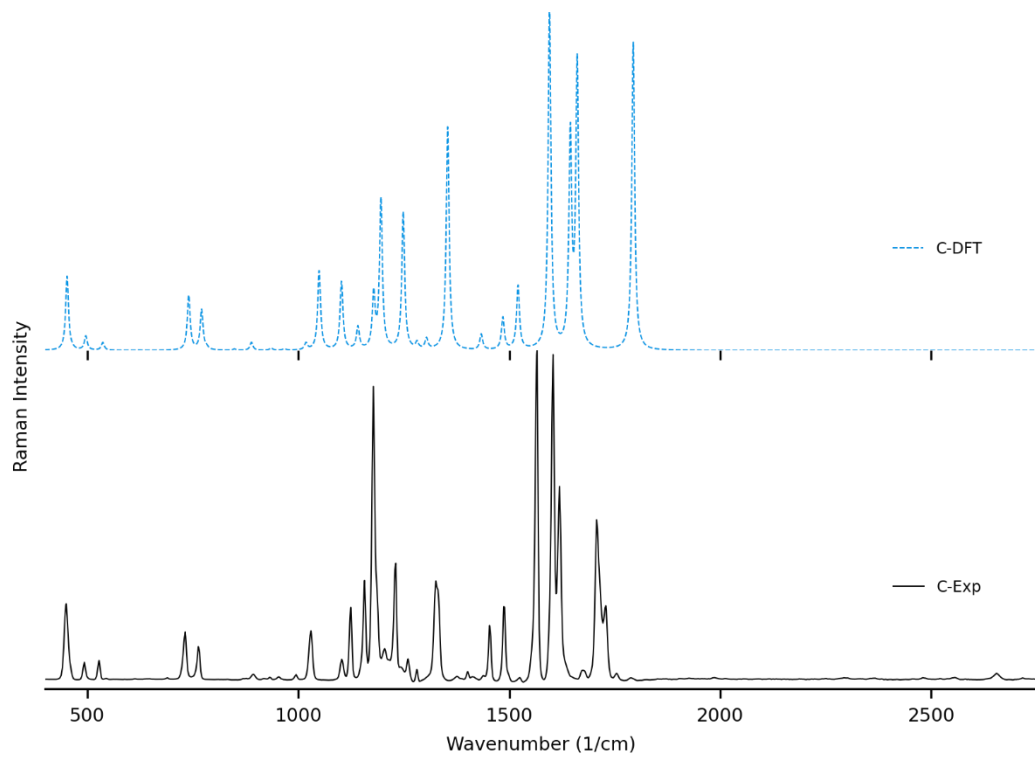
C	-2.88019800	0.69729000	-0.00022400
C	-1.68325800	1.41873900	0.00260000
C	-0.50412400	0.69793200	0.00145100
C	-0.50412500	-0.69793000	0.00145100
C	-1.68325700	-1.41873700	0.00260300
C	-2.88019800	-0.69728900	-0.00022200
H	-3.82168200	1.22866400	-0.00293400
H	-1.68143700	2.49949700	0.00348200
H	-1.68143500	-2.49949500	0.00348500
H	-3.82168300	-1.22866200	-0.00293000
C	0.90523700	-1.17786800	0.00549700
N	1.66927700	-0.00000200	0.06050000
O	3.03112300	-0.00000300	-0.11848800
H	3.42063900	0.00002300	0.76988500
C	0.90523700	1.17786800	0.00550800
O	1.34998300	-2.29917600	-0.02241000
O	1.34999000	2.29917300	-0.02241300

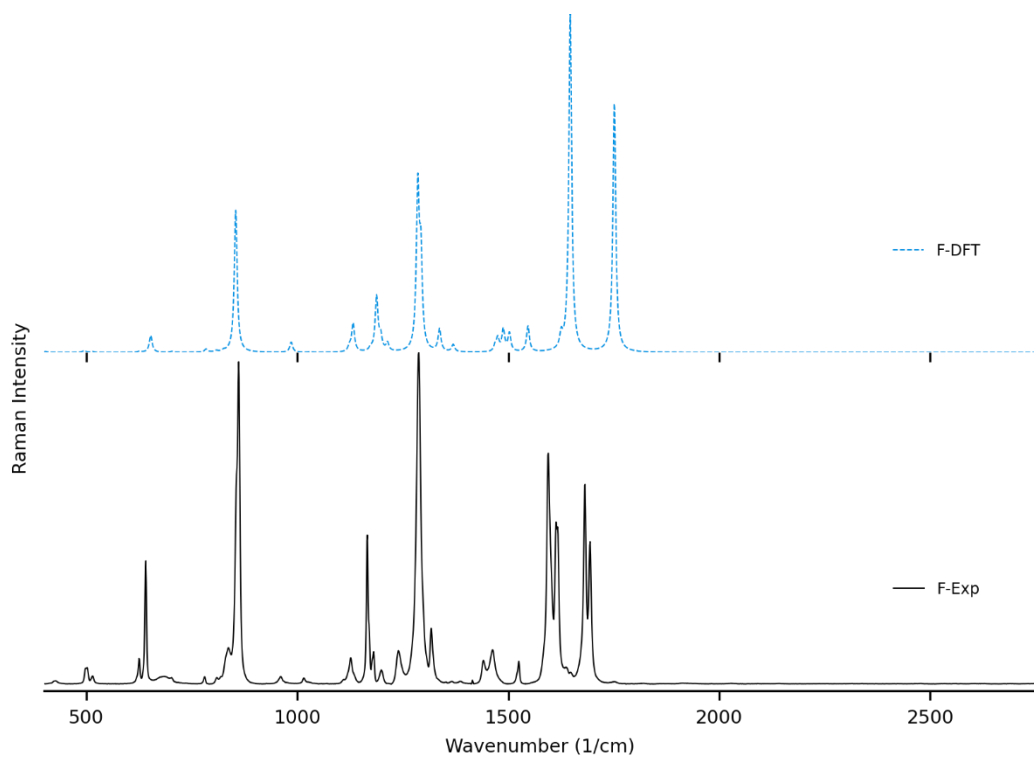
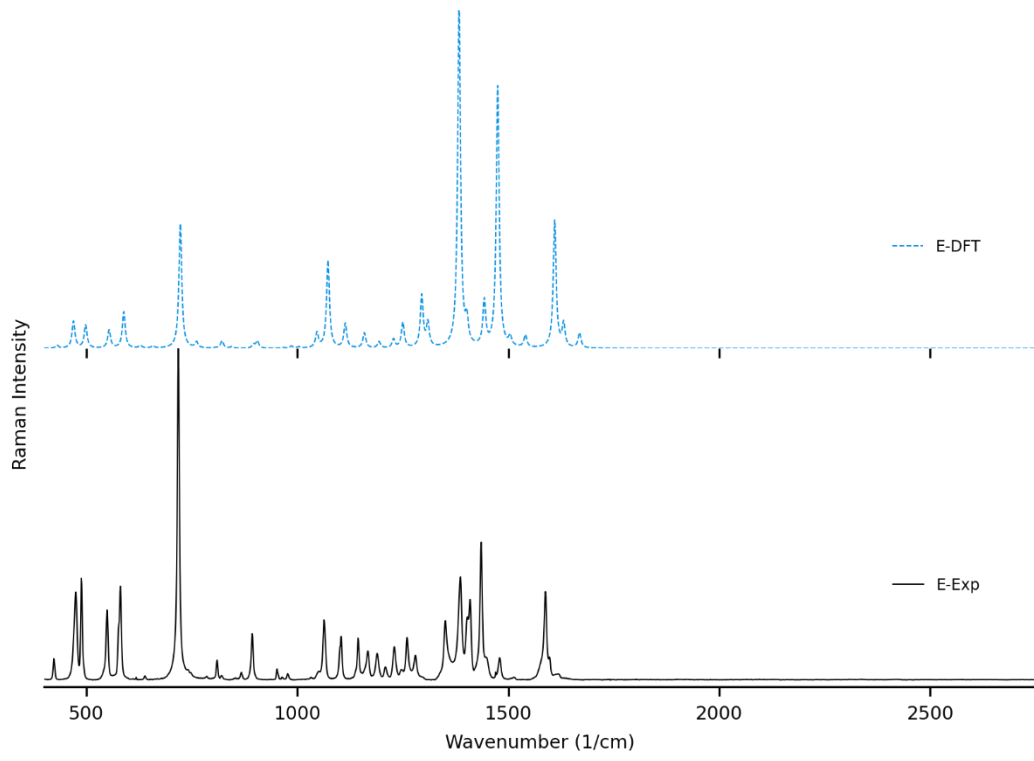
ANNEXE B

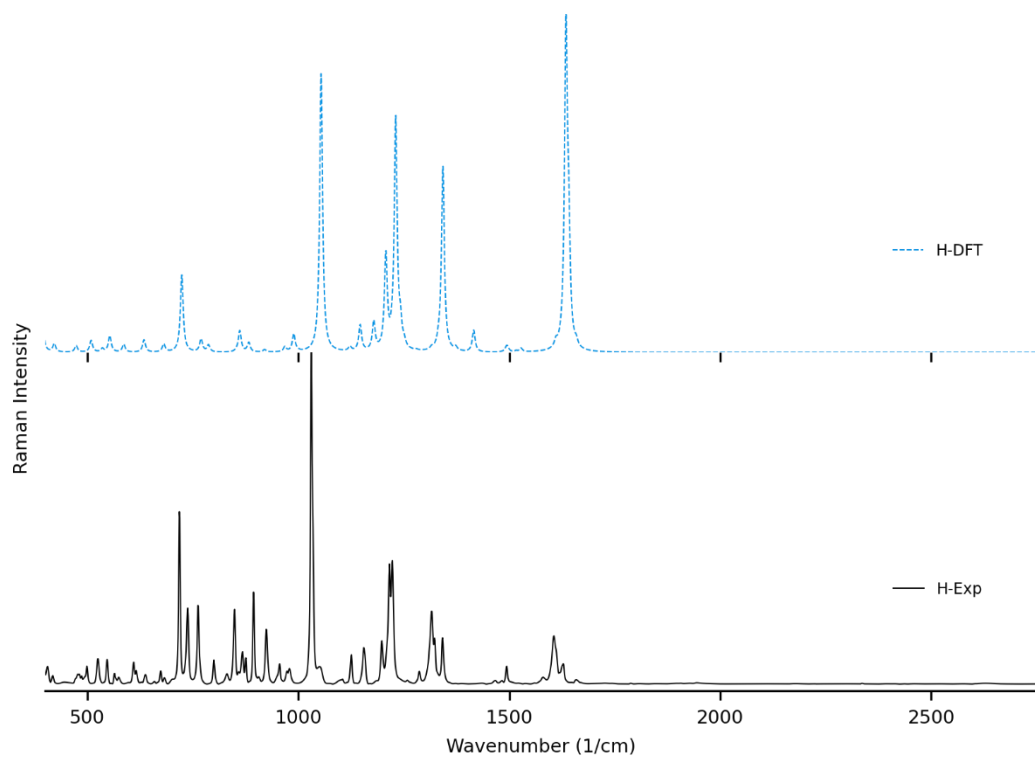
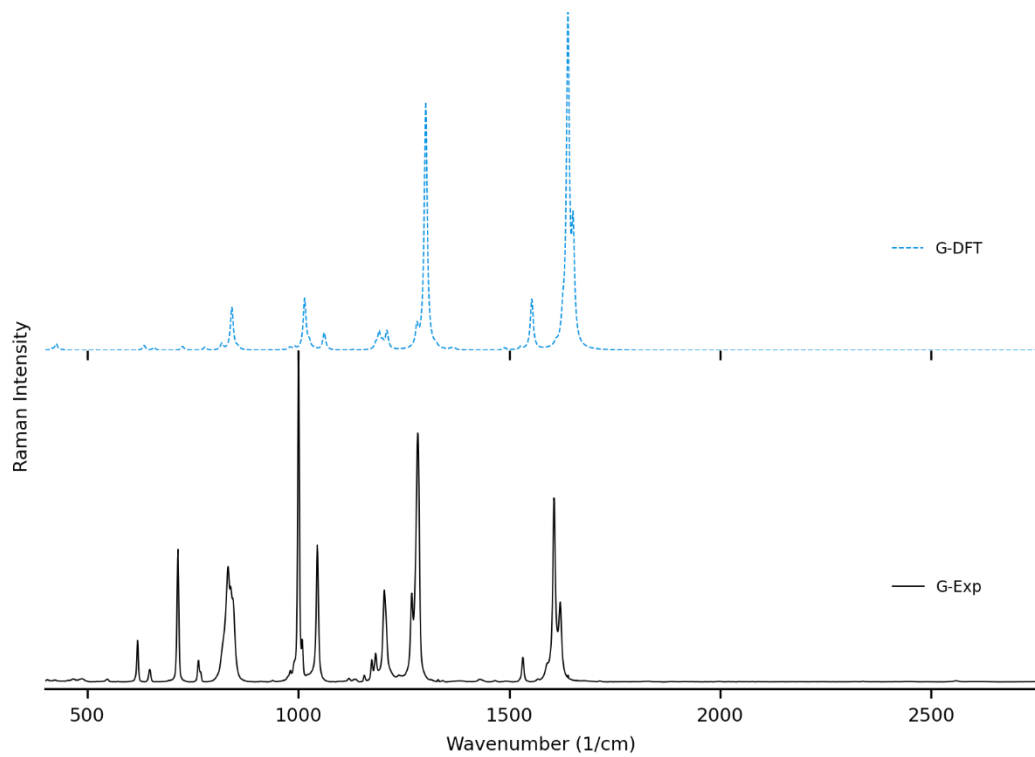
SPECTRES RAMAN THÉORIQUES ET EXPÉRIMENTAUX

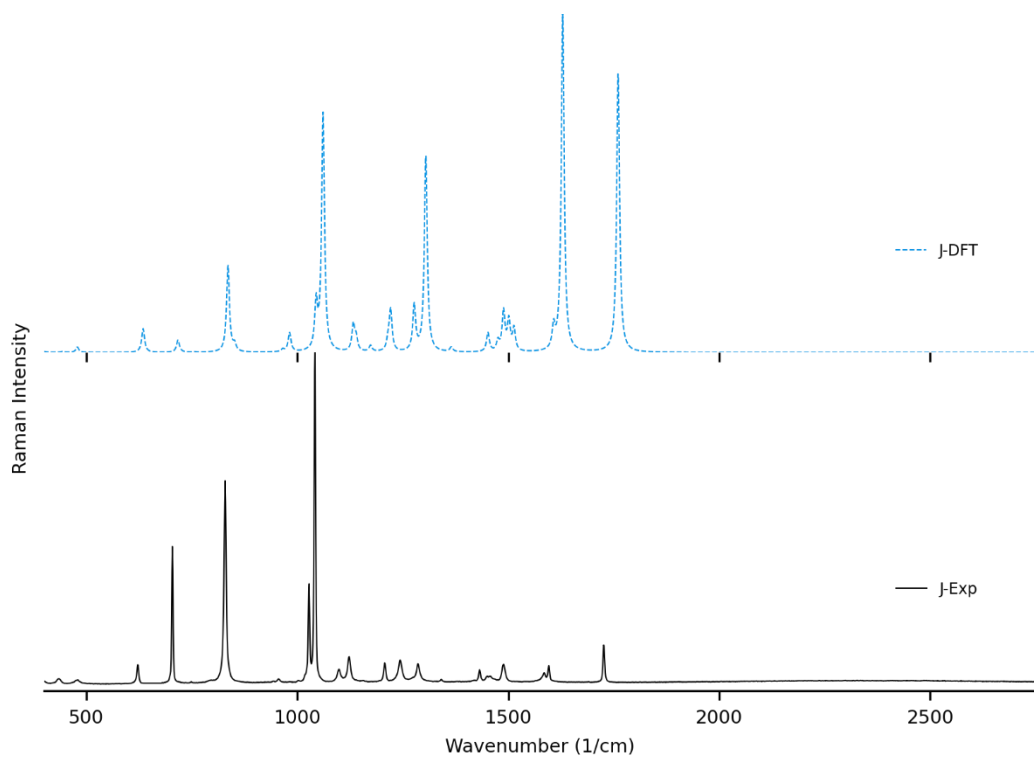
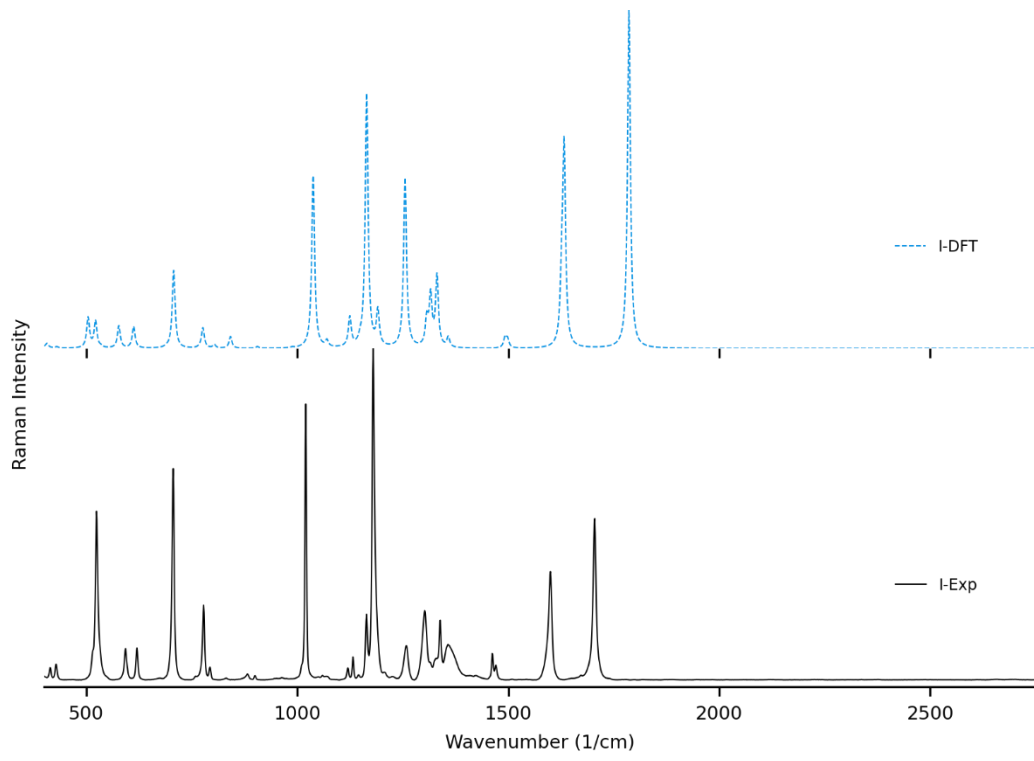
Échantillons solides, modélisés en phase gazeuse (facteur de correction 0.970)

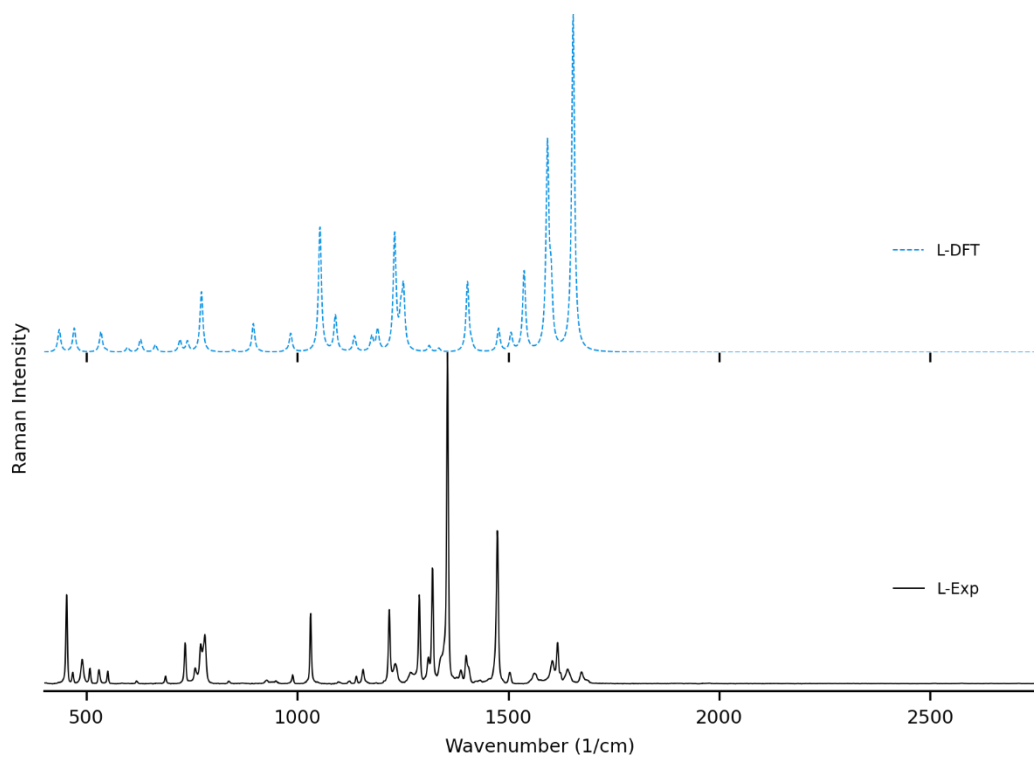
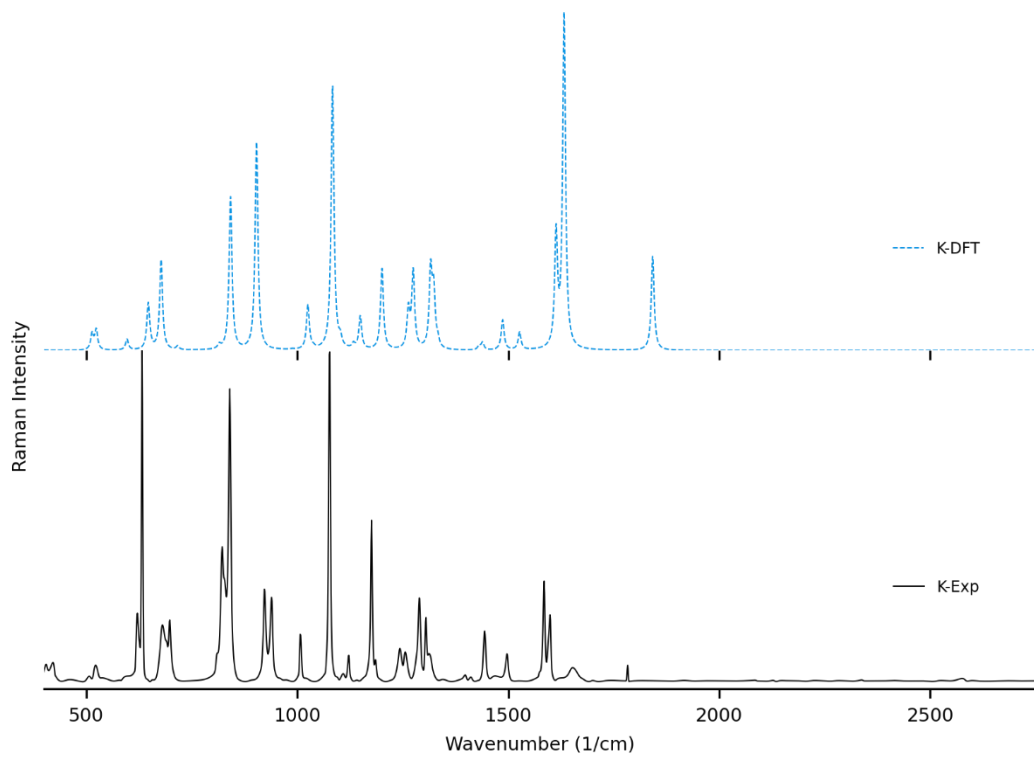


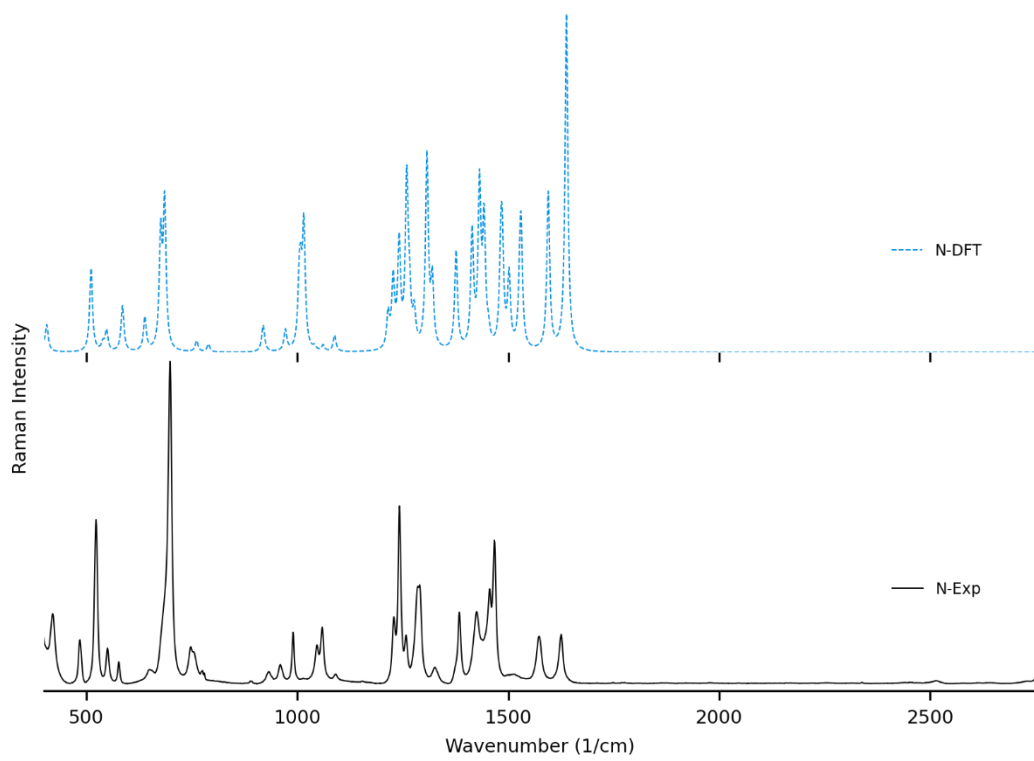
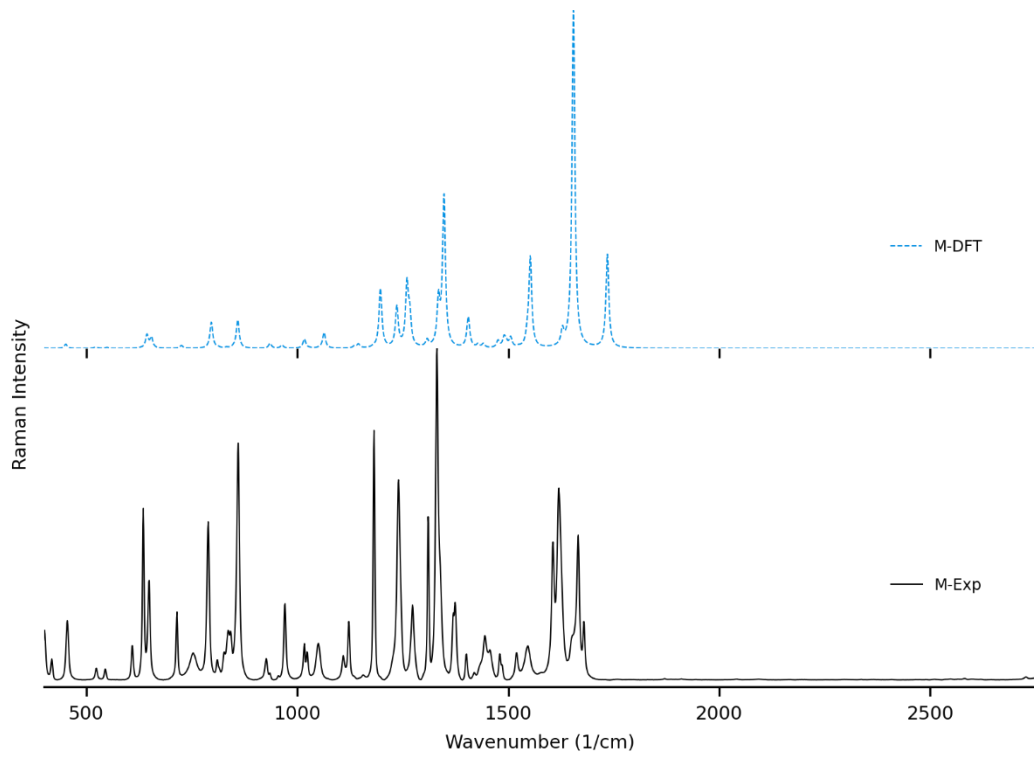


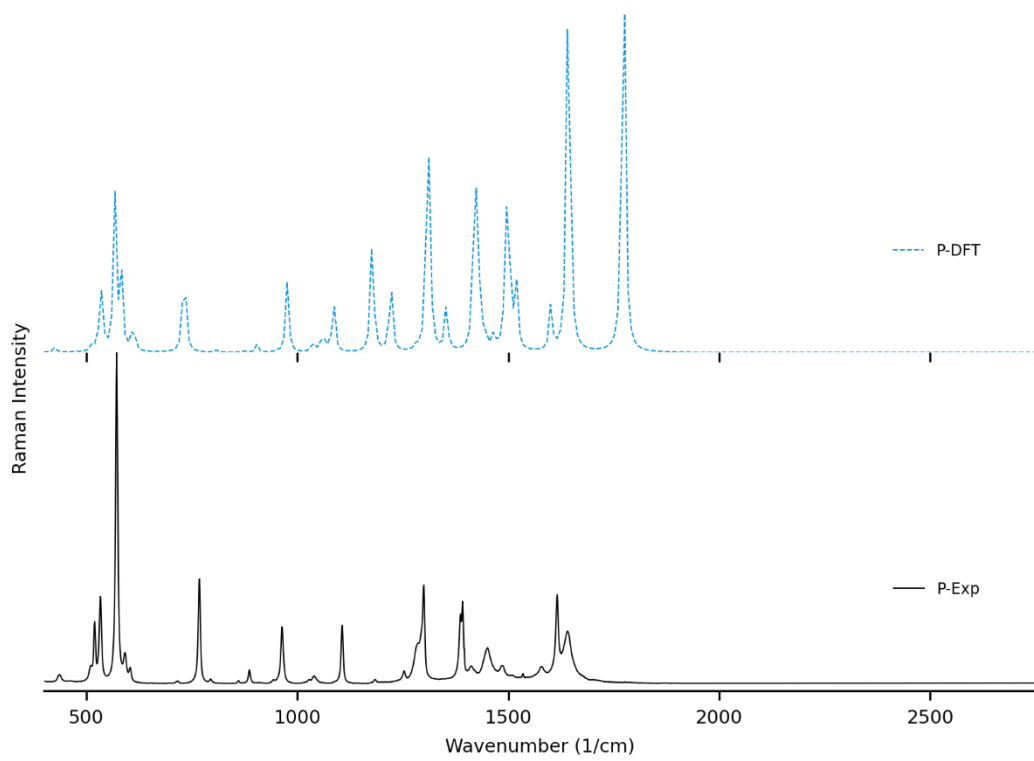
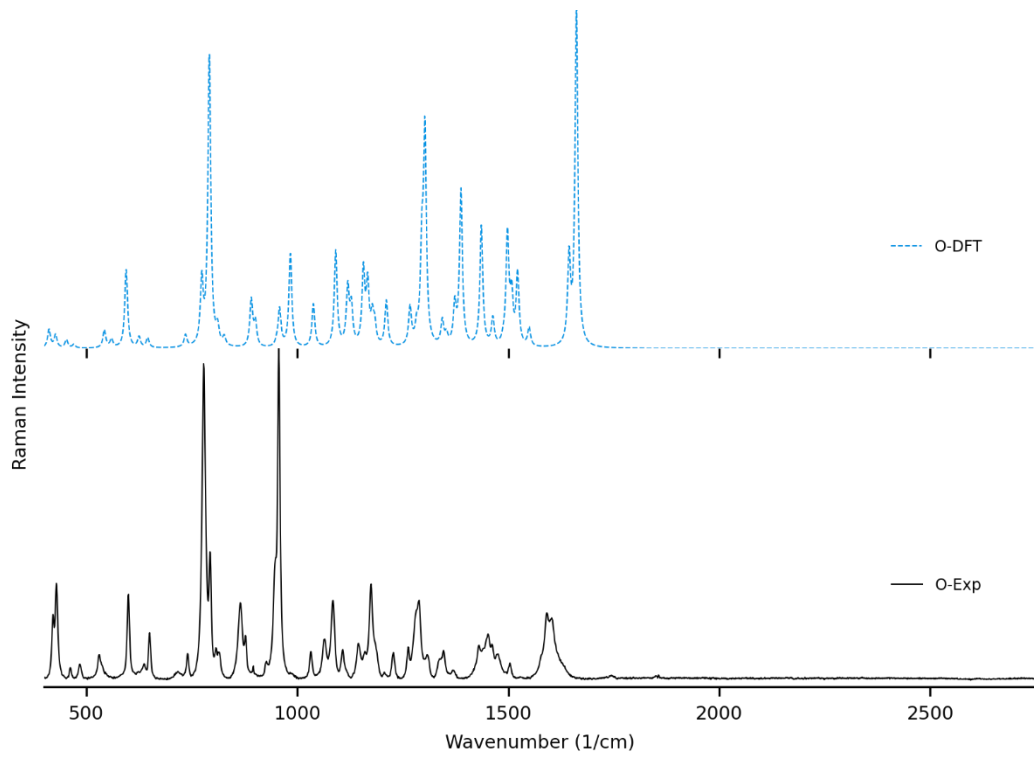


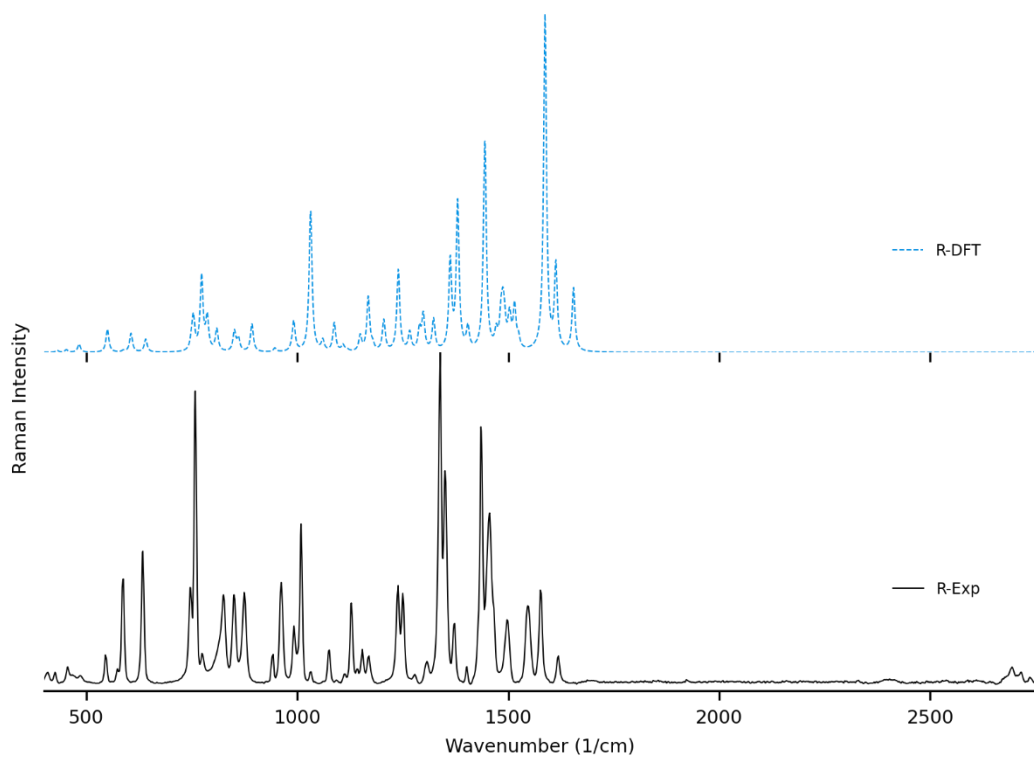
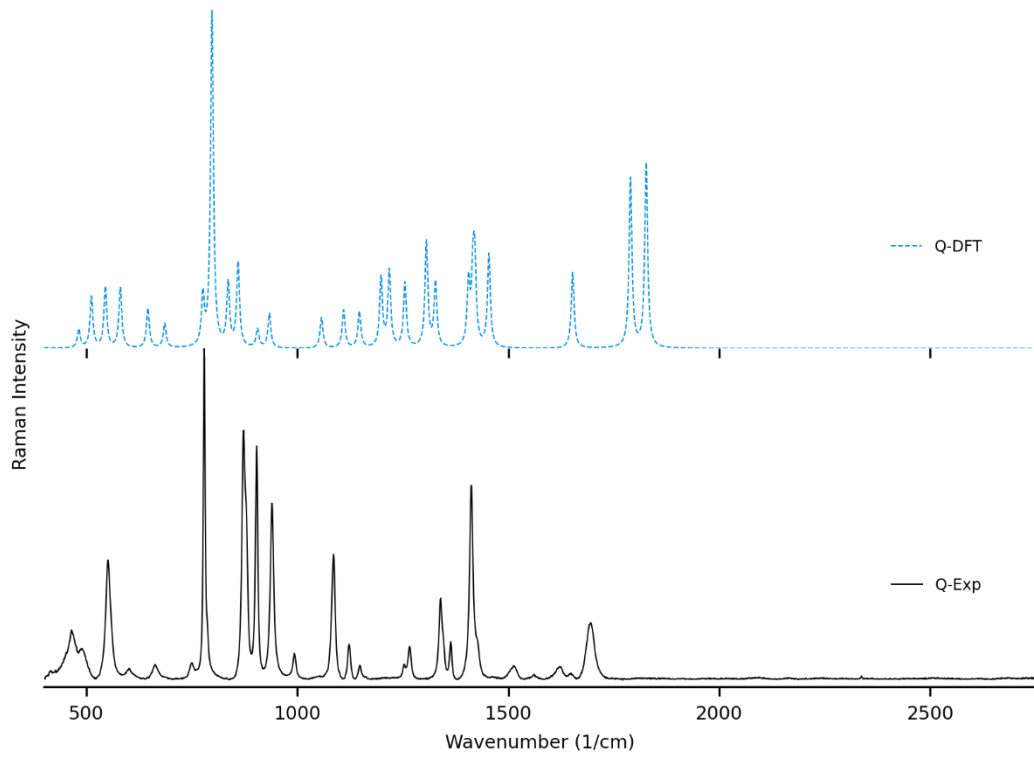


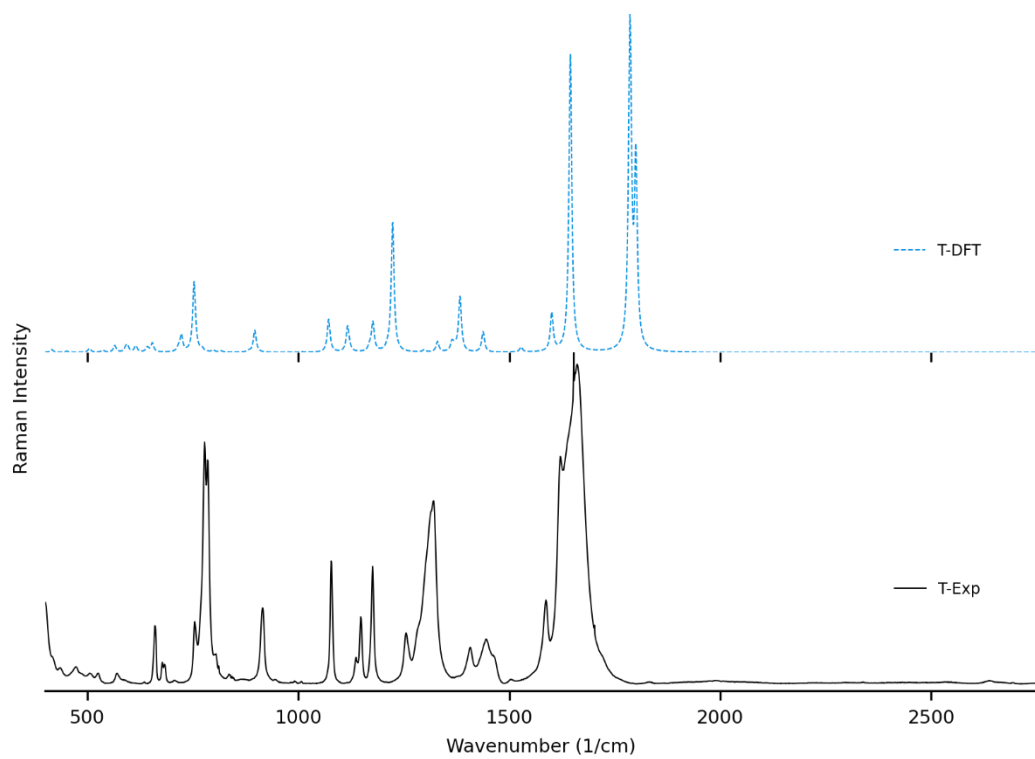
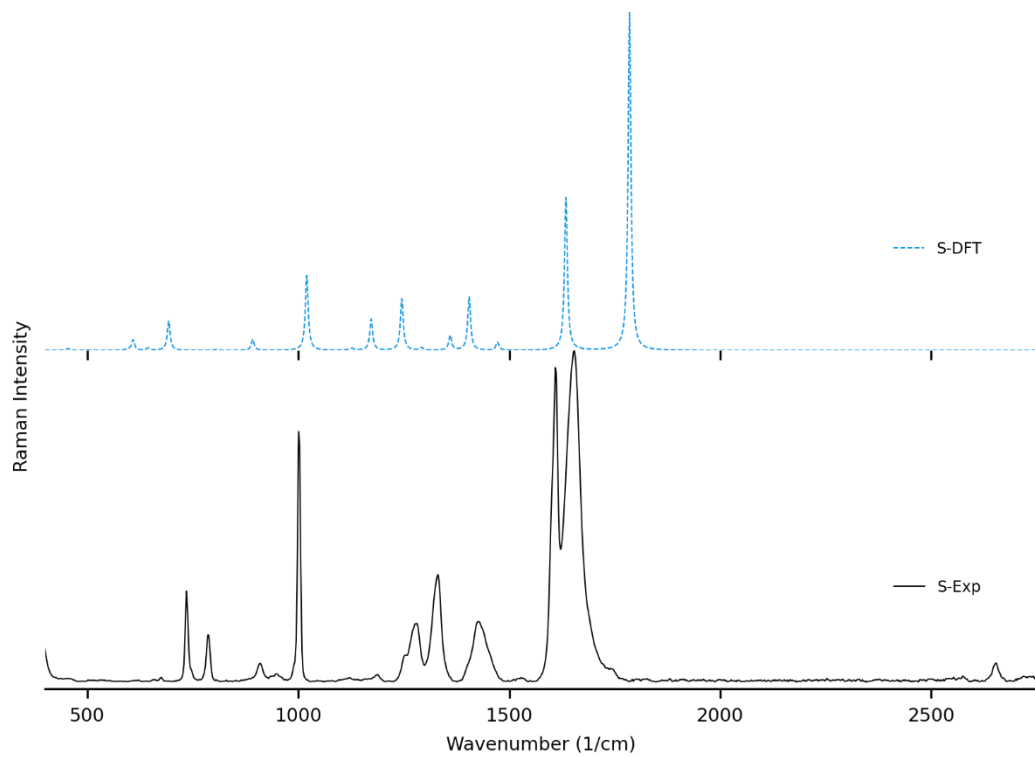




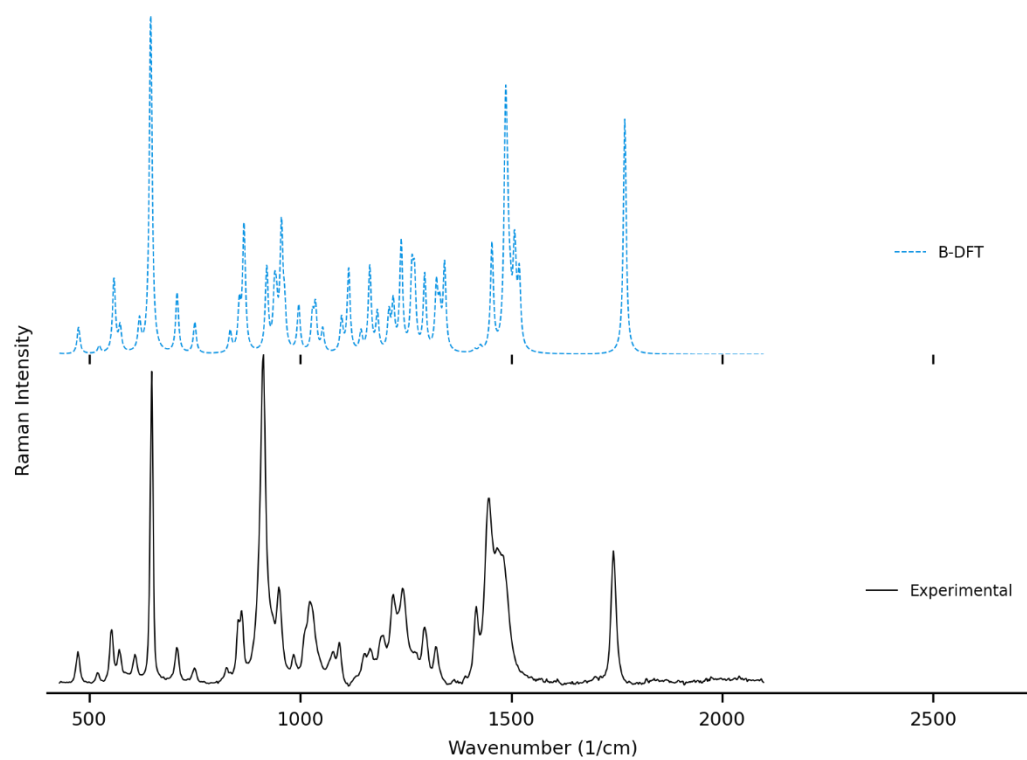
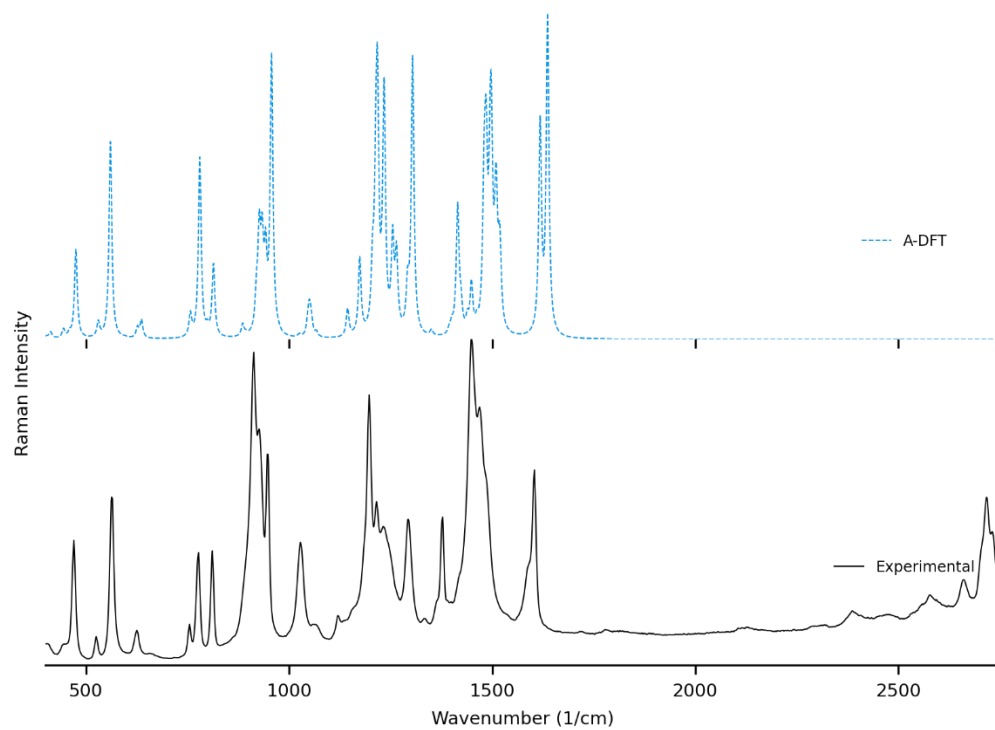


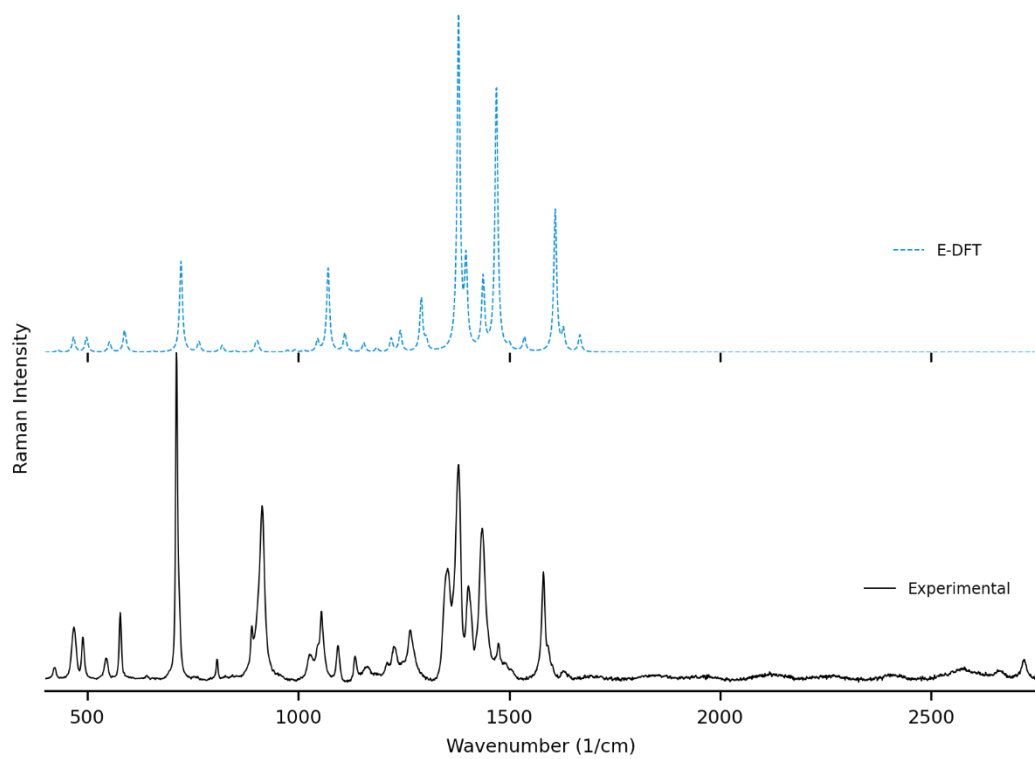
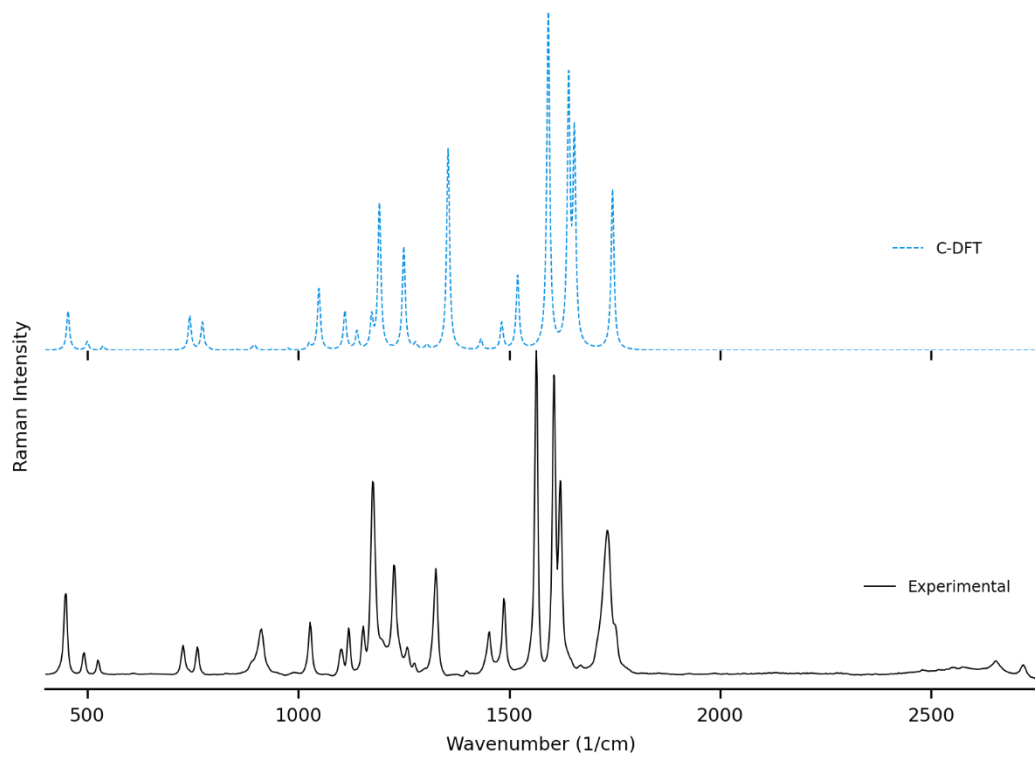


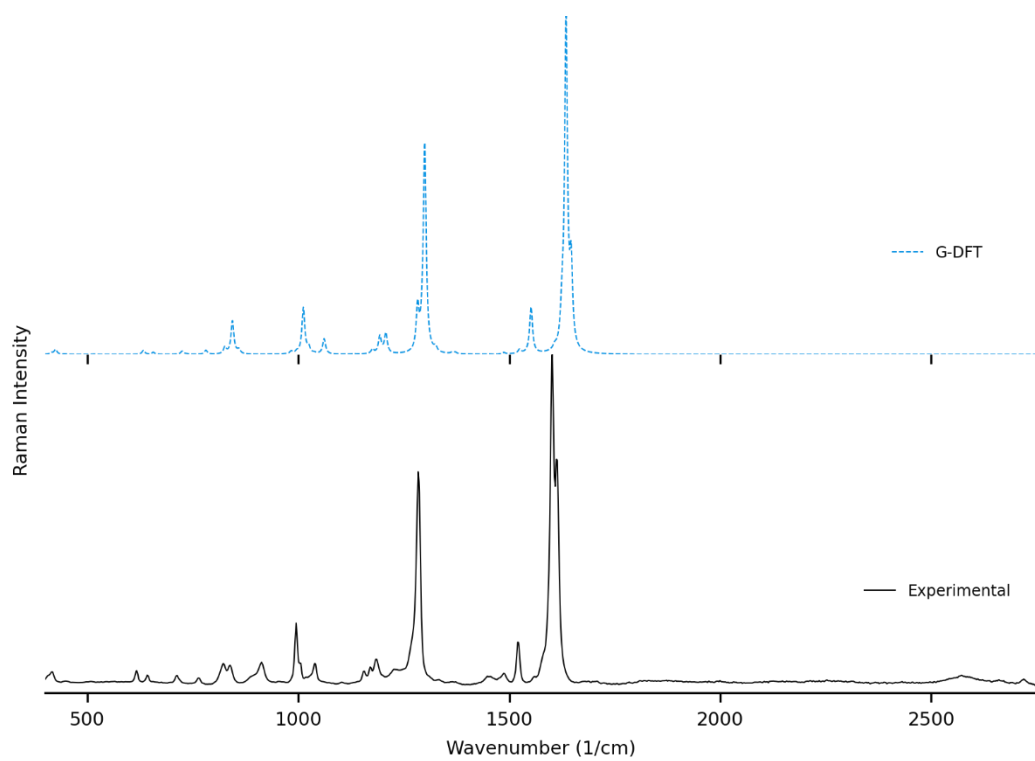
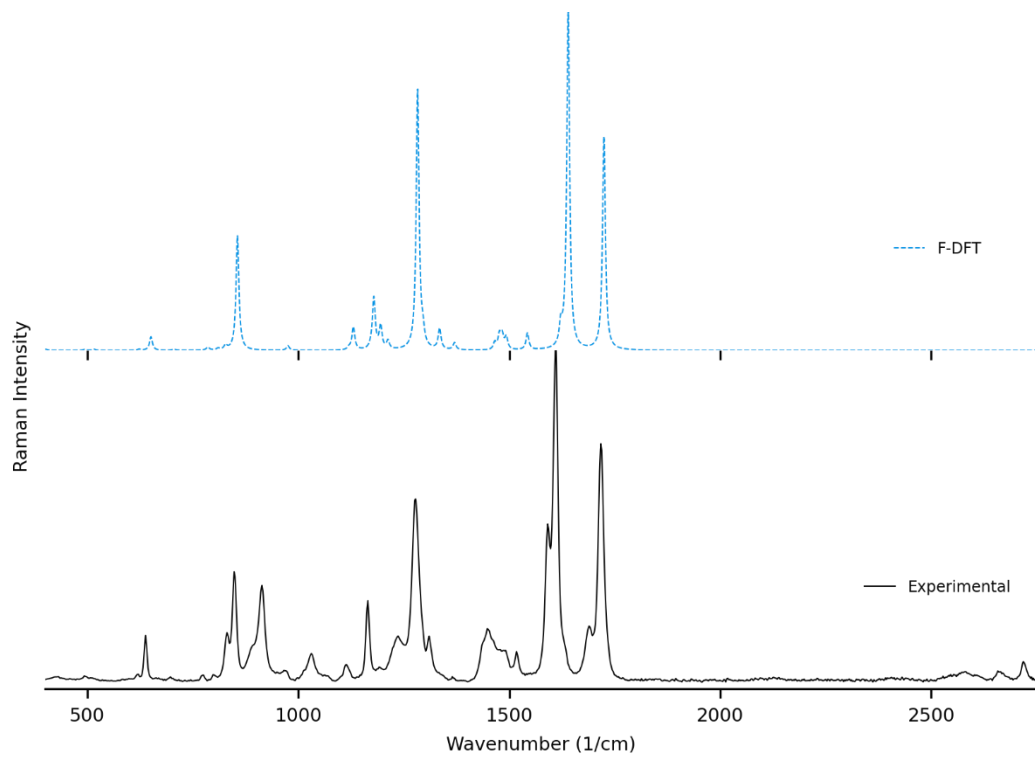


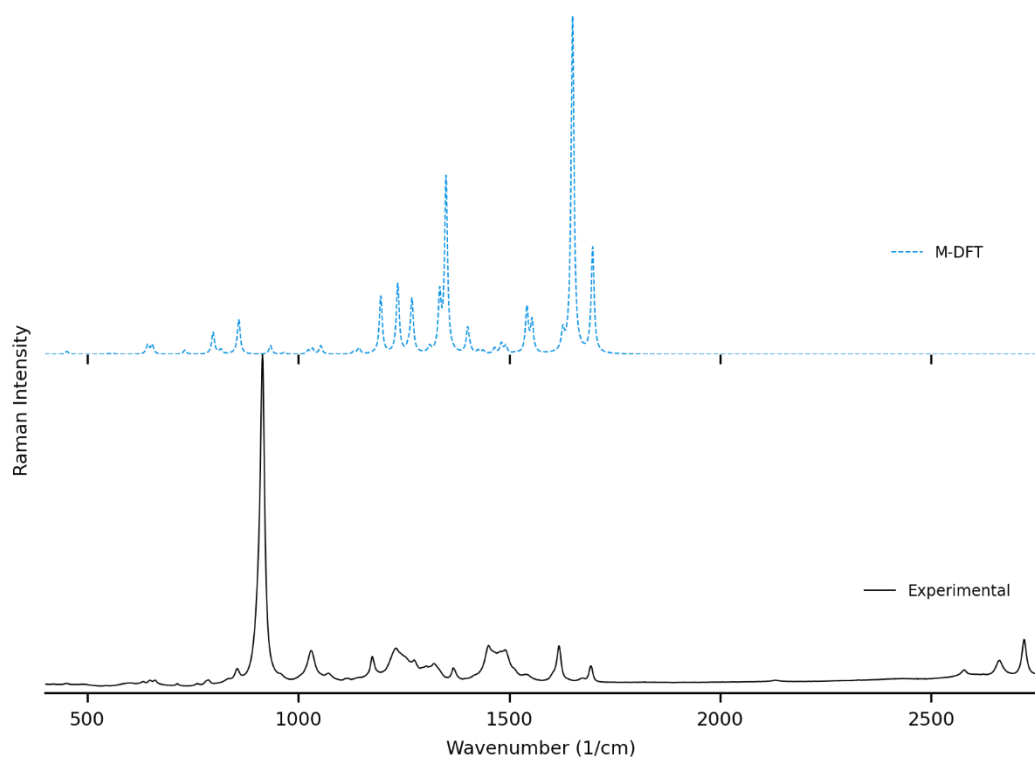
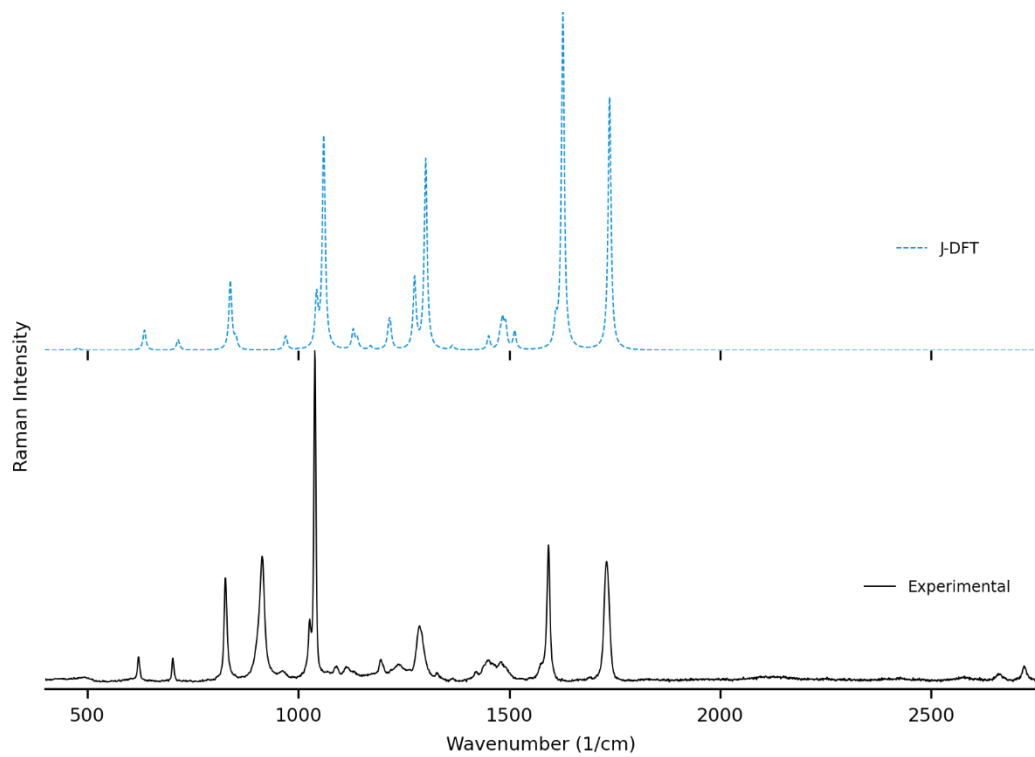


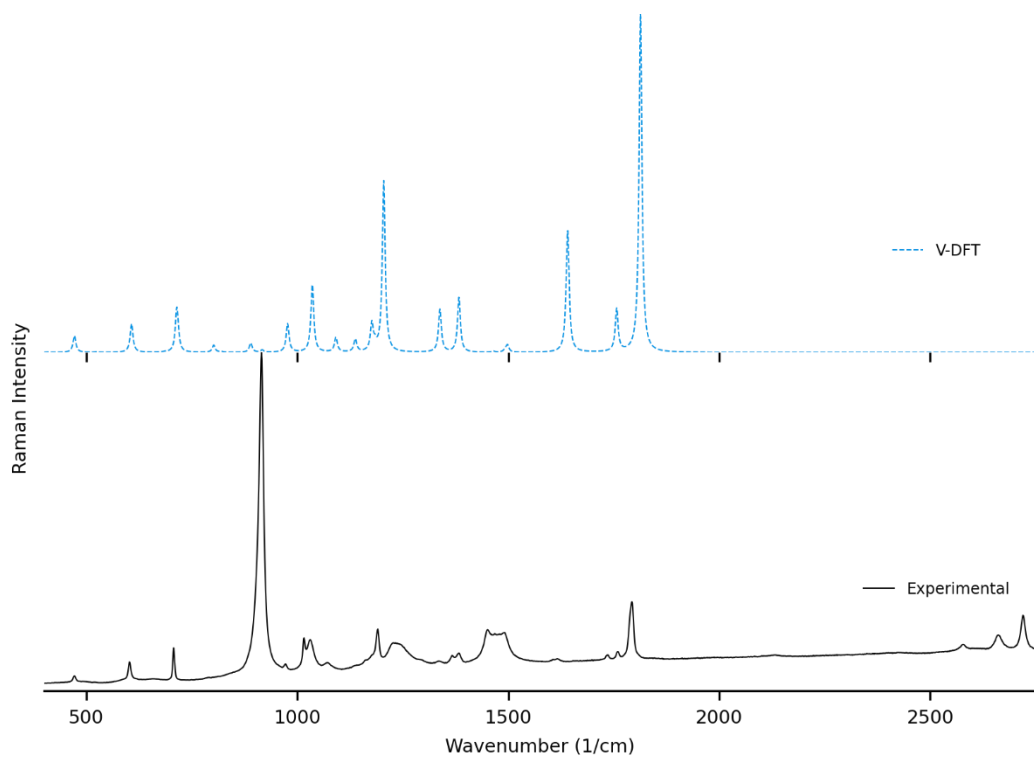
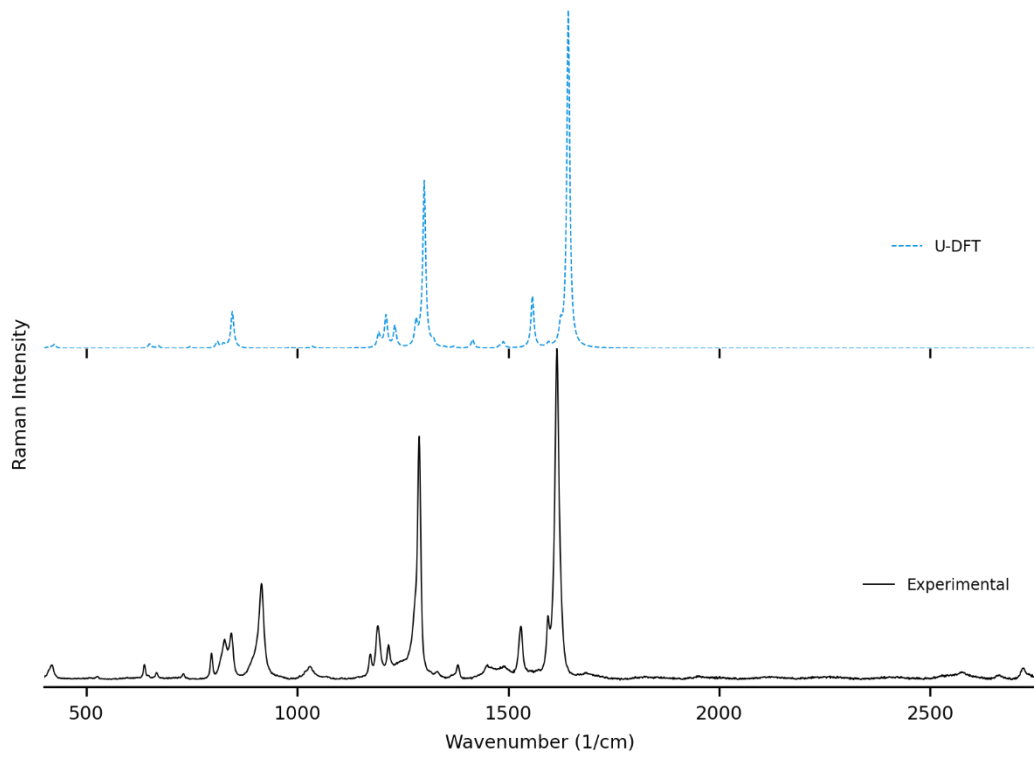
Échantillons en solution (THF), modélisés avec solvation SMD(THF), facteur 0.970











ANNEXE C

PARAMÈTRES D'ACQUISITION INSTRUMENTALE DES SPECTRES

Échantillons solides, mesurés sur une surface de quartz avec un objectif 50X

Identifiant	Nom	Laser (nm)	Intensité laser	Nb balayages	Temps balayages
A	BHT	532	1.0%	5x	120s
B	(+)-Camphre	785	1.0%	3x	100s
C	Coumarine	532	10.0%	10x	25s
D	Norcamphre	785	0.5%	2x	180s
E	8-Hydroxyquinoline	785	0.5%	4x	120s
F	Méthylparabène	785	5.0%	3x	100s
G	para-phénylphénol	785	5.0%	4x	100s
H	Xanthidrol	785	1.0%	3x	100s
I	Saccharine	785	0.5%	4x	100s
J	Nicotinate de méthyle	785	0.1%	4x	100s
K	Acide 4-bromophénylacétique	785	5.0%	4x	100s
L	2-Hydroxyquinoline	785	0.1%	4x	100s
M	Phénacétine	785	1.0%	2x	180s
N	Pyrixodine	785	5.0%	4x	100s
O	Adrénaline	785	1.0%	4x	100s
P	Acide 2,4,6-triméthylbenzoïque	785	0.5%	4x	100s
Q	Acide L-Aspartique	785	1.0%	4x	100s
R	Gramine	532	5.0%	6x	100s
S	Acide 1,3,5-benzènetricarboxylique	532	5.0%	6x	100s
T	Acide 1,2,4-benzènetricarboxylique	785	5.0%	4x	100s

Échantillons dans le THF, mesurés dans un tube RPE en quartz avec un objectif 5X

Rang	Identi- fiant	Nom	Laser (nm)	Intensité laser	Nb balayages	Temps balayages	Conc. (%m/m)
1	C	Coumarine	532	0.1%	8x	120s	14.8
2	G	para-phénylphénol	532	0.1%	8x	120s	23.0
3	B	(+)-Camphre	532	0.5%	2x	120s	48.1
4	E	8-Hydroxyquinoline	532	0.1%	8x	120s	38.7
5	U	4-(4-méthylphényl)phénol	532	0.1%	7x	120s	26.3
6	J	Nicotinate de méthyle	532	0.1%	5x	120s	47.2
7	A	BHT	785	0.5%	2x	120s	26.0
8	F	Méthylparabène	532	5.0%	2x	120s	3.04
9	M	Phénacétine	532	0.5%	6x	100s	7.15
10	V	N-Hydroxyphthalimide	532	1.0%	5x	120s	5.74

ANNEXE D
INFORMATION EN SUPPORT DE L'ARTICLE DU CHAPITRE 3

Supporting Information for:

Towards Routine Organic Structure Determination using Raman Microscopy

Jason Malenfant, Lucille Kuster, Yohann Gagné, Kouassi Signo, Maxime Denis, Sylvain Canesi*, Mathieu Frenette**a

Department of Chemistry, NanoQAM and Centre Québécois de Matériaux Fonctionnels (CQMF), Université du Québec à Montréal, Montreal, Quebec H3C 3P8, Canada

*frenette.mathieu@uqam.ca

Table of Contents

1.	Simple computational protocol for modelling the Raman spectrum of organic solids as dimers.....	109
2.	Molecular modelling strategy comparison.....	110
3.	DFT Data.....	111
3.1.	ORCA Input file example for pre-optimization.....	111
3.2.	ORCA Input file example for final frequency calculation.....	111
3.3.	Optimized atomic coordinates.....	112
4.	Raman spectra.....	166
5.	Materials.....	178
6.	Synthesis of Borneol/Isoborneol.....	178
7.	Measurement parameters.....	179
8.	SARA Compression function.....	179
9.	References.....	180

1. Simple computational protocol for modelling the Raman spectrum of organic solids as dimers

- 1- To model an organic molecule containing N_c carbons, we start by creating approximately $\frac{N_c}{4} + 1$ conformers (depending on rigidity) of a single molecule (monomer) in Avogadro and export each set of cartesian atomic coordinates to a .xyz file.
- 2- Then, each monomer geometry is pre-optimized with ORCA at the r²SCAN-3c level of theory in gas phase. An example input file can be used from section 3.1. The optimized monomers are ranked by their final single point energies (SPE), and the two most stable conformers are selected.
- 3- For each of the two pre-optimized monomers, we create, again with Avogadro, at least 8 dimer geometries with various relative positioning by means of rigid transformations. When applicable, we prioritize the hydrogen bonds between molecules followed by π -stacking placements.
- 4- The dimers are optimized with ORCA at the r²SCAN-3c level of theory in gas phase, with the TightOpt command.
- 5- The optimized dimers are again ranked by SPE, and the single most favorable geometry is used to calculate the Raman spectrum at the same level of theory still, with an input file using the commands shown in the example in section 3.2. If an imaginary frequency is found (in an ORCA output file, typically above all the vibration information, near the SPE), a new conformer must be generated. When a spectrum without imaginary frequencies is obtained, the spectral information can be obtained from the output file.
- 6- To extract the spectral data, the tool orca_mapspc is used with the following command:

```
orca_mapspc [OUTPUT FILE NAME] raman -cm -w04 -x00 -x14000 -n4000
```

And the generated .stk file is used for processing with the correlation software.

2. Molecular modelling strategy comparison

	1-DFT	2-DFT	3-DFT	4-DFT	5-DFT	6-DFT	7-DFT	8-DFT	9-DFT	10-DFT	11-DFT	12-DFT
1-Exp	94	49	27	41	56	69	43	33	50	44	55	65
2-Exp	40	79	34	23	53	56	43	22	34	24	31	66
3-Exp	13	16	83	16	48	20	18	20	14	6	13	29
4-Exp	47	25	32	66	44	24	34	31	44	47	40	38
5-Exp	36	32	50	38	62	40	28	35	33	29	34	43
6-Exp	53	58	48	28	59	75	31	20	28	41	38	61
7-Exp	52	52	21	38	49	56	59	46	63	51	65	66
8-Exp	30	25	13	44	30	21	26	54	56	40	34	25
9-Exp	30	24	40	22	42	34	18	41	80	37	38	39
10-Exp	44	35	36	52	48	43	44	29	54	72	54	53
11-Exp	49	27	27	68	48	46	52	55	66	68	73	48
12-Exp	46	53	38	28	51	45	23	13	26	24	23	71

Figure S1. Match scores for the twelve initial compounds, calculated as single molecules (“monomers”), with r^2 SCAN-3c in gas phase. When only one molecule is modelled, several false positives are obtained (in red). An average of 0.4 hours is required to generate the DFT spectra in the final frequency calculation on 16 cores. This figure is also shown in the main article as **Figure 5B**, but is reprinted here for ease of comparison.

	1-DFT	2-DFT	3-DFT	4-DFT	5-DFT	6-DFT	7-DFT	8-DFT	9-DFT	10-DFT	11-DFT	12-DFT
1-Exp	94	53	28	48	54	69	58	44	45	43	56	67
2-Exp	41	89	36	26	52	51	52	24	32	25	33	63
3-Exp	14	16	86	16	47	22	13	16	15	7	16	30
4-Exp	46	30	32	86	42	23	43	54	40	55	40	41
5-Exp	36	34	53	23	81	46	19	23	35	20	18	54
6-Exp	53	55	49	31	57	81	41	21	25	42	39	58
7-Exp	52	67	24	46	45	49	80	52	61	52	66	56
8-Exp	30	27	16	59	28	20	48	78	56	45	35	25
9-Exp	29	36	42	24	40	33	22	45	77	35	41	41
10-Exp	43	43	37	61	43	45	41	37	52	75	53	50
11-Exp	49	44	31	66	44	43	66	65	66	65	75	40
12-Exp	46	51	38	32	52	47	29	22	25	26	23	75

Figure S2. Match scores for the twelve initial compounds, calculated as pairs of molecules (“dimers”), with r^2 SCAN-3c in gas phase. A sizable improvement in spectral correlation can be observed. An average of 3.0 hours is required to generate the DFT spectra in the final frequency calculation on 16 cores. This figure is also shown in the main article as **Figure 5C**, but is reprinted here for ease of comparison.

	1-DFT	2-DFT	3-DFT	4-DFT	5-DFT	6-DFT	7-DFT	8-DFT	9-DFT	10-DFT	11-DFT	12-DFT
1-Exp	94	50	28	50	54	66	56	35	44	53	58	68
2-Exp	41	82	36	28	52	51	52	17	36	37	32	64
3-Exp	14	15	85	20	47	20	19	9	17	13	15	29
4-Exp	46	27	35	78	42	22	45	40	39	64	43	39
5-Exp	36	30	49	28	82	40	24	13	37	27	19	50
6-Exp	52	58	47	36	57	80	42	21	25	52	37	63
7-Exp	51	58	25	45	44	50	77	41	61	55	67	65
8-Exp	31	25	19	50	27	18	45	64	55	39	41	28
9-Exp	30	27	45	27	40	31	26	36	78	36	41	44
10-Exp	43	38	39	63	43	41	51	40	51	81	54	58
11-Exp	49	33	33	69	44	41	68	45	66	59	82	47
12-Exp	45	52	37	35	53	42	29	16	29	36	23	73

Figure S3. Match scores for the twelve initial compounds, calculated as four molecules (“tetramers”), with r^2 SCAN-3c in gas phase. Some correlation scores go down, which we hypothesize is because a complete conformational search for the position of 4 discrete molecules is harder to achieve, or the optimal conformer changes with the number of molecules modeled. An average of 14 hours is required to generate the DFT spectra in the final frequency calculation on 16 cores, which in the context of routine analysis we consider too high for molecules of this size.

3. DFT Data

3.1 ORCA Input file example for pre-optimization

```
# [MOLECULE IDENTIFIER]
!Opt r2SCAN-3c

%pal
  nprocs 16 # Number of cores
end

%maxcore 3000 # Memory allocation per core

*xyz 0 1 # Charge and multiplicity
  [CARTESIAN ATOMIC COORDINATES HERE]
*
```

3.2 ORCA Input file example for final frequency calculation

```
# [MOLECULE IDENTIFIER]
!Opt NumFreq r2SCAN-3c TIGHTOPT

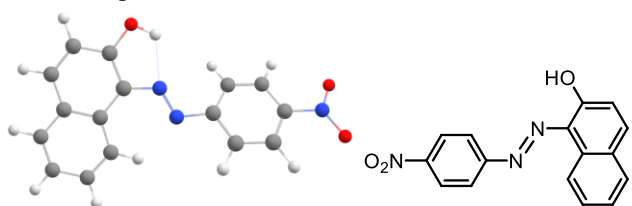
%elprop
  polar 1 # Calculate Raman spectrum
end

%pal
  nprocs 16 # Number of cores
end

%maxcore 3000 # Memory allocation per core

*xyz 0 1 # Charge and multiplicity
  [DIMER CARTESIAN ATOMIC COORDINATES HERE]
*
```

3.3 Optimized atomic coordinates



Main article molecule number: **PR1-enol**

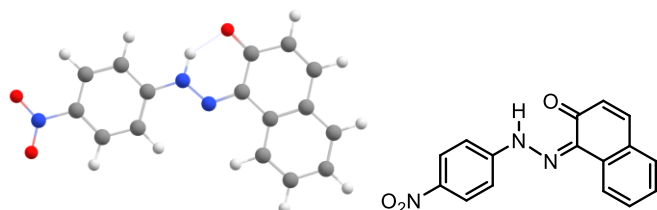
Name: PR1 (Pigment Red #1), enol form

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -2011.9500779559

C	3.40710546377955	-2.74019665988325	0.24975933141068
C	2.45373917678617	-1.74591465514957	0.13989257303273
C	2.83418603508477	-0.39118532590009	0.02363571147449
C	4.23376589228860	-0.08872767828272	0.03204070036504
C	5.17860675403714	-1.12934720557958	0.14266459455293
C	4.77659693244585	-2.44115044586260	0.24842743001514
H	3.08341201184807	-3.77332895814798	0.33874683495552
H	1.40480606901022	-2.00251721127269	0.13913946826118
C	1.91074720880214	0.71214304734943	-0.08636428655890
C	4.67229957495038	1.25912029986993	-0.06648823389195
H	6.23517432639257	-0.87339023597984	0.14526889617247
H	5.51077729768887	-3.23637985545891	0.33322312919138
C	3.79355383883327	2.30033365547036	-0.16516656127435
C	2.41274032129810	2.03215289282406	-0.17665835252368
H	5.74193429919738	1.45400031348476	-0.06082241898973
H	4.12421856842597	3.33068225283790	-0.24021265857454
O	1.58313048869499	3.07732754452983	-0.28668821891926
H	0.68185732972446	2.69163398841909	-0.28169064334013
N	0.53508552183785	0.71115378421091	-0.10142002600821
N	-0.10434158797317	-0.38258094315203	-0.13451842483643
C	-1.49993835936860	-0.23724919558637	-0.07080907323994
C	-2.24838311438980	-1.33517470620516	-0.51914745435471
C	-2.15764869961314	0.88821191171716	0.45444519867185
C	-3.63112767754637	-1.30232455846476	-0.48572967960133
H	-1.72020293274910	-2.20176053384143	-0.90433620940430
C	-3.53869550697343	0.92165672042448	0.50249520459345
H	-1.57595117461085	1.71619616366349	0.84373141292535
C	-4.25948630914698	-0.17095089013262	0.02503939359624
H	-4.22848096552034	-2.13343759082406	-0.84085728447605
H	-4.07067532060077	1.77294450742032	0.91092216318830
N	-5.72898263228212	-0.13237764887981	0.07613249686012
O	-6.33793450934108	-1.10868237876505	-0.34759121959730
O	-6.25548832101060	0.87478959514681	0.53687420632393



Main article molecule number: **PR1-keto**

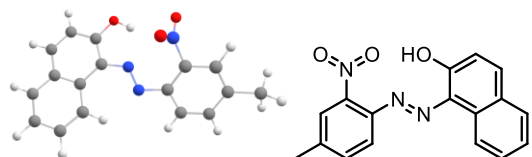
Name: PR1 (Pigment Red #1), keto form

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -2011.97185737517

C	-3.73803217608272	2.79644582182170	0.00032347321776
C	-2.70208442824400	1.87812996353103	0.00020631473691
C	-2.96707798080638	0.50079910108170	0.00001604084660
C	-4.31580519830430	0.06783671244378	-0.00005997829635
C	-5.35147324215531	1.01641472920444	0.00004756396960
C	-5.07003125470196	2.36964711762874	0.00024392380673
H	-3.51124080819599	3.85856570369603	0.00048656249223
H	-1.67231783842580	2.21702954936780	0.00028280862156
C	-1.90705773082755	-0.50069329060241	-0.00004814358735
C	-4.60919563535811	-1.33904153206707	-0.00021915955509
H	-6.38154627787650	0.66905879171839	-0.00000754113742
H	-5.87654340415677	3.09624883395094	0.00034285747513
C	-3.64422471092785	-2.28947392720325	-0.00026726844262
C	-2.24389508007566	-1.93761550186375	-0.00018413619433
H	-5.65678639179949	-1.63467344084587	-0.00027985483183
H	-3.87952984996040	-3.34924968299156	-0.00036425291398
O	-1.35306521316088	-2.82213975438781	-0.00011410818222
H	0.04574723450082	-1.91818707071323	0.00020793818606
N	-0.66229253939475	-0.04783741464092	-0.00000673040216
O	0.32023087742480	-0.91107874288996	0.00008827709932
C	1.64178355507320	-0.49160126539721	0.00006080254694
C	1.99641800595804	0.86541208232911	-0.00029193455791
C	2.63349732674631	-1.48446528248458	0.00037553499765
C	3.33207254972606	1.21891625075177	-0.00037423249412
H	1.22156441427697	1.62205507040455	-0.00054356278568
C	3.96845334858188	-1.12785246481931	0.00035384642771
H	2.34723158952584	-2.53250845716271	0.00065275635868
C	4.30542529002805	0.22240127786143	-0.00004353139527
H	3.63616927680979	2.25895190732266	-0.00067891013755
H	4.75177896220903	-1.87614667697618	0.00059997350355
N	5.72105017873793	0.60549161746403	-0.00014384048344
O	6.55403680421188	-0.29536504322336	0.00028289065899
O	5.98631534664387	1.80339701769107	-0.00064737954811



Main article molecule number: **PR3-enol**

Name: PR3 (Pigment Red #3), enol form

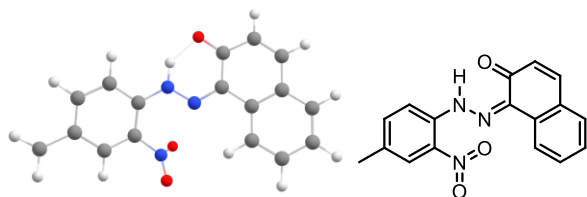
Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1045.254722988221

C	3.83350325769081	-1.98505362799704	-0.63570092924309
C	2.45087629640487	-1.93743441496700	-0.55925016930778
C	1.78574841271945	-0.77914031443440	-0.14226643944239
C	2.58599609985759	0.31382170092944	0.23202822453527
C	3.96608847422412	0.27744290040056	0.14627270169588
C	4.61601872315937	-0.87958414076050	-0.28900661953409
H	4.31798986061189	-2.89568169308996	-0.97801690906195
H	1.84318398687995	-2.78811392789774	-0.85240647302483
H	4.53646381270696	1.15418009004985	0.44174740241884
C	6.11615586863988	-0.93708395868272	-0.33837812761192
H	6.54242919678769	0.02927824414731	-0.62271238772094
H	6.52676311210315	-1.20051403116266	0.64412504349569
H	6.45973204610414	-1.69019800100610	-1.05238443955935
N	2.00115918583744	1.52865693475177	0.82399392018878

O	2.16300611547418	2.58170560613504	0.21426808513699
O	1.45892909728465	1.41079469847102	1.91031819304987
N	0.38633699105749	-0.81127886455383	-0.16127316069530
N	-0.14443462754687	0.33362599723614	-0.27640055984912
C	-1.51202438094917	0.47812651306116	-0.26607940324768
C	-2.53626849360701	-0.51068623950812	-0.02634817264347
C	-1.88599731607449	1.81772993716998	-0.52319345797481
C	-3.90214180508167	-0.08650404078573	-0.08440454415421
C	-2.28320056949519	-1.86705072180694	0.27556855346004
C	-3.23971842236293	2.20498956384680	-0.57328906403872
C	-4.21290773201424	1.27167304158910	-0.36268840860423
C	-4.93906313167526	-1.01420126860226	0.14348799470425
C	-3.32371522849189	-2.74830421329718	0.49834391881261
H	-1.26238089614053	-2.21445995620786	0.33111942212321
H	-3.47113418960621	3.24448191495299	-0.77955453299246
H	-5.25957278974974	1.56356127370292	-0.40122770277314
C	-4.66021439041693	-2.33091375720202	0.42946559268779
H	-5.96767382324791	-0.66580685129226	0.09243382272784
H	-3.09614140515152	-3.78452986181620	0.73203197302411
H	-5.46494896248062	-3.03833735048989	0.60480954165007
O	-0.98232393683036	2.77845712294952	-0.74233063194500
H	-0.09563843662109	2.37594469616683	-0.66187025628678



Main article molecule number: **PR3-keto trans**

Name: PR3 (Pigment Red #3), keto form, trans isomer

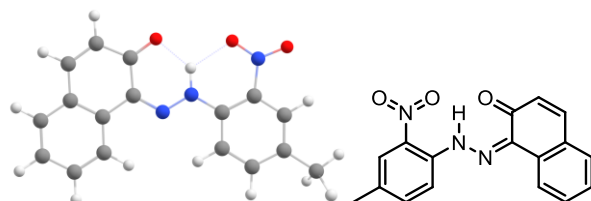
Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1045.261171350024

C	4.33900758045149	1.35645583538469	-0.48254878513140
C	3.01107325249910	1.71315065568553	-0.32080733350558
C	2.04119080029894	0.75257769245174	-0.01125763477684
C	2.46981260554732	-0.57194388834995	0.14474286404788
C	3.79195087641887	-0.93387570984655	-0.06081521450732
C	4.75449027902462	0.02556111369653	-0.36979507921816
H	5.06704134870723	2.12573860092076	-0.72579500328095
H	2.69778219386525	2.74390864035771	-0.46248946515902
H	4.06665319711176	-1.97902577320043	0.05105397491924
C	6.19733769237674	-0.35813842164589	-0.53794707199460
H	6.29684266686098	-1.38437148647312	-0.90170461249846
H	6.70342481023207	0.30762006722854	-1.24267524084853
H	6.73120831056124	-0.29299965905760	0.41806729749277
N	1.57160985812879	-1.63452138546651	0.63095604114301
O	1.47271920754376	-2.64062670392027	-0.05822394671940
O	1.04662007416167	-1.45797294688414	1.71888171759069
N	0.71781188541349	1.16321405357968	0.06598363023163
N	-0.25420820908415	0.30254789695397	-0.07164861725648
C	-1.50475127049524	0.73531382153840	-0.01245741327206
C	-2.54885229615939	-0.27055150958097	-0.15698951613713
C	-1.85337209038778	2.15480855617388	0.15820561568819
C	-3.90210626753603	0.14828733835061	-0.15016083624555
C	-2.26374000944177	-1.63783318160445	-0.29067106720453
C	-3.25612957686340	2.49404732660443	0.15376720505921
C	-4.21000189720562	1.54240480026044	0.00625641617452
C	-4.92587992708381	-0.80413075183279	-0.28719636564683
C	-3.28887392014655	-2.55815164505181	-0.42421453746060
H	-1.23177638198006	-1.97138464019052	-0.28104922524743

H	-3.50465460188056	3.54401893718803	0.27406063648838
H	-5.26037721175954	1.82859319667795	0.00748677074509
C	-4.62621778576094	-2.14599324246837	-0.42528489760995
H	-5.95997122924949	-0.46855427598355	-0.28017192375976
H	-3.04762417441399	-3.61228881434270	-0.52470185571960
H	-5.42271542701685	-2.87644624688646	-0.52914197593788
O	-0.96898328621856	3.04019387018229	0.28544644597371
H	0.43493192348037	2.16389187955087	0.17928600358374



Main article molecule number: **PR3-keto cis**

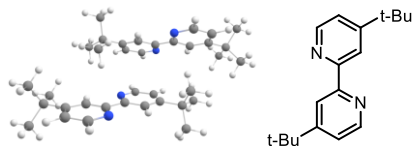
Name: PR3 (Pigment Red #3), keto form, cis isomer

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1045.266450183835

C	3.01991412208171	2.34467060230508	-0.00009811994353
C	1.83213343885746	1.64744792572658	-0.00008171470396
C	1.80807243969198	0.23970745521495	0.00000755375635
C	3.05935668469680	-0.41914667523495	0.00009149100297
C	4.25307001971423	0.30618306126110	0.00007883404872
C	4.26158291960598	1.68980956692127	-0.00001862723550
H	2.98961759203345	3.43196074922949	-0.00017961248501
H	0.88612952286026	2.17313026116152	-0.00014845957119
H	5.17980093617107	-0.25564324369968	0.00014728874195
C	5.54469262951433	2.47030865973733	-0.00001698616720
H	6.41332351635347	1.80786741714484	-0.00022058830518
H	5.61010930632036	3.11728878077405	-0.88206561962011
H	5.61029453635372	3.11697032365326	0.88225426566862
N	3.19000270105766	-1.87818919789268	0.00020145823534
O	4.31798654657478	-2.35666094004175	0.00042329455433
O	2.16051991196726	-2.55559315149844	0.00008071431580
N	0.59311963444391	-0.43117134547303	0.00001456056958
N	-0.52072750222078	0.26373848857767	0.00002797510022
C	-1.68459902940351	-0.35798572713481	0.00002094208747
C	-2.87370569998616	0.49161068317238	0.00003615385967
C	-1.82806590471647	-1.83157339936438	-0.00000802442298
C	-4.15172324574584	-0.11894169704452	0.00003482679803
C	-2.79909682940545	1.89202747869591	0.00005163489910
C	-3.17509040082200	-2.36604432204277	0.00001013245438
C	-4.25645968524830	-1.55469779079431	0.00002511634825
C	-5.30508463096369	0.68100050139109	0.00004745387104
C	-3.94899877083216	2.66338051110569	0.00006331896282
H	-1.82448731711376	2.36675062738588	0.00005579915663
H	-3.26479199299759	-3.44800253872164	0.00000663331033
H	-5.25647390172854	-1.98484165263206	0.00003241970644
C	-5.21051634806085	2.06061954788891	0.00006089545758
H	-6.27879836923448	0.19758095806551	0.00004740707674
H	-3.86720024929273	3.74644239154377	0.00007504116156
H	-6.10850200503577	2.67073264509247	0.00007062326962
O	-0.83368139076971	-2.58209518682397	0.00000204995868
H	0.54571381527935	-1.46553676764977	-0.00000213191755



Main article molecule number: 1

Name: 4,4'-di-t-butyl-2,2'-bipyridine

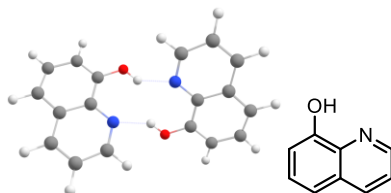
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1619.3857503721

C	3.92732784162321	9.54929003189110	33.20033250623817
H	3.23886478206480	10.06666608382867	32.52512130362345
H	3.35477505278079	9.17765217834772	34.05860011638485
H	4.33836760226128	8.68622535127841	32.66757412165111
H	3.66067391906175	12.18732150850289	33.86519669525948
N	1.66221558967236	10.88399472457393	30.30723592887676
N	3.25826407758058	14.11667704246955	30.44088013784869
C	0.787698806661516	10.16644801955850	31.01414061764351
H	0.51686280001496	9.19568388192338	30.59868027179347
C	0.23028476703497	10.58797476169592	32.21268291897461
H	-0.47724069015560	9.94163619345103	32.72313302403614
C	0.58972504748557	11.83443735260529	32.73862679798675
C	1.49962168308924	12.58436714969409	32.00010303642293
H	1.83128277635516	13.56392906265622	32.31982152805293
C	2.01089741836792	12.08229302494655	30.79676474228285
C	2.97236807396338	12.88131364925247	29.99378579130219
C	3.52992932016148	12.35446057877475	28.82945644490439
H	3.24524390348592	11.35012362149213	28.53977115173543
C	4.42722054943420	13.10921315157296	28.07386532356833
C	4.70747074394360	14.39417846893961	28.54095283394506
H	5.39430133158687	15.05220062014749	28.02052506101196
C	4.10350369842244	14.84165324854886	29.71274093698780
H	4.32146587006320	15.84344081574485	30.08515501552749
C	-0.01679979480530	12.32123348548710	34.05302904816630
C	-1.53953960507309	12.46774104036031	33.86966429961451
H	-2.00529693160297	11.51702305930308	33.59238020259446
H	-1.76654640339928	13.19721693555262	33.08496398616281
H	-1.999931568760713	12.81077730576584	34.80351952628487
C	0.27027501611497	11.28790035305850	35.15861223674412
H	-0.148118318888520	11.63275351878705	36.11089258533789
H	1.34875841854940	11.14836216350651	35.28642123095721
H	-0.17278644911449	10.31366516142448	34.93173612116314
C	0.55575831853399	13.67341481625121	34.49192834887318
H	0.09538282092344	13.97166677846626	35.43950092731398
H	0.34899629334410	14.46010231286341	33.75894872859575
H	1.63882745485557	13.62194040576363	34.64806361146146
C	5.02163972039393	12.53759167722860	26.78882381054548
C	3.90217091364209	12.46472991926320	25.73226333719658
H	4.29468771452005	12.04867720840374	24.79714967713632
H	3.49474446960929	13.46013047390505	25.52388810912371
H	3.08056277108655	11.82679379307464	26.07313217824501
C	6.16458755787136	13.40629279784976	26.25055744603586
H	5.82255866913780	14.40691717286916	25.96523442849474
H	6.58714559883312	12.93897106302160	25.35455834379075
N	7.90024300123907	8.32220191348973	29.99161919608978
N	7.29340384663029	11.87284822467585	30.22435013451179
C	8.55053166932933	7.47040333900539	29.19667378909623
H	8.47510959182810	6.41546068516544	29.46089934073406
C	9.28516459208564	7.85499802364125	28.08412688501257
H	9.77413649335920	7.09400439637318	27.48314828247010
C	9.37910944567460	9.21333201772829	27.75818411263845
C	8.71911182946275	10.10234407363104	28.59840894942218
H	8.74081101009796	11.17172081239259	28.43439492839077
C	7.98356344573418	9.62540126167030	29.69016131266196
C	7.23713032276790	10.57115519850042	30.55961972301571
C	6.51595941608235	10.11063221390583	31.65979907019074
H	6.52167799228037	9.04579206285336	31.86210583340066

C	5.81459090246707	11.00825719379011	32.46673409651690
C	5.86997488882811	12.35314449955816	32.10179456755662
H	5.33823511723810	13.11929329070660	32.65416360934975
C	6.61791909083039	12.72597158488989	30.98872954288132
H	6.66597700601248	13.77467230172975	30.69831494990771
C	10.16394436862274	9.66511539273860	26.52894461836866
C	11.63308646140573	9.22805574863086	26.67830903899846
H	11.72499028175592	8.14066691538984	26.76057107246909
H	12.08157761781574	9.67392989406880	27.57241891433226
H	12.21200448610129	9.54991560203931	25.80512607213641
C	9.55486223960484	9.00470028792325	25.27725297042235
H	10.10014121228538	9.32430075363432	24.38192762050193
H	8.5039865776576	9.29156911676078	25.16167702046500
H	9.60311170705988	7.91290154468486	25.32900962477180
C	10.11899342739919	11.18562689289432	26.34227628064096
H	10.68529310653122	11.46010083069045	25.44613139729599
H	10.56275059860532	11.71310521832080	27.19318606812434
H	9.09165240172280	11.54412114374681	26.21223511624424
C	5.03919739621977	10.50116424098170	33.67990557968609
C	6.00097504019648	9.74402517297979	34.61492217178948
H	5.45599003479090	9.38012058236544	35.49340525707478
H	6.80794316257197	10.40054975141166	34.95820733535640
H	6.45505321851941	8.88105043427545	34.11931870782682
C	4.39554892624607	11.64543576777819	34.47038340962619
H	5.14366678594258	12.357721666463205	34.83549697453594
H	3.87438734788532	11.23530483316456	35.34193916617919
C	5.57600693212437	11.12671048285094	27.05120428049451
H	6.32807887199777	11.15408471301702	27.84384716315487
H	6.04516040866783	10.73678640539101	26.14035365938110
H	4.79439211467618	10.42167575902082	27.34805175703604
H	6.96774691949400	13.50655939279380	26.99009390932919



Main article molecule number: 2

Name: 8-hydroxyquinoline

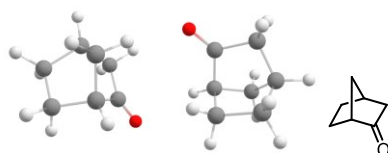
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -954.140341993345

N	0.65079015041470	2.14155279026096	0.00933918057436
N	-0.91382388426699	5.14247264676031	-0.06529371049949
C	0.23985400361215	1.00248220887415	-0.51502141308126
C	-0.49856873044556	6.25722333963756	-0.63650285467253
H	-0.83750493653250	0.88108169393396	-0.62261388647117
H	0.57968750519146	6.37430117648051	-0.73994588078942
C	1.11505617117783	-0.02808296811545	-0.90946848499761
C	-1.37033479246985	7.26860590507052	-1.08441764479521
H	0.71366630784736	-0.94007550835690	-1.33951419965126
H	-0.96539586437794	8.16064313711216	-1.55140023479644
C	2.46335215441996	0.14187229662392	-0.72352435589821
C	-2.72023990997119	7.10633821930022	-0.90303257790373
H	3.16958180088935	-0.63696406708152	-1.00120971027284
H	-3.42409470555948	7.87157541968830	-1.22149068762408
C	2.95113088345761	1.34143377558968	-0.15883627444774
C	-3.21275854598142	5.93288981449689	-0.28963960126299
C	4.32262992360076	1.57725750979035	0.08025937213655
C	-4.58631014329348	5.70712325913625	-0.05267723183947
H	5.04868728219456	0.81223974026076	-0.17877054388315

H	-5.31020831918346	6.45929877581343	-0.35212665435626
C	4.71428012773914	2.77090181319444	0.64012273503345
C	-4.98264542506188	4.53930196841603	0.55626976708543
H	5.76674990689308	2.95882887182068	0.83248907306520
H	-6.03675816713264	4.35932346823851	0.74725189243955
C	3.77971150816855	3.76870856646836	0.96208647546734
C	-4.05075691200732	3.55747830364732	0.93087505620887
H	4.10102703878726	4.71261188384605	1.39094862624384
H	-4.37600559394663	2.63353076765527	1.39849319890484
C	2.42939534858011	3.57747770291209	0.73340041131402
C	-2.69842650541020	3.73921985675801	0.70631409220629
C	1.98436093423721	2.33749064484516	0.18235547211837
C	-2.24879959945046	4.95368687135408	0.10441861056113
O	1.57505401220291	4.56413916347064	1.04680708868541
O	-1.84680464138685	2.76845917201198	1.07183201977354
H	0.66883403906593	4.40928931160490	0.68261034920390
H	-0.93722242200208	2.90529246848028	0.70837252622076



Main article molecule number: **3**

Name: Norcamphor

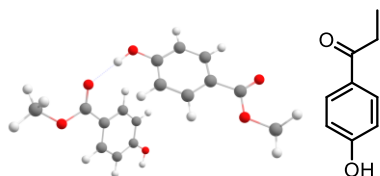
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -695.817074433485

C	0.67059079884431	7.76175366788943	0.21176661755461
C	-0.55968131128706	7.59761221878576	1.12133175050621
C	0.90012520504285	6.03964583300230	1.91528593423339
H	0.48017825773172	7.65579475807095	-0.86156766292281
H	1.16453654948527	8.73010652853500	0.36900724224012
C	0.11219059237563	7.24383519135258	2.46583116986971
H	-0.60508066192158	6.95369815697963	3.24086342820773
H	0.75543690926973	8.04348432280138	2.85083223531731
C	-1.24915550854154	6.26567553480214	0.76000986363676
H	-2.21484950363540	6.18017592678499	1.26868689253002
H	-1.43504551009863	6.18133056305818	-0.31581914323657
C	-0.24421809782257	5.19436931859655	1.28012041837435
H	-0.68207594177421	4.55825939801708	2.05455211240283
H	0.12687992588666	4.53249223267731	0.49153229888774
H	-1.22720166890098	8.46339150489338	1.13369488108591
H	1.55194664922275	5.47615565208124	2.58315893243579
C	1.60041087363834	6.65560984441803	0.71999665105638
O	2.65671513707297	6.32310227711151	0.22549897124595
C	4.02962094256689	3.30745845170722	1.24085307738477
C	4.39371444141828	2.76280533049345	-0.15152610077314
C	2.22601462724446	2.13119351949177	0.11015542055905
H	4.67507263455557	2.97568816257922	2.06117646128185
H	4.00032745164621	4.40548001526660	1.23842980793936
C	3.05519000228465	2.94980192791073	-0.89893609832560
H	3.05675593209710	2.50351730397258	-1.89969226075893
H	2.75226229937148	4.00010985424926	-0.96199553105604
C	4.49302032095997	1.22590895614829	-0.06162541536132
H	4.90979081716116	0.80816275263914	-0.98398052653984
H	5.13533614525899	0.90568298818093	0.76533789105770
C	3.00972811678317	0.78654809309990	0.13171650895933
H	2.66658021708217	0.15923282376965	-0.69746382321860
H	2.84319246267696	0.22950557733017	1.05899991670237
H	5.26133911912101	3.24678808969105	-0.60686594443106
H	1.15008588480489	2.02912805249208	-0.04390484575005
C	2.59848045590145	2.80582572814768	1.42099754055485

O 1.91269543447707 2.90663944297211 2.41528132834975



Main article molecule number: 4

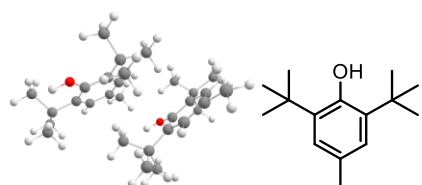
Name: Methylparaben

Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1070.537388950418

O	1.84011849141661	1.43325747605527	-1.18070123063178
C	2.81481306307643	-0.55911764425437	-2.08841651479450
C	3.14023754499612	-1.10696134185689	-0.83906479800833
H	2.97608054117070	-0.52560583290643	0.06222127003164
O	4.42452749743008	-4.36356735567636	-1.74987562125362
H	4.52698937143332	-4.78022522767449	-2.61389942791143
C	3.67854103338094	-2.37505388221170	-0.74159738120714
H	3.93151912885327	-2.80588191734635	0.22130742178443
C	3.89790385574062	-3.12130105428406	-1.90207829571823
C	3.58204386430482	-2.58768945024902	-3.15528985786515
H	3.75437227191117	-3.17190985303228	-4.05660132427918
C	3.04690265746999	-1.31430506446831	-3.24474426830727
H	2.79992858193524	-0.90182985104808	-4.21636643098912
C	2.22083622545037	0.78498519051726	-2.14749756640098
O	2.12144125553362	1.26080789555649	-3.40210747864238
C	1.52024457273756	2.56760180042719	-3.52373120943217
H	2.10969063992522	3.30945197560337	-2.97866579158915
H	1.51796312826759	2.78464498209233	-4.59186206063274
H	0.50118484813838	2.55653643624905	-3.12782999479039
H	1.43091295260771	1.05276687298956	0.54521347621047
O	-2.18378547702614	-4.12973213718691	3.38745084552851
C	-1.13280350567082	-2.47837240085718	2.01892988488071
O	-2.54382312772500	-4.15681935119150	1.15691838084187
C	-0.49536102028164	-1.86046880822046	3.10297554652448
H	-0.65294824774053	-2.26748769032066	4.09691111641474
O	1.31920865352240	0.82383222096837	1.49090801969915
C	0.31746397138300	-0.75962074643555	2.91346136742569
H	0.81656223263283	-0.27566916366452	3.74695684120387
C	0.50826645023616	-0.24794612107355	1.62383824558640
C	-0.12993737626411	-0.85366064419272	0.53330939738897
H	0.01339326965723	-0.45055776197518	-0.46520067664770
C	-0.93943652167160	-1.95991968319357	0.73363002700064
H	-1.42999635043072	-2.42694194138087	-0.11316322073686
C	-1.98566413772891	-3.65089329832304	2.28872262656921
C	-3.39032096651093	-5.30091241442103	1.36531732410119
H	-3.75580228652965	-5.57801704538095	0.37560259426339
H	-4.22400483729684	-5.04762530293783	2.02684667884477
H	-2.82321224833464	-6.12339186469486	1.81113208553804



Main article molecule number: 5

Name: BHT

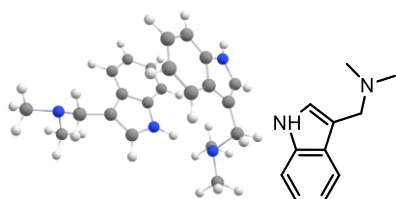
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1322.23997066399

C	-1.46032918050348	1.37493248902218	0.51712037350362
C	-2.35701639261029	0.42831816321459	1.33186730260362
C	-2.923594431483303	-0.69968636398341	0.44412429095265
H	-3.53839052239674	-0.31756981607467	-0.37113645831931
H	-3.54149666159723	-1.36807486785974	1.05590521072291
H	-2.10650386367196	-1.29684019481158	0.02550451513037
C	-3.50784762774561	1.22543317106351	1.97986530983250
H	-4.15429618715078	1.68928915711588	1.23322339180270
H	-3.10722534334127	2.01241938677036	2.62858168378347
H	-4.11831372761655	0.55439864538652	2.59558304047750
C	-1.57487328040454	-0.25588214837473	2.46334502086887
H	-2.24962741635714	-0.93222475918021	2.99901804660901
H	-1.18788941834182	0.46557736683900	3.19170406129916
H	-0.74253387981688	-0.85579019255112	2.07946746408340
C	-1.94572339658167	2.07843587955405	-0.60587371517953
O	-3.26574348723942	1.86972026050584	-0.93115910506778
H	-3.49916113739779	2.43720599160217	-1.67060694755342
C	-1.13139381784841	2.94445738009441	-1.36366242579171
C	-1.65392237879565	3.68173626687853	-2.60991742895566
C	-2.10383498912737	2.67471118184470	-3.69430404165472
H	-2.89289580679236	1.98383729522251	-3.38136366899107
H	-1.25240539434943	2.05686936830061	-3.99735812298564
H	-2.46842028476590	3.21348776444433	-4.57629351927844
C	-2.79395494393365	4.66160171455304	-2.24227064357664
H	-2.43687855129489	5.38459481530439	-1.50201051773052
H	-3.68981544253705	4.19784885927581	-1.81530619732076
H	-3.11476177662074	5.21092597687792	-3.13464196141423
C	-0.56140685136516	4.54263210627155	-3.26546335818807
H	-0.20050177413371	5.32831745435177	-2.59407585550659
H	-0.97852350021005	5.03110766555463	-4.15230536445335
H	0.29237853571774	3.93925064001246	-3.59009073505184
C	0.18749072209174	3.11802562209917	-0.93482180501531
H	0.84390154288359	3.78991381687430	-1.47609270945601
C	0.70269820395869	2.46463942518583	0.17786886611812
C	2.12751390224489	2.68916130123375	0.61143401695047
H	2.20215682778612	2.76610608149428	1.70095238691928
H	2.53134030149512	3.60917019859793	0.17841386613124
H	2.78196044791984	1.86286414793702	0.30693006229814
C	-0.13084922486469	1.59257506378262	0.87590778550535
H	0.28107729956183	1.06970693252936	1.73114495175361
C	1.63834186701143	-2.37329233613045	1.67863942912801
C	2.71058901890034	-1.89618220882620	2.67043591096426
C	2.52036304159470	-0.40344503893141	3.00303226822220
H	2.61888816859331	0.22169033070767	2.11581408292078
H	3.27472139875739	-0.08758362423461	3.73361642445469
H	1.53126602190022	-0.23376955334333	3.44408604702939
C	4.11745194376089	-2.13495098383150	2.08401278967821
H	4.25730814395584	-3.19746622493520	1.85599041573620
H	4.87813445556976	-1.84124284581595	2.81702386830416
H	4.28211622355826	-1.56210675066195	1.17074881831654
C	2.64122617506832	-2.66455011075237	3.99907833559119
H	3.43558420825281	-2.29995540012892	4.65881984026945
H	2.79472577803687	-3.74028816174566	3.86054608872470
H	1.68756315220471	-2.50965536490804	4.51565418099596
C	1.53859038582006	-1.84681536341115	0.37359785011183
O	2.41017541123448	-0.83286272450259	0.04681324280938
H	2.06859364363139	-0.38322988355568	-0.73397354165609
C	0.61846217343268	-2.34576879754830	-0.57255583321511
C	0.62249096600238	-1.88920295878472	-2.04502454980634
C	0.25018951124997	-0.39791739482120	-2.20428039650304
H	0.92220206403304	0.32626459613547	-1.72693058366767
H	-0.74105587869966	-0.20044150526029	-1.78627565419259
H	0.23463549500200	-0.14014080309130	-3.26968779149586
C	2.00004273539472	-2.18348569723776	-2.68062813220305
H	2.82876153283528	-1.67438957806737	-2.18269723222628

H	1.99617081482977	-1.87600373244182	-3.73280456957604
H	2.20698228223981	-3.25786963844916	-2.63821385637345
C	-0.41380099475106	-2.66113681741021	-2.87879319720629
H	-0.22606229166216	-3.73940567831784	-2.86436060510899
H	-0.35120555767154	-2.32662201701657	-3.91962436568825
H	-1.43660868216307	-2.47594305501492	-2.53352057397283
C	-0.26909090462292	-3.33257825649679	-0.13635631973383
H	-1.00752854074269	-3.73222589560249	-0.82218279243219
C	-0.24771439992353	-3.83728843695658	1.15811672531762
C	-1.26353024092298	-4.85208257967017	1.60867225445867
H	-0.83841602197821	-5.54622980941816	2.34040114307334
H	-1.64434807638741	-5.43708744752738	0.76597456369991
H	-2.12385570370498	-4.36293920142756	2.08326788738688
C	0.72620828837576	-3.36683076283999	2.03341520214989
H	0.75753126116802	-3.78908771149115	3.03153887442543



Main article molecule number: 6

Name: Gramine

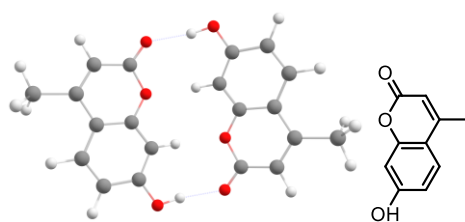
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1073.92746719103

H	13.04757828433263	24.44665970895246	12.72191219619079
H	14.11782145749662	26.12592564346862	10.71298729657407
H	14.17235202837920	25.71390244271148	16.74412377870799
H	14.87354753037177	26.62788062988737	18.09320453725962
N	13.94118478386574	24.63126479707890	13.14288790456415
N	16.27182189859538	25.68192168494286	16.83545043609470
C	14.33959305683936	24.19520405337746	14.38893090939262
C	15.58554469305796	24.69518597001975	14.67628027919826
C	17.11927137383335	26.24587827935832	13.24618565759556
C	17.15669468105306	26.94981359074009	12.05538475113147
C	16.08179538471734	26.90530365099699	11.14990623709319
C	14.95264541915674	26.14513627669567	11.40742916208035
C	14.92177798364654	25.43832543104135	12.60671717208425
C	15.98203237857461	25.48514611688424	13.54423584309121
C	16.37413251051169	24.51219706701642	15.92868471384551
C	17.34310427956223	25.65997245300085	17.83240927453649
C	14.96650415832547	25.71366391576337	17.49677942860826
H	13.70436038542643	23.54706824582973	14.97873673466268
H	17.94484613648965	26.30909229064219	13.95009436476253
H	18.02073144388463	27.56458412291483	11.82257502513978
H	16.13264635457255	27.48770406895347	10.23510307632499
H	17.43446007407753	24.39258361764607	15.67111219845809
H	16.06481106605488	23.58705823587121	16.45311262874082
H	14.81389581134586	24.84212664922105	18.16221522862821
H	17.26366395270151	26.54113280547870	18.47832973793021
H	18.31749358778276	25.68734173501541	17.33254108829709
H	17.30872885548916	24.75990651834343	18.47580517154903
H	16.08530152993795	27.43152476118932	15.84378960192137
H	13.40622533766367	27.06348272739402	14.99379984104313
H	15.83617683246013	32.12901298181630	15.32479693400909
H	15.95324336921017	33.82775554229995	14.81781679667184
N	15.94636098137521	28.30754890948309	15.32911161846013
N	16.73384970688533	32.44097312186991	13.45278595821554
C	16.96186693882186	29.18950629332214	15.02656868605802
C	16.57447246487224	30.04570970795682	14.02722056465342

C	14.31903223586372	30.13636726943290	12.72062067519740
C	13.09452863247878	29.49650303283189	12.60606294686531
C	12.76739613563774	28.39976283768276	13.42297141127007
C	13.65628825588041	27.91920747137173	14.37297514205563
C	14.88883848126843	28.55905476518927	14.48505685063100
C	15.24125250784031	29.66232304734370	13.66227159885843
C	17.39481296838333	31.12908606627705	13.41437558955358
C	17.54318182412780	33.43171590152757	12.75535871625148
C	16.47392243117136	32.86392995709773	14.82451838168105
H	17.90621188553528	29.14391710935429	15.55473621259908
H	14.56524357615770	30.99417813530045	12.10191378851971
H	12.36881707244236	29.85122128812066	11.87959081604543
H	11.79751281506822	27.92234682564852	13.31122331881759
H	17.58200901421282	30.89592482682608	12.35480812497736
H	18.38803361094766	31.15895967907968	13.91501622431026
H	17.40410815031526	32.97758589127604	15.42010466623900
H	17.02285584194614	34.39589651991339	12.74298431934440
H	17.70473565040327	33.11875603154991	11.71728071888154
H	18.53727217894446	33.58600929698616	13.22506566432404



Main article molecule number: 7

Name: Hymecromone

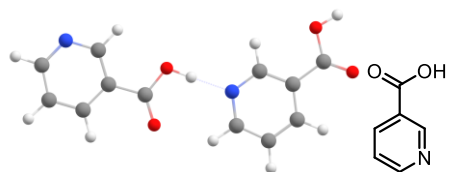
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1222.96523278743

O	-4.37430450465375	1.01608073030708	5.14835164077014
O	-3.09933360318085	0.69442563631339	6.93351621229102
O	-6.54273162797585	1.25616628804854	0.96379634533495
H	-5.65786482343705	1.57637249198308	0.68843242486251
C	-4.19561166316607	0.51262596228369	6.42437302718150
C	-5.32061186019267	-0.15905161196031	7.00677787170265
H	-5.19056673269285	-0.50276187329101	8.02687503987233
C	-6.47662387184019	-0.38431628240616	6.31471207166313
C	-6.57945956244029	0.05997973015200	4.95241878893605
C	-7.67804849215841	-0.16572742598165	4.09621519356632
H	-8.54819662199110	-0.69293334881926	4.47466461472914
C	-7.66505493751523	0.24602685131665	2.78064499487835
H	-8.50275799088449	0.04696420496985	2.12082987506864
C	-6.53830062934308	0.91752653219358	2.26412093751212
C	-5.46127739910663	1.20436055998523	3.09768235331179
H	-4.59272478598689	1.73174865967788	2.72564533828704
C	-5.48997897663009	0.75615252535484	4.40460712287227
C	-7.62099242204469	-1.10725680940372	6.95494702536887
H	-7.85224769843465	-2.02869612745632	6.40822683391987
H	-8.52546033781446	-0.48850980383313	6.94293180044973
H	-7.39228071801064	-1.36763694742501	7.98993412633128
H	-2.43926960826875	1.59314229062389	4.26267251113515
H	-1.32411209395111	0.89452997578162	6.14836399735342
O	-2.61783181310057	1.71558083906175	1.73735054430718
O	-3.88043318743935	1.86216701594112	-0.07896355796235
C	-2.75984555086976	1.65495807607290	0.36272101868488
O	-0.41091800824267	0.79582244549305	5.80467805334944
C	-1.57305845635266	1.36218266120485	-0.38712249226269
H	-1.68094310998196	1.36865539035428	-1.46597602127240
C	-0.38378152796799	1.06069841865571	0.21369271498735
C	-0.30766170853728	1.02921417454708	1.64795333103249
C	0.82792704043134	0.66544062795097	2.40229872126205
H	1.74873891078988	0.40838761942094	1.88805984366756

C	0.79155254024551	0.60465867207338	3.77919482066862
H	1.66105650654601	0.29965355068902	4.35162600690988
C	-0.39859899904386	0.91803245709730	4.46653493227894
C	-1.51986805782473	1.33367678940380	3.75443428940445
C	-1.46315563727842	1.35996061490907	2.37407095525587
C	0.82678565499766	0.74048288434322	-0.60733668521754
H	1.17924012271219	-0.27591926485636	-0.39761089817065
H	1.64918145790500	1.42398336534234	-0.36763678775840
H	0.61040478473208	0.81818145388082	-1.67431893656339



Main article molecule number: **8**

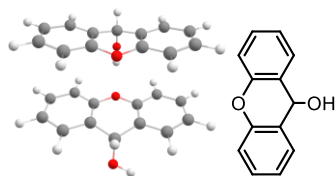
Name: Nicotinic Acid

Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -873.592372600774

O	-6.40408672599286	4.78979943938807	-3.52409961964615
C	-1.94603516538560	4.93935672793196	-2.81064106349096
C	-10.55975842706006	4.29349090756707	-4.01326565990257
C	-2.70121798627381	2.38675335386233	-2.19762002989032
C	-1.355680735252728	2.73643216602791	-2.08061037814413
C	-0.97588326894370	4.04224895977972	-2.39469073485245
C	-3.26537663151991	4.50763978864856	-2.90004082454224
C	-0.31534050276336	1.77712225512195	-1.63512863454443
N	-3.63306225439650	3.25488307851278	-2.59811231626288
O	-0.82138101408699	0.54813162275103	-1.37075072813919
O	0.85752867176543	2.04944551639860	-1.51729931888010
C	-8.99006963161404	2.25689392515312	-3.11513543785917
C	-8.39008806968889	3.48285536581119	-3.41762554284825
C	-9.19965785081520	4.52116452861284	-3.87690594944321
C	-11.05966538474811	3.03566633600597	-3.68456010720003
C	-6.92706478165247	3.71695987630341	-3.26587240614441
N	-10.29960951901221	2.02804697002320	-3.24268971797831
O	-6.26473335602791	2.64957880281236	-2.82531960824810
H	-0.07472054442874	-0.00359634171931	-1.08939323759222
H	-5.27994034673856	2.88381138894036	-2.74565543379950
H	-3.03788663463111	1.38021867739383	-1.96203196360244
H	0.06911005761403	4.32386650865486	-2.30596885827758
H	-1.69191399560703	5.96307349355706	-3.06465411691422
H	-4.07239670757573	5.16401272233134	-3.22110558163542
H	-8.38210249819262	1.43007253867767	-2.75495088255193
H	-12.12477072699487	2.82573796548659	-3.78077964191209
H	-11.22725429630601	5.07309385288435	-4.36744935150555
H	-8.74659167366554	5.47872957308096	-4.11534285419201



Main article molecule number: **9**

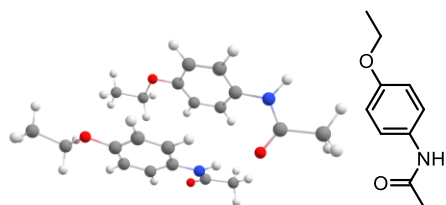
Name: Xanthinol

Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1303.46741017142

C	-3.14072364709137	-0.74885613291211	3.48547428933473
C	-1.93214362701649	-0.08227500314548	3.36550087789258
C	-0.74837653223062	-0.81611576513377	3.36093052051621
C	-0.74821376852985	-2.20319820987611	3.48191388333136
C	-1.98252345361588	-2.85255690347608	3.57486679574507
C	-3.17261789696799	-2.14010423708826	3.58318571024214
C	0.53700057623804	-2.96694611173558	3.61228283054632
C	1.71301545801404	-2.14185778602593	3.18786768671134
C	1.60935200350473	-0.76059778411021	3.06285514802684
C	2.71520051693115	0.03312665128705	2.77080346968928
H	2.58936538164514	1.10625726525980	2.66587780199435
C	3.95072514101886	-0.56980213564321	2.60049983443852
C	4.08178954688237	-1.95408734673375	2.71252927154407
C	2.96724370917775	-2.72609630981439	3.00307187238301
H	-4.06509049867697	-0.17965612681586	3.47991477325395
H	-1.88948681217963	0.99695713351908	3.25594606447800
H	-1.99803257002361	-3.93939856883625	3.64380447425146
H	-4.12139903996956	-2.66230415091203	3.65568691075122
H	4.81406562280221	0.04452515927728	2.36402787474326
H	5.05033893588824	-2.42433070980710	2.57472794493896
C	3.05762312590708	-3.80378460720459	3.11404471932091
O	0.41063547243361	-0.09069502301938	3.20661073922659
O	0.77857208388737	-3.33329796818990	4.99374041508418
H	-0.06052302008707	-3.63265530003321	5.36585933775269
H	0.48636496749086	-3.88744112726619	3.00622771090034
C	3.50611814726075	-0.88976343689239	-0.81445397455442
C	2.32166125923629	-1.53133670467691	-0.49496983912349
C	1.17132047100929	-0.77340776262256	-0.28515911824310
C	1.18677564175990	0.61734058261521	-0.37913934724557
C	2.39424215956035	1.23950728792910	-0.70752364724439
C	3.54687454922789	0.50138047581302	-0.92734471222215
C	-0.03965897798572	1.42953597644723	-0.06704061837934
C	-1.26938874826602	0.56384389574359	-0.07097983872326
C	-1.17172092076298	-0.82480010719416	0.00615069188305
C	-2.30560034879031	-1.63195346317988	0.07680557906501
H	-2.18132051497464	-2.70776670108761	0.14591629773869
C	-3.55868030789313	-1.04316826689893	0.06217916610025
C	-3.68504884221929	0.34465061252436	-0.02420622236349
C	-2.54588347503405	1.13217143769836	-0.08766313093585
H	4.40487632033769	-1.47753538394760	-0.97638581304625
H	2.26546237216478	-2.61060702349592	-0.39767563506711
H	2.41384826624517	2.32476598895472	-0.77430115024981
H	4.47511920575090	1.00223004639145	-1.18433863926481
H	-4.44389389220864	-1.67004780323557	0.11777020591264
H	-4.66834673761888	0.80437138730981	-0.04245051462056
H	-2.62753463165588	2.21537796501794	-0.13930456457189
O	0.03313500029724	-1.48361959645228	0.01207635757258
O	0.10170617483927	2.13668555541190	1.18275822223188
H	0.20837808032991	1.46655883840448	1.87576946484687
H	-0.15074392604229	2.23284829785881	-0.80677117659288



Main article molecule number: 10

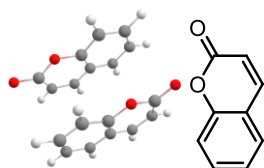
Name: Phenacetin

Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1187.98250256548

C	1.61270133690748	1.22189727766279	-3.25126991339474
C	1.50960448131933	0.72115292497353	-1.94415349021276
C	2.84570832060955	1.41440201891436	-3.84409234332266
H	0.70834572719137	1.48538361083367	-3.79559716533134
C	2.68052996665882	0.40283280381074	-1.25353795278419
N	0.21639616923067	0.57271125437691	-1.41078186703414
C	4.01762699591034	1.10614208577409	-3.14691981503592
H	2.92113858955651	1.81912332909000	-4.84839732523553
C	3.92606557619807	0.59347376553405	-1.85395622152481
H	2.61953941991568	0.01565637004037	-0.24587050643942
C	-0.12916168421283	0.34831740406292	-0.10338106572482
H	-0.54636182399791	0.78361788410559	-2.04840971728661
O	5.18897596079543	1.35153704041753	-3.81085931454523
H	4.81339907002097	0.33470112606159	-1.28765908153843
C	-1.62297579996014	0.29410682983367	0.15271492237174
O	0.67594476485642	0.19463822263810	0.80781942110753
C	6.41111884046577	1.03716747042156	-3.13958277855878
H	-1.86904850742011	1.04705903260002	0.90793421645451
H	-2.23297993005847	0.45091617648337	-0.73992274482510
H	-1.85724790584114	-0.68239704840171	0.58709614628423
C	7.54547240333833	1.41842747512625	-4.06480442542429
H	6.47228997808695	1.59819117942101	-2.19376172289256
H	6.43891193122129	-0.03723044806470	-2.90100663663969
H	7.51400929960003	2.48806824657825	-4.29225725306820
H	8.50557247717177	1.19140879505202	-3.59122705582020
H	7.480492371131065	0.86248518008997	-5.00443832616215
C	4.94624384007090	4.48225093402730	-1.75658889323874
C	3.46352811841813	4.33095670233407	-1.50492777833902
H	5.23736024619814	3.91047862066305	-2.64328811339489
H	5.50505529005464	4.10353688487865	-0.89473371186614
H	5.21381363609940	5.53188997998766	-1.91074113247194
O	2.78272407298702	4.83925286850397	-2.66591733778284
H	3.19963353792027	3.27429068843799	-1.36209508900347
H	3.14888355631827	4.89879046605079	-0.61638612781838
C	1.46425840185400	4.55486604947681	-2.79871264024375
C	0.66450155775858	4.01885113434436	-1.78524124174478
C	0.91019438399468	4.80832851709136	-4.05834186554063
C	-0.66128667226019	3.69928224745437	-2.05143522941416
H	1.06590909370795	3.81832437827423	-0.79889907546489
C	-0.40479386018233	4.46513056196356	-4.32030361920217
H	1.54393305545719	5.23126464476441	-4.83113349394792
C	-1.19331533472304	3.89305471749520	-3.32062774036309
H	-1.27901824704452	3.26906053426934	-1.26950345111353
H	-0.81606944302417	4.61431345043895	-5.31549562931018
N	-2.52866748493699	3.49051268430114	-3.61308055889095
C	-2.98373257543695	2.21239774396156	-3.46585835769084
H	-3.13506916443948	4.15611117914311	-4.06644869147221
C	-4.42182884969503	1.97904264063156	-3.86988371910545
O	-2.29186321216739	1.29701426631723	-3.02719501884311
H	-4.44842694115131	1.15999630796282	-4.59340492180259
H	-4.91360164580876	2.85745518764097	-4.29622942708719
H	-4.98050338884404	1.65217660214911	-2.98787511826311



Main article molecule number: 11

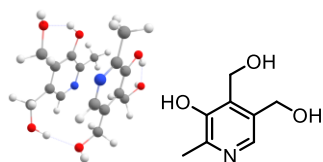
Name: -993.883543951551

Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh):

O	-0.04772284023907	-1.30566853911111	-3.25756047151263
O	-0.15736107692497	-2.06669659384363	-1.16354436938556
H	-0.93822447200873	0.09923796687487	-5.20613071446452
O	0.59988855387132	2.61954693938241	-4.32684820836740
O	0.97546599919684	3.14801677801348	-6.46018792335119
H	-0.90709593780863	2.49530323644956	-2.25969205342262
C	0.57664752742296	-1.22292401360447	-4.46539805463053
C	2.32629629402660	-2.69419752005861	-3.65327325597248
C	-0.01996466089352	-0.42944586336219	-5.43987675198879
C	1.77672047431559	-1.90988508308211	-4.71743731758737
C	2.36407323722935	-1.78981806326278	-5.98740217539219
C	0.46568321323555	-2.05791113695437	-2.19661489441952
C	1.70775696203782	-2.76133805770979	-2.45476687779348
C	0.58553228116340	-0.32328470469758	-6.68316910256057
C	1.77507514157772	-1.00607385275950	-6.96200287584185
H	2.09600772420678	-3.34477151345232	-1.62769778674841
H	0.13426598202597	0.31066253477397	-7.44006648449696
H	2.23537443401457	-0.91310932413631	-7.94033023615861
H	3.25358761214472	-3.23622705791370	-3.82763945584543
H	3.28947921583036	-2.32340498579863	-6.18969042760260
C	0.96748409516065	2.00008545128128	-3.16808886928854
C	3.09084579250067	1.39901168572549	-4.17537169873223
C	0.05134669742920	2.00597442372024	-2.12210981050455
C	2.21460505874525	1.36639801989551	-3.04360329556652
C	2.52634457123879	0.73361318880475	-1.82889358866005
C	1.41542351870460	2.62997108213046	-5.46200464599416
C	2.71665077261337	2.00723159924061	-5.32089237132251
C	0.38246878403898	1.36729775690257	-0.93706406842631
C	1.61988053504923	0.73336817978228	-0.78502956011450
H	3.34960979521482	2.03937657719126	-6.20002806581898
H	-0.33570300555867	1.34644403669069	-0.12375638449687
H	1.85538666695514	0.22508465855306	0.14396738730386
H	4.06329188354133	0.91713757391451	-4.09767681520179
H	3.48856916994161	0.23790462042011	-1.72808877563300



Main article molecule number: 12

Name: Pyridoxine

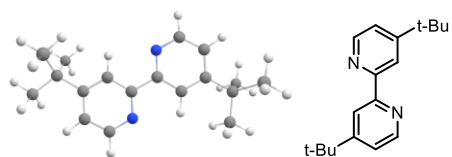
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1183.58286832443

N	-3.93631146687735	-2.14546602556160	-0.12808367331925
C	-2.89480456905776	-2.57903317796810	-2.69792769316578
C	-2.17537706864446	-1.87333333132334	-1.72927112706991
C	-2.73534589002683	-1.67225326465303	-0.45257446501483
C	-4.62742622507160	-2.82841512093701	-1.05320977226127
C	-4.16878450127788	-3.05550218170964	-2.34271016445823
C	-2.41685729522273	-2.77479192121181	-4.11388904173517
C	-5.04659716811544	-3.75835249488457	-3.35959861023272
C	-1.97002392858908	-0.88809446974912	0.57102467514352
O	-0.95166804422157	-1.32232482632227	-1.93704526003644
O	-0.98294860132456	-2.63189567644792	-4.21190931370637
O	-6.39000294159154	-3.91980525431291	-2.94994890528902
H	-0.65211510329741	-1.61498660385683	-2.81995670152384
H	-0.77077193397982	-2.36337425827883	-5.11171580867473
H	-6.78144141162483	-3.03296713640600	-2.88315481349777
H	-2.91237820795194	-2.03008024292411	-4.75689884474491

H	-2.69258757113692	-3.77851009654802	-4.46623246272930
H	-4.66005548019969	-4.76986540688629	-3.55115339781424
H	-4.99575787219460	-3.21418607834351	-4.31972500542170
H	-1.72137196882646	0.11129735274836	0.19429152459661
H	-2.56153268166979	-0.81180386581214	1.48638742493023
H	-1.01439967945732	-1.37211903552775	0.80242776290396
N	-3.80566567028876	0.29228824088542	-4.41760041135129
C	-4.457221222902759	0.97234469062825	-1.77351135100699
C	-3.24719256871377	1.36686498265429	-2.35157405931082
C	-2.95548043476050	1.00335392387870	-3.68290154543460
C	-4.96998824627867	-0.08283114490070	-3.86931112420285
C	-5.34286298522741	0.22393279539878	-2.57009314532195
C	-4.82472360947212	1.21581226322125	-0.33176591376606
C	-6.65880535053613	-0.24734481690578	-1.99574838387560
C	-1.64423998871607	1.40704370491340	-4.28383341937331
O	-2.28977249299111	2.07886212299779	-1.70793213214913
O	-4.01850560965108	2.25287441139090	0.26356247083320
O	-7.37914511387928	-1.14774350676676	-2.83832898612927
H	-2.64593064752269	2.31235859389564	-0.82861561656893
H	-3.90963259085514	2.05030347988559	1.19831520001000
H	-7.68503018534857	-0.67011442166051	-3.61771297685539
H	-5.88097324228453	1.51399017617642	-0.26041652518190
H	-4.70311632868335	0.27015166461950	0.22082902060915
H	-6.48169952635121	-0.80391394264870	-1.06764794537006
H	-7.29626421202155	0.61586253265966	-1.74198105658625
H	-5.59983578343738	-3.21194977073857	-0.75419907483985
H	-5.61917851434175	-0.68588973733494	-4.50046720114456
H	-1.5355759588429	2.49743444031629	-4.29384520580292
H	-1.58034086939347	1.02861484052830	-5.30585536360294
H	-0.81114759397280	1.01933759382242	-3.68655158045576



Main article molecule number: **1-monomer (Fig. S1)**

Name: 4,4'-di-t-butyl-2,2'-bipyridine

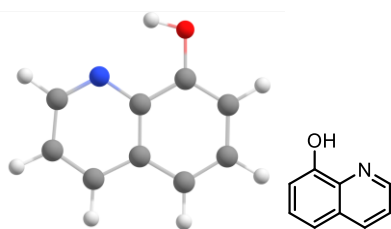
Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -809.684916603281

N	1.88568777204077	11.16674325459677	30.09362165116610
N	3.13137675777119	14.54722965174118	30.31045755615307
C	1.11646144034744	10.33467894347100	30.79752184589665
H	0.96680994578943	9.34270815102191	30.37097169714069
C	0.52049989003213	10.66662319665950	32.00553645771851
H	-0.09195624018654	9.92802184990462	32.51401017886221
C	0.71977082317072	11.94331441436106	32.54457579716895
C	1.52149921372921	12.81150355398342	31.81090385466650
H	1.73495226785471	13.81875320372122	32.14365303362268
C	2.08322538152808	12.39240686113806	30.59902025205684
C	2.94087902690333	13.31428038292739	29.80898853112479
C	3.51075025716888	12.89703223394663	28.60661761069812
H	3.30626370256081	11.88639509291513	28.27259282162241
C	4.31565326423896	13.76659335242277	27.86962098130160
C	4.50571537532551	15.04319808138093	28.39959499219610
H	5.11513879968868	15.78478513143355	27.89562935235732
C	3.89783014667104	15.37622320221315	29.60760455783595
H	4.04101612012284	16.37174442983880	30.02871783862115
C	0.07350399571924	12.33047453915944	33.87283823261480
C	-1.45699171823782	12.21581256846757	33.73836724819395

H	-1.76895746711460	11.19396520080117	33.50156682504579
H	-1.82714974529043	12.87515636343996	32.94615919755229
H	-1.93797607596896	12.50480295345621	34.67971309891372
C	0.56646840695189	11.36870230572828	34.97075560302318
H	0.11748666259253	11.63730862879424	35.93371063421596
H	1.65565221144579	11.42135938332619	35.07208760641943
H	0.29684604939891	10.33103454101757	34.75222136918677
C	0.41970199085348	13.76401971892045	34.28970406032006
H	-0.06555419713981	13.99186632646364	35.24448452222252
H	0.06976062559195	14.49562523769331	33.55389733394048
H	1.49844018158763	13.89765264533583	34.42288810086836
C	4.93819677999229	13.30134351105260	26.55503481350575
C	3.81382567016349	12.89583685964068	25.58290479420128
H	4.24378968044583	12.56294879953289	24.63139558520259
H	3.14870488178647	13.74268705939215	25.38284530151198
H	3.20886707088805	12.07792225428610	25.98481679969767
C	5.78378773609320	14.39506981759877	25.89340767433956
H	5.18476900981330	15.27906081812110	25.64885275983937
H	6.20657816431955	14.01243169517558	24.95870558116443
C	5.84041260322194	12.08352716423535	26.83147624626770
H	6.63870826668293	12.34288209392103	27.53501674431938
H	6.30074588752304	11.73927692583915	25.89837322000671
H	5.27304922583962	11.25026763986406	27.25622743423041
H	6.61847815808323	14.70573596105939	26.53134520298594



Main article molecule number: **2-monomer (Fig. S1)**

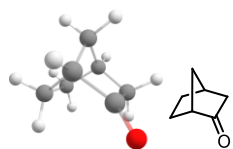
Name: 8-Hydroxyquinoline

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -477.062939536696

N	13.28273251818368	12.89134518786673	6.81453593511704
C	12.88796292487632	11.77263669800358	6.23976153215404
H	11.82874975959226	11.69346873295383	5.99674801425905
C	13.76338037725167	10.70807186846512	5.93812772293519
H	13.37534977168863	9.81223825785397	5.46351165881848
C	15.09534853718796	10.82592476010156	6.25193487010557
H	15.79509237615529	10.02317331343906	6.03206176811675
C	15.56660318346466	12.00606969872075	6.86914579661136
C	16.91428471531573	12.23005566299348	7.23298180543166
H	17.65665884626875	11.46180641193926	7.03703138114179
C	17.26534513916389	13.41912709825106	7.83002420662259
H	18.29992310181257	13.59455866773187	8.11070998894915
C	16.31746452665328	14.42764603867114	8.09250426970810
H	16.61287519601881	15.35922862513564	8.56428849956642
C	14.99740583726187	14.23506891447989	7.74812824599277
C	14.59653284865878	13.01604208042150	7.12778320936030
O	14.06831940262102	15.18354852363401	7.98716160405841
H	13.22673793782474	14.81502245933745	7.65903849105128



Main article molecule number: **3-monomer (Fig. S1)**

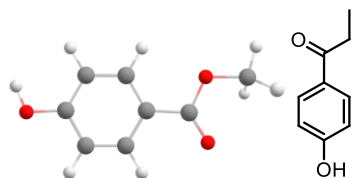
Name: Norcamphor

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -347.90373088891

C	0.51870075833866	1.29505445825362	-0.39453768030950
C	-0.92495754788521	1.03465627530763	0.06874340929238
C	0.16998631877058	-0.86668377367787	0.67221755402151
H	0.63446672574232	1.52602421493894	-1.45873785428489
H	0.98874215009472	2.10681631694529	0.17623588216690
C	-0.68192747819300	0.21885058138186	1.3576172774680
H	-1.60425147289126	-0.18029607841652	1.79359695772752
H	-0.13335838333374	0.77811974679659	2.12405322993522
C	-1.54476652979802	-0.02942049494759	-0.86051306452276
H	-2.61909149340622	-0.12780003296198	-0.67425340420700
H	-1.41548619400753	0.22645172733154	-1.91704818604160
C	-0.78552798816491	-1.33231136760998	-0.46420849389997
H	-1.46987902877844	-2.09086154918258	-0.07089112933918
H	-0.23850489554132	-1.78141582776017	-1.29874451835571
H	-1.52971544171501	1.93794202824448	0.18442672175257
H	0.58922445002852	-1.67139131947136	1.27810791781572
C	1.21922865676530	-0.02172300186194	-0.03658934359973
O	2.35938939397456	-0.32670290330998	-0.29600427589827



Main article molecule number: **4-monomer (Fig. S1)**

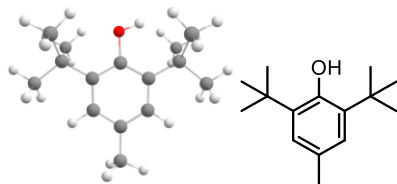
Name: Methylparaben

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -535.260594271769

H	17.90956636744076	9.49520852034183	15.05194227324298
O	14.11359178570963	14.35951895448102	18.15453291883977
C	15.29982396222303	12.88842516475296	16.69508811926987
O	14.01980065765853	14.70933454155191	15.92420649358539
C	15.79538791774179	12.10938325357829	17.74775242469777
H	15.52458709900070	12.37821143059897	18.76380394183927
O	17.75225890092073	9.61646460641000	15.99572165055156
C	16.61171993923992	11.02127791525273	17.50142409149296
H	17.00097942967662	10.41153180217309	18.31024556774631
C	16.94534317713782	10.69673440440707	16.18448806283508
C	16.45865845781015	11.46442350961071	15.12378400008806
H	16.72104077225615	11.20799866636236	14.09942029000077
C	15.64060348039528	12.55382978507080	15.38097209560105
H	15.26257231155943	13.14984960253397	14.55802741244321
C	14.43210409175771	14.03783018266603	17.02828402302531
C	13.16916803118794	15.84030926258792	16.19082516599598
H	13.69105671515362	16.57464225502608	16.81107310584027
H	12.92965405531146	16.26193489754682	15.21401364361216
H	12.25929123691867	15.52216824504740	16.70778971929216



Main article molecule number: **5-monomer (Fig. S1)**

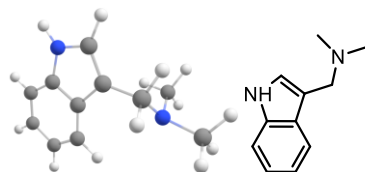
Name: BHT

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -661.112433195317

C	-1.24580697041107	0.16367639212499	-0.00010679246717
C	-1.20445657270553	1.55572807191057	-0.00080453667946
C	-0.00985039341904	2.27259016392019	-0.00075109008396
C	1.18305863314632	1.56310712060862	-0.00051425780852
C	1.22515379757971	0.16510136146893	0.00006135233277
C	-0.00813938501000	-0.51590265040974	0.00038746024206
H	-2.13243026275766	2.11709736619285	-0.00141187601352
H	2.10827114191947	2.12734891388937	-0.00088655887750
C	-2.58478851805513	-0.59012214578974	-0.00016155038682
C	-3.78141274150315	0.37394490401115	-0.00072402829036
H	-3.79673105366340	1.01194586801863	-0.89117998372294
H	-4.70578950382877	-0.21305410654552	-0.00063482976622
H	-3.79699096132860	1.01264937662691	0.88922172453737
C	-2.71039817027253	-1.46186402053977	1.26667841548347
H	-3.69125374659148	-1.95161492115767	1.28178816285126
H	-1.94095836292128	-2.23303220365891	1.30996995030626
H	-2.62777945721709	-0.83942644327613	2.16463068367479
C	-2.70990469511365	-1.46279017430281	-1.26640931562124
H	-1.94047832011909	-2.23401794346405	-1.30877472954036
H	-3.69077836794434	-1.95250209370662	-1.28158382119159
H	-2.62685739877959	-0.84101625296335	-2.16478082688006
C	2.57448960809506	-0.57764373899925	-0.00004335069709
C	2.73615390451046	-1.43536585796644	-1.27751985814350
H	3.71786682973408	-1.92245990438262	-1.27644844771343
H	1.98627383110858	-2.22273776043111	-1.40643084408553
H	2.67057015643334	-0.79517305642866	-2.16300707915255
C	2.73679960853744	-1.43457133322580	1.27788150356787
H	3.71849433208275	-1.92170037172939	1.27658494679014
H	2.67170902593086	-0.79383526282969	2.16301187738396
H	1.98698397041816	-2.22185804060822	1.40770813451686
C	3.76329312852984	0.39813519703836	-0.00062731780060
H	3.76802284799562	1.03578820408186	-0.89038394825499
H	3.76841112603178	1.03639470843287	0.88869193243740
H	4.69505916181898	-0.17718070525539	-0.00064146964117
C	-0.02384654889226	3.77827937431575	0.00100286349741
H	-0.52331892965077	4.17086081663782	0.89444940484081
H	0.99109960509531	4.18508413952823	-0.01999317487595
H	-0.56100126492576	4.17270695685432	-0.86929022927566
O	-0.06978880473339	-1.89015857978773	0.00085911355679
H	0.82390672087584	-2.24243836820282	0.00110139095097



Main article molecule number: **6-monomer (Fig. S1)**

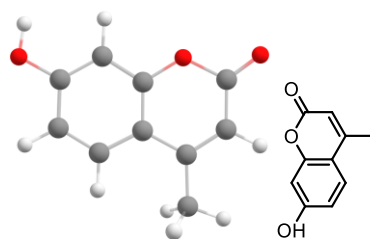
Name: Gramine

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -536.954162247783

H	10.59285936876690	14.19369082595759	13.31151605810053
H	10.72981204261871	16.03919265252571	11.16627002253114
H	10.47438549576812	17.42908906350500	16.60786037656080
H	10.74350828643377	18.60775540014236	17.90808065233962
N	11.11166600387260	14.92202888045987	13.76949747693277
N	12.47378915506655	17.64804989739854	17.21522120313990
C	11.56211514764697	14.88203878976847	15.07526006149389
C	12.24476573423588	16.03359673700090	15.36690940371926
C	12.71720151345687	18.10824839887639	13.85654792634735
C	12.50769121969003	18.60222890775338	12.57877405692853
C	11.79821275293730	17.86124628110738	11.61698141210320
C	11.28067657132078	16.60805518367982	11.91038802236325
C	11.49626362108889	16.11346945184651	13.19485609914137
C	12.21328285057136	16.84148334215299	14.17972411980827
C	12.91697029839091	16.36410163433528	16.65724687434695
C	13.24757003500001	17.96946427211195	18.40845282786870
C	11.04647885736318	17.62996806457759	17.51867376036748
H	11.36306528767385	14.02395938469365	15.70350985115468
H	13.24704416033834	18.68801345656849	14.60597342464837
H	12.89191728505301	19.58245647342229	12.31245726226900
H	11.649066614143388	18.27886388916488	10.62563708312026
H	14.00327308961579	16.43863583081911	16.49889609838060
H	12.75639904391386	15.52920438634463	17.37444279345797
H	10.78083120595422	16.85957657331297	18.27232963040199
H	12.94264016609086	18.94751560730103	18.79617108129209
H	14.31396831841748	18.01872305377206	18.16036866867697
H	13.11809834727978	17.22414456140104	19.22044275250493



Main article molecule number: **7-monomer (Fig. S1)**

Name: Hymecromone

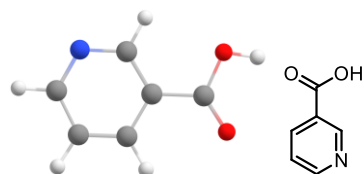
Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -611.466363427428

O	2.15904000166639	4.98358053503154	10.16201776117718
O	1.52962202350872	4.68939004107549	12.28165391202642
O	3.35407198238391	5.74405832116977	5.63455965849819
H	2.62391083365637	6.37052162932369	5.70426546463256
C	2.29672614231088	4.35596432819246	11.41265056857904
C	3.34929524840374	3.37459764885549	11.51709201556346
H	3.44251353947056	2.90028670022390	12.48803996590587
C	4.17133113444605	3.05589277477603	10.48449394972899
C	3.99322740183861	3.72441354800663	9.21975704727955
C	4.77619833438181	3.48604151865611	8.07608055578131
C	4.55516942596771	4.16092387082156	6.89396279683675
H	5.16012787437482	3.97567394109072	6.01325353799980
C	3.52556046251992	5.10970813614529	6.82541104524383
C	2.72884615683932	5.37490420702534	7.93272918589494
H	1.92513862159772	6.10538684615882	7.90343961621565
C	2.96906839015190	4.68265655447368	9.11442863931636
C	5.24919539552224	2.02829425781663	10.64450782912898

H	5.10795789610106	1.20202909064208	9.93819332617419
H	5.57514440932554	2.75305841924532	8.12163853129429
H	6.23526708871745	2.46258533622647	10.44325151396402
H	5.25170963681528	1.62030729504294	11.65688107875857



Main article molecule number: **8-monomer (Fig. S1)**

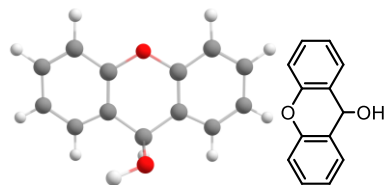
Name: Nicotinic acid

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -436.786352736965

O	7.38050596456261	11.45004572001977	4.54484546818390
C	4.96111726017770	14.89932575258003	4.18785695474100
C	7.62232553976778	15.08140517710174	4.74823406123930
C	7.05504565788738	13.81460162924352	4.57693781661914
C	5.69222280367489	13.72755297283969	4.29030335462650
C	5.61960247036313	16.11270161974874	4.37570814911084
C	7.83928889201111	12.56188872523937	4.68636818380568
N	6.92417580462425	16.21451025657821	4.65105502953557
O	9.14796314470295	12.78587293157785	4.96621562813345
H	9.56830633267468	11.91323568791007	5.01653571989991
H	8.68213185003068	15.18059648157923	4.97231512282394
H	5.07188367534410	17.05177810899120	4.30147839266988
H	3.89862430942547	14.88050502230856	3.96636983338080
H	5.23927329475322	12.75005891428193	4.15498228523006



Main article molecule number: **9-monomer (Fig. S1)**

Name: Xanthidrol

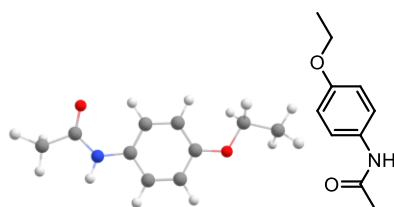
Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -651.722430671294

C	3.54478826820433	-0.99820354766330	-0.11966261171163
C	2.32559935481289	-1.59562997414977	0.16402544649925
C	1.17606394672911	-0.81038690804122	0.20807638815282
C	1.22929354308719	0.56671069832653	-0.01133990766099
C	2.46477975955328	1.14062640922864	-0.31939738218251
C	3.61955180235942	0.37182841203638	-0.37372767831104
C	-0.00121247107381	1.40161426227577	0.18704144451841
C	-1.23595331744550	0.57403097341898	0.00388170030116
C	-1.18120173566896	-0.80412902599834	0.20483913671649
C	-2.32736497139043	-1.59251188916570	0.14746251257621
H	-2.24098114563524	-2.66222993064053	0.30848594449069
C	-3.54769180243458	-0.99185646834371	-0.12307084727032
C	-3.62510405713376	0.38354698940263	-0.34476904273375

C	-2.47262508218134	1.15386169602819	-0.27836913966403
H	4.44207514705633	-1.60868505687596	-0.15810307459528
H	2.24183125787051	-2.66306009568245	0.34134400389428
H	2.51076215039854	2.21055927903382	-0.51496658914321
H	4.57221530993776	0.83277994170774	-0.61408966670233
H	-4.44430757978597	-1.60276261285127	-0.17078276660576
H	-4.58107094215269	0.84931714634219	-0.56166289985447
H	-2.52054598151108	2.22955030409813	-0.42862661720336
O	-0.00183901450828	-1.46553277110829	0.45556755422721
O	-0.05272696337571	1.92030553107120	1.53950311821217
H	0.83674246817270	2.22067918470082	1.76604870312307
H	-0.00250894388470	2.24675545284951	-0.52093372907309



Main article molecule number: **10-monomer (Fig. S1)**

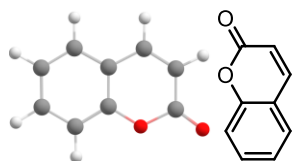
Name: Phenacetin

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -593.979783924289

C	17.74205714886844	12.91407312400948	14.59261173086514
C	16.39405072331122	12.64977126674120	13.95965680654202
H	17.77985507570187	13.92316748360957	15.01269818691488
H	17.92203033972672	12.19502421494327	15.39770904852145
H	18.54169147573750	12.81558136644415	13.85298141208641
O	16.21369633640752	13.61539820748727	12.91931594065203
H	15.58579792643600	12.74646253707105	14.70075524090359
H	16.35051886354021	11.63494217915780	13.53526601877704
C	15.05740922515344	13.56423321221342	12.19649699375934
C	14.02533678619379	12.64937520268952	12.40250465902209
C	14.93268959329149	14.51672245359415	11.18066738362537
C	12.88029744840581	12.67875253508221	11.60794555680787
H	14.08898421397797	11.89810604963173	13.18093145095160
C	13.79780891922145	14.54555626494458	10.39318994978223
H	15.73835748756083	15.22671165220991	11.02393170848223
C	12.75639489145823	13.62863370437497	10.59404928194543
H	12.08496429731820	11.96577631826251	11.77539062376106
H	13.71804935343379	15.29387995312334	9.60707607042250
N	11.63556792613866	13.72871046330153	9.74287837256709
C	10.48908710135239	12.97292966744049	9.73125776779359
H	11.69121618967087	14.45942089548078	9.05119585184063
C	9.48289074874491	13.37026331211415	8.66739161567248
O	10.26714547117619	12.05168683736181	10.50264463354835
H	8.55179080439371	13.65660852049737	9.16493649011610
H	9.80594854775526	14.18880917969942	8.01774594607100
H	9.26673310502335	12.49118239851413	8.05412325856838



Main article molecule number: **11-monomer (Fig. S1)**

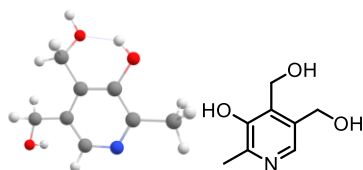
Name: Coumarin

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -496.934305122779

O	11.26467093682450	14.63339543011263	20.83493012569925
O	11.833393017688908	15.33752907975128	18.79539919070798
C	10.41548855206400	13.98174701798808	21.67731737076810
C	8.95999001099150	13.47122112870962	19.80347074821097
C	10.76122724031193	13.93810123077924	23.02444199389168
C	9.23794392663870	13.37742106980483	21.20616790753140
C	8.40585246733131	12.72122541603473	22.12697946029436
C	11.02281877766295	14.74231455425324	19.45876136917057
C	9.80562920459929	14.11842211106650	18.97404060124709
C	9.92061569303537	13.28306208970246	23.91172589073197
C	8.74209949135404	12.67330911165921	23.46760504769515
H	9.62611183798490	14.20807738633700	17.90884348911040
H	10.18452133726485	13.24511558888183	24.96434233247886
H	8.09457975413615	12.16455965929896	24.17424845712779
H	11.67885053458214	14.41619359316512	23.35074099630344
H	8.05291827921928	13.01017018720110	19.41845390924522
H	7.49350977910997	12.25197234525408	21.76790110978564



Main article molecule number: **12-monomer (Fig. S1)**

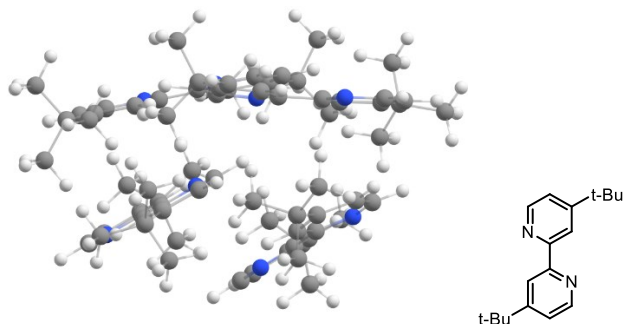
Name: Pyrixidine

Qty of copies of the molecule modelled: 1

Qty of imaginary frequencies: 0

Final single point energy (Eh): -591.777504554973

N	28.49391257705194	17.36563795917214	21.69068063465522
C	27.61849461937692	16.23863251192103	24.09486816985234
C	28.71567045457061	17.10244718278753	24.04895169607216
C	29.13247748762625	17.65063932213039	22.82369978562905
C	27.42939368913781	16.55428694859214	21.73008699413924
C	26.94131733522420	15.97023417639186	22.89459097170882
C	27.26641288277820	15.56229196712654	25.39006764208153
C	25.71960315790652	15.07406587222670	22.85250373414311
C	30.32030913729839	18.56217130669764	22.77313542317475
O	29.43895850876515	17.42912421823840	25.15555912008028
O	27.21254420769084	16.55363766615432	26.44293392544186
O	25.02517635859085	15.10266368938966	21.61400430866123
H	28.88619607353478	17.24080360780113	25.93545453026885
H	27.11263254498396	16.09314386556500	27.28416941811806
H	24.70851665962274	16.00272836267202	21.47181226776648
H	26.30491361226536	15.04110390848833	25.32087088293760
H	28.04185213082653	14.81399209494814	25.6252622425294
H	26.00922028107950	14.02470399968556	22.99649926860355
H	25.03921043411453	15.33645822918994	23.68012422461359
H	26.94729365087563	16.34712524016666	20.77793427735948
H	30.16983234752582	19.43581649762378	23.41697249922164
H	30.48189951071828	18.88813138109847	21.74461158550829
H	31.22025733843490	18.05630299193248	23.14078241570973



Main article molecule number: **1-tetramer (Fig. S3)**

Name: 4,4'-di-t-butyl-2,2'-bipyridine

Qty of copies of the molecule modelled: 4

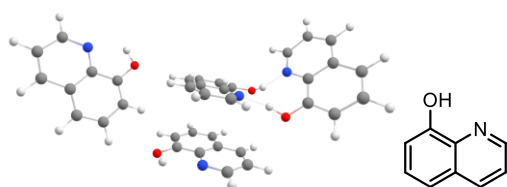
Qty of imaginary frequencies: 0

Final single point energy (Eh): -3238.803750464372

C	-0.44501146663687	-1.86880707826165	4.51364832781682
H	-1.21106191093488	-1.43894509416670	3.85935725044641
H	-0.82567884439991	-1.86481100401742	5.54167402556756
H	-0.29368496003424	-2.91353363918419	4.22695196892942
H	-0.13170253403681	0.89871500348957	4.36038928879017
N	-2.76049176969762	-0.76169659086299	0.37413771377015
N	-0.96114680598033	2.03766938360875	1.77097447530966
C	-3.73053294060553	-1.61361109283190	0.71447277237706
H	-4.02175815289421	-2.33226449584509	-0.05246112142927
C	-4.35166030373144	-1.60852950918879	1.95585220759394
H	-5.13495752419573	-2.33314865924540	2.15723773286191
C	-3.95708945718726	-0.67745797086857	2.92390050045997
C	-2.95363943916612	0.21613959971606	2.56036249159542
H	-2.59048761455896	0.98056114058824	3.23518014854695
C	-2.38469319915864	0.14766749412700	1.28416719831326
C	-1.34570287233019	1.12282268604180	0.86421963352945
C	-0.83505234065985	1.09193224386594	-0.43318702107794
H	-1.19387070450935	0.31700974561814	-1.10030699915116
C	0.10032912009500	2.03867190970264	-0.84681912917476
C	0.48832263097137	2.98699084409294	0.10143066287174
H	1.21219714020134	3.76066895004794	-0.12959317064100
C	-0.06299708828265	2.93789611243234	1.37888192236484
H	0.23937924832027	3.67182094137637	2.12695136200814
C	-4.62332241679728	-0.65954019701283	4.29798897014091
C	-6.11615630571613	-0.32193893623419	4.12097422857344
H	-6.62266291573343	-1.05763500815834	3.48853581258547
H	-6.23966395431864	0.66331088158850	3.65899362013484
H	-6.61643256984848	-0.31133410789572	5.09607679240158
C	-4.48715278163097	-2.04946160905636	4.94801891788937
H	-4.94535255457721	-2.04188759870048	5.94338782103877
H	-3.43363179964081	-2.32887970267566	5.05515878944911
H	-4.98206617440600	-2.82556922617466	4.35661299068018
C	-3.99214827635401	0.37817389455064	5.23287927207558
H	-4.49011205183015	0.34128203344816	6.20738973600152
H	-4.10004526628785	1.39545818586749	4.84233376478988
H	-2.92606459426085	0.18124579620416	5.39235434218157
C	0.61025680288683	2.03283204906695	-2.28638729061204
C	-0.54748777313690	2.48416958745078	-3.19886458821783
H	-0.23802436859197	2.44869990857379	-4.25042875841005
H	-0.85438963258713	3.50916585957278	-2.96373662764799
H	-1.42237811423875	1.83856284682697	-3.07221268029840
C	1.79508978836321	2.98533337160492	-2.48390391435983
H	1.52019694819587	4.02782177338922	-2.29198757793513
H	2.14317129277681	2.92423536487717	-3.52013018444871
N	1.64051964158834	-4.46097764892660	0.63580720558306
N	2.34090231313602	-0.94178309385611	0.35315069897019
C	1.59256456550477	-5.54206820996724	-0.14540828106220
H	1.40376150518573	-6.49083856719445	0.35664831772014
C	1.75596384409080	-5.51048011942569	-1.52227422994731

H	1.68212348304442	-6.43653370019876	-2.08476331222292
C	1.99448239532683	-4.28722506882561	-2.15906366632839
C	2.04479836752805	-3.16107564863671	-1.34559725512682
H	2.22212622942082	-2.17173974560264	-1.74579425278037
C	1.85977705877352	-3.28375585408114	0.03549538908404
C	1.87367092399243	-2.07688434746455	0.89985260755179
C	1.39755135860372	-2.13722885992806	2.20998192085375
H	1.05810925314295	-3.09697307448251	2.58002061327017
C	1.36680633708184	-0.99077040303667	3.00318395361393
C	1.86101139240793	0.17954945696932	2.42555917400038
H	1.86928313339611	1.12038252793442	2.96291071655720
C	2.33565858025880	0.14617138617131	1.11899073263690
H	2.72299516648355	1.05798257839512	0.66395448037666
C	2.18723695867606	-4.22127379950129	-3.67168096327383
C	3.41068105664428	-5.07450551490823	-4.05656737439877
H	3.27835410506777	-6.12239522019198	-3.77031098425144
H	4.31519626010700	-4.70061792472235	-3.56483197912750
H	3.56699001355685	-5.03815740361493	-5.14090902605866
C	0.93279950632612	-4.78076430193875	-4.36683276723199
H	1.05961994882177	-4.72699868368908	-5.45532941836665
H	0.03801358657664	-4.21507944483775	-4.08328956122076
H	0.75411523598946	-5.82663265055167	-4.09826081752594
C	2.41842508674930	-2.78772911785051	-4.16127274683713
H	2.54572365627680	-2.79056414499532	-5.24924011184407
H	3.31956663781467	-2.34715435902763	-3.72083005717978
H	1.56677048377760	-2.14043295285689	-3.92623928678722
C	0.85545111859793	-1.04977674244034	4.44105099513216
C	1.93363601713340	-1.73338462611271	5.30396145743728
H	1.60170850037347	-1.79187487699540	6.34727993342271
H	2.87288870909004	-1.17081380959472	5.27138576788556
H	2.13463712523716	-2.74954152271193	4.94866508661136
C	0.56940070472622	0.35025179760815	5.00056877069472
H	1.48309873065785	0.94539418779258	5.10232893100569
H	0.12865292778391	0.26130446280026	5.99942979322429
C	1.05692182416781	0.61434417818081	-2.68496935851873
H	1.83566163930091	0.24341704407079	-2.01032372424349
H	1.45413643459937	0.62786606408132	-3.70678515051220
H	0.22840676521788	-0.10121987738571	-2.66046175168136
H	2.63542351774781	2.71966330796372	-1.83268496107654
C	-5.79577278511310	-5.70784285654318	0.33406493911624
H	-4.70536529968786	-5.63697039295574	0.40686922672472
H	-6.17779351138768	-6.06980695339503	1.29508613897387
H	-6.19893712207879	-4.70171477128066	0.17208550505567
H	-4.58280541205933	-8.08927447586115	-0.46945738728471
N	-1.68339058123297	-5.10765092902843	-1.85902645915636
N	-2.35523170242839	-8.64920563682295	-1.99025445054631
C	-1.54345803639061	-4.15220072478273	-0.93927658172182
H	-1.33173380130974	-3.14888043272403	-1.30797338510261
C	-1.65565717831795	-4.37150365583107	0.42784011055476
H	-1.50845432958122	-3.53342015691084	1.10180892820577
C	-1.94000942509637	-5.65689802080090	0.89932767461355
C	-2.07659212300885	-6.65453740864432	-0.06140212858415
H	-2.28738981810576	-7.68174417722147	0.20538467019666
C	-1.94614519312231	-6.34661178234068	-1.41971032376110
C	-2.09820005438175	-7.41231930287546	-2.44343345000613
C	-1.95982832158816	-7.12226893880626	-3.80601019003795
H	-1.78249539582177	-6.09169623179198	-4.08652873348958
C	-2.05641340188373	-8.14121190514054	-4.75019890353616
C	-2.31519318534972	-9.42672552376470	-4.26115623258657
H	-2.39908808814559	-10.27978989474968	-4.92698570098226
C	-2.46556509872215	-9.62088677706268	-2.89543716477847
H	-2.68031585957862	-10.61697374814856	-2.50729137629259
C	-2.05369614244261	-5.94411704060752	2.39632250716083
C	-0.63569585475442	-5.94558096801985	2.99566487948624
H	-0.11608369893862	-5.00538650775469	2.79150345841878
H	-0.03220469297970	-6.74980882777670	2.56151689777708
H	-0.68411886531275	-6.09765692858167	4.08061051698948
C	-2.89630395442782	-4.85573160333041	3.08271472077604
H	-2.97615850048688	-5.06841641809586	4.15496489684983
H	-3.90848798549228	-4.82513084865671	2.66380975442173
H	-2.45910862869732	-3.85957326429763	2.97485008805795
C	-2.71169728998607	-7.30208120007281	2.67423262009693
H	-2.82233812715947	-7.43795126843763	3.75534323660564
H	-2.10970184505696	-8.13549757674246	2.29902844860730
H	-3.70715582723161	-7.36430262848555	2.22014350493278
C	-1.83481572396089	-7.90877143566444	-6.24424168067971

C	-0.44246628386243	-8.46169917398485	-6.60771528291489
H	-0.24941577480407	-8.32169794046452	-7.67778864903599
H	-0.36553783646032	-9.53069313922657	-6.38198866217361
H	0.33948986347491	-7.93904769356254	-6.04583784215548
C	-2.90468166674614	-8.64394749404404	-7.07074842717256
H	-2.85961886435145	-9.72898622230113	-6.94101878575700
H	-2.75575560099084	-8.43444557383871	-8.13564180025415
N	-3.80963824206083	-2.55617299288604	-2.81138031208969
N	-4.82385433736308	-5.44348854164559	-4.70857241515447
C	-3.14818397601610	-1.41003665733419	-2.99260542433202
H	-3.10682503516653	-0.74479932432262	-2.13041563941147
C	-2.52942750335716	-1.06073828216667	-4.18401419393103
H	-2.00732111083612	-0.11112056530150	-4.25216261726243
C	-2.59538970789561	-1.92921447822295	-5.27935753788841
C	-3.28836323313377	-3.11945469196258	-5.08972769361088
H	-3.40014278873404	-3.84770656613875	-5.88126976668116
C	-3.87188104934275	-3.40193695732846	-3.84834253667087
C	-4.59973041172099	-4.67800676750965	-3.63158857771438
C	-5.01899467281217	-5.04120236124511	-2.34698070817722
H	-4.78593904904911	-4.36650589422712	-1.53373064206870
C	-5.70001534750262	-6.23566041349553	-2.14514419670437
C	-5.92043181043813	-7.03032806153200	-3.27686051023801
H	-6.44042488431981	-7.98097228420125	-3.20869678989879
C	-5.46988464979300	-6.59478902734315	-4.51322641866915
H	-5.64290457756714	-7.20520332208315	-5.39966175197406
C	-1.93038570538370	-1.55053332872737	-6.60081777266329
C	-2.53196242337392	-0.22576965699949	-7.10699639878482
H	-2.36398449511570	0.59254029759592	-6.39973205358988
H	-3.61175013176584	-0.32334467810042	-7.26153688745580
H	-2.07044271773755	0.05327299630897	-8.06102698648134
C	-0.41985978437075	-1.37087055930165	-6.36321190370391
H	0.07750764954265	-1.08028419398407	-7.29592349193610
H	0.02720931456467	-2.30538767098259	-6.00954592921675
H	-0.21644233225145	-0.59721987460707	-5.61540773141895
C	-2.12932146190046	-2.62169164248879	-7.67793344107627
H	-1.63730191867018	-2.30468361148563	-8.60361092071214
H	-3.18997659430604	-2.78122218944483	-7.89984988378321
H	-1.69267126061686	-3.58045321596299	-7.37847031576686
C	-6.21984008355319	-6.67395972494808	-0.77824302587012
C	-7.75944412157359	-6.71675185779030	-0.83170573351720
H	-8.16064940033989	-7.02437075221010	0.14096957489382
H	-8.11577438176460	-7.42674843497144	-1.58467720424615
H	-8.16728956354003	-5.73026380681678	-1.07742439830788
C	-5.67691891894143	-8.07602737020033	-0.44671672013613
H	-6.03144964300846	-8.82648852701582	-1.15992648734069
H	-6.01311282509967	-8.38074754649106	0.55128138802104
C	-1.89453904549437	-6.42167837049354	-6.60891194904097
H	-2.86155661843794	-5.99158110237289	-6.32780891267242
H	-1.76229841043434	-6.30669633415843	-7.69047629186773
H	-1.10269961727973	-5.84836225895239	-6.11632666667277
H	-3.90858768981347	-8.30622270877899	-6.79405243550504



Main article molecule number: **2-tetramer (Fig. S3)**

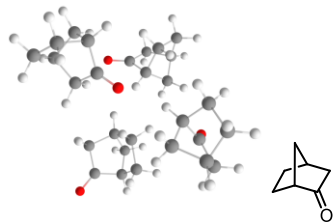
Name: 8-Hydroxyquinoline

Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1908.29156454211

N	4.78446973345379	-4.49078716841785	-2.58004498628625
N	-0.83919950326481	1.70222116546753	-0.00167193273208
N	-0.96605216384578	4.92943406991970	-0.07454304422958
N	-2.91380675258471	-1.90905318751997	0.90919987755132
C	5.36343187092662	-5.39505967290430	-3.34511949620321
C	-1.26753113750145	0.85933157272310	-0.92348919699547
C	0.03435380979349	5.74932090668890	-0.33276278825697
C	-3.88791816280399	-1.03076674845991	0.75889292973677
H	6.22050415400605	-5.06490702448891	-3.93132632055329
H	-2.24733303682478	1.07216995088924	-1.35096182871165
H	0.99615076039429	5.49119492907700	0.10998073229488
H	-4.69050149877722	-1.30396017041801	0.07390446115218
C	4.92425301535858	-6.73324222802480	-3.42702393093554
C	-0.53913013031128	-0.26917082115592	-1.33915947871102
C	-0.10273911221685	6.90974336068975	-1.11842527267088
C	-3.92075010852598	0.20973026218169	1.42556020366437
H	5.44510018631850	-7.43004362649255	-4.07627260474625
H	-0.94717877625403	-0.92230812535078	-2.10352540325664
H	0.76047293190188	7.54074060946609	-1.30371366652754
H	-4.73454106485789	0.90238186957111	1.24018443463576
C	3.84149035612612	-7.13100212780408	-2.68103774296902
C	0.67273172997671	-0.52737239596111	-0.74927688236246
C	-1.34106925065824	7.22682609692853	-1.61651308053135
C	-2.89818813264838	0.53934100408188	2.28188476030015
H	3.47935422834465	-8.15541602067778	-2.72316170717836
H	1.26966911613182	-1.38750853323169	-1.04098202045687
H	-1.49163926694592	8.12560118173323	-2.20978872671805
H	-2.88518788021129	1.50020717797617	2.78985065038476
C	3.18671038240160	-6.19853342181870	-1.84619932917729
C	1.16345417715633	0.33686672152149	0.25451402433169
C	-2.44477712733623	6.38439743161550	-1.35448174959184
C	-1.83981927836670	-0.37185525205330	2.48571198648392
C	2.06465316435682	-6.49979541216474	-1.03995909981301
C	2.38314117046103	0.10472806215948	0.92779193838402
C	-3.75269897208644	6.66805807547019	-1.80475377243695
C	-0.73831222714779	-0.13251529287011	3.33984035945875
H	1.65939078289444	-7.50766505238147	-1.04352098502918
H	2.98060593328655	-0.75968958650278	0.65518256991422
H	-3.93369723481226	7.56720972321829	-2.38632729067227
H	-0.67632331264076	0.80822338246987	3.87880652506237
C	1.50053839599547	-5.51313638547204	-0.26453745003997
C	2.78552547832670	0.97680248700406	1.91150330311254
C	-4.77551002609833	5.80281715704129	-1.49365420443367
C	0.24572307832154	-1.08466976230276	3.46016491485408
H	0.63569587344321	-5.72831021796236	0.35496333751230
H	3.71991041697149	0.80391246112918	2.43857404859683
H	-5.78631405685978	6.01712969328169	-1.82904689206751
H	1.09773816244300	-0.89800486248550	4.10682234696658
C	2.01465959043698	-4.20093159845740	-0.24627483698765
C	2.01833605792254	2.10589465049809	2.24447069372660
C	-4.54020090397861	4.62768425941101	-0.76106385398287
C	0.18989495307329	-2.30372370809542	2.75330884385888
H	1.54997739537430	-3.44293227712964	0.37567363390384
H	2.35026406776554	2.79672046427417	3.01361999975016
H	-5.35169889108319	3.94123544568033	-0.54020731280778
H	0.97262925605682	-3.04732641641851	2.86953147491588
C	3.10714623151032	-3.88330180953321	-1.02022764469378
C	0.82676436031687	2.37386201816384	1.59704763014565
C	-3.27019415764957	4.30474573097356	-0.31682414161932
C	-0.86729912218006	-2.55590937221438	1.90962738502728
C	3.71494487174080	-4.87601840922089	-1.83987455681681
C	0.36276312716688	1.47154131480810	0.59164360907986
C	-2.19483993534363	5.20789099520979	-0.58201379526013
C	-1.90478654721165	-1.59271979867252	1.75784356068485
O	3.62048169082516	-2.62987163308874	-1.02757747379745
O	0.11691958540562	3.45834202694718	1.96118182645984
O	-3.09172161574540	3.16827578549163	0.37329090067274
O	-0.95125623120914	-3.71255782108731	1.20905008678223
H	4.37209881288084	-2.65878027385515	-1.64978416084283
H	-0.43993002227618	3.79946973114327	1.20774425668716
H	-2.14830363642629	2.82378976002169	0.32383421489604
H	-1.77866663258195	-3.63912232023306	0.69512694011431



Main article molecule number: **3-tetramer (Fig. S3)**

Name: Norcamphor

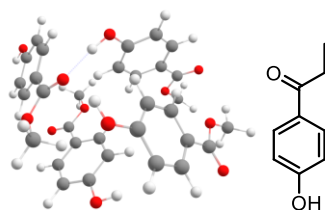
Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1391.6439489357

C	8.26373237254887	12.52156515022905	4.19815617658879
C	8.11998252597183	11.88367931899099	5.58995506372297
C	9.64509428867385	13.53945459587279	5.92529636642270
H	8.56632745243024	11.83413522280277	3.40124562831811
H	7.33763579280311	13.01711389945642	3.87783152409147
C	8.28047946559827	13.11754637102790	6.50503560563532
H	8.33911523469692	12.85550950561687	7.56714116983099
H	7.49778951838392	13.87109400926456	6.35908677328525
C	9.41234077694978	11.10417495713754	5.90250531966447
H	9.29775283585902	10.53011152320610	6.82639307305100
H	9.67734860698724	10.40229258608373	5.10333997261461
C	10.46951589224239	12.23099499055010	6.10197702582217
H	10.88913683071807	12.20111094453678	7.11339453339667
H	11.30059534327358	12.17597310642491	5.39110581744571
H	7.20107388495721	11.30486224107036	5.72248304815956
H	10.12574264950086	14.44160811019580	6.30701899922999
C	9.33616532237186	13.59077497540845	4.43647079214924
O	9.84724794356801	14.29951015739465	3.60057204006555
C	13.09045340700478	9.37529999467384	8.05239010216965
C	13.53497718535368	8.42720837602270	6.92471761816142
C	11.26817256280277	8.23846622542111	6.91109899012364
H	13.52476786428862	9.17430663219368	9.03759559459706
H	13.29842457941357	10.42468952104750	7.80268365876461
C	12.38135845009512	8.60887637308848	5.91419317870431
H	12.42305417295172	7.91654993139336	5.06790861946054
H	12.29473701863041	9.63371552641892	5.53611193644874
C	13.30807873597266	6.97638266063405	7.39916444938379
H	13.76356713791392	6.26571192397423	6.70214761942864
H	13.74504177452080	6.79604319589937	8.38693796397335
C	11.75458301516506	6.84394478745154	7.39868056467956
H	11.41624887015417	6.08917275930235	6.68157327857100
H	11.34506685773456	6.58290449759749	6.38001642376365
H	14.54066636481506	8.62413681589488	6.54401794164103
H	10.23883774675493	8.26338937327498	6.55889897638633
C	11.57182649126001	9.17534011848294	8.06408216982672
O	10.78975050559014	9.65687681226102	8.85608683025887
C	5.08856749986137	8.34288470025257	7.78143562370430
C	6.22010282336037	7.90121809394663	8.72717620010118
C	7.29736027498771	8.30938558747222	6.76732273410325
H	4.26729170854467	7.62804412371325	7.66050633509249
H	4.65188053594441	9.30142527305724	8.09222478767887
C	7.38518898002386	8.78466846533226	8.23113242404004
H	8.34919119061516	8.55297088515077	8.69440439420557
H	7.19095723894345	9.85778302501546	8.34659951566161
C	6.68924476900480	6.49657417811974	8.29479036439764
H	7.38274660096457	6.07791117290131	9.03094848201135
H	5.85180618534124	5.79825128687073	8.19424122504119
C	7.40877025258644	6.76690478159138	6.93954403284342
H	8.46805519736597	6.49284151914301	6.97696059601148
H	6.96112055567559	6.22482110972371	6.09971530458013
H	5.97591751624425	7.98724148251232	9.78905639864829
H	7.98293101766477	8.73061647993362	6.03318152575752
C	5.83424138727532	8.55764131579649	6.46221618405333
O	5.34084004847201	8.84807257433679	5.39213775776166

C	7.76329649894916	8.67575198351519	2.85643281737985
C	6.80300126406326	7.91809673715792	1.92222883830807
C	7.96825208080341	6.25598604751217	2.94687084221560
H	8.40927693945391	9.41535678190603	2.37078092483061
H	7.20298183589823	9.17293715368905	3.66154645180757
C	6.49307351264124	6.66954427784778	2.77556024293005
H	5.90643541308142	5.91620086851161	2.23836160345766
H	5.99465568536264	6.92177089499230	3.71717024977669
C	7.62583176914650	7.31449079791854	0.76546627336176
H	6.96757880412426	6.90632350659565	-0.00818607026409
H	8.27091547528071	8.06025000591655	0.28948965459911
C	8.43834863999714	6.18210222122711	1.46356704956690
H	8.18286242677939	5.19730275344703	1.05929991389235
H	9.52148007072203	6.30764409350816	1.36922939456791
H	5.93559918894538	8.50304124994775	1.60770517687248
H	8.20907956871739	5.37641161514450	3.54683218012133
C	8.58208731646921	7.54269183279500	3.46722864577060
O	9.54837024872447	7.65453393318853	4.19439707927001



Main article molecule number: **4-tetramer (Fig. S3)**

Name: Methylparaben

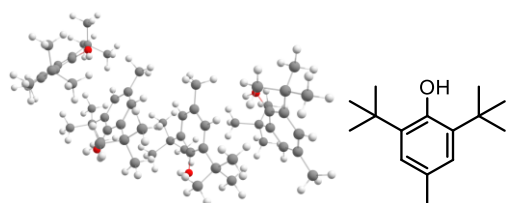
Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

Final single point energy (Eh): -2141.10037416641

O	-0.24220178189292	-1.47184632429158	-4.45674117200936
C	-0.57234941435552	0.29122265684313	-2.87549418256951
O	-0.22817569039748	-1.94541116792921	-2.24564350324148
C	-0.59519997737067	1.30200198437623	-3.84537031000551
H	-0.46656620657472	1.02762735351184	-4.88813285316328
C	-0.76051316781811	0.63617188993899	-1.53303113006920
H	-0.77274894197659	-0.13018679104377	-0.76931071042750
C	-0.34064381980519	-1.10099021795512	-3.30315615096065
C	-0.02025373601769	-3.32904684607657	-2.58164194016139
H	0.90660141239750	-3.44834108678056	-3.14950655355830
H	0.04920273159717	-3.85317807587376	-1.62751597849715
H	-0.85752161441554	-3.70963945532028	-3.17383399973355
O	3.25216938935806	1.66213179483105	-0.61729143953308
C	3.34105596753012	-0.01633247957198	-2.32640547305653
C	4.05478312408252	-0.95332534518647	-1.56566098872770
H	4.48140337077503	-0.66659032295615	-0.61001965931223
O	3.92119299822855	-3.89425746805147	-3.67022544010739
H	3.54244046907675	-4.02037811905287	-4.54903095544322
C	4.24892278576345	-2.24015505305599	-2.03012565844153
H	4.80617831641486	-2.96916784370430	-1.45135248805778
C	3.72085029471160	-2.61139336130251	-3.26964807982854
C	3.01194541086778	-1.68811584135819	-4.04153839984819
H	2.57609673441794	-1.98167377808033	-4.99345741080854
C	2.83200673954532	-0.39886560803076	-3.57337999180326
H	2.26704885590230	0.31090364921747	-4.16619259438210
C	3.11683678776832	1.33896354657331	-1.79875020010991
H	-1.11972367326873	4.35081786872717	-0.84920424475269
O	-1.07857675356225	4.26799349304298	-1.81559965280299
C	-0.77397512647614	2.62558771928454	-3.48434604822123
H	-0.78812114632910	3.41634791251754	-4.22786174213673
C	-0.93837395685709	2.95614730224899	-2.13493428286722

C	-0.94426161566738	1.95534853237915	-1.16083483135505
H	-1.09717010793981	2.20792312502526	-0.11546703468118
O	2.75268414495011	2.20183463802972	-2.74979729226654
C	2.48461240819920	3.55891423241229	-2.33626243214130
H	1.68020427008720	3.58396321608537	-1.60040340831015
H	3.38808906050959	4.01039451196904	-1.91609506278309
H	2.18155862911023	4.08001960901694	-3.24385896116375
H	3.30723282322993	0.45624695605374	0.78403432378782
O	-1.87690786165807	-2.62716187867581	4.12875587849635
C	-0.20463502671618	-1.90134372465721	2.58456587819053
O	-2.23902127767062	-2.87594744617136	1.91268912574486
C	0.68881615147750	-1.51340839506898	3.59093993638183
H	0.41736517535123	-1.69433324058238	4.62631565430227
O	3.33323801794305	0.06518144614492	1.68305265121170
C	1.880270780808181	-0.88710266802400	3.27334217203953
H	2.57026610480469	-0.56301961672180	4.04610998523015
C	2.19039083408384	-0.61921629853491	1.93477271192868
C	1.32192051882051	-1.03349332124538	0.91893955839157
H	1.56975449747645	-0.84449440622689	-0.12256725736943
O	0.13767024490063	-1.67423772795913	1.24732942876689
H	-0.53511788579800	-1.99224496460371	0.45832659484310
C	-1.49567916738698	-2.49536597617625	2.98193889076116
C	-3.51891269658408	-3.44827102784294	2.23783826794656
H	-3.97339503412564	-3.70834754645848	1.28125706089096
H	-4.13967514821081	-2.72295431887775	2.77213551574238
H	-3.39481865761691	-4.33732447013587	2.86261875901821
O	-2.55449765439269	1.02923175908753	3.49077165403427
C	-3.74613526886984	0.46176422657209	4.07118185618249
H	-4.08570233357019	1.07082437078662	4.91379085023625
H	-3.45603898268035	-0.53600955889010	4.40274446578655
H	-4.54494553154344	0.41374966658550	3.32476260259427
O	-3.71282959796181	2.92987008800056	3.08569158179152
C	-1.40617077546180	2.80644016412730	2.47378573811186
C	-1.37123291300983	4.10772806505318	1.95728075983470
H	-2.27453479097095	4.70907431905977	2.00380242421834
O	2.05778899256412	4.33456488667555	0.77777413669412
H	2.69224950502532	3.61587563787974	0.62635822214900
C	-0.20843752920203	4.61639712582200	1.39492181430973
H	-0.16170251917339	5.63253325771843	1.01159981968889
C	0.94239538062877	3.81802980361652	1.35322093640747
C	0.92471375578097	2.53360217048588	1.89784428920878
H	1.83119509298432	1.93789506457133	1.89441973451019
C	-0.24218663420396	2.03056821142109	2.44678759681631
H	-0.25024639367219	1.02683640455167	2.85711713337961
C	-2.67829636596225	2.29674111223020	3.03601650507945



Main article molecule number: **5-tetramer (Fig. S3)**

Name: BHT

Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

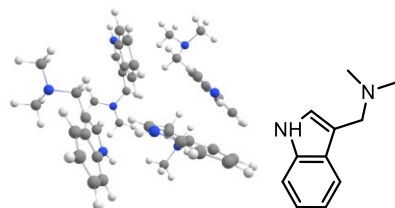
Final single point energy (Eh): -2644.49861619493

C	-6.64017921128843	6.11484631857510	-9.38778233464553
C	-7.03785002807618	4.72975654239040	-8.85387470810526
C	-6.80889586932946	3.64272000621556	-9.92439367725558
H	-7.42918339780533	3.79465053066593	-10.80805324596989
H	-7.04682705452542	2.65899549080674	-9.50191224159547

H	-5.75612501723615	3.62916122587467	-10.22631251262655
C	-8.51458829217383	4.74154939504453	-8.40881757889635
H	-9.19004335099479	4.97285683821538	-9.23401455774088
H	-8.66419743342473	5.48524943706230	-7.61762945986841
H	-8.78543921758224	3.75852722616285	-8.00562810997077
C	-6.19620893944857	4.33331684952436	-7.63230953903629
H	-6.51406350401628	3.34210685389727	-7.29092813135488
H	-6.33427480634088	5.02825792582250	-6.79599536308659
H	-5.12920660795241	4.27124620187904	-7.87136354856471
C	-7.24687740599630	6.67339636945365	-10.53435688124257
O	-8.21712790863282	5.91199584184534	-11.14220912945128
H	-8.58441424788747	6.40882653903107	-11.87900477883964
C	-6.88267470300546	7.93690299114668	-11.04304696315252
C	-7.48600864280706	8.50626181548137	-12.34061856364686
C	-7.13873268590888	7.58769294755628	-13.53530557706542
H	-7.50010730283759	6.55878714338051	-13.43914795387211
H	-6.05146112632415	7.52892992944419	-13.65053433379686
H	-7.56048290450016	8.00164626013961	-14.45829540128280
C	-9.01278569744148	8.72166127165437	-12.22179843437045
H	-9.22538608213854	9.42137078666130	-11.40796882665070
H	-9.60781116486760	7.82096275924524	-12.02911979722734
H	-9.39495458767252	9.15688820486770	-13.15166397248421
C	-6.90456813661832	9.88948448746459	-12.67740796912111
H	-7.13454132905674	10.62520825693186	-11.89949204848213
H	-7.35206414968421	10.24469581481456	-13.61194606766652
H	-5.82025717428377	9.84930207600722	-12.82449126112901
C	-5.91310879579785	8.65169979961646	-10.33502908843201
H	-5.62302067745049	9.63880264483618	-10.67595007865511
C	-5.29922584477283	8.14906896559528	-9.19503428601639
C	-4.27341457875776	8.96020677165489	-8.44903744315634
H	-4.67550421804780	9.33592152303475	-7.49869647190288
H	-3.94879245887434	9.82431671500774	-9.03624675853369
H	-3.38904726250012	8.36254739401617	-8.20478817238834
C	-5.66243926700251	6.87857638738027	-8.75146001628511
H	-5.15480914823009	6.47970814498983	-7.88024045119948
C	-2.32355526802591	4.08426257476777	-7.79928963635146
C	-1.82767927614356	4.74604817839623	-6.50343994175127
C	-2.78034954451903	5.87694239350137	-6.06839475920212
H	-2.83337902644063	6.66917237718644	-6.81521010977386
H	-2.42868687653943	6.31538067449675	-5.12691625343921
H	-3.78963428524466	5.48307272916251	-5.90255073164031
C	-0.40106384065598	5.29947695557601	-6.69998511458042
H	0.28082012105008	4.49491596648092	-6.99700473988111
H	-0.03668291092831	5.72353633558907	-5.75680461228694
H	-0.36639965476526	6.07846663433477	-7.46207851719405
C	-1.76314077903585	3.74398844467845	-5.34007311322161
H	-1.38545000059433	4.26019390738838	-4.45121643391178
H	-1.08431009948200	2.91094809441636	-5.55248255362027
H	-2.74962529057391	3.33809834020092	-5.09059121798122
C	-2.41061150300490	4.79162669504884	-9.01639135268396
O	-2.08863306762666	6.128829933553834	-8.97757843672380
H	-2.39722087357703	6.53953052216832	-9.79264541438734
C	-2.77584381806161	4.16613844018445	-10.22776057668083
C	-2.70406287046928	4.89315674962470	-11.58540263046378
C	-3.67219038816475	6.09448567661650	-11.67191592280764
H	-3.44951768401003	6.94356502077180	-11.01457826432679
H	-4.69620145082216	5.78980921565105	-11.43556002139847
H	-3.65547524265291	6.49873930198534	-12.69048807557326
C	-1.25059597373204	5.34169443219990	-11.85792473715770
H	-0.85084576939242	6.01005674032300	-11.09060761029490
H	-1.19542717674909	5.86136535698179	-12.82166569337666
H	-0.59268506292221	4.46756350554215	-11.90119305595441
C	-3.10033751036139	3.96388154533935	-12.74468841656754
H	-2.45075239574870	3.08502788013263	-12.80442690930952
H	-3.00163816336863	4.51206833470682	-13.68739421392066
H	-4.14035278879904	3.62962995916251	-12.66281646653195
C	-3.14536416150129	2.82109083674776	-10.16141433444327
H	-3.45426033450695	2.30445300695619	-11.06310894357517
C	-3.13480940448855	2.10029790527651	-8.97311410780859
C	-3.60674715629443	0.67198696473579	-8.93458893288918
H	-3.12647829416135	0.11430582111782	-8.12487150237840
H	-3.39597696722486	0.15481908600119	-9.87590041538604
H	-4.69091116375067	0.62166811348072	-8.77030361962964
C	-2.69595236370189	2.73990313401216	-7.81861001159091
H	-2.65720519498788	2.16216349678406	-6.90214173828535
C	-9.54378901182812	12.39064196970959	-10.46029618661168

C	-10.48864317025346	13.10697181770932	-11.43827905629339
C	-11.69974714465844	13.70913074579846	-10.69680465228376
H	-12.30338134804725	12.94864581624144	-10.20097235174140
H	-12.33677602824847	14.24122316660360	-11.41354253276211
H	-11.36071087838414	14.43724343538337	-9.95224915391926
C	-10.96268279362592	12.12135654826067	-12.52615506852435
H	-11.50406722317951	11.27261118824092	-12.10467784070660
H	-10.10350652954078	11.74028851679612	-13.09014882944300
H	-11.62582255538249	12.64008516700370	-13.22857480609596
C	-9.78492487667866	14.27027181452666	-12.15322802092206
H	-10.49867353991421	14.75261460101409	-12.82989536637174
H	-8.93818033356411	13.92730660000712	-12.75876389804373
H	-9.43266057136787	15.03282410886203	-11.45038965664859
C	-9.97869006276849	11.31258289671857	-9.65773537433499
O	-11.29638258229956	10.94477439755029	-9.79491529579002
H	-11.46539169834351	10.18078961357951	-9.23601043741381
C	-9.12105460182661	10.64210402757696	-8.76215062695746
C	-9.61223544960204	9.50094690334424	-7.85212400255019
C	-10.69519671504387	10.02337478607136	-6.87963964880795
H	-11.58369259576053	10.43983661908995	-7.36544802171083
H	-10.27624227976904	10.82050685446776	-6.25682017840051
H	-11.02812776063872	9.21185186964793	-6.22278505008950
C	-10.12456207931419	8.29367192591008	-8.67248559470765
H	-9.30758257853130	7.88827605438993	-9.27668711008773
H	-10.95131896531088	8.50319897457712	-9.36105767952253
H	-10.46699985515719	7.50515009856754	-7.99327410777626
C	-8.47770927848803	8.94447097920345	-6.97624900638002
H	-7.67108632259731	8.51670352138590	-7.58150220422842
H	-8.87752642583995	8.14386495938714	-6.34453751178923
H	-8.06140593629837	9.71138150328801	-6.31472085208482
C	-7.79090762712853	11.06793898075605	-8.71583798511186
H	-7.08949673894461	10.55931552640624	-8.06470946245005
C	-7.31658178058909	12.11952096044829	-9.48915156510463
C	-5.87146516367494	12.53595196210320	-9.42091482930532
H	-5.38306734646702	12.44795596879001	-10.39974781328822
H	-5.31791597263141	11.91444140270578	-8.71177973754859
H	-5.76456074147695	13.58160099712149	-9.11029902501317
C	-8.20832794572386	12.77106595742303	-10.34045617133849
H	-7.83596301642831	13.60379755192840	-10.92643584272539
C	-8.16630283763356	17.19899682410929	-10.14619815367112
C	-6.77983996314184	17.41683555582562	-10.77219505262273
C	-6.22062075157336	16.09388312500445	-11.33108960763964
H	-6.10254736364702	15.34338267675527	-10.54968619574810
H	-5.24065592613249	16.26975128827943	-11.79070345929428
H	-6.88923352495165	15.69380144213849	-12.10206122770010
C	-5.81016706918273	18.00405129860436	-9.72564401271632
H	-6.20771037820732	18.94301354696778	-9.32457902271783
H	-4.84331198532498	18.21777014746136	-10.19643399914771
H	-5.64303223346118	17.31862091989585	-8.89413109266329
C	-6.83258133603596	18.41023766632295	-11.94351387230311
H	-5.81992215190497	18.54262560153948	-12.33851669079979
H	-7.19834882958763	19.39461046354824	-11.63185219362504
H	-7.46005157291256	18.04543180284388	-12.76402773379029
C	-8.35558639603680	16.38088791117733	-9.01282031816362
O	-7.24691835011628	15.72492678770415	-8.52838010349912
H	-7.54787143371463	15.00953276055558	-7.95668007321800
C	-9.60690931543414	16.25644780163790	-8.37328790447829
C	-9.77630468863683	15.51879163502309	-7.02970394496661
C	-9.46261195772687	14.00868722333937	-7.12938186231441
H	-8.41806202157571	13.74249553263191	-7.33488564017620
H	-10.05803786035875	13.53697425878329	-7.91635852603515
H	-9.70539250970746	13.52759779327157	-6.17499592607199
C	-8.89074868161071	16.18998294732614	-5.95620626976532
H	-7.82707226887086	16.18264688115188	-6.20753318689139
H	-9.01636389633877	15.67713019324536	-4.99546506002968
H	-9.18838168911491	17.23590448107030	-5.82826192244611
C	-11.22453314131367	15.60318880155968	-6.51918500094613
H	-11.54693229444262	16.63935605107528	-6.37674625300912
H	-11.28702980761579	15.10006930098933	-5.54878389715940
H	-11.92685137221306	15.10483511586683	-7.19625930934913
C	-10.69248116095198	16.90079089548672	-8.97107265142072
H	-11.67825044003602	16.81493663592172	-8.52838096300068
C	-10.56539873407825	17.66387683075398	-10.12552338172961
C	-11.77143487673356	18.28801426237681	-10.774114171712196
H	-11.51319528362521	19.22144580255628	-11.28373797467312
H	-12.55218491535182	18.50639780237368	-10.03901764167132

H	-12.20699266523664	17.61493842955128	-11.52375305943176
C	-9.29689525554272	17.82280211113546	-10.67413122353622
H	-9.19944179131148	18.44919463887185	-11.55367398580910



Main article molecule number: **6-tetramer (Fig. S3)**

Name: Gramine

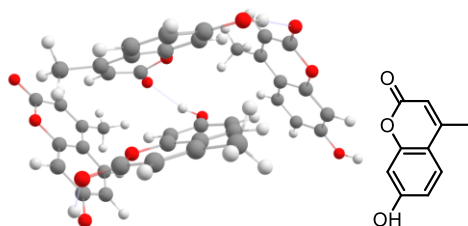
Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

Final single point energy (Eh): -2147.86968350666

H	-0.18832837142830	-6.97608920921248	-0.44879070368852
H	-2.81224591500327	-3.81991420161933	-5.73553422090333
N	0.17221754957231	-3.11608470816739	1.60522918252083
N	-3.03223607497139	2.30684465208807	-3.51174167265199
N	-0.71572541112898	1.06315698306302	3.74598403305834
N	2.42239271084893	-2.66978383858388	-1.73698495132613
N	3.03061240656733	0.51821747396324	1.83676589646774
N	1.42765134320265	1.35607528218982	-3.26129868727717
N	-2.71439317422310	4.19288693983983	0.98564781929301
N	-1.24537563836560	-3.42866730275829	-4.39118722280471
C	0.58165481272960	-2.01529592393975	0.88107651630357
C	-1.76852566185816	2.83882456408489	-3.67474612699376
C	-1.54121184966845	1.16755308157501	2.64446104007755
C	1.46816593305269	-2.01996596599636	-2.49680457990581
C	1.18065509370434	-1.10040403787783	1.70597959290922
C	-0.87612563908226	2.17290958097654	-2.87453056083146
C	-1.39062349318885	2.39756635833487	2.05687363233175
C	0.35248818426162	-2.80912714948669	-2.61618646571151
C	1.64023903578545	-1.24713763395849	4.27412024489814
C	-1.29553660160737	0.23316889900119	-1.18912539479263
C	0.19614168320849	4.36162211967253	2.71966189104573
C	-0.15015220077079	-5.13784665739015	-1.55168733201074
C	1.46214131025824	-2.07536143872809	5.37338382010682
C	-2.29735617977474	-0.56293171247789	-0.65903942379396
C	1.18059619685613	4.71305750769646	3.63103500513669
C	0.38943927734434	-6.09084229262555	-0.69889835915187
C	0.84015606593753	-3.33051899592377	5.24772176198646
C	-3.62936494154849	-0.44068696256024	-1.09314765347183
C	1.58149718430974	3.82959410065950	4.64963343747394
C	1.68264833390182	-5.94253619300800	-0.16262018123500
C	0.37770898995466	-3.78467690995712	4.02042737796456
C	-3.99076022229618	0.48835227482166	-2.05902208599848
C	1.00545711191449	2.57435904928807	4.78283541017891
C	2.46429130355458	-4.83649651308731	-0.46376090451240
C	0.54602595800049	-2.94775085838328	2.91952154401802
C	-2.97900167394363	1.29047143564482	-2.58310841925378
C	0.02005639531271	2.22092349602189	3.86293483861458
C	1.91554943220956	-3.87706070706683	-1.31097428662559
C	1.17353130031241	-1.67569790159071	3.02252463704584
C	-1.62821891725467	1.17607496399888	-2.16683616327522
C	-0.39358334271070	3.09405943808302	2.82038040061951
C	0.61208987924121	-4.00335569339554	-1.86294745697190
C	1.69284403326837	0.24483174393377	1.30711010791567
C	0.58383329093888	2.44877384166746	-2.74580075584591
C	-2.08176050124086	2.87694005706394	0.82551140057678
C	-0.90027548659842	-2.45692444442753	-3.34500795524855
C	3.38277869795905	1.91770276006495	1.63132204936846

C	2.83328601939593	1.64453044785300	-2.99071108769287
C	-3.28183059850325	4.61431773785286	-0.28942396098749
C	-2.53376075470016	-3.07870957330073	-4.97837495478563
C	4.01967377043246	-0.36626163937335	1.24162492694370
C	1.21415964746316	1.15089480793966	-4.69162778270704
C	-3.74527337715424	4.16005491164507	2.01808698295193
C	-0.21262702943157	-3.48618546767655	-5.41993356651856
H	-3.86697590445525	2.64966242030814	-3.95407673025637
H	-0.49382840216231	0.19676615616111	4.21353064039266
H	0.39745643795012	-1.96441129425291	-0.18314257475816
H	-1.59733820853685	3.66099668830864	-4.35722864091303
H	-2.17377138720070	0.34205594358569	2.34675563221300
H	1.62395258613535	-1.00867245246547	-2.85666282717567
H	2.15082056020910	-0.29195330787029	4.35752631585317
H	-0.27027579838658	0.14242811973539	-0.84690037715533
H	-0.12802871156178	5.04951466937027	1.94463064021127
H	-1.14034110997105	-5.26161576478266	-1.98029965424049
H	1.82221738122939	-1.75858945946957	6.34812843588975
H	-2.04902670018452	-1.29421949125413	0.10525293138071
H	1.64899937427975	5.69090445999906	3.56508124810582
H	0.72375472555758	-3.95850953012580	6.12601396330074
H	-4.39241477184937	-1.08044609737650	-0.65961121877332
H	2.35382763419776	4.13687424052274	5.34869846397993
H	2.07577737840732	-6.70722705333517	0.50066744884419
H	-5.02278522980535	0.58881179209414	-2.38359037950574
H	1.30844002072005	1.89578265210627	5.57513839093471
H	1.02270921535987	1.03096527735110	1.68056536932349
H	0.81986367783203	3.40755850501261	-3.25441213883001
H	-1.35217132854758	2.96952161847786	0.00717104423072
H	-1.74211281551376	-2.42204543555396	-2.63873160424856
H	1.68083077279221	0.32314242164894	0.19478767116881
H	0.83991878204544	2.58171496171673	-1.68399434109700
H	-2.81499793819019	2.10953624646340	0.50012414982746
H	-0.79889717854334	-1.42899376789866	-3.75490284216778
H	0.16643800505326	0.89643152575622	-4.87881426160057
H	-3.30254226555912	3.87026519673460	2.97597330743516
H	1.83750748715100	0.31998117892723	-5.04173203742824
H	-4.18682678069035	5.15665763423531	2.12815824353210
H	4.10990665123304	-0.21592520195893	0.14069747502236
H	1.46653660194452	2.04806898300315	-5.29305597939823
H	-4.55945994144487	3.44313289383063	1.78282547009600
H	-0.06430649696711	-2.51039186977832	-5.92911429255310
H	4.36586590542064	2.12254210902306	2.06949510085939
H	3.45810616806316	0.81168990297357	-3.33523787631910
H	-3.71023657987223	5.61866050151312	-0.19548423606798
H	2.64682813157853	2.56394406671489	2.12201050207474
H	2.98772537804555	1.76527356723090	-1.91229931332439
H	-2.49684497375517	4.64207716754450	-1.05488231105765
H	-3.30737326206312	-3.07226879869393	-4.20203904744705
H	3.42483561794263	2.19336974334245	0.55543115701414
H	3.18859350437437	2.56526703868186	-3.49551493856298
H	-4.07984800598329	3.92986881555041	-0.64912643454403
H	-2.52809681139872	-2.07852591415320	-5.46013774770574
H	-0.11308627414429	-3.98697319118584	1.18233049935979
H	3.25291454690551	-2.23460792212158	-1.37132025360689
H	-0.09918307920450	-4.75627863572099	3.92389569243539
H	3.45632529517464	-4.71683723399898	-0.03743954828366
H	3.74267043274088	-1.41011988895346	1.42703589052177
H	0.74094370848496	-3.78499837982525	-4.97389122583685
H	5.00360995929468	-0.18360435319827	1.68712451333662
H	-0.48815487917773	-4.22875903040103	-6.17652699731964



Main article molecule number: 7-tetramer (Fig. S3)

Name: Hymecromone

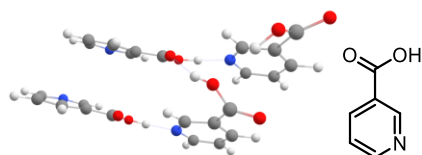
Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

Final single point energy (Eh): -2445.94062319554

H	-3.40015681972073	0.54514758614566	-1.63500877188130
H	-2.81038777850878	0.79657626720604	-4.05107473542904
H	-2.89067716958246	-0.58976564838418	-5.16088485626715
O	0.30921691547521	-2.09110876178714	-1.92672105667610
O	0.94823160498191	-3.22785393940469	-3.71999204928346
O	0.95476253678851	-4.98276863590051	-1.49454788160366
H	1.17592764665139	-4.61572050762967	-2.37352067822172
C	0.14864827682879	-2.42354380146493	-3.26034451767366
C	-0.92762530629654	-1.79035350877299	-3.96205253188562
H	-1.00230510674781	-2.01999750022682	-5.01914147173357
C	-1.81260379397577	-0.95811961037697	-3.34328859766427
C	-1.66472978569689	-0.70635710139884	-1.93639504470097
C	-2.53043037908869	0.09405626294932	-1.16737553907951
C	-2.27437433845126	0.35215892186631	0.16187941439241
H	0.29087210373950	-4.78401861254448	0.97437839730598
C	-1.12226637354838	-0.17245533771440	0.77196480778355
C	-0.28969752133169	-1.03679829721416	0.06531843544729
C	-0.56302904223171	-1.27438727336364	-1.27331086530800
C	-2.91180020500054	-0.29421735732497	-4.11025977222311
H	-3.89283063076697	-0.55389140366069	-3.69649373299050
O	1.17792430654322	1.68690319947795	0.12028723204466
O	1.98075229989981	0.56856757492890	1.86236644141840
O	-1.50394158953107	3.24144755617884	-3.45026783172879
H	-1.97109701606926	3.67395727163917	-2.70641780335304
C	2.07570254463933	0.77801335715267	0.65747446651640
C	2.99710666303722	0.16442654240438	-0.24718988723296
H	3.77657018902607	-0.43419027024221	0.20765497615729
C	2.84658268332278	0.24233301073038	-1.60168585182451
C	1.76002494044176	1.02029923780099	-2.12949857834720
C	1.41514852733954	1.10011686107525	-3.49288906890637
H	1.99965091936980	0.54593468578238	-4.22043340378896
C	0.34023139269276	1.84833239739612	-3.92298687845992
H	0.06968799992566	1.89301636909472	-4.97283068956427
C	-0.43636279150926	2.55898514154515	-2.99056450736150
C	-0.11078228007153	2.51546495329366	-1.63803980881441
H	0.57413137229064	-1.50731049595092	0.52536883818060
H	-0.71427173513815	3.01834564737770	-0.89489773218278
C	0.96059572006772	1.74389055181372	-1.22811888484792
C	3.75348036155158	-0.50846614835690	-2.52219531790826
H	4.17855078665924	0.15815222976950	-3.28065897618748
H	3.18941858925401	-1.28883847533237	-3.05003826533780
H	4.56769921139587	-0.98496963222936	-1.97447626472632
H	-6.37056294758323	2.40857865826606	1.17046254694816
H	-7.56616642042209	1.32865051293173	2.85398761742173
O	4.95622917752440	-2.78599097025933	-0.23723042049938
O	6.85077241221447	-1.71511968607999	0.24330015495551
O	-0.85487522847196	0.20258913998211	2.04448789764089
C	5.79879159431999	-2.12606012350029	0.67377530147026
C	5.29504817104756	-1.98749672929973	2.01554041757994
H	5.94114500041192	-1.44799083870568	2.69938223703745
C	4.08548784962502	-2.47136058626789	2.40728735016243
C	3.26869127379334	-3.16922310251133	1.45002585716775
C	2.01427913277543	-3.74047958418512	1.73492062097772
H	1.62106139665939	-3.68297750556496	2.74506996728698
C	1.26449965877678	-4.35816143945788	0.75622486235711
H	-2.93114642344887	0.98291567413776	0.74997765122664
C	1.74286294659412	-4.40607541226019	-0.56385260107687
C	2.99056755948560	-3.87216206783148	-0.87771481164942
H	3.38227060960547	-3.91347112458932	-1.88893891818214
C	3.73784421808219	-3.27646945690283	0.12712517004345
C	3.58672083725464	-2.25580584204404	3.80159589353314
H	3.35099223677475	-3.20805493844029	4.28985196551594
H	2.67070171489392	-1.65389355138544	3.78064807795613
H	4.32953284227801	-1.73006522309153	4.40427944420304
O	-4.22163139743133	3.46589923087900	0.33766204707684

O	-2.98318924314516	4.33458853379495	-1.28600477981376
C	-3.03255810658093	4.02954449268646	-0.10525281666530
O	-6.91936403856041	1.27677387713747	3.56800387924201
C	-1.98416302636243	4.17591379555639	0.86136193908603
H	-1.09084361099151	4.68724954642521	0.51999872416998
C	-2.06395652134732	3.65353973425702	2.11741189273120
C	-3.28782828847420	3.01678290295353	2.52113066500160
C	-3.50010776920161	2.42594919492168	3.78092486043764
H	-2.69225733650478	2.41544429248121	4.50447764155023
C	-4.70719842921855	1.85312354348584	4.11273786323355
H	-4.86979284961776	1.39604840588861	5.08256173912465
C	-5.75422377144898	1.85219669598552	3.17813252851058
C	-5.58147582008324	2.41149961867767	1.91705863113858
C	-4.35285520372005	2.97864260581985	1.60440811120175
C	-0.89898279120359	3.71662962334972	3.05021003309862
H	-0.54696282203579	2.70149057385930	3.26815945722318
H	-1.17471551028802	4.18763885555213	4.00016919633120
H	-0.07183567176651	4.27358635174571	2.60715110750192
H	0.11582166713647	0.18349201525435	2.18515984169055



Main article molecule number: **8-tetramer (Fig. S3)**

Name: Nicotinic acid

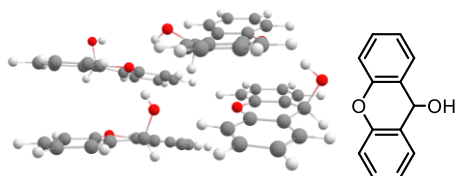
Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1747.20228741639

C	2.08604908674806	-0.16972819730039	-0.42357539944725
N	3.02392817160574	0.66572832365174	0.03664682670201
C	4.15186904256734	0.15505657874037	0.54063498391872
H	4.87785822672570	0.87162875843367	0.92076173270209
C	4.40750555884967	-1.21168960467991	0.57639021808388
H	5.34886503075044	-1.58073229789578	0.97042778556968
C	3.43727306734758	-2.07983422525584	0.10169851864608
H	3.57675160159891	-3.15680371217473	0.10420886814043
C	2.23906401456604	-1.55464829083813	-0.38394678738695
C	1.18419112108381	-2.50699199243203	-0.83817649664833
O	1.43416952239089	-3.64725957275473	-1.15439667226317
O	-0.09024677208545	-2.06298778052759	-0.85553189236326
H	-0.23772020044841	-1.23110705249151	-0.33310378088114
H	1.21002209680369	0.31255115354722	-0.85464831347846
C	0.71279747668751	4.80365535998595	-0.83708779138063
C	1.07194906145850	5.78586828002614	0.08994324157846
H	1.784731444103730	5.55501374426472	0.87815964200647
N	0.56469464333318	7.02130708320577	0.08190122318581
C	-0.32185988412609	7.31675767698716	-0.87443449773216
H	-0.71652449483893	8.33276931571320	-0.86380605439435
C	-0.74181329293737	6.41095623691880	-1.84609996281542
H	-1.46359506692712	6.71050838202128	-2.60003170816336
C	1.24807844153569	3.41767788051793	-0.78095916720478
O	2.25368825831083	3.26991098132141	0.07937295499517
H	2.56461832753219	2.30553776715070	0.05055997585395
O	0.78947044654810	2.51105583232186	-1.45944485679254
C	-0.21846758209377	5.12829403170017	-1.82243176238237
H	-0.51265656315449	4.36973659504706	-2.54176218582421
C	-0.17059317542353	-2.90961542129358	2.00773703503657
H	-1.07216608753344	-2.41985225992443	1.64514161102113
N	0.7004428527571	-2.12780595368303	2.65681596629585
C	1.81805279405305	-2.67794364463558	3.14508111642814
H	2.50739705890857	-2.00362647150927	3.64965711519483

C	2.09767916407257	-4.03356164888755	3.02078771215267
C	1.19796058737125	-4.84217249839252	2.34130365579679
H	1.36733046762484	-5.90710081797559	2.21316949178681
H	3.00898869270479	-4.44199404209312	3.44536905818603
C	0.04855370422794	-4.26815654794508	1.80071764352684
C	-0.92783041460601	-5.12934975743413	1.05563881631064
O	-1.56170941473870	-4.55954294479102	0.01587287067786
H	-1.11220831142073	-3.74584243850006	-0.28213504393025
O	-1.14993945601507	-6.26633573736150	1.39033860967291
C	-0.52827107390030	0.81037258021994	1.52411345018798
O	-0.88231840339310	0.01907747151929	0.64430823997013
O	0.25560674049457	0.48754118143190	2.53340342957447
H	0.44174821516505	-0.52636585157031	2.55259221658751
C	-0.98701656109567	2.21809042975906	1.53381683991412
C	-1.90410182727570	2.65077459619531	0.57514583096391
H	-2.24808573258714	1.95790479395595	-0.18590683070560
C	-2.34422135751557	3.96293175142316	0.62728204279643
H	-3.05770872486995	4.34447325571989	-0.09615791551247
C	-1.84017568901251	4.79713871546396	1.62250658021810
H	-2.15855832987157	5.83745676665419	1.68186680171967
N	-0.94975100162502	4.40266267370855	2.53966592745393
C	-0.53960539626690	3.13479024121537	2.49161745546892
H	0.18179989898281	2.82498032352557	3.24427163098154



Main article molecule number: **9-tetramer (Fig. S3)**

Name: Xanthhydrol

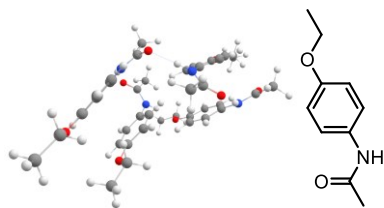
Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

Final single point energy (Eh): -2606.95219290775

C	-2.59641425692356	-7.90307608244394	2.80959184314864
C	-1.24339602276824	-7.62399473114304	2.91365538316070
C	-0.36375810463026	-8.64323709022311	3.26317235243801
C	-0.80367423788139	-9.94016442967315	3.50601422954516
C	-2.16857743373156	-10.20309367582202	3.35978434424925
C	-3.06502165687405	-9.19875565203448	3.02416758211918
C	0.13725110617374	-10.99758919686118	4.00722995536729
C	1.56861873575521	-10.60065069739054	3.81187582192401
C	1.91832427102364	-9.28141753704208	3.55146723767407
C	3.24569606624167	-8.87235180353452	3.47218419574708
H	3.47600989494462	-7.83131943427172	3.26547669935867
C	4.24830287603885	-9.81103156553367	3.65552799970445
C	3.92757292643806	-11.14415058819861	3.91022358666415
C	2.59748004895377	-11.52765738552631	3.98699395350128
H	-3.28567552258983	-7.10856181523951	2.54018037139225
H	-0.85723432161216	-6.62803455284122	2.71960595383804
H	-2.52541158587013	-11.21936437032122	3.52046431749591
H	-4.12201870827119	-9.42238692254298	2.92075551420439
H	5.28615651892162	-9.50013202739538	3.58909469126092
H	4.71513503750951	-11.87764152351343	4.05083326581605
H	2.33282064628857	-12.55855981630365	4.20773245708009
O	0.96927771429654	-8.29110753279854	3.34320229398582
O	-0.02981504588850	-11.20071838938207	5.43046812940516
H	-0.97625265611985	-11.27394583868903	5.60735736444580
H	-0.06251867308968	-11.94832559660340	3.48444203658734
C	4.39291880214412	-10.79276092719839	0.26325985771526
C	3.02056940439351	-10.95303947505163	0.35135279431346

C	2.19247489446361	-9.84496240096821	0.18236830479683
C	2.71580772579707	-8.57598683783193	-0.06121656945586
C	4.10385244927231	-8.44237574511123	-0.15063699253349
C	4.94196828828617	-9.53502500184865	-0.00553020542266
C	1.82044832030018	-7.37275707324437	-0.13415734944417
C	0.38528958933585	-7.77418944014115	-0.32039029818594
C	-0.02870019793755	-9.08511600913765	-0.08683085003532
C	-1.36664220520198	-9.45678664183140	-0.20355985225584
H	-1.64541145150231	-10.48750055323605	-0.00962057332149
C	-2.30610841705675	-8.50238105579011	-0.55633278522382
C	-1.91808711440137	-7.18202653173658	-0.79339485386966
C	-0.58252299889816	-6.82922309274303	-0.67347305837576
H	5.04056146407599	-11.65444785157513	0.39544419584150
H	2.57081129745643	-11.92017833961395	0.55113605085302
H	4.51677066123515	-7.45259573675018	-0.33090737821331
H	6.01796064460358	-9.41330722250754	-0.07037635165399
H	-3.34980101189730	-8.78922895921371	-0.64695040957157
H	-2.65488768872287	-6.43388433898615	-1.06945008818973
H	-0.27003859446944	-5.80051773980138	-0.83677518547885
O	0.84296221942625	-10.08668278866520	0.26708432902707
O	1.98124765928047	-6.54420114842190	1.04274041910311
H	1.69679489624030	-7.08203376444604	1.79724319992289
H	2.13436291966083	-6.71681486642927	-0.95651744294713
C	-2.61188967200425	-2.78664352435432	4.99858035327122
C	-1.694883355495439	-2.01328015622914	4.30309880946937
C	-0.49380625279240	-2.58623722021195	3.89814173471092
C	-0.18683033434863	-3.91893892417387	4.16254473389734
C	-1.10984065351302	-4.66478880604710	4.89547224671870
C	-2.31570311522840	-4.11268849854966	5.30835533712668
C	1.06971210584715	-4.53869630606783	3.59510537831923
C	2.08489321493107	-3.47451033223903	3.27388068231157
C	1.68518162786674	-2.15739701718981	3.06658362798893
C	2.58398459470826	-1.16921465352783	2.68016132341598
H	2.22526123739277	-0.15693730397684	2.52132595904502
C	3.91639916095104	-1.50508246094822	2.49416286583535
C	4.34471118970582	-2.81446365200343	2.70646380849739
C	3.43196749429321	-3.78603973005467	3.09365026145521
H	-3.55485455534600	-2.34717065877843	5.30959483563739
H	-1.89453529201933	-0.97324055839069	4.06382282841575
H	-0.87035450342694	-5.69380001138591	5.14763026102423
H	-3.02171009634780	-4.71433290828942	5.87248365171315
H	4.62166991243641	-0.74002539490146	2.18373090769565
H	5.38900310058451	-3.07632887589694	2.56641116185812
H	3.74674965420410	-4.81092389429922	3.26041491098611
O	0.36889906733335	-1.75897460519777	3.20718601610848
O	1.65183876748460	-5.50515263343252	4.47298067351774
H	1.31984879372221	-6.38045856871550	4.23687107474220
H	0.80655686451274	-5.02773786671942	2.63980177526874
C	2.81672749673242	-2.50691969344728	-0.80869629784497
C	1.84745155117400	-3.28654638341556	-0.20026798634255
C	0.60062350174435	-2.72540444415861	0.07212585992347
C	0.31129019704822	-1.39631467524676	-0.22669124328114
C	1.30680313503712	-0.63506787290440	-0.84618578649134
C	2.54639262505344	-1.17857244509347	-1.14465673863281
C	-0.98520278103406	-0.76299772773782	0.19977376232883
C	-1.98138821434731	-1.80514026320077	0.63056179790289
C	-1.59300892453239	-3.12161364158348	0.86878631636820
C	-2.49520892599676	-4.07389936757706	1.34182050219557
H	-2.14553763186286	-5.08893531894666	1.49995893038420
C	-3.80552781880319	-3.70398914564299	1.59142559842031
C	-4.22081297435576	-2.39070607958306	1.36127309483761
C	-3.31163268725277	-1.45937170690212	0.88444765623522
H	3.79424713101875	-2.93413993508137	-1.01208324031434
H	2.03916585535565	-4.31081492766974	0.10570352874515
H	1.09374249364695	0.40732742129386	-1.07081846823439
H	3.30675214462130	-0.57033765641907	-1.62473784958155
H	-4.50846460133485	-4.44033419485842	1.97029893644433
H	-5.24849308108914	-2.10017527651792	1.55560678063617
H	-3.61644302358999	-0.42903410313662	0.71720589149634
O	-0.31116720646941	-3.56734590109285	0.66732352600748
O	-0.76771150172965	0.20594287639332	1.24478949686894
H	-0.38323159608668	-0.28353596221570	1.98890556977434
H	-1.40155309116006	-0.16954721606078	-0.62480575523131



Main article molecule number: **10-tetramer (Fig. S3)**

Name: Phenacetin

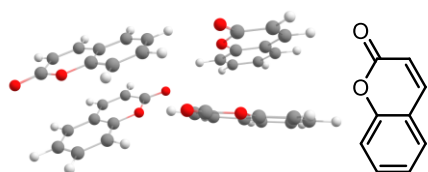
Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

Final single point energy (Eh): -2375.99581495721

C	-8.46280223048587	-4.29939387241979	0.55671470380065
C	-9.34319384904881	-4.35626397404574	1.64597457327447
C	-7.10990460924323	-4.10155437325201	0.75099016158775
H	-8.85666752569692	-4.38939776642934	-0.45288417158600
C	-8.834733592786736	-4.207167024110877	2.93596478940328
N	-10.69981830956513	-4.60454556317742	1.35850235495026
C	-6.59860387903365	-3.95570971201344	2.04384696515711
H	-6.42908455971076	-4.04576143159942	-0.09233872818007
C	-7.46863704839124	-4.00562409001812	3.13268204676560
H	-9.49961310700816	-4.25394351435895	3.78740849837345
C	-11.76930627338321	-4.49902536485619	2.18792703985869
H	-10.91699021565816	-4.73874297099427	0.37328231828283
O	-5.24762952993204	-3.76382572760992	2.13415358443436
H	-7.10535919615034	-3.89869794966299	4.14839612927564
C	-13.0747769937782	-4.99160765450581	1.60642504589341
O	-11.70266834670417	-4.08252676937708	3.35192330303721
C	-4.67282856870896	-3.68817846593644	3.44153157560480
H	-13.89299788888182	-4.36309213345244	1.96460453421253
H	-13.07086837682134	-5.02807296125448	0.51380969656015
H	-13.24592186354779	-6.00566904635643	1.98621099995741
C	-3.17303769284790	-3.58369977442923	3.27576200908901
H	-5.07307064199988	-2.81047261490682	3.97420968292855
H	-4.93998965670949	-4.58739011631972	4.01722543226357
H	-2.90365771317200	-2.68929023674998	2.70691977005874
H	-2.69411425680423	-3.52608157046005	4.25807393582333
H	-2.78310227495174	-4.45928641680557	2.74918484686588
C	-5.06686154353109	-0.16883824476250	2.07317635371195
C	-6.55219321789063	-0.37985045588142	1.88775181458290
H	-4.51979095256705	-1.05174147136965	1.73144158770640
H	-4.84295580699850	-0.00201475618473	3.13173946199929
H	-4.72281589923581	0.70011656143071	1.50483992908452
O	-6.76999153424730	-0.62342100003944	0.49167656061467
H	-6.90526221037251	-1.24595544782524	2.46735680925146
H	-7.12324155211122	0.50760158690891	2.20099408564974
C	-8.03074152640429	-0.91685419609105	0.08337264329421
C	-9.14871741760954	-0.94442125682422	0.920012074444660
C	-8.17156557503749	-1.21547964678890	-1.27698448692358
C	-10.39190359965782	-1.27632089149697	0.39095023729684
H	-9.06704852528367	-0.73611382686489	1.98005326247845
C	-9.41322263712058	-1.54444215176566	-1.79083657640366
H	-7.28951060016435	-1.19137090412300	-1.90901642486468
C	-10.53245390458469	-1.57636314118714	-0.95829000830172
H	-11.26191194691728	-1.29304484981975	1.04122986973274
H	-9.52545611341047	-1.78093816377857	-2.84502080026343
N	-11.82339137824151	-1.87204120264253	-1.50006575708939
C	-12.29076884904851	-3.13159396932191	-1.67825698222712
H	-12.39999833299488	-1.09356315640781	-1.80686770257754
C	-13.68738972962785	-3.22739493046222	-2.25314724173368
O	-11.65128330880547	-4.14603670616760	-1.39051855977558
H	-13.65253358204312	-3.82726421655225	-3.16691389329712
H	-14.13732279491629	-2.25420556927883	-2.46918380322756
H	-14.31998429183730	-3.76151254079956	-1.53716459152977
C	-13.04979947548747	1.63467024493939	0.97438489893417
C	-14.35922959131813	2.00068942415149	1.32016349186604
C	-12.74505464264869	1.22823890942037	-0.30932301916228

H	-12.27062133050900	1.65787646697378	1.73186772767236
C	-15.34435395010125	1.99714471557648	0.33103709546084
N	-14.58772089745495	2.30631523031263	2.67163766659314
C	-13.74992010186482	1.15703715053572	-1.27443362896971
H	-11.72775572982315	0.95497692684968	-0.57301008355003
C	-15.04141745546847	1.56720012649081	-0.96052653935748
H	-16.35177244377419	2.30158251107226	0.57874972985396
C	-15.76027394724420	2.67797376892032	3.26420456866898
H	-13.86519683596981	1.95173047891923	3.29606711756863
O	-13.35736176662063	0.64716936215644	-2.49577431816495
H	-15.83740437917508	1.54482030417595	-1.69604753735096
C	-15.68758488222848	2.76629959558690	4.77593638623035
O	-16.79617808963418	2.94292119642855	2.65967843446759
C	-14.31756360004735	0.64099211732834	-3.55843340779777
H	-16.53431463876088	2.21514702936175	5.19480074009390
H	-14.75409492861363	2.38090226795269	5.19522767688921
H	-15.80792887915745	3.81427364159832	5.06863534949404
C	-13.63061023625217	0.08188378158045	-4.78466336543260
H	-15.18569234357183	0.02529971870089	-3.27490745048688
H	-14.67217717116876	1.66693590455864	-3.73563481302982
H	-13.27134740877804	-0.93588355963903	-4.60092649934509
H	-14.33191042337842	0.05273939886536	-5.62399376646321
H	-12.77606619963989	0.70377827355041	-5.06563169086234
C	-19.77747159185600	0.38681076437676	0.35527512559916
C	-18.75295837216353	0.12855874061432	1.43664467049793
H	-19.33731368192174	0.96875308413573	-0.45990132410798
H	-20.61617884826549	0.95449375940445	0.76964208177505
H	-20.16043955266332	-0.55296902026705	-0.05350541896496
O	-17.68765019425228	-0.63151260291450	0.83938563673517
H	-18.35077785497053	1.06970670802759	1.83726923124412
H	-19.19400704580405	-0.44793737091601	2.26408735428009
C	-16.57171400278109	-0.84705459241079	1.58483185598156
C	-16.44299528968892	-0.50892471274707	2.93386612296388
C	-15.49535669625445	-1.43804930598128	0.91423841771166
C	-15.23030272120826	-0.71540324951172	3.58370183626415
H	-17.26614540534463	-0.06396396990389	3.47973929559411
C	-14.28935156950775	-1.62120437130985	1.56505342021572
H	-15.61198841021594	-1.69294955152756	-0.13423706188166
C	-14.14143412340390	-1.24644996821223	2.90338537145273
H	-15.12363077297221	-0.44061228607494	4.62847070091334
H	-13.44235113995727	-2.03137810025231	1.02118292447192
N	-12.88433741556944	-1.41724945584728	3.54609383095650
C	-12.20308060574670	-0.42562326559347	4.17478461491556
H	-12.46623245787268	-2.34863717089186	3.52700126496250
C	-10.89198441860913	-0.85439193227732	4.79731202205880
O	-12.59421024942108	0.74258443367142	4.25884546295933
H	-10.13883783943508	-0.09408713011302	4.57512032961035
H	-10.55286521745306	-1.83603385148523	4.45395017532319
H	-11.01520580759064	-0.88213492516222	5.88517833531889



Main article molecule number: **11-tetramer (Fig. S3)**

Name: Coumarin

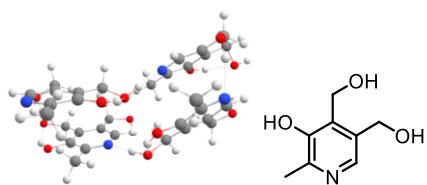
Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1987.78320779229

O	-0.52649978190481	-0.36654105560700	-2.32504607663384
O	-0.38017436948922	-2.19478362112380	-1.06633508072037
H	-1.09725583547388	2.12008902286366	-2.65211288538839

O	2.77589416248145	1.98924507241058	-3.76788130790817
O	3.30796897889251	2.08055757279055	-5.93180252188945
C	0.10476500543191	-1.14779021337929	3.24838053162052
C	-2.53240396189601	1.45482270404537	3.00544198185091
C	-0.12135112930784	-0.15037143552916	5.44866184400064
C	-2.11650605170932	0.09940933866618	1.03284712159611
C	-0.10772052611698	-2.15617776599420	2.31438887011444
C	-3.07736739384770	1.66676886602462	4.26800776230870
C	-0.36248676831252	-1.25372636575437	4.56809651344271
C	-2.74766459335310	0.25246001874719	2.31101626524080
C	-1.04965856188414	-2.41923951875119	4.94135686127675
C	-3.53342228029394	-0.73999882007977	2.91636238995546
C	1.04462797736089	1.04922395294888	3.66715026276647
C	-1.19188440903409	2.35805249635416	1.20526033544334
C	0.54543525409854	0.94413365824545	5.02270389595765
C	-1.36555800086029	1.09504239162387	0.51935292891905
C	-0.79484694321692	-3.29349411030613	2.71140805695164
C	-3.84715604962776	0.66691696263330	4.84427209906422
C	-1.26283324904434	-3.43152571013057	4.02362686312664
C	-4.07866494257265	-0.53728366280661	4.17106544095073
H	0.73947275996990	1.79805715169772	5.66138298913958
H	-0.84979746498931	0.99172660463297	-0.42592078940683
H	-0.97008985765604	-4.08199303905051	1.98531152828268
H	-4.27737729973875	0.82770953711782	5.82857017308603
H	-1.79647817224214	-4.32968627529846	4.31806575777632
H	-4.68129815016851	-1.31193782233365	4.63455084383129
H	1.84538344270126	3.02665026170160	-1.60357405571651
H	-0.49198007154340	-0.20761610200019	6.46982452066473
H	-2.22260113393789	-0.83743303328707	0.49049101868118
H	-1.41849843275746	-2.50678549873040	5.96004891042752
H	-3.68942008416146	-1.67607830932752	2.38716663001463
C	-0.36578606211484	0.29471886891478	-3.51126296442060
C	0.50097443944355	-1.74099178247596	-4.50072027643542
C	-0.72020458843723	1.63775559314437	-3.54829821163827
C	0.14354354420068	-0.36331565694886	-4.64159825642623
C	0.29695259645153	0.36956035383203	-5.83027705770879
C	-0.18657847469867	-1.70356056430826	-2.15843734486051
C	0.35188563499909	-2.37940453314260	-3.31813472372545
C	-0.55175277083404	2.33634305762988	-4.73411621284021
C	-0.04456034817383	1.70702294835918	-5.87618931117455
H	0.61533354377830	-3.42143981612780	-3.17783297429080
H	-0.80510884520054	3.39131201889046	-4.76805729372430
H	0.10665054831473	2.27539268529634	-6.78711407590896
H	0.90349837386102	-2.26551718472052	-5.36460971832378
H	0.71089053123385	-0.12575202677902	-6.70459784485584
O	0.77692937484176	-0.03909190330764	2.82647904558383
O	-1.78063446215130	2.46023150068974	2.47329177341446
O	1.66111554593926	1.97201038346254	3.19371417902523
O	-0.59423583915280	3.32071506822731	0.78877973878978
C	2.68886272240977	1.31295437856199	-2.58404316490694
C	3.66669424232699	-0.63854365873608	-3.64335765120076
C	2.16596382690408	1.99610398258243	-1.49082113051423
C	3.11504574717725	-0.02217832486660	-2.47672311291800
C	2.99286246685414	-0.66705179296916	-1.23401407468129
C	3.29149005622926	1.40900129948188	-4.92715757539090
C	3.75747256773854	0.04119746242234	-4.80760398003690
C	2.07143446875303	1.34072709656193	-0.27162131366365
C	2.47897315986018	0.00703707033418	-0.14324380504221
H	4.17649503705383	-0.39245298789831	-5.70858857513693
H	1.68516017602755	1.87087326849296	0.59387641058327
H	2.39306920106005	-0.48442494187134	0.81940091499270
H	0.23333861713023	-2.02772631688853	1.29191771956488
H	-2.88718010133113	2.60759274875760	4.77418737894086
H	4.01596221134924	-1.66695194345274	-3.57419606572852
H	3.31579825179622	-1.70175602998488	-1.14615293164734



Main article molecule number: **12-tetramer (Fig. S3)**

Name: Pyrixidine

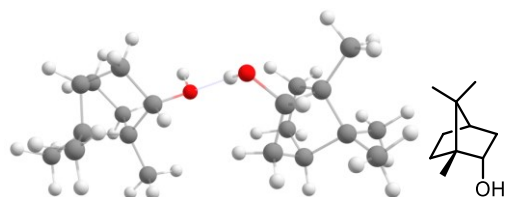
Qty of copies of the molecule modelled: 4

Qty of imaginary frequencies: 0

Final single point energy (Eh): -2367.20657566562

N	-1.17945630653381	1.89435185790149	-4.51262055297189
C	0.78006123610862	1.50799813520581	-2.54876054599901
C	-0.28169030954537	2.39775920660253	-2.35366114885554
C	-1.25844835054274	2.55293441283719	-3.35626413438995
C	-0.13284078426185	1.07967063360606	-4.72753061302894
C	0.85558669845152	0.84268667785810	-3.78504083843757
C	1.77415600721939	1.18742076575421	-1.45836539910358
C	1.97724949862428	-0.13628916109015	-4.06559053681466
C	-2.43092959204328	3.45130084965158	-3.11085812041065
O	-0.46134453820500	3.11199336238227	-1.22201936488371
O	1.83351258873699	2.19278777560951	-0.44806600306789
O	1.92126381394648	-0.74534743449902	-5.34373364011768
H	0.32772209793960	2.92598743392710	-0.65553247945325
H	1.92590055617451	1.73637488874379	0.43310620982293
H	1.21932064760365	-1.41107976070895	-5.35415974994496
H	1.47941648522853	0.22082115834093	-1.01436710389881
H	2.77582525315992	1.05450145502677	-1.89673918367703
H	2.94063515951769	0.39088849666247	-4.03035837171870
H	2.00723291941790	-0.89956811432212	-3.26751600117775
H	-0.08788013223727	0.58961916433818	-5.69655820913568
H	-2.09721358350511	4.47456011723476	-2.90611112435621
H	-2.99403604901875	3.13218444982508	-2.22585489699400
H	-3.09022970703259	3.45402445507809	-3.98118116802838
N	-1.26363769714941	-1.16147111580778	-1.62263589622344
C	-1.23645143344504	-2.40100203766176	-4.10710453874481
C	-0.35529003187024	-2.78774181095365	-3.09221767364881
C	-0.39171478771477	-2.14827091084365	-1.84570084201874
C	-2.10213765327992	-0.77056129828308	-2.59157881642076
C	-2.12738921245930	-1.34857201201528	-3.85473134585089
C	-1.20855283558343	-3.14719770002614	-5.41026116145790
C	-3.07949849372837	-0.81898621982091	-4.91504855868519
C	0.53843064555086	-2.57180012255867	-0.75183483359825
O	0.55553059261454	-3.78902705298914	-3.24155514472093
O	0.16565887940097	-3.17091063352944	-5.87252742539560
O	-3.64662738879161	0.42868224039285	-4.59801809456655
H	0.71114653589056	-3.92390069326296	-4.19200850989163
H	0.208613742120250	-3.62244541458744	-6.72358948103967
H	-2.92914416035289	1.09147811834123	-4.70072112304600
H	-1.84977383057545	-2.67361668557234	-6.15960436969799
H	-1.56238564032760	-4.17916994352448	-5.25468665432757
H	-3.91197068947262	-1.52569646641886	-5.04665129134593
H	-2.54732230335680	-0.77006183371784	-5.88049634725446
H	2.73269046550816	-0.77833895819123	0.69730127728322
H	1.58050044987643	-2.56928001512659	-1.08714252626838
H	0.43913538329080	-1.90912646631571	0.10986937725006
H	0.32798944856736	-3.60016884133919	-0.43569548191001
N	2.15416426148357	0.78495615912565	1.92474433772709
N	-2.65050834853385	2.59120550003258	4.71414373308759
C	2.50599921062175	-0.96627098414649	4.07320003086125
C	-1.67973780956685	0.84322081624155	2.72660075649730
C	2.13721260582967	0.36788993732541	4.26934364165899
C	-2.16257228938186	2.12248026957159	2.41248447275502
C	1.92527398587523	1.21775269907209	3.16399068889376
C	-2.62870136821993	2.97159863705765	3.43931563849680
C	2.55505867305379	-0.47990024739696	1.72616652766555
C	-2.20636062148440	1.36802012702805	5.01964292180131

C	2.72633800273670	-1.39426394181334	2.75263215263718
C	-1.70461457960999	0.46748315112061	4.08278787978361
C	2.70693922101096	-1.92589612932598	5.22919181783557
C	-1.19377662429165	-0.12264069838909	1.67080342605880
C	3.15223520610446	-2.81738699526616	2.45421026812565
C	-1.22275480469283	-0.87864366709818	4.56520544351372
C	1.45576575070015	2.62192000195444	3.37270464242335
C	-3.12849128718911	4.33952147785657	3.09064402368464
O	1.96693013256936	0.92102787401389	5.49998430487915
O	-2.21980344623964	2.62874394041728	1.16274320058881
O	2.02467677591837	-1.55650462367094	6.42731760020135
O	-0.84407763329152	0.53241690754054	0.45837882927010
O	3.49860928649605	-3.04081673640626	1.09462684879618
O	-0.57840956033139	-0.78409803637203	5.85276637508107
H	2.11388183500768	0.20657185822498	6.15935265496479
H	-1.74083393466740	2.01236812328137	0.56394496117012
H	1.06018651588002	-1.57938822123821	6.24531978084356
H	-1.02934559977505	-0.09772305024538	-0.30099132408048
H	4.33205383400171	-2.58899121831155	0.91749558664989
H	-0.21632620254526	0.11412500844942	5.90479172710643
H	3.77383987071710	-1.97322316094914	5.48951861751267
H	-0.31310509722579	-0.66077409397073	2.05867273221466
H	2.40859421898927	-2.94027269426022	4.92739246255108
H	-1.97519291802451	-0.88046245876594	1.48296443866255
H	2.32045945913703	-3.50502724628354	2.66024725966658
H	-2.06348225782330	-1.57223638969583	4.69709829710601
H	3.97936956455786	-3.11371027119275	3.11987266530624
H	-0.52819553954114	-1.33385184121156	3.84652723636437
H	-2.76771575740714	0.05748347514813	-2.36469144563761
H	-2.26242463548587	1.08720158640028	6.07058787373915
H	0.53640955644605	2.63825781721596	3.96917463892983
H	-3.97448232456434	4.28024619976447	2.39663134676437
H	2.20170531210992	3.20442856865107	3.92531467697144
H	-2.35305135972069	4.91908712641935	2.57737222326757
H	1.26457086648749	3.09990981957054	2.41000375693007
H	-3.43594950248856	4.85651983756588	4.00079297843247



Main article molecule number: 14

Name: (+)-Borneol

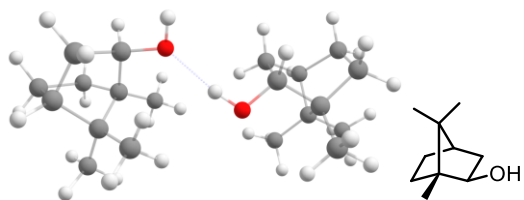
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -934.027737487291

C	8.56048217056175	30.47305324463541	12.37325635533960
C	7.65649321051954	31.36766116624389	11.47197315959708
C	7.43002756524891	32.62838470327893	12.32967848435894
C	6.44269339981910	32.23567350322636	13.44662587763084
C	9.98246368186839	33.02582801031762	12.19113326820860
H	8.13775840708660	31.59258611325870	10.51441918359997
H	6.71040634384473	30.86207944629359	11.24730760193095
H	5.57029788710522	31.71353957272662	13.04093100936392
H	6.89924906527913	30.30398158762947	14.44692357908500
H	10.06588394460707	32.34152233320101	11.34485591637998
C	8.69438417540663	31.30780257375048	13.67896417552949
C	7.30607502957606	31.31630613575802	14.35750388379800

C	8.77507006491868	32.75938936745310	13.09870884231950
C	8.77492924074204	33.85774056812726	14.16703323675632
C	9.82005631691821	30.84067509412919	14.57787279523328
H	6.07818676919420	33.11435920867336	13.98879242920977
H	7.38238004023267	31.70649804244302	15.37814053603473
H	9.92170785229649	34.04224003964940	11.78442123313732
H	10.91443328160982	32.96109940963032	12.76455592065151
H	10.78670803461034	30.87908916800247	14.06189549820201
H	9.88930708095932	31.45413032157741	15.48373188285598
H	9.65062582420213	29.80219816815357	14.88382208631840
H	7.99188340360725	33.74175243244760	14.91887767031668
H	9.73739311942518	33.87894805801076	14.69131021142737
H	7.13125466604429	33.51510610509534	11.75895190428879
H	8.63948776737630	34.83821031865843	13.69459812795904
O	8.10846970018531	29.13955743952584	12.55233732258850
H	7.16150526062860	29.15412027006018	12.76567025055556
H	9.55714352083867	30.35654237305275	11.93145288792804
C	5.04981458756098	28.19227701512389	11.45438212475766
C	5.08135658311395	26.65342619462755	11.65242731037923
C	3.79224152610478	26.18389343553724	10.94999132627492
C	2.61414014366452	26.58289568840383	11.86102290608539
C	4.81308775342076	27.10800089173679	8.75972727445817
H	5.98567142383107	26.21319000636325	11.22211047685279
H	5.08283055084879	26.39220351242298	12.71786539841439
H	2.78207676823199	26.25997023562440	12.89444183357953
H	2.70365503640400	28.66626524465398	12.65821578638097
H	5.80851032958885	27.15286105110331	9.20525797467408
C	3.71206731293107	28.44482880622726	10.71510267656745
C	2.57271521297378	28.13145759976707	11.71102499317823
C	3.66500747819697	27.19374505729930	9.77296847271717
C	2.36034674078584	27.06192426910820	8.98135768967545
C	3.62580013803420	29.81731057732508	10.0807563046844
H	1.67716066357753	26.12227436671456	11.53351143614235
H	1.61397594465266	28.47509926580674	11.30956874338280
H	4.74512285339090	26.16469236758413	8.20577412102291
H	4.73958230992084	27.92032866490067	8.02778625644841
H	4.44635637591382	29.98259251299320	9.37287657909183
H	2.67850074583268	29.95138810080203	9.54665465554755
C	3.69064731646458	30.59623195441455	10.84934151191773
H	1.46145926686149	27.16240800357377	9.59237680978661
H	2.31286967655887	27.82249615834663	8.19385865568348
H	3.79703636982020	25.12767200638795	10.66035689905386
H	2.32011373985843	26.08225499288401	8.49076287046202
O	5.21707354939251	28.95612829236482	12.65904806137662
H	4.68273591282034	28.55718542607142	13.35533818787065
H	5.88323686452961	28.52863352681429	10.82891760714153



Main article molecule number: **15**

Name: (+)-Isoborneol

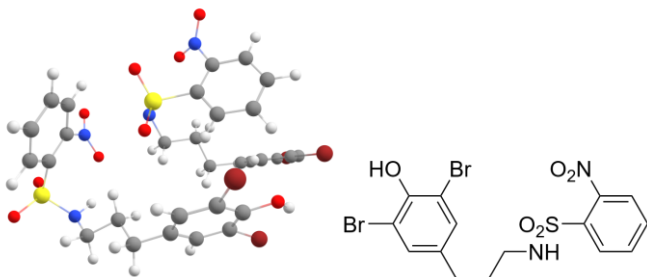
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -934.025723151241

C	13.16711733145648	27.94752308737340	13.26509871140271
C	12.190695236607037	29.05517012729147	12.76516999292383
C	12.87540655074234	30.36220897405079	13.20826682490767
C	12.77163038304533	30.41922902632165	14.74655140274940

C	14.74426854354847	29.81148691241129	11.50758217962979
H	12.09323337350808	28.99235585761980	11.67722189337511
H	11.18603143326856	28.95347294153969	13.19067742209957
H	11.75013975282783	30.21278062051438	15.08363007314196
H	14.18064217330507	29.00095279351467	11.04788126780018
C	14.34297238916690	28.75876957483163	13.86762419698520
C	13.78056241858907	29.32273411057212	15.19644762822868
C	14.37845927957110	30.04862728063859	12.97696198099962
C	15.31887362841794	31.14126743923477	13.49943303193528
C	15.62823874511635	27.97279403141717	14.02140735184507
H	13.04420778460521	31.40523241510059	15.13575350688941
H	14.57955465131247	30.73110793826987	10.93233263418993
H	15.80694739503671	29.55716003264143	11.41883410634767
H	15.98627735013311	27.60469249121816	13.05639695759752
H	16.40925571587130	28.59215157607495	14.47742652755548
H	15.47251718278322	27.10072398098184	14.66769900963684
H	15.15752243524331	31.40594163531200	14.54614467079378
H	16.36347028348776	30.82744860186197	13.39281346652466
H	12.50112531539080	31.25741218482311	12.69880934497513
H	15.19226425125420	32.05319833701252	12.90392225505944
H	12.70416422819401	27.34116887074482	14.06250390886731
H	13.30999700025211	28.53322491030895	15.79347695713954
H	14.58896032596070	29.73969266048510	15.80581825595058
O	13.62489043416799	27.07187956926482	12.23664177143517
H	12.84772625042733	26.60369696757209	11.89140741861217
C	10.58899414491020	25.32625110478326	10.11539372207280
C	10.69723461969735	23.77943599926870	10.00667855981688
C	11.47754424798450	23.56628968179102	8.69291214826218
C	10.52898911910855	23.96169350645390	7.54472377413767
C	13.52167078788849	24.64928182143627	9.84989712231122
H	11.24395998488742	23.38587950946492	10.86944597860780
H	9.71632732165844	23.29175175028178	9.98360426555748
H	9.55157519208377	23.47811434105747	7.64482657917821
H	13.08074821583374	24.48549585378561	10.83258276842633
C	11.38465347192303	25.82851516286760	8.88905159864182
C	10.45087247797282	25.50845066817593	7.69032256851892
C	12.49728855621168	24.73715629454363	8.71346591983448
C	13.29733585091493	24.88196021189543	7.41287043593468
C	11.82060673041962	27.27781786021556	8.95525466328080
H	10.93438896566851	23.66901576988773	6.57148771636447
H	14.21829660803344	23.82645363180356	9.64971882313529
H	14.11304819241019	25.56951321976087	9.91110169654728
H	12.52675476150101	27.45521025131794	9.77142439736641
H	12.30264805918918	27.57386267224062	8.01691262335745
H	10.95867333698926	27.93768137002885	9.11059084876091
H	12.68159695357871	25.01259287830501	6.52106140513255
H	13.97381178749692	25.74149083462595	7.47916064232099
H	11.91329948624096	22.56615629708613	8.59488475985262
H	13.91871820125684	23.99188411721767	7.26001423410646
H	9.54624821349499	25.67016143746765	10.04288459698796
H	9.43158668873754	25.86634166458322	7.87464019538008
H	10.80458343147183	26.02015326243139	6.79035171996834
O	11.13263471759468	25.85535350857861	11.34283821252131
H	10.61526803208558	25.49495036963629	12.07162327401639



Main article molecule number: **16**

Name: SKS-1-5-2

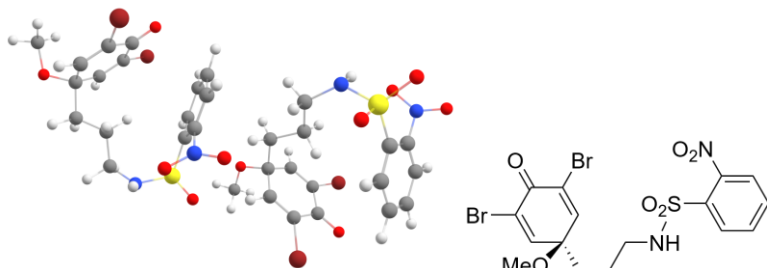
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -13223.5757911229

C	-0.85193535931600	0.89378917220941	0.14755668836490
C	-1.95011365619027	0.28843249584645	0.76049482605755
C	-2.58865561143782	0.91227675818816	1.82053837700147
C	-2.17183273498423	2.15398795612077	2.31806267033621
C	-1.08157478479790	2.75254663548603	1.67455884459515
C	-0.42478366573148	2.13367005021288	0.62078365283097
H	-2.32840983683399	-0.66187642130709	0.39489886519853
H	0.42042525827255	2.63353964345575	0.15880511942159
Br	-0.52227149649453	4.48460107447228	2.23544636517167
Br	-4.14344032388674	0.10725190979009	2.59436977091923
O	-2.76004278960910	2.79028961599816	3.35242226614882
H	-3.47434096569497	2.23695576179381	3.71036152175848
C	-0.17567142216637	0.22742316639762	-1.02161800494060
H	-0.94899029676470	-0.15398658855270	-1.70281229057363
H	0.39559576545648	0.97435368822482	-1.58543050131503
C	0.75558339917735	-0.92930253809368	-0.63061005648524
H	0.22370616503645	-1.61305845056852	0.04692501694226
H	1.62341198211515	-0.54491414395319	-0.07913942450915
C	1.19567573044717	-1.69560365612835	-1.87782555961987
H	0.30757489050835	-2.06601527331537	-2.41001243954101
H	1.72007570072836	-1.03643916990736	-2.57980655336517
N	2.06054276583004	-2.84922207657980	-1.59841739511682
H	1.89798535771350	-3.29492553397718	-0.69957682021111
S	3.70872068536175	-2.72676368068678	-1.86608292387087
O	3.86081523311111	-2.12947675609976	-3.19603206145025
O	4.28744277172884	-4.02283635616696	-1.51789223079703
C	4.41977069976195	-1.50578842257409	-0.71306255329534
C	4.81705937474224	-0.28871218770368	-1.25580273267248
C	4.73303865727083	-1.77103073534250	0.62470761462924
C	5.48338324571958	0.65132827373656	-0.47289712612955
H	4.62576253001009	-0.10296983710480	-2.30717549560254
C	5.44437069174048	-0.86294729540931	1.39326486822214
C	5.80341679635549	0.36241785470310	0.84635172572960
H	5.76866622571367	1.60334046419151	-0.90968457319103
H	5.67216136697196	-1.10289929415856	2.42416999597643
H	6.32239385919530	1.08670069674473	1.46523100493627
N	4.27428076290030	-2.98272895417395	1.32606564310109
O	3.13191880079151	-3.37295809849153	1.09432233339937
O	5.02732071938489	-3.48264114250221	2.14133094880801
C	-1.91069575650641	-1.79569622340024	5.24662737421605
C	-2.51146311250706	-2.77795199111612	6.02809507420818
C	-3.39608930927970	-2.41523880849109	7.03454034782552
C	-3.72580489880765	-1.08242728826187	7.30275540859090
C	-3.10762413344023	-0.11344601821799	6.50212906001042
C	-2.21513177718333	-0.45686011015375	5.50394733930567
H	-2.29769012436618	-3.82795003287743	5.85510801000762
H	-1.75931675142337	0.32798363689054	4.90634882325837
Br	-3.55103070504247	1.72167120311072	6.77411180899666
Br	-4.22144266732017	-3.77023570413350	8.10685471811603
O	-4.58894596652364	-0.69631901747691	8.26241505904382
H	-4.92327467950449	-1.48693286471562	8.71520192920816
C	-0.93232765083641	-2.12895912205557	4.15155110791400
H	-0.98426996904607	-3.19656839146858	3.90825278996687
H	-1.22208979068752	-1.58082062370212	3.24285793232336
C	0.50013439433376	-1.73273651668566	4.53472956422333
H	0.88231930523074	-2.40986155052778	5.30999863099922
H	0.47431395105726	-0.73405878593336	4.98109027920243
C	1.43430671389547	-1.72728391144423	3.32651299993275
H	1.54890743826507	-2.73855547758468	2.91679016529931
H	1.02298409037575	-1.10479660523188	2.52224199010134
N	2.79436659245204	-1.23329154844444	3.61349398938235
H	3.21329942492617	-1.63822123703900	4.44891948037749
S	3.06222960205476	0.42394484166146	3.61478386011045
O	2.51324360134662	0.92693153072712	2.34992580317842
O	4.46597068403099	0.62364960488159	3.98394220665998
C	2.06066247441151	1.196520229253566	4.92534348039070
C	0.95374326973953	1.93364032597549	4.51981029604816
C	2.37158270102738	1.15073135665719	6.28627312045301
C	0.14193680947571	2.56580363186843	5.45721211601244
H	0.73979018692425	2.00696179877425	3.45934288895946
C	1.58919536862253	1.80817014446185	7.22557026376812
C	0.45580610061402	2.49636849691616	6.80951700198863
H	-0.73073725585383	3.11581157067835	5.11949795961757
H	1.86979338436921	1.75896054666577	8.27221213449638

H	-0.17728025933572	2.98312189452393	7.54429561100953
N	3.50191181749293	0.36659072450847	6.82305190652509
O	3.62597675991634	-0.78754362141583	6.42783218614082
O	4.19456569496450	0.90174854776606	7.66904848526773



Main article molecule number: 17

Name: SKS-i-2-2

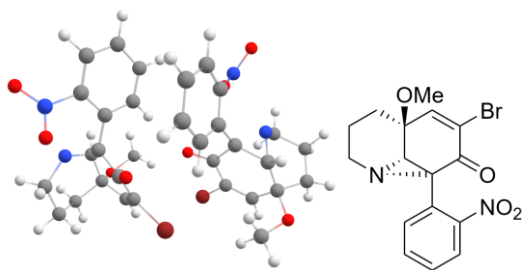
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -13452.501833357

C	-0.99831891434629	0.75662932461464	-0.90905409954719
C	-1.17445155697410	0.40268312848495	0.53836505052463
C	-2.33627776803058	0.54632188738800	1.17853851053344
C	-3.59067463665019	0.98295962553739	0.50578663244178
C	-3.43646614262081	1.25541571274941	-0.94685345219158
C	-2.27806107607380	1.13112197428844	-1.59713634589055
H	-0.28172931111782	0.05734638096798	1.05523056121315
H	-2.20097275252674	1.30517570219826	-2.66789383038023
Br	-5.02411998904610	1.77239038787763	-1.86245687965390
Br	-2.48918624605881	0.18181779915107	3.04091108215969
O	-4.64761238943339	1.09536870431699	1.09327066185498
C	-0.01307133617409	1.94744606570974	-1.01677352484253
H	0.98147865134253	1.55850175774097	-0.76346110330264
H	0.01366675457345	2.24173593851903	-2.07189507613136
C	-0.36880581675572	3.14659596042534	-0.14606639116751
H	-0.16954780250060	2.92849767946848	0.91134465234597
H	-1.44100803958477	3.35682293955380	-0.22636343426953
C	0.42545393680618	4.38385847346334	-0.57372134254410
H	1.48493024275350	4.11997766959010	-0.70177952864999
N	0.06797651267793	4.75372504060417	-1.54145569471482
H	0.40364463961730	5.50071256767492	0.37932649046737
H	0.75708107300931	5.26405491362952	1.30357211949156
S	-0.98663322337963	6.42394335622353	0.54941254420701
O	-1.45698487746359	6.69915480196818	-0.81072043716221
O	-0.66885040314779	7.48191887765415	1.50700761636929
C	-2.29592155846403	5.41323038303032	1.32383146309914
C	-3.37867562410657	5.06022071309973	0.52220091203087
C	-2.32964490298412	5.06487699649785	2.67588439474721
C	-4.44087114698427	4.33317736920742	1.04887704548823
H	-3.37558007146353	5.36638776308577	-0.51888943544278
C	-3.39981000869957	4.36299407156653	3.21604808406031
C	-4.44747015886405	3.97556501304835	2.39232542179282
H	-5.26002579774709	4.03560314272180	0.40264811229049
H	-3.38796294555251	4.11231534192514	4.27108235886503
H	-5.26324319702566	3.38661685546809	2.79713899027606
N	-1.23060272942238	5.37306114450112	3.61276711861767
O	-0.08476884563208	5.12128353026200	3.25199097263633
O	-1.54219349867098	5.79885524673838	4.70978357812517
O	-0.34239767863638	-0.30687451475766	-1.60817841003065
C	-1.12188789809857	-1.50474588327681	-1.72878095264451
H	-2.01093249202381	-1.34119572565181	-2.35073419948018
H	-0.47534478408355	-2.23138252164851	-2.22206229893806
H	-1.42591491985449	-1.88347479262309	-0.74334651781399

C	4.91322370652267	0.27586415891201	-8.09381473796130
C	4.82510641507354	1.76885081636794	-7.97018325598933
C	5.10516459250134	2.42004666720222	-6.83917284965553
C	5.60994173958686	1.74276674315598	-5.61366194930642
C	5.751337443350526	0.26702325478213	-5.74651168379808
C	5.46318333939948	-0.39420151186077	-6.86876567867309
H	4.48349859053143	2.29861884545215	-8.85696008863692
H	5.60377169201899	-1.47025124686104	-6.94312925858635
Br	6.40964647131374	-0.64033780617986	-4.20917385380023
Br	4.89538240743013	4.30482796760919	-6.70657001351023
O	5.89122925559329	2.34186196004543	-4.59631586358455
C	3.49853239758409	-0.28983685362194	-8.37948164976947
H	3.23641115632588	0.01650283557102	-9.39962462019280
H	3.58389454011000	-1.38429043751512	-8.38619455796837
C	2.43896251255257	0.16938597875638	-7.38156041928825
H	2.17409655031479	1.21891312465478	-7.56462971266810
H	2.84540297630573	0.12422370405652	-6.36454107014639
C	1.17810144105971	-0.69470705113768	-7.46149549775049
H	0.85935703370029	-0.80339934110114	-8.50689279731703
H	1.37380189741442	-1.70265399626738	-7.08001734567820
N	0.01355066818230	-0.15871890273589	-6.74472218063865
H	-0.31154687906776	0.74492493497387	-7.08130968080880
S	-0.02916906367251	-0.17998469426589	-5.07235961925996
O	0.47131616943405	-1.49558071507763	-4.66707150644931
O	-1.35681076461475	0.31700823428938	-4.70010548715612
C	1.16326866830052	1.03159889286249	-4.42186773080853
C	2.27674918862140	0.54492181805818	-3.74240340522002
C	0.97430438292925	2.41131936994743	-4.48691143030869
C	3.20572238776710	1.42356701865079	-3.19740021522296
H	2.39534533515890	-0.52863815498601	-3.63956888146164
C	1.88717273283257	3.29827425881316	-3.93154462819784
C	3.02221677750124	2.79855901581401	-3.30784445848862
H	4.08471821902155	1.03280484511920	-2.69547926175715
H	1.70237133834814	4.36554447345682	-3.99913985807288
H	3.76463577077426	3.48249389605722	-2.91042446515600
N	-0.21573792089237	3.01689127853922	-5.11670310841214
O	-0.44520371858056	2.74466372992258	-6.28941054289351
O	-0.86061841845389	3.78722309567770	-4.42790154834885
O	5.68172806047845	-0.09236701784907	-9.24418238264798
C	7.05207837090086	0.32078594060636	-9.20450338165005
H	7.59039785207681	-0.13553306364739	-8.36310650980642
H	7.49370264131762	-0.02163945137850	-10.14189566282920
H	7.14385077221057	1.41329258213567	-9.13896856896220



Main article molecule number: **18**

Name: SKS-2-1-1b

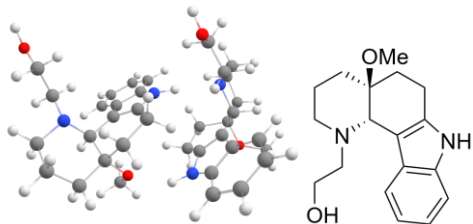
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -7205.86349706478

C	2.72071818388470	0.22691083897602	0.17560485636420
C	1.84586214245310	1.32439820404413	0.81943187319624

C	0.39420844606795	1.07590788359933	0.42422261649091
C	0.80503187036447	-1.43362753822783	0.23439366506722
C	2.26664564555350	-1.15983522001787	0.62312186982083
H	2.89575297401142	-1.92246636020930	0.15079235134698
H	3.76849477046882	0.40807644945869	0.43920229664450
H	0.78069930332005	-1.82015663625455	-0.79266052807498
H	0.35734488782930	-2.20059932662683	0.87406171554512
H	2.41028860742825	-1.26724637859871	1.70340584081285
H	2.62638245921221	0.34288946604927	-0.91240592572093
H	0.00536487745744	1.80729343868783	-0.28661537930365
C	-0.56554131475196	0.45227722601753	1.39547187130762
C	2.09182276237660	1.27082973769317	2.30211064273930
H	3.05993568894320	1.64473651319906	2.62973551073513
C	1.26118736086945	0.67519468834346	3.15797838359238
C	-0.06830099669428	0.13419974264559	2.79085742980855
O	-0.76245059970383	-0.47640510931739	3.57900059656229
N	-0.09731222607907	-0.27159251018968	0.21493034293769
Br	1.76001215414067	0.45063398913881	4.99170774646023
O	2.30559281473710	2.56179306503475	0.29054797557681
C	1.63683676135448	3.71284527993103	0.82082768336870
H	1.64030117140278	3.72045153723041	1.91802131557710
H	2.18519345209109	4.57991119037872	0.44671339797436
H	0.59585991226200	3.77746348876871	0.47409211664527
C	-1.99512543230603	0.89973938334862	1.34554199784264
C	-3.04232323677935	0.13846853515134	0.82712238664864
C	-2.28019820623651	2.21470311498423	1.71938272305834
C	-4.31820093586007	0.66022269310085	0.64493273461397
C	-3.55523342577592	2.74434678064486	1.56841513991944
H	-1.48262150292852	2.82916409005895	2.12887729547476
C	-4.57627563874830	1.96913767243602	1.02530782123558
H	-5.08850948492635	0.03163138369293	0.21192950102061
H	-3.74316507409789	3.77241972219552	1.86428330749731
H	-5.57177862270517	2.38223200731475	0.89397479561694
N	-2.84542863020175	-1.27337733581127	0.45376318270747
O	-3.43045418591478	-1.67380288746111	-0.54502435662651
O	-2.14964151251849	-1.96098795916550	1.18762772505936
C	0.11098826820177	5.32949967462677	8.99207433463277
C	0.27932912993338	4.04923851632933	8.13209015109306
C	-1.06013006985235	3.73009578519800	7.49861419737971
C	-1.07931541015562	6.13371579511497	6.86454909181308
C	-0.76400373977348	6.42283486049194	8.33267226231365
H	-1.71464814951248	6.50159398587635	8.87115982172046
H	1.10827800575990	5.70715176447080	9.24020091144649
H	-1.79808828134997	6.86095613477145	6.47354992331957
H	-0.17338375033693	6.21062933126853	6.25180804462707
H	-0.27689342569516	7.40015568009967	8.40568505055742
H	-1.033911224089222	4.99368595762917	9.93203869297807
H	-1.70676559829888	3.09264050677977	8.10716147279481
C	-1.24689649524085	3.68548684160488	5.99631171414182
C	1.38555529105090	4.26377002889668	7.13867142673639
H	2.34961233804765	4.52264158463926	7.57241238947634
C	1.24786810581547	4.14817467280421	5.81566196961271
C	-0.01647600128428	3.77655614676826	5.13817109516061
O	-0.05787009248045	3.52339063822457	3.94671854454568
N	-1.72353523303933	4.81230683117785	6.81225847015737
Br	2.74223830474327	4.45491027074305	4.66722746176943
O	0.67357666296371	3.03675963199450	9.06150301112830
C	0.92117805155627	1.75569821143633	8.47943167341315
H	1.78066501379334	1.77726656452857	7.79644004053936
H	1.13790060881578	1.08361717813788	9.31206033393637
H	0.04464803477879	1.37886291781626	7.93359018425403
C	-2.34777909961393	2.83556658427920	5.44412266404536
C	-3.57532154703242	3.35833292474638	5.03429027557939
C	-2.20782442364028	1.45077912158799	5.43292445940197
C	-4.64827233128002	2.55039020930321	4.68547007231316
C	-3.25037136460413	0.62409416374106	5.02580098148965
H	-1.26052721483484	1.01134564796155	5.73362959346042
C	-4.47795937574150	1.17221586675012	4.67215292852737
H	-5.58610158441709	3.01096880311156	4.39436041766537
H	-3.09325674710427	-0.44817462613514	4.98641409682265
H	-5.29816947611627	0.53116792839719	4.36480994306588
N	-3.75809118375857	4.80839473853307	4.87296394206846
O	-4.82307625389309	5.29175422341376	5.23071785293401
O	-2.84815394551173	5.42699404263517	4.32937598353314



Main article molecule number: **19**

Name: SKS-4-3b-3

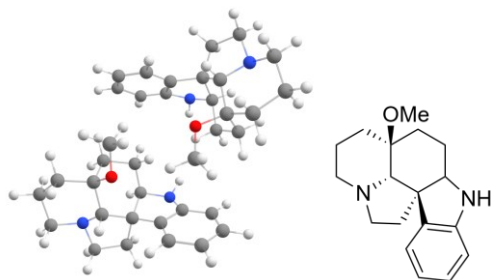
Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

Final single point energy (Eh): -1920.21246543518

C	-22.11418468397797	-8.95506958918705	-3.55869499148777
C	-20.77833605991088	-8.67533408598246	-4.25036655605066
C	-20.86961208667495	-7.36494883495269	-5.08348996923334
C	-22.60775404045678	-6.49827984276868	-3.60806682053364
C	-22.59542052695710	-7.75437489596320	-2.75083092954668
H	-21.94790967438528	-7.59806308393069	-1.88036160734446
H	-22.86451838260614	-9.18834299877060	-4.32637009724833
H	-23.37616107130831	-6.59816242064323	-4.40512656225415
H	-22.89626088514920	-5.64027671633360	-2.99054362518524
H	-23.60618880267143	-7.93496029790376	-2.36557543682290
H	-22.00887671746670	-9.84569919517547	-2.92793776434146
H	-21.64913011261863	-7.49789020247935	-5.86184262976198
C	-19.63418065541424	-8.57190128316609	-3.23752134393380
H	-19.61460713522309	-9.49640975033470	-2.64935867309614
H	-19.85028589103887	-7.75815607384629	-2.54391458997424
C	-19.54283533697224	-7.13020948140104	-5.72546477591616
C	-18.25818856993362	-8.35136178112627	-3.88899678229504
H	-17.58806832069357	-7.85041327652579	-3.17546937767182
H	-17.80039667316241	-9.31914356730765	-4.13252924810687
C	-18.37908088489037	-7.56700231402785	-5.14531279122179
O	-20.41610266238690	-9.80130508128364	-5.06812742985120
C	-21.06733180955444	-9.95416190940310	-6.32027045246056
H	-20.73011440696326	-9.20517327316485	-7.05128685783745
H	-20.78756969746578	-10.94617231041938	-6.68496476328942
H	-22.16379477942542	-9.91386112744724	-6.24019175869397
N	-21.28818912701956	-6.22477125741621	-4.20388277358203
C	-19.19249863130060	-6.53766372091795	-6.98639240609385
H	-16.37291207075747	-7.15947037395569	-5.60508991283247
N	-17.31034244113823	-7.28891529142642	-5.97196693926662
C	-17.7749981462300	-6.61964876589606	-6.08831206533702
C	-19.91102554763434	-5.94210856101615	-8.03038541163833
H	-20.99504072409677	-5.87531315445372	-7.99022667330609
C	-17.07764514520909	-6.09122017297761	-8.16791247388451
H	-15.99443094573251	-6.15744774415622	-8.22429703959764
C	-17.81584104567454	-5.49116582853269	-9.18081745557221
H	-17.30077099179808	-5.07234412003878	-10.04008086415792
C	-19.21649207605245	-5.42678506151654	-9.11585752496919
H	-19.76427976929091	-4.95996050533959	-9.92910214081885
C	-21.34140520666907	-4.97973973354015	-4.98689452955751
H	-20.36038922565165	-4.81324490207643	-5.44324092667724
H	-22.07398114517961	-5.05402300698180	-5.81469217719730
C	-21.65917683382055	-3.75236505099400	-4.13903385905961
H	-22.74225607429096	-3.65107592258551	-3.97100551639607
H	-21.16553119860879	-3.83810111566410	-3.15636855575149
O	-21.16305252142318	-2.62594623015122	-4.86724947344008
H	-21.43559934409316	-1.82594818952076	-4.40643001350467
C	-16.56857089145154	-4.89817650596681	-2.11596564750669
C	-17.56837098976315	-4.78755062651118	-3.04755046743450
C	-17.41246442768307	-4.30093936674790	-4.43990435561020
H	-17.96094706293725	-3.35829952181614	-4.57312842629923
H	-17.88107186670491	-5.01077346302538	-5.12847040763900
C	-15.92561565487928	-4.08734720608932	-4.80534469984352
H	-15.76664397130265	-4.37062834431218	-5.85309634672830

H	-15.68077827265186	-3.02712838669352	-4.71283332976700
C	-14.91792188718067	-4.85834368548494	-3.94597219584493
C	-13.47855268960000	-4.51748031035501	-4.34377008498868
H	-13.36405254839416	-4.65964093554592	-5.42559054656125
C	-13.06935168993834	-3.10810970203511	-3.92014460384740
H	-13.58341079247193	-2.35512402240579	-4.52887595321408
H	-11.99296321689143	-2.96922672189378	-4.07942779058306
C	-13.40356074073571	-2.88215731600418	-2.45343351897724
H	-13.15804666570684	-1.85374946700802	-2.16580755884554
N	-14.82830322538010	-3.11951218781078	-2.20201644104021
C	-15.20986202247881	-2.64370285254651	-0.86934186705386
C	-15.56735756885707	-1.16521669649214	-0.91920147038169
H	-16.48101107221214	-1.04476373877491	-1.52140398701356
H	-14.76416126955374	-0.58742102007696	-1.40457287538862
O	-15.76302498751379	-0.73099432840978	0.42922318757547
H	-16.07097077193166	0.18021872301247	0.41123501040336
H	-16.08942228031245	-3.19097924616408	-0.52070168076708
H	-14.40852537194972	-2.80247886211820	-0.12217915959993
H	-12.77724769428589	-3.55099828258368	-1.82209759830642
H	-12.80755390909222	-5.23493259523484	-3.85223608601584
O	-15.15665951559712	-6.25888472598450	-4.24165460406804
C	-14.63505013752436	-7.22936547829754	-3.33983164610283
H	-14.75946852771361	-8.19994180405036	-3.83052169691171
H	-13.56550529094694	-7.08496590597455	-3.13478904310066
H	-15.18824496738185	-7.24210148695938	-2.39074709290183
N	-18.75899702042045	-5.23070138862708	-2.51125388612343
H	-19.58182931736405	-5.49198615661776	-3.05777934504621
C	-18.52473471691116	-5.67692176408540	-1.22433136944919
C	-19.40146050035491	-6.21827701817919	-0.28967666141428
H	-20.45879638466477	-6.33667938739520	-0.51084733487882
C	-18.88230130018658	-6.58722677004995	0.94580393960574
H	-19.54198070366229	-7.01375794408661	1.69546163694806
C	-17.52259901616246	-6.40638092908443	1.24402191805662
C	-16.65197940981174	-5.84942855376115	0.31749096894594
H	-15.60471626885951	-5.69760497607368	0.56780870494824
H	-17.15051759757181	-6.69805158724689	2.22171164856228
C	-15.15605259774306	-4.54656103340850	-2.43969341812938
H	-14.47138761401243	-5.18043409116330	-1.83665817905907
C	-17.14538582386250	-5.47972022715267	-0.93978107179929



Main article molecule number: **20**

Name: SKS-5-1-2b

Qty of copies of the molecule modelled: 2

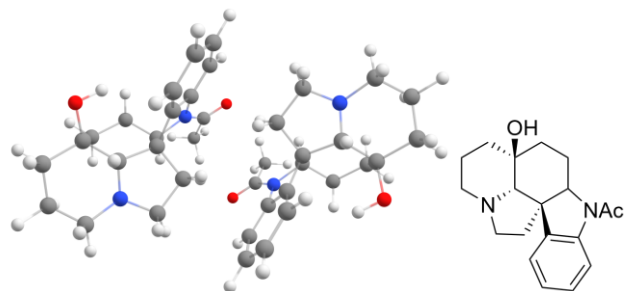
Qty of imaginary frequencies: 0

Final single point energy (Eh): -1769.78290724427

C	14.30152252368839	5.01364895789819	-0.03086045474253
C	13.00829814012873	5.38059801308495	0.71547421339173
C	12.36496350024472	4.09299382615927	1.24810039585762
C	14.48987881418910	2.89931254220254	1.30625043168652
C	15.23636885370614	4.12128785286855	0.78677641546796
H	15.65134547154734	4.67712668426341	1.63607251211245
H	14.82545449295341	5.92753729344620	-0.33571064431339
H	14.20069870214374	2.24516964134095	0.45334195573514

H	15.13446473322009	2.30408700412866	1.96647369669993
H	16.08239142387596	3.80419108089308	0.16532002044200
H	14.00857099149216	4.48436155405367	-0.94943025570619
H	12.02816798467543	3.50749575707216	0.36271672633232
C	13.25893976840136	6.35811401300754	1.87131659124065
H	13.71187424802155	7.28026193967424	1.48861184057294
H	13.97798697780133	5.90620637606449	2.56429335497193
C	11.18087034111075	4.24722282797955	2.21542904089236
C	11.95530991486508	6.66529304949381	2.60258211600426
H	12.12644982666744	7.41639661975718	3.38711889702391
H	11.23229365703199	7.08120702812815	1.88857582661065
C	11.36316939785837	5.40339124625880	3.23732322562504
O	12.03061450159571	5.91333395865819	-0.19288636074071
C	12.47484808131322	7.00161937883365	-1.00388454390082
H	13.07941369229062	6.65218654243204	-1.85356563518684
H	11.58478101471109	7.50430280685423	-1.38468042560913
H	13.05079272542049	7.74043000851663	-0.43718929322074
N	13.32098910878376	3.32932989456805	2.05559903958642
C	9.80430708438798	4.54999756550170	1.67654383546426
H	9.80673234577148	6.49531335530968	4.17631198442470
N	9.97514469058320	5.61479038697791	3.70734594661503
C	9.12759323717914	5.31626759992039	2.63566092206264
C	9.14309958299678	4.10067333259145	0.54942360236438
H	9.66660982827045	3.51094971724906	-0.19884512579471
C	7.78014529015207	5.63045302266510	2.48377134446303
H	7.24972564803663	6.21290135955889	3.23279911970171
C	7.12366595708707	5.17401995520364	1.33990421098125
H	6.07324225775864	5.41526017658691	1.20058023869284
C	7.79234927217230	4.42337387746878	0.37525419877952
H	7.26389898847183	4.08423992952827	-0.51079386997957
C	12.54401682048467	2.23373410996083	2.63380731850237
H	12.42362615693609	1.40602716400462	1.90558873564648
H	13.05261210370935	1.82572448488371	3.51604471719050
C	11.17667918001220	2.86382053791484	2.95057448993848
H	10.36072207159010	2.24085393657911	2.57446402830891
H	11.01305378030892	3.00481122896258	4.02264803043408
H	12.00061803080988	5.07608896591606	4.06888041675467
C	7.14599109132755	8.58320879858537	-3.06329395466015
C	8.22274432823280	8.74686840667159	-1.97456603831861
C	8.63035973626073	10.22610319859336	-1.88797727242930
C	6.51615654623491	10.99606316329102	-2.83381987212165
C	5.98760986024239	9.57234725096079	-2.93169742777680
H	5.32173256191694	9.47800451416062	-3.79795038001870
H	6.75736675709734	7.55815979324484	-3.05102589045998
H	7.00209342040414	11.28432054735769	-3.79273822682320
H	5.69496375846481	11.70400831296654	-2.66015079693924
H	5.38999281697050	9.34959184872263	-2.03940806536703
H	7.64377552488328	8.73474747015591	-4.03180434840336
H	9.17250419079431	10.46585522992356	-2.83109239186386
C	7.75258594263711	8.24326770934323	-0.60910804963680
H	7.47915405181440	7.18401613527352	-0.67153089423636
H	6.84559172511838	8.79402255456111	-0.33186404064411
C	9.50297919958226	10.65393217685064	-0.69188740339383
C	8.83146391992047	8.45872505196955	0.44582101912564
H	8.52160464848663	8.01490298220522	1.39996868834802
H	9.75636086411441	7.95194304896366	0.13888910024211
C	9.11659647429876	9.94747544449549	0.63631415553610
O	9.43116541702489	8.08222896908138	-2.39415158910614
C	9.36636766743008	6.66113117834001	-2.47404438042023
H	8.51874018499928	6.31440794251564	-3.08062050655852
H	10.28962619697417	6.34120728932241	-2.96388770455441
H	9.31940359034612	6.18683740116879	-1.48346220133337
N	7.45102056392088	11.08343491594114	-1.72486923433298
C	10.99606300438251	10.42622236497867	-0.71587138243154
H	10.40946707100047	9.55999306252193	2.26599085502416
N	10.31850587957297	10.17906979666245	1.47014453542815
C	11.41615201966504	10.21055535500563	0.60536682514458
C	11.92680163810368	10.54732905148279	-1.73040706063516
H	11.60691109751626	10.70790078734532	-2.75688688110945
C	12.76828598176140	10.13046309936530	0.92669587767278
H	13.09317039946908	9.97989955504852	1.95312864464620
C	13.69829346746120	10.25864253704037	-0.10631526146358
H	14.75876329169476	10.19658273680555	0.12374770394111
C	13.28952419844420	10.45352337225131	-1.42422186230585
H	14.02946089869604	10.53885291449963	-2.21400584779232
C	7.98802377547829	12.42375476912784	-1.50197232136483

H	8.30171737767838	12.88916571077714	-2.45893294047113
H	7.22762928821501	13.07348408816203	-1.05136411440520
C	9.20961379634171	12.19207835432366	-0.59999227173328
H	10.07270091168902	12.76458821521939	-0.94973483692969
H	9.02499880188304	12.47838843357098	0.43923849411699
H	8.24229484698687	10.43576785672028	1.08673671943063



Main article molecule number: **21**

Name: SKS-6-2bs-1

Qty of copies of the molecule modelled: 2

Qty of imaginary frequencies: 0

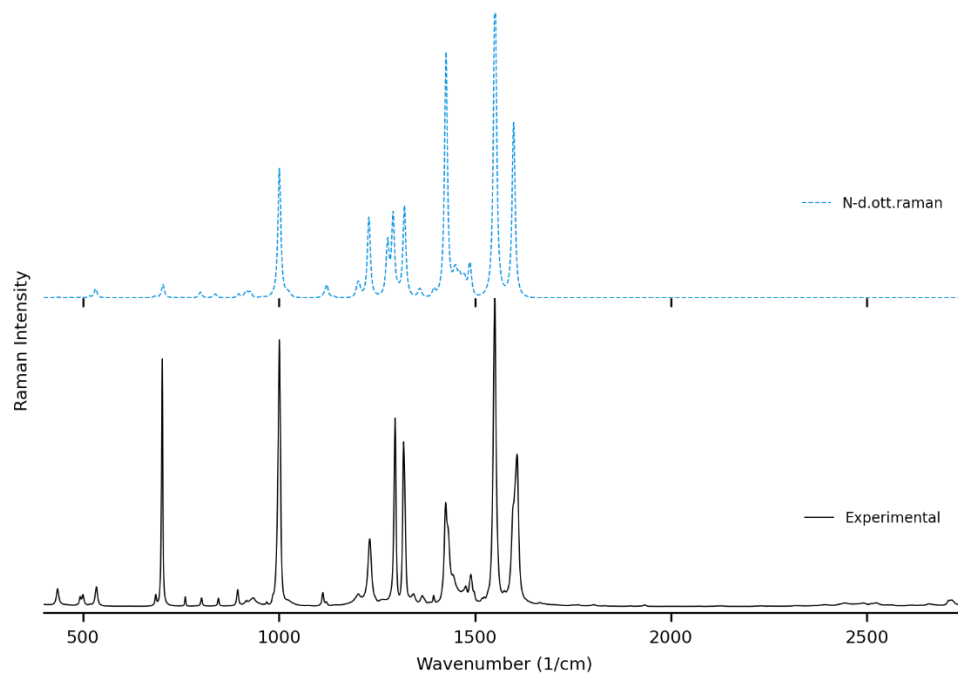
Final single point energy (Eh): -1996.51147290755

C	-15.25463573562393	14.49806279602931	3.82426150526327
C	-15.08341861118351	13.60227767807450	2.59457547137627
C	-15.59615659526143	12.19915427372811	2.93786398595844
C	-17.11824102729275	13.01916934981835	4.64271953996757
C	-16.67317628082852	14.45816437131415	4.39587985609947
H	-16.70643160159658	15.01945555386329	5.33703109592127
H	-14.96002680354561	15.52231037386198	3.56457777519357
H	-16.51231891195144	12.57641898313809	5.46417628239541
H	-18.16825948213992	12.98977284319437	4.96125792558754
H	-17.38103445509231	14.94111404777342	3.71156137886167
H	-14.54382484798694	14.14164327093533	4.58189883167855
H	-14.91175861973369	11.78095410506606	3.71211314184084
C	-15.81942325652576	14.15438886430660	1.36952824571916
H	-15.39066329267939	15.13298755231405	1.12356268082642
H	-16.87700558566617	14.29989681297379	1.61699595688476
C	-15.70268954851107	11.19951304341702	1.78008845186355
C	-15.69990432277138	13.21104083819817	0.17676280154931
H	-16.20642903024152	13.63352373052539	-0.69922939636358
H	-14.64519844576497	13.09763667343198	-0.10976020762160
C	-16.27969559110652	11.82521782104602	0.48127477979048
O	-13.66813113019455	13.55930023088163	2.35416350828729
N	-16.97676013398068	12.24567431040861	3.41835433439681
C	-14.44602111247630	10.55141870945724	1.25318212567546
N	-15.83002221891694	10.85658726007919	-0.54881627779085
C	-14.60406991723219	10.30628023558798	-0.11705885643281
C	-13.30405799303021	10.12128347331612	1.91026094117597
H	-13.17662089566653	10.29972858659010	2.97469381172091
C	-13.63109048305158	9.64315811032727	-0.85576193341897
H	-13.76917186910539	9.45354758086712	-1.91222856792716
C	-12.48385950827494	9.22298854166000	-0.18161235272934
H	-11.70710392958118	8.70561469503881	-0.73767910379320
C	-12.31553279010101	9.45397711116598	1.18145238482230
H	-11.41074125682435	9.11980530608848	1.67996777235748
C	-17.39694419975190	10.84518454532879	3.52766730434263
H	-17.06733133956997	10.40267676768621	4.48835951925672
H	-18.48945446183394	10.76350451021407	3.48209148468054
C	-16.69058689108369	10.13870626292172	2.35397204755356
H	-16.14778576360532	9.25485972830032	2.70133429654989

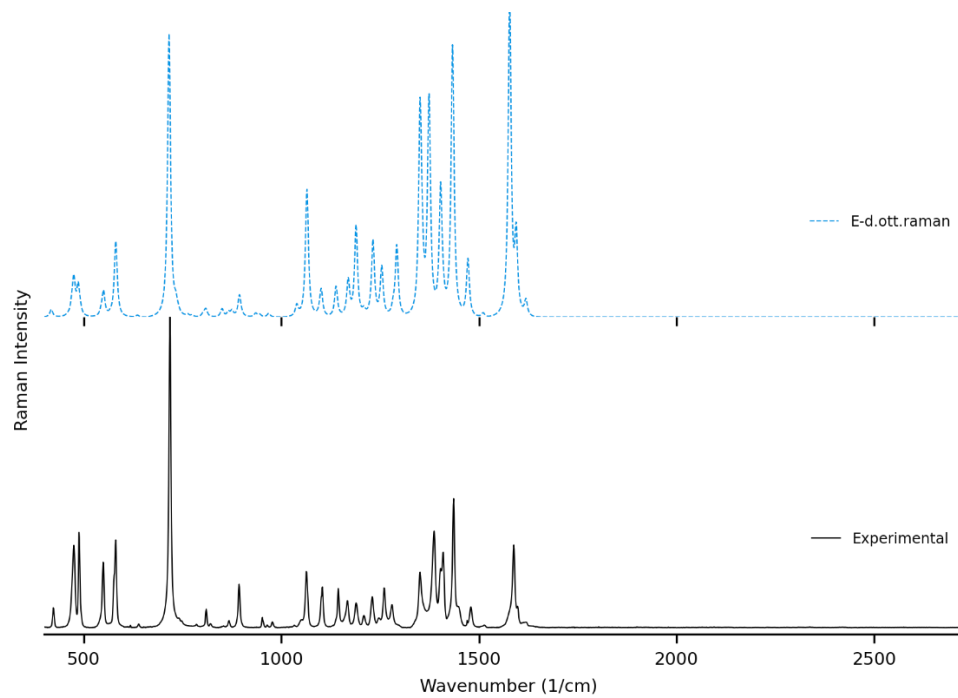
H	-17.39155382816112	9.80697381415758	1.58260739943783
H	-17.37107491928097	11.85994826196847	0.52901283697199
H	-13.47657698011701	12.84458785891957	1.73272404145714
C	-16.53935531717891	10.45345216857552	-1.65218805024559
C	-17.84694320664672	11.16570533831619	-1.91038790319337
H	-18.62150787750259	10.77651305019556	-1.23699442623784
H	-17.77553780572079	12.24717665650015	-1.76649850912011
H	-18.14324457263127	10.95296806947592	-2.93826853566328
O	-16.14134327147519	9.54436628694583	-2.37829911901282
C	-16.41390426539246	3.30724280366423	0.69754247298348
C	-17.46274194559311	3.39560585629293	-0.41615941168970
C	-18.35960437100284	4.61810913680028	-0.24142478510740
H	-19.11890132634679	4.65310105930141	-1.03190790872377
H	-18.91108685459936	4.54726568929813	0.70602971811747
C	-17.55753763120465	5.92446405424215	-0.24951714436001
N	-18.39699560254910	7.03275858135429	0.27049794073076
C	-19.06316697346589	7.97003815892777	-0.47792893544507
O	-19.60105895663207	8.94637868473146	0.04120362899855
C	-19.08033458980642	7.75289006183070	-1.97361516420792
H	-19.87442437954136	8.37409142772600	-2.38983789487917
H	-19.25055514907292	6.70833789045808	-2.24827194464931
H	-18.12451288990805	8.07426381203785	-2.40746021834375
C	-18.23252064644770	7.07532349700656	1.67186627746845
C	-19.03924915562144	7.67477309562191	2.63181332678093
H	-19.91925077990275	8.23633623543270	2.34637040883134
C	-18.66338213216343	7.53573746382941	3.96856317109779
H	-19.27989963266778	7.99163274368199	4.73817602953404
C	-17.52087948602883	6.82766030694291	4.33304310671511
H	-17.25414886940696	6.73230040822627	5.38116543096728
C	-16.71830684579507	6.23335328068562	3.35510096817664
H	-15.82257969607615	5.68574116407953	3.63636149087450
C	-17.08365035325775	6.35487860351147	2.02366268032616
H	-17.19386328594764	6.15459737473364	-1.25421326577267
C	-16.38074291860099	5.91593280626136	0.76240227295433
C	-15.29149471301284	6.94232573077470	0.33601677724597
H	-15.71659745615906	7.70884318282709	-0.31863177488325
H	-14.88518793041848	7.45381091409278	1.21378048421024
C	-14.20128891863938	6.10343820855243	-0.36044444634601
N	-14.76700371570701	4.75176774686946	-0.43605582214263
C	-13.78576998462100	3.68670417185712	-0.57594156965426
C	-14.49902201933548	2.34098970142624	-0.68189356295628
H	-15.04412794233019	2.29733098876309	-1.63244616335065
H	-13.75821517402889	1.53313093546041	-0.70206678130607
H	-13.18402796220127	3.87462812682703	-1.47453423716811
H	-13.08809964092121	3.67480108926857	0.29071597951815
H	-13.95156908998736	6.47415634742062	-1.36193772899319
H	-13.27248255638650	6.11795968970761	0.24232886769661
H	-18.06429215928740	2.47919715866950	-0.39642307250412
H	-16.95710168660153	3.44933618695987	-1.38685977669871
O	-17.04577782288064	3.06897675648821	1.96498103063800
H	-17.47928891615283	3.88231590563541	2.25563322545834
C	-15.59467558301986	4.60032523966766	0.75972447275823
H	-14.97127035829699	4.55081965698717	1.68281866809389
C	-15.45215333978435	2.13324772313819	0.49699195933119
H	-16.03223981507102	1.20962294276928	0.37958340400422
H	-14.87225875962361	2.03268694517289	1.42431753975661

4. Raman spectra

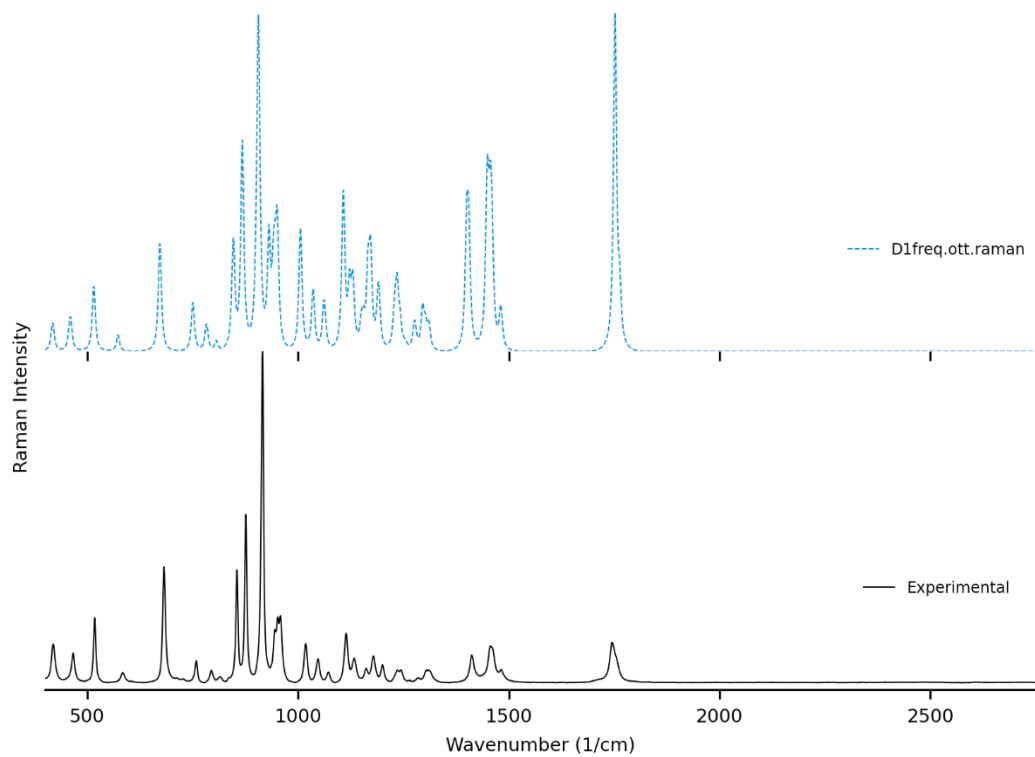
Molecule 1 : 4,4'-di-t-butyl-2,2'-bipyridine



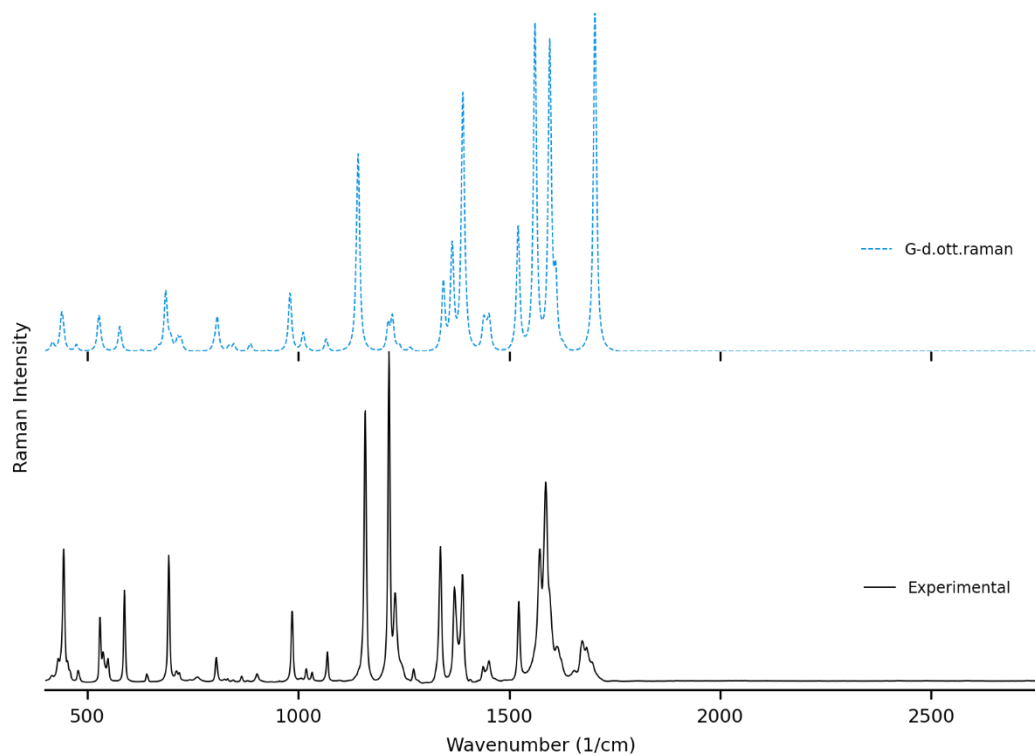
Molecule 2: 8-Hydroxyquinoline



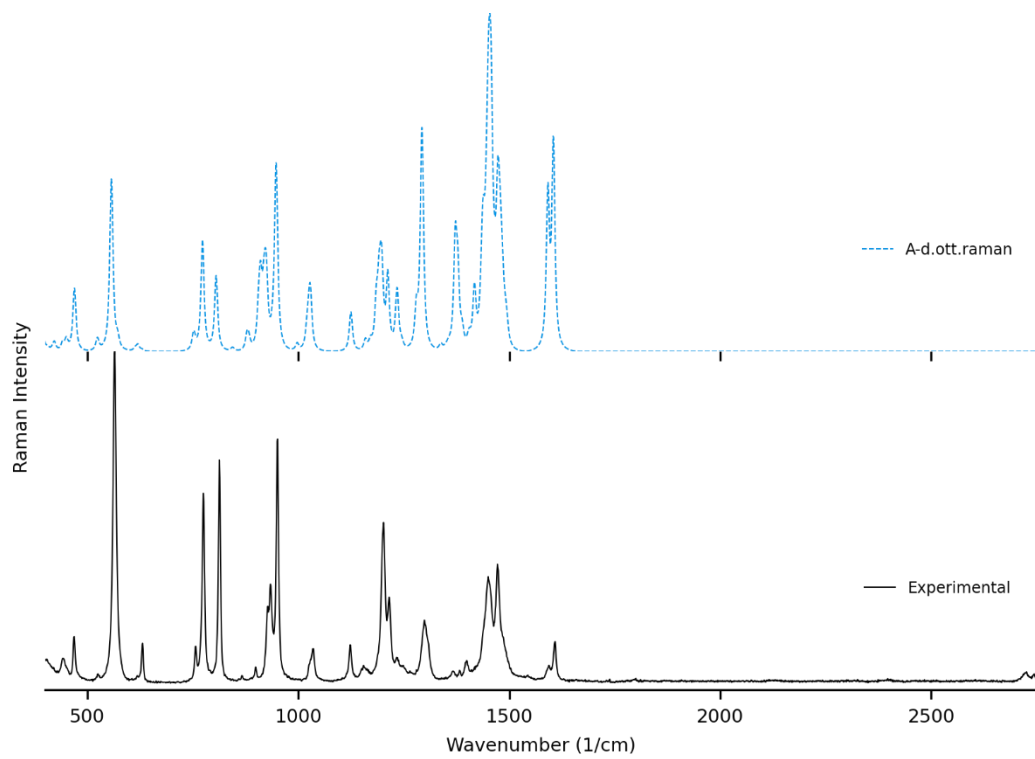
Molecule 3: Norcamphor



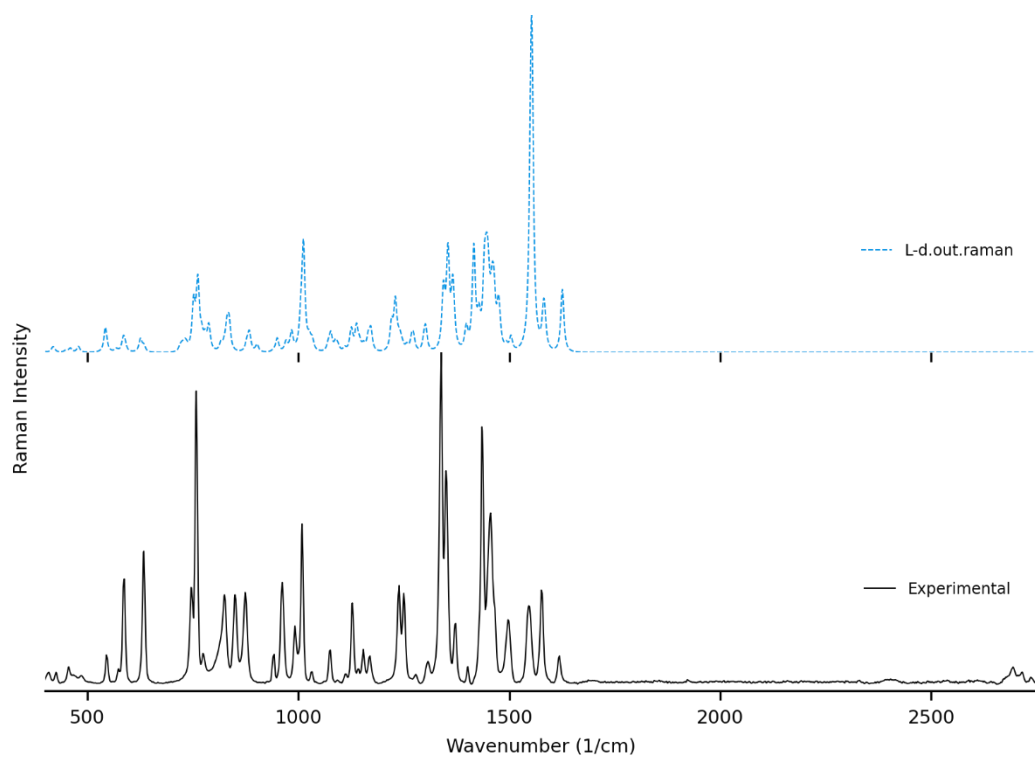
Molecule 4: Methylparaben



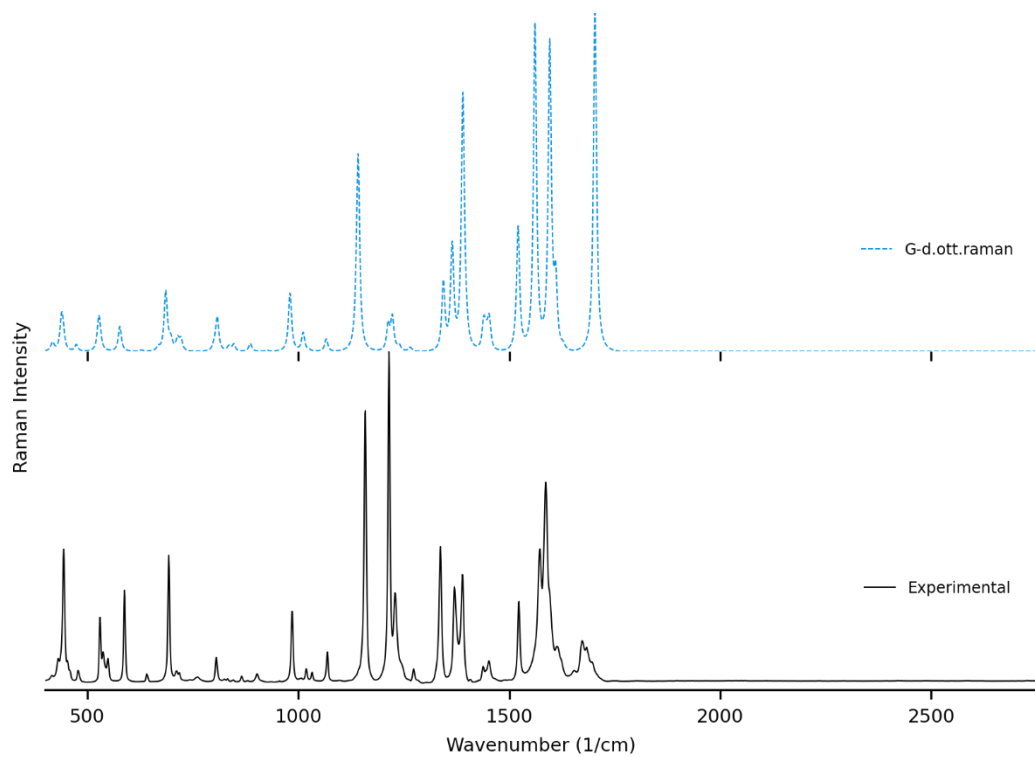
Molecule 5: BHT



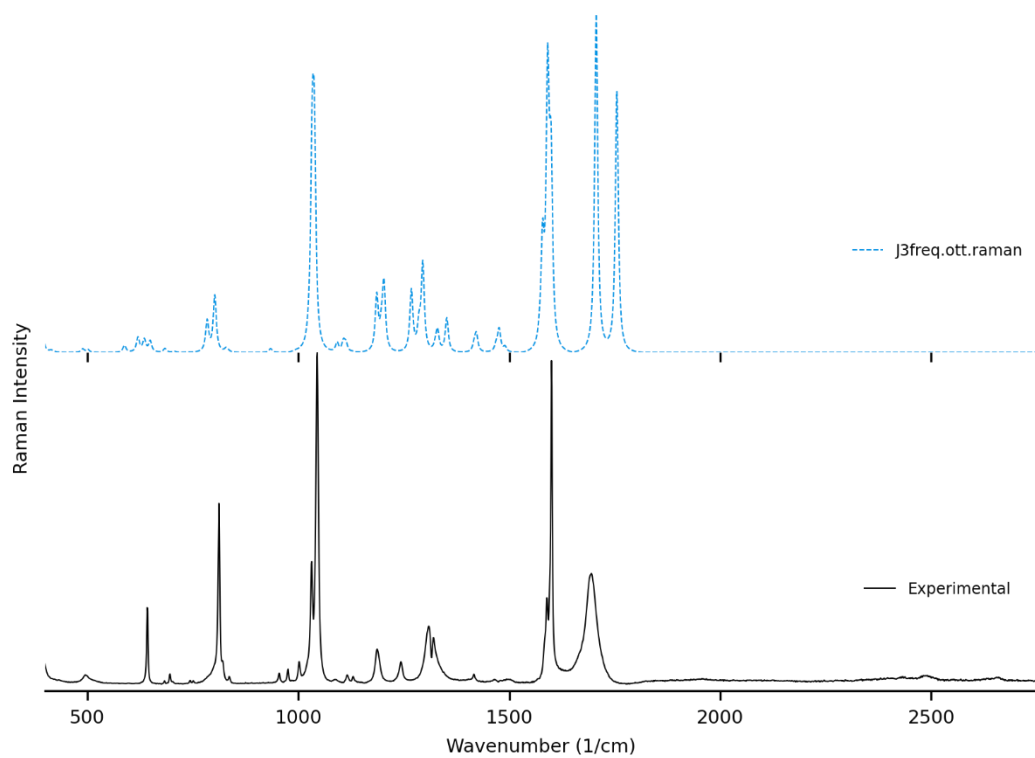
Molecule 6: Gramine



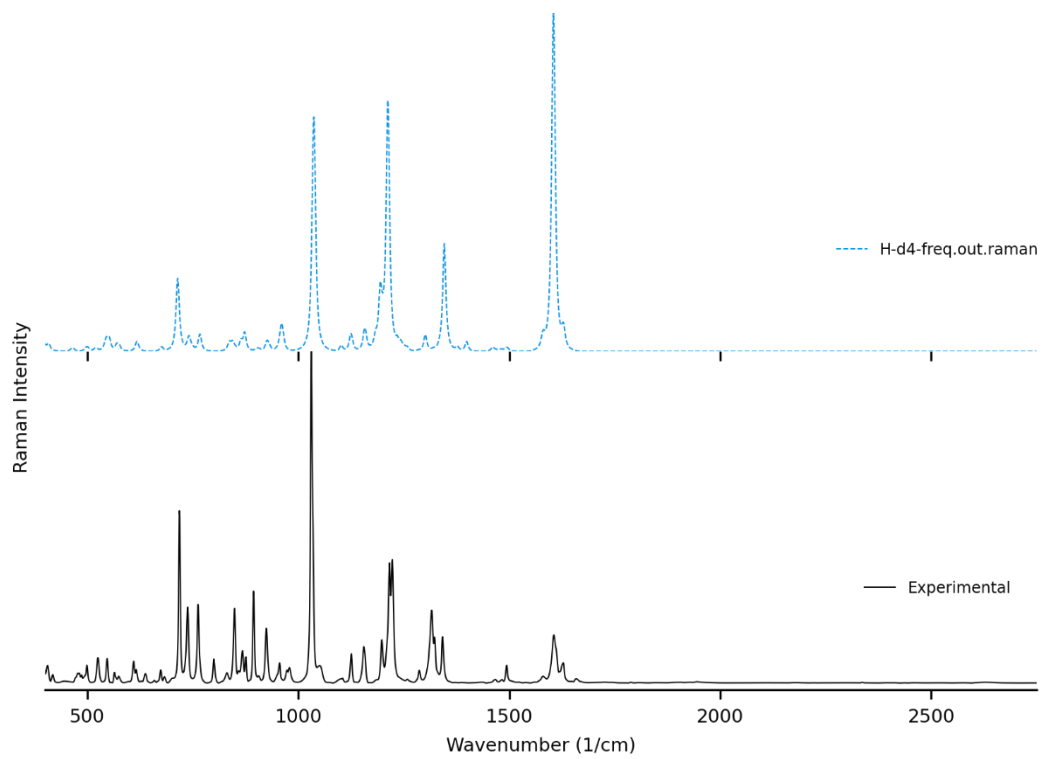
Molecule 7: Hymecromone



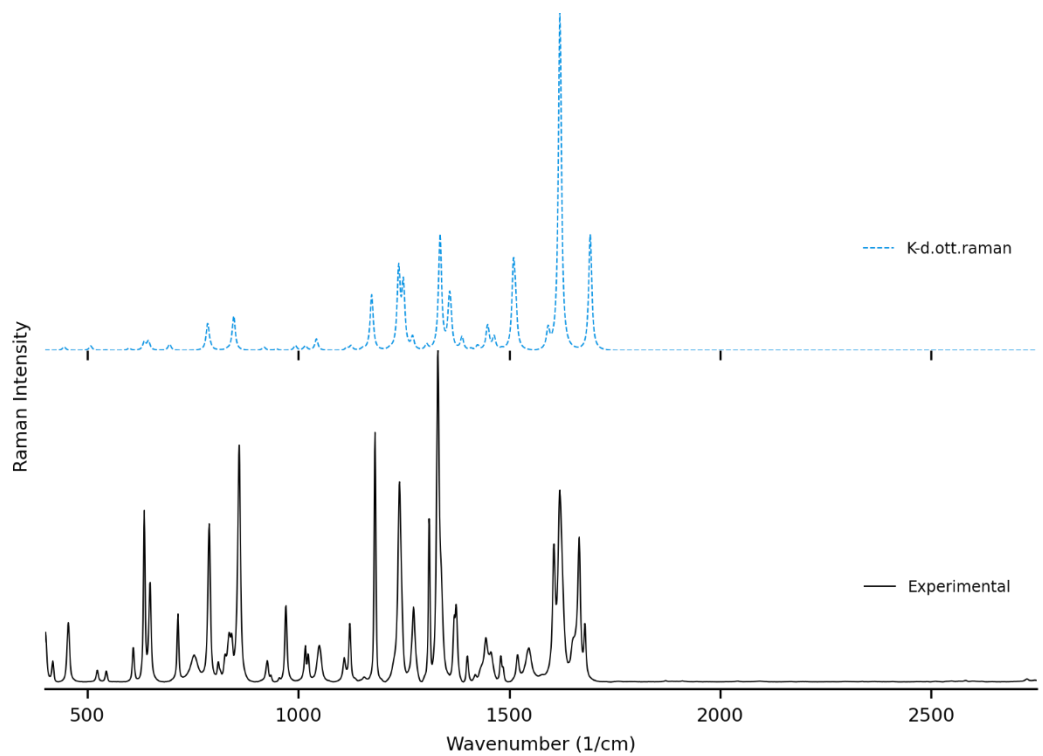
Molecule 8: Nicotinic acid



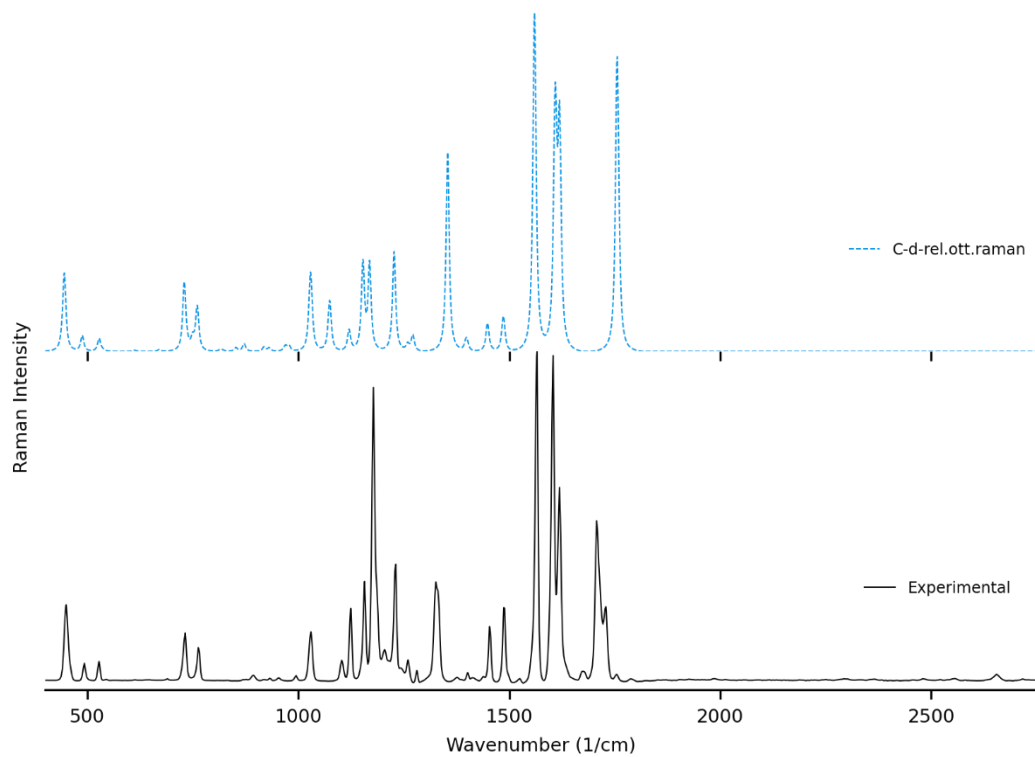
Molecule **9**: Xanthinol



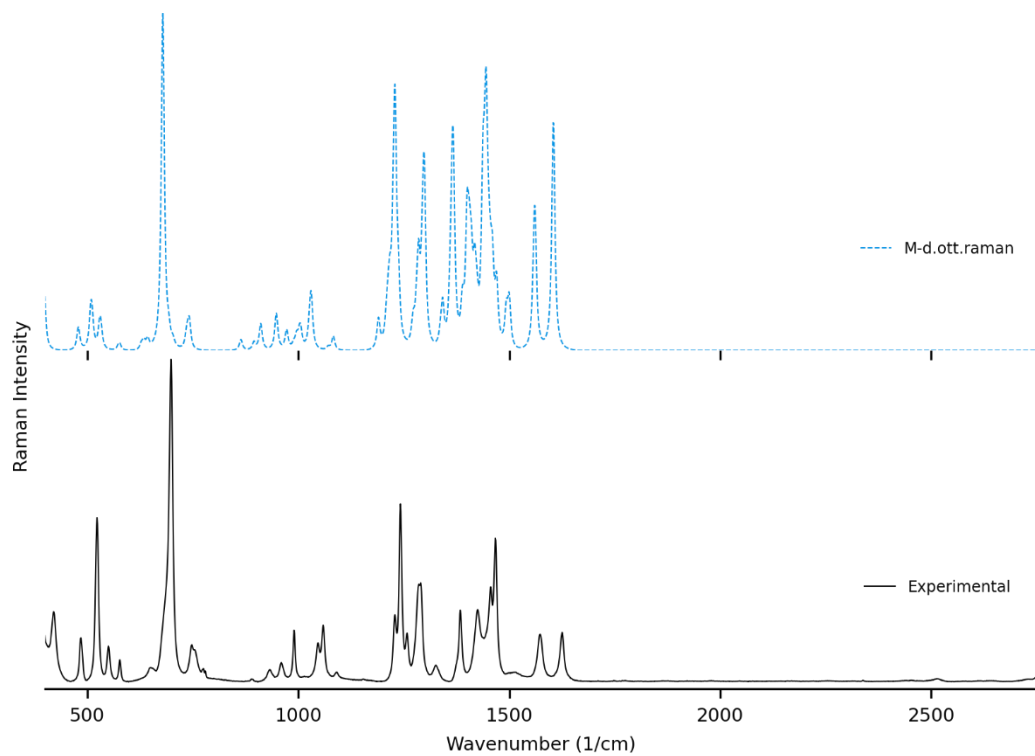
Molecule **10**: Phenacetin



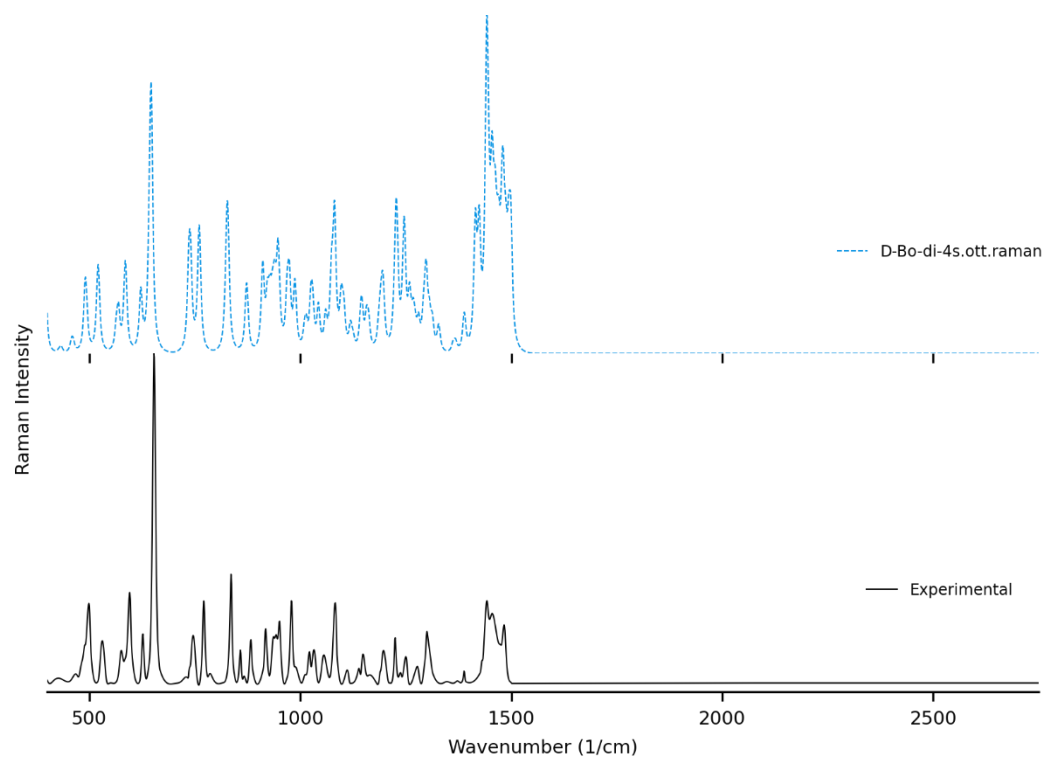
Molecule **11**: Coumarin



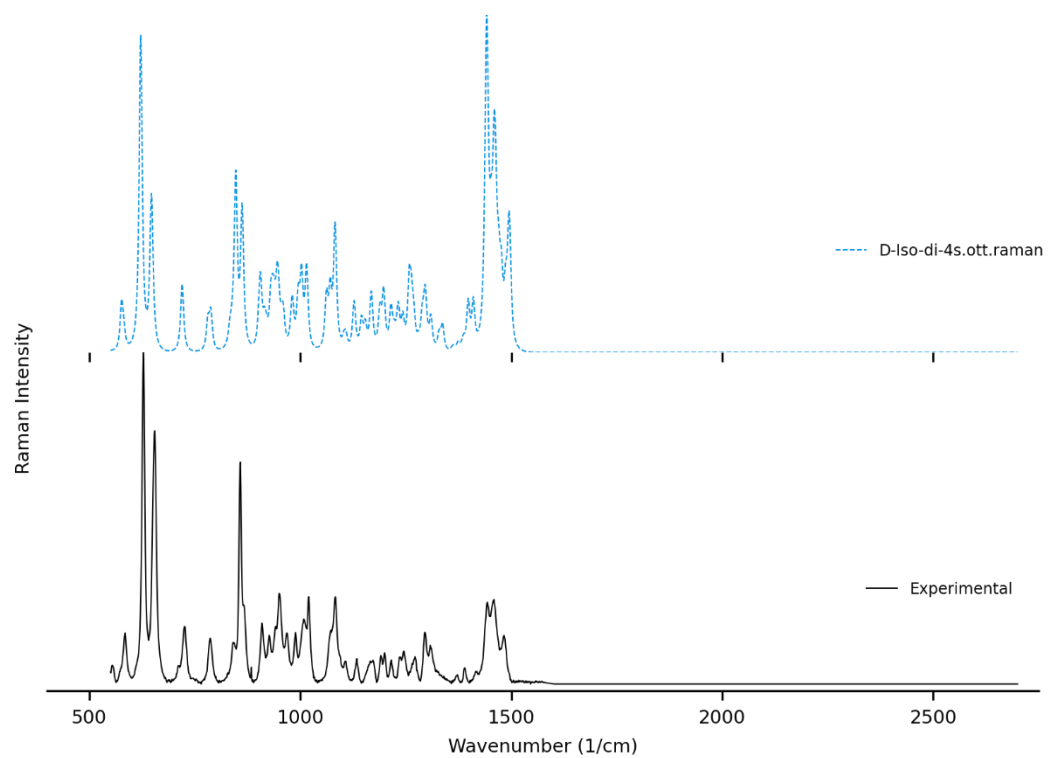
Molecule **12**: Pyrixdine



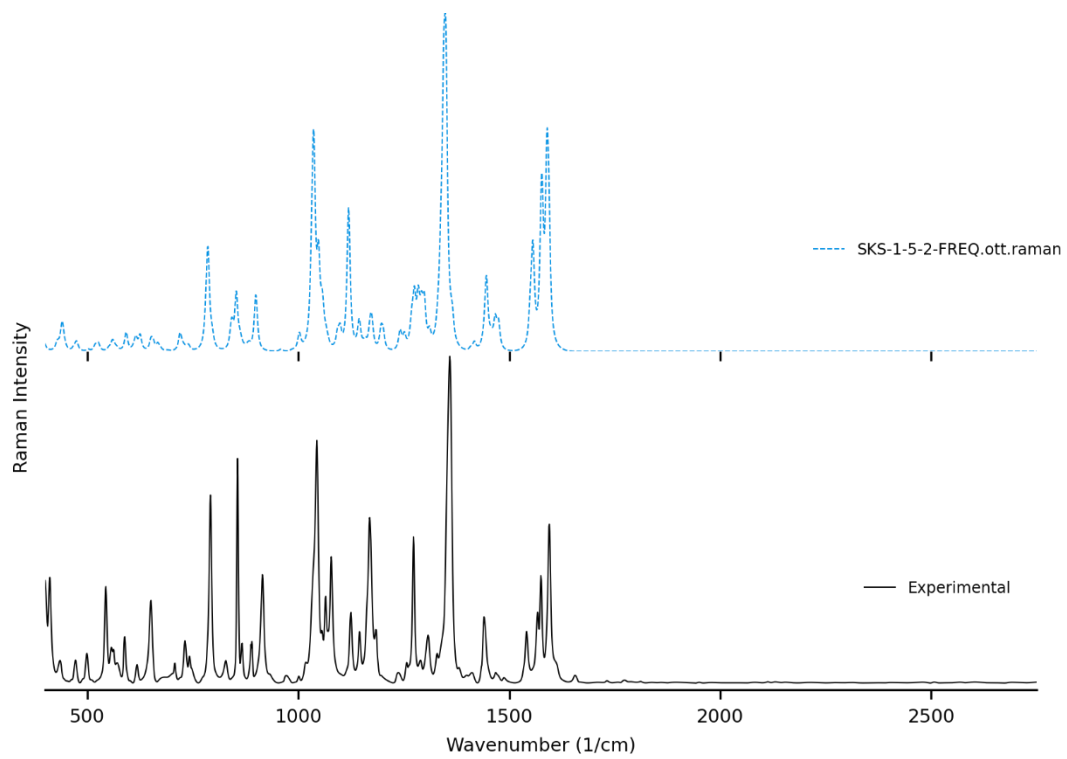
Molecule **14**: (+)-Borneol



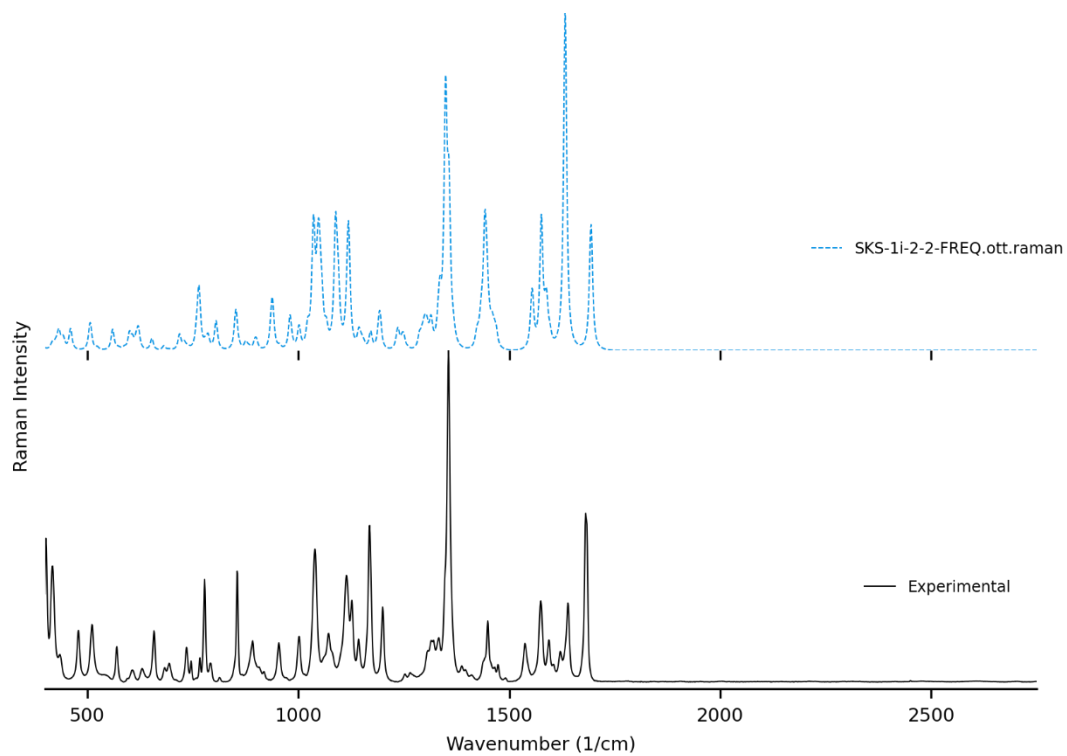
Molecule **15**: (+)-Isoborneol



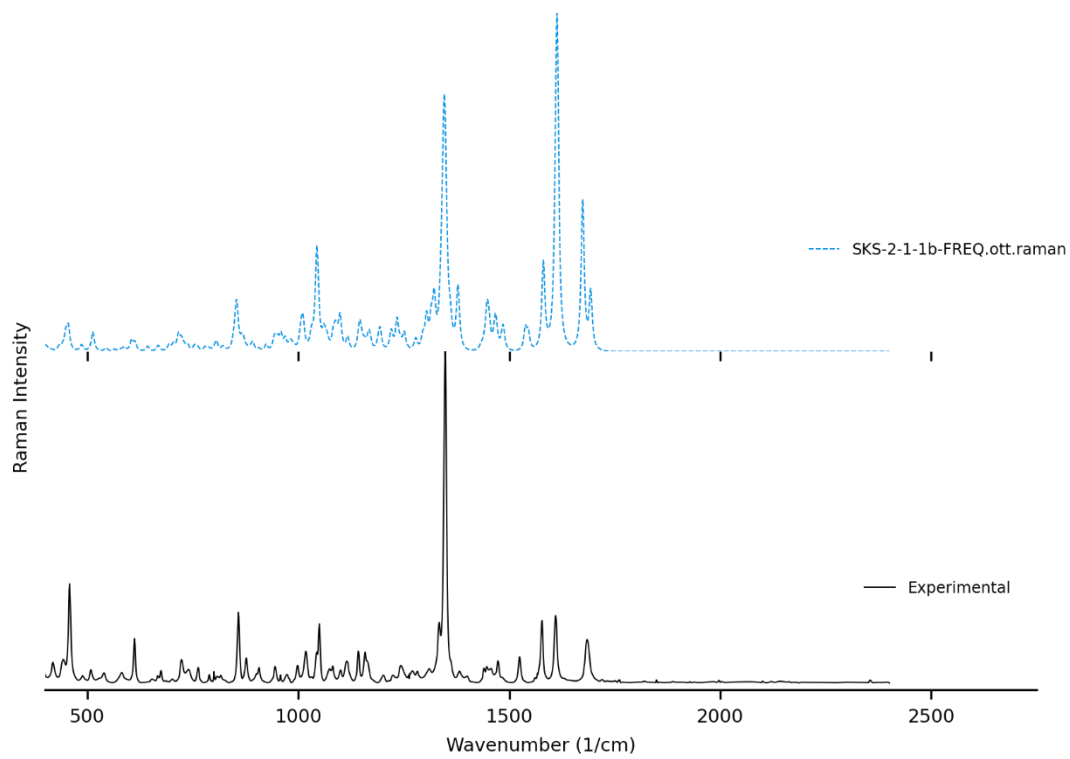
Molecule 16



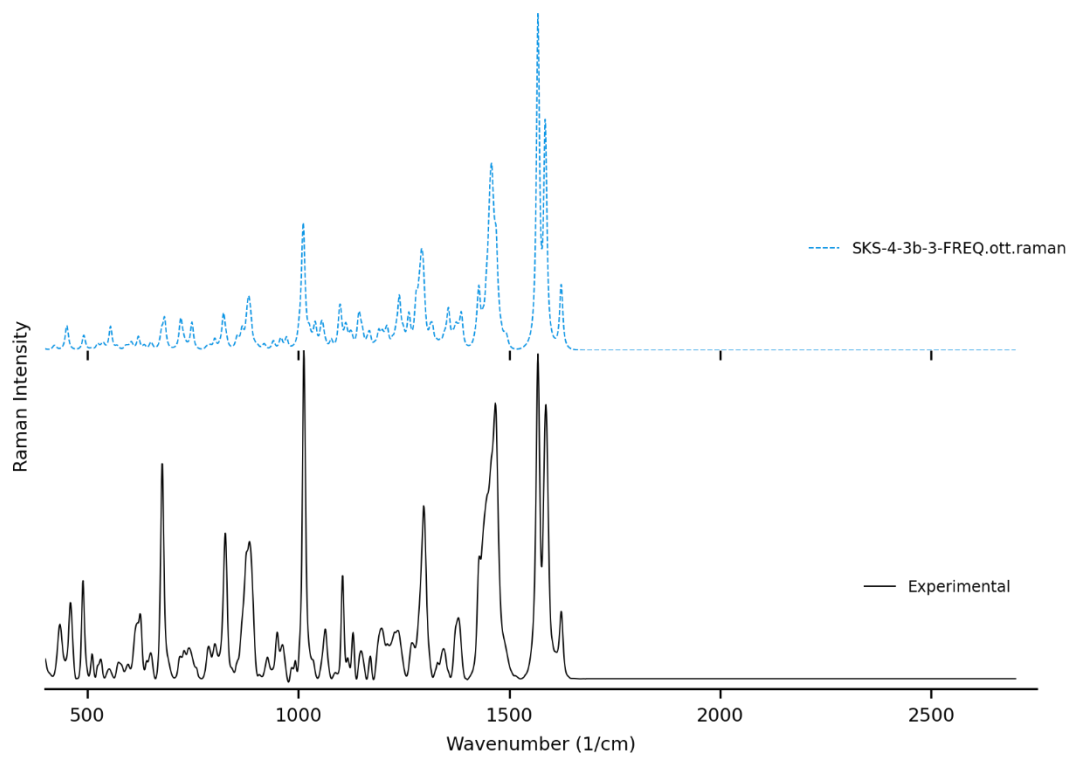
Molecule 17



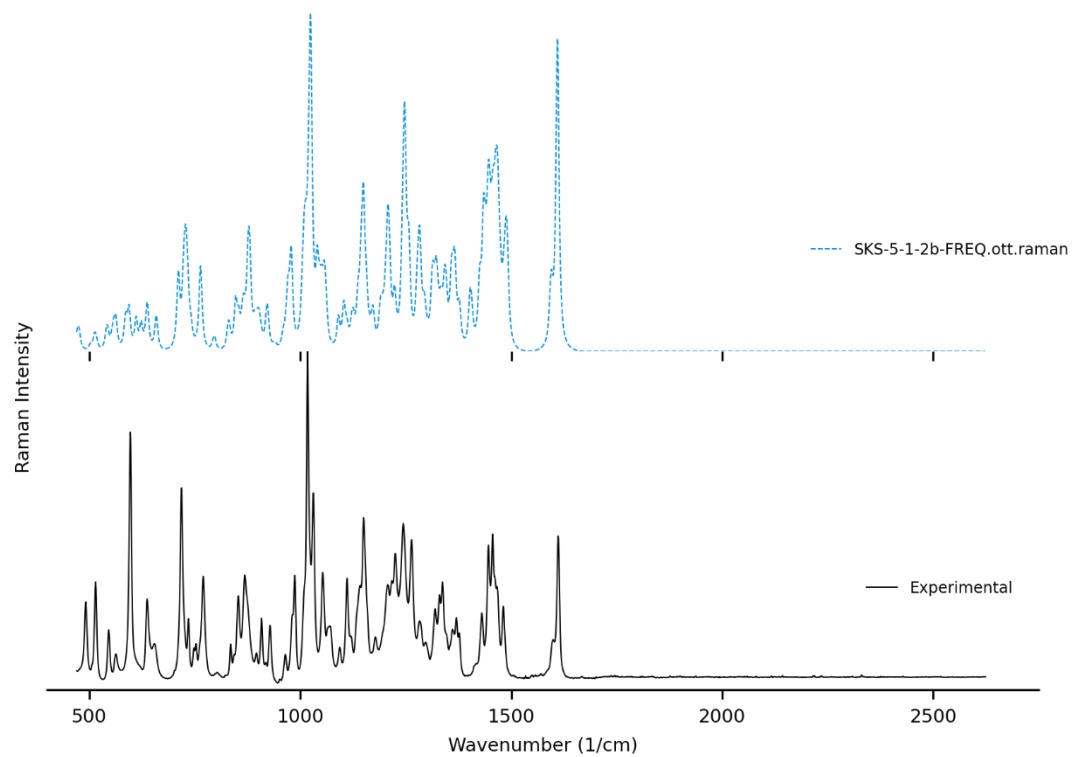
Molecule 18



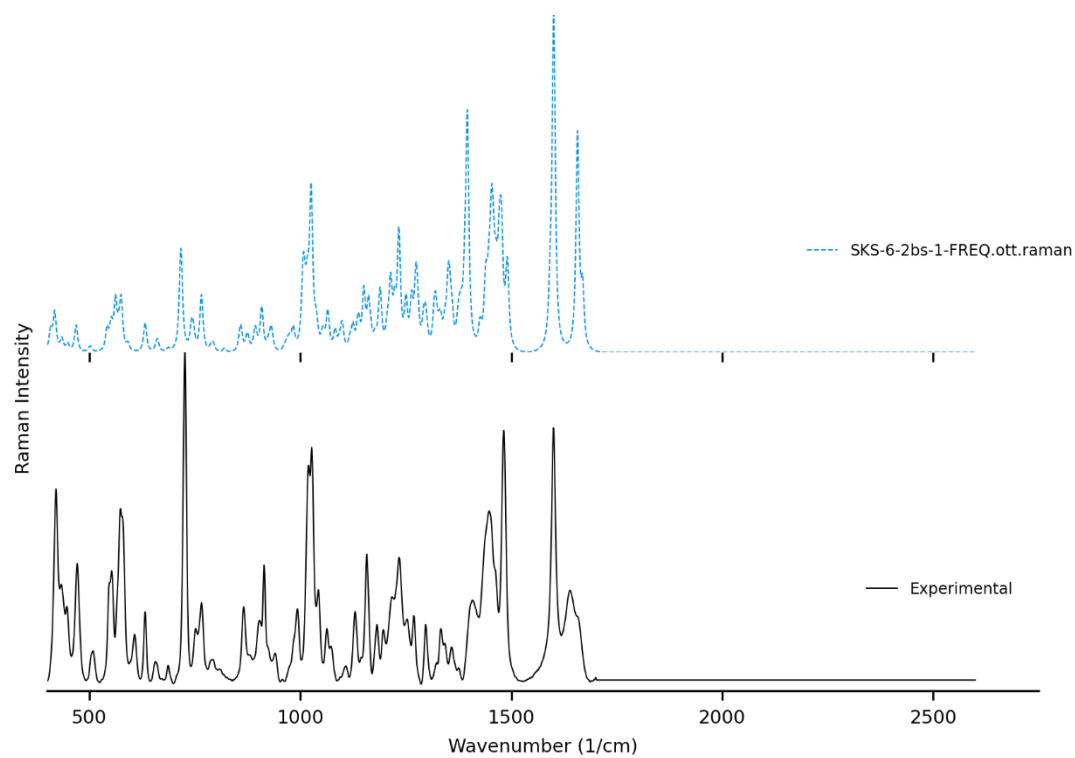
Molecule 19



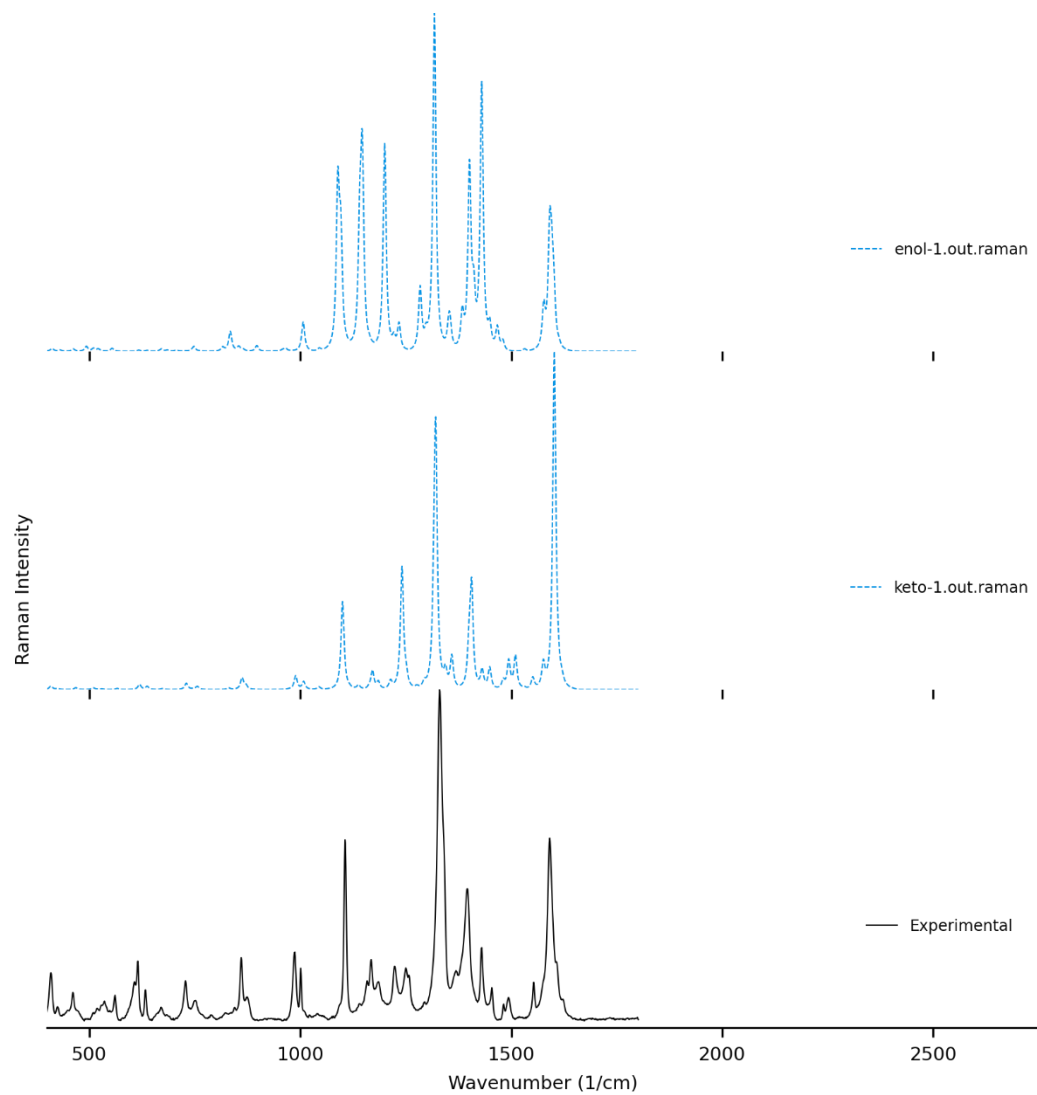
Molecule 20



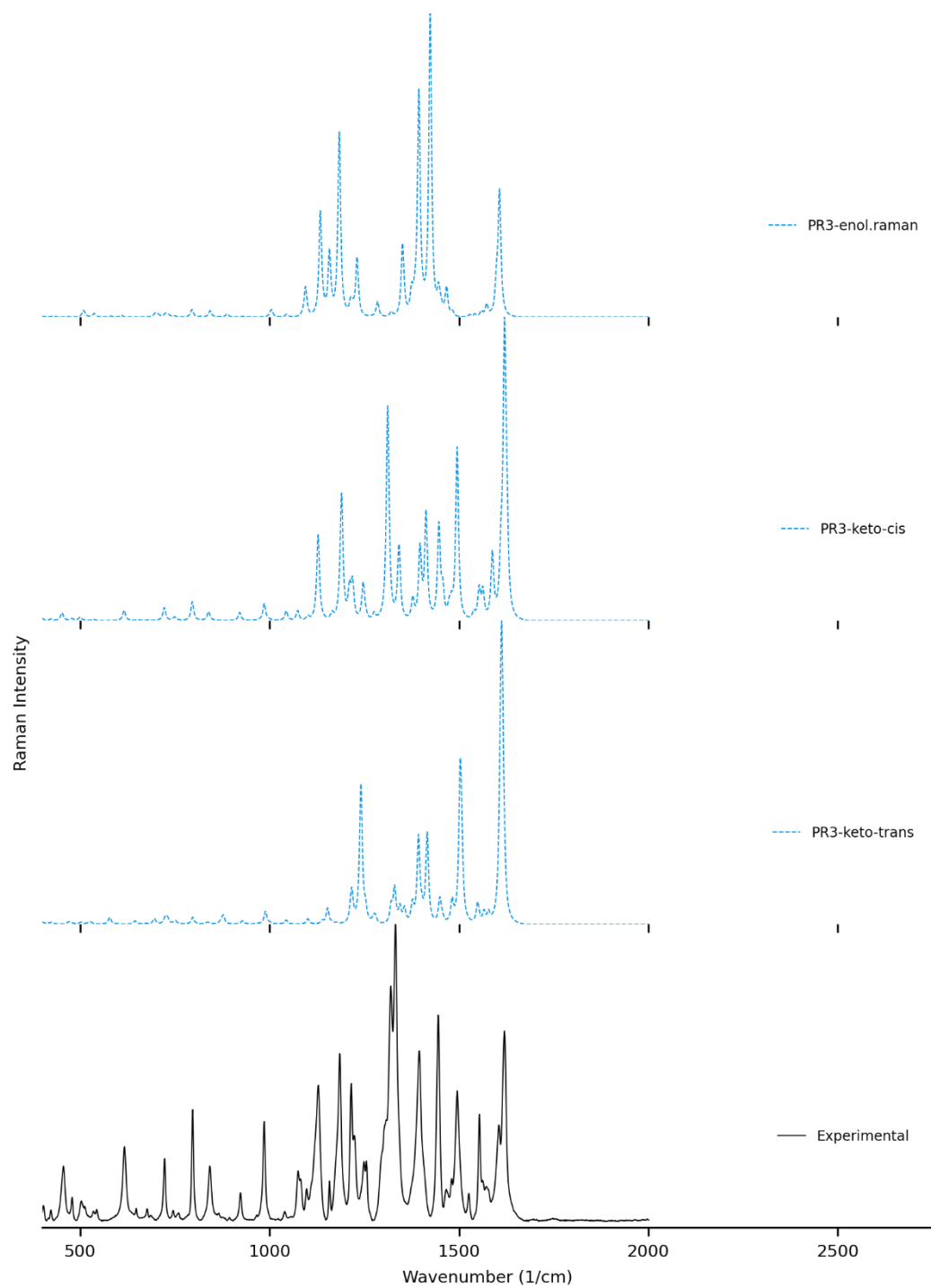
Molecule 21



Molecule **PR1**



Molecule **PR3**



5. Materials

The instrument used was a Renishaw inVia microscope (United Kingdom). The following table details the origin of the samples.

Table S1. Suppliers of analyzed powders

Molecule number	Name	Supplier
1	4,4'-di-t-butyl-2,2'-bipyridine	Sigma
2	8-Hydroxyquinoline	EMD (MERCK)
3	Norcamphor	Aldrich
4	Methylparaben	Fisher
5	BHT	Aldrich
6	Gramine	Aldrich
7	Hymecromone	Sigma-Aldrich
8	Nicotinic acid	J.T. Baker
9	Xanthinol	Anachemia
10	Phenacetin	Anachemia
11	Coumarin	Eastman
12	Pyridoxine	Sigma
13	(+)-Camphor	Aldrich
14	(+)-Borneol	Synthesis
15	(+)-Isoborneol	Synthesis
16	16	Published Synthesis ¹
17	17	Published Synthesis ¹
18	18	Published Synthesis ¹
19	19	Published Synthesis ¹
20	20	Published Synthesis ¹
21	21	Published Synthesis ¹
PR1	PR1	Database ²
PR3	PR3	Database ²

6. Synthesis of Borneol/Isoborneol

To a solution of (+)-Camphor (9.9mg, 0.065mmol) in MeOH (0.5mL) at room temperature is added NaBH₄ (1éq, 2.5mg). The resulting mixture was stirred for 1hour (monitored by TLC). After completion, 1mL of saturated aqueous solution of NH₄Cl was added followed by a dimethyl ether extraction (3x5mL). The organic layers were combined and washed with Brine and dried. Solvent removal in vacuo and column chromatography (8 : 2 Hexane : AcOEt) of the residue afforded the respective alcohols.

Isoborneol/borneol : white solid; m.p. 195,8°C; ¹H NMR (300 MHz, CDCl₃): d=0.81 and 0.84 (2 s, 3H), 0.86 (s, 3H), 0.92 (s, 3H), 1.27–1.38 (m, 3H), 1.62–1.78 (m, 2H), 1.86 (s, 1H), 2.03–2.08 (m, 1H), 2.27–2.36 (m, 1H), 3.43–3.61 (isoborneol) and 3.95–3.97 ppm (borneol) (two m, 1H); ¹³C NMR (75 MHz, CDCl₃): d=11.3, 13.3, 18.6, 20.0, 20.4, 25.8, 27.2, 28.2, 29.6, 33.8, 38.9, 40.3, 45.0, 46.3, 48.9, 49.4, 77.3, 79.8 ppm.

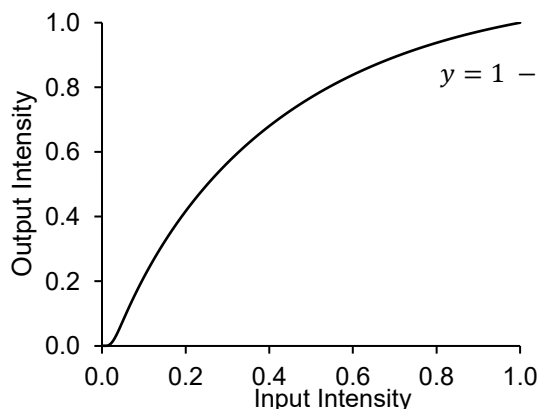
7. Measurement parameters

All the solid powders were placed on a quartz surface and measured using the 50X objective of a Renishaw inVia confocal Raman microscope.

Table S2. Instrumental parameters for the measurement of Raman spectra

Molecule number	Name	Laser wavelength	Laser intensity	Nb scans	Scan time
1	4,4'-di-t-butyl-2,2'-bipyridine	532 nm	1%	7x	6 m
2	8-Hydroxyquinoline	785 nm	0.50%	4x	2 m
3	Norcamphor	785 nm	0.50%	2x	2 m
4	Methylparaben	785 nm	0.50%	3x	100 s
5	BHT	532 nm	1%	5x	2 m
6	Gramine	532 nm	5%	6x	100 s
7	Hymecromone	785 nm	0.10%	5x	1600 s
8	Nicotinic acid	532 nm	5%	3x	16 s
9	Xanthidrol	785 nm	1%	3x	100 s
10	Phenacetin	785 nm	1%	2x	3 m
11	Coumarin	532 nm	10%	10x	25 s
12	Pyrixodine	785 nm	5%	4x	100 s
13	(+)-Camphor	785 nm	1%	3x	100 s
14	(+)-Borneol	785 nm	1%	2x	100 s
15	(+)-Isoborneol	785 nm	1%	2x	100 s
16	16	785 nm	1%	3x	60 s
17	17	785 nm	1%	8x	100 s
18	18	785 nm	0.05%	20x	800 s
19	19	785 nm	5.00%	14x	2m
20	20	785 nm	0.50%	5x	10 m
21	21	785 nm	5%	35x	100 s

8. SARA Compression function



$$y = 1 - 1.16(5x + 1.1)^{-0.7(x+1)} + 0.11(46^{-350x^2 - 9.8x - 0.0686}) \quad (S1)$$

The compression function is generated in SARA by generating 2000 equally-spaced points along equation (S1) and normalizing each axis with the min-max algorithm. The resulting points shown in **Figure S4** are used by interpolating the spectral input intensity to the corresponding output intensity.

Figure S4. Intensity compression curve used by SARA in the course of its spectral pre-processing.

9. References

1. Signo K., Canesi S., *Org. Lett.* 2022, 24, 27, 4939–4942.
2. Fremout, W.; Saverwyns, S., *Journal of Raman spectroscopy*, **2012**, 43.11, 1536-1544.

APPENDICE A

iROMA (Code du programme pour Python 3)

```
# PROGRAM HEADER
"""
Filename: iROMA.py
Author: Jason Malenfant
Date created: 2020-06-20
License: CC BY-NC-SA 4.0
"""

import base64
import io
import re

import numpy as np
import pandas as pd
import dash
import plotly.express as px
import plotly.graph_objects as go

from scipy import special
from dash import dcc, html, dash_table
from dash.dependencies import Input, Output, State
from dash.exceptions import PreventUpdate

version = 1.82
score_algo_version = 15

# -----
# Functions and placeholders
# -----

machineStateString = 'Idle.'
tableEntryList = list()
removeEntryList = list()
exportDir = r"C:\Users\hadew\Documents\Google Drive\2021 Research\DFT-Raman\iROMA"

def generate_spectrum_from_vibrations(df):
    half_width = 60
    width = (half_width * 2) + 1
    wave_max = max(df['waveData']) + 200
    wave_min = max(0, (min(df['waveData']) - 200))

    generatedSpectrum = pd.Series(index=range(wave_min, (wave_max + 1)), dtype='float64',
name='intens')
    generatedSpectrum = generatedSpectrum.fillna(value=0)

    for vibration, mu in enumerate(df['waveData']):
        if (mu - half_width) > 0:
            wavePointsArray = np.linspace((mu - half_width), (mu + half_width), width)
            curvePoints = np.linspace((0 - half_width), (0 + half_width), width)

            voigtFactorsArray = special.voigt_profile(curvePoints, 1, 4)

            factorsArray = voigtFactorsArray

            maxFactor, minFactor = np.amax(factorsArray), np.amin(factorsArray)
            normalizationRatio = maxFactor / minFactor

            inputIntensity = df['intensData'][vibration]
            outputIntensities = inputIntensity * ((factorsArray - minFactor) / normalizationRatio)
            peak = pd.Series(data=outputIntensities, index=wavePointsArray, name='intens')
            generatedSpectrum = generatedSpectrum.add(peak, fill_value=0)
```

```

return generatedSpectrum

def interpolate(spectrum, resolution):
    rangeMin = int(np rint(min(spectrum['wave'])))
    rangeMax = int(np rint(max(spectrum['wave'])))
    usrRange = range(rangeMin, rangeMax, resolution)

    outputNum = np.interp(usrRange, spectrum['wave'], spectrum['intens'])

    outputFrame = pd.DataFrame(usrRange, columns=['wave'])
    outputFrame['intens 0'] = pd.DataFrame(outputNum)

    return outputFrame

def truncate(experimental_spectrum, theoretical_spectrum):
    minExpWave = min(experimental_spectrum['wave'])
    maxExpWave = max(experimental_spectrum['wave'])
    minTheoWave = min(theoretical_spectrum['wave'])
    maxTheoWave = max(theoretical_spectrum['wave'])

    if minExpWave < minTheoWave:
        minSlice = minTheoWave
    elif minTheoWave < minExpWave:
        minSlice = minExpWave
    else:
        minSlice = minExpWave

    if maxExpWave > maxTheoWave:
        maxSlice = maxTheoWave
    elif maxTheoWave > maxExpWave:
        maxSlice = maxExpWave
    else:
        maxSlice = maxExpWave

    cutExpSpec = experimental_spectrum[(experimental_spectrum['wave'] <= maxSlice) &
                                        (experimental_spectrum['wave'] >= minSlice)]
    cutTheoSpec = theoretical_spectrum[(theoretical_spectrum['wave'] <= maxSlice) &
                                       (theoretical_spectrum['wave'] >= minSlice)]

    return cutExpSpec, cutTheoSpec

def normalize(input_spectrum):
    output = pd.DataFrame()
    for col in input_spectrum.columns:
        minValue = input_spectrum[col].min()
        maxValue = input_spectrum[col].max()
        if (not minValue) & (not maxValue):
            output[col] = 0.001
            print('Zero-only column detected. Corresponding sub-spectrum filled with 0.001.')
        else:
            output[col] = (input_spectrum[col] - minValue) / (maxValue - minValue)
    output.fillna(0.001)
    return output

def parse_spectra(usr_exp_spec, usr_theo_spec, scale_factor):
    preExpSpec = usr_exp_spec.sort_values(by=['wave'])
    preTheoSpec = usr_theo_spec.sort_values(by=['wave'])

    preTheoSpec['wave'] = preTheoSpec['wave'] * scale_factor

    rez = 1
    interpExpSpec = interpolate(preExpSpec, rez)
    interpTheoSpec = interpolate(preTheoSpec, rez)

    cutExpSpec, cutTheoSpec = truncate(interpExpSpec, interpTheoSpec)

```

```

wavesArray = np.array(range((min(cutExpSpec['wave'])), ((max(cutExpSpec['wave'])) + 1)))
del cutExpSpec['wave']
del cutTheoSpec['wave']

norExpSpec = normalize(cutExpSpec)
norTheoSpec = normalize(cutTheoSpec)

finalExpSpec = pd.DataFrame()
finalTheoSpec = pd.DataFrame()
finalExpSpec['wave'] = wavesArray
finalTheoSpec['wave'] = wavesArray
finalExpSpec['intens'] = norExpSpec['intens 0'].reset_index(drop=True)
finalTheoSpec['intens'] = norTheoSpec['intens 0'].reset_index(drop=True)

return finalExpSpec, finalTheoSpec

def score_math(sim_table_frame, lone_peaks_series):
    if sim_table_frame.empty:
        calcDataSeries = pd.Series(dtype='float64')
    else:
        nbLonePeaks = lone_peaks_series.size // 2

        n = len(sim_table_frame.index)

        thresholdDisplace, thresholdIntensity = 30, 0.65
        weightDisplace, weightLonePeaks, weightIntensity = 0.40, 0.55, 0.05

        pointsPoolDisplace, pointsPoolIntensity = n*thresholdDisplace, n*thresholdIntensity
        displaceCostArray = np.empty(0)

        for delta in sim_table_frame['displace']:
            penalty = thresholdDisplace * (1-(((thresholdDisplace ** 2 - delta ** 2) **
0.5).real/thresholdDisplace))
            displaceCostArray = np.append(displaceCostArray, penalty)

        costDisplacement, costIntensity = displaceCostArray.sum(), sim_table_frame['int'].sum()

        score_decimal = ((weightDisplace * ((pointsPoolDisplace - costDisplacement) /
pointsPoolDisplace)) +
            (weightLonePeaks * (1 - ((nbLonePeaks - 0.5) / n))) +
            (weightIntensity * ((pointsPoolIntensity - costIntensity) /
pointsPoolIntensity)))

        score = score_decimal * 100

        calcDataSeries = pd.Series(data=[score, nbLonePeaks, costDisplacement, costIntensity, n],
            index=['SCORE', 'Nb Lone Peaks', 'Total raman shift cost',
'Total intensity cost',
            'Number peak pairs'])

    return calcDataSeries

# -----
# Information for Dash
# -----
external_stylesheets = ['customstyle.css']

iroma = dash.Dash(__name__, external_stylesheets=external_stylesheets)

colors = {
    'background': '#FFFFFF',
    'text': '#000000'
}

# -----
# Dash App Layout
# -----
iroma.layout = html.Div([

```

```

html.Div(
    style={'backgroundColor': colors['background'], 'width': '85%', 'margin': '0px auto'},
    children=[html.H1(children='iROMA', style={
        'textAlign': 'center', 'color': colors['text'], 'margin-top': '2px', 'margin-bottom':
'0px'}),
        html.H3(children=str('The interactive Raman Organic Match Assessment v' +
str(version)),
            style={'textAlign': 'center', 'color': colors['text'], 'margin-top':
'2px',
                'margin-bottom': '15px'}),
        html.Div(style={'columnCount': 2, 'margin-bottom': '5px'},
            children=[
                dcc.Upload(['Drag or Click to upload WiRE Spectrum'],
                    id='uploadExpDisplay',
                    style={
                        'width': '80%',
                        'height': '60px',
                        'lineHeight': '60px',
                        'borderWidth': '1px',
                        'borderStyle': 'dashed',
                        'borderRadius': '5px',
                        'textAlign': 'center'
                    },
                ),
                dcc.Upload(['Drag or Click to upload Gaussian or ORCA Spectrum'],
                    id='uploadTheoDisplay',
                    style={
                        'width': '80%',
                        'height': '60px',
                        'lineHeight': '60px',
                        'borderWidth': '1px',
                        'borderStyle': 'dashed',
                        'borderRadius': '5px',
                        'textAlign': 'center'
                    },
                ),
            ]),
        dcc.Graph(
            id='combiSpecDisplay',
            style={},
            figure=(px.scatter(height=300))
        ),
        html.Div(id='parseExpHide', style={'display': 'none'}),
        html.Div(id='parseTheoHide', style={'display': 'none'}),
    ]),
html.Div(
    style={'columnCount': 2, 'display': 'flex'},
    children=[
        html.Div( # Column 1
            style={'display': 'inline-block', 'width': '100%', 'margin-right': '8px'},
            children=[
                html.Button('Add/del pair', id='addButtonDisplay', n_clicks=0,
                    style={'padding-top': '15px', 'padding-bottom': '15px', 'padding-
left': '30px',
                        'padding-right': '30px', 'font-size': '20px', 'margin-left':
'15%',

```

```

        'margin-right': '30px', 'margin-top': '0px', 'margin-
bottom': '10px'})
    ),
    html.Button('Calculate Score!', id='autoCalcButtonDisplay', n_clicks=0,
disabled=True,
    style={'padding-top': '5px', 'padding-bottom': '5px', 'padding-
left': '15px',
    'padding-right': '15px', 'font-size': '16px', 'margin-left':
'30px',
    'margin-right': '30px', 'margin-top': '0px', 'margin-
bottom': '10px'})
    ),
    html.Button('Export Results', id='exportScoreButton', n_clicks=0,
disabled=False,
    style={'margin-left': '30px', 'margin-right': 'auto', 'margin-top':
'0px',
    'margin-bottom': '10px'})
    ),
    dash_table.DataTable(
        id='pairTableDisplay',
        style_cell={'text-align': 'center'},
        columns=[{'id': 'wexp', 'name': 'Experimental peak (1/cm)', 'type':
'numeric'},
        {'id': 'iexp', 'name': 'Exp. intensity', 'type': 'numeric'},
        {'id': 'wtheo', 'name': 'Calculated peak (1/cm)', 'type':
'numeric'},
        {'id': 'itheo', 'name': 'Calc. intensity', 'type': 'numeric'},
        ],
        editable=False,
        style_as_list_view=True
    ), ]
),
html.Div(
    style={'display': 'inline-block', 'width': '100%', 'margin-left': '8px'},
    children=[
        html.Button('RESET PEAK PICKING', id='resetPairsButton',
style={'display': 'block', 'align': 'right', 'border-color': 'red',
'margin-left': 'auto', 'background-color': '#FFCCAA',
'padding-top': '4px',
'padding-bottom': '4px', 'padding-left': '5px', 'padding-
right': '5px',
'font-size': '13px'}),
        html.Div(id='machineStateDisplay',
children=(html.H1(machineStateString)),
style={'margin-block-start': '0em', 'padding-top': '0px', 'margin-
top': '0px'})
    ),
    dcc.Dropdown(
        id='modeSelectDisplay',
        options=[
            {'label': 'Create peak PAIR', 'value': 'PAR'},
            {'label': 'Pick LONE peak', 'value': 'LON'},
            {'label': 'DELETE peak pair', 'value': 'DEL'},
        ],
        value='PAR'
    ),
    html.Br(style={'font-size': '5px'}),
    html.Div(id='miscDataDisplay',
children=[
        'List of lone peaks: ',
        html.Div(id='lonePeaksDisplay', children="[No peaks]",
style={'word-break': 'break-word'})
    ],

```

```

                style={'font-size': '22px'}
            ),
            html.Div(id='checklist1Display', children='Checklist placeholder.',
style={'font-size': '15px'}),
            html.Div(style={'columnCount': 2, 'display': 'flex'},
children=[
                html.Div(children=[
                    html.Div('Last selected peaks for pairing: ',
style={'font-size': '15px'}),
                    html.Div(id='tableEntryDisplay', children='[No peaks]',
style={'font-size': '15px'}),
                    ], style={'margin-left': '0px', 'margin-right':
'50px'}),
                html.Div(children=[
                    html.Div('Last selected peak for removal: ',
style={'font-size': '15px'}),
                    html.Div(id='removalEntryDisplay', children='[No
peaks]',
style={'font-size': '15px'}),
                    ], style={'margin-left': '50px', 'margin-right':
'auto'})
            ]),
            html.Br(),
            html.Hr(),
            html.Div(style={'display': 'block', 'width': '80%', 'margin-left': 'auto',
'margin-right': 'auto'},
children=[dcc.Slider(id='scaleFactorSlider', min=0.80, max=1.10,
value=0.98, step=0.01,
marks={
                1: {'label': '1.00'},
                1.10: {'label': '1.10'},
                0.80: {'label': '0.80'},
                0.98: {'label': '0.98'}
            }
),
            html.H3(id='sliderIndicatorDisplay',
style={'text-align': 'center', 'margin-block-start':
'0em'})
        ],
    ),
    html.Br(),
    html.Div(id='tableDataHide', style={'display': 'none'}),
    html.Div(id='autoCalcDataHide', style={'display': 'none'}),
    ])
    ],
    ),
    html.Div(style={'position': 'relative'},
children=[
        html.Hr(),
        html.H3('Jason Malenfant, 2020-2022', style={'margin-block-start': '0em'})
    ])
    ],
    style={'width': '98%', 'margin-left': 'auto', 'margin-right': 'auto'}
)

```

```

# Dash App Callbacks
# -----
@iroma.callback(
    [Output('machineStateDisplay', 'children'),
     Output('autoCalcButtonDisplay', 'disabled')],
    [Input('combiSpecDisplay', 'clickData'),
     Input('addButtonDisplay', 'n_clicks'),
     Input('resetPairsButton', 'n_clicks'),
     Input('modeSelectDisplay', 'value'),
     Input('tableEntryDisplay', 'children'),
     Input('pairTableDisplay', 'selected_cells'),
     Input('autoCalcDataHide', 'children'),
     Input('exportScoreButton', 'n_clicks'),
     Input('lonePeaksDisplay', 'children')],
    [State('lonePeaksDisplay', 'children'),
     State('tableDataHide', 'children'),
     State('removalEntryDisplay', 'children'),
     State('autoCalcButtonDisplay', 'disabled'),
     State('autoCalcDataHide', 'children'),
     State('uploadExpDisplay', 'filename'),
     State('uploadTheoDisplay', 'filename'),
     State('scaleFactorSlider', 'value')]
)
def update_machine_state(
    combi_click_data, add_btn_clicks, reset_btn_clicks, mode_value,
    table_entry_children, pair_table_selection, auto_calc_data_children, export_btn_clicks,
lone_peaks_children,
    lone_peaks_state, table_data_children, removal_entry_children, state_calc_disabled,
    auto_calc_data, exp_filename, theo_filename, scale_factor_value):
    calcBtnDisabled = state_calc_disabled
    stateMessage = 'NullMessage'
    ctx = dash.callback_context
    triggerID = [p['prop_id'] for p in dash.callback_context.triggered][0]
    if not ctx:
        stateMessage = 'Idle. Start by clicking on an experimental peak.'
        calcBtnDisabled = True
    elif 'tableEntryDisplay' in triggerID:
        calcBtnDisabled = False
        if mode_value == 'PAR':
            try:
                tableEntryRead = pd.read_json(table_entry_children, orient='records', typ='series')
            except ValueError:
                tableEntryRead = pd.DataFrame()
            if len(tableEntryRead) == 0:
                stateMessage = 'No peak or spectra present. Please upload spectra.'
                calcBtnDisabled = False
            elif len(tableEntryRead) == 2:
                expWave = str(int(tableEntryRead[0]))
                stateMessage = 'Experimental peak at ' + expWave + 'cm-1 parsed. Please add
calculated peak.'
                calcBtnDisabled = state_calc_disabled
            elif len(tableEntryRead) == 4:
                pairText = str(int(tableEntryRead[0])) + ' / ' + str(int(tableEntryRead[2]))
                stateMessage = 'Ready to add pair ' + pairText + 'cm-1 to the table. Click \"Add/del
pair\".'
                calcBtnDisabled = state_calc_disabled
            else:
                'Warning: tableEntryRead has an weird number of entries.'

        elif mode_value == 'LON':
            numLonePeaks = int(len(pd.read_json(lone_peaks_state, orient='records',
typ='series').values) // 2)
            stateMessage = 'Peak should be in lone peaks list. Currently ' + str(numLonePeaks) + '
lone peaks total.'
            calcBtnDisabled = False

        elif mode_value == 'DEL':
            try:
                removalEntryRead = pd.read_json(removal_entry_children, orient='records',
typ='series')

```

```

        removalEntryReadList = removalEntryRead.to_list() # Screw learning JSON libs.
    except ValueError:
        removalEntryReadList = list()

    try:
        remWave = removalEntryReadList[0]
        stateMessage = 'Ready to delete ' + str(int(remWave)) + 'cm-1. Click \"Add/del
pair\" button.'
        calcBtnDisabled = False
    except IndexError:
        stateMessage = 'ERROR: Experimental peak required to delete a pair.'

    else:
        stateMessage = 'Peak pairing operation not supported with this mode.'

    elif 'addButtonDisplay' in triggerID:
        if mode_value == 'PAR':
            try:
                pairTableDataParsed = pd.read_json(table_data_children, orient='columns',
typ='frame')
                stateMessage = 'Added peaks to data table. When ready, click \"Calculate Score!\",'
            except ValueError:
                raise PreventUpdate
            calcBtnDisabled = False

        elif mode_value == 'LON':
            stateMessage = 'Button does nothing in LON mode. Peak is added by clicking on datapoint.'

        elif mode_value == 'DEL':
            try:
                pairTableDataParsed = pd.read_json(table_data_children, orient='columns',
typ='frame')
                stateMessage = 'Removed selected pair from data table.'
            except ValueError:
                raise PreventUpdate
            except KeyError:
                raise PreventUpdate
            calcBtnDisabled = state_calc_disabled

        else:
            stateMessage = 'Idle/No message.'
            calcBtnDisabled = False

    elif 'resetPairsButton' in triggerID:
        stateMessage = 'All tables RESET. Ready to start anew.'
        calcBtnDisabled = True

    elif 'lonePeaksDisplay' in triggerID:
        try:
            numLonePeaks = int(len(pd.read_json(lone_peaks_children, orient='records',
typ='series').values) // 2)
            stateMessage = 'Initialized.'
        except ValueError:
            raise PreventUpdate
        calcBtnDisabled = False

    elif 'combiSpecDisplay' in triggerID:
        if mode_value == 'PAR':
            stateMessage = ''
            calcBtnDisabled = False

        elif mode_value == 'LON':
            numLonePeaks = int(len(pd.read_json(lone_peaks_children, orient='records',
typ='series').values) // 2)
            stateMessage = 'Peak should be in lone peaks list. Currently ' + str(numLonePeaks) + '
lone peaks total.'
            calcBtnDisabled = False

        elif mode_value == 'DEL':
            stateMessage = ''

```



```

elif 'calcButtonDisplay' in triggerID:
    calcBtnDisabled = True
    stateMessage = 'Calculating score...'

elif 'autoCalcDataHide' in triggerID:
    calcBtnDisabled = True
    if not auto_calc_data_children:
        stateMessage = 'Idle. Start by clicking on an experimental peak.'

    else:
        try:
            calcReportSeries = pd.read_json(auto_calc_data_children, orient='records',
            typ='series')
            score_string = str(round(calcReportSeries[0], 2))
            stateMessage = 'Score: ' + score_string + '%'
        except ValueError:
            stateMessage = 'Unable to provide score.'
        except KeyError:
            raise PreventUpdate

elif 'modeSelectDisplay' in triggerID:
    if mode_value == 'PAR':
        stateMessage = 'Select experimental peak to begin.'

    elif mode_value == 'LON':
        stateMessage = 'Click on peak to add to lone peaks list.'

    elif mode_value == 'DEL':
        stateMessage = 'Select an experimental peak on graph to delete its pair.'

    else:
        stateMessage = 'CRITICAL ERROR: Mode unrecognized.'

    calcBtnDisabled = state_calc_disabled

elif 'exportScoreButton' in triggerID:

    try:
        calcResultsRead = pd.read_json(auto_calc_data, orient='records', typ='series')

        calcResultsList = calcResultsRead.to_list()
        calcExportString = ''
        exp_filename_string = str(exp_filename)
        theo_filename_string = str(theo_filename)
        calcResultsList.append(exp_filename_string[:-4])
        calcResultsList.append(theo_filename_string[:-4])
        for items in calcResultsList:
            calcExportString = calcExportString + str(items) + ','
        calcExportString = calcExportString + str(scale_factor_value)

        stateMessage = calcExportString

        pairTableDataRetrieve = pd.read_json(table_data_children, orient='columns',
        typ='frame')
        pairTableDataRetrieve['displace'] = pairTableDataRetrieve['wexp'] -
        pairTableDataRetrieve['wtheo']
        pairTableDataRetrieve['int'] = pairTableDataRetrieve['iexp'] -
        pairTableDataRetrieve['itheo']
        pairTableDataRetrieve = pairTableDataRetrieve.abs()
        tableDataExportJson = pairTableDataRetrieve.to_json(orient='columns')

        correlationName = exp_filename_string + "-" + theo_filename_string

        pairFilePath = exportDir + "\\\" + correlationName + "-par.json"
        pairFileInstance = open(pairFilePath, "w")
        pairFileInstance.write(tableDataExportJson)
        pairFileInstance.close()

        loneFilePath = exportDir + "\\\" + correlationName + "-lon.json"
        loneFileInstance = open(loneFilePath, "w")
        loneFileInstance.write(lone_peaks_children)

```

```

        loneFileInstance.close()

    except ValueError:
        raise PreventUpdate

    else:
        stateMessage = 'Idle/No specific message.'
        calcBtnDisabled = True

    if stateMessage == '':
        stateMessage = 'Idle/no message.'

    return html.H1(stateMessage), calcBtnDisabled

# -----
# End of Machine State refresh function
# -----

@iroma.callback(
    Output('sliderIndicatorDisplay', 'children'),
    [Input('scaleFactorSlider', 'value')]
)
def update_slider_indicator_text(scale_factor_value):
    return 'A correction factor of {} is applied to the theoretical spectrum
wavenumbers.'.format(scale_factor_value)

@iroma.callback(
    [Output('parseExpHide', 'children'),
    Output('uploadExpDisplay', 'children')],
    [Input('uploadExpDisplay', 'contents')],
    [State('uploadExpDisplay', 'filename')]
)
def input_exp_spec(upload_exp_contents, upload_exp_filename):
    print(upload_exp_filename)
    print('Processing spectrum...')
    regex = re.compile('^.*?,'')
    try:
        uploadRead = regex.sub('', upload_exp_contents)
    except TypeError:
        raise PreventUpdate
    decodedBytes = base64.b64decode(uploadRead)
    decodedString = decodedBytes.decode('utf-8')
    fakeFile = io.StringIO(decodedString)

    try:
        if upload_exp_filename.endswith('.txt'):
            if decodedString.startswith('#Wave          #Intensity'):
                parseExpWrite = pd.read_csv(fakeFile, names=['wave', 'intens'], skiprows=(0,),
                    sep='\t', dtype='float64')
            else:
                parseExpWrite = pd.read_csv(fakeFile, names=['wave', 'intens'],
                    dialect="excel", comment='#', dtype='float64')

        elif upload_exp_filename.endswith('.peaks'):
            parseExpWrite = pd.read_csv(fakeFile, names=['wave', 'intens'], comment='#',
                usecols=[1, 2], sep='\t', dtype='float64')

        else:
            parseExpWrite = pd.read_csv(fakeFile, names=['wave', 'intens'],
                dialect="excel", comment='#', dtype='float64')

    except Exception:
        raise PreventUpdate

    parseExpWriteJson = parseExpWrite.to_json(orient='columns')
    outputString = f'{upload_exp_filename} processed as experimental spectrum!'
    return parseExpWriteJson, outputString

```

```

@iroma.callback(
    [Output('parseTheoHide', 'children'),
     Output('uploadTheoDisplay', 'children')],
    [Input('uploadTheoDisplay', 'contents')],
    [State('uploadTheoDisplay', 'filename')]
)
def input_theo_spec(upload_theo_contents, upload_theo_filename):
    print(upload_theo_filename)
    print('Processing spectrum...')
    regex = re.compile('^.*?,' )
    try:
        uploadRead = regex.sub(',', upload_theo_contents)
    except TypeError:
        raise PreventUpdate
    decodedBytes = base64.b64decode(uploadRead)
    decodedString = decodedBytes.decode('utf-8')
    virtualFile = io.StringIO(decodedString)

    if upload_theo_filename.endswith('.dat'):
        parseTheoWrite = pd.read_csv(virtualFile, names=['wave', 'intens'], sep='\s+',
dtype='float64')

    elif upload_theo_filename.endswith('.stk'):
        parsePeakData = pd.read_csv(virtualFile, names=['waveData', 'intensData'], sep='\s+',
dtype='float64')
        parsePeakData['waveData'] = parsePeakData['waveData'].astype(int)

        print("Generating actual spectrum from peak data...")
        parseTheoSeries = generate_spectrum_from_vibrations(parsePeakData)
        parseTheoWrite = pd.DataFrame()
        parseTheoWrite['wave'] = parseTheoSeries.index.values
        parseTheoWrite['intens'] = parseTheoSeries

    elif upload_theo_filename.endswith('.txt'):
        if decodedString.startswith('# Raman Activity Spectrum'):
            parseTheoWrite = pd.read_csv(virtualFile, names=['wave', 'intens'], sep='\s+',
dtype='float64', comment='#', usecols=[0, 1])
        else:
            parseTheoWrite = pd.read_csv(virtualFile, names=['wave', 'intens'],
dialect="excel", comment='#', dtype='float64')

    elif upload_theo_filename.endswith('.peaks'):
        parseTheoWrite = pd.read_csv(virtualFile, names=['wave', 'intens'], comment='#',
usecols=[1, 2], sep='\t', dtype='float64')

    else:
        parseTheoWrite = pd.read_csv(virtualFile, names=['wave', 'intens'],
dialect="excel", comment='#', dtype='float64')
        parseTheoWriteJson = parseTheoWrite.to_json(orient='columns')
        outputString = f'{upload_theo_filename} processed as theoretical spectrum!'
        return parseTheoWriteJson, outputString

@iroma.callback(
    Output('combiSpecDisplay', 'figure'),
    [Input('parseExpHide', 'children'),
     Input('parseTheoHide', 'children'),
     Input('autoCalcButtonDisplay', 'n_clicks'),
     Input('combiSpecDisplay', 'clickData'),
     Input('scaleFactorSlider', 'value')],
    [State('tableDataHide', 'children'),
     State('combiSpecDisplay', 'figure'),
     State('modeSelectDisplay', 'value'),
     State('scaleFactorSlider', 'value')]
)
def create_combi_fig(parse_exp_children, parse_theo_children, auto_calc_clicks, combi_spec_click,
scale_factor_value,
                    table_data_children, combi_fig, mode_select_value, scale_factor_state):
    triggerID = [p['prop_id'] for p in dash.callback_context.triggered][0]
    figCombiSpectra = combi_fig

```

```

scaleFactor = scale_factor_state

if 'parse' in triggerID or 'scaleFactorSlider' in triggerID:
    try:
        expRead = pd.read_json(parse_exp_children, orient='columns', typ='frame')
        theoRead = pd.read_json(parse_theo_children, orient='columns', typ='frame')
    except Exception:
        raise PreventUpdate

    processedExpSpec, processedTheoSpec = parse_spectra(expRead, theoRead, scaleFactor)
    combiSpectra = pd.DataFrame()
    combiSpectra['wave'] = processedExpSpec['wave']
    combiSpectra['Experimental'] = processedExpSpec["intens"]
    combiSpectra['Calculated'] = processedTheoSpec["intens"]
    figCombiSpectra = px.scatter(
        combiSpectra, x="wave", y=["Experimental", "Calculated"],
        labels=dict(index="Shift", value="Relative Raman Intensity", variable="Spectra"),
        title="Compared spectra")
    figCombiSpectra.update_layout(margin=dict(l=10, r=10, t=0, b=0), height=400,
        xaxis_title="Wavenumber (1/cm)")

elif 'autoCalcButtonDisplay' in triggerID:
    print('AUTOCalc click nb ' + str(auto_calc_clicks))
    try:
        tableDataRead = pd.read_json(table_data_children, orient='columns', typ='frame')
    except ValueError:
        raise PreventUpdate

    expRead = pd.read_json(parse_exp_children, orient='columns', typ='frame')
    theoRead = pd.read_json(parse_theo_children, orient='columns', typ='frame')
    processedExpSpec, processedTheoSpec = parse_spectra(expRead, theoRead, scaleFactor)
    combiSpectra = pd.DataFrame()
    combiSpectra['wave'] = processedExpSpec['wave']
    combiSpectra['Experimental'] = processedExpSpec["intens"]
    combiSpectra['Calculated'] = processedTheoSpec["intens"]
    figCombiSpectra = px.scatter(
        combiSpectra, x="wave", y=["Experimental", "Calculated"],
        labels=dict(index="Shift", value="Relative Raman Intensity", variable="Spectra"),
        title="Compared spectra")
    figCombiSpectra.update_layout(margin=dict(l=10, r=10, t=0, b=0), height=400,
        xaxis_title="Wavenumber (1/cm)")

    lineColorString = 'rgb(198,0,251)'

    for i, row in tableDataRead.iterrows():
        waveDots = list()
        waveDots.append(row['wexp'])
        waveDots.append(row['wtheo'])
        intDots = list()
        intDots.append(row['iexp'])
        intDots.append(row['itheo'])
        line_name = "Pair " + str(i)

        figCombiSpectra.add_trace(go.Scatter(x=waveDots, y=intDots, name=line_name,
mode='lines',
line=dict(color=lineColorString)))

else:
    raise PreventUpdate

return figCombiSpectra

@iroma.callback(
    [Output('checkboxList1Display', 'children'),
    Output('tableEntryDisplay', 'children'),
    Output('lonePeaksDisplay', 'children'),
    Output('removalEntryDisplay', 'children')],
    [Input('combiSpecDisplay', 'clickData'),
    Input('resetPairsButton', 'n_clicks')],

```

```

        Input('uploadTheoDisplay', 'children'),
        Input('uploadExpDisplay', 'children']],
        [State('modeSelectDisplay', 'value'),
        State('lonePeaksDisplay', 'children'),
        State('tableEntryDisplay', 'children'),
        State('removalEntryDisplay', 'children']]
    )
def reaction_to_peak_click(click_data, reset_btn_clicks, up_theo_btn_content, up_exp_btn_content,
                           mode_value, lone_peaks_children, table_entry_children,
removal_entry_children):
    checklist1 = []
    tableEntryJson = table_entry_children
    lonePeaksWriteJson = lone_peaks_children
    removeEntryJson = removal_entry_children
    triggerID = [p['prop_id'] for p in dash.callback_context.triggered][0]
    if 'reset' in triggerID:
        print(f'Resetting lone peaks list. \#{reset_btn_clicks}')
        lonePeaksWrite = pd.Series(dtype='float64')
        lonePeaksWriteJson = lonePeaksWrite.to_json(orient='records')
        return checklist1, tableEntryJson, lonePeaksWriteJson, removeEntryJson

    elif 'uploadTheoDisplay' in triggerID:
        print(f'Resetting lone peaks list. due tu spectrum upload.')
        lonePeaksWrite = pd.Series(dtype='float64')
        lonePeaksWriteJson = lonePeaksWrite.to_json(orient='records')
        return checklist1, tableEntryJson, lonePeaksWriteJson, removeEntryJson

    elif 'uploadExpDisplay' in triggerID:
        print(f'Resetting lone peaks list. due tu spectrum upload.')
        lonePeaksWrite = pd.Series(dtype='float64')
        lonePeaksWriteJson = lonePeaksWrite.to_json(orient='records')
        return checklist1, tableEntryJson, lonePeaksWriteJson, removeEntryJson

    if click_data is not None:
        selWave = click_data['points'][0]['x']
        selInt = click_data['points'][0]['y']

        if mode_value == 'PAR':

            if click_data['points'][0]['curveNumber'] == 0:
                if not tableEntryList:
                    tableEntryList.append(selWave)
                    tableEntryList.append(selInt)
                    message = "Selected " + str(int(selWave)) + "cm-1 with intensity " +
str(round(selInt, 4)) + "."
                    listReady = False
                else:
                    tableEntryList.clear()
                    tableEntryList.append(selWave)
                    tableEntryList.append(selInt)
                    message = "Replaced exp value with " + str(int(selWave)) + "cm-1."
                    listReady = False

                print(message)

            elif click_data['points'][0]['curveNumber'] == 1:
                if not tableEntryList:
                    message = "Error. Please select an experimental peak first."
                    print(message)
                    print()
                    listReady = False
                else:
                    if len(tableEntryList) == 2:
                        tableEntryList.append(selWave)
                        tableEntryList.append(selInt)
                        message = "Added " + str(int(selWave)) + " cm-1 as calc peak. READY to
append"
                        listReady = True
                    else:
                        message = "Unable to add peak because of list length."
                        listReady = True

```

```

        print(message)

    else:
        print('ERROR: Curve number does not exist.')
        listReady = False

    readinessFeedback = list()
    if tableEntryList:
        readinessFeedback.append('EPL')
        if listReady:
            readinessFeedback.append('CPL')

    checklist1 = dcc.Checklist(
        options=[
            {'label': 'Experimental Peak Loaded', 'value': 'EPL', 'disabled': True},
            {'label': 'Calculated Peak Loaded', 'value': 'CPL', 'disabled': True}
        ],
        value=readinessFeedback,
        id='checklist1'
    )

    tableEntrySeries = pd.Series(tableEntryList, dtype='float64')
    tableEntryJson = tableEntrySeries.to_json(orient='records')

    elif mode_value == 'LON':

        lonePeaksSeries = pd.Series([selWave, selInt])
        if lone_peaks_children == '[No peaks]':
            lonePeaksWrite = lonePeaksSeries
            print('Created lonePeaks Series and appended peak ' + str(selWave) + ' , ' +
str(selInt))
        else:
            print('Not not peaks')
            lonePeaksRead = pd.read_json(lone_peaks_children, orient='records', typ='series')
            lonePeaksWrite = lonePeaksRead.append(lonePeaksSeries, ignore_index=True)

            lonePeaksWriteJson = lonePeaksWrite.to_json(orient='records')

    elif mode_value == 'DEL':

        if click_data['points'][0]['curveNumber'] == 0:
            if not removeEntryList:
                removeEntryList.append(selWave)
                removeEntryList.append(selInt)
                message = "Selected " + str(int(selWave)) + "cm-1 with intensity " +
str(round(selInt, 4)) + "."
            else:
                removeEntryList.clear()
                removeEntryList.append(selWave)
                removeEntryList.append(selInt)
                message = "Replaced exp value with " + str(int(selWave)) + "cm-1."
            print(message)

        elif click_data['points'][0]['curveNumber'] == 1:
            message = "Error. An experimental peak must be selected."
            print(message)
            print()

        else:
            print('ERROR: Curve number does not exist.')

        removeEntrySeries = pd.Series(removeEntryList, dtype='float64')
        removeEntryJson = removeEntrySeries.to_json(orient='records')

    else:
        print('Warning: Mode Undefined, or app is booting up.')

return checklist1, tableEntryJson, lonePeaksWriteJson, removeEntryJson

```

```

@iroma.callback(
    Output('tableDataHide', 'children'),
    [Input('addButtonDisplay', 'n_clicks'),
     Input('resetPairsButton', 'n_clicks'),
     Input('uploadTheoDisplay', 'children'),
     Input('uploadExpDisplay', 'children')],
    [State('modeSelectDisplay', 'value'),
     State('tableEntryDisplay', 'children'),
     State('tableDataHide', 'children'),
     State('removalEntryDisplay', 'children')],
    prevent_initial_call=True
)
def update_hidden_data_table(add_btn_clicks, reset_btn_clicks, up_theo_btn_content,
up_exp_btn_content, mode_value, table_entry_children, table_data_children,
removal_entry_children):
    tableDataWriteJson = []

    if add_btn_clicks != 0:
        triggerID = [p['prop_id'] for p in dash.callback_context.triggered][0]
        if 'addButtonDisplay' in triggerID:
            if mode_value == 'PAR':
                try:
                    tableDataRead = pd.read_json(table_data_children, orient='columns',
typ='frame')
                    tableDataWrite = tableDataRead
                except NameError:
                    print('Creating new tableDataWrite.')
                    print()
                    tableDataWrite = pd.DataFrame(columns=['wexp', 'iexp', 'wtheo', 'itheo'])
                except ValueError:
                    print('Creating new tableDataWrite.')
                    print()
                    tableDataWrite = pd.DataFrame(columns=['wexp', 'iexp', 'wtheo', 'itheo'])

                try:
                    tableEntryRead = pd.read_json(table_entry_children, orient='records',
typ='series')
                    tableEntryReadList = tableEntryRead.to_list()
                except ValueError:
                    raise PreventUpdate

                if len(tableEntryReadList) < 4:
                    print('Not enough peaks to add. Did not append table.')
                    print()
                elif len(tableEntryReadList) == 4:
                    if tableEntryReadList[0] not in tableDataWrite['wexp'].values:
                        if tableEntryReadList[2] not in tableDataWrite['wtheo'].values:
                            print('Appending data table with new peak.')
                            print()
                            tableDataWrite.loc[len(tableDataWrite)] = tableEntryReadList
                        else:
                            print()
                            print('Calculated value detected in list. Please delete it first to
reassign.')
                            print()
                            tableDataWriteJson = table_data_children
                    else:
                        print()
                        print('Experimental value detected in list. Please delete it first to
reassign.')
                        print()
                        tableDataWriteJson = table_data_children
                else:
                    print('Warning: entries in tableEntryRead is over 4. Did not append table.')

                tableDataWriteJson = tableDataWrite.to_json(orient='columns')

            elif mode_value == 'DEL':
                try:

```

```

        tableDataRead = pd.read_json(table_data_children, orient='columns',
typ='frame')
        tableDataWrite = tableDataRead
    except NameError:
        print('Creating new tableDataWrite.')
        print()
        tableDataWrite = pd.DataFrame(columns=['wexp', 'iexp', 'wtheo', 'itheo'])
    except ValueError:
        print('Creating new tableDataWrite.')
        print()
        tableDataWrite = pd.DataFrame(columns=['wexp', 'iexp', 'wtheo', 'itheo'])

    try:
        removalEntryRead = pd.read_json(removal_entry_children, orient='records',
typ='series')
        removalEntryReadList = removalEntryRead.to_list()
    except ValueError:
        raise PreventUpdate

    if len(removalEntryReadList) < 2:
        print()
    elif len(removalEntryReadList) == 2:
        if removalEntryReadList[0] in tableDataWrite['wexp'].values:
            print('Removing pair from data table.')
            print()
            tableDataWrite = tableDataWrite[tableDataWrite.wexp !=
removalEntryReadList[0]]
            tableDataWrite.reset_index(drop=True, inplace=True)
        else:
            tableDataWriteJson = table_data_children
    else:
        print('Warning: entries in removalEntryRead is over 2. Did not append table.')

    tableDataWriteJson = tableDataWrite.to_json(orient='columns')

    else:
        tableDataWriteJson = table_data_children

    elif 'resetPairsButton' in triggerID:
        tableDataWriteJson = []
        print(f'Reset button press triggered erasing tableDataHide. Click number
{reset_btn_clicks}')

    elif 'uploadExpDisplay' in triggerID:
        tableDataWriteJson = []

    elif 'uploadTheoDisplay' in triggerID:
        tableDataWriteJson = []

    else:
        tableDataWriteJson = table_data_children

    return tableDataWriteJson

@iroma.callback(
    Output('pairTableDisplay', 'data'),
    [Input('tableDataHide', 'children'),
    Input('resetPairsButton', 'n_clicks')]
)
def update_pairings_table(table_data_children, reset_pairs_clicks):
    triggerID = [p['prop_id'] for p in dash.callback_context.triggered][0]
    if 'resetPairsButton' in triggerID:
        print(f'Reset button action provoked pairTableDataWrite erasing. Click number
{reset_pairs_clicks}.')
        pairTableDataWrite = []

    elif not table_data_children:
        pairTableDataWrite = []

    else:

```



```

        dataRead = pd.read_json(table_data_children, orient='columns', typ='frame')
        pairTableDataWrite = dataRead.to_dict('records')

    return pairTableDataWrite

@iroma.callback(
    Output('autoCalcDataHide', 'children'),
    [Input('autoCalcButtonDisplay', 'n_clicks')],
    [State('lonePeaksDisplay', 'children'),
     State('tableDataHide', 'children')]
)
def auto_score_calc(auto_calc_btn_clicks, lone_peaks_children, table_data_children):
    triggerID = [p['prop_id'] for p in dash.callback_context.triggered][0]

    if 'autoCalcButtonDisplay' in triggerID:
        print(f'AUTO-CALCULATING SCORE! Iteration #{auto_calc_btn_clicks}!')

        try:
            pairTableDataParsed = pd.read_json(table_data_children, orient='columns', typ='frame')
        except ValueError:
            print("Tried to calculate score but no data present!")
            raise PreventUpdate

        pairTableDataParsed['displace'] = pairTableDataParsed['wexp'] -
pairTableDataParsed['wtheo']
        pairTableDataParsed['int'] = pairTableDataParsed['iexp'] - pairTableDataParsed['itheo']
        pairTableDataParsed = pairTableDataParsed.abs()

        try:
            lonePeaksParsed = pd.read_json(lone_peaks_children, orient='records', typ='series')
        except ValueError:
            lonePeaksParsed = pd.Series(dtype='float64')

        calcDataWrite = score_math(pairTableDataParsed, lonePeaksParsed)
        print(calcDataWrite)

    else:
        calcDataWrite = pd.Series(dtype='float64')

    calcDataWriteJson = calcDataWrite.to_json(orient='records')

    return calcDataWriteJson

# -----
# Server start statement
# -----
if __name__ == '__main__':
    print('👋 🌟🌟 👋 hello!')
    iroma.run_server(debug=False)

# -----
# Server shutdown
# -----
print('👋🌟🌟 👋 server closing')
```

RÉFÉRENCES

- Aprà, E., Bhattarai, A., Baxter, E., Wang, S., Johnson, G. E., Govind, N. et El-Khoury, P. Z. (2020). Simplified Ab Initio Molecular Dynamics-Based Raman Spectral Simulations. *Applied Spectroscopy*, 74(11), 1350-1357. <https://doi.org/10.1177/0003702820923392>
- Aroca, R. (2006). *Surface-Enhanced Vibrational Spectroscopy*. John Wiley & Sons Ltd.
- ATSM, I. (2015). Standard Guide for Use of Spectral Searching by Curve Matching Algorithms with Data Recorded Using Mid-Infrared Spectroscopy. <https://doi.org/10.1520/e2310-04r15>
- Bajusz, D., Racz, A. et Heberger, K. (2015). Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J Cheminform*, 7, 20. <https://doi.org/10.1186/s13321-015-0069-3>
- Barone, V. et Cossi, M. (1998). Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *The Journal of Physical Chemistry A*, 102(11), 1995-2001. <https://doi.org/10.1021/jp9716997>
- Batista, A. J., Vianna, P. G., Ribeiro, H. B., Matos, C. J. S. et Gomes, A. S. L. (2019, Jun 20). QR code micro-certified gemstones: femtosecond writing and Raman characterization in Diamond, Ruby and Sapphire. *Sci Rep*, 9(1), 8927. <https://doi.org/10.1038/s41598-019-45405-7>
- Becke, A. D. (1993a, 1993/04/01). Density - functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, 98(7), 5648-5652. <https://doi.org/10.1063/1.464913>
- Becke, A. D. (1993b). A new mixing of Hartree - Fock and local density - functional theories. *The Journal of Chemical Physics*, 98(2), 1372-1377. <https://doi.org/10.1063/1.464304>
- Bhartia, R., Beegle, L. W., Deflores, L., Abbey, W., Razzell Hollis, J., Uckert, K., Monacelli, B., Edgett, K. S., Kennedy, M. R., Sylvia, M., Aldrich, D., Anderson, M., Asher, S. A., Bailey, Z., Boyd, K., Burton, A. S., Caffrey, M., Calaway, M. J., Calvet, R., Cameron, B., Caplinger, M. A., Carrier, B. L., Chen, N., Chen, A., Clark, M. J., Clegg, S., Conrad, P. G., Cooper, M., Davis, K. N., Ehlmann, B., Facto, L., Fries, M. D., Garrison, D. H., Gasway, D., Ghaemi, F. T., Graff, T. G., Hand, K. P., Harris, C., Hein, J. D., Heinz, N., Herzog, H., Hochberg, E., Houck, A., Hug, W. F., Jensen, E. H., Kah, L. C., Kennedy, J., Krylo, R., Lam, J., Lindeman, M., McGlown, J., Michel, J., Miller, E., Mills, Z., Minitti, M. E., Mok, F., Moore, J., Nealon, K. H., Nelson, A., Newell, R., Nixon, B. E., Nordman, D. A., Nuding, D., Orellana, S., Pauken, M., Peterson, G., Pollock, R., Quinn, H., Quinto, C., Ravine, M. A., Reid, R. D., Riendeau, J., Ross, A. J., Sackos, J., Schaffner, J. A., Schwochert, M., O Shelton, M., Simon, R., Smith, C. L., Sobron, P., Steadman, K., Steele, A., Thiessen, D., Tran, V. D., Tsai, T., Tuite, M., Tung, E., Wehbe, R., Weinberg, R., Weiner, R. H., Wiens, R. C., Williford, K., Wollonciej, C., Wu, Y.-H., Yingst, R. A. et Zan, J. (2021). Perseverance's Scanning Habitable Environments with Raman and Luminescence for Organics and Chemicals (SHERLOC) Investigation. *Space Science Reviews*, 217(4). <https://doi.org/10.1007/s11214-021-00812-z>
- Boys, S. F. (1950). Electronic wave functions - I. A general method of calculation for the stationary states of any molecular system. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 200(1063), 542-554. <https://doi.org/10.1098/rspa.1950.0036>
- Brehm, M. et Thomas, M. (2017, Jul 20). Computing Bulk Phase Raman Optical Activity Spectra from ab initio Molecular Dynamics Simulations. *J Phys Chem Lett*, 8(14), 3409-3414. <https://doi.org/10.1021/acs.jpcllett.7b01616>
- Brehm, M. et Thomas, M. (2019, Jul 9). Computing Bulk Phase Resonance Raman Spectra from ab Initio Molecular Dynamics and Real-Time TDDFT. *J Chem Theory Comput*, 15(7), 3901-3905. <https://doi.org/10.1021/acs.jctc.9b00512>
- Brehm, M., Thomas, M., Gehrke, S. et Kirchner, B. (2020). TRAVIS—A free analyzer for trajectories from molecular simulation. *The Journal of Chemical Physics*, 152(16), 164105. <https://doi.org/10.1063/5.0005078>

- Caldeweyher, E., Bannwarth, C. et Grimme, S. (2017). Extension of the D3 dispersion coefficient model. *The Journal of Chemical Physics*, 147(3), 034112. <https://doi.org/10.1063/1.4993215>
- Caldeweyher, E., Ehlert, S., Hansen, A., Neugebauer, H., Spicher, S., Bannwarth, C. et Grimme, S. (2019). A generally applicable atomic-charge dependent London dispersion correction. *The Journal of Chemical Physics*, 150(15), 154122. <https://doi.org/10.1063/1.5090222>
- Cheeseman, J. R. et Frisch, M. J. (2011). Basis Set Dependence of Vibrational Raman and Raman Optical Activity Intensities. *Journal of Chemical Theory and Computation*, 7(10), 3323-3334. <https://doi.org/10.1021/ct200507e>
- Chu, X.-L., Li, J.-Y., Chen, P. et Xu, Y.-P. (2014). Algorithms, Strategies and Application Progress of Spectral Searching Methods. *Chinese Journal of Analytical Chemistry*, 42(9), 1379-1386. [https://doi.org/10.1016/s1872-2040\(14\)60768-4](https://doi.org/10.1016/s1872-2040(14)60768-4)
- Ciubuc, J. D., Bennet, K. E., Qiu, C., Alonzo, M., Durrer, W. G. et Manciu, F. S. (2017, Sep 28). Raman Computational and Experimental Studies of Dopamine Detection. *Biosensors (Basel)*, 7(4). <https://doi.org/10.3390/bios7040043>
- Clemente, F. R., Vreven, T. et Frisch, M. J. (dir.). (2010). *Getting the Most out of ONIOM: Guidelines and Pitfalls*. Wiley-VCH Verlag GmbH.
- Császár, P. et Pulay, P. (1984, 1984/01/01/). Geometry optimization by direct inversion in the iterative subspace. *Journal of Molecular Structure*, 114, 31-34. [https://doi.org/https://doi.org/10.1016/S0022-2860\(84\)87198-7](https://doi.org/https://doi.org/10.1016/S0022-2860(84)87198-7)
- Curutchet, C., Muñoz-Losa, A., Monti, S., Kongsted, J., Scholes, G. D. et Mennucci, B. (2009). Electronic Energy Transfer in Condensed Phase Studied by a Polarizable QM/MM Model. *Journal of Chemical Theory and Computation*, 5(7), 1838-1848. <https://doi.org/10.1021/ct9001366>
- de Gelder, R., Wehrens, R. et Hageman, J. A. (2001). A generalized expression for the similarity of spectra: application to powder diffraction pattern classification. *Journal of Computational Chemistry*, 22(3), 273-289. [https://doi.org/10.1002/1096-987x\(200102\)22:3<273::Aid-jcc1001>3.0.Co;2-0](https://doi.org/10.1002/1096-987x(200102)22:3<273::Aid-jcc1001>3.0.Co;2-0)
- de Souza, B., Farias, G., Neese, F. et Izsak, R. (2019, Jun 7). Efficient simulation of overtones and combination bands in resonant Raman spectra. *J Chem Phys*, 150(21), 214102. <https://doi.org/10.1063/1.5099247>
- Edwards, H. G. M., Brooke, C. J. et Tait, J. K. F. (1997). Fourier Transform Raman Spectroscopic Study of Pigments from English Mediaeval Wall Paintings. *Journal of Raman Spectroscopy*, 28(2-3), 95-98. [https://doi.org/10.1002/\(sici\)1097-4555\(199702\)28:2/3<95::Aid-jrs74>3.0.Co;2-0](https://doi.org/10.1002/(sici)1097-4555(199702)28:2/3<95::Aid-jrs74>3.0.Co;2-0)
- Fairchild, S. Z., Bradshaw, C. F., Su, W. et Guharay, S. K. (2009, Jul). Predicting Raman spectra using density functional theory. *Appl Spectrosc*, 63(7), 733-741. <https://doi.org/10.1366/000370209788700991>
- Ferrari, B. C. et Bennett, C. J. (2019). A Comparison of Medium-Sized Basis Sets for the Prediction of Geometries, Vibrational Frequencies, Infrared Intensities and Raman Activities for Water. *Journal of Physics: Conference Series*, 1290(1). <https://doi.org/10.1088/1742-6596/1290/1/012013>
- Furness, J. W., Kaplan, A. D., Ning, J., Perdew, J. P. et Sun, J. (2020). Accurate and Numerically Efficient r²SCAN Meta-Generalized Gradient Approximation. *The Journal of Physical Chemistry Letters*, 11(19), 8208-8215. <https://doi.org/10.1021/acs.jpcllett.0c02405>
- Galeev, R. R., Saveliev, A. A., Galeeva, E. V., Falaleeva, T. S., Aryslanov, I. R., Semanov, D. A. et Davletshin, R. R. (2019). Application of the normalized relative error distribution analysis for non - destructive quality control of drugs by Raman spectroscopy. *Journal of Raman Spectroscopy*, 50(10), 1614-1623. <https://doi.org/10.1002/jrs.5658>
- Garg, A., Enright, C. G. et Madden, M. G. (2015, May 26). Improving spectral library search by redefining similarity measures. *J Chem Inf Model*, 55(5), 963-971. <https://doi.org/10.1021/acs.jcim.5b00077>
- Giridhar, G., Manepalli, R. R. K. N. et Apparao, G. (2017). Confocal Raman Spectroscopy. Dans S. Thomas, R. Thomas, A. K. Zachariah et R. Kumar Mishra (dir.), *Spectroscopic Methods for Nanomaterials Characterization* (vol. 2, p. 141-161). Elsevier. <https://doi.org/10.1016/B978-0-323-46140-5.00007-8>.

- Goerigk, L. (2017). Chapter 6 - A Comprehensive Overview of the DFT-D3 London-Dispersion Correction. Dans A. Otero de la Roza et G. A. DiLabio (dir.), *Non-Covalent Interactions in Quantum Chemistry and Physics* (p. 195-219). Elsevier. <https://doi.org/10.1016/B978-0-12-809835-6.00007-4>.
- Gomes da Costa, S., Richter, A., Schmidt, U., Breuninger, S. et Hollricher, O. (2019, Mar). Confocal Raman microscopy in life sciences. *Morphologie*, 103(341), 11-16. <https://doi.org/10.1016/j.morpho.2018.12.003>
- Grimme, S. (2006). Semiempirical hybrid density functional with perturbative second-order correlation. *The Journal of Chemical Physics*, 124(3), 034108. <https://doi.org/10.1063/1.2148954>
- Grimme, S., Antony, J., Ehrlich, S. et Krieg, H. (2010). A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics*, 132(15), 154104. <https://doi.org/10.1063/1.3382344>
- Grimme, S., Hansen, A., Ehlert, S. et Mewes, J.-M. (2021). r²SCAN-3c: A “Swiss army knife” composite electronic-structure method. *The Journal of Chemical Physics*, 154(6), 064103. <https://doi.org/10.1063/5.0040021>
- Groenhof, G. (2013). Introduction to QM/MM simulations. *Methods Mol Biol*, 924, 43-66. https://doi.org/10.1007/978-1-62703-017-5_3
- Gryniewicz-Ruzicka, C. M., Rodriguez, J. D., Arzhantsev, S., Buhse, L. F. et Kauffman, J. F. (2012, Mar 5). Libraries, classifiers, and quantifiers: a comparison of chemometric methods for the analysis of Raman spectra of contaminated pharmaceutical materials. *J Pharm Biomed Anal*, 61, 191-198. <https://doi.org/10.1016/j.jpba.2011.12.002>
- Hill, J. G. (2013). Gaussian basis sets for molecular applications. *International Journal of Quantum Chemistry*, 113(1), 21-34. <https://doi.org/10.1002/qua.24355>
- Hohenberg, P. et Kohn, W. (1964). Inhomogeneous Electron Gas. *Physical Review*, 136(3B), B864-B871. <https://doi.org/10.1103/physrev.136.b864>
- Hopmann, K. H., Ruud, K., Pecul, M., Kudelski, A., Dračinský, M. et Bouř, P. (2011). Explicit versus Implicit Solvent Modeling of Raman Optical Activity Spectra. *The Journal of Physical Chemistry B*, 115(14), 4128-4137. <https://doi.org/10.1021/jp110662w>
- Huang, Y., Luo, Y., Liu, H., Lu, X., Zhao, J. et Lei, Y. (2020). A Subcutaneously Injected SERS Nanosensor Enabled Long-term in Vivo Glucose Tracking. *Engineered Science*. <https://doi.org/10.30919/es8d1161>
- Jensen, F. (2013). Atomic orbital basis sets. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(3), 273-295. <https://doi.org/10.1002/wcms.1123>
- Jimenez-Hoyos, C. A., Janesko, B. G. et Scuseria, G. E. (2008, Nov 28). Evaluation of range-separated hybrid density functionals for the prediction of vibrational frequencies, infrared intensities, and Raman activities. *Phys Chem Chem Phys*, 10(44), 6621-6629. <https://doi.org/10.1039/b810877c>
- Johnson, E. R. et Becke, A. D. (2005). A post-Hartree–Fock model of intermolecular interactions. *The Journal of Chemical Physics*, 123(2), 024101. <https://doi.org/10.1063/1.1949201>
- Johnson, E. R., Becke, A. D., Sherrill, C. D. et DiLabio, G. A. (2009). Oscillations in meta-generalized-gradient approximation potential energy surfaces for dispersion-bound complexes. *The Journal of Chemical Physics*, 131(3), 034111. <https://doi.org/10.1063/1.3177061>
- Kirchner, B., Blasius, J., Esser, L. et Reckien, W. (2021). Predicting Vibrational Spectroscopy for Flexible Molecules and Molecules with Non - Idle Environments. *Advanced Theory and Simulations*, 4(4). <https://doi.org/10.1002/adts.202000223>
- Kohn, W. et Sham, L. J. (1965). Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review*, 140(4A), A1133-A1138. <https://doi.org/10.1103/physrev.140.a1133>
- Komornicki, A. et McIver, J. W. (1979, 1979/02/15). An efficient ab initio method for computing infrared and Raman intensities: Application to ethylene. *The Journal of Chemical Physics*, 70(4), 2014-2016. <https://doi.org/10.1063/1.437627>

- Kopec, M. et Abramczyk, H. (2018, Jun 5). Angiogenesis - a crucial step in breast cancer growth, progression and dissemination by Raman imaging. *Spectrochim Acta A Mol Biomol Spectrosc*, 198, 338-345. <https://doi.org/10.1016/j.saa.2018.02.058>
- Kruse, H. et Grimme, S. (2012). A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. *The Journal of Chemical Physics*, 136(15), 154101. <https://doi.org/10.1063/1.3700154>
- Kudelski, A. (2008, Jun 30). Analytical applications of Raman spectroscopy. *Talanta*, 76(1), 1-8. <https://doi.org/10.1016/j.talanta.2008.02.042>
- Kuhar, N., Sil, S., Verma, T. et Umapathy, S. (2018). Challenges in application of Raman spectroscopy to biology and materials. *RSC Advances*, 8(46), 25888-25908. <https://doi.org/10.1039/c8ra04491k>
- Larkin, P. (2011). Infrared and Raman Spectroscopy Principles and Spectral Interpretation. Dans *Infrared and Raman Spectroscopy*. Elsevier. <https://doi.org/10.1016/b978-0-12-386984-5.10001-1>.
- Lasch, P. (2012). Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging. *Chemometrics and Intelligent Laboratory Systems*, 117, 100-114. <https://doi.org/10.1016/j.chemolab.2012.03.011>
- Li, J., Hibbert, D. B., Fuller, S. et Vaughn, G. (2006). A comparative study of point-to-point algorithms for matching spectra. *Chemometrics and Intelligent Laboratory Systems*, 82(1-2), 50-58. <https://doi.org/10.1016/j.chemolab.2005.05.015>
- Lin, C. K., Huang, Q. R., Hayashi, M. et Kuo, J. L. (2021, Nov 24). An ab initio anharmonic approach to IR, Raman and SFG spectra of the solvated methylammonium ion. *Phys Chem Chem Phys*, 23(45), 25736-25747. <https://doi.org/10.1039/d1cp04451f>
- Luber, S., Iannuzzi, M. et Hutter, J. (2014, Sep 7). Raman spectra from ab initio molecular dynamics and its application to liquid S-methyloxirane. *J Chem Phys*, 141(9), 094503. <https://doi.org/10.1063/1.4894425>
- Lussier, F., Thibault, V., Charron, B., Wallace, G. Q. et Masson, J.-F. (2020). Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering. *TrAC Trends in Analytical Chemistry*, 124. <https://doi.org/10.1016/j.trac.2019.115796>
- M. Edwards, H. G. et Chalmers, J. M. (dir.). (2005). *Raman Spectroscopy in Archaeology and Art History*. The Royal Society of Chemistry.
- Mardirossian, N. et Head-Gordon, M. (2017). Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Molecular Physics*, 115(19), 2315-2372. <https://doi.org/10.1080/00268976.2017.1333644>
- Marenich, A. V., Cramer, C. J. et Truhlar, D. G. (2009). Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *The Journal of Physical Chemistry B*, 113(18), 6378-6396. <https://doi.org/10.1021/jp810292n>
- Mennucci, B., Tomasi, J., Cammi, R., Cheeseman, J. R., Frisch, M. J., Devlin, F. J., Gabriel, S. et Stephens, P. J. (2002). Polarizable Continuum Model (PCM) Calculations of Solvent Effects on Optical Rotations of Chiral Molecules. *The Journal of Physical Chemistry A*, 106(25), 6102-6113. <https://doi.org/10.1021/jp020124t>
- Merrick, J. P., Moran, D. et Radom, L. (2007). An Evaluation of Harmonic Vibrational Frequency Scale Factors. *The Journal of Physical Chemistry A*, 111(45), 11683-11700. <https://doi.org/10.1021/jp073974n>
- Novitskiy, I. M. et Kutateladze, A. G. (2022, Apr 1). DU8ML: Machine Learning-Augmented Density Functional Theory Nuclear Magnetic Resonance Computations for High-Throughput In Silico Solution Structure Validation and Revision of Complex Alkaloids. *J Org Chem*, 87(7), 4818-4828. <https://doi.org/10.1021/acs.joc.2c00169>
- Orlando, A., Franceschini, F., Muscas, C., Pidkova, S., Bartoli, M., Rovere, M. et Tagliaferro, A. (2021). A Comprehensive Review on Raman Spectroscopy Applications. *Chemosensors*, 9(9), 262. <https://doi.org/10.3390/chemosensors9090262>

- Park, J. K., Park, A., Yang, S. K., Baek, S. J., Hwang, J. et Choo, J. (2017, Jan 16). Raman spectrum identification based on the correlation score using the weighted segmental hit quality index. *Analyst*, 142(2), 380-388. <https://doi.org/10.1039/c6an02315k>
- Parr, R. G. et Weitao, Y. (1995). *Density-Functional Theory of Atoms and Molecules*. Oxford University Press. <https://doi.org/10.1093/oso/9780195092769.001.0001>
- Pavia, D. L., Lampman, G. M., Kriz, G. S. et Vyvyan, J. R. (2009). *Introduction to Spectroscopy* (Fourth Edition éd.). Brooks/Cole.
- Perdew, J. P., Ernzerhof, M. et Burke, K. (1996, 1996/12/08). Rationale for mixing exact exchange with density functional approximations. *The Journal of Chemical Physics*, 105(22), 9982-9985. <https://doi.org/10.1063/1.472933>
- Perdew, J. P. et Schmidt, K. (2001). Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conference Proceedings*, 577(1), 1-20. <https://doi.org/10.1063/1.1390175>
- Perera, A. S., Thomas, J., Poopari, M. R. et Xu, Y. (2016). The Clusters-in-a-Liquid Approach for Solvation: New Insights from the Conformer Specific Gas Phase Spectroscopy and Vibrational Optical Activity Spectroscopy. *Front Chem*, 4, 9. <https://doi.org/10.3389/fchem.2016.00009>
- Raman, C. V. (1953). A new radiation. *Proceedings of the Indian Academy of Sciences - Section A*, 37, 333-341.
- Rappoport, D. et Furche, F. (2010, Oct 7). Property-optimized gaussian basis sets for molecular response calculations. *J Chem Phys*, 133(13), 134105. <https://doi.org/10.1063/1.3484283>
- Rodriguez, J. D., Westenberger, B. J., Buhse, L. F. et Kauffman, J. F. (2011, Jun 1). Quantitative evaluation of the sensitivity of library-based Raman spectral correlation methods. *Anal Chem*, 83(11), 4061-4067. <https://doi.org/10.1021/ac200040b>
- Santra, G., Calinsky, R. et Martin, J. M. L. (2022). Benefits of Range-Separated Hybrid and Double-Hybrid Functionals for a Large and Diverse Data Set of Reaction Energies and Barrier Heights. *The Journal of Physical Chemistry A*, 126(32), 5492-5505. <https://doi.org/10.1021/acs.jpca.2c03922>
- Schlegel, H. B. (2011). Geometry optimization. *WIREs Computational Molecular Science*, 1(5), 790-809. <https://doi.org/10.1002/wcms.34>
- Shah, N. C., Lyandres, O., Walsh, J. T., Jr., Glucksberg, M. R. et Van Duyne, R. P. (2007, Sep 15). Lactate and sequential lactate-glucose sensing using surface-enhanced Raman spectroscopy. *Anal Chem*, 79(18), 6927-6932. <https://doi.org/10.1021/ac0704107>
- Sholl, D. S. et Steckel, J. A. (2009). DFT Calculations of Vibrational Frequencies. Dans *Density Functional Theory* (p. 113-130). John Wiley & Sons, Inc. <https://doi.org/https://doi.org/10.1002/9780470447710.ch5>.
- Smith, W. E. et Dent, G. (2005). *Modern Raman Spectroscopy – A Practical Approach*. John Wiley & Sons, Ltd.
- Sousa, S. F., Fernandes, P. A. et Ramos, M. J. (2007, Oct 25). General performance of density functionals. *J Phys Chem A*, 111(42), 10439-10452. <https://doi.org/10.1021/jp0734474>
- Steindal, A. H., Ruud, K., Frediani, L., Aidas, K. et Kongsted, J. (2011). Excitation Energies in Solution: The Fully Polarizable QM/MM/PCM Method. *The Journal of Physical Chemistry B*, 115(12), 3027-3037. <https://doi.org/10.1021/jp1101913>
- Stiles, P. L., Dieringer, J. A., Shah, N. C. et Duyne, R. P. V. (2008). Surface-Enhanced Raman Spectroscopy. *Annual Review of Analytical Chemistry*, 1(1), 601-626. <https://doi.org/10.1146/annurev.anchem.1.031207.112814>
- Thomas, M., Brehm, M., Fligg, R., Vohringer, P. et Kirchner, B. (2013, May 14). Computing vibrational spectra from ab initio molecular dynamics. *Phys Chem Chem Phys*, 15(18), 6608-6622. <https://doi.org/10.1039/c3cp44302g>
- Tomasi, J., Mennucci, B. et Cammi, R. (2005). Quantum Mechanical Continuum Solvation Models. *Chemical Reviews*, 105(8), 2999-3094. <https://doi.org/10.1021/cr9904009>
- Vidic, J., Chaix, C., Manzano, M. et Heyndrickx, M. (2020). Food Sensing: Detection of Bacillus cereus Spores in Dairy Products. *Biosensors*, 10(3), 15. <https://doi.org/10.3390/bios10030015>

- Wei, D., Chen, S. et Liu, Q. (2015). Review of Fluorescence Suppression Techniques in Raman Spectroscopy. *Applied Spectroscopy Reviews*, 50(5), 387-406. <https://doi.org/10.1080/05704928.2014.999936>
- Weigend, F. et Ahlrichs, R. (2005). Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7(18), 3297. <https://doi.org/10.1039/b508541a>
- Wiens, R. C., Maurice, S., Robinson, S. H., Nelson, A. E., Cais, P., Bernardi, P., Newell, R. T., Clegg, S., Sharma, S. K., Storms, S., Deming, J., Beckman, D., Ollila, A. M., Gasnault, O., Anderson, R. B., André, Y., Michael Angel, S., Arana, G., Auden, E., Beck, P., Becker, J., Benzerara, K., Bernard, S., Beyssac, O., Borges, L., Bousquet, B., Boyd, K., Caffrey, M., Carlson, J., Castro, K., Celis, J., Chide, B., Clark, K., Cloutis, E., Cordoba, E. C., Cousin, A., Dale, M., Deflores, L., Delapp, D., Deleuze, M., Dirmyer, M., Donny, C., Dromart, G., George Duran, M., Egan, M., Ervin, J., Fabre, C., Fau, A., Fischer, W., Forni, O., Fouchet, T., Fresquez, R., Frydenvang, J., Gasway, D., Gontijo, I., Grotzinger, J., Jacob, X., Jacquino, S., Johnson, J. R., Klisiewicz, R. A., Lake, J., Lanza, N., Laserna, J., Lasue, J., Le Mouélic, S., Legett, C., Leveille, R., Lewin, E., Lopez-Reyes, G., Lorenz, R., Lorigny, E., Love, S. P., Lucero, B., Madariaga, J. M., Madsen, M., Madsen, S., Mangold, N., Manrique, J. A., Martinez, J. P., Martinez-Frias, J., McCabe, K. P., McConnochie, T. H., McGlown, J. M., McLennan, S. M., Melikechi, N., Meslin, P.-Y., Michel, J. M., Mimoun, D., Misra, A., Montagnac, G., Montmessin, F., Mousset, V., Murdoch, N., Newsom, H., Ott, L. A., Ousnamer, Z. R., Pares, L., Parot, Y., Pawluczyk, R., Glen Peterson, C., Pilleri, P., Pinet, P., Pont, G., Poulet, F., Provost, C., Quertier, B., Quinn, H., Rapin, W., Reess, J.-M., Regan, A. H., Reyes-Newell, A. L., Romano, P. J., Royer, C., Rull, F., Sandoval, B., Sarrao, J. H., Sautter, V., Schoppers, M. J., Schröder, S., Seitz, D., Shepherd, T., Sobron, P., Dubois, B., Sridhar, V., Toplis, M. J., Torre-Fdez, I., Trettel, I. A., Underwood, M., Valdez, A., Valdez, J., Venhaus, D. et Willis, P. (2021). The SuperCam Instrument Suite on the NASA Mars 2020 Rover: Body Unit and Combined System Tests. *Space Science Reviews*, 217(1). <https://doi.org/10.1007/s11214-020-00777-5>
- Wong, M. W. (1996, 1996/07/05/). Vibrational frequency prediction using density functional theory. *Chemical Physics Letters*, 256(4), 391-399. [https://doi.org/10.1016/0009-2614\(96\)00483-6](https://doi.org/10.1016/0009-2614(96)00483-6)
- Wu, J.-B., Lin, M.-L., Cong, X., Liu, H.-N. et Tan, P.-H. (2018). Raman spectroscopy of graphene-based materials and its applications in related devices. *Chemical Society Reviews*, 47(5), 1822-1873. <https://doi.org/10.1039/c6cs00915h>
- Zhang, R., Shang, Z., Lu, S., Jia, N., Jiang, X., Pu, Z., Du, Y. et Hu, Y. (2021). Spectral partition correlation based on Voigt function for Raman spectral library search. *Chemometrics and Intelligent Laboratory Systems*, 215. <https://doi.org/10.1016/j.chemolab.2021.104353>
- Zhao, Y., Lynch, B. J. et Truhlar, D. G. (2004). Doubly Hybrid Meta DFT: New Multi-Coefficient Correlation and Density Functional Methods for Thermochemistry and Thermochemical Kinetics. *The Journal of Physical Chemistry A*, 108(21), 4786-4791. <https://doi.org/10.1021/jp049253v>
- Zvereva, E. E., Shagidullin, A. R. et Katsyuba, S. A. (2011, Jan 13). Ab initio and DFT predictions of infrared intensities and Raman activities. *J Phys Chem A*, 115(1), 63-69. <https://doi.org/10.1021/jp108057p>