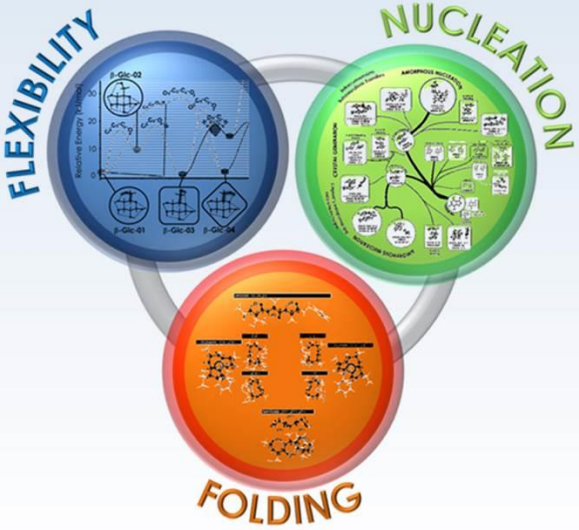
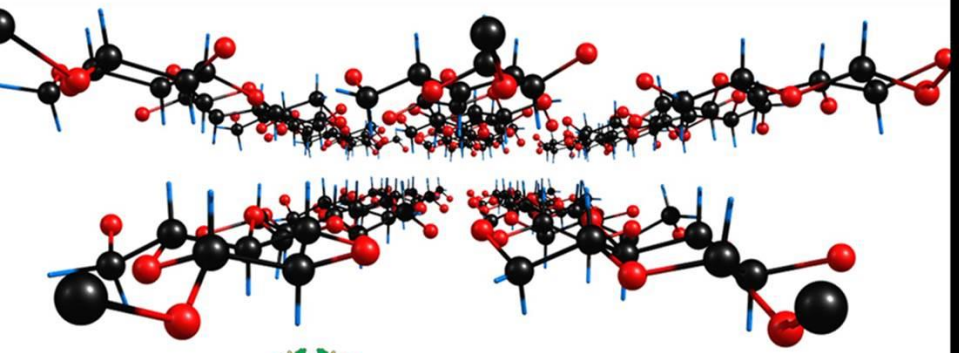




EUSKO JAURLARITZA
GOBIERNO VASCO

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea
ZIENTZIA ETA TEKNOLOGIA FAKULTATEA
FACULTAD DE CIENCIA Y TECNOLOGÍA

Structure and Intermolecular Interactions of Relevant Molecules
Molekula Nabarmenen Egitura eta Interakzio Intermolekularrak

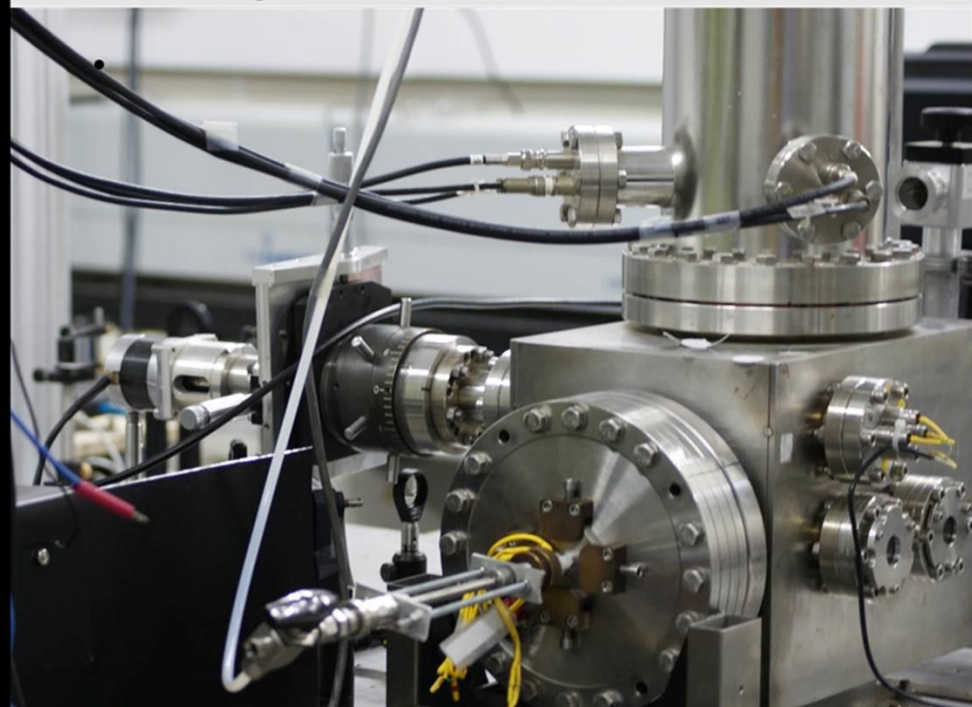
I. USABIAGA
2017

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea
ZIENTZIA ETA TEKNOLOGIA FAKULTATEA
FACULTAD DE CIENCIA Y TECNOLOGÍA

Structure and Intermolecular Interactions of Relevant Molecules

Molekula Nabarmenen Egitura eta Interakzio Intermolekularrak

Imanol Usabiaga Gutierrez 2017



Structure and Intermolecular Interactions of Relevant Molecules
Molekula Adierazgarrien Egitura eta Interakzio Intermolekularrak

Imanol Usabiaga Gutierrez

2017

Supervisors:

José Andrés Fernández González
Emilio José Cocinero Pérez

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea
ZIENTZIA ETA TEKNOLOGIA FAKULTATEA
FACULTAD DE CIENCIA Y TECNOLOGÍA

I would like to thank all the people who have helped me in this thesis project.

The first version of this thesis was written in a combination of English and Euskera. Next, the sections in Euskera were translated back into English to help the examiners to evaluate it. For clarity, two indexes for the two languages were included and the section in Euskera precedes the English version of the manuscript. To help the reader to identify each section, "Orrialdea|x" , where x = page number appears at the bottom of the pages in the Euskera version, while, "Page|x" appears at the bottom of the pages in the English. Likewise, "Appendix |x" was used for the appendixes.

Tesiaren %50 baino gehiago euskaraz egin da baina tesia internazionala izateko eta aztertzaile internazionalek ulertu dezaten tesi osoa ingelerara itzuli da. Atal bakoitza bereizteko orri-oinak desberdindu dira: Euskaraz dauden atalak "Orrialdea|x", Ingeleraz dauden atalak "Page|x" eta eranskinak "Appendix |x".

AURKIBIDEA

1. KAPITULUA: Sarrera eta Helburuak

2. KAPITULUA: Sistema Esperimentala eta Metodologia

MALGUTASUNA

3. KAPITULUA: Glukosa eta Fenol Molekulen Arteko Elkarrekintza: "*Bizirik Dirauten Iaioa*"

4. KAPITULUA: β-D-glukopiranosa Ekuatoriala: Propietate Estrukturalak Karbohidrato Ugarienean.

7. KAPITULUA: Ondorioak

1. Sarrera eta Helburuak

1.1. Sarrera

Sistema molekularren fundamentuen ikerketa, adibidez: protoien transferentzia, azukre baten egitura, polimero baten egitura eta interakzioak, proteina eta sendagaia den molekularen arteko konplexuaren kristala, eta abar, orokorrean neurketetan erabilitako teknikagatik mugatuta dago. Askotan neurtzeko prozeduraren ezaugarriak eta mugak eragozpenak ditu, hauek simulatzen den sisteman aldaketak eragiten dituzte, jasotako emaitzak izan behar zintezkeenen desberdinak sortuz. Edozein kasutan, teknika bakoitzak bere gaitasun propioak ditu eta beti ere zenbait informazio interesgarri sortzen dute. Benetan ona den informazioa teknika desberdin askoren emaitzak batzen direnean lortzen da, bakoitza molekulen ezaugarri batzuk detektatzen dituztenez, guztiak batzean, ikuspuntu orokor bat eraiki daiteke sistema molekular osoa ulertuz.

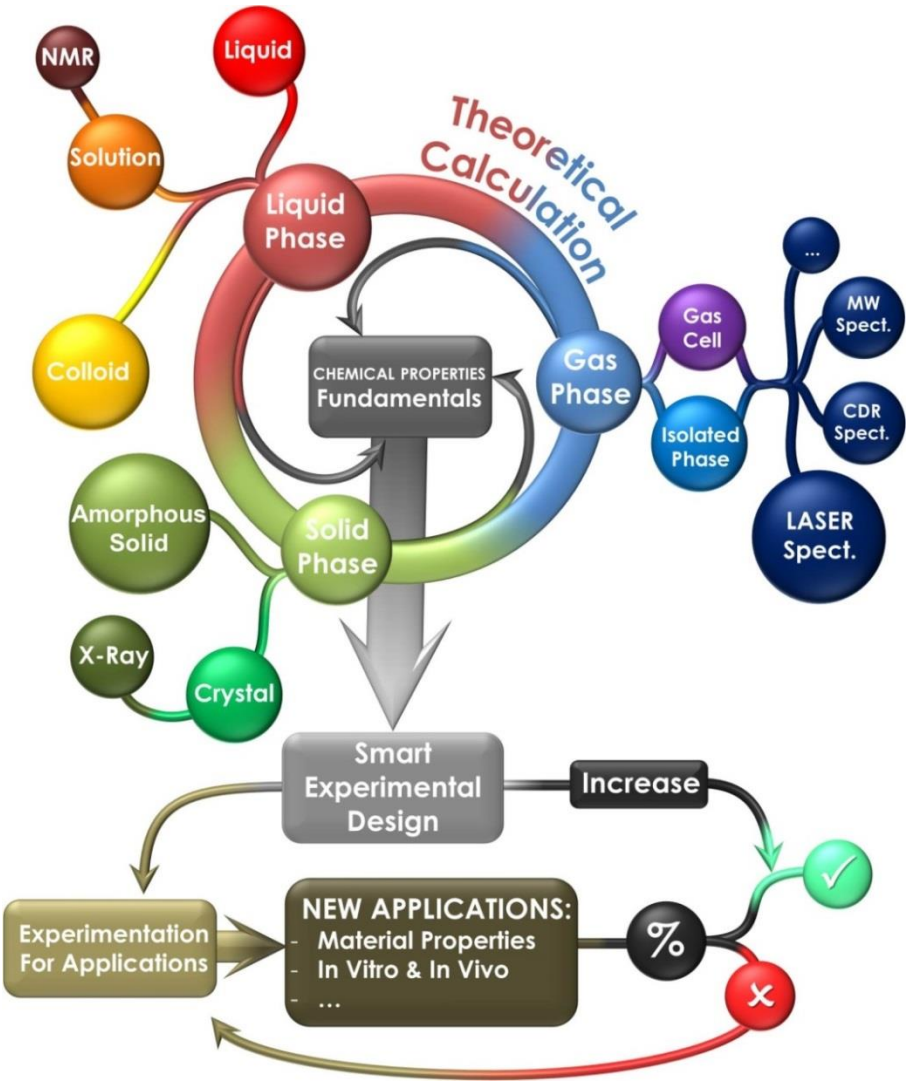
Materiaren propietateak ikertzeko teknika esperimental asko daude, batzuk prozesu dinamikoak ikertzen dituzte, beste batzuk propietate mekano-kuantikoen determinazioan ardazten dira, badaude konposaketa ere neurtzen dutenak, eta abar, baina alor garrantzitsuenetarikoen artean molekulen hiru dimentsiotako egitura determinazioa da. Askotan sistema molekular baten egiturak eta bere propietateek harreman estua dute, horregatik egitura neurtzen duten teknikak asko garatu dira, gaur egun multzo handia eratu dutelarik. Beste era batera esanda, egitura eta propietateak zuzenean konektatuta daude, molekula sinpleenetik hasita -adibidez monomero bat, sistema zailenera, biokimikan orekan dauden erreakzio-sare erraldoietaraino. Badaude adibide asko egitura eta egituren arteko elkarrekintzak propietateekin lotzen dutenak, beste askoren artean adibide bat: polimero batetan kateen egitura eta kateen arteko elkarrekintzek materialari elastikotasuna, erresistentzia, eta abar ematen dio.[1] Grafitoan pilatutako karbono geruzek π-lainoa eratzen dute egitura osoan zehar eta honek elektrizitatea garraiatzeko gaitasuna ematen dio materialari.[2] Biologikoa den ingurumen kimikoari dagokionez egitura eta interakzioen konplexutasuna maila askoz handiagoa da, honek funtzio asko eta oso desberdinak betetzeko gaitasuna ematen die: lotura gune izan daitezke,[3] egituren euskarri,[4,5] informazioa kodifikatu eta garraiatu dezake mRNA eta DNA gertatzen den eran, errekonozimendu molekularreko prozesuetan parte hartu[6] eta abar.

Lehenago aipatu den bezala badaude teknika asko egituraren determinaziorako balio dutenak, hauek erabili ohi diren fasearen arabera (solido, likido, gas) sailkatu daitezke. Solido egoeran kristalografia da teknikarik erabiliena, likido egoeran NMR oso ezaguna da eta gas egoeran teknika espektroskopiko asko aurkitu daitezke.

Gas erak, molekula sistema handi hauentzako giro arraroa dela alde batera utziz, badu zalantzarik gabeko abantaila, batez ere espantsio supersonikoa erabiltzen denean. Espantsio supersonikoak molekulak asko hozten ditu, honen barnean dauden molekula guztiak oinarrizko maila ro-bibrazionalera (baita maila elektronikoa ere) jezten dituelarik. Espantsio supersonikoan gas garraiatzailea molekulen kontra talka egiten du, prozesu honek tenperatura rotazionala eta bibrazionala tenperatura translazionalean bihurtzen du, baina molekula guztiak zuzen eta haien artean talkarik egin gabe hegan doanez sistemaren barruan tenperatura translazionala ere baxua dela esan daiteke. Azkenean orokorrean tenperatura asko jezten da molekulak izoztuta eta isolatuta geratzen direlarik, eta esperimentua egiten den unean honela aurkitzen dira, horregatik egoera honi hotz isolatutako fasea deituko zaio, CI fasea alegia. Teknika honek interferentziarik gabe molekulen ezaugarriak ezagutzea ahalbidetzen du, hemen ez baitago disolbatzailerik ezta kristalaren matrize ekarpenik, horregatik molekulen benetako jokaera jakiteko aukera ematen duela esan ohi da. Teknika honek hoztu eta espektroan aurkituko litzatekeen trantsizioen kopurua murrizten du, hau populazio guztia egitura gutxi batzuetan pilatuta geratu delako gertatzen da, egitura egonkorrenetan alegia. Trantsizio gutxiago izateak baditu bere ondorioak, seinalea ere trantsizio gutxi batzuetan pilatuta geratzen da, hauetako bakoitzaren intentsitatea[7,8] asko igoz, honek sentikortasunaren handitzea dakar eta laser askok erabiltzen dituzten teknikak erabiltzea ahalbidetzen du.[9,10] Laserrak erabili beharrean mikrouhinak erabiltzen badira hiru dimentsioetako egitura zehaztasun izugarriaz neurtu daitezke, baina teknika hau sistema ez oso handietara murrizturik dago.[11,12] Teknika bakoitzak bere mugak ditu, adibidez: egitura kristalografikoen ikerketak matrize kristalinoak mugatzen ditu. Kristal batek badu bere egitura errepikakorra eta simetrikoa.[13] Hau ahalbidetzeko molekulak beste era batera lotzera bultzatu ditzake baita euren egitura aldatzera. Disoluzioan RMN-ek du esparru oso handia, teknika honek multiplizitate egokia duten atomoak detektatzen ditu, baita atomoen arteko elkarrekintzak ere. Askotan teknika hau interakzio guneetako informazioa eskuratzeko erabiltzen da,[14,15] baina disoluzioaren natura dinamokoa dela eta bakarrik batezbesteko egiturak lor daitezke.

Proiektu honetan erabiliko den teknikak espantsio supersonikoa eta masetan bereizgarria den laser espektroskopia bateratzen ditu, honen bidez espektro elektronikoa eta bibrazionala neurtu daitezke molekula eta konplexu zehatzetan. Askotan isomero bakoitzaren informazioa bananduta eskuratzea ahalbidetzen du. Zehazki espantsio supersonikoa "*time-of-flight*" masen espektrometroaren ionizazio kameran barruan gertatzen da, bertan egoera resonante batetik zenbait fotoien

bidez ionizatzen duen teknika erabiltzen da (MPI "*multi-photon ionization*"). Badaude molekula asko berotzean deskonposatzen direnak, hauek ere gas eran aurkitu ahal izateko laserraren bidez solidoak gasak bihurtzeko teknika konektatuta du espantsioan. Bigarren kapituluan zehaztasun gehiago emango da teknika honi buruz. Laser espektroskopiak aurrerakuntza handia aurkezten du beste teknika espektroskopiko askorekin alderatuta,[9] oso selektiboa da informazioa eskuratzen, konplexutan mailaren arabera bereizten ditu eta.[16] Egiturak eta elkarrekintzak ezagutzeak propietate kimikoen ezagutzan sakontzen laguntzen du, gero eta hobeto ezagutu sistema molekularren oinarrizko funtzionamendua orduan eta errazagoa da hauetan eragitea erabilera praktiko berriak emateko.



1.1 Eskema: Diagrama honek propietate kimikoen oinarriak eta hauek eskuratzeko beharrezkoak diren teknika esperimentalak konektatzen ditu. Tesi honetan lortutako emaitzak, zenbait propietate kimiko azaltzen ditu, aplikazio berrien inspiraziorako eta bere garapenerako lagungarri izan daitezke, ezagutza honek esperimentuaren diseinu efektiboago bat sustatzen duelarik.

Tesi honetan erabilitako prozedurak 1.1 eskema jarraitzen du, bertan ikus daitekeenez gas fasetik lortutako emaitzak (tesi honen kasuan IC fasean laser espektroskopiaren bidez) beste faseetatik lortutako emaitzekin alderatzen dira kalkulu teorikoen bitartez. Oinarrizko edozein ikerketaren asken helburuak zerbaiti buruzko ezagutzan sakontzean datza, baina objektiboa ez da ezagutza soila hartzea, gauza praktiko eta erabilgarriekin zubia eraikitzea da helburua, nahiz eta konexioa ez hain agerikoa izan. Zenbat eta hobeto ezagutu oinarriak orduan eta errazago da hauek proiektu zehatz bat aurrera eramateko erabiltzea, azkenean arrakasta izateko probabilitatea handituz.

Orokorrean ondo pentsatutako esperimentuen diseinua beharrezkoa da aplikagarriak diren proiektuak aurrera eramateko, baina arrakasta izateko faktore asko kontutan izan behar dira. Adibidez medikamendu berri baten aurkikuntza egiteko[17] ausaz egindako esperimentuen bitartez lortu daiteke, zuzenean "*in vivo*" laginetan nahi den efektua lotu dela ikusi arte. Gauzak horrela egiteak ekonomikoak eta moralak diren arazoak dakarte, adibidez izaki bizidunekin esperimentuak eta abar. Azken hauek bezalako esperimentuak beharrezkoak dira baina haien erabilera ahalik eta gutxiena izan behar da, dirutza balio dute eta adibidez gizakiekin esperimentuetan ibiltzea ez da gauza egokia. Horregatik molekulen funtzionamenduaren oinarriak ondo ezagutu behar dira, aplikatzean beharrezkoak diren esperimentuak murrizten, adibidez "*in vivo*" egiten direnak, saiakera gutxiago behar izateko, diru eta arazoak aurreztuz. Proposatutako eskemak ez du esan nahi aurrerapen zientifiko eta teknologiko guztiak ondo pentsatutako esperimentuen ondorioak direla, "serendipia" hor egongo da beti, baina arrakasta izateko probabilitateak asko handitzen ditu.

1.2. Helburuak

Tesi honen helburu orokorrak oso desberdinak diren sistema molekular sorta baten ezaugarriak, propietate kimikoak eta funtzionamendua aurkitzean datza, helburu hau lortzeko erabiltzen diren esperimentu zein kalkuluetan behar diren egokitzapenak eginez.

Ikertutako sistemak hiru taldeetan sailkatu daitezke: Molekula malguak, egituraren eraldaketa oso erraz egiten dutenak ia edozein girotan (3 eta 4 kapituluak); foldameroak, hauen ezaugarri nagusia egitura anitzak izan dezaketela da baina behin egitura horretan tolestuta daudenean oso zaila da egituraren eraldaketaren bat egotea (5 kapitulua); molekula zurrunak, hauetan aldakortasuna erreakzio kimikoen bidez ematen da, adibidez tautomeria edo protoi baten

transferentziaren bidez (6 kapitulua). Sistema bakoitzean hoberen moldatzen diren prozedura teoriko eta esperimentalak identifikatu eta erabili izan dira. Helburua ez da soilik molekulen propietateak aurkitzea baizik eta kapitulu bakoitzean biologia edo material berriei buruz sortutako galdera batzuei erantzutea ere.

Zehazki, 3. kapituluko helburu nagusia emaitza esperimentalak ondoen azaltzen duen energia kalkulu identifikatzea da (energia erlatibo VS lotura energi), eta horretarako glukosa·fenol zenbait dimero eratuko dira. 4. Kapituluan ikerketa β-glukosaren propietate ekuatorialetan zentratzen da: molekula hau hidroxilo guztiak plano ekuatorialean dituen piranosa bakarra da, era berean naturan ugarienetarikoen artean dago. Ikerketa honen xedea β-glukosaren hidroxilo ekuatorialak molekulari nolabaiteko gaitasun berezia ematen dion determinatzea da. 5. kapituluan foldamero baten atal bat (tripeptido bat eta oligomero bat) erabiltzen da hurbilketa moduan tolesteko prozesua egiturak eta honen ondorioz lortutako propietate kimiko-fisikoak determinatzeko. Azkenean 6. kapituluan molekula zurrun baten propietateak deskribatuko dira (teobromina) eta datu hauekin metaketa edo nukleazio prozesuan emango diren lehenengo urratsak ikertuko dira. Bukatzeko datu guztiak batuko dira (monomeroa, dimero eta trimeroa) baita bibliografiatik ateratako datu kristalografikoak eta kalkulu teoriko guztiak nukleazio prozesuaren zuhaitz genealogikoa eraikitzeko. Zuhaitz honen bitartez irudi batetan era errazean laburbiltzen da prozesuak hartu ditzakeen bideak eta jokabide orokorra.

Lehen aipatu den bezala erabiliko den teknika espektroskopikoa espantsio supersonikoko masan bereizgarria den laser espektroskopia da, kimika konputazionala laguntza giza erabiliz. Sistema bakoitzak bere metodologia konputazionala berezia behar izango du eta hauxe garatua eta azaldua izango da. Esperimentazioari dagokionez egindako ekarpenak laginaren pilaketa protokoloetara, laserren doiketara eta jasotze aktibo teknikaren garapenera zuzenduko dira.

Ondorio moduan esan daiteke aurkitutako berezitasun kimikoak aztertu eta laburbildu dira etorkizuneko proiekturen bati lagungarri izan daitezen. Arrazoi honengatik ahalegin handia egin da emaitzen sinplifikazioan (adituentzako soilik diren taula eta emaitza mordoa "*Appendix*" edo eranskineetan jarriz) eta laburbildutako irudi orokorretan bildu dira, erabilgarriak diren ondorioak batuz eta ahalik eta zientzialari gehiengoari irisgarria egiteko nahiarekin.

1.3. Erreferentziak

(1) Van Krevelen, D. W.; Te Nijenhuis, K. *Properties of Polymers: Their Correlation with Chemical Structure; Their Numerical Estimation and Prediction from Additive Group Contributions*; Elsevier, 2009.

(2) Sengupta, R.; Bhattacharya, M.; Bandyopadhyay, S.; Bhowmick, A. K. A Review on the Mechanical and Electrical Properties of Graphite and Modified Graphite Reinforced Polymer Composites. *Prog. Polym. Sci.* **2011**, *36* (5), 638–670.

(3) Dominguez, C.; Boelens, R.; Bonvin, A. M. J. J. HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **2003**, *125* (7), 1731–1737.

(4) Harris, H. Actinomycetales and Related Groups. **1958**, No. 1958, 173–189.

(5) Schleifer, K. H.; Kandler, O. Peptidoglycan Types of Bacterial Cell Walls and Their Taxonomic Implications. *Bacteriol. Rev.* **1972**, *36* (4), 407–477.

(6) MacCallum, R. M.; Martin, a C.; Thornton, J. M. Antibody-Antigen Interactions: Contact Analysis and Binding Site Topography. *J. Mol. Biol.* **1996**, *262* (5), 732–745.

(7) Oswald, S.; Wallrabe, M.; Suhm, M. A. Cooperativity in Alcohol–Nitrogen Complexes: Understanding Cryomatrices through Slit Jet Expansions. *J. Phys. Chem. A* **2017**, acs.jpca.7b01265.

(8) Gottschalk, H. C.; Altnoder, J.; Heger, M.; Suhm, M. A. Control over the Hydrogen-Bond Docking Site in Anisole by Ring Methylation. *Angew. Chemie - Int. Ed.* **2016**, *55* (5), 1921–1924.

(9) Levy, D. H. Laser Spectroscopy of Cold Gas-Phase Molecules. *Annu. Rev. Phys. Chem* **1980**, *31*, 197–225.

(10) Zwier, T. S. Laser Spectroscopy of Jet-Cooled Biomolecules and Their Water-Containing Clusters: Water Bridges and Molecular Conformation. *J. Phys. Chem. A* **2001**, *105*, 8827–8839.

(11) Balle, T. J.; Flygare, W. H. Fabry-Perot Cavity Pulsed Fourier Transform Microwave Spectrometer with a Pulsed Nozzle Particle Source. *Rev. Sci. Instrum.* **1981**, *52* (1), 33–45.

(12) Pate, B. H.; De Lucia, F. C. Broadband Molecular Rotational Spectroscopy Special Issue. *J. Mol. Spectrosc.* **2012**, *280* (1), 1–2.

(13) Finger, L. W. Physical Properties of Crystals, Their Representation by Tensors and Matrices. *Eos, Trans. Am. Geophys. Union* **1983**, *64* (45), 643.

(14) Kogelberg, H.; Solís, D.; Jiménez-Barbero, J. New Structural Insights into Carbohydrate-Protein Interactions from NMR Spectroscopy. *Curr. Opin. Struct. Biol.* **2003**, *13* (5), 646–653.

(15) Mayer, M.; Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chemie - Int. Ed.* **1999**, *38* (12), 1784–1788.

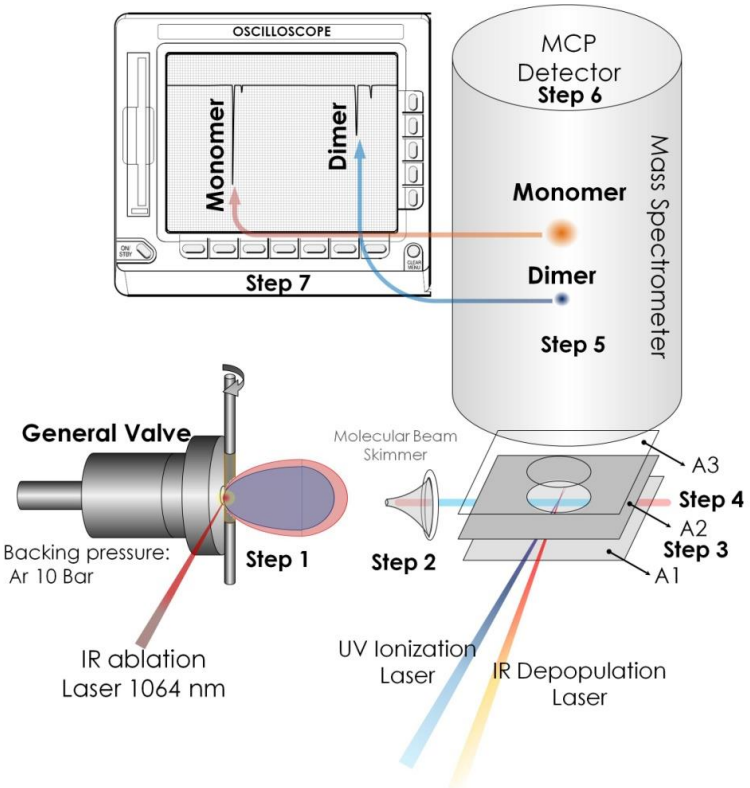
(16) Speranza, M.; Satta, M.; Piccirillo, S.; Rondino, F.; Paliadini, A.; Giardini, A.; Filippi, A.; Catone, D. Chiral Recognition by Mass-Resolved Laser Spectroscopy. *Mass Spectrom. Rev.* **2005**, *24* (4), 588–610.

(17) Marinelli, L.; Lavecchia, A.; Gottschalk, K.-E.; Novellino, E.; Kessler, H. Docking Studies on avb3 Integrin Ligands: Pharmacophore Refinement and Implications for Drug Design. *J. Med. Chem.* **2003**, *46*, 4393–4404.

2. Sistema Esperimentala eta Metodologia

2.1. Sistema Esperimentala

Atal honetan sistema esperimentalaren atal nagusiak deskribatuak izango dira. Helburua deskribapen zehatzak egitea da emaitza esperimentalak errepikakorrak izan daitezen. Sarreran aurreratu den bezala erabilitako teknika masatan zehaztasuna duen laser espektroskopia da, eta honi espantsio supersonikoa gehitu zaio espektro garbiak izateko.[1–3] A2.3 (A; Appendix atalean dagoela esan nahi du) irudian laborategi osoaren diagrama aurkitu daiteke eta 2.1.1 irudian masen espektrometroaren diagrama eta funtzionamendua azaltzen da.

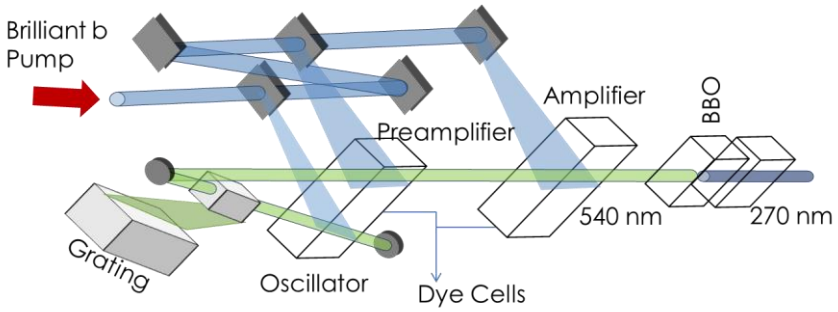


2.1.1 Irudia: Esperimentuak espantsio supersoniko bat eratzean datza, hau huts altuko kamera baten barruan egiten da, espantsioak 10Hz-tan funtzionatzen duen balbula sortzen du. Balbula honen irteeratik hurbil lagina eramango duen zilindroa dago, balbula irekita dagoen bitartean zilindroaren kontra laser batek jotzen du espantsio supersonikoaren hasieran lagina sartuz (**Step 1**). Gas garraiatzaileak lagina hoztu eta ionizazio gunera garraiatzen du. Konoarekin gas espantsioaren barruko aldea aukeratzen da (hotz dagoena) (**Step 2**). Laserraren pultsua (Q-switch) molekulen ionizazio gunera ailegatzearekin sinkronizatzen da, sortutako ioiak tentsio diferentzia txiki baten bitartez azelerazio gunera mugitzen dira (A1= 4000 V, A2 = 3700 V) (**Step 3**). Azelerazio gunean tentsio diferentzia handiaren eraginez azeleratu egiten dira (A2 = 3700 V, A3 =lurra) (**Step 4**). Ioiek aske hegan egiten dute espektrometroan zehar, masa eta kargaren arabera abiadura bat izango dute, karga bat denez gehiengo kasuetan masaren arabera bananduko dira (**Step 5**). Ioiek detektagailuaren kontra jotzen dute seinale elektriko bat sortuz (**Step 6**). Seinale hori osziloskopio batekin jasotzen da, osziloskopioaren ardatza tenporalaren arabera masa desberdinen ailegatzea bereiziz (**Step 7**).

2.1.1. Laserrak

Laserrak oso erabilgarriak diren erremintak dira, eta tesi honetan oso desberdinak diren funtzioak beteko dituzte. Laser argindarrak hurrengo erabilerak izango ditu: (i) Molekulak kitzikatu eta ionizatzeko; (ii) Despopulatzeko erresonantzia bikoitzeko esperimentuetan; (iii) tenperaturarekin deskonposatzen diren molekulak baporatzeko. Funtzio hauek hurrengo ataletan azalduko dira.

2.1.1.1. Kitzikatze eta Ionizazio Laser Argindarrak (Ultramore (UV) Laserrak)



2.1.1.1 Irudia: Koloratzaile laserren sinplifikatutako diseinu optikoa.

Koloratzaile mota askotako UV laser aurkitu daitezke baina guztiak diseinu optiko berean oinarritzen dira. 2.1.1.1 irudian eskema orokorra ikusgai dago. Laserra kameraren barruko aldera garraiatzeko kuartzoko prismak erabiltzen dira. Laserra kameran ardazteko luzaera desberdinetako lente batzuk (CaF2 eta kuartzoa) erabiliko dira. Laserrak duen energiaren arabera gehiago edo gutxiago ardaztea behar izango da, zenbat eta energia txikiagoa orduan eta ardazketa handiagoa.

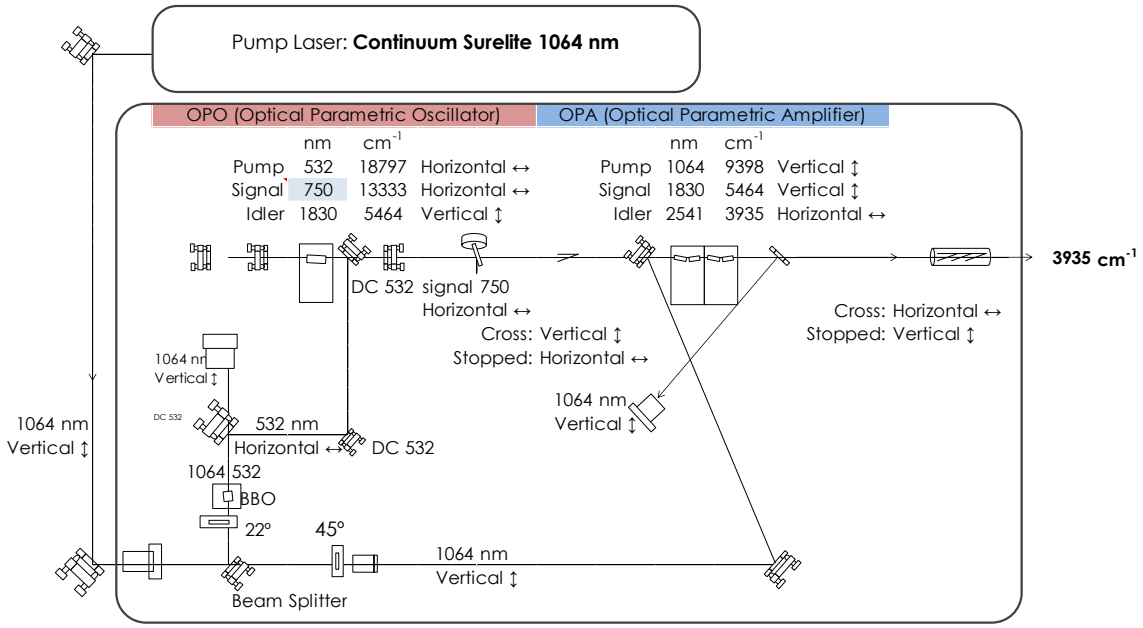
Fine Adjustment koloratzaile laserra: Brilliant b Nd:YAG laserra $\omega 2$ eta $\omega 3$ harmonikoen moduluekin, 120 mJ 355 nm-tan argindarra. Laser hau tarte honetan erabili da 250-270 nm, eta erabilitako koloratzaileak hurrengoak dira: Coumarin 540A eta Coumarin 500. Batez beste sortutako argindarra ~300 μJ da.

Scan Mate koloratzaile laserra: Brilliant b Nd:YAG laserra $\omega 2$ harmonikoaren moduluarekin, 350 mJ 532 nm-tan argindarra. Laser hau tarte honetan erabili da: 285-350 nm, eta erabilitako koloratzaileak hurrengoak dira: Rhodamine (575, 590 eta 610).

2.1.1.2. Despopulazio Laserrak (Ultramore/Infragorri (UV/IR) Laserrak)

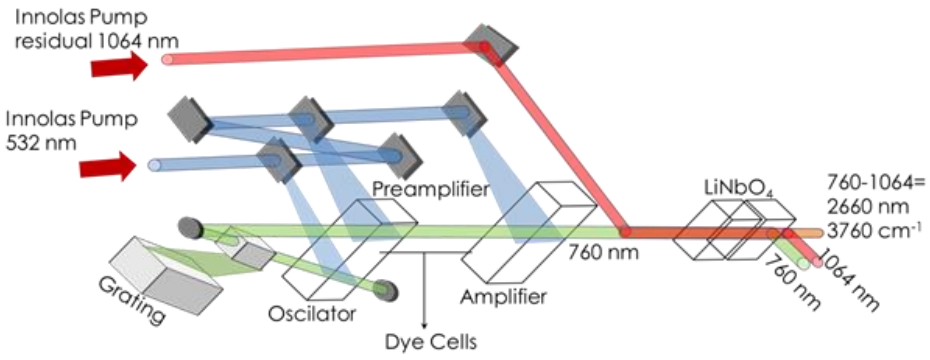
Erresonantzia bikoitzeko esperimentuak aurrera eramateko gutxienez bi laser dira beharrezkoak, eta hauetako bat gutxienez UV laserra izan behar da molekulak ionizatu ahal izateko (teknika hauen xehetasunak aurrerago azalduko dira). Bigarren laserra UV edo IR izan daiteke. UV laserrak lehen azaldu dira eta horren ondorioz IR laserrei buruz arituko da atal honetan, OPO/OPA IR eta koloratzaile IR laserrak.

Laser Vision OPO (Optical Parametric Oscillator) / OPA (Optical Parametric Amplifier) Laserra: argindarra Continuum Surelite Nd:YAG 1064 nm 800 mJ. Laserraren bukaeran sortutako argindarra 10 mJ sintonizagarria 2500-4000 cm^{-1} tartean. Uhin zabalera ~10 cm^{-1} (ikusi 2.1.1.2.1 irudia).



2.1.1.2.1 Irudia: Laser Vision OPO/OPA eskema

Fine Adjustment koloratzaile Laserra: Innolas SpitLight 1200 Nd:YAG laserra argindar moduan duelarik: 532 nm, 600 mJ. Laser honen funtzionamendua UV laserraren berdina da. Desberdintasun bakarra bigarren koloratzaile kubetaren ondoren BBO kristala izan beharrean LiNbO4 kristala duela da (2.1.1.2.2 irudia). IR erradiazioa sortzeko erabilitako koloratzaileak LDS nahasketak dira, hauek energia tarte honetan sortzen dute: 3100 and 3800 cm^{-1}.



2.1.1.2.2 Irudia: Koloratzaile laserren sinplifikatutako diseinu optikoa.

2.1.1.3. Ablazio laserra Nd:YAG (1064 nm Laserra)

Baporaketa prozesurako erabiltzen den energia 1 mJ-koa da, erabiltzen den laserra Quantel Ultra da eta 30 mJ-ko energiarekin (2.1.1.3 argazkia) emititzen duen uhin luzaera 1064 nm da. Laserra ablazio gunera ispilu hauekin eramaten da: 1064nm/45º eta CaF_2 35 cm lentearekin zuzenean laginaren zilindroan ardazten da.

2.1.1.3 Argazkia: Quantel Ultra Laserra.
http://www.bmisurplus.com/products/45549-big-sky-laser-ultra
(azkenengo sarrera 25, 2015).

2.1.2. Masen Espektrometroa

2.1.2 Argazkia: Hegaldi denborako masen espektrometroa.

Erabilitako masen espektrometroak Jordan TOF Products, Inc. oinarrizko diseinua du (2.1.2 argazkia), zenbait aldaketekin. Espantsio kameran laser desortzioko sistema bat du, honek funtzionatzeko beste bi leiho gehiago ditu, hauetatik ablazio laserra sartzeko, lagina mugitzeko pausoko motorra du (diseinua Prof. J. Simmon-en taldeak eginda da Oxford-eko unibertsitatean). Detektagailuak Jordan 18 mm bi plater aktibodun modeloa du eta 2500 V konektatuta dago. Hutsa bi ponpa turbomolekularren bidez lortzen da: lehenengoa detektagailu ondoan (Leybold Turbovac 151 N_2: 145 l/s) eta bigarrena espantsio kameraren azpian (Leybold TW 701 N_2: 680 l/s). Ponpa turbomolekularrak ponpa rotatorioetara konektatuta daude (TRIVAC D25B; 25,7 m^3/h) hauek tubomolekularren irteeran daude hasierako hutsa sortuz $2 \cdot 10^{-3}$ mbar. Bukaerako hutsa kameraren barruan $1 \cdot 10^{-7}$ mbar-koa da gas pultsurik gabe eta $1 \cdot 10^{-4}$ mbar-ekoa da balbula pultsatzen dabilenean. Neurtzeko egoera normala $5 \cdot 10^{-5}$ mbar-ekoa da. Hutsa neurtzeko Penning neurgailua dago (kameraren barruan) eta Pirani (rotatoriak baino lehen).

2.1.3. Tresna Elektronikoak

Orokorrean tresneri elektronikoak haien antzekotasunen edo erabileraren arabera sailkatuko dira. Bi talde nagusiak energi iturriak eta kontrolagailuak dira.

2.1.3.1. Energia Iturriak

2.1.3.1 Taula: Energia Iturriak

Funtzioa	Modeloa	Tentsioa
Erauzte platera	SRS STANFORD RESEARCH SYSTEMS, IMC. Ξ MODEL	4000 V
Azelerazio platera	PS350/5000V-25W Ξ HIGH VOLTAGE POWER SUPPLY	3700 V
Detektagailua	BRANDENBURG; Photomultiplier power Supply model 476R	2500 V
Biraketa platera (X_{1-2}&Y_{1-2})	Jordan TOF	0V

Energia iturrien propietateak (2.1.3.1 Taula) oso garrantzitsuak dira masen espektrometroa ondo funtzionatzeko: Plateretan dagoen tentsioa oso egonkorra izan behar da seinale eta zehaztasun handiena izateko. Arrazoi honengatik atal

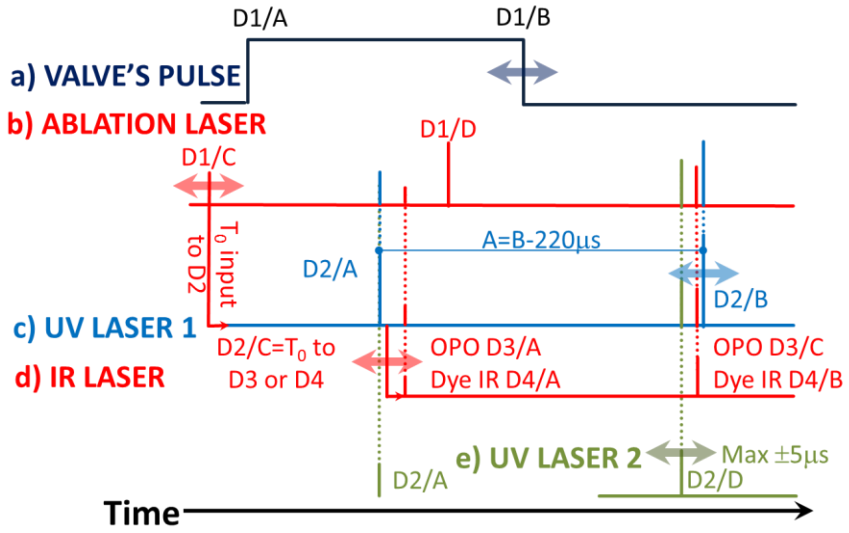
hauek oso kalitate onekoak izan behar dira; taulan aurkitu diren onenak erakusten dira.

2.1.3.2. Kontrolagailuak

2.1.3.2 taulan beharrezkoak diten kontrolagailu guztiak ikusgai daude, betetzen duten funtzioaren deskribapen labur batekin.

2.1.3.2 Taula: Kontrolagailuak

Elektronika	Egilea	MODELOA	Deskribapena
LAU KANALEKO DENBORA KONTROLAGAILUA	SRS STANFORD RESEARCH SYSTEMS, IMC.	DG535	Esperimentuan gertatzen diren gertakizun guztiak sinkronizatzeko
BI KANALEKO PULSO DIBISOREA	ISMO Workshop	ISMO SE032	Neurketa aktiboa ahalbidetzen duen tresna da
BALBULA KONTROLAGAILUA	GENERAL VALVE Corporation	IOTA ONE	28 tentsioko pultsu elektriko karratua
PAUSOKAKO MOTOREAREN KONTROLAGAILUA	OXFORD Workshop	KV0084b	Kontrolagarria den pausoz pausokako dotorea RS 440-442
OZILOSKOPIOA (1 aukera)	TEKTRONIX	TDS 3032	300 MHz eta 2.5 GS/s seinalearen jarraipena egiteko
OZILOSKOPIOA (2 aukera)	AGILENT Technologies	DSO9104A	1 GHz eta 20 GS/s seinalearen jarraipena egiteko
ORDENAGAILUA	3RSYSTEM	CPU: E5500 RAM: 4GB	Windows 7 (64 bits) profesionala Agilent VEE kontrolatzeko programarekin



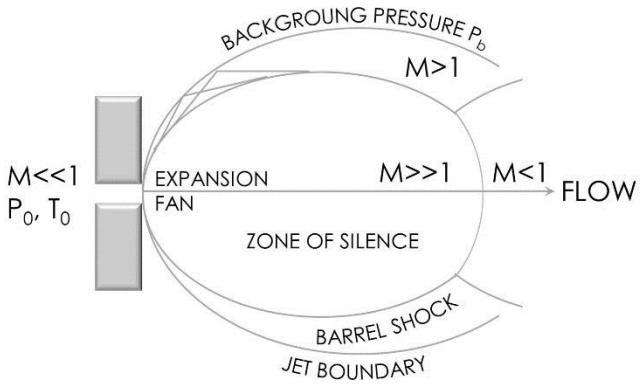
2.1.3.2.1 Eskema: Seinalearen optimizaziorako denboraren kontrol zehatza ezinbestekoa da a) balbula kontrolatzen duen pultsu elektrikoa; D1/B kanalarekin kontrolatu daiteke pultsuaren luzaera. b) D1/C kanalean ablazioa gas pultsuarekiko optimizatu daiteke. Kanal hau beste kanal guztiekin konektatuta dago laser guztiak batera mugitzeko. c) Ionizazio laserra; D2/B kanalean egokitu daiteke, IR eta UV laserrei konektatuta dago. d) IR Laserra; D2/C kanalean optimizatu daiteke eta normalean ionizazio laserra baino 50 ns lehenago kokatzen da. e) UV bigarren laserraren Q-Switch;

antolaketa honetan ez dago flash lamp-arekin konektatuta. Energia galerak ekiditeko (Fash lamp eta Q-switch)kanal honetan gehienez mugitu daitekeen tartea 5 μs-koa da.

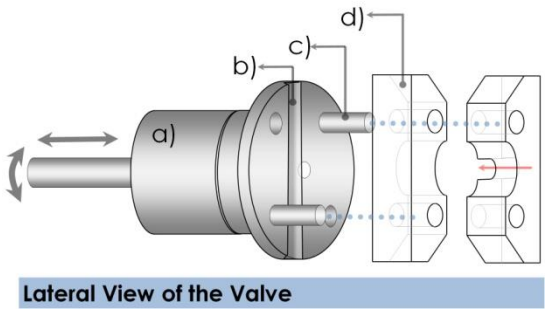
Denbora kontrolatzaileak garrantzitsuenetariko tresnen artean daude, sinkronizazio bera era askotan lortu daiteke. 2.1.3.2.1 eskeman esperimentuan gertatuko diren gertakari guztien sekuentzia erakusten da. Diagramaren ulermenerako eta informazio guztia eskuragarria izan dadin Table A2.1-an xehetasun guztiak erakusten dira.

2.1.4. Laginaren prestakuntza

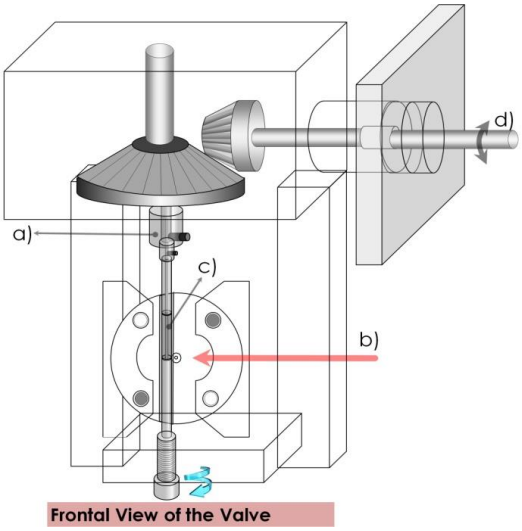
Isolatutako fasean egiten diren ikerketek lagina prestatzeko beharrezkoak diren bi tresna dituzte: balbula bat, espantsio supersonikoa sortzeko (2.1.4.1 irudian azaltzen da hauen funtzionamendua) eta ablazio laserra tenperaturarekin apurtzen diren molekulak baporatzeko (2.1.4.2 eta 2.1.4.3 irudiak ikusi). 2.1.4.2 irudiak balbula kokatzeko diseinatutako sistema berria erakusten du, honek laser ablazioa eta doiketa asko errazten du. Doiketarako beste bi hobekuntza egin dira konoaren atzeko leihoa eta 45º dagoen bideokamera.



2.1.4.1 Irudia: Presio altuan dagoen gas tutu estu batetik pasatzen da eta bertan azeleratu egiten da. Gasa tututik pasatu baino lehenagoko egoeratik (P_0, T_0), eta jarritako presio diferentzia (P_0-P_b) azelerazioa ematen da. Gas fluxua soinuaren abiadurara ailegatzen da (esan nahi duena, Mach 1 abiadura hartu duela) irteeran, P_0/P_b presio erlazioa balio kritiko hau gainditu behar du: $G\equiv((\gamma+1)/2)\gamma/(\gamma-1)$, 2.1 baino txikiagoa dela gas guztientzako, (γ fluido espezien propietate da). Erlazioa balio kritiko hori baino txikiagoa bada subsonikoa izango da eta ez da espantsiorik egongo. Normalean honelako iturrientzako fluxua isoentropikoa dela esan daiteke eta biskositatearen efektuak eta beroaren konduktibitatea mesprezatu daitezke. Jet espanditzen den heinean talken abiadura murriztu eta prozesu zinetikoak gelditzen dira, partaide guztian norabide eta abiadura berdinekin doazela, izoztuta eta isolatuta egongo bailiran (CI-Fasea), (xehetasun gehiagorako erref.[4]).



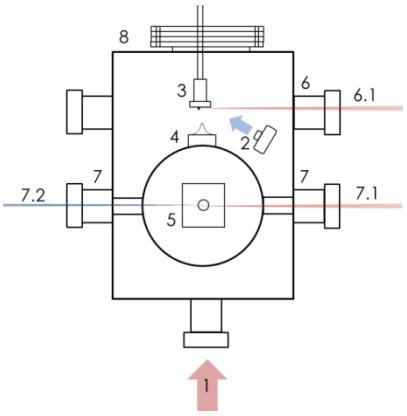
2.1.4.2 Irudia: Alboko balbularen konexioaren eskema. Diseinu berri hau egin behar izan da behin hutsa egina dela gas pultsua optimizatu ahal izateko. Ezkerreko geziak nola konektatzen eta deskonektatzen den azaltzen du. a) Pultsatutako Parker Balbula: Miniaturizatutako abiadura altuko eta huts alturako balbula, 28 V irekitzen den balbula Kel-F poppetarekin. b)Laginari lekua egiteko zuloa. c) Konektatzeko makila. d) Bukaerako horma.



2.1.4.3 Irudia: Aurretik ikusita balbularen eskema, laser ablaziorako tresna guztiekin. a) Laginaren zilindroa pausokako motorrarekin konektatzen duen ardatza. b) Ablazio Laserra. c) Grafitozko laginaren zilindroa. d) Pausokako motorra.

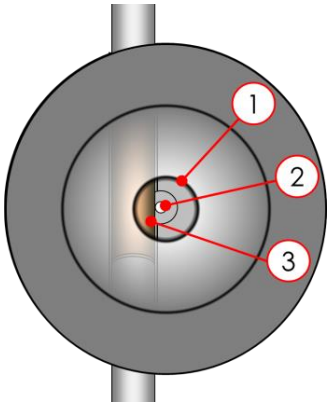
2.1.5. Ablazio laserraren doiketa

Sistema esperimentalaren doiketak arazo oso larria du, seinaleari begira egiten bada ablazio laserra posizio zuzenean jartzea oso zaila dela. Laser honek mugitzean lagin berria ablazionatzen du seinalea handituz, hau doitzeko beste metodologia bat egin behar izan da. Hurrengo irudietan azaltzen da nola doitu behar den. Xehetasunak 2.1.5.1 eskeman eta 2.1.5.2 eta 2.1.5.3 irudietan.

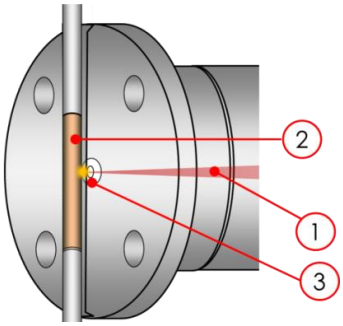


1) **Doiketarako 1º elementua (Leioa)**
2) **Doiketarako 2º elementua (Bideo kamera)**
3) Balbula eta laginaren zilindroa
4) Konoa
5) Masen espektrometroa erauzketa eta azelerazio platerak
6) Ablazio laserraren leihoa (6.1)
7) Ionizazio (UV, 7.2) eta despopulazio (UV edo IR, 7.1) laserrentzako leihoa
8) Espantsio kamera

2.1.5.1 Eskema: Espantsio kamerak goitik begiratuta, mota honetako edozein espektrometroren oiko diseinu du, ablazio laserraren doiketa egiteko bi tresna berri ditu (1, 2). Lehenengo tresna kameraren atzealdeko leihoa da (**Leioa 1**), bertan balbula eta ablazioa konoarekiko ondo kokatuta dagoen egiaztatu daiteke (2.1.5.2 irudia). Prozesu hau egiten den bitartean laser guztiak itzalita mantentzen dira segurtasuna mantentzeko. **2º** hobekuntza jarritako bideokamera da: Tresna hauei esker behin balbularen zuloa ondo kokatuta dagoela zuloaren aurrean ablazio laserra zuzen jartzea posiblea da (2.1.5.3 irudia).



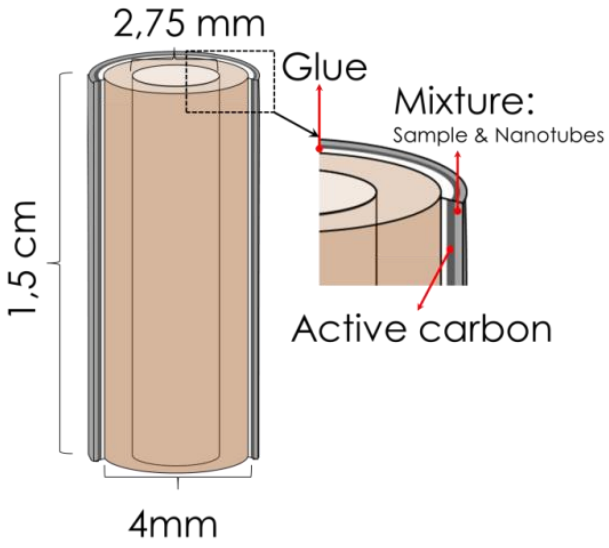
2.1.5.2 Irudia: 1) Konoaren zuloa. 2)espantsiorako zuloa. 3) laginaren zilindroa. Irudi honek doiketa zuzena erakusten du.



2.1.5.3 Irudia: 1) Ablazio laserra. 2) Laginaren zilindroa. 3) Balbularen zuloa eta ablazio laserra ondo doituta. Irudi honek bideokameratik ikus daitekeena erakusten du. Prozedura hau egiten den bitartean laserra martxan dabil baina erabiltzailea babestuta dago irudia kamerak hartzen du eta.

2.1.6. Laginaren Eransketa

Laginaren zilindroa grafitoz eginiko tutu bat da, lagina bertan jartzeko lehengo eta behin asken hau karbonozko nano-tutuekin nahastu behar da, gero zilindroarekin hautsaren gainean presioa eginez geruza mehe bat jartzen zaio zilindroari. Eransketa hobetzeko askotan lagina jarri baino lehen karbono aktiboaren geruza bat itsasten da, gero lagina hobeto itsatsiko da karbono aktiboaren gainean. 2.1.6.1 irudiak geruzak nola geratu behar diren erakusten du. Lagin likidoak ere tanta bat gehituta egin daitezke.[5]



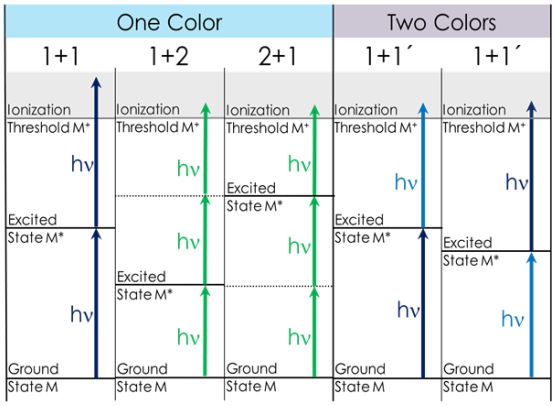
2.1.6.1 Irudia: Laginaren zilindroa eta lagina jartzeko geruzak.

2.2. Metodologia Esperimentala

Erabiliko diren datu esperimentalak ionizazio prozesuan zehar lortu egiten dira (**Step 3**, 2.1.1 irudian) hauek laser konbinaketa desberdinen artetik datoz. Atal honetan erabiliko diren teknika espektroskopikoak azalduak izango dira, erresonantzia bakarreko tekniketatik hasita, laser bakarrarekin aski dutenak, erresonantzia anitzeko tekniketara. Azken honetan badago era bat baino gehiago informazio berdina lortzeko. Nahiz eta erresonantzia hirukoitza eta laukoitza[6] konformeroak bereizten erabilgarritasun handia erakutsi izana, proiektu honetan erresonantzi bakarreko eta erresonantzia bikoitzeko teknikak erabiliko dira denbora eta dirua aurrestuz.

2.2.1. *"Resonance Enhanced Multi-Photon Ionization"* (REMPI)

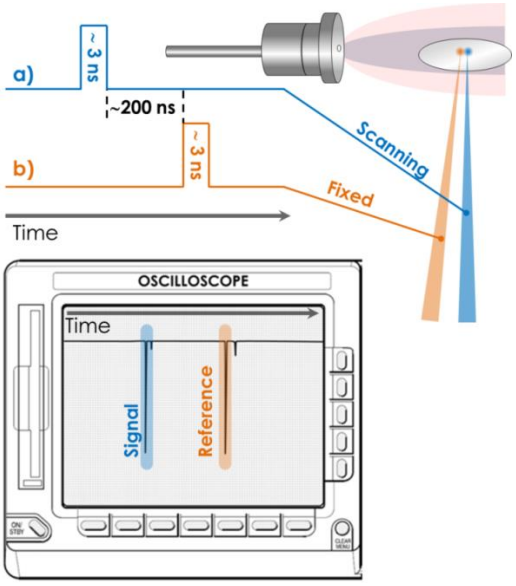
Teknika honetan zenbait fotoien adsortzioaren bidez ionizatzen ditu molekulak, fotoi baten energia bitartekari maila batekin bat doanean ionizazioan igoera handia ematen da. Fotoiak askoz posibilitate gehiago ditu adsortzioa emateko honen energia kitzikatutako maila batekin bat egiten badu. 2.2.1.1 irudiak zenbait adibide erakusten ditu, konbinazio asko daude ionizatzeko eta. Neurketa honen bidez espektro elektronikoa egiten da, kitzikatutako mailei buruzko informazioa lortuz. Baina proiektu honetarako REMPI espektroskopia konformeroen aurrean bereizteko soilik erabiliko da. Konformero desberdinek REMPI espektro desberdina dutenean hauen artean bereizi daitezke, uhin luzaera ondo aukeratzen bada, bata edo bestearen IDIRS espektroa lortu daiteke. Azken teknika honek (IDIRS) erresonantzia bikoitza erabiltzen du eta aurrerago azalduko da.



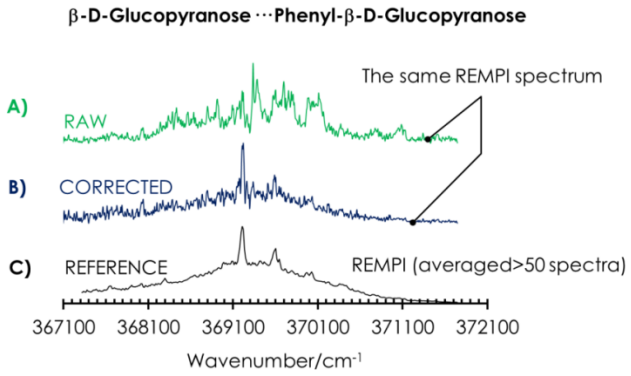
2.2.1.1 Irudia: REMPI espektroskopiaren fotoien adsortzioaren eskema. Bi talde nagusi daude: kolore bat edo fotoi bakarrekoak (1+1) eta bi kolorekoak edo bi fotoi desberdin erabiltzen duena (1+1´). Molekula dituen kitzikatutako mailen arabera prozesua era desberdin askotan eman daiteke: 1+1 kasua, fotoi bat kitzikatzeko eta beste bat ionizatzeko erabiltzen du (bi fotoiak berdinak dira). 1+2 kasua, lehenengo fotoiarekin kitzikatu egiten da baina gero ionizatzeko bi fotoien adsortzioa behar da. 2+1 kasua, lehenengo bi fotoiak kitzikatzeko erabiltzen ditu eta bat ionizatzeko. Kasu batzuetan ionizazioa apurketa mailatik aurrerago doa, hau gertatzen denean konplexu handiak apurtu eta txikiagoen espektroan ekarpenak izan ditzakete. Egoera hau gertatzen denean bi koloreko teknikak erabiltzen dira: lehenengo kolorearen intentsitatea murrizten da soilik fotoi bakarraren absortzioa ahalbidetzeko, bigarren laserraren fotoiaren energia egokitzen da ionizazio eta apurketaren artean egoteko eta ez gorago (cases 1+1´).

2.2.1.1. Zuzenketa aktiboa REMPI espektroskopian

REMPI espektroskopiaren arazorik handiena kasu askotan ikus daitekeen seinale/zarata proportzioa da, batez ere ablazio laserra eta molekula handien espektroskopia egitean. Arazo honi aurre egiteko zuzenketa aktiboko teknikak erabili daitezke. Proposatutako metodo honek beste laser bat erabiltzen du momentuoro espantsioan molekulen kontzentrazioa ebaluatzeko (2.2.1.1.1 irudian). Erreferentziako seinalea erabilita espektro gutxiagoekin REMPI on bat lortzea errazagoa da, esperimentu guztiak egiteko lagin gutxi (miligramo batzuk) dagoenean hau oso erabilgarria da. 2.2.1.1.2 irudian adibide bat erakusten da, bertan tratatu gabeko espektroa (RAW) tratatuarekin eta erreferentzia batekin alderatzen dira.



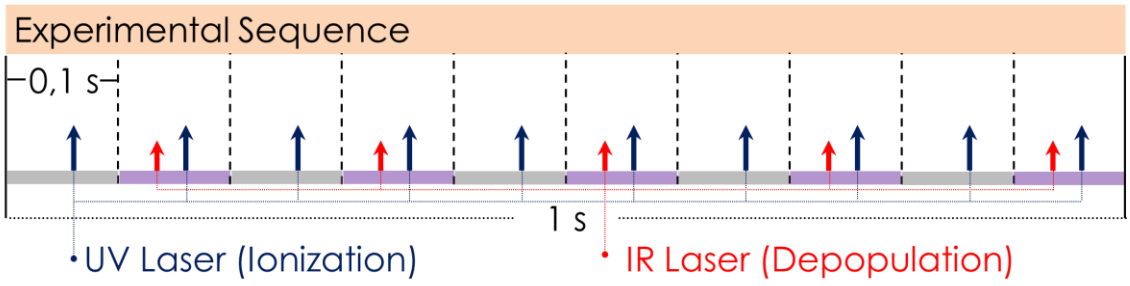
2.2.1.1.1 Irudia: (a) eta (b) laserren arteko Q-Switch sinkronizazioa erakusten du. a) REMPI espektroa jasotzeko erabiliko den laserra espazioan erreferentzia laserraren atzetik dago (balbularen irteeratik urrunago) baina denboran 200 ns lehenago kokatuta dago. Konfigurazio honetan nahiz eta bi zonaldeak ia berdinak izan laser batek ez du bestearengan influentziarik izango. b) Erreferentzia laserra uhin luzaera finko batean dago, uhin luzaera hau aukeratzeko aldez aurretik frogako REMPI bat egitea gomendagarria da adsortzioa erakusten duen zonalde bat aukeratzeko. Garrantzitsua da bi laserrak intentsitate "spot" tamaina bera izatea, bestela teknika honek ez du ondo funtzionatuko. Kontuz ibili behar da denbora luzean gertatzen diren aldaketekin, laser batek intentsitatea irabazi eta bestea galdu dezakete.



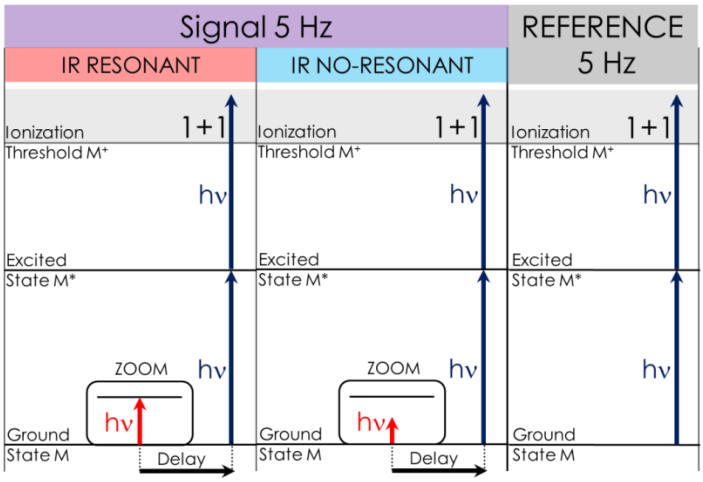
2.2.1.1.2 Irudia: A) REMPI espektro normala. B) Zuzendutako REMPI espektroa, erreferentziako ionizazioa 36100-37100 cm^{-1} tartean finkatu delarik; zuzenketa egiteko RAW REMPI spectrum erreferentziarengatik zatitzen da. C) REFERENCE erreferentzia 50 REMPI espektroen batezbestekoa da.

2.2.2. Infragorrien Ioien despopulaketaren espektroskopia (IDIRS)

Teknika honek zuzenean zuzenketa aktiboa egiteko aukera ematen du, baina zuzenketa aktibo hau egiteko era asko daude. Teknikaren oinarrizko funtzionamendua hurrengoa da: UV laserra REMPI espektrotik aukeratutako trantsizio batetan finkatzen da 10Hz. 5Hz-tara IR laserra gehitzen da hau molekula berdinekin eragiteko doituko da espazioan eta denboran, gero IR laser hau 50-100 ns lehenago sinkronizatuko da. 2.2.2.1 irudian sekuentzia ikus daiteke, antolaketa honek seinale a eta erreferentzia batera hartzea baimentzen du. Espantsio supersonikoan molekula guztiak oinarrizko maila bibrazionalean daude, IR trantsizio batekin bat doanean molekula kopurua bat maila horretara eramango ditu, gero UV laserra datorrenean populazio gutxiago ikusiko du eskuragarri oinarrizko mailan eta horren ondorioz ioi gutxiago sortuko dira 2.2.2.2 irudia ikusi. Azkenean seinalearen datuak erreferentziarengatik zatitzen dira.



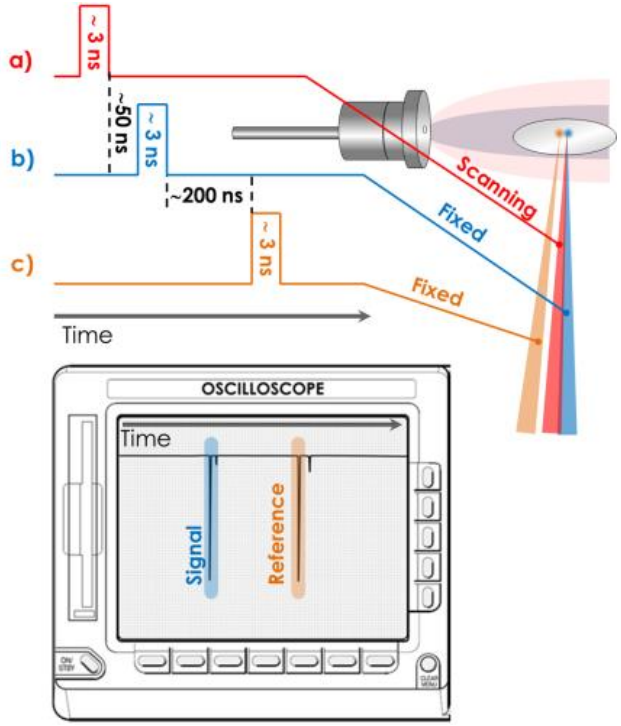
2.2.2.1 Irudia: Gezi urdina: UV laserra 10 Hz-tan funtzionatzen du. Gezi gorria: IR despopulaketa laserra 5 Hz-tan funtzionatzen du. Grisezko atalak: erreferentzia, bertan soilik ionizazio laserra dago. Morezko atalak: seinalea, bi laserren konbinaketaren ondorioa da.



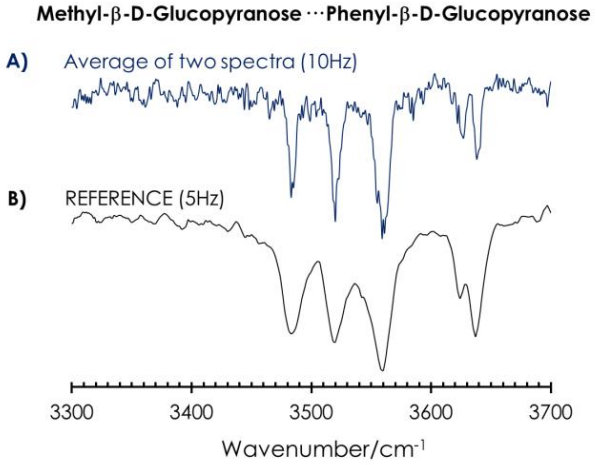
2.2.2.2 Irudia: Erreferentzia grisez: adsortzio multifotonikoa finkatutako uhin luzaeran. Morean seinalea (5 Hz): bi posibilitate daude, lehenengoa IR laserra erresonantea ez denean ez da ezer gertatzen eta seinalea konstante mantentzen da (urdinez adierazita); bigarren aukera IR laserra erresonantea denean ematen da, populazioa beste maila batera mugitzen du eta horren ondorioz honen ondoren datorren UV molekula gutxiago ditu ionizatzeko seinalea txikiagotuz (gorriz adierazitako atala).

2.2.2.1. 10 Hz IDIRS Espektroskopia

Lehenago REMPI espektroskopian azaldutako (2.2.1.1 atalean) zuzenketa aktiboko teknika IDIRS espektroskopian ere erabili daiteke. Desberdintasun bakarra UV laserra IR laser bategatik aldatzen dela, eta azken hau da espektroa lortzeko erabiliko den laserra. 2.2.2.1.1 irudian antolaketa erakusten da eta 2.2.2.1.2 irudian adibide bat.



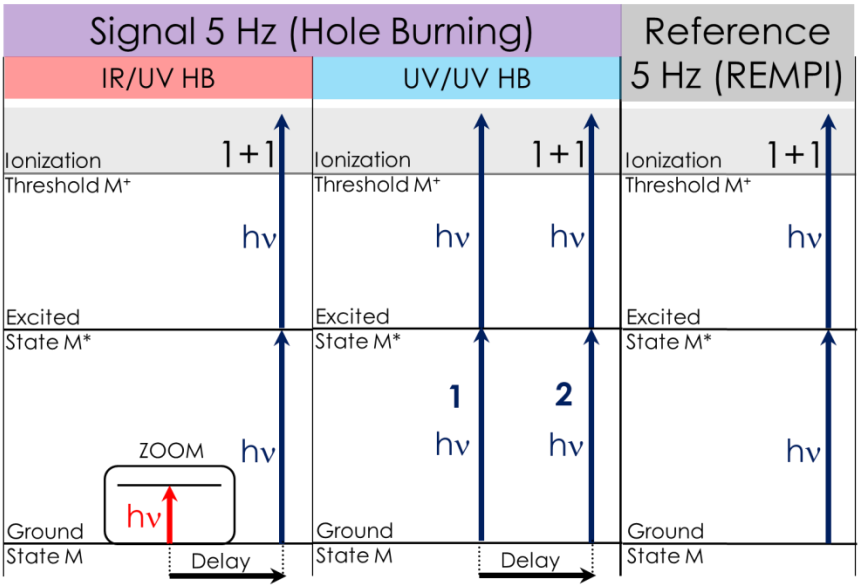
2.2.2.1.1 Irudia: Antolaketa esperimentala 10 Hz-ko IDIRS zuzenketa aktiboko esperimenturako. a) IR laserra ekorketa egitean ioien sorkuntzaren murrizketa sortzen du erresonantea denean, 50 ns beranduago UV laserra dator. b) Finkatutako UV laserra, IR laserra erresonantea denean honek ioi gutxiago sortzen ditu. c) Finkatutako bigarren UV laserra, hau erreferentziaren papera betetzen du: laser hau berdina den zonalde bat ionizatzen ari da baina ez du beste IR eta UV laserren influentziarik nabaritzen.



2.2.2.1.2 Irudia: A) 10 Hz esperimentuan egindako bi espektroen batezbestekoa. B) Erreferentzia 10 espektroen batezbestekoa.

2.2.3. *"Hole Burning"* Espektroskopia (HB)

Teknika honek konformero bakoitza REMPI espektroan duen ekarpena bereizteko balio du. Beste kasuetan bezala era bat baino gehiago dago esperimentu hau aurrera eramateko, gehien erabilitako metodoak IR/UV eta UV/UV dira. IDIRS espektroak lortzeko antolaketa berdina du baina denboran lehenengoa den laserra aldatu daiteke (2.2.3.1 irudia). Esperimentuen sekuentzia 2.2.2.1 irudian azaldutakoaren berdina da baina kasu honetan 5 Hz-tan funtzionatzen den IR laserra finko dago trantsizio batetan. (5 Hz) IR laserra UV (5 Hz) laser batengatik aldatu daiteke baina hau REMPI espektroaren trantsizio batetan jarri behar da (urdinez adierazitako UV/UV HB 2.2.3.1 irudian). Bi antolaketetan informazioa bigarren laserrarekin REMPI espektroa batzean lortzen da. Maiztasunean desberdinduta bi informazio desberdin lotzen dira, (5 Hz) soilik ekorketa egiten ari den laserra dagoenean REMPI espektroa lortzen da, (5´ Hz) bi laserrak daudenean erreketa laserra finkatuta dagoen trantsizioa duten konformero guztiak seinalearen murrizketa izango dute baina beste guztiak ez. Bien arteko diferentzia egitean erre den edo erre diren konformeroen REMPI espektroa lortzen da. Teknika hau oso erabilgarria da zenbat korfomero dauden jakiteko eta gero IDIRS espektroak egiteko konformazio bakarra ionizatzen duen uhin luzaera aukeratzeko.



2.2.3.1 Irudia: UV laserra 10 Hz funtzionatzen du (2.2.2.1 irudian bezala) baina kasu honetan ere ekorketa egiten dabil. Erreferentzia, 5 Hz-tan funtzionatzen, REMPI espektroa neurtzen du. Seinalea, moreaz margotuta, 5 Hz-tan REMPI espektroa neurtzen ari du baina lehenago dagoen laserra konformero bat edo batzuk erretzen ari ditu, hauei dagokion bandak txikiagoak izango dira.

2.3. Metodologia Teorikoa

Kalkulu teorikoen helburu nagusia egitura egonkorrak aurkitzea eta hauen energia erlatiboa zein den determinatzean da. Egiturak esleitzeko iragarpen teoriko eta neurketa esperimentalak alderatu egiten dira. Kalkulu teorikoak ondo egiteko prozedura, egitura posibleen bilaketatik hasita, energiaren, egituraren optimizazioraino eta asken esleipenerarte, hainbat urratsez osatutako prozesua da. Metodologia zuzenena aukeratzeko lehenengo egin behar dena eragiten ari diren aldagaiak identifikatzean da. Hiru talde nagusitan sailkatu daitezke prozedurak: 2.3.1) Aldakortasun Kimikoa; Konplexuak eta molekulen malgutasunari lotuta dagoena. 2.3.2) Kalkuluen zehaztasuna; egitura posibleen bilaketan eta honen ondoren datorren DFT kalkuluei zuzenduta dago. 2.3.3) Giro Esperimentala; IC-fasean dauden berezitasunei zuzentzen da. Aldagai hauen garrantzia orokorrean handia da, kasuaren arabera kalkuluak egiteko era guztiz aldatu daiteke.

2.3.1. Aldakortasun Kimikoa

Atomoen konbinazioak molekula mota izugarriak sortu ditzake, hauetako bakoitza bere berezitasun eta ezaugarriekin: ziklo zurrunetatik hasita malgutasun handiko molekula linealetara. Aniztasun hau atomoen arteko elkarrekintzen ondorioa da eta bi talde nagusietan banandu daitezke: 2.3.1.1) Aldakortasun intramolekularra, molekulak berez duen ezaugarriak mugatzen duena eta 2.3.1.2) Aldakortasun intermolekularra, hauek ez-kobalenteak diren elkarrekintza ahulek arautzen dituzte.

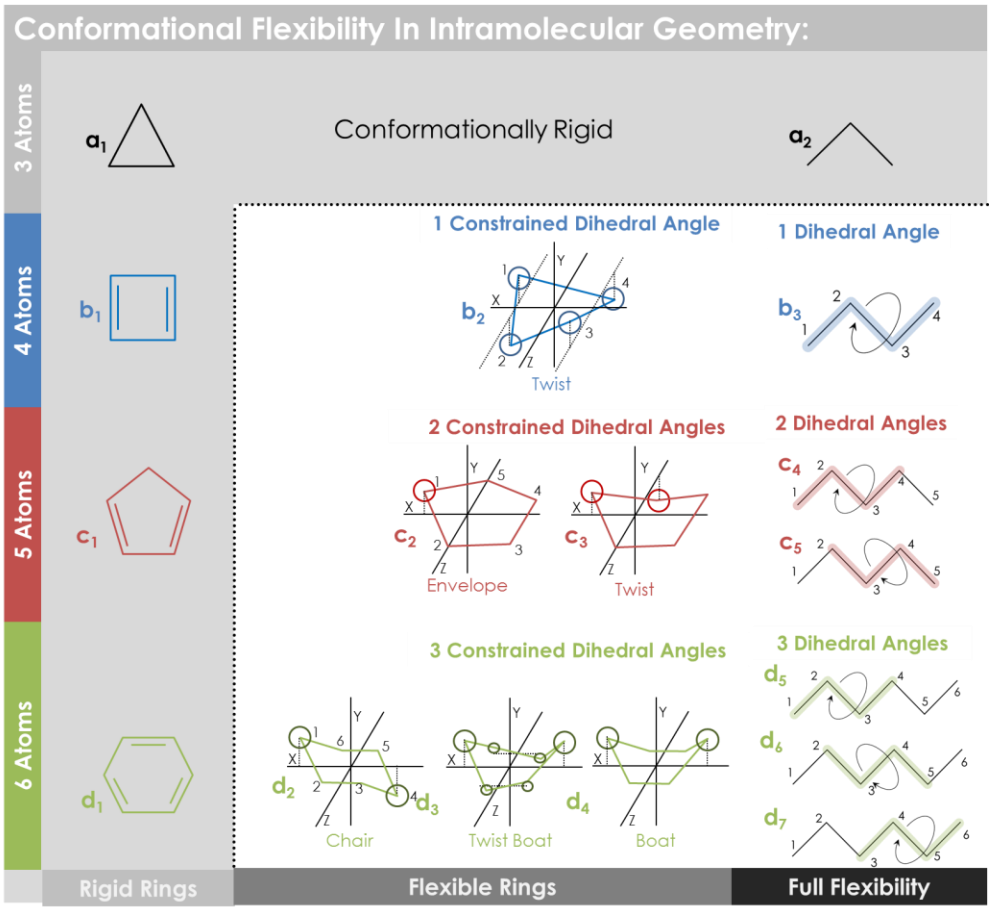
2.3.1.1. Aldakortasun Intramolekularra

Malgutasun intramolekularra aldagai askok sortzen dute baina aldaketa esanguratsuenen artean angelu diedro baten aldaketa da. Badaude beste aldaketa batzuk tautomeriak eta protoien transferentziak eta horrelako prozesuak non lotura kobalenteak apurtu eta berregiten diren. Azken hauek ziklo zurrunetan ere egon daitezke. Aldaketa hauek ez dute energia asko behar eta horrek kontutan izateko beharra eragiten du.

2.3.1.1.1. Angelu Diedroen Malgutasuna

Molekulen propietateak atomoen karga dentsitateari (intramolekularrak diren hidrogeno loturak eta abar sortzeko gaitasuna ematen diona) eta lotura

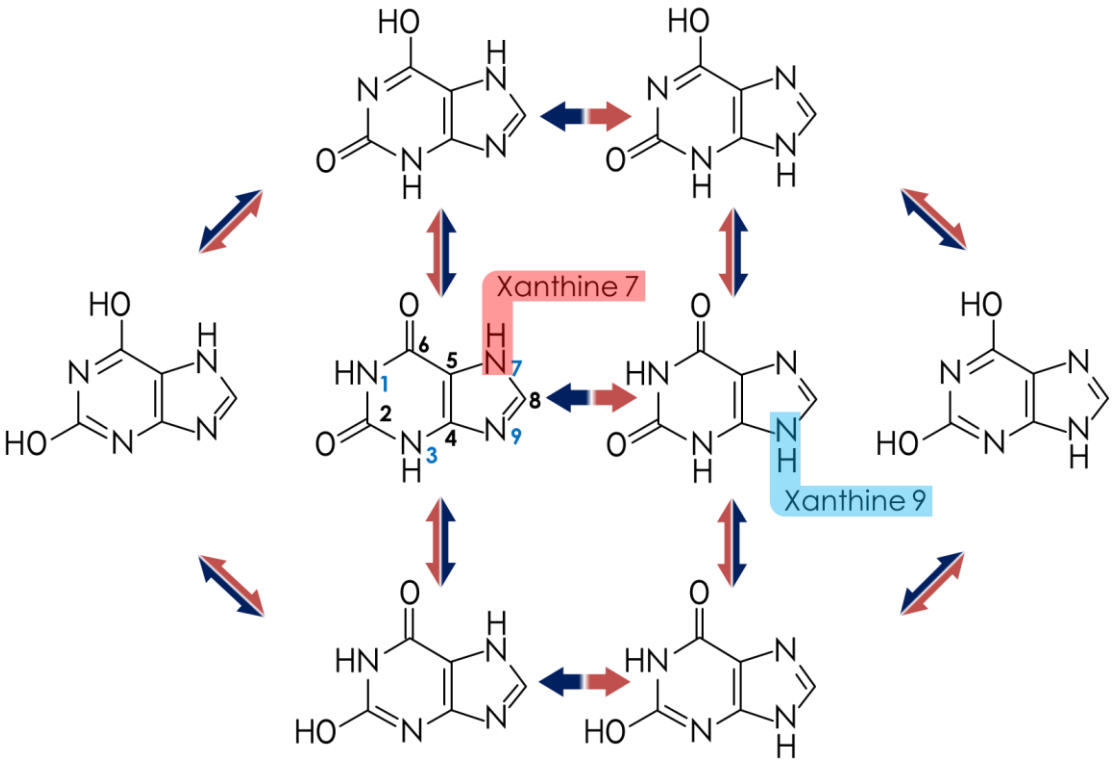
kobalenteen malgutasunari lotuta daude. Lehenago aipatu den bezala angelu diedro baten biraketa da hau arautzen duen ezaugarri nagusia. 2.3.1.1.1 irudian ikus daiteke orokorrean propietate honi dagokion diagrama, atomo kopuruaren arabera antolatuta dagoelarik. Irudian ikus daitekeenez zurrunak diren konposatuetatik hasita aldaketa txiki batzuekin konplexutasuna asko igotzen da. Figure A2.2, A2.3, eta A2.4 (A dutenak apendizean daude) irudietan ikus daiteke familia hauei dagokien eraldaketa hesiak 2.3.1.1.1 irudiko b2, c2-3 eta d2-3-4. Gainera Figure A2.5 irudian nola kate batean karbono atomo bat gehitzeak isomero kopuruaren biderkatzea eragin dezakeen ikus daiteke alkohol linealak eredu gisa hartuta.



2.3.1.1.1 Irudia: Ziklo zurrunak: a1, b1, c1 eta d1hauentzako ez dago isomerorik. **Ziklo malguak:** b2, c2, c3 eta d2, d3, d4 Zenbait isomero izan ditzakete. Borobilduta dauden atomoak *XZ*-planotik kanpo daudenei dagokie. Ziklo eran egoteak angelu diedroaren mugimendua mugatzen du, gehienez 90º angelua izan dezaketelarik. **Mugarik gabeko Malgutasuna:** b3, c4, c5 eta d5, d6, d7. Lau atomo duten molekula linealetatik hasita (b3) non angelu diedro bakarra dagoen, angelu berri bat gehitu daiteke atomo berri bat gehitzen den bakoitzean (c4 eta c5: 5 atomo eta 2 angelu diedro; d5, d6 eta d7: 6 atomo eta 3 angelu diedro, eta honela jarraitu daiteke). Molekula organikoetan 360º errotazioa 3N konformazio egonkor sortzen ditu, N karbono kopurua minus 3 da. Aldarapen esterikoa dela eta alkilo katea hiru egoera posible ditu, anti, gauch eta gauche, hauek angelu hauei dagokie: 180°, 60° and -60°, molekula luzeetan beste fenomeno batzuk gerta daiteke, adibidez gauche (60°) bi egitura egonkor berritan banatzen da (~30° eta ~90°) honen adibidea Figure A2.5 irudiko AG_EG2-AG_EG2d & G_EG2G-G_EG2G2d konformazioak dira.

2.3.1.1.2. Isomerizazio Intramolekularra

Beste aldakortasun kimiko mota bat molekula barnean eman daitezkeen isomerizazio prozesuak dira, baita molekula zurrunei eragiten dietenak. Esanguratsuenen artean oreka tautomerikoa dago. Hauetan normalean NH edo OH-en hidrogenoa molekula barnean beste leku batera mugitu daiteke. 2.3.1.1.2 irudian adibide bat erakusten da: Xantinaren (base purikoa) isomerizazioa, N7 eta N9 artean protoia mugitu daiteke eta amida-imida artean ere tautomerizatu daiteke(laktan-laktin pareak posizioen artean Ox-Nx eta Oy-Ny) .

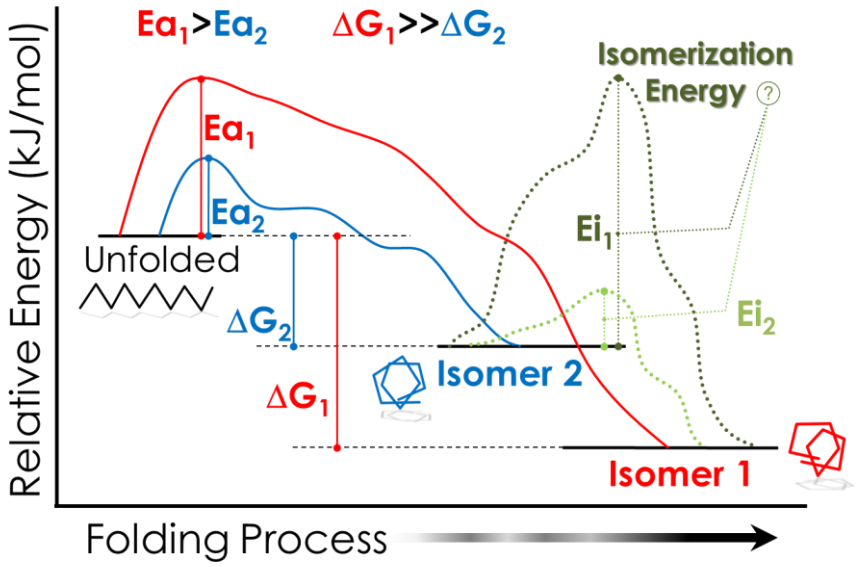


2.3.1.1.2 Irudia: Aldakortasun kimikoa molekula zurrunetan: xantina. Tautomero posible guztiak irudian erakusten dira.

2.3.1.1.3. Toleste Prozesu Intramolekularra

Prozesu hau ulertzea oso zaila da, molekula luzea denez momentu berean talde funtzional asko ditu elkarri eragiten. Molekula egingo dituen lotura intramolekularrak eragin handia izango dute bukaerako egituran. Toleste prozedura ulertzeko zenbait pausotan banandu daiteke: hasiera batean haien

artean hurbil dauden atalak lotura ahulak eratzen dituzte, honen ondoren zonalde hauek elkarrekintzak eratzen dituzte molekula egonkortuz eta toleste prozesuari bukaera emanez. Tolesteko prozedurak egitura egonkorrenera ez ailegatzea eragin dezake bertara ailegatzeko bidea energetikoki debekatuta dagoelako. 2.3.1.1.3 irudian dagoen adibidean ikus daiteke molekula tolesteko urrats bakoitzak energian duen eragina, horregatik askotan bitartekari egonkorrak (Ea_2) dituen egiturak probabilitate gehiago izango ditu eratzeko. Egitura egonkorrenak bitarteko egitura ezegonkorretatik pasatu beharra (Ea_1) blokeatzen du, baina beste mekanismo bat dago egitura egonkorrena hartzeko, behin tolestuta dagoela eraldatzeko energia hesia altuegia ez bada gerta liteke egitura egonkorrenera eraldatzea. Askotan eraldaketa bidea oso konplexua da (Ei_1 or Ei_2) eta kalkuluen bitartez iragarpen on bat egitea ezinezkoa da. Gai honi buruz bosgarren kapituluan hitz egingo da.



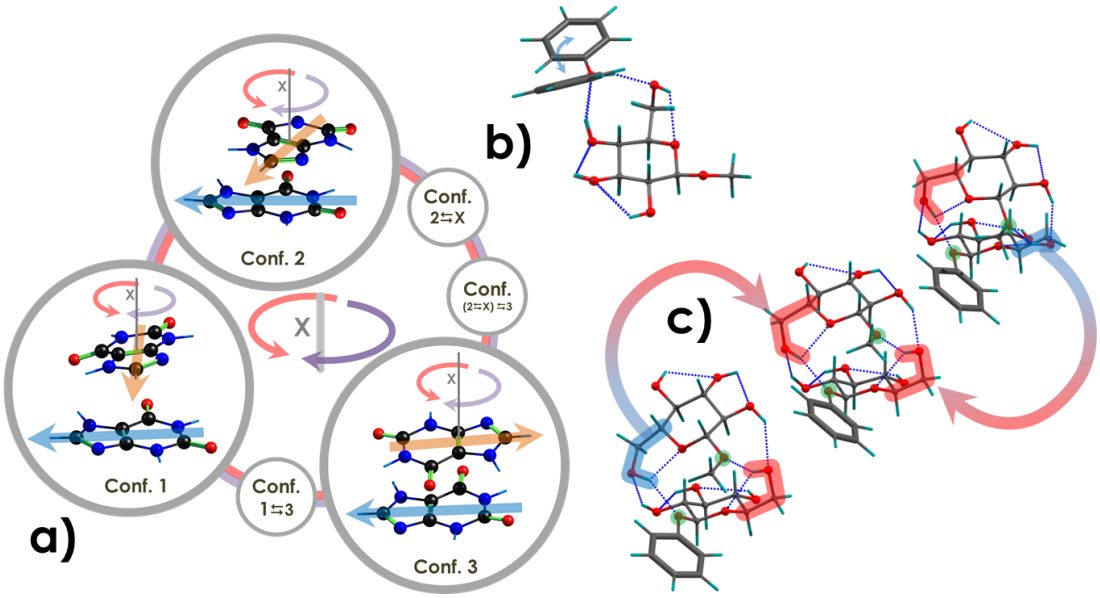
2.3.1.1.3 Irudia: Toleste prozesua bi isomeroentzako: Isomer 1 isomeroak aktibazio energia altuagoa du (Ea) baina sortzen duen egitura bigarren isomeroa baino (isomer 2) egonkorragoa da. Isomerizazio energiak (Ei) zein egitura lortuko den determinatzen du: (Ei_2) energia baxua bada, 2 isomerotik lehenengo isomeroa lortu daiteke 1. (Ei_1) Hesia altua bada 2 isomeroa lortuko da.

2.3.1.2. Aldakortasuna Intermolekularra

Aldakortasun intermolekularrak konplexua eratzen denean egitura aldatzeko gaitasunari buruz erreferentzia egiten du. Hiru talde nagusitan sailkatu daitezke: 2.3.1.2.1) Lotura guneko malgutasuna; 2.3.1.2.2) Lotura guneen arteko lehiaketa; eta 2.3.1.2.3) Bi molekula baino gehiagoko konplexuetan dagoen efektu estatistikoa.

2.3.1.2.1. Lotura Guneko Malgutasuna

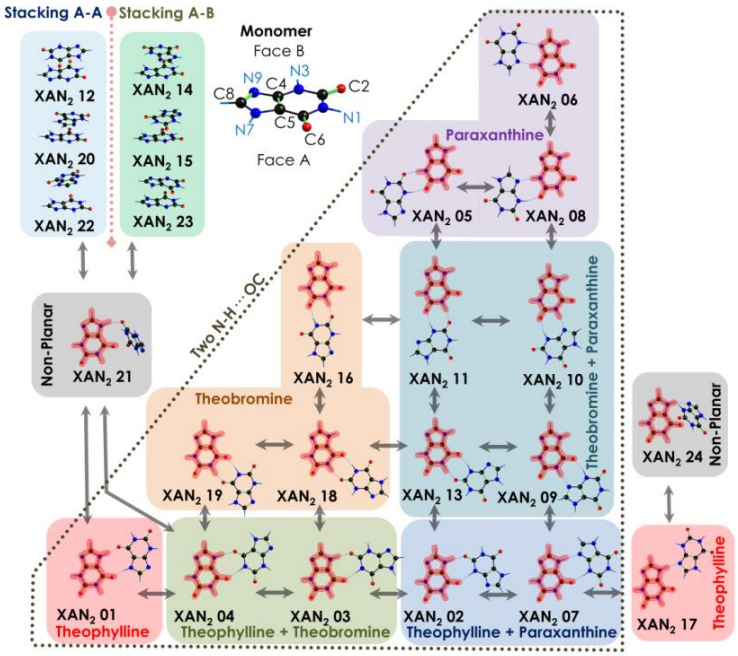
Lotura guneko malgutasuna elkarrekintza eta hau kokatzeko duen posibilitateekin lotuta dago. Loturaren ezaugarriak elkarrekintzaren (PES) energi potentzialaren gainazala ulertzeko balio du. Lotura ahul baten malgutasuna nolakoa den jakiteko bi ezaugarri izan behar dira kontutan, alde batetik molekula bien arteko elkarrekintzarako dauden egitura egonkor posibleen kopurua eta hauek konektatzen duten eraldaketa bideak. Orduan elkarrekintza baterako egitura egonkor asko daudenean eta hauek energi hesi txikiengatik konektatuta daudenean elkarrekintza oso malgua dela esan daiteke. Malgutasuna handia denean hautagai asko egongo dira, batez ere hesi txikiengatik konektatutakoen artean. Teknika eta sistemaren arabera batzuetan hauen artean bereizi daiteke baina beste kasu askotan ezin da. 2.3.1.2.1 irudian zenbait adibide erakusten dira: a) pilaketa-konformazioan dagoen xantina dimeroaren malgutasuna; b) Fenol eta glukosaren arteko dimeroan fenolak duen malgutasuna zikloa elkarrekintzarik izan edo ez izateko; c) Monomeroen egituraren efektua dimeroaren eraketan, (Fenil/metil-β-D-glukopiranosa). Adibide hauek azalduko diren kapituluetan ikusgai egongo dira.



2.3.1.2.1 Irudia: a) Xantinaren pilatutako dimeroak, pilaketa ardatza irudikatu da konplexu bakoitzaren goialdean; beheko xantina beti posizio berean mantentzen da, goiko x ardatzaren arabera biratzen du. Egituren artean bitarteko egituraren bat egon badaiteke borobil txiki bat gehitu da, bat baino gehiago badaude bi jarriko dira. Esperimentalki hauen artean bereiztea oso zaila da eta oso erraza da akatsak izatea. b) Fenola eta metil-β-D-glukopiranosaren arteko elkarrekintza: fenolaren eraztun aromatikoa glukosaren gainean egon daiteke pilatuta edo beste aldera begira glukosarekin ukitu gabe, kasu honetan ezin izan dira desberdindu. c) Glukosa dimero: pilota berdeek karbono anomerikoa markatzen dute, urdinez azpimarratutakoa GA geometria adierazteko erabiltzen da eta gorria GG kasurako (G= gauche, A=anti), adibide honetan bibrazioetan eragin handia duenez bien artetik desberdindu daiteke neurketa esperimentala egin ondoren.

2.3.1.2.2. Lotura Guneen Arteko Lehiaketa

Molekula baten lotura gunea edozein lekutan egon daiteke, eta guztiak kontutan izateko molekula bakoitzaren inguruan esfera bat jarri beharko litzateke; molekula baten esferak bestearekin elkarrekintzak egin ditzake eta euren arteko konbinaketak posibilitate kopurua gehiegizko lanean bihurtzen dute. Gainera molekula malgua bada eta egitura bat baino gehiago badu honek faktore batengatik biderkatzen du konformazio kopurua. Egoera nahasi honi aurre egiteko zenbait sinplifikazio egin daitezke, aldarapen indarrak eta simetria bi sinplifikazio garrantzitsu dira baina benetakoa lotura gune garrantzitsuen identifikaziotik dator. Lotura gune garrantzitsuenak besteekiko egonkortze handiago bat eragiten dute, lotura guneen artean lehiaketa egon daiteke baina kasu gehienetan bat edo gehienez bi dira nabarmenki nabaritzen diren erakarpenak. 2.3.1.2.2 irudiak xantinaren dimeroaren lotura-gune posibleen diagrama erakusten du.

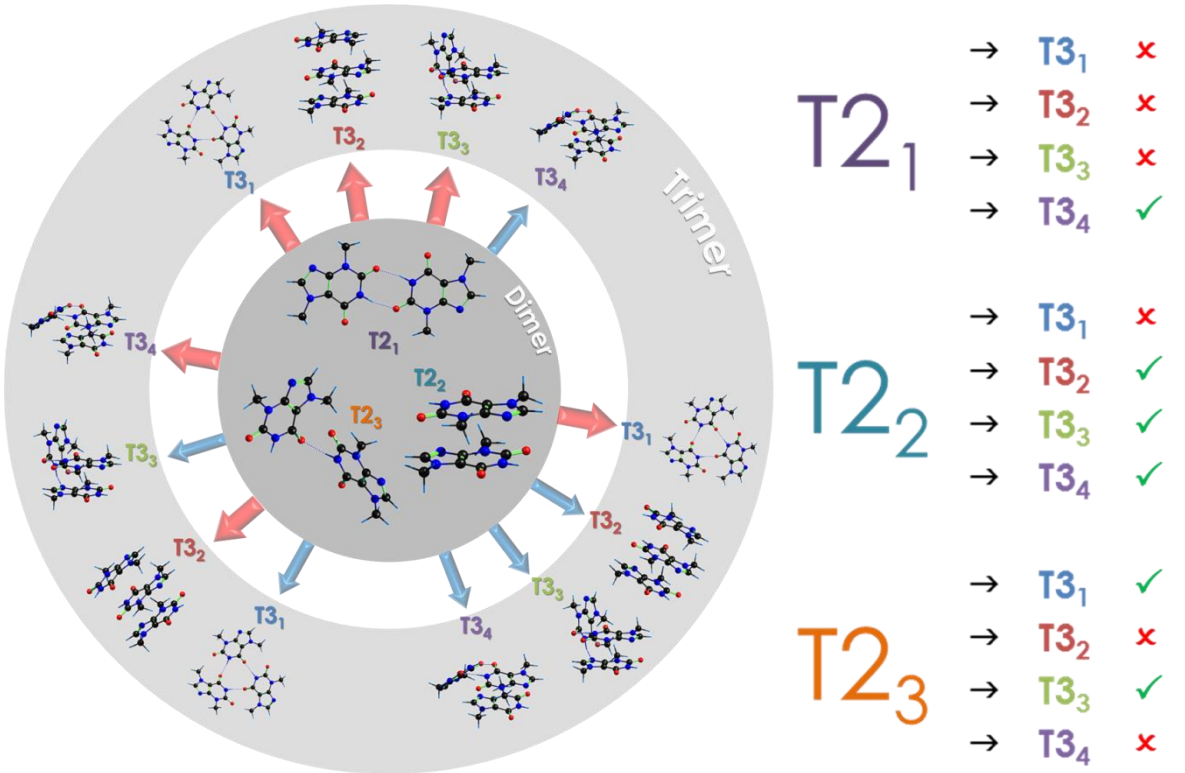


2.3.1.2.2 Irudia: 7 Xantinaren lotura-gune garrantzitsuenen diagrama orokorra.

2.3.1.2.3. Bi Molekula Baino Gehiagoko Konplexuetan Dagoen Efektu Estatistikoa

IC-fasean konplexuaren eraketa egiten denean kontutan izan beharreko parametroak daude, hauek posibleak diren konplexuak mugatzen dituzte nahiz eta egonkorrago den bat egon. Trimero bat sortzeko dimero eta monomero baten

batuketa behar da, dimero egonkorrak ez du zertan trimero egonkorrenaren aitzindari izan behar. Tetrameroari begira gehiago zailtzen da zeren trimero eta monomero baten konbinaketaz edo bi dimeroren batuketaz sor daiteke. Ikusitako arau hauek trimeroak mugatzen dituzte baita tetrameroak eta hortik aurrera dauden konplexu guztiak. 2.3.1.2.3 irudian teobrominaren dimerotik sortu daitezkeen trimeroak irudikatu dira. (seigarren kapituluan benetako emaitzak erakusten dira)



2.3.1.2.3 Irudia: Konbinaketa estatistikoak, dimero eta monomerokoak batzen direnean zein trimero sortu dezaketen diagrama. Trimeroaren eraketa eraldaketa energia hesiengatik mugatuta dago. Diagrama honetan adibide moduan erakusten dira, benetako emaitzak ikusteko seigarren kapitulura jo.

2.3.2. Kalkuluen zehaztasuna

Atal hau bi urratzetan banandu daiteke: 2.3.2.1) Bilaketa Konformazionala eta 2.3.2.2) Egituraren eta energiaren optimizazioa eta frekuentzien kalkulua. Lehenengoak Mekanika molekularra erabiltzen du egitura posible guztiak aurkitzeko. Bigarren urratzean lortutako egiturak "*ab-initio*" edo DFT metodoekin egituraren optimizazioa eta energiaren balio esanguratsuak lotzen dira. Gehien erabilitako kalkulu metodoa DFT M06-2X[7] da, erabilitako programa Gaussian 09.[8]

2.3.2.1. Bilaketa Konformazionala

Erreferentziarik gabe oso zaila da mekanika molekularraren bidez egindako kalkuluen zehaztasuna jakitea. Horregatik lehenengo ataletan azaldutako ezaugarri kimikoak, kalkuluekin hasi baino lehen jakitea, oso garrantzitsua da. Adibidez bilaketa konformazionalak ez ditu kontutan hartzen tautomerizazioaren ondorioz eman daitezkeen aldaketak, horregatik tautomero bakoitzarentzako kalkulu bat egin behar da. Informazio asko dago bilaketa konformazionalei buruz, berez oso ohikoak diren kalkuluak dira. Tesi honetarako jarraitutako metodologia sistema desberdin askorekin osatu da eta erabilitako programa MacroModel software[9,10] da. Erabilitako indar eremu garrantzitsuenak hurrengoak dira: OPLS 2005,[11] MMFFs,[12] eta AMBER (atomoen karga zuzentzeko GLYCAM).[13,14]

Kasu askotan hautagaien zerrenda luzeegia da, eta hau murrizteko aurre aukeraketa bat egin behar da: Lehenengo energia tarte txiki bateko egitura egonkorrenak hartzen dira, lehenengo hogeitik berrogeiraino tartean eta zuzenean optimizatzen dira Gaussianen.[8] Gehienetan indar eremua ondo parametrizatua badago egitura egonkorrena aurkitzen du. Beste alde batetik akats handia egin daiteke hurbilketa honekin eta hau ez gertatzeko bilaketa osoaren lagin bat tratatu egin behar da. Maestro software (Schrödinger Release 2017-1, LLC, New York) programak "clustering" edo egituren antzekotasunaren arabera taldeetan banatzeko erreminta du, horrela familietan banatzen dira eta bakoitzetik bat edo bi ordezkari optimizatzen dira. Ahalik eta errealitateko (energi potentzialaren gainazala) PES-etik hurbilago egoteko, prozedura osoa zenbait alditan, indar eremua aldatuz, errepikatzea gomendagarria da.

2.3.2.2. Egituraren Optimizazioa eta Frekuentzien Kalkulua

Bilaketa konformazionalean aurkitutako geometriak optimizatu egiten dira kalkulu zehatzago batekin, ab-initio edo DFT, gehienetan zehaztasun nahiko dutelako eta

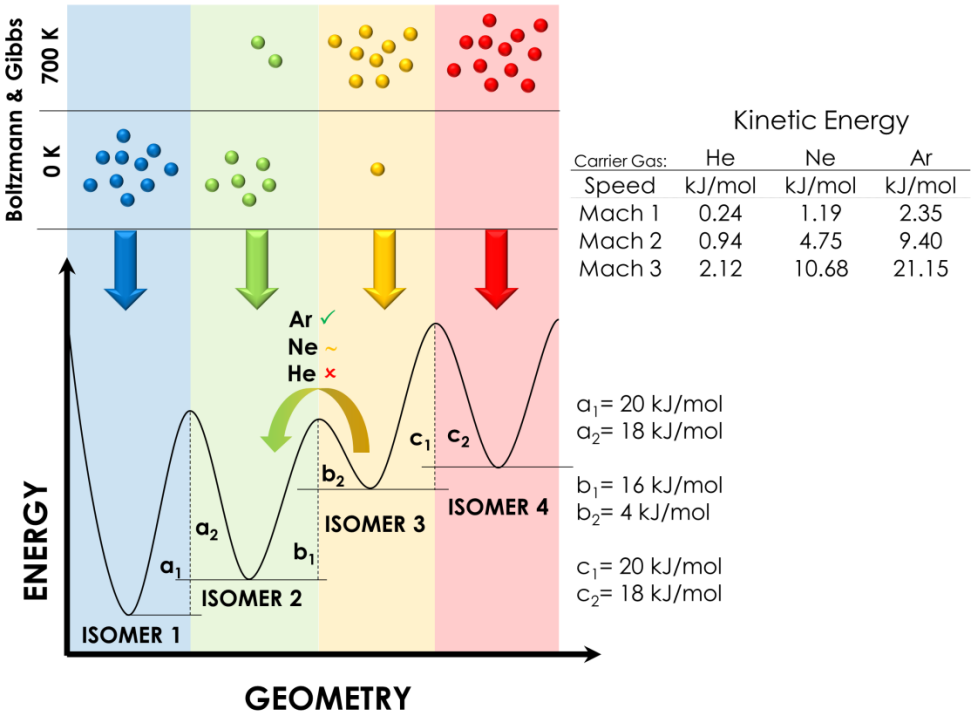
azkarragoak direlako DFT metodoak erabiliko dira. Hauen artean erabiliena M06-2X da eta kasu batzuetarako B3LYP Grimes Dispertsioarekin ("*Empirical Dispersion*"=GD3).[7,15] Atomoen orbitalak definitzeko erabilitako funtzio baseak z bikoitza (6-31+G(d)) eta hirukoitza (6-311++G(d,p)) dira. Optimizatutako egiturei beste kalkulu bat egiten zaie, hurbilketa harmoniko erabiltzen duen frekuentzien estimazioa. Azkenean esleipena simulatutako espektroak espektro esperimentalekin alderatzean egiten da. Gehienetan aurkitutako egitura egonkorrena espektro esperimentala ondo erreproduzitzen duena da baina gerta daiteke horrela ez izatea, beste efektu batzuk baitaude hau aldatu dezaketenak. Izan ere, espantsioaren hozketa prozesuan zehar zenbait fenomeno lehiatzen dira egoera dinamiko bat sortuz; bertan egitura bakoitzaren egonkortasuna eta eraldatzeko behar den energia mugatzen dute aurkituko diren egiturak. Hurrengo atalean espantsioak duen eraginak azaltzen dira.

2.3.3. Giro Esperimental Dinamikoaren Eragina

CI-fasea eratzea oso prozesu konplexua da gertatzen diren gertaerei dagokionez.[4] Pultsatutako balbula erabilita bi era desberdin daude: (i) Lagina gas garraiatzailearekin batera jarri eta hau berotu gas eran egoteko; eta (ii) Lagina ablazio zilindroan jarri balbularen zuloaren bestaldean hutsa dagoen lekuan, eta laserraren bidez balbularen pultsuaren bitartean gasa bihurtuz. Bi egoeratan molekulen tenperatura nahiko altua da edozein hesi energetiko gainditzeko eta gehienetan desiragarriak ez diren molekulen apurketa prozesuren bat izateko. Espantsioa ematen den bitartean gas garraiatzailearekin egiten dituen talkak molekula rotazional, bibrazional eta translazional eran hozten ditu. Prozesu hau oso azkar ematen da eta gehienetan produktu zinetikoa hobesten du, horrela hoztu baino lehenagoko egiturak harrapatuta gelditzen dira (2.3.1.1.3 irudia).[16–19] Hau ondo ulertzeko posiblea den informazio guztia ikertu behar da eta bi propietate nagusitan laburtu daitezke: i) Konformazioen banaketa hozte prozesua baino lehen[20] eta ii) gainditu daitezkeen energi hesiak gas garraiatzailearen talka batengatik.[21] 2.3.3 irudian bi propietateak azalduak dira.

Egoera zailtzen da laser ablazioa gehitzen denean: sortzen den ablazio luman tenperatu ezezaguna duten molekula asko daude.[22,23] Horregatik Gibbs-en energiaren diagramatik tenperatura bat aukeratzea oso zaila bihurtzen da . Arazo honi aurre egiteko Gibbs-en energiaren diagrama tenperatura tarte osoan kalkulatu egiten da 0K-etik 700K-era, azken tenperatura honetan materia organikoa deskonposatu egiten da. Diagrama honekin tenperatura bakoitzean

zeintzuk diren egitura egonkorrenak jakin daiteke eta entropiaren eragina barneratzen du. Tenperatura bibrazionala tenperatura konformazionalean eragin handiena duena da.[20] Onartuta dauden hesi energetikoak 2.3.3 irudian daude eta energi zinetikoen taulan.[4,24]



2.3.3 Irudia: Eragina duten bi faktore garrantzitsu daude hozketa prozesuan, lehenengoa hozketa baino lehen dagoen konformazio banaketa eta bestea zein den gas garraiatzailea da. Irudian aurkezten diren bi adibideak (0K and 700K): lehenengoaren kasuan (0K) populazio gehiena pilota edo konformazio urdina eta berdean pilatuta dago. Bigarren kasuan (700K), gehien bat eratuko diren konformazioak horiak eta gorriak dira. Bigarren adibide honetan gas garraiatzaileak aldaketak eragin ditzake, kasu batzuetan energi hesia gainditu eta konformazio bat beste egonkorrago batera eraldatzen da. Espantsio supersonikoaren abiadura kontutan izanda (Mach ~2) eta gas garraiatzailearen energi zinetikoa abiadura horretan, talka bat egitean zenbat energia jokoan egongo den jakin daiteke. Irudiari begira eta taulari begira kasu honetan gerta daitekeena iragarri daiteke. Energi hesi guztiak debekatuta daude isomero 3 isomero 2 sortzeko norabidean izan ezik (b_2). Hesi hau gainditu ko da Ar gas garraiatzailea bada (~10kJ/mol), Ne gasaren kasuan (~5kJ/mol) gutxi gorabehera hesia eta talkaren energia bat datoz, kasuaren arabera hesia gainditu ko da eta beste batzuetan agian ez. He gas garraiatzaile dagonenean ezingo da gainditu eta soilik isomero 3 eta 4 ikusiko dira.[4]

Gibbs energiaren diagramaren kalkulua nahiko espezifikoa da eta sistema bakoitzak bere hurbilketa behar du.[25] Bi taldetan banatzea posiblea da: 2.3.3.1) Gibbs energia erlatibo askearen kalkulua monomeroentzako eta 2.3.3.2) Konplexuentzako. Bigarren talde honetan berriro ere bi talde txikiagoetan sailkatu daiteke Gibbs-en energi erlatiboa eta Gibbs-en lotura energi erlatiboan.

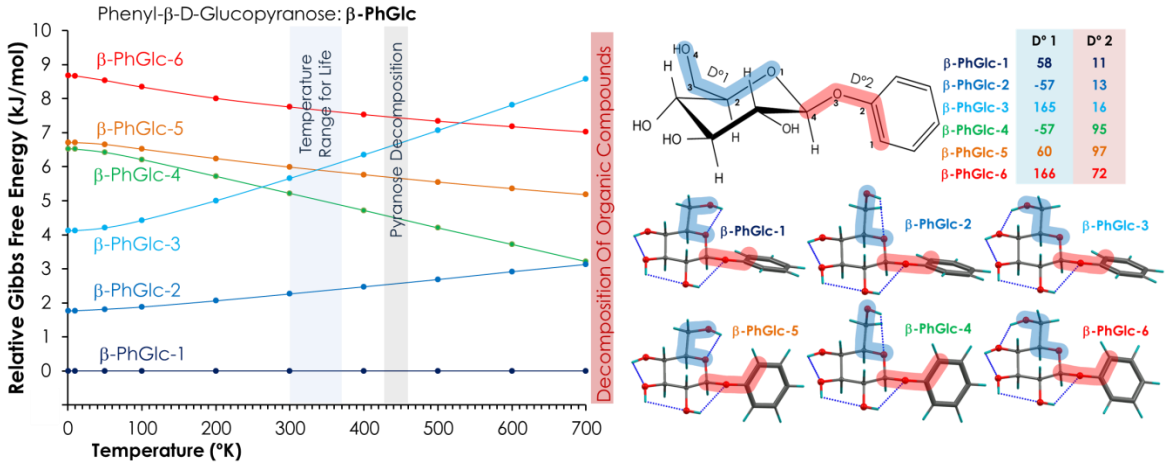
2.3.3.1. Gibbs-en Energi Erlatiboa Monomeroetan

Honen kalkulurako hurrengo ekuazioak erabiltzen dira:

$$1)\quad \Delta G_{ISOMER\,i,T} = \Delta E_{Electronic\ i} + \Delta E_{ZPE\,i} + \Delta E_{ddH_{i,T}} - T \cdot \Delta S_{dS_{i,T}}$$

$$2)\quad \Delta G_{Relat.ISOMER\,i,T} = \Delta G_{ISOMER\,i,T} - \Delta G_{ISOMER\,GM,T}$$

Lehenengo (1) Ekuazioak Gibbs-en energia kalkulatzen du *i* isomeroarentzat tenperatura (T) zehatz batean, eta horretarako energi elektronikoa ($\Delta E_{Electronic\ i}$), zero puntuko energiaren zuzenketa ($\Delta E_{ZPE\,i}$), entalpia tenperatura bakoitzean ($\Delta E_{ddH_{i,T}}$)eta entropia ($T \cdot \Delta S_{dS_{i,T}}$) kontutan ditu. Isomeroen arteko energi erlatiboa lortzeko bigarren (2) ekuazioa erabiltzen da, bertan egitura egonkorrenaren energia ($\Delta G_{ISOMER\,GM,T}$) isomero bakoitzari kentzen zaio ($\Delta G_{ISOMER\,i,T}$). Diagramaren adibide bat 2.3.3.1 irudian ikusgai dago.



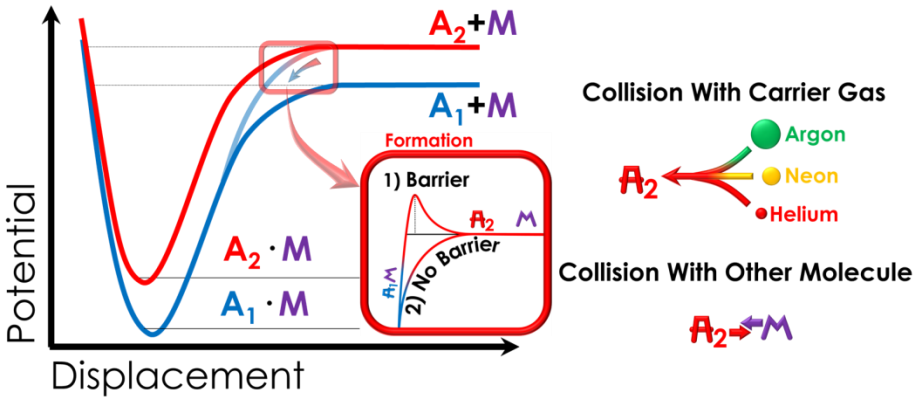
2.3.3.1 Irudia: Fenil-β-D-Glukopiranosa (β-PhGlc) molekularentzako Gibbs-en energi erlatibo askearen diagrama.

2.3.3.2. Gibbs-en Energi Erlatibo Askea Konplexuetan

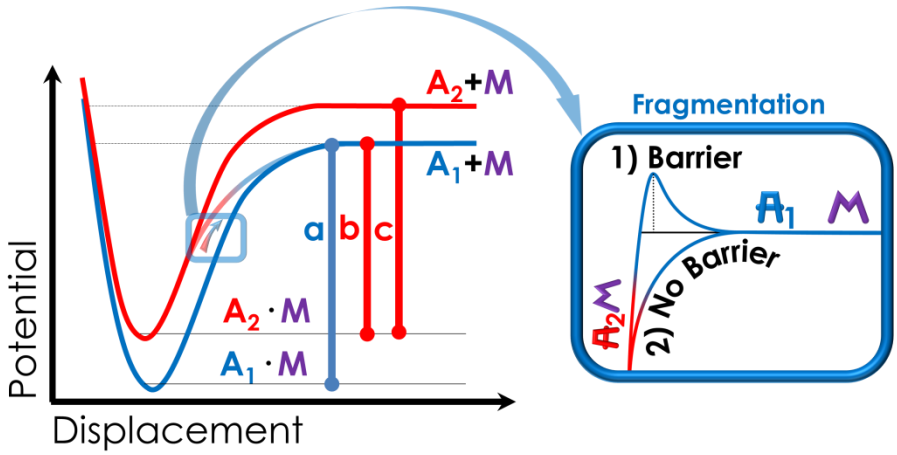
Konplexuetan gertatzen dena ulertzeko eraketa eta apurketa prozesuen ikerketara zuzendu beharra dago. 2.3.3.1.1 irudian eraketan eman daitezkeen bideak erakusten dira, eraketarako bi bide nagusi daude, oikoena: A_i+M→A_iM izanik eta gerta litekeen beste bide bat: A_i+M→A_JM da. Lehenengo mekanismoak A_i isomero M molekularekin batzen ditu eraldaketarik gabe eta bigarrenean lotu baino lehen edo bitartean eraldaketa bat gertatzen da A_i isomeroan A_J

isomeroan bihurtuz. Hau gasaren talkek mugatuko dute, efektu hauek 2.3.3.2.1 irudian laburbiltzen dira xehetasun guztiekin.

Lotura indarra kalkulatzeko apurketa diagrama aztertu behar da. Lehen eraketan gertatzen den eran bi eraketa bide nagusi daude: lehenengoa konplexuan dauden monomeroak, aske dagoenean hartu dezaketen antzeko egitura egonkorrenarekin lotzen dituena da (MS). Bigarrena konplexuko monomeroen egiturak zuzenean monomeroen artetik egitura egonkorrenarekin lotzen ditu nahiz eta desberdinak izan (GM). 2.3.3.2.2 irudian bi lotura mota edo apurketa era laburbiltzen dira.[26]



2.3.3.2.1 Irudia: Energia potentzialaren grafikoa bi konplexu isomeriko desberdinentzako. Konplexuen eraketa zuzenean eman daiteke (A_1M eta A_2M) erreakzio hauek jarraituz: $A_1+M \rightarrow A_1M$, eta $A_2+M \rightarrow A_2M$.Konplexuaren eraketa mekanismoa isomerizazio baten bidez ere eman daiteke: $A_2+M \rightarrow A_1M$. Konplexua eratzeko bigarren bide hau energi hesia izan dezake edo ez (normalean energi hesia dago),hesi hau gainditzeko bi talka mota eman daitezke: Gas garraiatzailearekin egindako talkak (menderatzen duen mekanismoa da) eta molekula bat bestearen kontra talka egiten dutenean (A_x M-rekin). Teorian bigarren motako talkekin oso energi hesi altuak gainditu daitezke, energi zinetikoa (abiadura berean) masarekin handitzen baita (espantsioa beti Mach 1 eta Mach 2 tartean dago). Bigarren mekanismo hau ez dago guztiz argi eta egotekotan oso proportzio txikian egongo da. Ondorio moduan energi hesia gainditzeko orokorrean arautzen duen fenomenoa gas garraiatzailearekin talkak kontutan duena da.



2.3.3.2.2 Irudia: Energia potentzialaren grafikoa apurketa prozesurako. Bi bide nagusi daude: Minimo orokorraren bidez "***Global Minimum***" **(GM)** erreakzioa jarraitzen duela: $A_1M \rightarrow A_1 + M$ (a) energi diferentziarekin, eta $A_2M \rightarrow A_1 + M$ (b) energi diferentziarekin. Bigarren kasu honetan hesia agertu daiteke (ikusi 1) "*Barrier*" eta 2) "*No Barrier*" kasuak, kutxa urdinaren barruan); orokorrean hesia dagoela suposatzen da eta honen altuera jakiteko beste kalkulu baten beharra dago. Beste posibilitatea monomeroaren antzekoen duen egituraren bitartez ematen den apurketa da "***Most Similar***" **(MS):** $A_1M \rightarrow A_1 + M$ (a) lehengo kasuaren berdina da, baina beste konplexuarena desberdina da $A_2M \rightarrow A_2 + M$ (c) energi diferentzia izango du, eta hau handiagoa da edo indartsuagoa da lotura indarra.

Bi diagrama hauetan oinarrituta (2.3.3.2.1 eta 2.3.3.2.2 irudiak) diseinatu daitezke zenbait energi kalkulu. Konplexuei buruz ari garenez lehenengo eta behin egin behar dena "*Counter Poise*" zuzenketa da($\Delta E_{BSSE_{A_iM}}$).[27] 3 ekuazioan Gibbs energi askea erakusten da edozein tenperaturarako T eta AM konplexurako. A molekula i konformazioan dagoelarik eta M molekula egitura edo isomero bakarra duen molekula izanda($\Delta G_{A_iM,T}$).

$$3) \quad \Delta G_{A_iM,T} = \Delta E_{Electronic\ A_iM} + \Delta E_{ZPE_{A_iM}} + \Delta E_{BSSE_{A_iM}} + \Delta E_{ddH_{A_iM,T}} - T \cdot \Delta S_{dS_{A_iM,T}}$$

3 ekuazioan oinarrituta Gibbs-en lotura energi askea lortu daiteke. Lotura energia kalkulatzeko hurbilketa bat baino gehiago dago, era bakoitzak ezaugarri eta propietate batzuk nabarmentzen dituelarik. MS metodologia konplexuarekin alderatuta monomeroaren antzekoen eta egonkorra den egitura erabiltzen du kalkulurako ($\Delta G_{A_i,T}$) eta horrela lotura energia lortzen da ($\Delta G_{(MS)A_iM,T}$), ikusi 4 ekuazioa. Ekuazio honen errepresentazio grafikoa egiten bada zerotik pasatuko den funtzio bat ikusten da, kasu horretan (ΔG=0) konplexuaren eraketa eta apurketa berdin bananduta daude. Tenperatura honetatik gora($T_{\Delta G_{(MS)A_iM}=0}$) erreakzioa konplexuen apurketara doa. 5. ekuazioa minimo orokorra den isomeroarena da eta ekuazioa hau 8. ekuazioaren parekoa da, GM metodoarekin kalkuluarena, isomero egonkor orokorra bi kasuetan berdina baita ($\Delta G_{(MS)A_{GM}M,T}$=$\Delta G_{(GM)A_{GM}M,T}$). Gibbs-en loturak energia erlatiboa lotura indarra eta

energia potentzialari buruzko informazioa ematen du, hau 6. ekuazioan kalkulatzen da, eta 2.3.3.2.1.2 irudian diagrama ikusgai dago.

$$4)\quad \Delta G_{(MS)A_iM,T} = \overset{3\ Ekuazioa}{\Delta G_{A_iM,T}} - (\overset{1\ Ekuazioa}{\Delta \boldsymbol{G_{A_i,T}}} + \Delta G_{M,T})$$
$$\boldsymbol{T_{\Delta G_{(MS)A_iM=0}}} \equiv Orekako\ Tenperatura\ (MS)$$

$$5)\quad \Delta G_{(MS)A_{GM}M,T} = \Delta G_{A_{GM}M,T} - (\Delta G_{A_{GM},T} + \Delta G_{M,T})$$

$$\boldsymbol{6)}\quad \boldsymbol{\Delta G_{Relat.(MS)A_iM,T} = \Delta G_{(MS)A_iM,T} - \Delta G_{(MS)A_{GM}M,T}}$$
$$\boldsymbol{Gibbs\ Lotura\ Energi\ Erlatiboa\ (MS)}$$

Kalkulua egiteko bigarren era GM da, honek Monomeroaren egitura egonkorrenaren energia kentzen dio dimeroari lotura indarra kalkulatzeko, 7. ekuazioa. Metodologia honen bidez lortutako orekako tenperatura (ΔG=0) ez da guztiz zuzena, kasu askotan apurketa isomerizazio batekin lotuta dagoen prozesua delako, monomeroaren egitura egonkorrena konplexuaren parte ez denean alegia. GM lotura energiak ez du era honetan informazio interesgarri gehiegirik erakusten baina erlatiboa egiten denean 11. ekuazioa konplexuen energi erlatiboa lortzen da eta energiaren ikuspuntu orokor batean (eta ez MS metodoan bezala non soilik lotura indarra kontutan artzen dela) zein den isomero egonkorrena esaten du, diagrama 2.3.3.2.1.2 irudian ikusgai dago.

$$7)\quad \Delta G_{(GM)A_iM,T} = \overset{3\ Ekuazioa}{\Delta G_{A_iM,T}} - (\overset{1\ Ekuazioa}{\Delta \boldsymbol{G_{A_{GM},T}}} + \Delta G_{M,T})$$
$$\boldsymbol{T_{\Delta G_{(GM)A_iM=0}}} \equiv Orekako\ Tenperatura\ (GM)$$

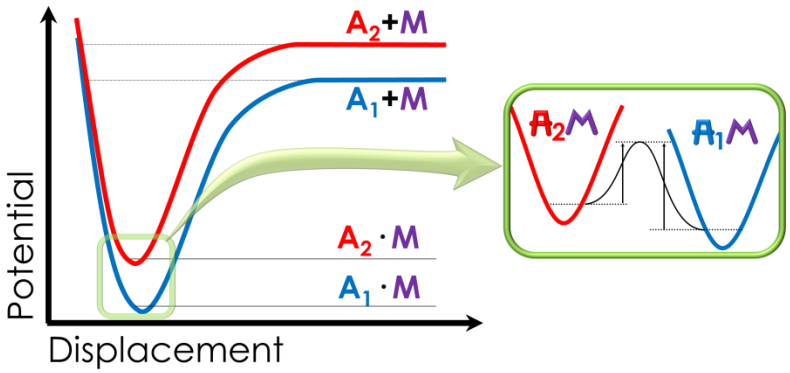
$$8)\quad \Delta G_{(GM)A_{GM}M,T} = \Delta G_{A_{GM}M,T} - (\Delta G_{A_{GM},T} + \Delta G_{M,T})$$

$$9)\quad \Delta G_{Relat.(GM)A_iM,T} = \Delta G_{(GM)A_iM,T} - \Delta G_{(GM)A_{GM}M,T}$$

$$10)\quad \Delta G_{Relat.(GM)A_iM,T} = \overset{7\ Ekuazioa}{\left(\Delta G_{A_iM,T} - (\Delta G_{A_{GM},T} + \Delta G_{M,T})\right)} - \overset{8\ Ekuazioa}{\left(\Delta G_{A_{GM}M,T} - (\Delta G_{A_{GM},T} + \Delta G_{M,T})\right)}$$
$$7\ eta\ 8\ Ekuazioak\ 9\ ekuazioan\ ordezkapena$$

$$\boldsymbol{\Delta G_{Relat.(GM)A_iM,T} = \Delta G_{A_iM,T} - \Delta G_{A_{GM}M,T}}$$
$$\boldsymbol{11)\quad 10\ ekuazioaren\ sinplifikazioa}$$
$$\boldsymbol{Gibbs\ Energi\ Erlatiboa\ (GM)}$$

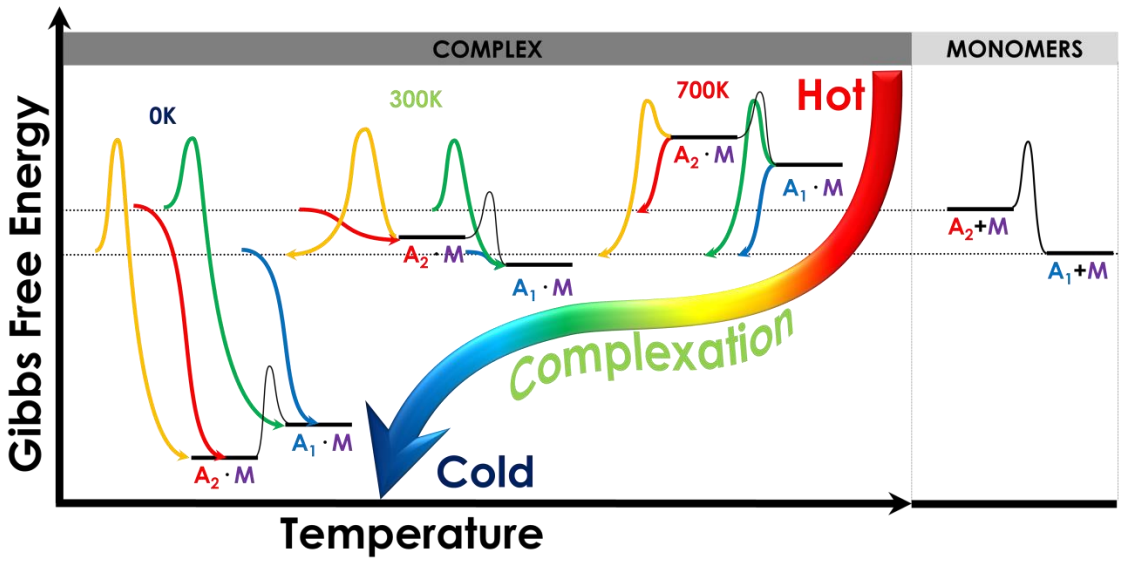
Isomerizazio energi hesia monomeroetan aldatu edo txikiagotu daiteke konplexuan bigarren molekularen bitartekariaren egitura egonkortzen duelako. Efektu hau faktore askok mugatzen dute baina garrantzitsuenen artean molekulak beste molekularekiko duen kokapena da. Molekula bat beste molekulak eraldatu behar duen atalean kokatzen bada bitartekaria den egitura egonkortzeko posibilitateak daude eta honen ondorioz gas garraiatzailearen talka baten ondorioz hasiera baten monomeroan gaindiezina zen energi hesia konplexuan gainditzeko aukera dago. 2.3.3.2.3 irudian fenomeno hau erakusten da, gainera 4. kapituluan β-β dimerorako honelako mekanismo bat proposatzen da (Figure A4.4 ikusi).



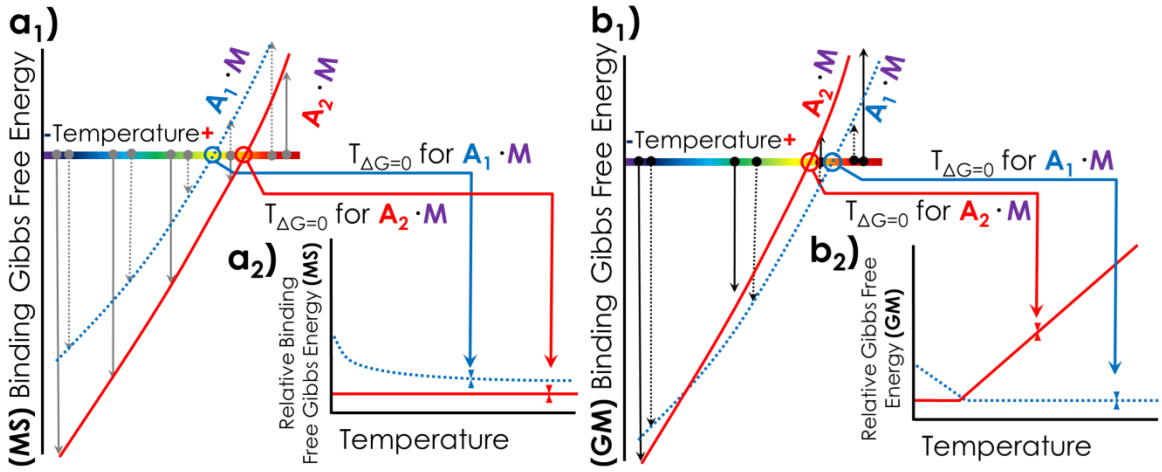
2.3.3.2.3 Irudia: Eraldaketa behin konplexua eratuta dagoela.

2.3.3.2.1. Tenperaturaren Efektua

Kontutan izango den azken propietatea tenperatura da. 2.3.3.2.1.1 irudian, monomero askeen batuketa, konplexuarekin alderatuta, eta eratu daitekeen diagrama ikus daiteke, baita tenperaturak duen eragina. Bertan ikus daiteke nola tenperatura altuetan konplexuak ezegonkorragoak direla, oreka tenperaturatik hurbil (ΔG=0), dimeroa eratzen hasten da eta tenperatura baxuagoetan eraketa hobesten da. Espantsio supersonikoak hozketa adiabatikoa eragiten du eta orekan kalkulatutako Gibbs-en energia ez dute hau erreproduzitzen.[25] Gibbs-en diagramak tenperatura bakoitzean dagoen konformeroen banaketa erakusten du eta hauen artean zein banaketa izoztu den jakiteko gaitasuna ematen du.



2.3.3.2.1.1 Irudia: Tenperaturaren eraginaren diagrama baina monomeroak konstante mantentzen dira hurbilketa moduan tenperatura tarte osoan zehar. Diagrama hau espantsioko hozte prozesuaren hurbilketa bat da. A2+M eta A1+M monomero askeen batuketa da. A2M eta A1M konplexuen energia da. Diagrama monomeroak eta konplexuak zenbait tenperaturatan konparatzen ditu.



2.3.3.2.1.2 Irudia: Bi diagrama mota GM (orokorra) eta MS (lotura). **a_1)** MS Gibbs-en lotura energiaren diagrama 4 ekuazioa erabilita. **a_2)** MS Gibbs-en lotura energi erlatiboaren diagrama, 6 ekuaziotik eraikia; **b_1)** GM Gibbs-en lotura indarra 7 ekuazioarekin eraikita; **b_2)** GM Gibbs-en energia erlatiboaren diagrama 11 ekuazioaren kalkulutik egina. Normalean artzen den tenperatura tartea 0K-700K da, materia organikoa 700K ingurua deskonposatu egiten baita. Diagrama isomero guztiei buruz edo esanguratsueni buruz egin daiteke.

4. eta 7. ekuazioak erabilita posiblea da lotura energia eta lotura energi erlatiboaren diagrama eraikitzea (2.3.3.2.1.2 Irudia: a_1 eta b_1). Normalean MS beste erabilera bat du, konplexu edo molekula desberdinen arteko lotura indarrak ere alderatu ditzake. Orokorrean konplexuek duten egonkortasun erlatiboa jakiteko erabili behar diren ekuazioak 6. eta 11. dira. Lortutako diagramak Lotura

MS eta egonkortasun erlatiboa GM tenperaturarekin nolakoak diren erakutsiko du, 2.3.3.2.1.2 Irudia: a_2 eta b_2.

2.4 Erreferentziak

(1) Demtröder, W. *Laser Spectroscopy. Basic Concepts and Instrumentation*; Springer, 2003.
(2) Levy, D. H. Laser Spectroscopy of Cold Gas-Phase Molecules. *Annu. Rev. Phys. Chem* **1980**, *31*, 197–225.
(3) De Vries, M. S.; Hobza, P. Gas-Phase Spectroscopy of Biomolecular Building Blocks. *Annu. Rev. Phys. Chem* **2007**, *58*, 585–612.
(4) Scoles, G. Atomic and Molecular Beam Methods. In *Atomic and Molecular Beam Methods*; Oxford University Press, 1988; Vol. 1, pp 14–53.
(5) Camiruaga, A.; Usabiaga, I.; Insausti, A.; León, I.; Fernández, J. A. Sugar–peptidic Bond Interactions: Spectroscopic Characterization of a Model System. *Phys. Chem. Chem. Phys.* **2017**.
(6) Leon, I.; Millan, J.; Cocinero, E. J.; Lesarri, A.; Castano, F.; Fernandez, J. A. Mimicking Anaesthetic-Receptor Interaction: A Combined Spectroscopic and Computational Study of Propofol···phenol. *Phys. Chem. Chem. Phys.* **2012**, *14* (25), 8956–8963.
(7) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120* (1–3), 215–241.
(8) Gaussian 09 Revision D0.1 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, "; M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. M.; G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. H.; A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. H.; M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. N.; Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J.; J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. B.; K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. N.; K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. T.; M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. C.; et al. Gaussian 09 Revision D0.1. *Gaussian, Inc., Wallingford CT* **2013**.
(9) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. Macromodel—an Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics. *J. Comput. Chem.* **1990**, *11* (4), 440–467.
(10) Leon, I.; Montero, R.; Longarte, A.; Fernandez, J. A. Influence of Dispersive Forces on the Final Shape of a Reverse Micelle. *Phys. Chem. Chem. Phys.* **2015**, *17* (3), 2241–2245.
(11) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* **2005**, *26* (16), 1752–1780.
(12) Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17* (5–6), 490–519.
(13) CASE, D. A.; CHEATHAM, T. E.; DARDEN, T. O. M.; GOHLKE, H.; LUO, R. A. Y.; MERZ, K. M.; ONUFRIEV, A.; SIMMERLING, C.; WANG, B.; WOODS, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688.
(14) KIRSCHNER, K. N.; YONGYE, A. B.; TSCHAMPEL, S. M.; GONZÁLEZ-OUTEIRIÑO, J.; DANIELS, C. R.; FOLEY, B. L.; WOODS, R. J. GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates. *J. Comput. Chem.* **2008**, *29* (4), 622–655.
(15) Grimme, S. Density Functional Theory with London Dispersion Corrections. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1* (2), 211–228.
(16) Campargue, R. Progress in Overexpanded Supersonic Jets and Skimmed Molecular Beams in Free-Jet Zones of Silence. *J. Phys. Chem.* **1984**, *88* (20), 4466–4474.
(17) Patterson, D.; Doyle, J. M. Cooling Molecules in a Cell for FTMW Spectroscopy. *Mol. Phys.* **2012**, *110* (15–16), 1757–1766.
(18) Deckers, J.; Fenn, J. B. High Intensity Molecular Beam Apparatus. *Rev. Sci. Instrum.* **1963**, *34* (1), 96–100.

(19) Smalley, R. E.; Wharton, L.; Levy, D. H. Molecular Optical Spectroscopy with Supersonic Beams and Jets. *Acc. Chem. Res.* **1977**, *10* (4), 139–145.
(20) Kudoh, S.; Takayanagi, M.; Nakata, M. Conformational Cooling in a Supersonic Jet of 1,2-Dichloroethanestudied by Matrix Isolation Infrared Spectroscopy. *Chem. Phys. Lett.* **1998**, *296* (November), 329–334.
(21) Ruoff, R. S.; Klots, T. D.; Emilsson, T.; Gutowsky, H. S. Relaxation of Conformers and Isomers in Seeded Supersonic Jets of Inert Gases. *J. Chem. Phys.* **1990**, *93* (5), 3142–3150.
(22) Srinivasan, R.; Braren, B. Ultraviolet Laser Ablation of Organic Polymers. *Chem. Rev.* **1989**, *89* (6), 1303–1316.
(23) Lesarri, A.; Mata, S.; López, J. C.; Alonso, J. L. A Laser-Ablation Molecular-Beam Fourier-Transform Microwave Spectrometer: The Rotational Spectrum of Organic Solids. *Rev. Sci. Instrum.* **2003**, *74* (11), 4799–4804.
(24) Desyatnyk, O.; Pszczolkowski, L.; Thorwirth, S.; Krygowski, T. M.; Kisiel, Z. The Rotational Spectra, Electric Dipole Moments and Molecular Structures of Anisole and Benzaldehyde. *Phys. Chem. Chem. Phys.* **2005**, *7* (8), 1708–1715.
(25) Godfrey, P. D.; Brown, R. D. Proportions of Species Observed in Jet Spectroscopy−Vibrational Energy Effects: Histamine Tautomers and Conformers. *J. Am. Chem. Soc.* **1998**, *120* (41), 10724–10732.
(26) Çarçabal, P.; Cocinero, E. J.; Simons, J. P. Binding Energies of Micro-Hydrated Carbohydrates: Measurements and Interpretation. *Chem. Sci.* **2013**, *4* (4), 1830–1836.
(27) Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19* (4), 553–566.

3. Glukosa eta Fenol Molekulen Arteko Elkarrekintza: "*Bizirik Dirauten Iaioa*"

3.1. Sarrera eta Helburuak

Biologia mailan gertatzen diren oreka kimikoek sare konplexua eratzen dute,[1] honen ondorioz oso saila da erreakzio sarearen eskema egitea. Azken aldian argitaratutako zenbait ikerketek matematikako algoritmo konplexuak eta aurreratuen "*imaging*" espektroskopia nahasten dute oreka biologikoen eskema eraikitzen saiatzeko.[2] Aipatutako teknika hauek sortzen duten probabilitate sarea aldez aurretik ondo ezagutzen diren oreka biokimikoekin alderatu egiten dira analisiaren baieztapena lortzeko.[3] Gaur egun sistema hauen kalkulu konputazional osoak ezinezkoa dira, gainera badaude atal dinamiko asko kontutan izan behar direnak eta gehiengoak ezagutu gabe jarraitzen dute.[4] Nahaspila honetan etengabe ari dira konplexuak sortzen eta banantzen alegia.

Natura izaera ez-kobalentea duten molekulen arteko elkarrekintzetan oinarritzen da, bizitza izenarekin ezagutzen dugun oreka kimiko multzo hau zuzentzeko.[5] Molekulen arteko lotura ahul hauek izaki bizidunen metabolismoa funtzionarazten dute energia kostu txikienarekin. Adibidez, badira prozesu asko non prozesua hasteko bi molekulen arteko elkarrekintza eman behar den, seinalea daraman molekula eta errezeptorea elkartzen dira konplexu bat eratuz eta errezeptorean egitura aldaketa bat eragiten du. Honen ondorioz seinalearekin lotutako prozesuak aktibatzen dira. Bi molekulak lotura ahulen (Van der Waals eta hidrogeno loturak) bidez elkarrekintzan dabilenez, momenturen batean bananduko dira erreakzioei bukaera emanez, eta denboratxo bat pasa ondoren bere onera bueltatu eta beharrezkoa denerako berriro prozesuekin hasteko prest jarriko dira. Seinalea eta errezeptorea lotura kobalenteekin lotuta egongo balira, sistema bere onera bueltatzea saila edo ezineko gerta izango litzateke. Badaude toxikoak diren konposatu asko, hauek erakusten duten toxikotasuna itzulezinak diren erreakzioak sortuak izan daitezke. Erakarpen kimiko handia duten konplexuek askotan oreka kimiko fisikoetan eragin dezakete eta izaki bizidunengan ondorio latzak eragin.[6]

Aipatu den bezala, orokorrean, erreakzio biologikoak erreakzio batetik beste erreakzio baten arteko trantsiziozko erreakzioz osaturik daude. Adibide bat propofola (anestesikoa) egiten duten erreakzioak dira: Burmuinean dauden $GABA_A$ kanaletan finkatzen denean hauek irekita uzten ditu, eta honek minaren seinalea gelditzen du, Cl^- "post-synapses"-etara sartzen da Na^+ kanalen (minaren seinalea Na kanalak irekitzean ematen da) irekitzea konpentsatuz. Odoleko kontzentrazioa jetzi eta propofola gune aktibotik joan bezain laster efektua bukatzen da.[7] Paziente bati dosi zuzena emanez lotarazten du, baina gaindosia

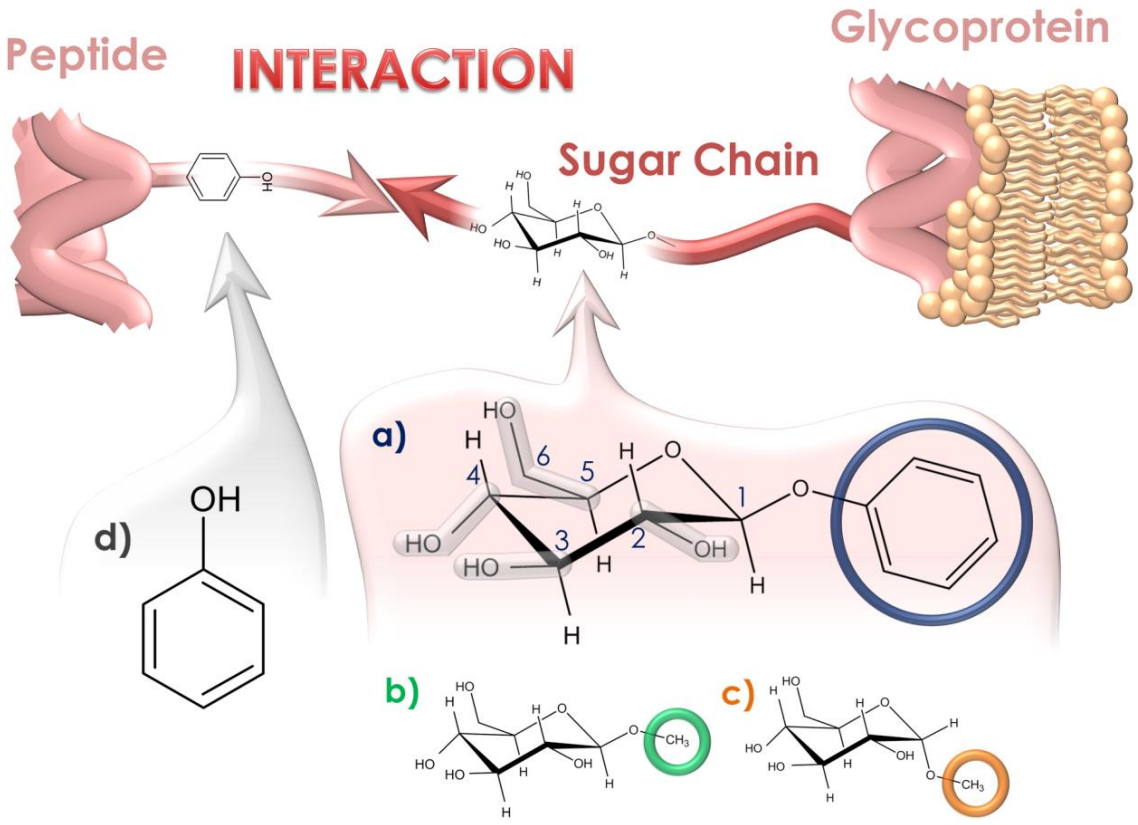
ematen bazaio heriotza eragiten du. Honen arrazoiak propofolaren desiragarriak ez diren efektuak dira: lehenengo propofol molekulak $GABA_A$ kanaletara bideratzen dira baina kontzentrazioa handiegia denean beste errezeptore batzuetan eragina izaten dute, hauen artean errezeptore adrenergikoak daudelarik. Farmako berri bat aurkitzeko, adimenez diseinatutako esperimentua proposatzen denean, atal zailena ez da gune aktiboan lotzen den molekula bat aurkitzea, benetako zailtasuna gorputzean dauden beste ehun mila proteinekin ez lotzean datza. Beste interakzio guztien artetik egonkortasun apur bat handiagoa izateak izan daiteke molekula bati bere funtzio biologikoa ematen diona.[8] Erreakzio Kimiko hauek ondo ulertzea guztiz beharrezkoa da, xedea kontrolatzea bada, eta honela zenbait gaixotasun sendatuko dira, adibidez ormona baten falta zuzenduz edo birusak haien zelula dianan sartzea ekidituz.

Zenbait teknika esperimental existitzen dira honelako informazio espezifikoa ahalbidetu dezaketenak. Haien artean, kristalografia da ziur aski gehien erabili dena; honek molekulen egitura determinatzea ahalbidetzen du baita molekulen arteko elkarrekintzak ere harrapatu daitezke kristalaren barnean.[9,10] Teknika honek baditu bere mugak, adibidez kristal eran egon behar dela; horrek beste efektu eta indarrak gehitzen ditu, hauek molekularen egituran eta erakutsitako interakzioetan islatu daitezke. Azken finean teknika honen bitartez informazio asko lortu daiteke baina biologikoa den ingurumenaz oso urrun dago eta batzuetan lortutako informazioa ez da guztiz zehatza, baita lortutako distantziei buruzko informazioa askotan ez da guztiz zuzena. Zuzenagoa da sistema hauek disoluzioan ikertzea, baina disoluzioak giro dinamikoa du, molekulak etengabeko mugimenduan baitabiltza, eta ezaugarri honek zehaztasuna murrizten du, soilik molekulen batezbesteko egituraren informaziora lortu daiteke eta.

Hasiera batetan badirudi sistema biologikoen ikerketa hotz eta isolatutako fasean ez dela posible, berez ingurugiroak oso desberdinak dira. Ahala ere hauei buruzko zenbait informazio[11] lortu daiteke, espantsio supersonikoa erabilita hoztu eta isolatu daitezke molekulak eta konplexuak, hauetan lortutako egitura ez da beste faktoreen menpe egongo berez dituztenen menpe baizik. Teknika hau ingurugiroaren influentziarik gabe molekulen benetako egitura erakusteko gai da. Neurketa hauek kalkulu teorikoen beharra dute egituren esleipena egiteko, eta mekanika kuantikoan oinarritzen dira. Esperimentu hauen bidez molekulen jokabidea determinatu daiteke. Gainera, beste ingurugiroetan antzekoak diren oreka kimikoak espantsio supersonikoan ere ematen dira. Hasiera batean molekulak asko berotzen dira laserraren absortzioa dela eta, jasotako energiarekin egitura batetik bestera eraldatzen hasten dira, orokorrean tenperatura altuetan

egonkorragoak diren egiturak hartuz. Bat-batean (orekatik kanpo) gas eramailea sartzen da egoera dagoen bezala izoztuz,[12] normalean harrapatutako egiturak egonkorrenak dira baina gerta daiteke energia erlatiboan altua den egituraren bat detektatzea, hau egituren arteko hezien menpe dago. Aurkitutako egiturak oso baliotsuak dira, hozte prozesuaren informazioa eta baita konplexu horretan arautzen duten propietateei buruzko informazioa ematen dutelako. Egitura bakoitza disoluzioan hartu dezakeen egituren definizio altuko argazkien bezalakoak dira. Hasiera batean ezegonkorragoak diren egitura horietatik, batzuk lotura indartsuagoa erakusten dutela ikustea Interesgarria da, baina egonkortasun honek monomeroen egituraren ezegonkortzea dakar gehienetan. Fenomeno hau molekulen arteko erakarpen altuagoaren ondorioa izan daiteke, eta honek espantsioan ematen diren talken (banandu gabe) bizirauteko probabilitateak igotzen ditu eta beste ingurugiro batean abantaila bat bezala ikus daiteke, ingurune biologikoetan gertatzen den moduan, konplexuaren erdibizitza apur bat luzatuz. Molekulen arteko lotura energian ematen den aldaketak, prozesu zinetikoan eta oreka biologikoetan eragin garrantzitsua du, horregatik honelako konplexuak benetan posibleak diren jakiteko batzuk isolatzearen saiakera egingo da.

Aurrekoa kontutan izanda kapitulu honetan fenol eta glukosaren arteko erakarpen intermolekularretan ardaztuko da ikerketa.[13,14] Nahiz eta sistema molekularra sinplea iruditu, molekulen arteko elkarrekintzari dagokionez nahiko konplexutasun handia erakusten du, horregatik kalkulu teorikoak moldatu behar izan dira. Glukosa naturan aurkitzen diren karbohidrato ugarienen artean dago eta konformazio ugari ditu, eta hauek hidroxiloen eta hidroxilmetiloen kokagunean edo norabidean desberdintzen dira.[15] 3.1 eskemak molekulen egiturak erakusten ditu, Metil-D-Glukopiranosaren bi anomeroak (α/β-MeGlc) eta β anomeroaren Fenil-D-Glukopiranosa (β-PhGlc) anomero bakarrarena baita fenolaren (P) egitura ere. Metilatutako bi anomeroek karbono anomerikoa duen eragina bereizteko balio dute eta fenilglukosa metiloa bolumen handiko talde batengatik (feniloa) aldatzen denean, eragiten duen eragina jakiteko balio izango du.



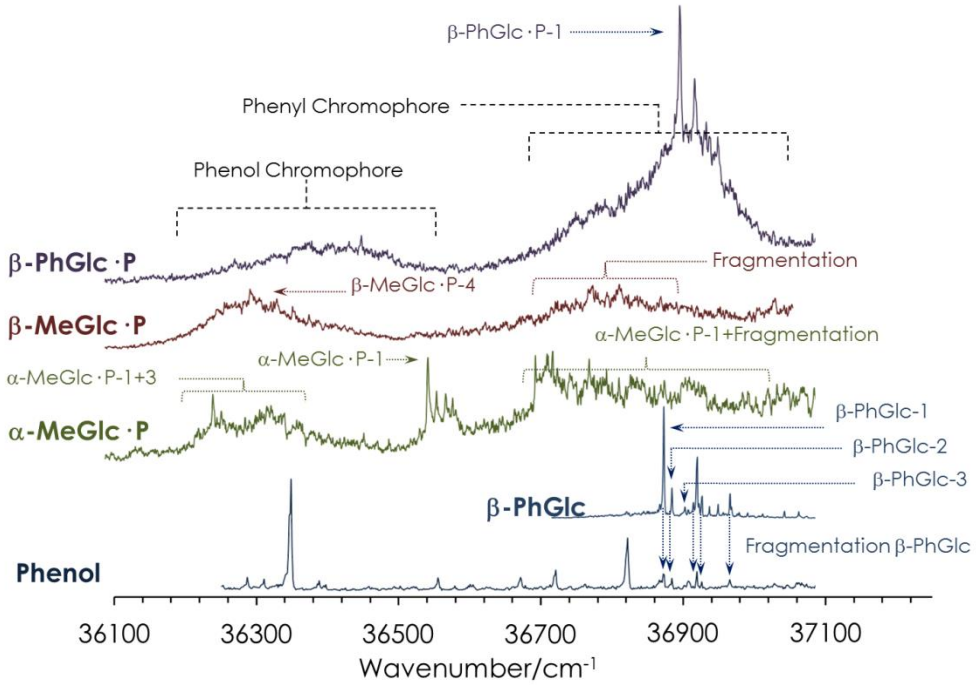
Eskema 3.1: Azukre eta fenolaren arteko elkarrekintza: **a)** fenil-β-D-Glukopiranosa (β-PhGlc), zenbaki urdinak karbono atomoenak dira. Grisez nabarmenduta hidroximetiloak eta hidroxiloak dira. **b)** Metil-β-D-Glukopiranosa (β-MeGlc). **c)** Metil-α-D-Glukopiranosa (α-MeGlc). **d)** Fenola (P).

3.2. Emaitza Esperimentalak

3.2.1 irudian REMPI espektroak kromoforodun bi monomeroak erakusten ditu, P eta β-MeGlc, eta hauekin eraikitako heterodimeroak. Esperimentuan erabilitako baldintzak bigarren kapituluan azaldutakoak dira. Bi kromoforo daudenez REMPI espektroa bi zonaldetan banatu daiteke, kromoforo bakoitzarentzako (bata feniloa eta bestea β-PhGlc-ren fenil taldea) zona bat. Pentsa daiteke seinale gehiengoa P-tik etorriko dela baina β-PhGlc⋯P dimeroan ikus daitekeenez β-PhGlc-ak seinale gehiago sortzen du 36700-37100 cm^{-1} tartean. Azken tarte hau β-PhGlc-ren 0_0^0 transizioari dagokio.

Bi kromoforoen arteko intentsitate erlatiboa zehazki ikertzen bada (P) ~36300 cm^{-1} aldean dagoena (β-PhGlc) ~36900 cm^{-1} aldekoarekin argi ikusten da konplexu berdinaren barruan seinaleen arteko desberdintasuna. ~36800 cm^{-1} tartean Ikusitako absortzioak, metil-β-D-glukopiranosa · fenol (β-MeGlc·P) eta metil-α-D-glukopiranosa · fenol (α-MeGlc·P) konplexuentzat, masa altuagoko sistemen

apurketaren ondorioz sortuak dira (Figure A3.1-b) eta Figure A3.2-b behatu), honen ondorioz tarte hauek konplexu hauetan saihestuak izan dira. Xehetasun gehiagorako dagoeneko argitaratutako artikulura jo daiteke.[16]



Irudia 3.2.1: REMPI espektroak: Fenola (P), Fenil-β-D-Glukopiranosa (β-PhGlc), Metil-α-D-Glukopiranosa · Fenol (α-MeGlc·P), Metil-β-D-Glukopiranosa · Fenol (β-MeGlc·P), eta Fenil-β-D-Glukopiranosa · Fenol (β-PhGlc·P). P eta β-PhGlc monomeroen espektroak aurretik beste autoreak argitaratutako espektroekin bat doaz.[17,18]

3.2.1 irudiko REMPI espektroa ikusita zenbait uhin luzaera aukeratzen dira eta haietan kokatzen da UV laserra, egoera honetan IR laserrarekin IDIRS espektroa jasotzen da, eta hauek 3.2.2 irudian ikusgai daude. Teoriarekin simulatutako espektroak hurbilketa harmonikoaren bidez eginda daude; honek daukan akatsa zuzentzeko faktoreak erabiltzen dira, faktoreak kalkulatzeko nahastezina direnzenbait molekula eta monomeroak erabiltzen dira. Talde bakoitzak anharmonizitate desberdin bat erakusten du eta honen ondorioz bakoitzarentzako faktore bat optimizatzen da doiketarako espektroak erabilita. Lortutako faktoreak hurrengoak dira: CH-entzako 0.9530 eta OH-entzako 0.9385. Balio hauek ondo funtzionatzen dute aske dauden kasuetarako, baina gero eta lotura indartsuagoa orduan eta anharmonizitatearen eragin handiagoa, eta are handiagoa onartutako akatsa.

Monomeroen ikerketa oso erabilgarria da, eta hau emandako emaitzatan oinarritu daiteke gero dimeroen analisia egiteko. Berez dimeroaren REMPI espektroa monomero bakoitzaren REMPI espektroa gehi interakzioek eragiten duten perturbazioa dela esan liteke. IR espektroan gauza bera gertatzen da, bi monomeroen trantsizioak ditugu eta molekulen arteko loturari zenbait trantsizio zehatzetan eragiten duen perturbazioa gehitu behar zaio.

OH askeak dituen zonalde bat bereizi daitezke 3.2.2 irudiko espektroan: 3610-3640 cm^{-1} tartean glukosaren OH askeei dagokie. Talde honek ez ditu beste molekularekin hidrogeno loturarik eratzen baina haien artean hartutako norabidea bata bestearengan duen influentzia dela eta hartzen dute, sare bat eratuz.

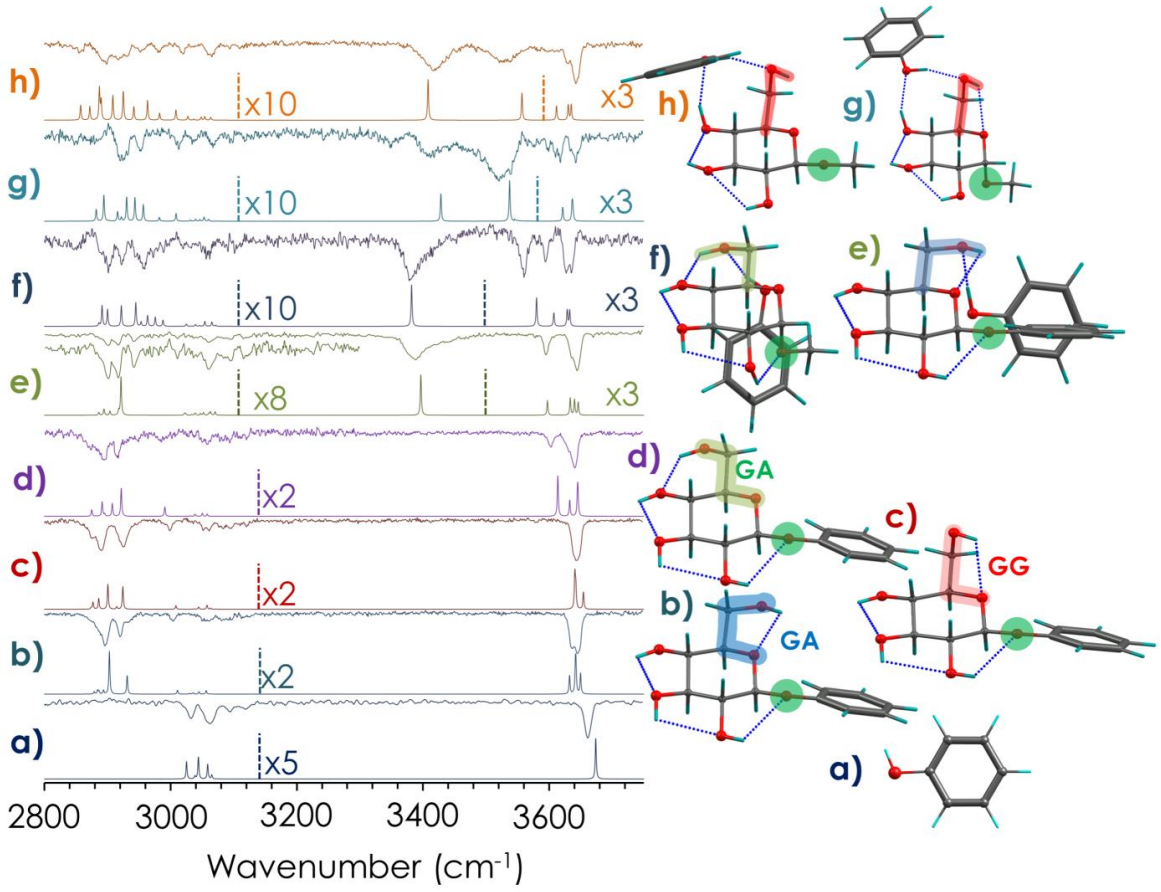
Fenolaren OH askea ~3650 cm^{-1} maiztasunean agertzen da, eta hau aurretik beste autora batzuk argitaratutakoarekin bat dator.[17] Fenolaren OH-ak duen edozein interakzioak gorrirantz desplazatuko du bere trantsizioa espektroan. Erakusten duen desplazamendua interakzioaren indarrarekin proportzionala izango da. Dimeroan fenolaren trantsizioak espektroan (3.2.2-e,h irudia) gorrirantz doan desplazamendua erakusten du, eta honek interakzio erdi indartsua duela esan nahi du.

Faltatzen diren trantsizioei dagokionez 3.2.2 irudiko espektroak, beti OH banda kopuru bera du, tontorrak agertzen diren posizioak molekulak eragiten duen interakzioaren menpe daude. Trantsizioen desplazamenduak informazio oso garrantzitsua ematen du gero espektro teorikoekin alderatzeko eta esleipena egiteko.

Kalkulu asko egin dira garrantzitsua den minimoren bat ez galtzearren, testuan bakarrik esleitutako egiturak eta kalkuluak erakusten dira baina datu guztiak apendizeetan aurki daitezke (Figures: A3.1-d), A3.2-c), A3.3-d)). Espektro esperimentalen eta teorikoen konparaketa osoa ere bertan aurki daiteke. Kapituluan zehar ikusgai dauden konparaketak zuzenean esleitutako egiturekin egin dira errazago ulertzeko eta laburbiltzeko.

α-MeGlc·P konplexuaren bi IDIRS espektro neurtu dira, hauek UV laserraren ionizazio uhin luzaeraren arabera gertatzen da (Figure 3.2.2-f), -g)). Espektro esperimentala eta teorikoa alderatzen direnean bi isomero daudela bereizi daiteke, bakoitzari espektro bat egokituz α-MeGlc·P-1 and -3. Bi familiak OH-en arteko elkarrekintzan bereizten dira: lehenengoa 6 oxigenoarekin elkarrekintzan dabilena ($O_{Ph}H \cdot O6$) (α-MeGlc·P-1, 3.2.2-f)eta bigarrena hidrogeno lotura oxigeno

4 eta 6 artean kokatzen den familia (O4H· O_{Ph}H· O6) (α-MeGlc·P-3, 3.2.2-g). Bi familia hauek hidrogeno lotura eta hidrogeno sare intramolekular desberdinak sortzen dituzte. (α-MeGlc·P-3, 3.2.2-g irudia) g konformeroa intramolekularrak diren glukosaren OH-ak erloju orratzen norantza jarraitzen dute eta (α-MeGlc·P-1, 3.2.2-f irudia) f konformeroa aldiz, kontrakoa. Hidrogenoen elkarrekintza hau molekulen arteko interakzioaren sustatzaile moduan funtziona dezake.



Irudia 3.2.2: IDIRS espektroak: **a)** Fenola **b)** β-PhGlc-1 **c)** β-PhGlc-2 **d)** β-PhGlc-3 **e)** β-PhGlc·P-1 **f)** α-MeGlc · P-1 **g)** α-MeGlc · P-3 **h)** β-MeGlc · P-4.

β anomeroa daukaten bi sistema ikertu dira: β-PhGlc·P (3.2.2-e irudia)) eta β-MeGlc·P (3.2.2-h irudia)). Ez da berezitasunik ez aldaketa handirik ikusi UV laserra erabiltzen den REMPI espektroetan. Gainera isomero bakarra behar da bietako edozein azaltzeko (Figure A3.2 eta A3.3 ikusi, xehetasun gehiago izateko). β-PhGlc·P kasuan egitura egonkorrenarekin esleitu daiteke, honek lehenengo familiaren interakzio mota erakusten du eta baita IDIRS espektroak alderatzen direnean antzekotasunak ikus daitezke (O_{Ph}H· O6; (β-PhGlc·P-1, 3.2.2-e eta α-MeGlc·P-1, 3.2.2-f) bi kasuetan fenola interakzioa erakusten duen. Ziurrenik P-ren lotura hau guztiz beharrezkoa da egitura hau egiteko.

β-MeGlc·P kasuan esleitutako familia ez da egitura egonkorrena, α anomeroan ikusitako bigarren familiarekin bat egiten du, bi hidrogeno zubi eratzen duena (O4H· O_{Ph}H· O6). Honek ere α anomeroaren IDIRS espektroarekin antzekotasunak ditu (O4H· O_{Ph}H· O6; (β-PhGlc·P-4, 3.2.2-h eta α-MeGlc·P-3, 3.2.2-g). Azkenengo bi sistema hauen datu guztiak apendizeetan aurki daitezke (Figure A3.2 eta Figure A3.3).

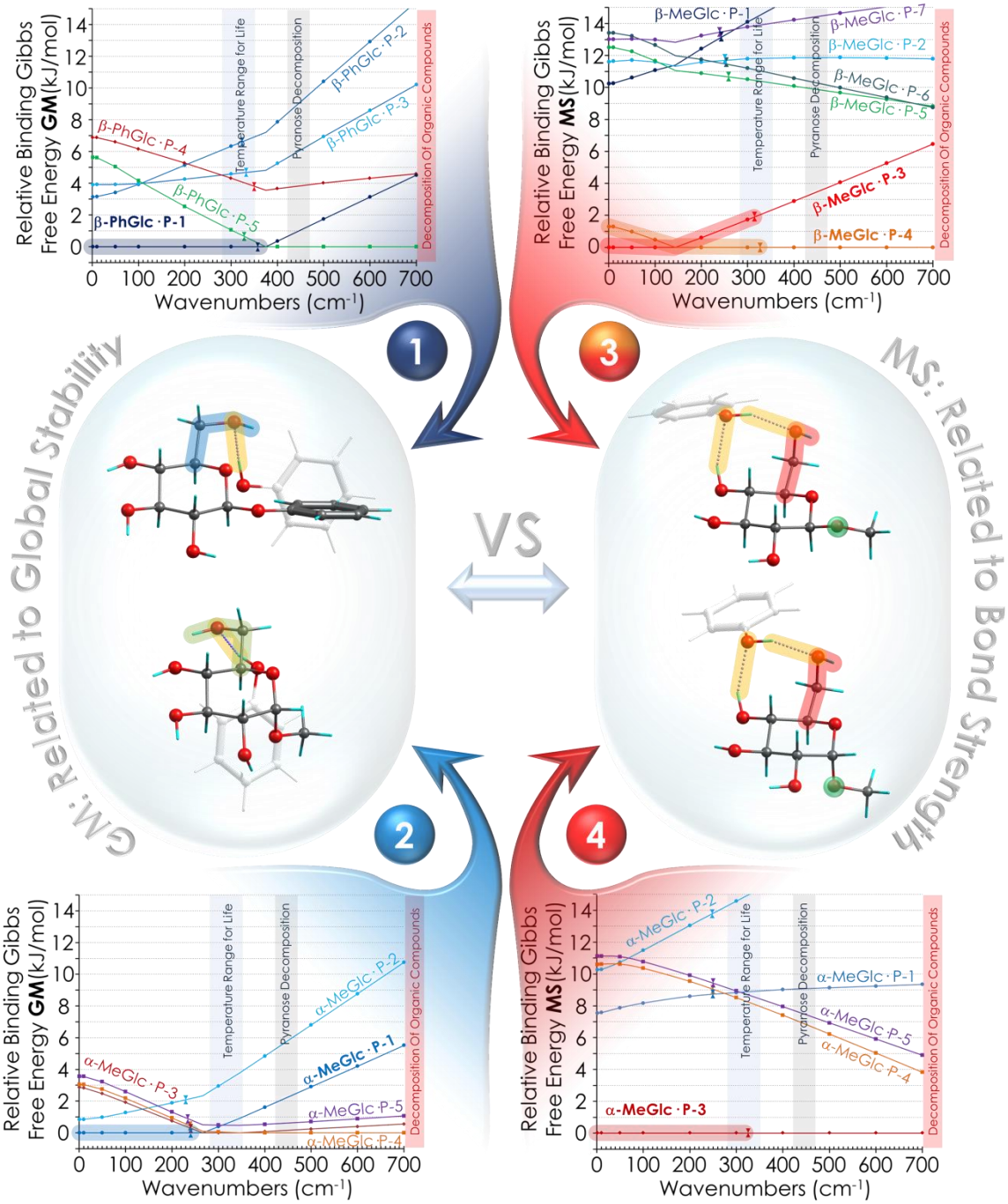
3.3. Emaitzen Eztabaida

Lehengo atalean aipatu den bezala β-MeGlc·P konplexuan esleitutako familia egonkortasunean bigarrena da (O4H· O_{Ph}H· O6) eta arraroa da egitura egonkorrena ez izatea. Hau Franck-Condon parametroengatik gerta daiteke, hau da, egoera kitzikatuan egituraren aldaketa bat gertatzen denez trantsizioa emateko probabilitatea asko jezten da trantsizioa desagertuz. Beste alde batetik minimo orokorraren familia (O_{Ph}H· O6) beste konplexuetan detektatua izan da; arraroa da hain egitura antzekoen artean aldaketa hain handia gertatzea, eta horrek beste arrazoi batengatik izan daitekeela pentsatzera eramaten gaitu.

Figure A3.4 irudian Gibbs energia askearen diagramak erakusten dira. Bigarren kapituluan azaldu den moduan konplexu batetan energia hau bi eratan kalkulatu daiteke: lehenengoa Gibbs-en energia erlatiboa da, hau da egitura bakoitzaren energia minus minimo orokorra (GM izena du, ikusi Figure A3.4). Bigarren erak Gibbs lotura energian kalkulatzean datza, hau egiteko dimeroaren energiari monomeroena kendu behar zaio, baina kontuz monomeroa dimeroan duen egiturarena izan behar da, gero hau ere erlatiboan bihurtzen da, Gibbs-en lotura energia erlatiboa lortuz (MS izena du, ikusi Figure A3.4). Eraikitako diagramek dimeroek tenperaturarekin duten dependentzia erakusten dute; askotan gertatzen da, minimotik hurbil agitura asko badaude, minimo orokorra tenperaturarekin aldatzen dela. Energia erlatiboan minimo asko daudenean baina lotura energi erlatiboan minimo bakarra dagoenean eta besteekin alderatuta diferentzia nahikorekin (~8 kJ/mol adibidez), kasu honetan gerta daiteke gogorren lotuta dagoen egitura ikustea eta ez egonkorrena.

3.3 irudiak eratutako lau konplexuen Gibbs energiak erakusten ditu, diagrama guztiak apendizeko Figure A3.4 aurkitu daitezke. 3.3 irudian ikus daitekeen datuen arabera bi egoera desberdin daude. β-PhGlc·P dimeroaren kasuan badu energi erlatiboan beste egitura guztiek baino egonkorrago den egitura bat (3.3-1 irudia begiratu, urdin ilunez azpimarratuta dago minimoa). Nahiz eta lotura energiari

begiratzean β-PhGlc· P-3 konformeroaren antzekoa izan (MS diagrama Figure A3.4-c$_2$ dago) bakarrik lehenengo isomeroa detektatu da, kasu honetan arautzen ari dena energia erlatiboa da.



Irudia 3.3: Lehenengo Familia 1&2) β-PhGlc· P eta α-MeGlc·P dimeroak eta Gibbs energia askearen kalkuluak (GM); **Bigarren Familia 3&4)** β-MeGlc·P eta α-MeGlc·P dimeroak eta Gibbs lotura energia askearen diagrama (MS).

β-MeGlc·P dimeroan egoera kontrakoa da: GM diagramak (Figure A3.4-b_2) antzeko energia duten egitura asko erakusten ditu, lehengo konplexuan MS diagraman gertatzen zen eran. Dena den kasu honetan MS diagraman egitura egonkorrenetatik β-MeGlc·P-3 & -4 besteetara energia diferentzia handia da, ~10 kJ/mol tenperatura tarte osoan zehar (ikusi 3.3-3 irudia, laranja eta gorriz margotuta). Energia hau metodologia honen akats tipikoa(~4 kJ/mol)[19] baino handiagoa da. Espektro esperimental eta teorikoaren konparaketa egitean β-MeGlc·P-3 & -4 esleitzeko hautagai onenak direla ikus daiteke. Bi isomero hauen arteko diferentzia bakarrak fenola azukrearekiko duen angeluan datza, beti lotura berdinak dituztelarik, bi egituren arteko hezia txikia izango da eta espantsioan egonkorrenera erlaxatuko da. Laburbilduz esperimentua eta teoriaren konparaketa gehi lotura energiaren kalkuluek esleipena baieztatzen dute. Konplexu honetan arautzen duen parametroa molekulen arteko Gibbs-en lotura energia izanik. Lotura indar berezi honek hozte prozesuaren talkak bizirauteko probabilitateak handitzen ditu.

α-MeGlc·P konplexuan egoera korapilatsuagoa da: alde batetik GM diagraman α-MeGlc·P-1konformazioa beste egiturekiko nahiko egonkorragoa dela ikusten da, nahiz eta egituraren bat hurbil izan(3.3-2, irudiko urdinez nabarmenduta). Baina MS diagraman argi eta garbi agertzen da α-MeGlc·P-3 konformeroa egonkorragoa dela beste guztiekin alderatuta(3.3-4, irudia gorriz nabarmenduta), bigarren egitura egonkorra baino 7 kJ/mol egonkorragoa da tenperatura tarte osoan zehar. Kasu honetan espantsioan egitura biak detektatuak izan dira, egitura bakoitza familia bati dagokio eta biak ikusten badira horrek isomerazioa debekatuta dagoela esan nahi du, eta hezia altua izango da.

Emaitza guztiak laburbilduz ikusitako joera guztiak 3.3 irudian batu dira; bertan β-XxGlc·P dimeroak α-MeGlc·P bi konformazioekin erlazionatuta daude, eta esleitutako egitura guztiak bi familietan bananduta daude: lehenengoa ($O_{Ph}H\cdot O6$) non egitura egonkorrena duten azukre dimeroak dauden (3.3 irudiko esker aldea), eta bigarren familia ($O4H\cdot O_{Ph}H\cdot O6$) lotura indar handiena duten egiturei dagokio (3.3 irudiko eskuin aldea)). Lehengo familia egitura egonkorrenei dagokio, edozein kasutan, bai monomeroak bananduta edo dimeroa eratzen badaude, molekula guztiak egitura egonkorrenean daude; bigarren familian aldiz, lotura indar altuena izatea da bere ezaugarri nagusia, eta horretarako monomeroan ez da egonkorren hartu lezaketen egitura, konplexua egonkortzea intramolekularra galdu eta intermolekularra irabazten du. Azken hau IC fasean lehenengoz isolatzea lortu baita.

3.4. Ondorioak

Konplexu baten energia erlatiboa gas egoeran egindako ikerketetan, kontutan izan beharreko parametro garrantzitsuenetariko bat da, beti ere esleipena zuzen egin nahi bada. Ikerketa honen bidez kontutan izan behar den beste ezaugarri bat aurkitu da, eta honek energia erlatibo baxuen eta lotura indar altuen duten konplexuen artean ematen den lehiaketan datza. Molekula baten barruko atomoen arteko elkarrekintzak lotura kobalente indartsuen bidez ematen dira baina konplexu batetan lotura kobalenteak ez diren molekulen arteko interakzio ahulen bidez ematen dira. Konplexua eratzeko jokatzen duten bigarren mailako indar hauen maximizatzeak konplexuari espantsioan ematen diren talkei bizi irauteko gaitasuna ematen dio, eta zenbait kasu partikularretan konplexu egonkorrenarekin batera detektatu daiteke (α anomeroarekin ikusi den moduan). Molekulen arteko elkarrekintza metodo hau konplexu kimikoek erakusten duten oreka kimikoen oinarria izan daiteke, biokimikan ematen diren oreka kimikoen parte garrantzitsua izanez. Berez elkarrekintza energia altuago bat izateak egoera dinamiko desberdin askotan konplexua erakusten duen biziraupenean efektua izango du, konpetentzi handiko giroetan bizirik irauteko lagungarria izango delarik. Horrela "Bizirik Dirauten Iaioak" diren konplexuak direla esan daiteke. Ikerketa honetan mota honetako konplexu bat lehenengoz aurkitu eta isolatu egin da.

3.5. Erreferentziak

(1) Dill, K.; Bromberg, S. *Molecular Driving Forces. Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience*, 2nd ed.; Taylor & Francis, 2010.

(2) Gamez-Pozo, A.; Berges-Soria, J.; Arevalillo, J. M.; Nanni, P.; Lopez-Vacas, R.; Navarro, H.; Grossmann, J.; Castaneda, C. A.; Main, P.; Diaz-Almiron, M.; Espinosa, E.; Ciruelos, E.; Fresno Vara, J. A. Combined Label-Free Quantitative Proteomics and microRNA Expression Analysis of Breast Cancer Unravel Molecular Differences with Clinical Implications. *Cancer Res.* **2015**, *75* (11), 2243–2253.

(3) Dweep, H.; Sticht, C.; Pandey, P.; Gretz, N. MiRWalk - Database: Prediction of Possible miRNA Binding Sites by " Walking" the Genes of Three Genomes. *J. Biomed. Inform.* **2011**, *44* (5), 839–847.

(4) Young, D. C. *COMPUTATIONAL DRUG DESIGN A Guide for Computational and Medicinal Chemists*; WILEY, 2009.

(5) Riley, K. E.; Hobza, P. Noncovalent Interactions in Biochemistry. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1* (1), 3–17.

(6) Nelson, D. L.; Nelson, D. L.; Lehninger, A. L.; Cox, M. M. *Lehninger Principles of Biochemistry*; W.H. Freeman: New York, 2008.

(7) Trapani, G. M.; Altomare, C.; Sanna, E.; Liso, G. B. and G. Propofol in Anesthesia. Mechanism of Action, Structure-Activity Relationships, and Drug Delivery. *Current Medicinal Chemistry*. 2000, pp 249–271.

(8) Jorgensen, W. L. The Many Roles of Computation in Drug Discovery. *Science* **2004**, *303* (5665), 1813–1818.

(9) Garman, E. F. Developments in X-Ray Crystallographic Structure Determination of Biological Macromolecules. *Science (80-.).* **2014**, *343* (6175), 1102 LP-1108.

(10) Glaeser, R. M. Review: Electron Crystallography: Present Excitement, a Nod to the Past, Anticipating the Future. *J Struct Biol* **1999**, *128* (1), 3–14.
(11) De Vries, M. S.; Hobza, P. Gas-Phase Spectroscopy of Biomolecular Building Blocks. *Annu. Rev. Phys. Chem* **2007**, *58*, 585–612.
(12) *Jet Spectroscopy and Molecular Dynamics*; Hollas, J. M., Phillips, D., Eds.; 1995.
(13) León, I.; Millán, J.; Cocinero, E. J.; Lesarri, A.; Castaño, F.; Fernández, J. A. Mimicking Anaesthetic-Receptor Interaction: A Combined Spectroscopic and Computational Study of Propofol···phenol. *Phys. Chem. Chem. Phys.* **2012**, *14* (25), 8956–8963.
(14) Usabiaga, I.; Gonzalez, J.; Arnaiz, P. F.; Leon, I.; Cocinero, E. J.; Fernandez, J. A. Modeling the Tyrosine-Sugar Interactions in Supersonic Expansions: Glukopiranosa-Phenol Clusters. *Phys. Chem. Chem. Phys.* **2016**, *18* (18), 12457–12465.
(15) Alonso, J. L.; Lozoya, M. a.; Peña, I.; López, J. C.; Cabezas, C.; Mata, S.; Blanco, S. The Conformational Behaviour of Free D-Glucose—at Last. *Chem. Sci.* **2014**, *5* (2), 515.
(16) Usabiaga, I.; González, J.; Arnáiz, P. F.; León, I.; Cocinero, E. J.; Fernández, J. A. Modeling the Tyrosine-Sugar Interactions in Supersonic Expansions: Glukopiranosa-Phenol Clusters. *Phys. Chem. Chem. Phys.* **2016**, *18* (18), 12457–12465.
(17) Billes, F.; Mohammed- Ziegler, I. Vibrational Spectroscopy of Phenols and Phenolic Polymers. Theory, Experiment, and Applications. *Appl. Spectrosc. Rev.* **2007**, *42* (4), 369–441.
(18) Talbot, F. O.; Simons, J. P. Sugars in the Gas Phase: The Spectroscopy and Structure of Jet-Cooled Fenil β-D-Glucopyranoside. *Phys. Chem. Chem. Phys.* **2002**, *4* (15), 3562–3565.
(19) Marianski, M.; Supady, A.; Ingram, T.; Schneider, M.; Baldauf, C. Assessing the Accuracy of Across-the-Scale Methods for Predicting Carbohydrate Conformational Energies for the Examples of Glucose and a-Maltose. *J. Chem. Theory Comput.* **2016**, *12*, 6157–6168.

4. β-D-glukopiranosa Ekuatoriala: Propietate Estrukturalak Karbohidrato Ugarienean.

4.1. Sarrera eta Helburuak

Azukreak oso molekula erabilgarriak dira, leku askotan aurki daitezke funtzio oso desberdinak betetzen, adibidez: egituren oinarri giza zelulen horman[1] edo DNAn,[2,3] eta errekonozimendu molekularren prozesuetan, zelula bereizle moduan. Berez, askotan, zelulen mintz plasmatikoaren bestaldean aurki daitezke proteina edo lipido bati itsatsirik.[4,5,6] DNAn eta proteinetan gertatzen den eran, non nukleotido edo aminoazidoak kodea idazteko beharrezkoak diren hizkien funtzioa betetzen duten, azukreek ere gaitasun hau erakusten dute. Hauen kasuan hidroxiloen posizioa da hizkiak desberdintzen dituena, posizio bakoitza azukre zehatz bati dagokiolako. Molekula hauek bi motatako zikloak eratu ditzakete, bost eta sei atomodunak. Molekula hauek epimero familiaren barruan sailkatzen dira eta beraien artean bereizgarri nagusia hidroxiloen sekuentziak dira, bai ekatorial/axial posizioetan nahastuta aurkitu daitezkeelako. Epimeroak beste aldaketa asko jasan ditzateke azukre berriak eratuz eta hizkien kopurua handituz. Epimero hauen artean, garrantzitsuenetariko bat β-D-glukopiranosa da, hau naturan ugarienetarikoen artean dago eta funtzio asko eta desberdinetan parte hartzen du.

Kapitulu honen helburua β-D-glukopiranosa beste karbohidratoengandik bereizten dituen propietateen ebaluazioa da. Oso bereizgarria da, azukre guzti hauen artean, zikloari lotutako hidroxiloak, ekuatorial planoan kokatuta duen piranosa bakarra izatea.

Esperimentalki balioztatutako metodologia erabilita molekularen energi potentzialaren gainazala (PES) ikertua izango da. Glukosak berez duen malgutasuna dela eta, normalean egiten den hiru dimentsioetako errepresentazio grafikoa, ez da posible. Arazo honi aurre egiteko beste modu berri bat proposatzen da, behin egitura egonkorrenak kalkulatuta, hauek antzekoak diren egiturekin konektatuko dira, azkenean egitura batetik bestera pasatzeko beharrezkoa den energi hezia kalkulatuz. Nahiz eta honelako errepresentazioak benetako PESetik urrun egon oso informazio erabilgarria emango digu.

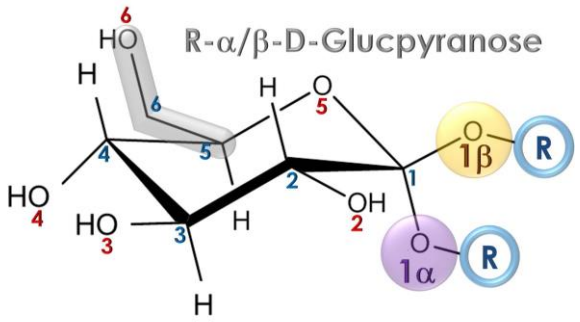
Behin monomeroaren propietate kimikoak determinatuta eta informazio honetan oinarrituz, konplexuak nola eratzen dituen ikertzera bideratuko da ikerketa. β-D-glukopiranosaren konplexatzeko gaitasunak argitzea helburu gisa, zenbait dimero ikertuko dira. α, β eta hauen arteko nahasketak eratu eta laser espektroskopiaren bidez hidroxilo ekuatorialen eta ez ekuatorialen efektua bereiziko da. Molekula

hauen artean lotura indartsuagoak erakusten dituen dimeroren bat dagoen bereizteko, Gibbsen energia askearen grafikoa kalkulatuko da.

Azkenengoz egitura kristalografiko asko datu baseetatik hartu eta deskonposatu egingo dira. Egituren artean monosakaridoak, disakariadoak eta polisakaridoak egoteko aukeratu dira. Benetako esfortzua zelulosaren egituraren deskonposaketan jarriko da, zelulosa oso interakzio intermolekular interesgarriak eratzen baititu kristalean. Polisakarido hau oso arrunta da, egurrean kontzentrazio altuetan aurki daiteke eta orokorrean landareetan funtzio estruktural bat betetzen du. Askotan zelulosa kristala bata bestearen gainean jarritako geruzez eratuta dago, hauek hidrogeno lotura sare handia sortzen dute plano ekuatorial osoan zehar, gaitasun hau egitura osoa egonkortzen duelarik. Irudi orokor honetatik hasita, hidroxiloen kokatze ekuatoriala eta hauek propietate makromolekularretan duten eragina determinatzea da helburua. Azkenean monomerotik, dimeroetatik eta kristaletatik ateratako informazio guztia kontutan hartuta hidroxilo ekuatorialekin duten erlazioa ezagutaraziko da.

4.2. β-D-glukopiranosa Monomeroak: Egituraren Analisia

PESaren analisia oso arrunta da,[7] normalean ikertutako molekularen propietate kimikoei buruzko informazioa eskuratzeko erabiltzen da. Hala ere, erakusten duten konplexutasunarengatik oso molekula sinpleetara murrizten dira, gehienez bi askatasun gradu (DF) dituzten molekuletara. Konplexutasun handiagoko pare bat eredu propofola[8,9] eta 1,2-butanodila[10] molekulak dira, hauetan PES errepresentazioa egiteko grafiko bat baino gehiago beharrezkoak dira, honek egoera asko nahasten du. Zenbat eta sistema molekular korapilatsuagoa ikertu orduan eta sinplifikazio gehiago behar izango dira informazio baliogarriak eskuratzeko.



Eskema 4.1: zenbaki urdinak karbonoak izendatzen dituzte, gorriak oxigenoak eta karbono anomerikoa morez edo horiaz bereizita daude.

Glukopiranosari dagokionez lau hidroxilo aurki daitezke bere zikloari lotuta, gainera bostgarren karbonoari lotuta hidroximetil talde bat du. Esandako atal bakoitzak zenbait DF gehitzen ditu (ikusi eskema 4.1 eta xehetasun gehiagorako apendizeko Scheme A4.1 begiratu) izugarrizko

konbinaketa kopurua posibleak direlarik. Arazo honi aurre egiteko molekula ataletan banatu behar da, zati bakoitzaren egonkortasuna ikertzeko eta behar diren kalkuluen kopurua murrizteko.

Molekularen deskonposaketan lehenengo parametroa heteroatomikodun sei partaidez osatutako zikloa da, bi posizioetan egon daitezke "*twist*" kiribilduta (txalupa posiziotik egonkorra den geometria hurbilena da) edo aulki "*chair*" egoeran. Oso energi diferentzia handia dago kiribilduta eta aulki posizioen artean horrexegatik sinplifika daiteke aulki egoerara, askoz egonkorragoa dena (xehetasun gehiago kapitulu biko apendizeetan aurkitu daitezke zehatz mehatz Figure A2.3an). hidroxiloei dagokionez bi dira egin daitezkeen sinplifikazioak: lehenengoa hidroxilo guztiak alde batera (β-Glc-03,04) edo bestera (β-Glc-09,10) kokatuta dituzten egiturak, bitartean dauden egitura egonkorren bidez (β-Glc-05-08) konektatzean datza. Isomerizazio energia asko jeisten da eraldaketa hidrogeno eklipsatutako posiziotik gertatzen denean (diagrama 4.1, laranjaz margotutako zutabea), karbono alifatikoekin eklipsatzen denean beharrezko energia askoz handiagoa da (diagraman zutabe urdinak). Bigarren sinplifikazioa eraztuneko karbonoekin lotuta dago, hauek tetraedro konformazioa daukate beren hiru minimoekin: Gauche-Anti 1 (GA 1), Gauche-Anti 2 (GA 2) eta Gauche-Gauche (GG). GG posizioa ez da benetako minimoa (diagramako zutabe gorria), honen ondorioz bi minimo eta eraldaketa mekanismo bat soilik izan behar dute kontutan, GA1 eta GA 2 (diagramaren bi zutabe berdeak)minimoak eta hidrogenoarenekin eklipsatuta duen eraldatze bidea alegia.

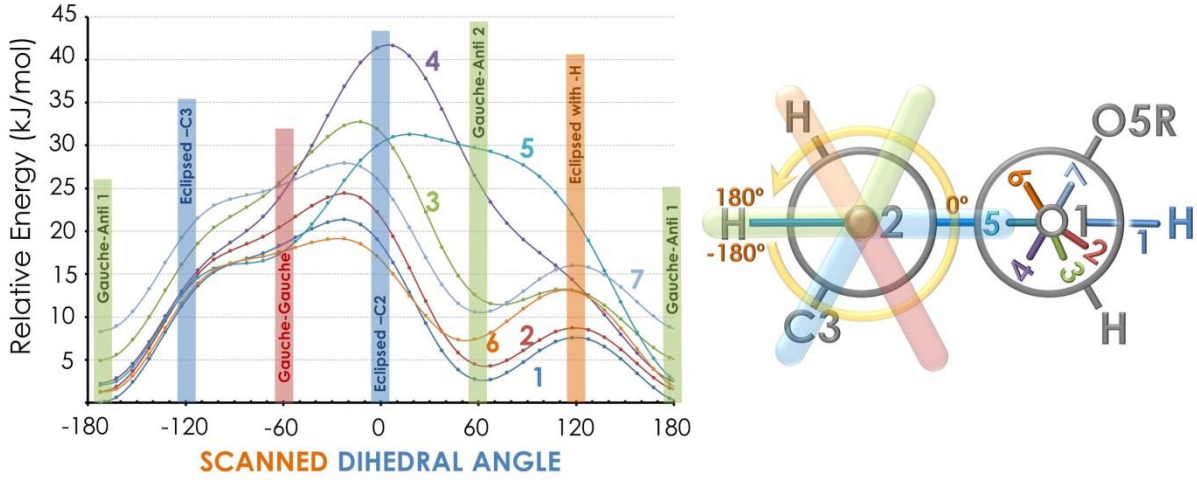
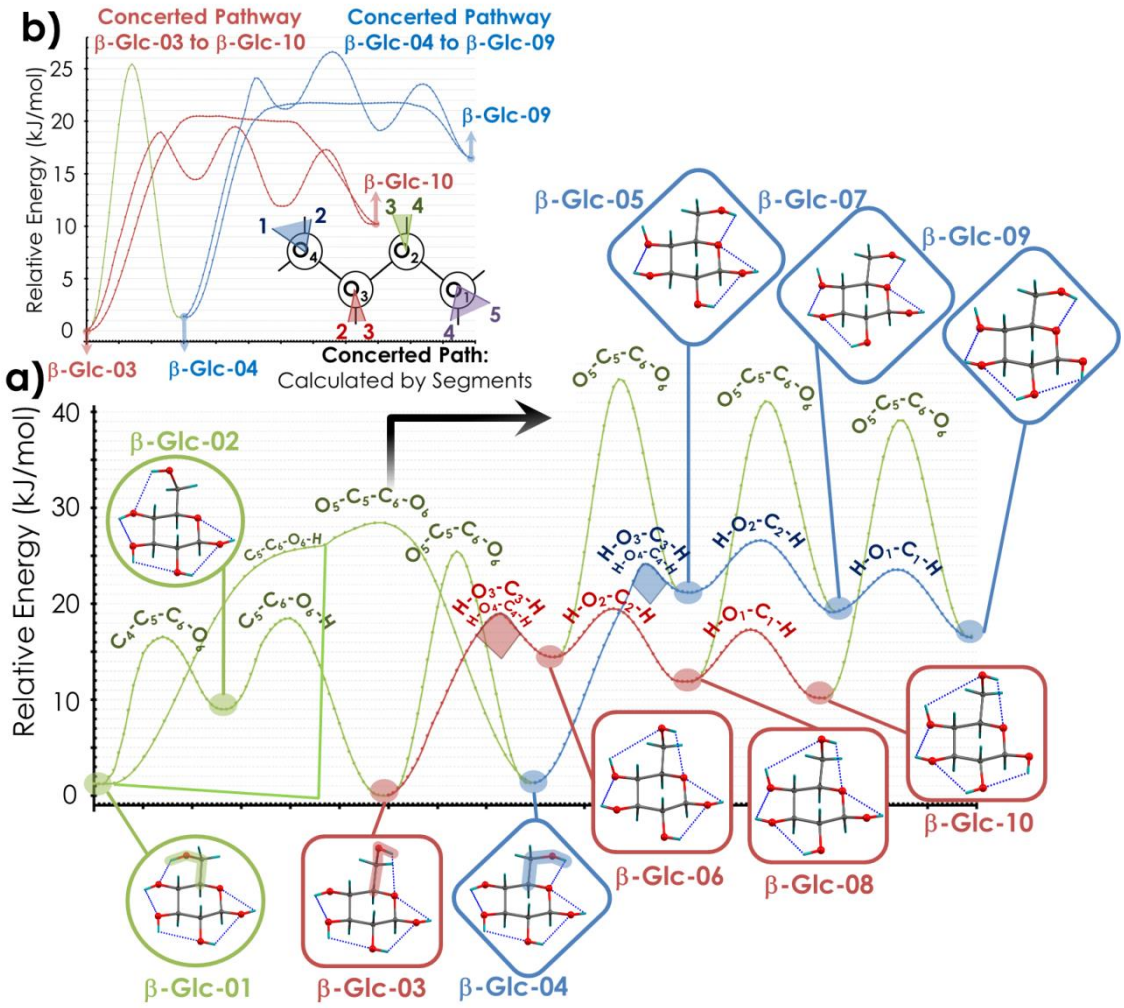


Diagrama 4.1: Newmanen proiekzioa O2-C2 eta O1-C1, O5-C1-O1-H angelu diedroa finkatuz bigarrena eskaneatzen da (C1-C2-O2-H). zazpi posizioetan finkatuta eskaneatzen da zazpi grafiko lortuz. Kalkuluak metodo eta mahia honetan eginda daude: M06-2X/6-311++G(d,p). Xehetasun eta kalkulu gehiagorako apendizeak begiratu, Figure A4.2.



Irudia 4.2.1: β-D-glukopiranosa molekularen SPES diagrama. a) Minimoak konektatzen dituen isomerizazio diagrama; b) Bat-bateko eraldaketa: DF asko aldatzen dira (Gorriz eta urdinez adierazitako "*plateaudun*" heziak) a atalean erakutsitakoekin konparatuta. Ikus daiteke nola hidroxilo guztiak parte hartzen duten eraldaketa prozesuan, β-Glc-03&04 egituretatik hasita β-Glc-10&09 egituretan bukatuz. Kalkuluen mahaia: M06-2X/6-311++G(d,p). Xehetasun gehiagorako behatu Table A4.1 jo.

Emaitza hauetan oinarrituz PESa aztertzeko metodo berria proposatu da. Behin egitura egonkor guztiak aurkituta PES grafikoa sinplifikatutako PES (SPES) grafikoarengatik aldatu daitezke, grafiko berri honetan minimoak kokatu eta haien arteko energi heziak kalkulatu dira egitura guztiak konektatuz. Kalkulu teoriko guztiak bezala hauek ere estimazioak dira eta kontuz erabili behar dira, zehazki β-glukosarentzako eraiki den diagrama 4.2.1 irudian ikusgai dago. Diagraman energian minimoak diren hiru egitura bereizi daitezke, beste minimo batzuk ere badaude baina 10 kJ/mol ezegonkorragoak dira. Gainera bi energia hezi ikus daitezke, alde batetik $\lesssim$ 10 kJ/mol energiadunak eta bestetik $\gtrsim$ 20 kJ/mol energiadunak, azkenengo hauek, hiru egitura egonkorrenak (β-Glc-01, 03, 04)

konektatzen dituztenak diren bitartean beste egitura guztiak ≲ 10 kJ/mol energiadun hezien bidez lotuta daude. Egituren arteko energia erlatiboa eta eraldaketa heziak kontutan izanda, espantsio supersonikoko oste prozesuari biziraungo dioten egitura bakarrak, hiru egitura egonkorrenak izango direla iradoki daiteke. Espantsioei buruz egindako zenbait kalkulu, egoera supersoniko baten batezbesteko abiaduran (Match 2, Figure A4.4-c) doan Ar atomo (Ar gas eramailea da lagina hoztu eta garraiatzen duena) batek gainditu ditzakeen energi altuera ≲10 kJ/molekoa da, ≳20 kJ/mol heziak kolisio hauen eraginenaren gainetik daudenez gainditzea ez da posible izango. Gas eramaileak egituren eraldaketa eragingo du hiru egitura egonkorretan erori arte, bertan heziak handiegiak direnez egitura horretan harrapatuta geratuko dira. Kalkulatutako datu hauek bat doaz lehenago beste autore batzuk argitaratutakoekin: hiru egitura egonkor hauek mikrouhinen espektroskopiaren teknikaz detektatutakoen berdinak dira,[11] gainera fenil-β-D-glukopiranosa molekularentzat aurkitutako egituren baliokideak dira.[12] Laser espektroskopiaren bidez esperimentu hau egin ahal izateko molekula honen beharra dago, teknika honek ionizatzeko kromoforo bat erabiltzen baitu. Lan honetan fenilglukosaren hiru egitura egonkorrenak berriro ikertu dira, batutako espektro esperimentala CHen eremura luzatuz, lortutako aparteko informazio aldez aurretik argitaratutako informazioa baieztatzen duena (Figure A4.1).

Molekularen hidroxiloen elkarrekintza zaiagotzen du geometrien jokabidea ulermena, posible egiten baitu bat-bateko eraldaketa, egitura batetik (β-Glc-03&04) aldaketa asko dituen beste egitura batzuetara pasatuz(β-Glc-10&09), ikusi 4.2.1 irudia. Mekanismo honek "plateau" egitura bat sortzen du (4.2.1.b irudia), hau lehenengo DF edo eskanaren koordenatua maximora ailegatzen denean hasten da, egoera horretan DF edo koordenatua aldatuz doa, energia konstante mantentzen den bitartean. Koordenatuak banan-banan begiratuta lau angelu diedroetatik pasatzen da bat-bateko eraldaketa hau: 4.2.1.b irudiko 1-2 angelu diedroa (tarte urdina) ordezkatua da 2-3 (tarte gorria), eta 3-4 (tarte berdea) angeluengatik bukaerako angelura heldu arte 4-5 (tarte morea), egitura hauek "plateau" forma ematen diete egitura isoenergetikoak direlako. Ikus daiteke nola gorri eta urdinen arteko energi diferentzia txikiagoa den "plateu" kasuen artean beste bi kasuen artean baino.

Atal honetako datu guztiak kontutan izanda monomeroaren jokabide konformazionala laburbildu daiteke: Alde batetik hidroximetiloa hiru posizio desberdinetan egon daiteke (GA β-Glc-01; GG β-Glc-03 eta AG β-Glc-04); beste aldetik, nahiz eta zenbait bitarteko egitura egonkor existitu (β-Glc-05, 06, 07 eta 08)

zikloari lotutako hidroxiloak batera funtzionatzen dute, hau bata bestean duten eraginarengatik ematen da.

4.3. β-D-glukopiranosa Dimerizazio Prozesua isolatuta eta hotz dagoen fasean (CI)

Zenbait faktore kontutan izan behar dira dimerizazio prozesua aztertzeko, alde batetik oxigeno anomerikoa non kokatuta dagoen, aldaketa txiki honek guztiz aldatu dezake molekulen arteko lotura. Orokorrean α anomeroa egonkorragoa da eta zikloaren apurketaren bidez batetik bestera eraldatu daiteke (α eta β), populazio guztia α anomeroan pilatuta aurkitu ohi delarik. Anomeroen arteko eraldaketa blokeatu daiteke oxigeno horri zerbait atxikituz metil/fenil-α/β-D-glukopiranosen kasuan bezala. Beste alde batetik, atxikitutako atal horrek influentzia izan dezake molekulen arteko elkarrekintzan: D-glukopiranosa, metil-α/β-glukopiranosa eta fenil-α/β-glukopiranosa kontsideratu dira, hauen arteko dimeroak eratzeko, karbo-oxigeno anomerikoaren simetrian eta geometrian duten eragina ikertzeko. Fenil-α-D-glukopiranosaren kasuan salgai aurkitu ez daitekeenez α-α dimeroen ikerketa kalkulu teoriko soilen bidez egin da.

Azkenengo, fenil-β-D-glukopiranosa·D-glukopiranosaekin egingo da, D-glukopiranosa bi anomeroen nahasketa sortzen du gas egoeran. Dimero honen eraketak hautagai asko ditu espektro esperimentalari esleitzeko. Beste aldetik,aukera ematen du anomeroen arteko lehiaketan zeinek irabazten duen ikertzeko.

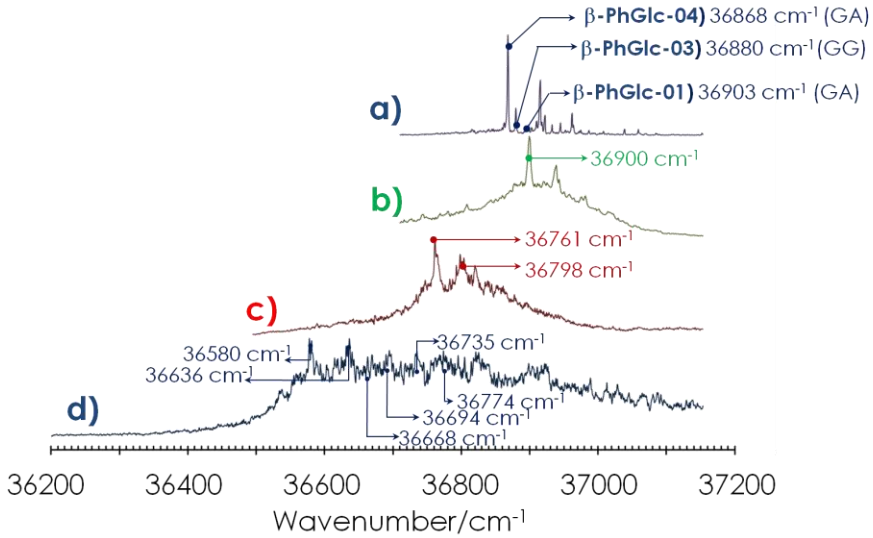
Lehengo emaitza teoriko eta esperimentalengatik hasiko da ikerkuntza, hauek baitira zehatzenak, soilik teoriak direnekin bukatzeko. Bukaeran zenbait tendentzia isolatu eta zehaztuak izango dira.

REMPI espektroak 4.3.1 irudian ikusgai daude, hauek kitzikatutako mahaien espektroa lortzeko eta IDIRS esperimentuak egiteko beharrezkoa den ionizazio uhin luzaera egokiena aukeratzeko erabiliko dira. Nahiz eta REMPI espektroengandik egiturari dagokion informazioa eskuratzeko prozesua oso zaila izan, horrenbesteko zailtasuna ez duen zenbait datu lortu daitezke. Fenil-β-D-glukopiranosaren espektroa 4.3.1 irudian ikus daiteke, hau erreferentziaz balio du: molekula hiru egituratan aurkitzen da, egitura hauek gradu desberdinetan interakzionatzen dute kromoforoarekin, zenbat eta interakzio handiagoa egon orduan eta efektu nabarmenagoa 0_0^0 trantsizioan. Monomeroaren 0_0^0 trantsizioaren desplazamendua konplexuenarekin alderatuz kromoforoak interakzioan duen paperari buruzko

informazioa atera daiteke. Fenil-β-D-glukopiranosa·D-glukopiranosa dimeroaren espektroak zenbait tontor mehe aurkezten ditu absortzio zabal baten gainean, intentsitate handiena duen tontorra 36900 cm^{-1} maiztasunean agertzen da. Trantsizio hau 0_0^0 trantzisioarena dela esaten badugu ondorioztatu daiteke nola molekulen arteko elkarrekintza kromoforotik urrun gertatzen dela. Izan ere tontor hau ez du monomeroen 0_0^0 trantzizioarekin alderatuta desplazamendu nabarmenik erakusten.

4.3.1 irudiari begiratuta (c) espektroa fenil-β-D-glukopiranosa·metil-β-D-glukopiranosari dagokio, kasu honetan nahiko zabala dela esan daiteke lehengo espektroarekin alderatzen bada, gainera energia gutxiagoko maiztasunetara desplazatuta du 0_0^0 trantsisioa (gorrira desplazatuta), kromoforoarekin interakzio zuzena duelako izan daiteke. Espektroaren bandak zabaltzen direnean askotan bi molekulen arteko lotura dakarren frekuentzia baxuko bibrazioengatik gerta ohi da.[13] Kromoforoa interakzio indartsuan parte hartzen ari denean efektu hau gehieago nabarmentzen da.

Azkenengo espektroa (d) fenil-β-D-glukopiranosa·metil-α-D-glukopiranosarena dagokio, honek erakusten duen egitura monomeroenarekiko oso bestelakoa da: zabala eta determina ezina den egitura aurkezten du, gainera monomeroen 0_0^0 trantsisiotik ~ 300 cm^{-1} gorrirantz desplazatuta du hasiera. Desberdintasun hau zenbait jatorri desberdin izan ditzake, alde batetik, konformazio bat baino gehiago egotea eta bestetik, gutxienez konformazio horietatik bat kromoforoarekin elkarrekintza gogorreko egitura bat hartu izana.



Irudia 4.3.1: REMPI espektroak a) β-D-glukopiranosa b) fenil-β-D-glukopiranosa· D-glukopiranosa c) fenil-β-D-glukopiranosa ·metil-β-D-glukopiranosa d) fenil-β-D-glukopiranosa·metil-α-D-glukopiranosa.

Behin REMPI espektroetatik informazioa guztia aterata bigarren urratsarekin jarraitu daiteke, IR espektroen interpretazioa. Informazio zehatzena espektro teorikoak eta esperimentalen konparaketatik dator. Espektro teoriko eta esperimentalen arteko esleipena ahalik eta ondoen egiteko, REMPI eta IR espektroengatik datorren informazioa, Gibbs energia askearen balioekin erlazionatuko dira. Ahalik eta era argienean azaltzeko banan-banan tratatuko dira 4.3.2 irudian dauden konparaketak.

Lortutako IDIRS espektroak β-β dimeroarentzat (fenil-β-D-glukopiranosa⋯metil-β-D-glukopiranosa) 4.3.2 irudiaren(a) atalean aurkitu daitezke, proposatutako egitura teorikoarekin. Xehetasun gehiago lortzeko Figure A4.3 eta Table A4.2etara jo apendizeetan. Bertan ionizazio uhin luzaera desberdinetan lortutako espektroak daude (a) atala, teoriaren bidez eta esperimentalki lortutako espektroen arteko konparaketa (b) atalean erakusten dira. Gibbsen energiaren ikerketa osoa (c) grafikoetan daude eta egitura egonkor garrantzitsuenak (d) atalean.

Irudian (4.3.2 (a)) ikusten den konparaketan oso antzekotasuna handia bereizten da egitura egonkorrenaren espektroa esperimentalarekin alderatzen denean. Benetan ona den esleipena Gibbsen energiari begiratzean hobetu egiten da, egitura egonkorrena beste egitura guztiak baino 10 kJ/mol egonkorragoa da tenperatura tarte osoan zehar. Nahiz eta diferentzia hau izan, monomeroetan hiru egitura egonkor izanda, zerbait gerta liteke hezietan egitura bakar bat hain garbi ikusteko. Monomeroaren heziak ikusitako hiru egoeren artean eraldatzeko gutxi gora behera $\gtrsim$ 25 kJ/mol altuerakoak dira (Figure A4.4 (a), konplexuan hezi hori murriztu egiten da, $\lesssim$10 kJ/moletaraino jezten da minimo orokorraren egitura eskuratzeko norabidean, kontrako norabidean altua izaten mantentzen da ($\gtrsim$ 20 kJ/mol). Balio hauek gas eramaileak transmititu dezaketen energia zinetikoarekin alderatzen direnean (He eta Ar match 2ko abiadurarekin, xehetasun gehiago Figure 4.4 (c) atalean) ikus daiteke nola Ar gas eramaile moduan erabiltzen denean konplexuaren hezia minimoaren norabidean ahalbidetuta dagoen baina ez beste zentzuan, honen arrazoia dimero egoeran dagoelako gertatzen da, bigarren molekula zentzu bakar batean egonkortzen du bitarteko egitura, monomeroetan hezi hori gaindi ezina izaten jarraitze du eta horregatik monomeroaren neurketetan hiru isomero bereizten dira.

Gainera dimeroen isomerizazio hezian altueran tenperatura duen eragina kontutan izatean (Figure A4.4-d) dimeroak nahiko ahalbidetuta dute minimo orokorrerako bidea eta ez beste biderik. Egindako esleipen ona eta bestelako froga guzti hauek kontutan izanda ondorioztatu daiteke lortutako emaitza

benetan zehatza dela eta fenil-β-D-glukopiranosa· metil-β-D-glukopiranosaren 01 dimeroaren konformazioa dela dudarik gabe detektatutako egitura, xehetasun gehiagorako ikusi laugarren kapituluko apendizeetako 4.3 atala).

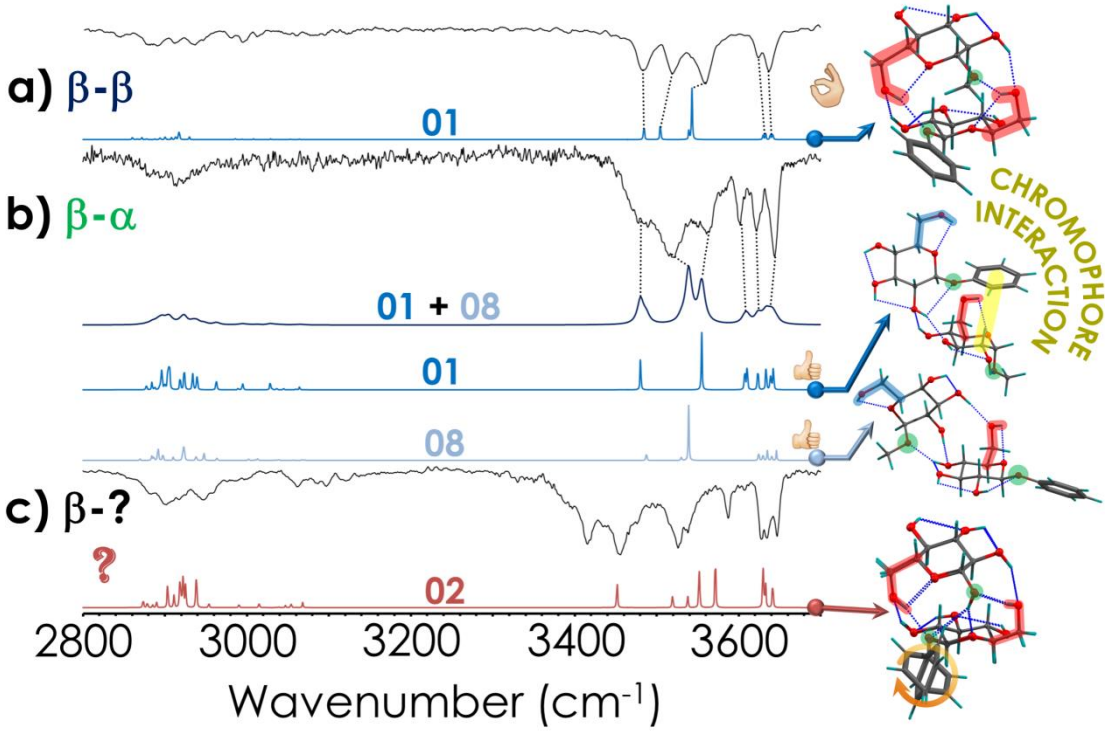


Figure 4.3.2: IDIRS espektroak: a) fenil-β-D-glukopiranosa · metil-β-D-glukopiranosa; b) fenil-β-D-glukopiranosa · metil-α-D-glukopiranosa; c) fenil-β-D-glukopiranosa · D-glukopiranosa. Espektro esperimentalak beltzez margotuta daude, koloredunak teoriaren bidez simulatutako espektroak dira, kalkuluen metodoa eta mahaia: M06-2x/6-311++G(d,p).

Fenil-β-D-glukopiranosa·metil-α-D-glukopiranosa dimeroa α-β anomeroak nahasten ditu egoera guztiz aldatuz. Egonkortasun erlatiboari buruz egindako kalkuluak diferentzia txikia erakusten du lehenago eta bigarren egituren artean (Table A4.3 beatu). 4.3.2 irudiko (b) atalean erakusten den espektro esperimentala nahiko ondo azaldu daiteke egitura egonkorrenarekin, baina kasu honetan ez da guztiz ulertzen gutxienez beste egituren baten ekarpenik gabe. Entalpiari begira egonkorrena den bigarren egitura ez du egokitze onena erakusten, Gibbsen energiari begiratuz espazioaren "konformazio-tenperaturatik" hurbil (100-200K, Figure A4.5) beste egitura bat egonkorrena da eta hau bai, nahiko ondo egokitzen da falta den tontorra azaltzeko. Konformazio bat baino gehiago dagoen jakiteko beste froga bat egin da, IDIRS espektroak ionizazio uhin luzeera desberdinetan neurtu dira, lortutako espektroak konparatzen direnean ikus daiteke nola tontorren altuera erlatiboa aldatzen den espektro batetik bestera. Fenomeno honek isomero bat baino gehiago dagoela adierazi nahi du.

Dimeroaren egitura garrantzitsuenak Figure 4.5-d-en ikusgai daude, 01 konformazioa kromoforoarekin interakzionatzen ari da baina esleitutako beste egitura, 08 ez du horrelakorik. 0_0^0 trantsizioarekiko ikusitako desplazamendua egitura hauen ondorioz gertatu da baina biak ez dute eragin bera eta seguruenik haien artean ere trantsizio honen kokapenean desberdintasun handiak dituztela nahiz eta ezin daitekeen egiaztatu.

Anomeroen arteko lehiaketa ikertzeko fenil-β-D-glukopiranosa···D-glukopiranosa dimeroa erabili da, honetan bi anomeroetatik bat β da baina bestea α edo β izan daiteke. Lehenago azaldu den bezala, badaude zenbait bide anomero bat bestean eraldatzeko, gainera α anomero egonkorrena dela aipatu izan da. Beste alde batetik, kontutan izan behar da fenil-β-D-glukopiranosa laserraren bidez gasean bihurtzean zenbait apurketa prozesu izan ditzakeela eta horien bidez β anomeroa sortu. Sistema honen datu guztiak apendizeetan daude (Tables A4.4, A4.5, Figures A4.6 eta A4.7)

IR-HB "*Hole-Burning*" teknika erabili da isomeroen kopurua jakiteko, honen bidez bi posibilitate sortzen dira, konformazio bakarra egotea edo bat baino gehiago, beti ere REMPI espektro berarekin. Beste aukera bat dago, nagusi den bat egotea eta espektroaren behealdean ikusten den absortzio zabala, nagusi ez diren beste konformero batzuena izatea, hauek IR espektroan duten ekarpena kendu ezinik. Baina gauzak gehiago nahasteko α anomeroa egonkorragoa da eta lagin askotan proportzio handiagoan aurki daiteke. Beste alde batetik, nahiz eta laginean kontzentrazio baxuetan egon, β anomeroa ablazio prozesuan sortu daiteke, gainera egitura askoz egonkorragoak ahalbidetzen ditu, β-β dimeroan ikusitakoen baliokideak alegia. Dena den suposatzen badugu espektro guztia konformero bakar batetik datorrela, azaldu dezakeen egitura bakarra β-β 02 egitura da (β-PhGlc · β-Glc, behatu Figure A4.6, 01 baliokidea da). Konformero hau trantsizio kopuru berdina duen bakarra da eta atalaren hasieran β-β fenil eta metil dimeroaren egitura baliokidea du. Informazio guztia kontutan izanda esan beharra dago dimero honen esleipena ez dela guztiz nahastezina.

α-α dimeroaren ikerketa (fenil-α-D-glukopiranosa···metil-α-D-glukopiranosa) soilik kalkulu teorikoetan oinarritzen da, hauen mahia lehengo sistemetan erabilitako prozedura jarraitu du (Tables A4.6 eta A4.7). Dimero honek kromoforoarekin elkarrekintzan dabilen egitura asko ditu, egitura hauek egonkorragoak dira eta ez dute azukreen beretako elkarrekintza bereizten usten(Figure A4.8-a), Gibbsen energi diagraman argi ikus daiteke kromoforoarekin lotura dutenak nola

egonkorragoak diren (behatu Figure A4.8-b eta c). Arazo honi aurre egiteko beste kalkulu batzuk egin dira, metilatutako dimeroarenak alegia ((metil-α-D-glukopiranosa)2) Table A4.8-k laburbiltzen ditu energiari buruzko emaitzak. Gibbs energiaren grafikoak eta egitura egonkorrenak Figures A4.9 (a) eta (b) ataletan aurki daitezke, (c) atalean konparaketa bat egiten da (metil-α-D-glukopiranosa)2 eta (fenil-α-D-glukopiranosa·metil-α-D-glukopiranosa) dimeroen artean, familietan antolatu dira erakusten duten loturaren arabera. Bien artean konparaketarako kromoforoarekin elkarrekintzan dabilen egiturak alde batera utzi dira, nahiko ondo egokitzen dira bi kalkuluak baina metilatutako dimeroak pare bat familia berri azaleratu ditu.

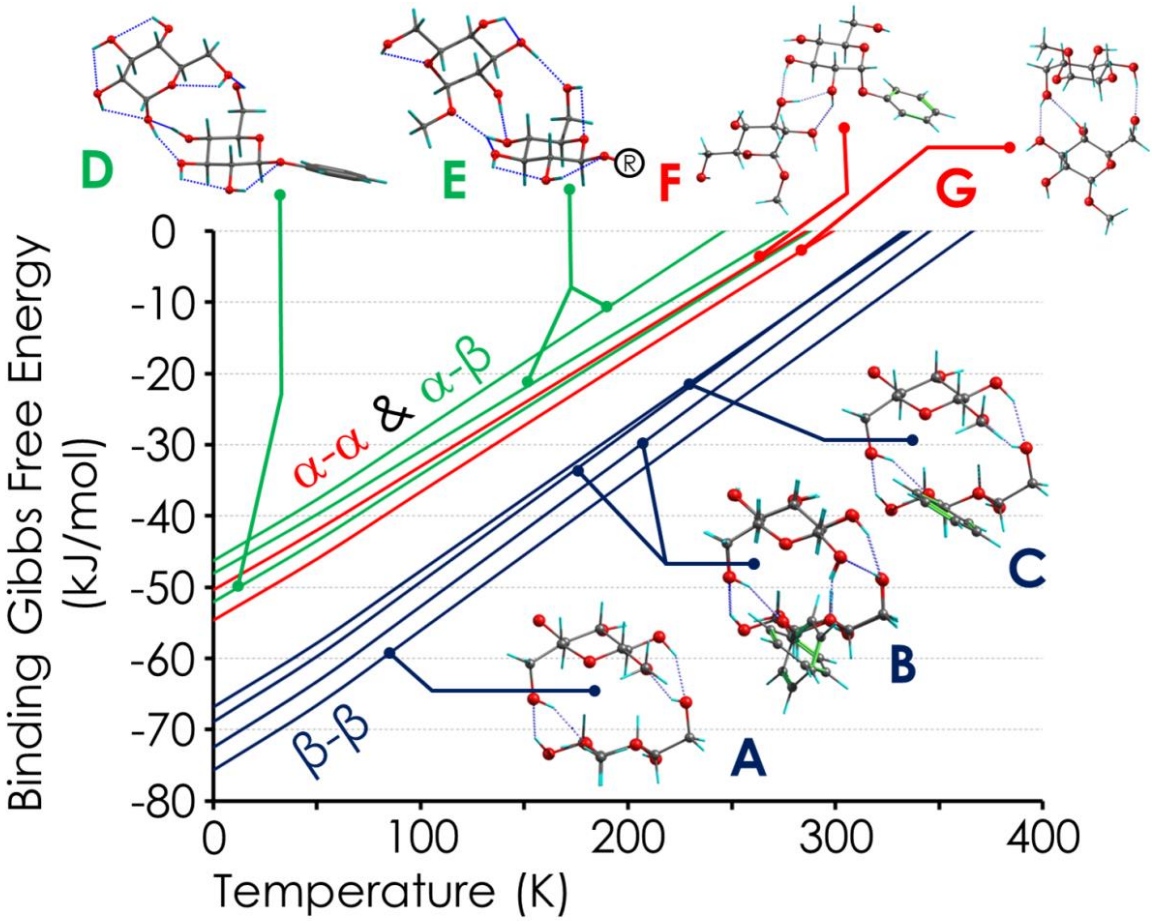


Figure 4.3.3: (MS) Gibbs-en lotura energiak: **A)** β-MeGlc · β-MeGlc (01 egituraren β-PhGlc · β-MeGlc metilatutako bertsioak) **B)** β-PhGlc · β-Glc (hautagai esperimentalak: 01 eta/edo 02) **C)** β-PhGlc · β-MeGlc (esperimentalki aurkituta: 01) **D)** β-PhGlc · α-Glc (04 konformeroa kromoforoarekin elkarrekintzarik gabe) **E)** β-Ph/MeGlc · α-MeGlc (08 egokitutako egitura eta bere metilatutako bertsioa) **F)** α-PhGlc · α-MeGlc (48 egitura, kromofororik gabeko egitura) **G)** α-MeGlc · α-MeGlc (01egitura teorikoa).

Dimero guzti hauen emaitzak batera jarri dira, kromoforodun egituretan kromoforoa metilo batengatik aldatu da eta kalkulatu dira egitura berri horiek (adibidez: fenil-β-D-glukopiranosa·metil-β/α-D-glukopiranosa → metil-β-D-glukopiranosa· metil-β/α-D-glukopiranosa). Datu guzti hauen konparaketa 4.3.3 irudian ikusgai daude, bertan dimero desberdinen konparaketa egiten da euren gibbsen lotura energia alderatuz. Diagraman ikus daiteke nola konplexu guztiak gutxienez hiru hidrogeno loturaz osatuta dauden. Badago salbuespen bat, β-β familia, dimero honek lau hidrogeno zubiz osatutako sarea eratzen du era simetrikoan, bi alde batetik eta beste bi bestetik (bost dituen bat badago, β-D-glukopiranosak osatzen duen dimeroa, oxigeno anomerikoa aske baitu beste lotura bat egiteko).

β-β konplexuaren hidrogeno loturen sarea eraikitzeko ekuatorial planoan dauden hidroxiloekin guztiz loturik daude. Honek loturen sarea eratzeko gaitasuna ez du hidroxiloetan aldaketa handirik eragiten ez du. Hidroximetilo bakoitza bi hidrogenozko lotura eratzen dituzte, badaude beste bi konplexu antzekoak eta interakzio simetrikoak duten egiturak, eta beste bi simetrikoak ez direnak. Egitura guzti hauen artean (4.3.4 irudia) soilik bat ikusi da (a) honen zergatia jakiteko beste kalkulu batzuk egin dira, beste lau egiturak eskuz eraiki eta egituraren optimizazioa eta frekuentzien kalkulua egin da (kalkuluen metodoa eta mahaia: M06-2X/6-311++G(d,p)). 4.3.4 irudia egitura hauek erakusten ditu, duten antzekotasunen bidez kokatuta daude, erdian urdinez simetrikoak direnak, alde batetik esleitutakoa (a) eta beste bi egitura simetriko posibleak (c eta e), energian 20 kJ/mol ezegonkorragoak direnak. Berez (c) eta (d) egituren hidrogeno loturen distantzia zehaztasunez neurtzen badira benetan bi hidrogeno lotura soilik eratzeko gai direla ikus daiteke. Beste bi egitura asimetrikoei dagokionez, haien asimetria hidrogeno lotura hurbilagoak eratzeko gaitasuna usten die baina inoiz (a) egituran ikusitako lau loturetara ailegatzea. Hiru ondorio atera daitezke emaitza hauetatik: (i) Alde batetik, hidroxiloen axial/ekuatorial kokagunea efektu handia duela osatutako dimeroengan eta hauek eratu dezaketen hidrogeno loturen sarean (ii) Bigarren mailako erakarpen indarren influentzia, CH eta oxigenoen arteko hidrogeno zubi ahulak dimeroa egonkortzen ditu (iii) Aldarapen indarren efektua CH-HC eta "bikote bakarti-bikote bakarti"-en artean, hauek azkeneko egitura mugatzen dute, ikusitako kasu honetan beste lau egiturei lau hidrogeno zubi egitea saihestuz. Azken finean, bikote bakartiek eta CHak hidroxiloen beste eragiten duten egituran, bertara ailegatzeko bidea blokeatu dezakete.

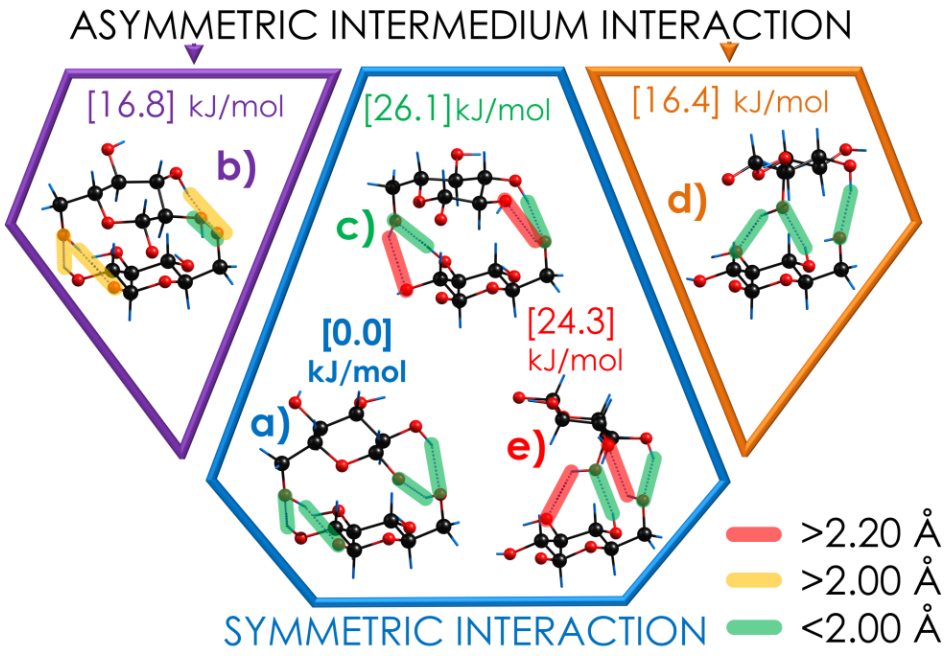


Figure 4.3.4: Esleitutako energi minimoaren egitura eta hidroximetil bakoitzeko bi hidrogeno lotura sortzen dituen beste egitura posibleak. Egituren optimizazioa metodo eta kalkulu maila: M06-2X/6-311++G(d,p) a)esleitutako minimo gorenaren egitura: P2M6P1-M2P6M1 b) Beste lotura mota: P2M6P1-M3P6M2 c) Beste lotura mota: P3M6P2-M3P6M2 d) Beste lotura mota: P4M6P3-M3P6M2 e) Beste lotura mota: P4M6P3-M4P6M3. Loturen kodeak apendizeetan azalduta dago, Scheme A4.1.

4.4. Poly-1,4-β-D-glukopiranosaren Kristalaren Eraketa (Zelulosa I)

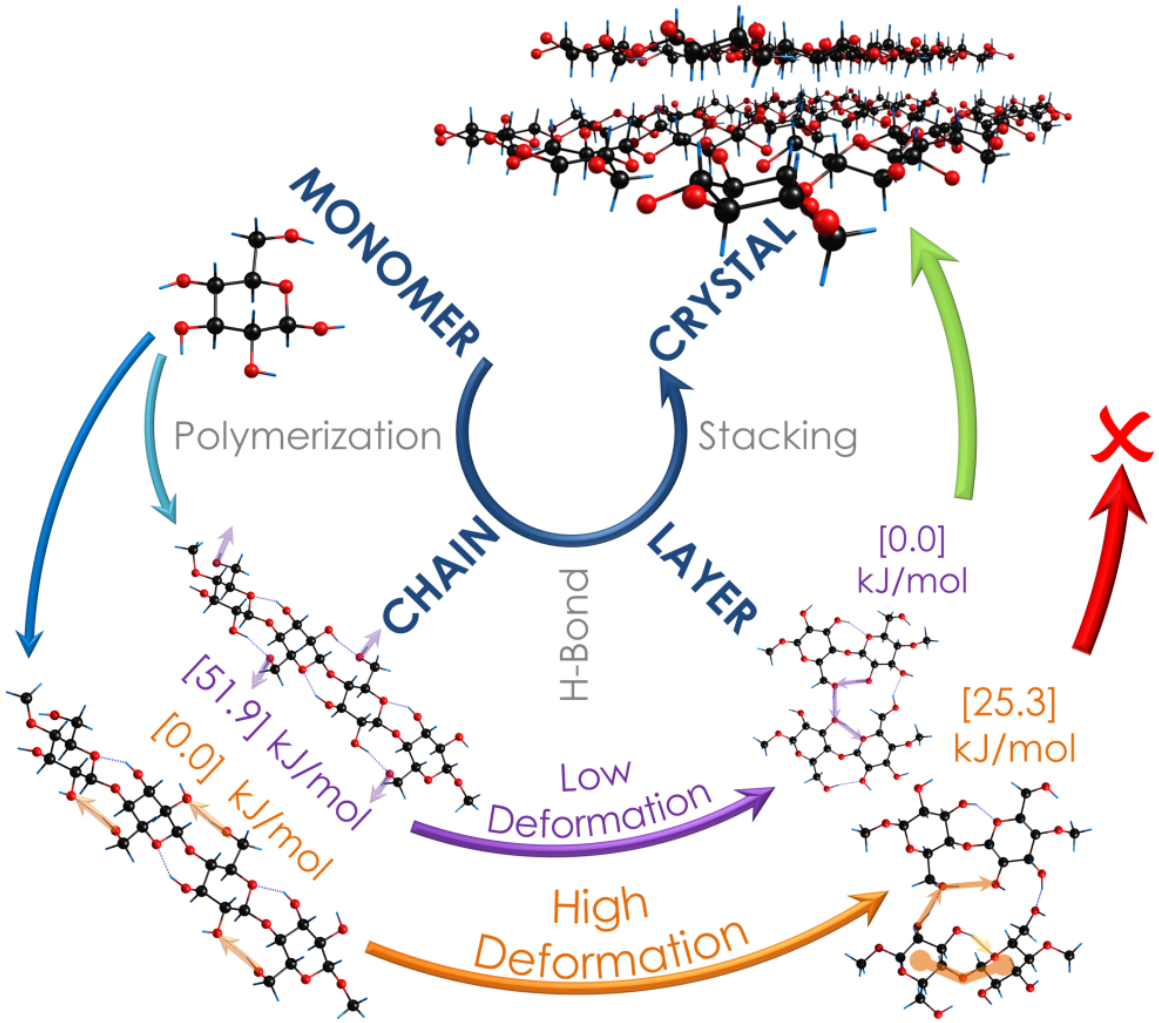
Egitura kristalografiko askoren artean[14,15,16,17,18,19,20,21,22,23,24,25] (behatu Figure A4.10) egindako ikerketa zelulosa I egituraren propietate bereizgarriak agertu dira. [26,27,28] Zelulosaren unitate monomerikoak bata bestearekin interakzionatzen dute hidroxilo ekuatorialen bidez, elkarrekintza hau kristalaren plano ekuatorial osoan zehar banatzen da laminak sortuz (xehetasun gehiagorako irudia hau behatu: Figure A4.11). hain zuzen ere, propietate honi lehenago dimeroetan ikusitako elkarrekintza simetrikoa eta bereizten duen egonkortasun handia lotuta dago. Kristalaren kasuan, hidroxiloen kokatze ekuatoriala zelulosa kateak marra zuzenetan egonkortzeko gaitasuna ematen die, hauek euren artean konplexatzen direnean laminak sortzen dituztelarik. Kate zuzenetan kokatzeko gaitasuna da geroko laminak sortzekoa ahalbidetzen diotena eta honek ere maximizatu egiten ditu molekulen arteko hidrogeno zubien kopurua. Kristalean gertatzen dena hobeto ulertzeko, atal bakoitzak duen influentzia ikertzeko, zatitan bananduko da, laburpena 4.4.1 irudian aurkezten da.

Zatirik txikiena monomeroa da eta hau dagoeneko aztertua izan da lehengo atalean. Kateetan egoteak eragiten duten efektua ikusteko zenbait kalkulu egin dira. Apendizeetan dagoen Figure A4.12-an aurkezten dira egindako zenbait froga, bertan ikus daiteke bi kate mota daudela, bata glukosa guztiak alde bererantz begira ditu (Figure A4.12-2): (d) glukosa guztiak alde negatibora tolestuta, (f) alde positiborantz tolestuta dituena eta (e) aurreko bien nahasketa. Gerta daiteke zelulosa katea tolestuta duten egitura kristalinoren bat aurkitzea baina egitura hori ez du plano ekuatorialen berezitasunik erakutsiko, beraz ez ditu lehen aipatutako interakzio intermolekular kopuru handirik izango. Gainera ikerketaren gune den kristalaren egiturarekin ez doaz bat, horrexengatik alde batera utziko dira. Beste hiru egitura lineal kalkulatuak dira (Figure A4.12-1) hauek monomeroak bata aurrekoarekiko 180º biratuta ditu kristalean ikus daitekeen moduan. Egitura hauei begiratuta (c) letrarekin markatutakoa hiru arrazoiengatik alde batera utzi behar da: lehenengoa, bere egonkortasun erlatiboa oso baxua da, bigarrena, ez du behar beste zuzentasunarik erakusten eta hirugarrena, bere aldarapen esterikoa, kristalean kokatuta egongo balitz, kateen hidrogenoak (HO6) beste kateen hidrogenoarekin batera (HO3) leku bera okupatuko lukete, eta hau ezinezkoa da. Gehienez egitura hau uretan edo beste eratako lotura egiten aurkitu liteke.

Egindako selekzioa bi hautagairekin husten gaitu: Egonkorrena, hidrogeno intramolekular mordoa eratzen ari dena (Figure A4.12 (a) atala eta laranjaz adierazitako katea 4.4.1 irudian) eta beste egitura bat, hau ez da hain egonkorra baina OH askeak ditu gero beste kateekin lotzeko (Figure A4.12 (b) atala eta morez adierazita katea 4.4.1 irudian).

Hurrengo urratsa laminak nola eratzen diren ulertzean datza, honetarako kristalean dagoen kateen arteko elkarrekintza irudikatzeko beharrezkoa den zati txikiena hartu da. Hiru kokatze posible aurkitu dira, bi aurreko kateekin erlazioa dutenak eta beste berri bat. Lehenengo eta behin energien optimizazio kalkulua (M062X/6-31+G(d)) egin da atomo guztiak kristalaren geometrian finko utziz, hidroxiloen hidrogenoak izan ezik (Figure A4.13 (1) atala). Koloreak bat datoz kateen koloreekin, eta hauekin konparatuta, lehen egonkorrena (laranja) zena, orain alderantziz da, eta morezadierazitako kasua da energi minimoa erakusten duena. Honen ondoren kalkuluak berregin dira, baina kasu honetan atomo guztiak aske utziz, optimizazioaren ondoren egiturak konparatu dira, helburua zenbat aldatu diren ikustea da. Zenbat eta gehiago aldatu orduan eta urrunago egongo dira minimoaren egituratik eta probabilitateak txikiagoak izango dira egitura hau

kristalak erakusten duen izateko. Azkenean, aldaketa gutxien erakusten duena, kristaleko benetako egitura izango da. (Figure A4.13 (2) atala)



Irudia 4.4.1: Zelulosa I^{26} motako kristalaren egitura zatitan bananduta. Kalkuluen metodoa eta mahaia: M062X/6-31+G(d).

Bigarrena aukera onena dela nahiko argi bereizten da, goiko irudian more kolorez adierazitakoa, honek oso ondo erreproduzitzen du kristala bere energia minimoarekin. Gero geruzen arteko elkarrekintza dator, honek bi ezaugarri mugatzen ditu: lehenengoa aldarapen indarrak minimizatzeko, oxigenoaren elektroi bikote bakartiak urruntzen dituena (O···O) baita CHak urruntzen dituenak (CH···HC) eta bigarrena, erakargarriak diren elkarrekintzak CH eta bikote bakartien artean (CH···bikote bakarti) hidrogeno lotura ahulak izenarekin ere ezagunak.

Azkenengoz aipatzea benetan zelulosan oinarritutako kristal asko daudela (zelulosa I α edo β) laminak era askotan pilatu ditzaketenak,[26] baina guztiak azaltzea izugarrizko lana da eta kapitulu honen ildotik kanpo dago. Izan ere, kapitulu honen helburua zenbait propietate zehatzak ulertzean eta azaltzean datza.

4.5. Ondorioak

CI-fasean, datu kristalografikoen eta mekanika kuantikoaren erabilera sinergiakoa β-D-glukopiranosa propietateak isolatu eta ulertzeko atea ireki du, berez hiru propietate nagusi ikusi dira: (i) glukosaren egitura egonkorrenak aurkitu dira (GA, GA eta GG) bere PES diagraman kokatuz (ii) hidroxiloen jokabidea deskribatu da lehenago, hauek bakoitza bere aldetik mugituz eta zenbait egitura egonkor emanez (bitarteko egiturak β-Glc-05,06,07 eta 08)eta, lehentasuna erakutzi duen jokabidea, guztiak batera mugitzen direnean bateratutako mekanismoaren bidez (bat-bateko geometria aldaketa 03 eta 04 egituretatik 09 eta10 egituretara (*iii*) Bigarren mailako indarren (CH eta bikote bakartien erakarpen/aldarapen indarren) influentzia. Baina efekturik garrantzitsuena hidroxiloen kokapen ekuatorialean agertu da, honek egonkortze handiago bat lortzea ahalbidetzen du beste egituretatik bereiziz. Egitura kristalografikoari dagokionez hidroxilo ekuatorialek zuzen egotea eta planaritatea ahalbidetzen dute, honek kateen arteko interakzioa intermolekularren zenbakia maximizatzen du eta pilaketa efizienteago honi esker dentsitate handiagoko polimeroak sortzen ditu.[29,30] Propietate hau dimeroetan ere ikusgai dago: ekuatorialki kokatutako hidroxiloak lau hidrogeno lotura indartsu eskuratzea onartzen du. Ikerketa honen bidez berezia den egonkortze indarra aurkitu da β-D-glukopiranosa epimeroan, ulergarria da gaitasun hau izanda naturan duen funtzio garrantzitsua ulertzea, egitura sendoak eratzen baititu.

4.6. Erreferentziak

(1) Carpita, N. C.; Gibeaut, D. M. Structural Models of Primary Cell Walls in Flowering Plants: Consistency of Molecular Structure with the Physical Properties of the Walls during Growth. *Plant J.* **1993**, *3* (1), 1–30.

(2) Wang, a H.; Nathans, J.; van der Marel, G.; van Boom, J. H.; Rich, a. Molecular Structure of a Double Helical DNA Fragment Intercalator Complex between Deoxy CpG and a Terpyridine Platinum Compound. *Nature* **1978**, *276*, 471–474.

(3) Wang, A. H.; Fujii, S.; van Boom, J. H.; van der Marel, G. A.; van Boeckel, S. A.; Rich, A. Molecular Structure of r(GCG)d(TATACGC): A DNA--RNA Hybrid Helix Joined to Double Helical DNA. *Nature* **1982**, *299* (5884), 601–4.

(4) Varki, A. Biological Roles of Oligosaccharides: All of the Theories Are Correct. *Glycobiology* **1993**, *3* (2), 97–130.

(5) Dwek, R. A. Glycobiology: Towards Understanding the Function of Sugars. *Chem. Rev.* **1996**, 96, 683–720.

(6) *CARBOHYDRATE RECOGNITION Biological Problems, Methods, and Applications*; Boons, B. W. and G.-J., Ed.; WILEY, 2001.
(7) Marquardt, R.; Quack, M. Handbook of High-Resolution Spectroscopy; 2011; Vol. 1, pp 539–549.
(8) León, I.; Millán, J.; Cocinero, E. J.; Lesarri, A.; Castaño, F.; Fernández, J. A. Mimicking Anaesthetic-Receptor Interaction: A Combined Spectroscopic and Computational Study of Propofol···phenol. *Phys. Chem. Chem. Phys.* **2012**, *14* (25), 8956–8963.
(9) Leon, I.; Cocinero, E. J.; Millán, J.; Jaeqx, S.; Rijs, A. M.; Lesarri, A.; Castaño, F.; Fernández, J. A. Exploring Microsolvation of the Anesthetic Propofol. *Phys. Chem. Chem. Phys.* **2012**, *14* (13), 4398–4409.
(10) Calabrese C., Maris A., Marcelino N., Vigorito A., M. S. High Resolution Free Jet Millimeter Wave Absorption Spectroscopy: A Bridge to Astrochemistry. In *The 24th International Conference on High Resolution Molecular Spectroscopy PRAHA*; 2016.
(11) Alonso, E. R.; Pen, I.; Cabezas, C.; Alonso, J. L. Structural Expression of Exo-Anomeric Effect. *J. Phys. Chem. Lett.* **2016**, *7*, 845–850.
(12) Talbot, F. O.; Simons, J. P. Sugars in the Gas Phase: The Spectroscopy and Structure of Jet-Cooled Phenyl β-D-Glucopyranoside. *Phys. Chem. Chem. Phys.* **2002**, *4* (15), 3562–3565.
(13) Nibbering, E.; Dreyer, J.; Kühn, O. Chapter 7. Vibrational Dynamics of Hydrogen Bonds. *Anal. Control Ultrafast Photoinduced React. Chem. Phys.* **2007**, 619–687.
(14) Brown, G. M.; Levy, H. A. A-D-Glucose : Precise Determination of Crystal and Molecular Structure by Neutron-Diffraction Analysis. *Science (80-.).* **1995**, *147* (9), 1038–1039.
(15) Jefrey, S. S. C. C. G. A. The Refinament of the Crystal Structures of B-D-Glucose and Cellobiose. *Acta Crystallogr. Sect. B* **1967**, *B24*, 830–838.
(16) Ferrier, W. G. The Crystal and Molecular Structure of B-D-Glucose. *Acta Crystallogr.* **1963**, *16*, 1023.
(17) Kimura, F.; Oshima, W.; Matsumoto, H.; Uekusa, H.; Aburaya, K.; Maeyama, M.; Kimura, T. Single Crystal Structure Analysis via Magnetically Oriented Microcrystal Arrays. *CrystEngComm* **2014**, *16* (29), 6630.
(18) Prusov, A. N.; Prusova, S. M.; Zakharov, A. G. Interaction of Cellulose and Lignocellulosic Polymers with Water and Aqueous Systems *. *Russ. Chem. Bull. Int. Ed.* **2014**, *63* (9), 1926–1945.
(19) Melekh, N. V; Frolova, S. V; Aleshina, L. A. POLYMERS X Ray Analysis of Powdered Celluloses Obtained with the Use of Lewis Acids. **2014**, *56* (2), 134–141.
(20) Moncrief, J. W.; Sims, S. P. Absolute Configuration Determination Using the Anomalous Scattering of Cu-K, X-Rays by Oxygen Atoms: Cellobiose. *Chem. Commun.* **1969**, 914–915.
(21) Kalenius, E.; Kekäläinen, T.; Neitola, R.; Beyeh, K.; Rissanen, K.; Vainiotalo, P. Size- and Structure-Selective Noncovalent Recognition of Saccharides by Tetraethyl and Tetraphenyl Resorcinarenes in the Gas Phase. *Chem. - A Eur. J.* **2008**, *14* (17), 5220–5228.
(22) Popov, D.; Burghammer, M.; Bule, A.; Montesanti, N.; Putaux, J. L.; Riekel, C.; Synchrotron, E.; Facility, R.; Cedex, G. A-Amylose Single Crystals: Unit Cell Refinement from Synchrotron Radiation Microdiffraction Data. **2006**, 3704–3706.
(23) Popov, D.; Bule, A.; Burghammer, M.; Chanzy, H.; Montesanti, N.; Putaux, J. Crystal Structure of A-Amylose: A Revisit from Synchrotron Microdiffraction Analysis of Single Crystals. *Macromolecules* **2009**, *42* (4), 1167–1174.
(24) Nishiyama, Y.; Mazeau, K.; Morin, M.; Cardoso, M. B.; Chanzy, H.; Putaux, J. Molecular and Crystal Structure of 7-Fold V-Amylose Complexed with 2-Propanol. *Macromolecules* **2010**, *43* (20), 8628–8636.
(25) Sarko, A.; Biloski, A. CRYSTAL STRUCTURE OF THE KOH-AMYLOSE COMPLEX*. *Carbohydr. Res.* **1980**, *79*, 11–21.
(26) O'Sullivan, A. C. Cellulose: The Structure Slowly Unravels. *Cellulose* **1997**, *4* (3), 173–207.
(27) VanderHart, D. L.; Atalla, R. H. Native Cellulose: A Composite of Two Distinct Crystalline Forms. *Science (80-.).* **1984**, *223*, 283–284.
(28) Nishiyama, Y.; Langan, P.; Chanzy, H. Crystal Structure and Hydrogen-Bonding System in Cellulose I from Synchrotron X-Ray and Neutron Fiber Diffraction. **2002**, 9074–9082.
(29) Sun, C. True Density of Microcrystalline Cellulose. *J. Pharm. Sci.* **2005**, *94* (10), 2132–2134.
(30) Sarko, A.; Wu, H.-C. H. The Crystal Structures of A-, B- and C-Polymorphs of Amylose and Starch. *Starch/Stärke* **1978**, *30* (3), 73–78.

7. Ondorioak

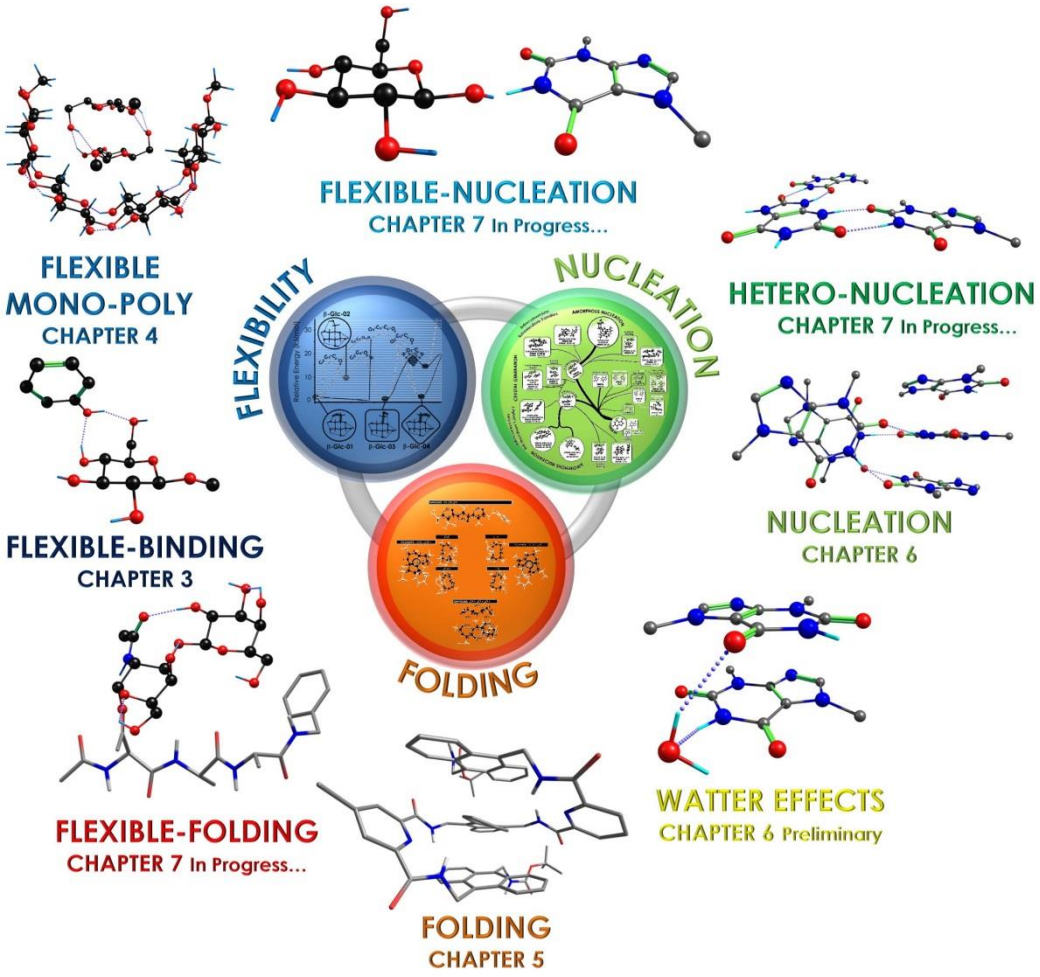
7.7. Ondorioak

Esperimentuen diseinu adimentsua propietate kimikoetan eta hauek arautzen dituzten elementuen erabileran oinarritzen da. Molekulen ezaugarri kimikoen ezagutza oso tresna indartsua da gero oreka kimikoetan aldaketak egin nahi badira, baita aplikazio edo ideia berriak izateko.[1]Ideiak proiektuetan bihurtzen dira eta proiektu hauek ahalik eta era egokienean aurrera eramateko oinarrien ezagutza guztiz beharrezkoa da, azkenengo helburua molekulak nola eragingo duten jakitea izanik, adibidez gaisotasun bat sendatzeko molekula aktibo berria sortuz. [2]

Tesi honetan zenbait sistemen egitura eta elkarrekintzak ikertu dira, prozesu hauetan arautzen duten indarrak determinatuz eta mekanismoak deskribatuz. Ikerketa hau zenbait sistema molekular garrantzitsuen eta erreakzio biologiko esanguratsuekin zerikusia duten beste zenbait sistemen ulermenera zuzendu da. Lortutako emaitza esperimentalen kantitate handia kalkulu teorikoekin sinergikoki nahasten dira, emaitza esperimentalen ulermen soiletik harago dagoen konprenimendura eramanez, sistemen propietate kimikoen ikuspuntu orokor bat eratuz. Azkenean bibliografian eta datu baseetan dagoen informazioarekin alderatzen da emaitzen fidagarritasuna ebaluatuz eta ikerketak ematen duen ikuspuntu orokorra are gehiago handituz.

Emaitza asko lortu direnez hauen ulermena errazteko multzotan banandu dira. 7.7 irudiaren erdian ikusgai dauden taldeetan sailkatuko dira sistemak, eta esanguratsuenak hauek dira: Sistema Malguak edo flexibleak, foldameroak eta nukleazio prozesuak. Nahiz eta sailkapen hau egin, ez da ahaztu behar guztiak molekulez osaturik daudela eta taldeen artean erlazio asko/ugari dagoela. "*In Progress*…" hitzekin markatutakoak sailkapen desberdineko molekulak nahasten dituzte kalkuluak asko zailtzen dituelarik. Hauetarako talde desberdinetako metodologiak batu beharko dira eta agian metodologia berriren bat diseinatu.

Lehenengo propietatea molekulen malgutasuna da, molekula hauek ia edozein egoeran bere egitura aldatzeko gaitasuna bereizten dute. Talde honen barruan azukreak ikertzea oso interesgarria da, hauek eginkizun askotan parte hartzen baitute naturan. Garrantzitsuak egiten duten beste ezaugarri bat da, beren egituraren barruan aldakorrak diren atal asko daudela: bost karbonoz eta oxigeno batez osatutako zikloa, zikloari ilaran lotutako hidroxiloak eta hidroximetiloa, azken honek erakusten duen malgutasuna askotan lotura gune nagusian bihurtzen du.



Irudia7.7: Ikertutako elkarrekintzen laburpena

Malgutasun molekularrak molekulen elkarrekintzan duen efektua hirugarren kapituluan isolatzea lortu da. Bertan molekula malguen egonkortzea lekuz aldatzeko gaitasuna bereizi da, zonalde intramolekularretatik molekulen arteko elkarrekintza gunera mugituz eta indar handiagoko loturak eratuz. Bi energia mota alderatu egiten dira: konplexuaren energia erlatiboa eta molekulen arteko lotura indarra. Propietate berri bat isolatu da, energia erlatiboan zenbait minimo lehiaketan dabilenean; hauetako batek besteekin alderatuta lotura energia handiagoa baldin badu honen eraketa sustatuko da besteenaren ordez. Ondorio moduan esan daiteke kasu batzuetan lotura indarra arautzen den indarra dela.

Ingurugiro makroskopikoan asko dira honen egonkortzean parte hartzen duten propietateak. Oso interesgarria da mundu mikroskopikotik ezaugarri bat hartzea eta mundu makroskopikoan ekarpenik duen egiaztatzea, honekin batera nola ematen den ulertzea garrantzitsua da. Laugarren kapituluan honelako ikerketa aurrera eraman da, β-glukosaren hidroxilo guztiak ekuatorial eran kokatuta

daudela ikusi eta dimeroan eta zelulosa I kristalean egonkortzean duen ekarpena determinatu da. Ondorio moduan plano ekuatorialak β-β dimeroan eta zelulosa kate eta lamina zuzenetan eragiten duela esan daiteke. Ikerketa honen bidez ulertu daiteke zergatik naturak glukosa hautatu duen funtzio estrukturala betetzeko eta ez beste piranosaren bat.

Sistema flexibleek eta foldameroek oso aberatsa den egituren panorama erakusten dute, eta egituren artean eraldatzeko gaitasuna sortutako lotura intramolekularrez mugatuta dago. Foldameroen kasuan, egitura posibleetan aberatsenak dira, horregatik ze egitura hartuko duten jakitea oso erronka zaila da. Molekula erraldoi hauetako toleste prozesuan efektu askoren lehiaketa ematen da, eta bosgarren kapituluan foldamero baten zatitxo baten ikerketa egiten da. Lortutako datuek ez dute disoluzioan lortu litekeen egituraren ereduari erantzungo, disoluzioan ingurugiroa guztiz desberdina izanda egitura oso desberdinak sortuko dira eta. Baina teknika honen bitartez, peptido txiki honek nahiago duen egitura lortuko du kanpoko efektuen inolako influentziarik gabe. Modu berean egitura hau ingurumen hidrofobo batean hartu dezaken egituraren antzekoa izan daiteke. Beste makromolekula bat ere ikertu da, antrazeno unitateetan oinarritutako polimero bat. Makromolekula honetan bere egiturak elektrizitatea garraiatzeko gaitasuna ematen dio, eta egindako neurketen bidez egiaztatu egin da nola antrazeno unitateak bata bestearen gainean pilatzen direla, grafitoan gertatzen den bezala, π laino luze bat sortuz eta eroalea bihurtuz. Azken neurketa hau materiak berri batean izan da eta honen bidez bere propietateak nondik datozen jakitea lortu da, honen ondorioz garrantzi handiko artikulua argitaratuz.

Nukleazio edo agregazioen prozesuak ikertzeko espantsioan konplexu batzuk sortu dira. Bi molekula edo gehiagoren arteko elkarrekintza azaltzea ez da gauza makala, egitura mota asko eratu dezakete eta. Sistema hauekin hasteko garrantzitsua da zenbait sinplifikazio egitea, adibidez molekula zurrunak erabiltzea, honek monomeroan egitura asko izatearen ondorioz egin beharreko kalkulu guztiak aurrezten baititu. Seigarren kapituluan molekula zurrun baten nukleazio prozesuaren lehenengo momentuetan sorturiko konplexuak isolatu dira. Hautatutako molekula teobromina da, honek base purikoen deribatua da baina timina base pirimidinikoaren talde funtzionala du, horregatik gorputzean zenbait efektu eragiten ditu, lotura biologikoak eratzeko gai delako. Gainera teobromina gizakien gorputzaren barruan zenbait nukleazio prozesuen tartean sartzen dela pentsatzen da, azido urikoaren kristalen eraketan alegia. Kapituluan teobromina batetik hiru teobromineko konplexutara ailegatzen da baina ikerketa osoa bost teobrominetaraino helduko da, baita dimeroan eta trimeroan urak duen

influentzia behatuz. Dena den lortutako emaitzak espantsioan orekatik kanpo dauden egiturak dira, espantsioaren hozketa azkarra amorfoa eratzeko joera nagusitzen da eta. Kalkulu teorikoen bidez zuhaitz genealogiko osoa eraiki da momerotik trimerora (6.4.2 irudia ikusi). Diagrama honek ikuspuntu orokor bat ematen du eta espantsioko egiturak bertan kokatu daitezke. Beste alde batetik datu baseetatik lortutako teobrominaren egitura kristalografikoaren dimeroak eta trimeroak kokatu dira zuhaitzean. Azkenean, amorfotik antolatutako kristalera ailegatzeko, zuhaitzean zehar beharrezkoak diren eraldaketak ikus daitezke, ikuspuntu orokor bat emanez.

Ikerketekin jarraituta beste sistema asko neurtu dira, hauetako asko konplexuagoak dira eta adibidez nukleazioa molekula malguekin nahasten dituzte. Azkenean lehen bananduta egin diren sistemak (molekula malguan, foldameroak eta nukleazioa) orain nahastu egingo dira eta hauek tratatzeko metodologia teoriko berriak egin/asmatu beharko dira. Kalkuluak oraindik ordenagailuan martxan daude baina bukatuta daudenean hauek sailkatzeko eta prozesatzeko metodologia oso astuna eta neketsua ideiatu beharko da.

Tesi honen helburua ez da propietate kimikoen determinazioan soilik mugatzen, lagungarria izan daiteken erreferentzia bat izatea ere bilatzen du, zenbait adibideren tratamendua eginez (glukosa, foldameroak eta base purikoen deribatuak) eta metodologia hoberena azalduz.

Orokorrean oinarrizko ikerketak aplikatzetik oso urrun daudela pentsatzen da. Tesi honetan egindako irudietan, adituentzako diren taula eta diagramak laburbildu eta sinplifikatu dira, hauen ulermena ahalik eta zientzialari gehiengoari hurbiltzeko. Benetako asmoa da etorkizunean lan hau beste lanetarako inspirazioa eta esperimentu adimentsuagoak eratzeko balio dezala.

7.8. Erreferentziak

(1) De Vries, M. S.; Hobza, P. Gas-Phase Spectroscopy of Biomolecular Building Blocks. *Annu. Rev. Phys. Chem* **2007**, *58*, 585–612.

(2) Byrn, S.; Pfeiffer, R.; Ganey, M.; Hoiberg, C.; Poochikian, G. Pharmaceutical Solids: A Strategic Approach to Regulatory Considerations. *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists*. 1995, pp 945–954.

INDEX

CHAPTER 1: Introduction

CHAPTER 2: Experimental Setup and Methodology

FLEXIBILITY

CHAPTER 3: Intermolecular Interactions in Glucose·Phenol Complexes: the "Proven Survivors"

CHAPTER 4: Equatorial β-D-glucopyranose: Structural Properties in One of the Most Abundant Carbohydrates

FOLDAMERS

CHAPTER 5: Analysis of Folding Structures

NUCLEATION

CHAPTER 6: Analysis of the Nucleation Process in Theobromine

CHAPTER 7: Additional Experimental Results and Conclusions

1. Introduction & Objectives

1.1. Introduction

The isolation and characterization of chemical properties represent the starting point to understand fundamental processes in Nature.[1] Such knowledge is compulsory in most cases to produce ground-breaking technological progresses.

Among the chemical properties, the intra- and inter-molecular interactions represent one of the main forces involved in the stabilization of molecular systems. In particular, non-covalent interactions are ubiquitous in Nature and govern fundamental processes that shape the world and even life as we know it on Earth. From condensation of gases into a liquid phase, to interaction between proteins or their ligands to perform complicated tasks inside a cell, to proton transfer… etc, all those processes are controlled by what is usually termed as non-covalent interactions.[2] They are of quantum chemical nature and have their origin in the interaction between the electronic clouds of neighbour molecules or between the electronic cloud and point charges. They are usually classified into electrostatic, charge transfer, dispersion and exchange-repulsion. Hydrogen bond deserves special mention, because it has some covalent character, due to the overlapping between the orbitals of the interacting molecules and it is usually the strongest intermolecular interaction. However, even in the case of hydrogen bond (H-bond hereafter), the magnitude of these interactions is small, compared with ordinary covalent bonds. Thus, while a typical C-H or C-C bond present interaction energy values of ~ 435 and 368 kJ/mol respectively,[3] the H-bond between two isolated water molecules is only ~ 17 kJ/mol, as computed at M06-2X/6-311++G(d,p) level.[4] Thus, the strength of the non-covalent interactions is around a 5 % of that of a true covalent bond. However, they are relatively long-range forces and usually present additive properties. As a result, their collective effect goes beyond the relatively small importance of the individual interactions.

Returning to the nature of the non-covalent interactions, their small module is a major problem for their characterization. Any perturbation will result in an incorrect measurement of their properties and especially, of the effects they produce in the structure of a given molecule or a molecular aggregate. Thus, the information that can be obtained from each system, is usually limited by the technique used. In many cases, the characteristics and limitations of the technique interfere so strongly with the measures that the observations cannot be separated from the perturbation, and the information extracted is of very limited use. Certainly, by definition, the act of measuring itself always introduces a perturbation on the system.

On the other hand, each technique is better suited to measure some specific aspects of the non-covalent interactions, in a given environment. Thus, to completely characterize a molecular system, several techniques are usually required, to obtain data on the three most relevant aspects: the structure of the system, the strength of the interactions and the dynamics of spontaneous processes, such as isomerization or proton transfer.

Probably, the structure of the molecular systems is the most vastly explored aspect and this is reflected in the wide variety of techniques that aim at structural determination. The structure of a molecular system is directly connected to the chemical properties exhibited by a system, from the simplest monomeric unit, to the most complicated biochemical equilibrium. In particular, the studies of molecular structures and intermolecular interactions are used to determine many chemical properties. There are many examples that connect the structure/interactions with the properties: e.g. in polymers the structure of the chains and the interactions between them conferee to the material its elasticity, strength, etc.[5] In graphite, the stacked layers produce a π-cloud along the entire sheet making it electrically conductive.[6] As regards the biological environment, the level of complexity is orders of magnitude higher, in part as a consequence of the heterogeneity of the functions fulfilled by the biochemical systems: e.g. docking centers,[7] structural supports,[8,9] storage and transfer of the information like in mRNA and DNA, molecular recognition,[10] etc.

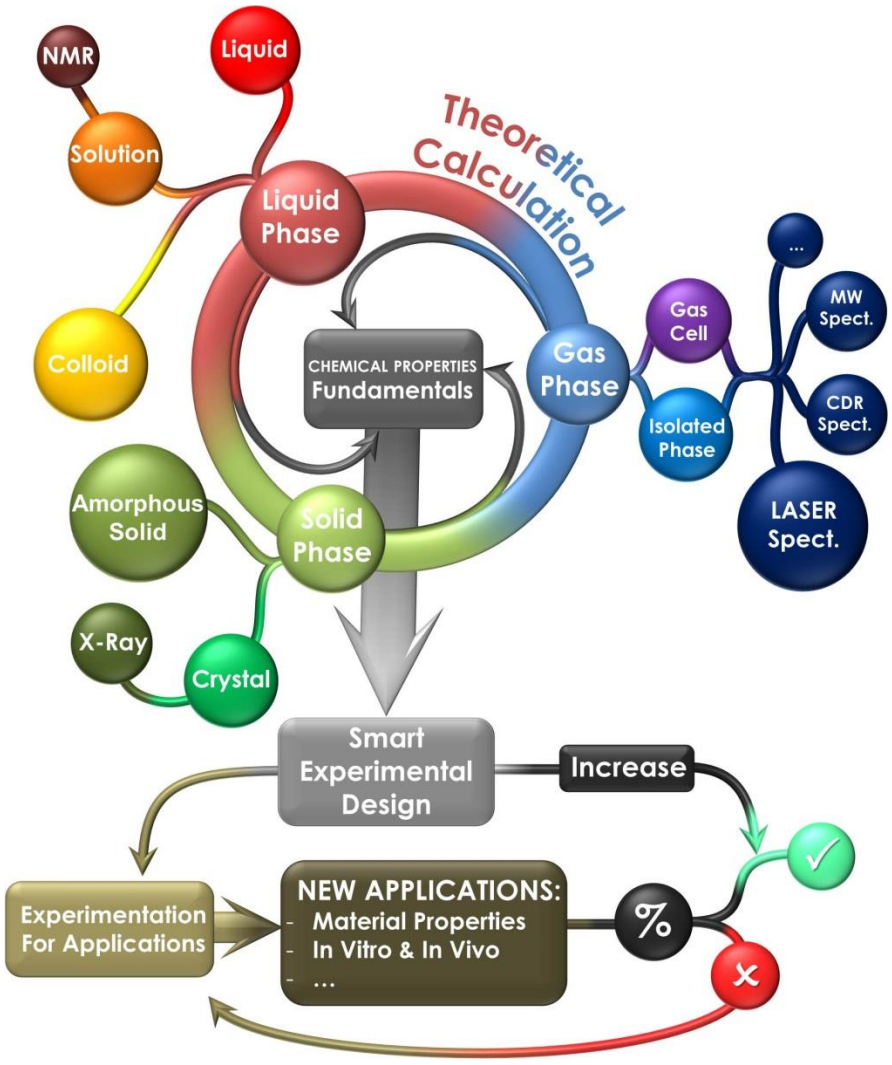
As aforementioned, there are many techniques aimed at structural determination. They can be classified attending to the phase in which they are applied: solid, liquid or gas phase. In the solid phase, crystallography is the dominant technique; in liquid phase, NMR is one of the most commonly used, and in the gas phase many different spectroscopic techniques have been described.[11] Despite the unusual environment that the molecules encounter in gas phase, such studies present unquestionable advantages, especially when supersonic expansions are employed. Indeed, supersonic expansion are used as a way to cool the molecules until most of the population lies in the ro-vibrational (and, of course, electronic) ground state: the collisions that form the expansion transform the internal energy of the molecules in translational energy, favoring at the same time aggregation processes and creating a cold-isolated (CI) phase.[12] Once the expansion is completed, the species travel isolated, without interactions or interferences from a solvent, permitting the study and characterization of the native aggregation and conformation preferences of the systems. The use of supersonic expansions was a

true revolution in the spectroscopic study of complex systems: the cooling not only simplified the spectra by elimination of hot bands, but also resulted in increased signal intensity,[13,14] opening the door to the development of sophisticated techniques combining several lasers in different spectral regions[15,16] that allowed extracting precise information from the systems, or even the use of microwave spectroscopy, able to determine the molecular structure with unprecedented resolution.[17,18] Nevertheless, each experimental technique in each phase presents its own limitations: for example, the crystallographic studies are limited by the modulation introduced by the matrix, which can interfere in the structure determination of isomers, and by the interactions between the monomeric units for which the symmetry and the periodicity of the crystal have to be fitted.[19] In solution, NMR represents probably the most powerful technique, able to reveal detailed structural information about the interaction sites,[20,21] by exploiting the coupling between carbons, hydrogens and any other atom with the appropriate multiplicity. NMR analysis yields structural information on the interacting parts of a molecular system, and produce dynamic but averaged structures, concomitantly with the nature of solution.

In this thesis project, the application of supersonic expansion coupled with several mass-resolved laser-based spectroscopic techniques was extensively used to record the vibronic spectra of different types of selected systems. In particular, the supersonic expansion takes place inside the ionization chamber of a time-of-flight mass spectrometer where the sample is ionized resonantly using multi-photon ionization (MPI). The system is completed with a laser vaporization device. Certainly, most of the molecules studied in this work, are thermolabile, and therefore cannot be transferred into gas phase by simple warming, and a more sophisticated methodology was required, as described in Chapter 2. This technique presents many advantages with respect to other spectroscopic methodologies[15] and the combination with the mass spectrometer allows one to record selective information at different aggregation levels.[22] Especially, the ability of unravelling intra- and inter-molecular non-covalent interactions which rule the structural arrangements of molecules, permits to deepen in the knowledge of the fundamental principles that control a system, and set the grounds to learn how to modulate the molecular systems in a more selective way.

As stated above, thanks to the simplification in the experimental spectra introduced by the use of jets, it is possible to tackle nowadays the spectroscopy of very large molecular systems. However, interpretation of the spectra of those systems is by no means univocal, due to their intrinsic complexity and they may be

interpreted in the light of massive calculations, usually conducted at several levels: from the fast but inaccurate molecular mechanics that enables a fast exploration of the potential energy surfaces, to the slow but very accurate quantum mechanic calculations that produce results that may be directly compared with the experimental data. Furthermore, the calculations, once validated by comparison with the experimental results, may be used to build bridges between the CI-phase and the crystals or even solution, as it will be demonstrated along this work.



Scheme 1.1: Connection of the fundamental chemical properties with the experimental techniques used in their characterization. These results, obtained from basic science, like in this thesis, can set the grounds on which to base new application in the development of a smart experimental design.

The approach used during this thesis follows Scheme 1.1, where the CI-phase data are produced and compared to those from solid and liquid phases, already in the

literature. Moreover, the scheme shows how the experimental data coming from different sources can be connected through theoretical calculations. The main goal is to achieve a complete and deep understanding of a molecular system by connecting synergistically all these kinds of analysis. Achieving a deep understanding of the fundamentals that govern the chemical and biochemical systems is useful to generate basic knowledge in very diverse scenarios, which are part of the application of smart experimental design.

In general, the smart experimental design refers to how to deal with the generation of any new application, and its completion strongly depends on a deep knowledge of the fundamental principles and procedure on which it is based. One of the most clear examples of this experimental design is the production of a new drug.[23] Generally, the pharmaceutical companies start the process by testing their libraries of compounds against a receptor or protein that wants to be activated (agonist molecules) or deactivated (antagonist molecules). This implies testing thousands or even hundreds of thousands of compounds in a semi-automated way, but commonly using radioactive isotopes to track the pharmacokinetics of the interaction. This process is extremely long and costly, as the chemicals used are very expensive and produces a large amount of radioactive waste. The next step is even more expensive and even morally questionable, as it involves the "*in vivo*" tests of the candidate molecules. In this example, the smart experimental design would involve a deep knowledge on the non-covalent interactions that govern the docking of the drug on the protein. Certainly, considerations such as the existence of hydrophilic or lipophilic groups in the active centre of the protein, among others, may allow one to design a reduced number of molecules (or families of molecules) that may fit in such active centre.

Therefore, a basic knowledge, in principle away from any practical application, is easily connected to the technological advancement that in turn can produce a measurable economic impact. Furthermore, the advantage of the basic knowledge is that any progress achieved may affect other fields and influence implications that are still unknown: the same knowledge on non-covalent interactions that may help in speeding the development of new drugs, can serve to design new polymeric materials, as the same non-covalent interactions are involved in the structural arrangement of polymeric chains or it may help to predict the 3D structure of a protein.

The proposed scheme does not mean that the smart experimental design is the only way to get new advances (e.g. serendipity), but in many cases it can strongly

increase the likelihood of success, underlying the importance of the generation of basic knowledge.

1.2. Objectives

The main objective of this thesis is to analyze and characterize different types of relevant molecular systems in order to outline and improve new approaches for the identification of particular chemical properties, which characterizes their behavior.

The investigated systems can be divided in three classes: flexible molecules, able to easily modify their structure (Chapter 3 and 4); foldamers, characterized by a very complex conformational landscape but their interconversion is limited by the barriers to abandon a previously adopted folded structure (Chapter 5), and rigid molecules, for which the variability comes from chemical reactions like tautomerization (Chapter 6). For each class, several new computational strategies were developed in order to fully characterize the selected molecular systems. Each of the chapter of this thesis does not only describe the experiments performed for the characterization of the particular system, but also addresses a specific biochemical or physical-chemical question.

In particular, the main objective of chapter 3 is to determine which is the energy calculation that fits better with the experimental results (relative energy VS relative binding energy), making use of different glucose-phenol clusters. In chapter 4, the study is centered on the properties of β-glucose: this is the only pyranose carbohydrate with all the hydroxyl groups in the equatorial plane, and at the same time, it is one of the most abundant sugars in nature. The investigation aimed at analyzing its molecular properties, in order to determine if this all-equatorial arrangement confers to this biomolecule a specific ability. The chapter 5 approaches the analysis of small sections of foldamers (a tripeptide and an oligomer), in order to analyze, in the most accurate way, which are the preferred folded conformations and their implications on their physical-chemical functions. Finally, chapter 6 describes the properties of a rigid molecule: theobromine, and based on the conclusions extracted, the initial stages of its nucleation process are analyzed. Additionally, all the experimental data are compared to crystallographic ones in order to build a genealogical tree (monomer, dimer, trimer) of its nucleation process, by means of theoretical simulations. This kind of analysis aims at better understanding the forces that rule the starting stages of the nucleation process of this molecule.

As aforementioned, the molecular systems were investigated in supersonic expansions, using a combination of mass-resolved laser-based spectroscopic techniques and computational chemistry. As it will be demonstrated, each system presents its own peculiarities that will require of refined methodologies or completely different computational approaches. Moreover, different improvements on the setup were done, in order to tackle the spectroscopy of some of the systems, mainly regarding sample deposition, laser alignment and active subtraction experiments. Thanks to such improvements, we present here the spectra of some of the largest and more complex systems studied up to date.

It is important to mention that the detected chemical properties are summarized and analyzed focusing on the main characteristics of each system, and how they can be further explored in future works. For this reason, a special effort was put in simplifying the results obtained (presenting most of the data as appendix, which consist of large numbers of tables and diagrams), summarizing them through the chapters using condensed figures, so as to make the data as accessible as possible to a broader scientific audience.

1.3. References

(1) De Vries, M. S.; Hobza, P. Gas-Phase Spectroscopy of Biomolecular Building Blocks. *Annu. Rev. Phys. Chem* **2007**, *58*, 585–612.

(2) Hobza, P.; Muller-Dethlefs, K. *Non-Covalent Interactions*; RSC Theoretical and Computational Chemistry Series; The Royal Society of Chemistry, 2009.

(3) Blanksby, S. J.; Ellison, G. B. Bond Dissociation Energies of Organic Molecules. *Acc. Chem. Res.* **2003**, *36* (4), 255–263.

(4) León, I.; Montero, R.; Castaño, F.; Longarte, A.; Fernández, J. A. Mass-Resolved Infrared Spectroscopy of Complexes without Chromophore by Nonresonant Femtosecond Ionization Detection. *J. Phys. Chem. A* **2012**, *116* (25), 6798–6803.

(5) Van Krevelen, D. W.; Te Nijenhuis, K. *Properties of Polymers: Their Correlation with Chemical Structure; Their Numerical Estimation and Prediction from Additive Group Contributions*; Elsevier, 2009.

(6) Sengupta, R.; Bhattacharya, M.; Bandyopadhyay, S.; Bhowmick, A. K. A Review on the Mechanical and Electrical Properties of Graphite and Modified Graphite Reinforced Polymer Composites. *Prog. Polym. Sci.* **2011**, *36* (5), 638–670.

(7) Dominguez, C.; Boelens, R.; Bonvin, A. M. J. J. HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **2003**, *125* (7), 1731–1737.

(8) Harris, H. Actinomycetales and Related Groups. **1958**, No. 1958, 173–189.

(9) Schleifer, K. H.; Kandler, O. Peptidoglycan Types of Bacterial Cell Walls and Their Taxonomic Implications. *Bacteriol. Rev.* **1972**, *36* (4), 407–477.

(10) MacCallum, R. M.; Martin, a C.; Thornton, J. M. Antibody-Antigen Interactions: Contact Analysis and Binding Site Topography. *J. Mol. Biol.* **1996**, *262* (5), 732–745.

(11) Schermann, J.-P. *Spectroscopy and Modeling of Biomolecular Building Blocks*; Elsevier: Amsterdam, 2008.

(12) Scoles, G. Atomic and Molecular Beam Methods. In *Atomic and Molecular Beam Methods*; Oxford University Press, 1988; Vol. 1, pp 14–53.

(13) Oswald, S.; Wallrabe, M.; Suhm, M. A. Cooperativity in Alcohol–Nitrogen Complexes: Understanding Cryomatrices through Slit Jet Expansions. *J. Phys. Chem. A* **2017**, acs.jpca.7b01265.
(14) Gottschalk, H. C.; Altnoder, J.; Heger, M.; Suhm, M. A. Control over the Hydrogen-Bond Docking Site in Anisole by Ring Methylation. *Angew. Chemie - Int. Ed.* **2016**, *55* (5), 1921–1924.
(15) Levy, D. H. Laser Spectroscopy of Cold Gas-Phase Molecules. *Annu. Rev. Phys. Chem* **1980**, *31*, 197–225.
(16) Zwier, T. S. Laser Spectroscopy of Jet-Cooled Biomolecules and Their Water-Containing Clusters: Water Bridges and Molecular Conformation. *J. Phys. Chem. A* **2001**, *105*, 8827–8839.
(17) Balle, T. J.; Flygare, W. H. Fabry-Perot Cavity Pulsed Fourier Transform Microwave Spectrometer with a Pulsed Nozzle Particle Source. *Rev. Sci. Instrum.* **1981**, *52* (1), 33–45.
(18) Pate, B. H.; De Lucia, F. C. Broadband Molecular Rotational Spectroscopy Special Issue. *J. Mol. Spectrosc.* **2012**, *280* (1), 1–2.
(19) Finger, L. W. Physical Properties of Crystals, Their Representation by Tensors and Matrices. *Eos, Trans. Am. Geophys. Union* **1983**, *64* (45), 643.
(20) Kogelberg, H.; Solís, D.; Jiménez-Barbero, J. New Structural Insights into Carbohydrate-Protein Interactions from NMR Spectroscopy. *Curr. Opin. Struct. Biol.* **2003**, *13* (5), 646–653.
(21) Mayer, M.; Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chemie - Int. Ed.* **1999**, *38* (12), 1784–1788.
(22) Speranza, M.; Satta, M.; Piccirillo, S.; Rondino, F.; Paliadini, A.; Giardini, A.; Filippi, A.; Catone, D. Chiral Recognition by Mass-Resolved Laser Spectroscopy. *Mass Spectrom. Rev.* **2005**, *24* (4), 588–610.
(23) Marinelli, L.; Lavecchia, A.; Gottschalk, K.-E.; Novellino, E.; Kessler, H. Docking Studies on avb3 Integrin Ligands: Pharmacophore Refinement and Implications for Drug Design. *J. Med. Chem.* **2003**, *46*, 4393–4404.

2. Experimental Setup and Methodology

2.1. Experimental Setup

In this section the main components of the experimental setup are shown and described. The goal of this chapter is to explain all the details, allowing the reproducibility of the experimental measurements. As already anticipated in the introduction of the thesis, the application of several mass-resolved laser-based spectroscopic techniques, coupled with supersonic expansion, was extensively used to record the vibronic spectra of the different selected systems.[1–3] In Figure A2.1 (A refers to the Appendix), a general scheme of the experimental set up can be found, and in Figure 2.1.1 the sequence of the experiments is depicted and described.

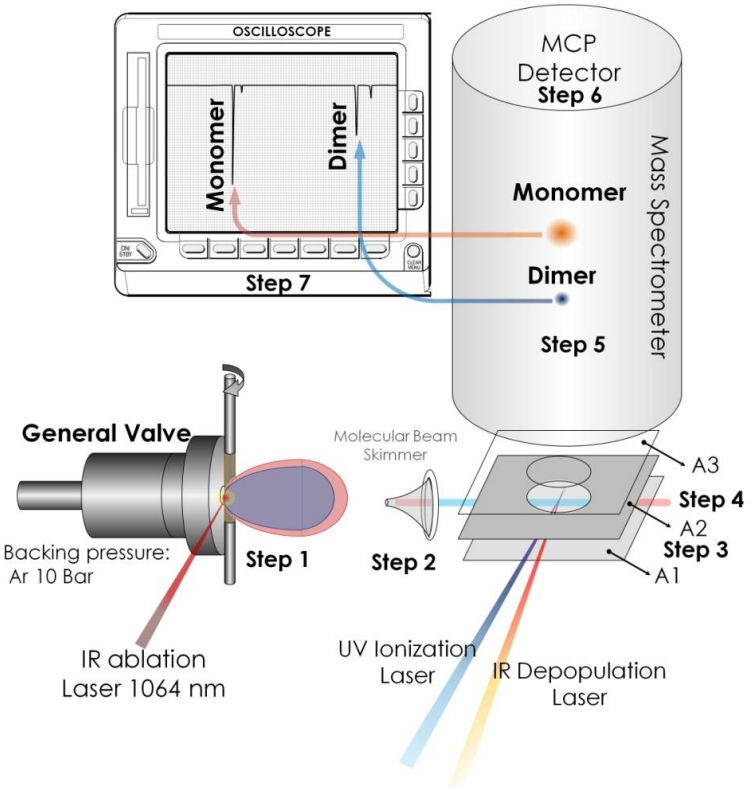


Figure 2.1.1: The experiment consists of a pulsed valve that generates a supersonic jet expansion into a high vacuum chamber, at 10 Hz. A solid sample rod attached to a stepper motor (rotating and translating to constantly refresh the sample) is vaporized by an ablation pulsed-laser (**Step 1**). The desorbed molecules are cooled and transported by the carrier gas (usually argon) to the ionization region of a linear time-of-flight (TOF) mass spectrometer. The molecular beam is skimmed selecting the coldest part (**Step 2**). The ionization tunable-dye laser is synchronized with the arrival of the molecular beam to an inter-plate space, where it excites and ionizes the molecules, which are subsequently extracted by the electric field produced by two plates set at A1= 4000 V and A2 = 3700 V, and moved to a second inter-plate region (Ionization Region, **Step 3**). There, a high voltage difference (A2 = 3700 V, A3 = grounded) accelerates the ions proportional to their mass-to-charge ratio (m/z) (**Step 4**). The ions are separated according to their m/z during the free flight in the mass spectrometer's tube (**Step 5**). Subsequently, the ions hit the detector producing an electric signal at different times (**Step 6**). These electrical signals are monitored by an oscilloscope in order to optimize the state of each of the masses in real time (**Step 7**).

2.1.1. Laser Sources

Since the laser sources are very versatile tools, various are the functions that they have played in the experiments performed in this thesis. In particular, the laser sources were used: (*i*) for the excitation and ionization of molecules or clusters; (*ii*) to depopulate molecular systems in double resonance techniques and (*iii*) in order to generate gases of thermolabile molecules. All these aspects are described more in details in the following sections.

2.1.1.1. Excitation and Ionization Sources (UV tunable and pulsed dye Lasers)

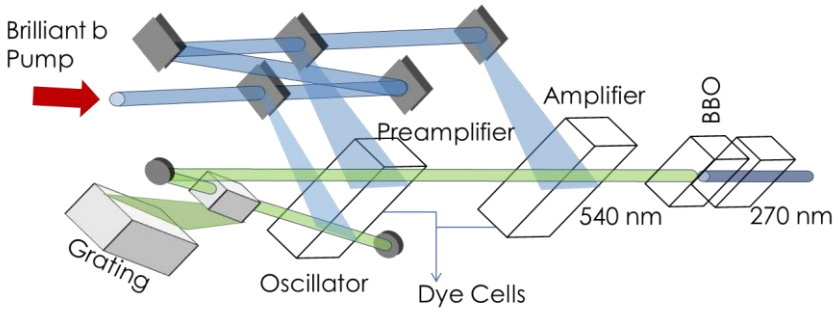


Figure 2.1.1.1: Simplified Optical Layout of a Dye Laser.

Regarding UV tunable laser, different types were used during this work, but all of them are based on the same optical layout. In Figure 2.1.1.1, a general scheme is presented. The UV pulsed laser is reflected on a quartz prism, collimated and focused by convex lenses of f = 1 or 2 m. The choice of the focal length depends on the fluence delivered by the laser source: for low fluence values (~100 μJ), a short focal length lens may be used in order to get the maximum density of photons.

Fine Adjustment tunable UV Dye Laser: Pumped by Nd:YAG laser (Brilliant b; 120 mJ/pulse at 355 nm) with ω2 and ω3 modules, in the range 250-270 nm. The dyes used were: Coumarin 540A and Coumarin 500. The averaged energy produced by the laser in the UV region was ~300 μJ/pulse.

Scan Mate tunable UV Dye Laser: Pumped by Nd:YAG laser (Brilliant b; 350 mJ/pulse at 532 nm) with ω2 module. This laser was used in the range 285-350 nm, employing Rhodamine dyes (575, 590 and 610).

2.1.1.2. Depopulation Lasers (UV/IR Lasers)

Double resonance techniques (UV-UV, Ion-dip) allow for a selective discrimination between different conformations. This kind of experiments require of at least two lasers. One of them must be a UV laser in order to ionize the sample (the details of this technique will be found below). The second laser can be an IR (Ion-dip experiments) or a UV (UV-UV hole-burning experiments) laser. Regarding the IR laser, two different types of lasers were used: an OPO/OPA IR and an IR dye laser.

Laser Vision OPO (Optical Parametric Oscillator)/OPA (Optical Parametric Amplifier) Laser: Pumped by Nd:YAG laser (Continuum Surelite; 800 mJ/pulse at 1064 nm). The laser final power is about 10 mJ/pulse from 2500 to 4000 cm^{-1}. The bandwidth of the laser is ~10 cm^{-1}. The operation system is described in the figure 2.1.1.2.1.

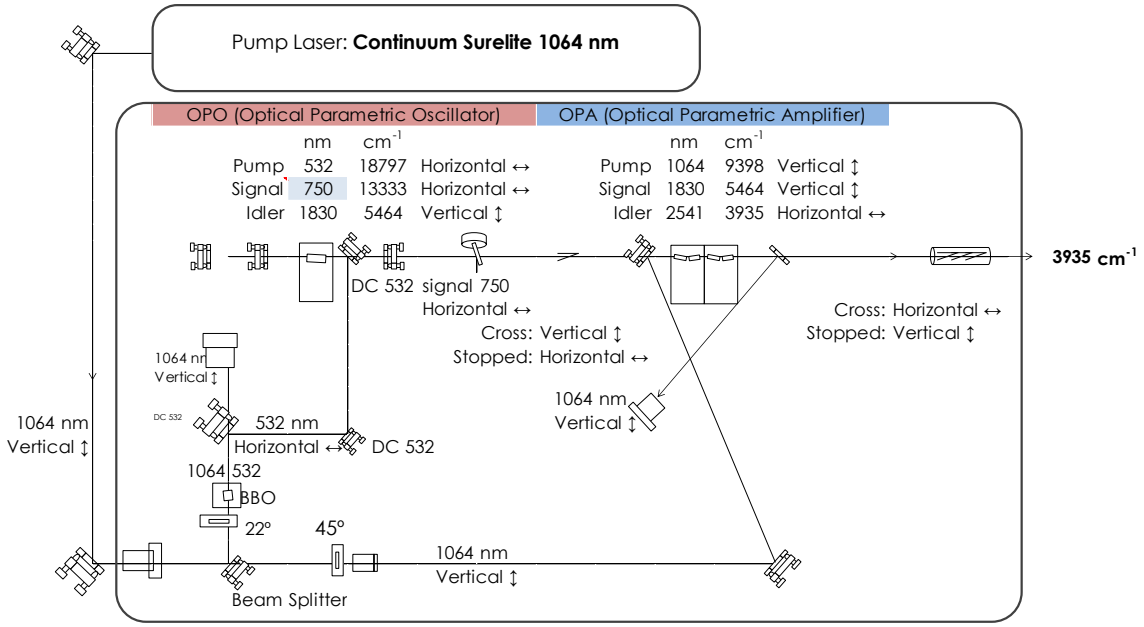


Figure 2.1.1.2.1: Laser Vision OPO/OPA Scheme

Fine Adjustment tunable IR Dye Laser: Pumped by Nd:YAG laser (Innolas SpitLight 1200; 600 mJ/pulse at 532 nm). The design of this laser is very similar to the Fine Adjustment UV described before. The main difference is that, after the second dye cell, instead of the BBO doubling crystal, it uses a LiNbO4 crystal (see Figure

2.1.1.2.2). To generate the IR radiation, a mixture of LDS dyes was used, generating radiation in the 3100 and 3800 cm^{-1} region.

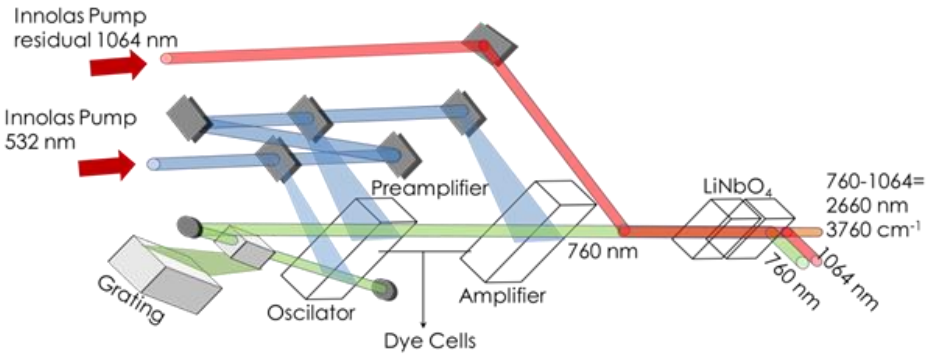


Figure 2.1.1.2.2: Simplified Optical Layout of Dye Laser.

2.1.1.3. Ablation Laser (Nd:YAG Lasers)

A Nd:YAG laser was used to vaporize samples, avoiding the thermal decomposition promoted by conventional heating methods, which has prevented the study of amino acids and carbohydrates in the past. The operating power values are around 1 mJ and the laser employed was the Quantel Ultra with a maximum energy of 30 mJ/pulse at 1064 nm (see Photo 2.1.1.3). Appropriate dichroic mirrors (1064nm/45°) were used to drive the laser to the target and a CaF_2 35 cm focus lens was used to focus the radiation into the sample rod.

Photo 2.1.1.3: Laser Quantel Ultra.
http://www.bmisurplus.com/products/45549-big-sky-laser-ultra
(last accessed October 25, 2015).

2.1.2. Mass Spectrometer

The mass spectrometer used was a modified Wiley-McLaren lineal time-of-flight (Jordan TOF Products, Inc., Photo 2.1.2). The expansion chamber was equipped with two extra windows to allow the entrance of the ablation laser radiation and a stepper motor for the rotation and translation movement of the sample rod (designed by Prof. J. Simmons' group at Oxford University). The detector was a Jordan 18 mm active side dual plate model, operating at 2500 V. The vacuum was achieved through two turbomolecular pumps: the first one closer to the detector (Leybold Turbovac 151 N_2: 145 l/s) and the second one below the expansion chamber (Leybold TW 701 N_2: 680 l/s). The turbomolecular pumps were connected to a rotary pump (TRIVAC D25B; 25,7 m^3/h) that produced a preliminary vacuum of $2 \cdot 10^{-3}$ mbar. The final vacuum reached inside the chamber was around ~$1 \cdot 10^{-7}$ mbar, and ~$1 \cdot 10^{-4}$ mbar when the valve was pulsing. The normal operating conditions are $5 \cdot 10^{-5}$ mbar. The vacuum was checked with a Penning (in the chamber) and a Pirani (before the rotary pump) gauges.

Photo 2.1.2: Time of Flight Mass Spectrometer.

2.1.3. Electronic Devices

Several electronic devices were used in this work, which mainly can be divided in power supplies and controllers.

2.1.3.1. Power Supplies

Table 2.1.3.1: Power Supplies

Function	Model	Voltage
Extraction plate	SRS STANFORD RESEARCH SYSTEMS, IMC. Ξ MODEL	4000 V
Accelerating plate	PS350/5000V-25W Ξ HIGH VOLTAGE POWER SUPPLY	3700 V
Detector	BRANDENBURG; Photomultiplier power Supply model 476R	2500 V
Deflecting plate ($X_{1\text{-}2}$&$Y_{1\text{-}2}$)	Jordan TOF	0V

The properties of the power supplies (see Table 2.1.3.1) are very important in order to optimize and reach the best conditions in the mass spectrometer: a stable voltage generated by the plates is in turn related to obtain a better and most

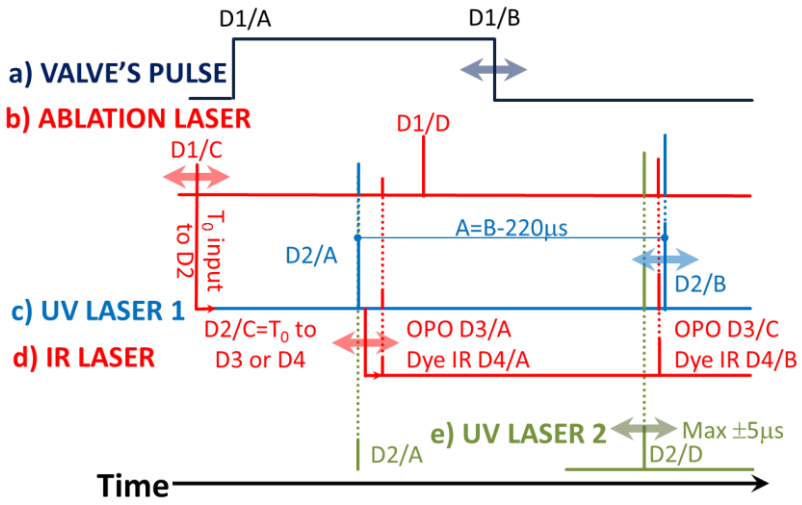
stable signal. For this reason, these components must be of high quality. In the tables the apparatuses that allowed obtaining the best results are indicated.

2.1.3.2. Controllers

In Table 2.1.3.2 all the controller devices necessary for the experiment are described.

Table 2.1.3.2: Controllers

Electronical Device	Made by	MODEL	Description
FOUR CHANNEL DIGITAL DELAY/PULSE GENERATOR	SRS STANFORD RESEARCH SYSTEMS, IMC.	DG535	For the synchronization of all the electronic devices
TWO CHANNEL DIGITAL PULSE DIVISOR	ISMO Workshop	ISMO SE032	It divides the signal and allows the active correction measurements
VALVE PULSE DRIVER	GENERAL VALVE Corporation	IOTA ONE	It generates a square pulse of 28 V to control the valve
STEPPING MOTOR CONTROLLER	OXFORD Workshop	KV0084b	Adjustable control of speed for RS 440-442 stepping motor
OSCILLOSCOPE (option 1)	TEKTRONIX	TDS 3032	300 MHz and 2.5 GS/s oscilloscope used for signal monitoring
OSCILLOSCOPE (option 2)	AGILENT Technologies	DSO9104A	1 GHz and 20 GS/s oscilloscope used for signal monitoring
COMPUTER CONTROL	3RSYSTEM	CPU: E5500 RAM: 4GB	Widows 7 (64 bits) professional with Agilent VEE Control Software



Scheme 2.1.3.2.1: synchronization sequence. The optimization of the signal requires the control of the timing of the events. a) Valve Pulse Driver; with the channel D1/B is possible to adjust the duration of the gas pulse. b) The ablation shot is adjusted in D1/C with respect to the molecular pulse. This channel is linked to all the other channels, in order to move all the lasers together. c) Ionization Laser; adjustable in D2/B and linked to IR and UV second lasers. d) IR Laser; optimizable in D2/C and usually fixed at 50 ns before the ionization laser. e) Q-Switch of the UV second laser; it is not linked with the flash lamp. In order to avoid the energy gap caused by the delay between the flash lamp and the Q-Switch, the maximum range available is 5 µs.

The pulse sequence is controlled by a pulse generator. The delay generator represents one of the most important devices. The same timing distribution can be obtained in different ways. In Scheme 2.1.3.2.1, the specifications for all the events that are taking place during the experiment are shown. For the proper understanding of the scheme, and in order to have all the information need for replicating the experiment, all the details for the delays of the controller devices are reported in Table A2.1.

2.1.4. Sample Preparation

Samples preparation for isolated phase studies requires of two main tools: a pulsed valve which generates the supersonic jet (see Figure 2.1.4.1), and a laser ablation to vaporize thermolabile molecules in the case of solid samples (see Figures 2.1.4.2 and 2.1.4.3). Figure 2.1.4.2 reports the design of the valve in detail, which was designed to simplify the refill of the sample and the laser ablation alignment. Regarding the optimization of the laser ablation, two other refinements were done: a window behind the skimmer and a video camera set at 45° with respect to the nozzle.

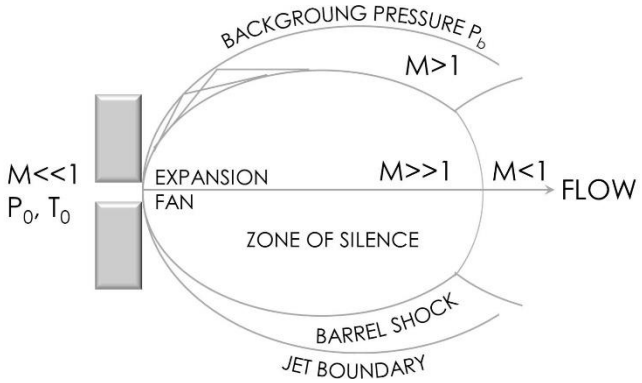


Figure 2.1.4.1: the nozzle is a truncated-cone-shaped pinhole. The gas starts from a stagnation state (P_0, T_0), and with an imposed pressure difference (P_0-P_b) accelerates, as the area of the nozzle decreases, toward the source exit. The flow may reach sonic speed (that is, mean velocity equal to the local speed of sound or Mach number equal to 1) at the exit, provided the pressure ratio P_0/P_b exceeds a critical value: $G \equiv ((\gamma+1)/2)^{\gamma/(\gamma-1)}$, which is less than 2.1 for all gases, (γ represents a property of the fluid species). If the pressure ratio is less than this critical value, then the flow will exit subsonically, with exit pressure nearly equal to P_b, without any further expansion. Typically for this kind of sources, the flow can be considered isoentropic and viscous effects and heat conduction can be neglected. As the jet expands, the velocities of collision decrease and the kinetic processes, at a certain point, stop. This *frozen point* is the aim for this kind of spectroscopic purposes (CI-phase), and it can be used for creating and observing unstable species during the jet expansion (more details can be found in the chapter 2 of ref.4).

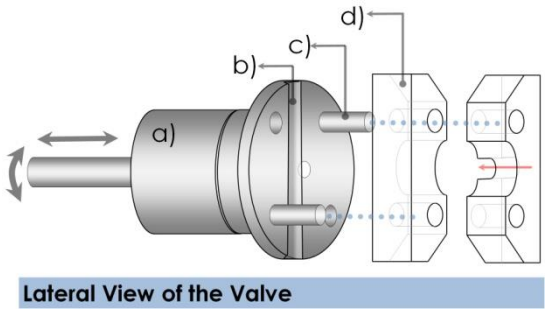


Figure 2.1.4.2: lateral view of the valve. The original design was improved in order to optimize the gas pulse without opening the chamber. The arrows on the left side, show the coupling process to insert the valve in the right position. a) Parker Pulsed Valve: Miniature High Speed high vacuum dispense valve, 28 V opening valve with Kel-F poppet. b) Groove for the sample rod. c) Spike. d) End flange.

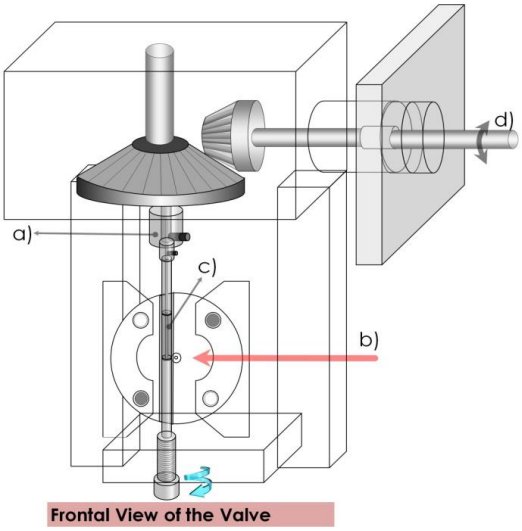
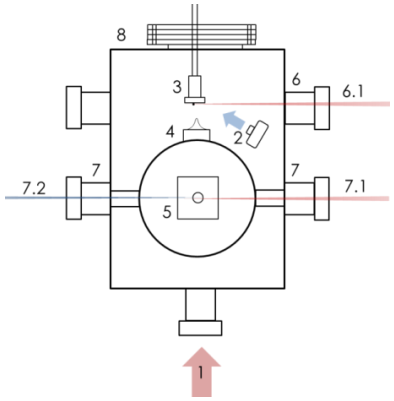


Figure 2.1.4.3: frontal view of the valve with the laser ablation device. a) Coupling between the sample rod axis and the stepping motor axis. b) Ablation Laser. c) Graphite sample rod. d) Stepping motor location.

2.1.5. Alignment of the Laser Ablation

The alignment process is quite complex, especially because, until the laser beam hits fresh sample, any movement of the ablation laser produces a temporary increase of the signal, even if the final alignment get worse. To solve this problem a new specific methodology was developed and it is described in the following pictures.



1) **Alignment 1 (window)**
2) **Alignment 2 (video camera)**
3) Valve and Sample rod
4) Skimmer
5) Mass spectrometer extraction and acceleration plates
6) Window for ablation laser (6.1)
7) Windows for ionization (UV, 7.2) and burning lasers (UV or IR, 7.1)
8) Chamber

Scheme 2.1.5.1: Upper view of the chamber, which presents the typical design of this kinds of mass spectrometers. Two components were added to facilitate the laser alignment (1, 2). The first one is a window at the end of the chamber (**Window 1**), through which is possible to check if the nozzle is perfectly aligned with the skimmer and the mass spectrometer (see Figure 2.1.5.2.). During this procedure, for safety, all lasers are switched off. The 2nd component is a video camera: thanks to this device, once that the nozzle lays in the correct position, it is possible to adjust properly the ablation laser in the center of the sample rod with respect to the nozzle, switching it on (see Figure 2.1.5.3).

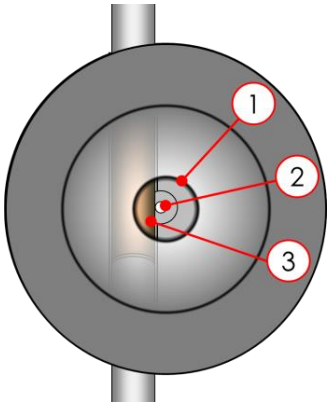


Figure 2.1.5.2: 1) Skimmer hollow. 2) Nozzle. 3) Sample rod. This figure shows the right alignment.

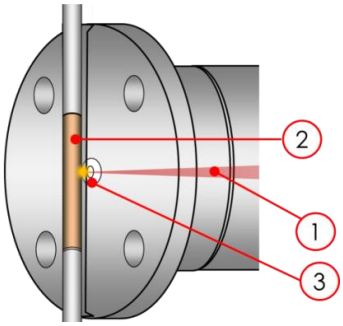


Figure 2.1.5.3: 1) Ablation Laser. 2) Sample rod. 3) Nozzle correctly aligned with the ablation plume. This figure shows the view given by the video camera. During this procedure, the laser is switched on but the operator is safe because the image is taken by the camera.

2.1.6. Deposition of the sample

The sample rod is a graphite cylinder where the sample, after being mixed with carbon nanotubes, is deposited on its surface. In order to optimize the desorption process, it is a good habit to deposit an active carbon layer on the rod before the sample deposition, which makes easier to stick the sample and helps to obtain a more homogeneous distribution of the sample avoiding fluctuations in the signal. In Figure 2.1.6.1, a diagram of the rod and all the layers is shown. It is worth noting that also liquid samples can be prepared using the same deposition method, by adding a drop of the sample on the cylinder.[5]

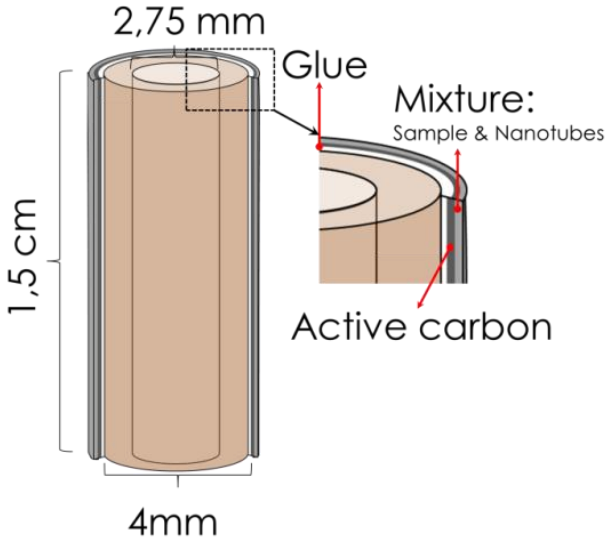


Figure 2.1.6.1: sample preparation. Solid rod layers structure

2.2. Experimental Methodology

The experimental study of the structure and conformation of neutral (bio)molecules (or their complexes) requires of high resolution spectroscopic technique which enables to obtain detailed and accurate structural information for each of the possible conformers. In this thesis, laser spectroscopy coupled to double resonance methods were used.[6] These techniques are very versatile and encompasses several experimental configurations. In the following pages, the most relevant ones will be described.

2.2.1. Resonance Enhanced Multi-Photon Ionization (REMPI)

The REMPI technique is based on the signal amplification produced by a resonant state during the excitation plus ionization process. In Figure 2.2.1.1, some schemes of the absorbance process are reported, where several combinations can take place. Scanning the energy of the excitation photon and recording the ion current generated at each wavenumber, the electronic excitation spectrum is obtained. REMPI provides high sensitivity and, in combination with mass spectrometry, is mass selective: species of different stoichiometry appear in different mass channels and can be studied without interferences. Furthermore, REMPI spectroscopy is useful to obtain information on the different conformational arrangements that a molecule or its complexes can adopt. This is true only when different geometries produce different spectra. In order to separate the information related to each geometry, double resonance techniques are required.

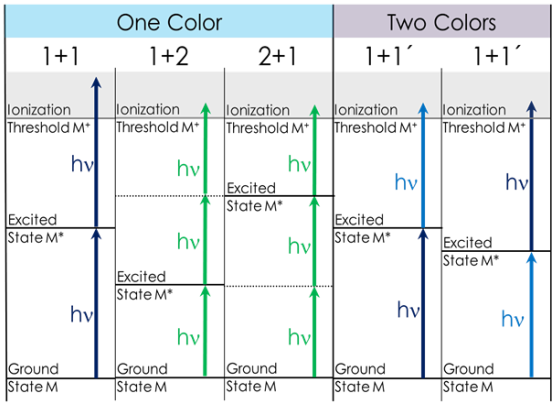


Figure 2.2.1.1: REMPI spectroscopy photon absorption schemes. There are two main groups: one color or single photon source (1+1) and two colors or double photons source (1+1´). Depending on the properties of the molecule, the process can be different: in 1+1 case, one photon is used to reach the excited state and the second one to ionize. In 1+2 case, the first photon is used to excite the system as before, and the next two photons to ionize. Instead, in 2+1 the first two photons are for the excitation of the molecule and the third one for the ionization. In some cases, the ionization process is above the fragmentation threshold: this produces contributions in the spectra of lower mass species and a loss of the signal. To avoid this situation, two colors techniques are used: the first color laser is attenuated in order to produce the absorption of only a single photon; with the second laser, the energy of the second photon is adjusted in the range between the ionization and the fragmentation threshold (cases 1+1´).

2.2.1.1. Active Correction for REMPI Spectroscopy

REMPI is a powerful methodology to access a broad range of information about structure and non-covalent interactions in isolated molecules and their intermolecular complexes. In this work, the scientific challenge was to push up spectroscopy limits towards larger (bio)systems, a path that was not free from difficulties. In fact, these larger systems present in general a low signal-to-noise (S/N) ratio, that makes very difficult to record a reliable REMPI spectrum. In order to overcome this limitation, the active correction technique was used in this work. This method consists of adding an additional laser to evaluate the molecular concentration in the supersonic expansion at each molecular pulse, see Figure 2.2.1.1.1. This signal is used as reference. In this way, it is possible to get the spectrum faster and this is crucial for synthetic samples, for which only few milligrams are available. In Figure 2.2.1.1.2, an example of a normal REMPI spectrum (RAW) is shown and compared to the corresponding corrected one and to the reference.

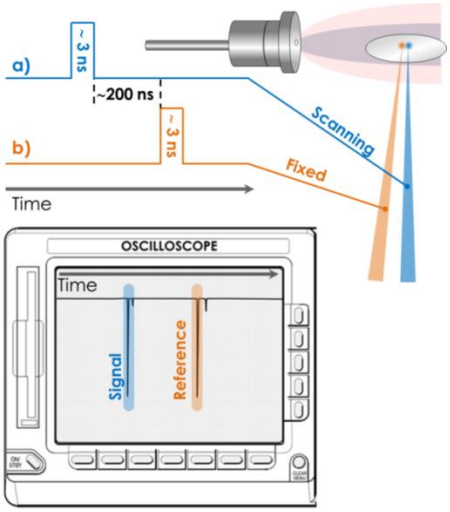


Figure 2.2.1.1.1: (a) and (b) show the synchronization of the laser's Q-Switch. a) The REMPI laser spectrum is scanned in a region which is located spatially next to the region probed by the reference laser, but it is synchronized 200 ns earlier. In this way, even if the two regions are close and equivalent, they do not interfere one with each other. b) This laser is fixed in a wavelength and it is necessary for registering a preliminary REMPI spectrum in order to choose an ionization region where the absorption takes place. It is important to get similar intensity and spot-size between both lasers, otherwise the technique does not work. Moreover, attention must be paid to the changes of the intensities during long measurements, as the energies of the lasers slowly change with time.

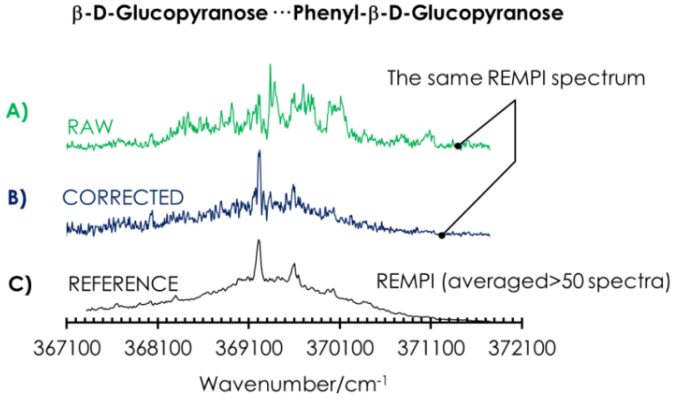


Figure 2.2.1.1.2: A) Normal REMPI spectrum. B) Corrected spectrum fixing the reference between 36100-37100 cm^{-1}; the correction is done by dividing the RAW REMPI spectrum by the reference signal. C) REFERENCE represents the average of more than 50 spectra.

2.2.2. Ion Dip Infrared Spectroscopy (IDIRS)

REMPI has many advantages. However, this technique is not able, for all the molecular systems, to resolve the spectra belonging to different conformers, making compulsory the employment of double resonance techniques. These techniques make use of the intrinsic properties of the complexes to discriminate between different structures. The great potential of IDIRS experiments is that they are conformational selectively. Actually, by adding a second IR laser in the REMPI scheme, the specific IR spectrum of each conformer can be obtained, without interferences due to other molecular structures. This technique gives the possibility to obtain signals using directly the active correction method, using an internal reference sampling. By fixing the UV ionization laser in a wavelength that corresponds to one absorption line, a constant signal is produced; this laser operates at 10 Hz. At 5 Hz, an IR laser is added, which is spatially overlapped but ~50-100 ns earlier in time, see Figure 2.2.2.1. With this arrangement, both the signal and the reference are recorded every 5 Hz, and this enables a real time correction. Indeed, during the supersonic expansion most of the molecules lie in the ground vibrational state, thus when the IR laser is resonant, it excites part of these molecules to another state. For this reason, the ionization laser (UV one) finds less population available compared to when the IR laser is not resonant, producing a depopulation, see Figure 2.2.2.2. The final data are obtained by dividing the recorded signal by the reference one.

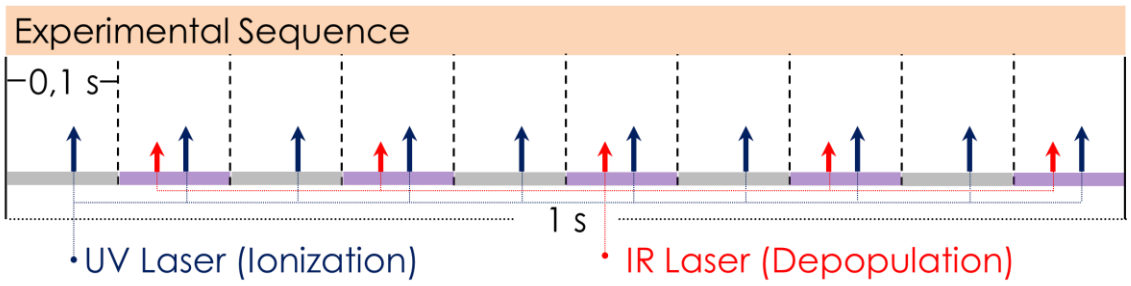


Figure 2.2.2.1: Blue arrow: UV ionization laser operating at 10 Hz. Red arrow: IR depopulation laser operating at 5 Hz. Grey section: reference, given by the ionization laser. Purple section: signal, formed by both ionization and depopulation lasers.

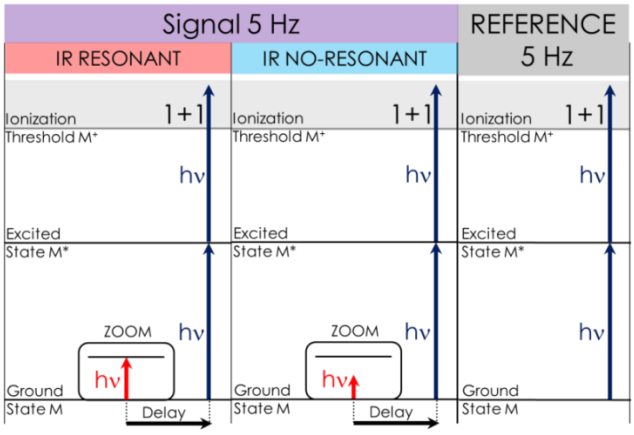


Figure 2.2.2.2: Reference in grey: multiphoton absorption at a fixed wavelength. Signals in purple (5 Hz): two different possibilities. When the IR laser is not resonant, the ionization signal is not affected (blue section); when the IR laser is resonant, part of the population of the conformer is moved to another state, producing a dip in the ionization signal (red section).

2.2.2.1. 10 Hz IDIRS Spectroscopy

The same active correction technique reported in section 2.2.1.1 for REMPI spectroscopy can be employed to obtain faster IR spectra. The only difference is that the IR laser is added and used to obtain the IR spectrum, as is shown in Figure 2.2.2.1.1. The results of this kind of experiment are displayed in Figure 2.2.2.1.2.

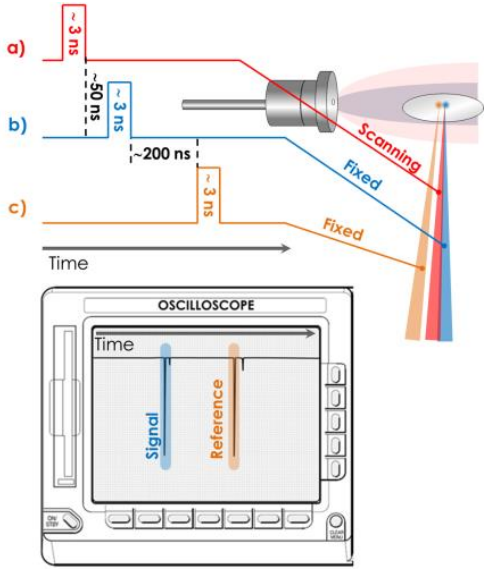


Figure 2.2.2.1.1: the experimental arrangement for 10 Hz active correction IDIRS. a) IR laser configured to scan producing an ion depletion (resonant transitions) on a fixed 50 ns later UV laser. b) Fixed UV laser, when the a) laser arrives to a resonant transition the ions generated by this UV laser undergo a dip of the signal. c) Second fixed UV laser, this works like a reference: when an ion depletion occurs and this is not related to the IR laser, it is corrected by this reference laser.

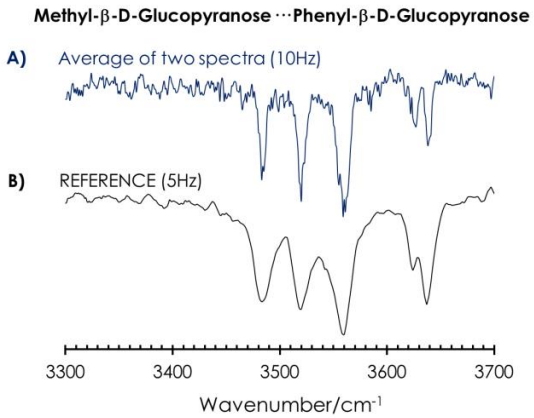


Figure 2.2.2.1.2: A) average of two spectra recorded at 10 Hz. B) Subtracted reference recorded at 5 Hz , average of 10 spectra.

2.2.3. Hole Burning Spectroscopy (HB)

This double resonance technique is analogous to the IDIRS experiments, but replacing the IR laser by a tunable UV laser. In this case, the specific electronic spectrum of each conformer can be obtained. This technique is used to identify the contribution of each conformer in the REMPI spectrum. More than one method can be used to carry out these types of experiments; the most employed approaches are IR/UV and UV/UV. The setup is the same shown previously for the IDIRS experiment, but in this case the first laser can vary (see Figure 2.2.3.1). The time sequence of the experiment is similar to that in Figure 2.2.2.1, but in this case the IR laser operating at 5 Hz is fixed in one of the transition of the IR spectrum. It is also possible to replace the IR laser (5 Hz) with a UV one (5 Hz), see the blue box (UV/UV HB) in Figure 2.2.3.1. In both cases the information is obtained scanning the REMPI laser. Working at 5 Hz, it allows one to achieve two different types of information: when only one laser is recording, the result is the REMPI spectrum, while when there are both lasers, the depopulated REMPI spectrum is recorded. The difference between the REMPI and the depopulated REMPI spectra, produces a specific REMPI spectrum of the species related to the depopulate transition. This kind of experiment is very useful to identify how many isomers are present and also to choose the ionization transition where only one isomer is ionized.

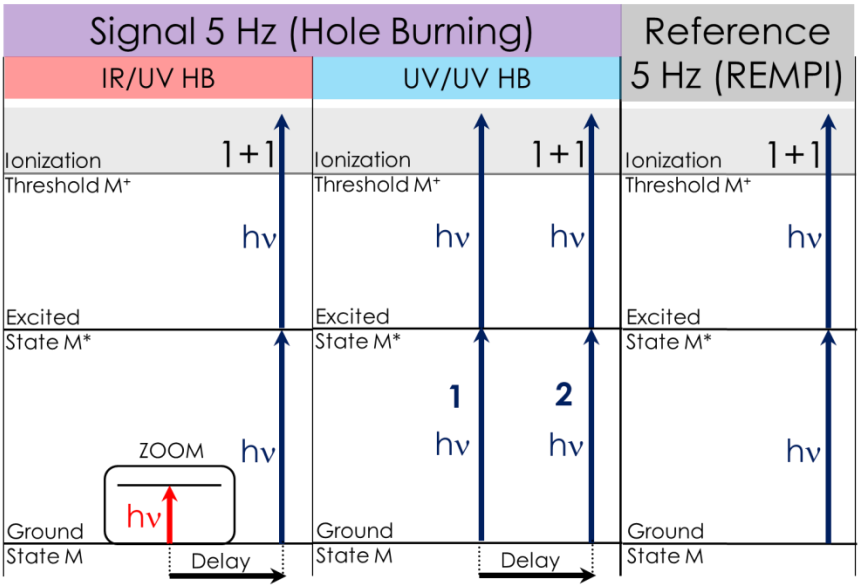


Figure 2.2.3.1: The UV laser operates at 10 Hz (as in Figure 2.2.2.1) but in this case it is also scanning. The reference, operating at 5 Hz, is recording the REMPI spectrum. The signals (purple section) at 5 Hz are recorded combining REMPI spectra with some dip in the lines; all the lines associated with the isomer that is depopulated undergo a depletion of the signal.

2.3. Theoretical Methodology

The main objective of the theoretical analysis regards the structural and energy determination of the molecules and their aggregates. The conformations are assigned through the comparison between the experimental spectra and the theoretical predictions. The theoretical approach, from the conformational search until the final assignment, involves several steps. In order to select the most accurate methodology for each molecular system, first of all it is necessary to define which the main variables are. In general, they can be distinguished in three different groups: 2.3.1) chemical variability, referred to the flexibility of the molecules and clusters; 2.3.2) computational accuracy, focused on the approach used for the conformational search and the following Density Functional Theory (DFT) calculations, and finally 2.3.3) dynamics of the formation of the aggregates during cooling. Moreover, the importance of these variables in affecting the predictions is also analyzed with a more general chemical point of view, considering possible constraints or benefits that have to be included in the calculations.

2.3.1. Chemical Variability

The huge amount of atomic combinations that can be achieved, produce an enormous number of molecules, characterized by a wide range of conformational properties: from rigid rings to linear high flexible molecules. The origin of this heterogeneity arises from the interaction between atoms and it can be divided into two groups: 2.3.1.1) intramolecular variability, produced by the intrinsic properties of the molecule, by non-covalent and long-range interactions; and 2.3.1.2) intermolecular variability, governed by the weak non-covalent intermolecular interactions alone.

2.3.1.1. Intramolecular Variability

The intramolecular conformational flexibility is caused by different kinds of variables, but it can be divided into two categories: those that involve orientations of part of the molecule, such as changes in the orientation of an hydroxymethyl group, and changes that involve internal rearrangement processes, such as tautomerization.

2.3.1.1.1. Intramolecular Dihedral Flexibility

The rotation of a dihedral angle is the main degree of freedom behind this kind of flexibility. In Figure 2.3.1.1.1, a general description of the conformational flexibility is reported, ordered according to the number of atoms. It is worth noting how, starting from very simple rigid compounds, the complexity related to the molecular flexibility increases very fast, even with just small changes. Some examples of interconversion paths in molecules belonging to the families b2, c2-3 and d2-3-4 of Figure 2.3.1.1.1 can be found in Figures A2.2, A2.3, and A2.4. Moreover, in Figure A2.5, the representation of the increased conformational isomerism related to the insertion of a single carbon atom in a molecular system is done (linear alcohols case).

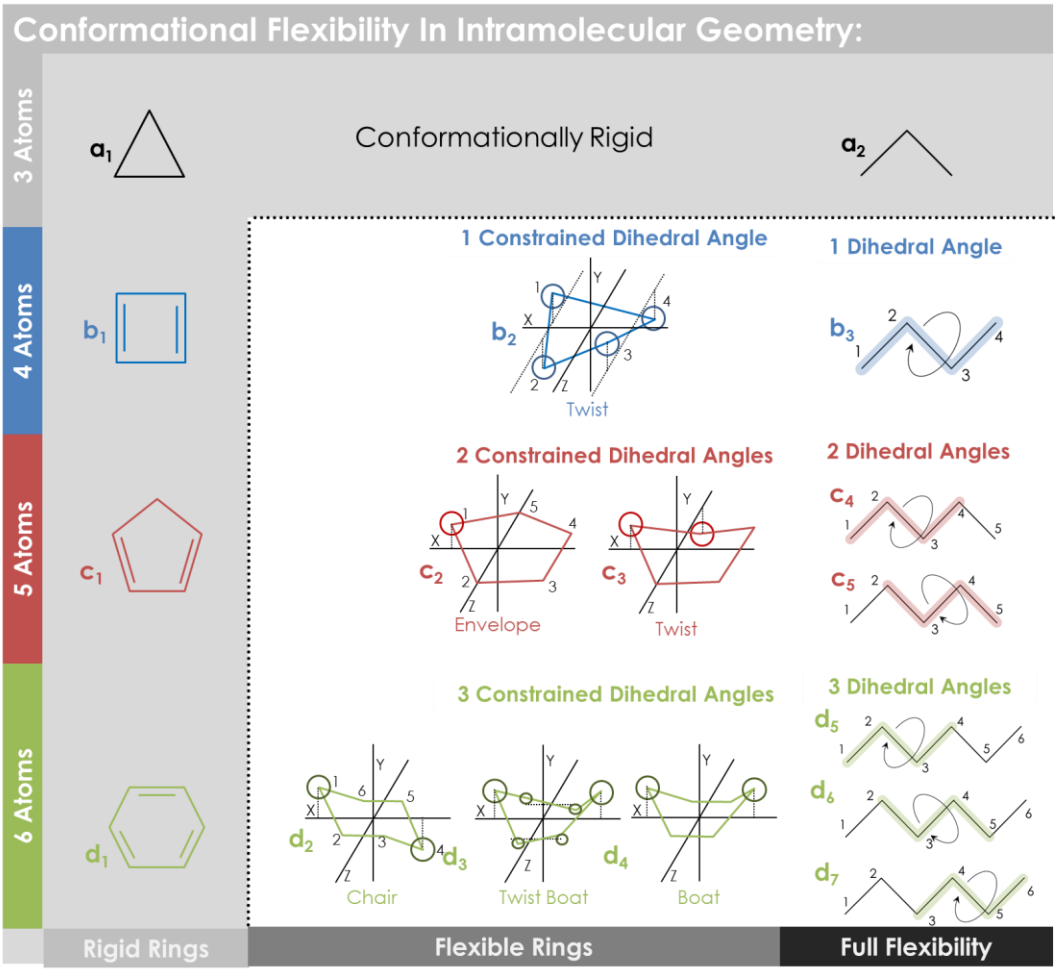


Figure 2.3.1.1.1: Rigid Rings: a1, b1, c1 and d1 which present a single isomer. **Flexible Rings:** b2, c2, c3 and d2, d3, d4 for which several conformational isomers are allowed and they may interconvert through puckering movements. The circular markers denote the atoms out of the *XZ*-plane. For these dihedral angles, the rotational degrees of freedom are limited by the ring, hindering torsions bigger than 90°. **Full Flexibility:** b3, c4, c5 and d5, d6, d7. Starting from linear molecules with 4 atoms (b3) where only one dihedral angle can be defined, a new dihedral angle appears each time that one atom is added to the system (c4 and c5: 5 atoms and 2 dihedral angles; d5, d6 and d7: 6 atoms and

3 dihedral angles, and so on). As regards organic compounds, the rotation of 360° around each dihedral angle produces 3N conformational minima, where N represents the number of heavy atoms (C, O, N, etc.) – 3, of the molecule. Because of steric hindrance, these minima for the alkyl chains usually lie at the three staggered orientations, namely *anti*, *gauche* and *gauche*'. The corresponding dihedral angle values are approximately: 180°, 60° and -60°, respectively. For long chains another phenomenon can occur: e.g a *gauche* (60°) minimum splits into two new minima (~30° and ~90°) because of the stabilization that can occur along the longer chain (see Figure A2.5: conformations AG_EG2-AG_EG2d & G_EG2G-G_EG2G2d).

2.3.1.1.2. Intramolecular Isomerization

Intramolecular isomerization is another type of chemical variability. It may affect also to rigid molecules, which do not show the conformational flexibility presented in the previous section. One of the most representative reactions is the tautomeric equilibrium, and the prototropic one represents one of the most common form of tautomerism. In this case protons, generally those in OH and NH groups, can migrate from one position of the molecule to another. In Figure 2.3.1.1.2 an example of this isomerization is depicted: the purine base xanthine presents a relocation of a proton from position N7 to position N9 and an amide-imidic acid tautomerism (lactam-lactim pairs for positions Ox-Nx and Oy-Ny).

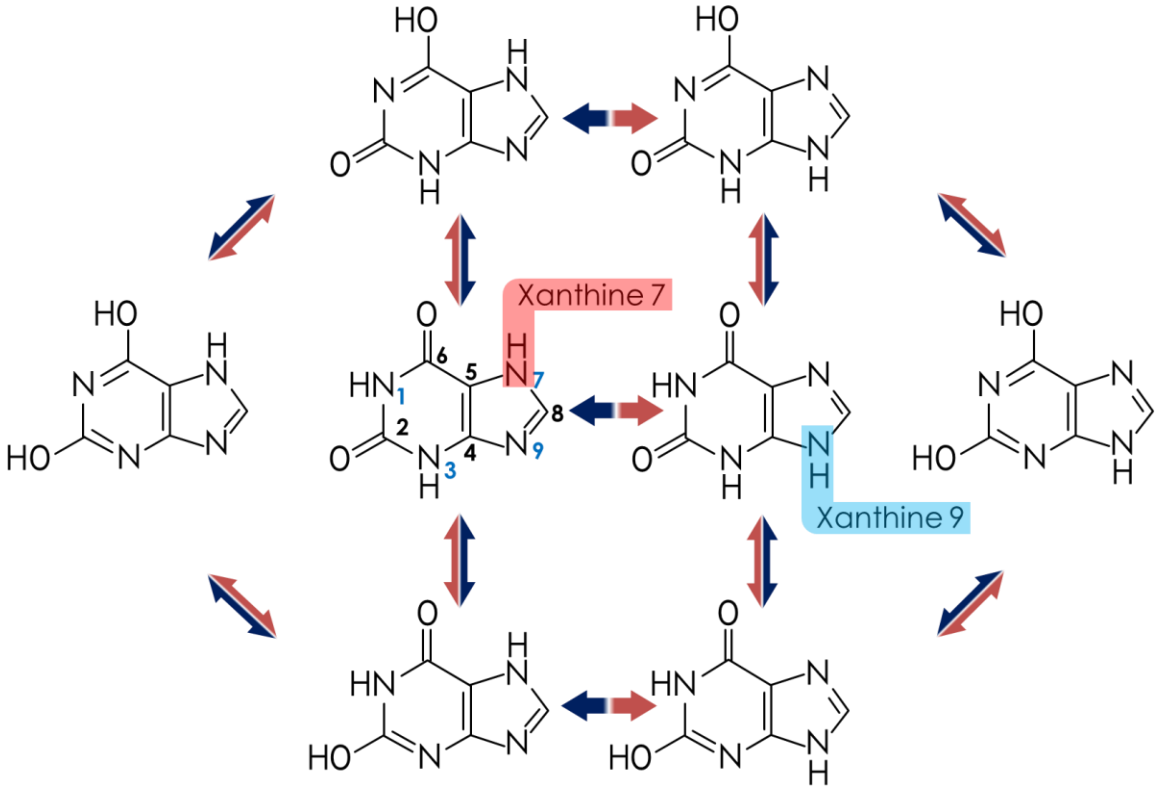


Figure 2.3.1.1.2: The chemical variability in a rigid molecule: xanthine. The possible tautomers for this purine base are reported.

2.3.1.1.3. Intramolecular Folding

This is a special case of molecular flexibility that takes place in long-chain molecules, such as peptides and it is the most complicate process due to the combination of many effects that occur simultaneously. The folding process is guided by non-covalent intramolecular interactions. The kinetic of this mechanism can be divided in several steps: firstly, some parts of the molecule interact rapidly forming limited folded sections; then, these sections interact with each other producing the final 3D structure. The geometries adopted in the initial step introduce constraints that limit the number of conformations available during the final folding stage. This means that in some cases the formation of the most stable structure is not favored by the folding paths, see Figure 2.3.1.1.3. In fact, during the folding some molecular portions present stronger interactions than others and this enhances the formation of the most stable intermediate (Ea_2), blocking sometimes the possibility to reach geometries coming from the less stable intermediate (Ea_1) and that would result in more stable final structures. Although, this does not prevent the formation of the Isomer 1 through an interconversion pathway, such way significantly complicates to reach that conformation. The interconversion process between isomers is a very complicated task to deal, making difficult to know which is the corresponding path and if it is allowed (isomerization energy: Ei_1 or Ei_2). A deeper study of this topic will be presented in Chapter 5.

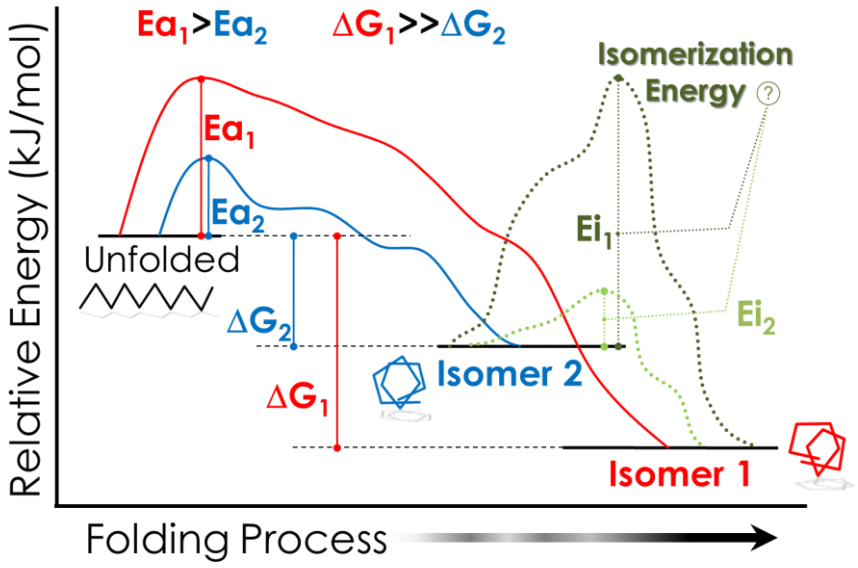


Figure 2.3.1.1.3: an example of folding paths for two isomers: Isomer 1 presents a higher activation energy (Ea) despite that it is more stable than Isomer 2. The isomerization energy (Ei) determines the resulting conformation: if this is low (Ei_2), after the formation of Isomer 2, it will evolve into Isomer 1. Otherwise, if the barrier is too high (Ei_1), mainly Isomer 2 will be formed.

2.3.1.2. Intermolecular Variability

The intermolecular variability refers to the many ways in which molecules can interact when forming molecular aggregates. The different factors responsible of this process can be divided in three main groups: 2.3.1.2.1) flexibility of the binding site; 2.3.1.2.2) competition between binding sites; and 2.3.1.2.3) statistical effects on formation of large clusters.

2.3.1.2.1. Flexibility of the Binding Site

The flexibility that characterizes the chemical bond is related to the number of stable geometries and connection pathways. When the number of connected geometries is high and the energy barriers enable interconversion, the number of candidates for the assignment of the experimental spectra grows up. Depending on the experimental technique and the results achieved, it is impossible in many cases to carry out a univocal assignment and only a family of molecules or species can be proposed. In Figure 2.3.1.2.1 some examples are presented: a) the flexibility that characterizes the stacking structures of xanthine dimer; b) the flexibility of the phenol unit in its interaction with methyl-β-D-glucopyranose; c) the effect of the monomer's structure of phenyl/methyl-β-D-glucopyranose inside the dimer. All these examples will be treated in the following parts of this thesis.

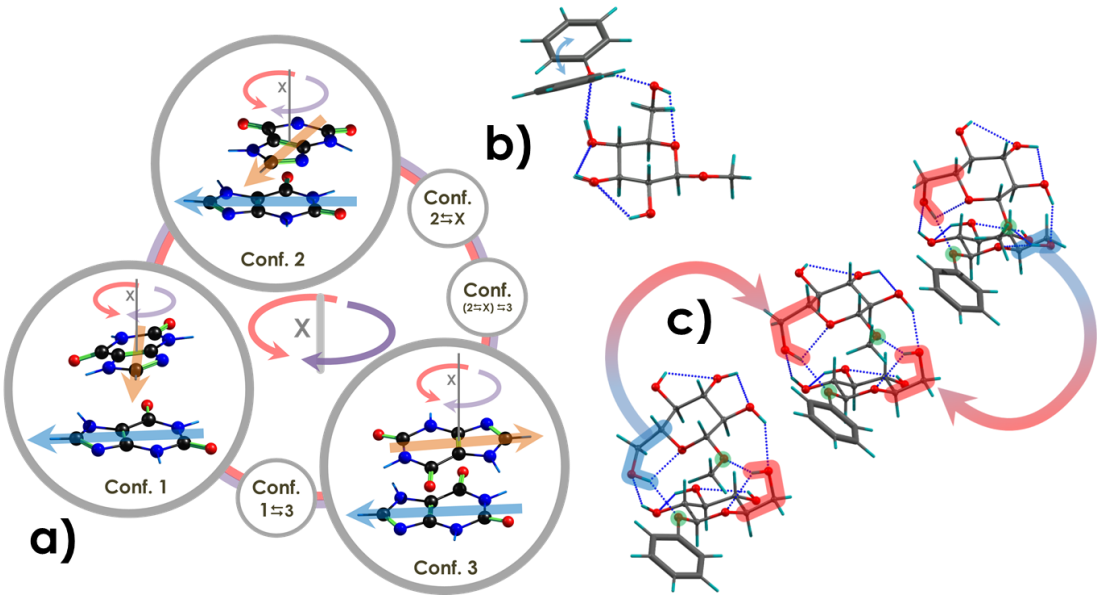


Figure 2.3.1.2.1: a) xanthine dimer displayed from the side to show the stacking; the lower monomer was kept fixed, while the upper one was rotated around the rotation axis x in order to simplify the interconversion pathway between all the stacked isomers. The little grey circles identify the biggest structural change for the interconversion between isomers: the change between conformers 1 and 3

(Conf.1⇆3) is ~90°, while for passing from conformer 2 to 3 requires a rotation of ~230°. With this big difference, it is possible that more than one stable intermedium isomer exists, and in order to mark this possibility, two little circles are displayed in the diagram. As regards these stacked arrangements, the assignment is complicated and at the same time it is easy to overlook intermediate structures. b) Phenol interacting with methyl-β-D-glucopyranose: the phenyl moiety can approach the sugar by forming a stacking interaction or, from the other side, avoiding it. c) Glucose dimer: the green balls mark the anomeric oxygens; the blue lines refer to the hydroxymethyl substituent in GA (G= gauche, A=anti)conformation and the red lines indicate the hydroxymethyl group in GG arrangements. In this case each isomer presents a unique vibrational signature, due to the perturbation that the interaction introduced in the vibrational modes of the groups involved in the intermolecular H-bonds, enabling to propose very accurate assignments.

2.3.1.2.2. Competition between Binding Sites

Two molecules can interact in a large number of ways. The most simplistic way to tackle this problem would be to visualize the molecules as spherical surfaces. Thus, the total amount of possible interactions would be given by the product of their corresponding interacting coefficients. In order to simplify the situation, different kinds of parameters can be used, like the symmetry, repulsive interactions or chemical intuition. The strongest simplification comes from the identification of the main binding sites. The preferential binding sites present an extraordinary stabilizing effect with respect to the other parts of the molecule, which can have different competing sites, each one able to act as an anchor point, but usually one of them is the preferred interaction site. Figure 2.3.1.2.2 shows the procedure followed to simplify the preferential binding sites of xanthine 7H for dimerization.

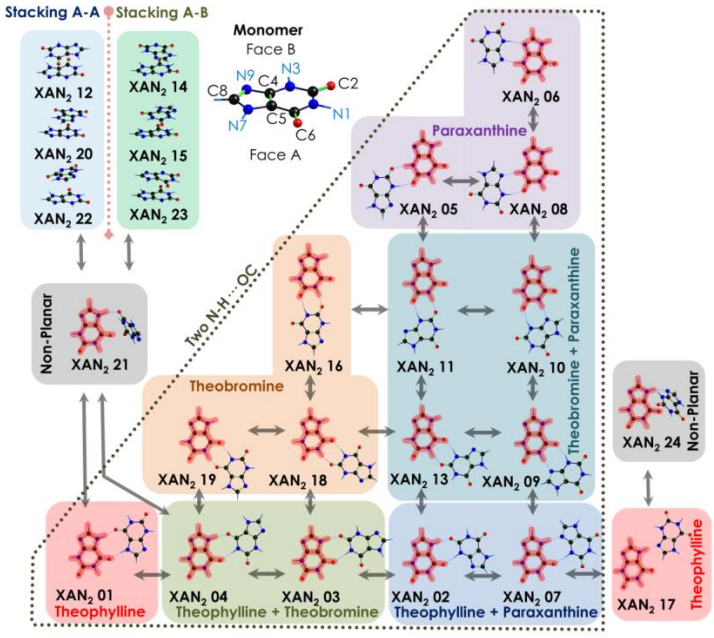


Figure 2.3.1.2.2: Overview of the main interaction sides of xanthine 7H in its dimer formation.

2.3.1.2.3. Statistical Effects in the Formation Process of Large Clusters

The complexation in the CI-phase is mainly determined by some boundary conditions, which drives the formation of the clusters towards a given direction. For example, the trimer is mainly generated by the combination of the dimer with a third molecule. Likewise, the tetramer can be formed by combining the trimer with another monomer or through the combination of two dimers. For this reason, the preferential geometry adopted by the dimer affects the structural arrangement chosen by the trimer, and therefore, also for the tetramer. This tendency can also be extended to the generation of bigger clusters. In Figure 2.3.1.2.3, the possible trimers that can arise from each dimer of theobromine are presented (see Chapter 6 for the processing of this topic).

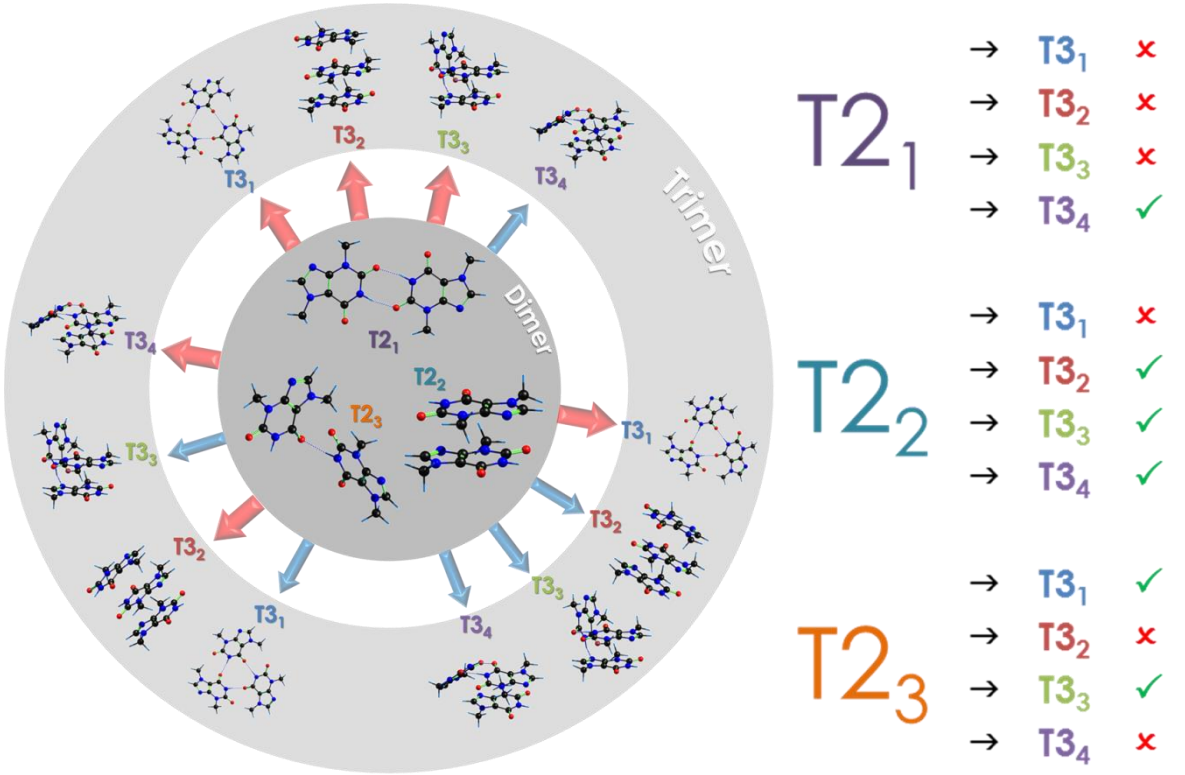


Figure 2.3.1.2.3: statistical combination of the theobromine dimers in order to generate trimers. Formation of the trimers is determined by the geometry of the dimer and the interconversion pathways. In this diagram, only some possibilities are shown, since the complete genealogical tree with the experimental assignments are reported in Chapter 6.

2.3.2. Computational Accuracy

This section is divided in two parts: 2.3.2.1) conformational search, and 2.3.2.2) structure optimization and frequency calculations. The first one aims at defining a list of all the possible geometries that the molecule or the cluster can adopt. In the second part, the structures obtained are optimized and minimized using several theoretical methods, which are selected considering the properties of each system.

Briefly, the exhaustive conformational search required to screen all possible molecular structures is done using fast molecular mechanics methods, without imposing any restriction on the systems.[7] The structures were generated through advanced Monte Carlo methods and low-mode conformational search algorithms on large scale systems (for more details, see section 2.3.2.1). Then, the most relevant geometries, attending to relative energies and representative structures of each family, were fully re-optimized using DFT and/or ab-initio methods. The DFT computational method mainly used in this thesis is the dispersion-corrected M06-2X,[8] while the ab-initio one is the MP2,[9] both implemented in Gaussian 09 program package.[10] All optimizations were conducted imposing no structural constrains and a normal mode analysis was carried out to test the absence of negative frequencies. Finally, the frequencies of the OH, NH and CH stretch modes were corrected for anharmonicity using different multiplier factors. Binding energy values include also the Boys and Bernardi correction to the basis set superposition error (BSSE).[11]

2.3.2.1. Conformational Search

Conformational searches represent fast approaches with which a general vision of a problem can be quickly gained, although these methods are not very accurate. For this reason the chemical analysis described above is fundamental in order to analyze the system and recognize the possible absence of important conformational families. In fact, several situations can be hardly included in the conformational search, such as the existence of tautomeric structures, for which a specific conformational search has to be carried out. A wide literature is available describing the operation of the conformational search. In particular, in this thesis the simulation procedure employed was already tested successfully for similar systems, and makes use of the MacroModel software.[12,13] The force fields that result more interesting for this thesis are the following: OPLS 2005,[14] MMFFs,[15] and AMBER (also with the atomic charge correction for sugar GLYCAM).[16,17]

In many cases, the list of possible candidates is very large. In order to simplify it, a pre-selection procedure can be applied: firstly choosing a small energy window, then the lowest-lying energy conformers can be optimized by using the Gaussian 09 program package.[10] Usually, this approach enables finding the most stable minima, since the force fields are generally developed using data bases with the structure of real molecules and if the parametrization is good enough, the candidates obtained constitute a representative set of structures. On the other hand, this approach can also be a source of mistakes, so it is always necessary evaluate all the whole set of candidates. Moreover, through the clustering procedure which is available in Maestro software (Schrödinger Release 2017-1, LLC, New York), to group the geometries based on their similarities, it is possible to select one or two representative candidates for each family and subsequently optimize them. For this reason, in order to reach the most realistic picture of the PES, it is necessary to repeat all these processes using several force fields, since some structures can be missed otherwise.

2.3.2.2. Structure Optimization & Frequency Calculations

The geometries coming from the conformational search are energetically optimized to the most stable conformation. As already mentioned, in this work the M06-2X functional was mainly used, but in some cases the B3LYP functional with Grimes Dispersion (keywords: Empirical Dispersion=GD3) was also used.[8,18] The basis sets usually employed are Pople's double and triple ζ (6-31+G(d) and 6-311++G(d,p), respectively). For the final optimized geometry the corresponding vibrational frequencies were calculated to validate the structure as a minimum and to estimate the harmonic frequencies for comparison with the experimental IR spectra. The final assignment arises from the comparison of the calculated spectra with the experimental measurement. Usually, the first trial assignment focuses on the global minimum conformation, but in some cases there are factors that influence the experimental measurement and so the assignment can be directed to evaluate the presence of other stable geometries. In fact, during the generation of the isolated molecules and complexes in the supersonic jet, different processes are competing and this gives rise to a dynamic situation in which the equilibrium depends on the relative energies and the barriers related to the formation pathways, as described below.

2.3.3. Dynamical Effects of the Expansion Cooling

The formation of the CI-phase is a complex non-equilibrium process.[4] It can proceed in two different ways: (*i*) the sample is evaporated by thermal heating and seeded in the carrier gas used to create the expansion; or (*ii*) the sample is evaporated by a laser shot outside the nozzle and the vapor is picked by the molecular pulse. In both cases the temperature of the sample is high enough to enable isomerization, and also to produce (usually undesired) fragmentation products. During the expansion, the enormous amount of collisions produced by the carrier gas cool the sample rotationally, vibrationally and translationally. This cooling process is produced suddenly and it preferentially enhances the formation of the kinetic products, trapping the stable conformations at the temperature before the cooling (Figure 2.3.1.1.3).[19–22] To better understand this process, it is necessary to analyze two main aspects: *i*) the conformer distribution before the supersonic expansion,[23] and *ii*) the allowed barriers to overcome by the collisions produced by the carrier gas.[24] In Figure 2.3.3 these two aspects are explained.

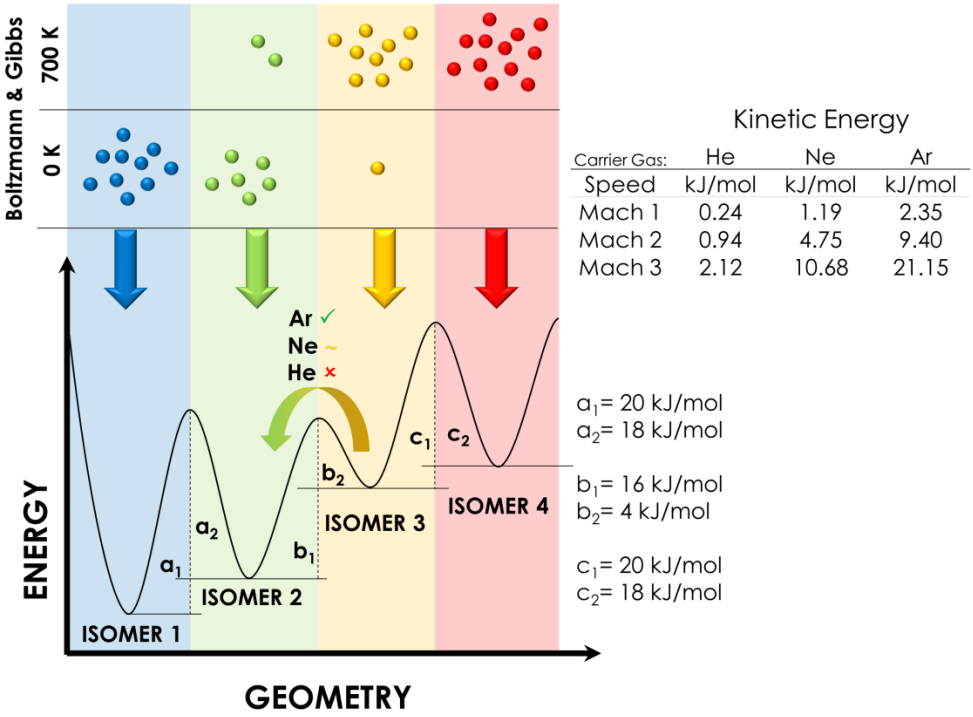


Figure 2.3.3: There are two main aspects affecting the cooling process during the supersonic expansion: the distribution of conformers before the cooling and the barriers for isomerization relative to the collisions with the carrier gas. Let us consider the two cases: 0 and 700 K, reported in the upper part of the figure. In the first one, the species represented by blue and green balls are mostly at 0 K. In the second case, at 700K, the species mainly formed are the yellow and red ones. In this second situation, the collisions with the carrier gas may enable overcoming the interconversion barriers, driving the population into a different, more stable isomer. For example, considering the speed of the supersonic expansion (Mach ~2) and the kinetic energy transferred during a direct collision with a

single atom of the carrier gas, the barrier between the isomer 3 and the isomer 2 (b_2) can be surpassed in the direction of isomer 2, if the carrier gas is argon (trapping green and red isomers). In the case of neon, the collision energy lies on the limit, so the crossing of the barrier is not that straightforward, while for helium this process is forbidden, and yellow and red isomers will be trapped.[4]

The situation complicates when the laser ablation is used: the plume formed contains many molecules at an undetermined, but high, temperature.[25,26] This situation in turn complicates the selection of a single temperature for the Gibbs free energy diagram. In order to solve this problem, the full temperature diagram was calculated, from 0 K up to 700 K, the temperature at which the organic matter decomposes. By exploiting these diagrams, it is possible to determine the most stable species at each temperature, and this permits to choose all the candidates for the assignment, incorporating the possible entropic contributions. The vibrational temperature produces the main differences between conformers.[23] The barriers allowed by the collisions produced by each carrier gas are displayed in Figure 2.3.3, in the table of kinetic energies.[4,27]

The calculation of the Gibbs free energy is more specific and each system needs of a specific approach.[28] This type of calculation can be divided in two groups: 2.3.3.1) the first regards the relative Gibbs free energy for monomers, while the second, 2.3.3.2), is for the complexes. This second group can in turn be divided in the Relative Gibbs Free Energy and in the Relative Binding Gibbs Free Energy.

2.3.3.1. The Relative Gibbs Free Energy in Monomers

For the calculation of the relative Gibbs free energy, the following equations were used:

$$1)\quad \Delta G_{ISOMER\,i,T} = \Delta E_{Electronic\ i} + \Delta E_{ZPE\,i} + \Delta E_{ddH_{i,T}} - T \cdot \Delta S_{dS_{i,T}}$$

$$2)\quad \Delta G_{Relat.ISOMER\,i,T} = \Delta G_{ISOMER\,i,T} - \Delta G_{ISOMER\,GM,T}$$

Equation (1) gives the Gibbs free energy for the *i*-th isomer at the specific temperature T, taking into account the electronic energy of the isomer ($\Delta E_{Electronic\ i}$), the zero-point energy correction ($\Delta E_{ZPE\,i}$), the contribution of the enthalpy at each temperature ($\Delta E_{ddH_{i,T}}$) and the entropic contribution ($T \cdot \Delta S_{dS_{i,T}}$). In order to obtain the relative Gibbs free energy of the isomers, equation (2) was used, where Gibbs free energy of the global minimum ($\Delta G_{ISOMER\,GM,T}$) is subtracted from the energy of the selected isomer ($\Delta G_{ISOMER\,i,T}$). An example of a final diagram is reported in Figure 2.3.3.1.

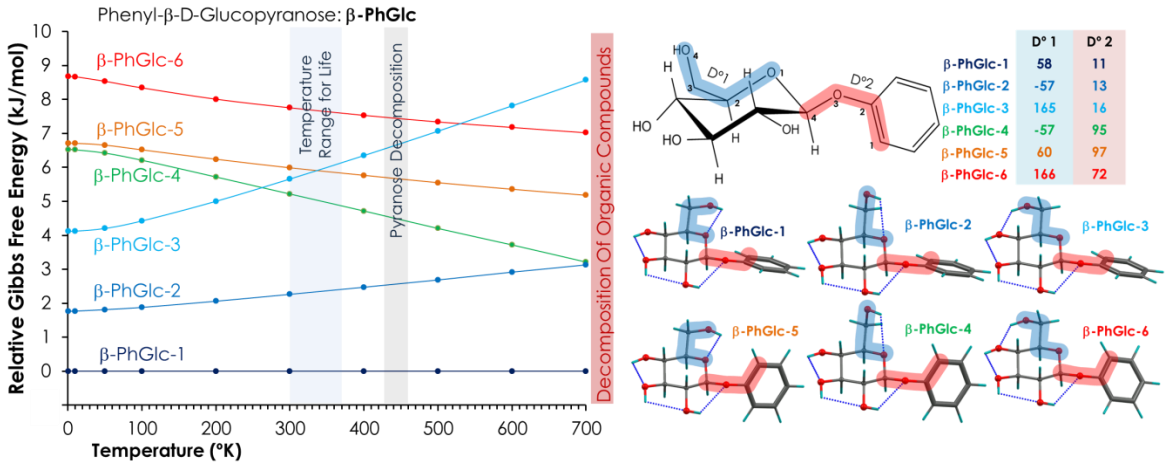


Figure 2.3.3.1: Relative Gibbs free energy diagram for phenyl-β-D-Glucopyranose (β-PhGlc).

2.3.3.2. The Relative Gibbs Free Energy in Complexes

Regarding the complexes, in order to clearly understand their Gibbs free energy, it is necessary to analyze the formation and the fragmentation processes. In Figure 2.3.3.1.1 the possible formation pathways for a complex are shown. There are two main pathways; the most common one follows this general reaction: $A_i+M \rightarrow A_iM$. However, depending on the conditions, another path is also possible: $A_i+M \rightarrow A_JM$. This second mechanism depends mainly on the barrier between the conformations. As explained previously, the collisions and the carrier gas usually influence the overcome of these barriers. All these effects are summarized in Figure 2.3.3.2.1.

In order to calculate the binding energy, it is necessary to consider the fragmentation diagram. In general, there are two main pathways: one is connected with the most similar conformation in the monomer (MS) or , regardless of the conformation adopted in the complex, with the global minimum conformation (GM). In Figure 2.3.3.2.2 these two binding energy calculations are explained.[29]

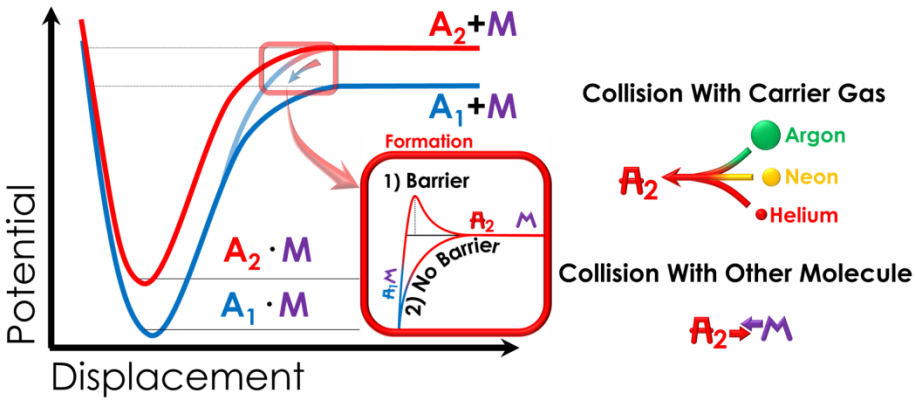


Figure 2.3.3.2.1: Potential energy profiles for the formation of two isomeric complexes. The formation of the complex A_1M or A_2M is possible through a direct reaction: A_1+M to form A_1M, and A_2+M to form A_2M. It is also possible that an aggregation process occurs, through interconversion, giving rise to: A_2+M which forms A_1M. This aggregation pathway can have or not a barrier (but usually a barrier is present), and in order to overcome this barrier, two types of collisions are required: collisions with the carrier gas like in the monomer case (the dominant ones) and collisions with other molecules (M). In theory, this second type of collision are the ones that can overcome big barriers because the kinetic energy increases with the mass (the expansion speed keeps always similar, between Mach 1 and 2). This latter process is not well clear, anyway these kinds of collisions are really improbable, so the typical observed effects useful for passing a barrier are produced by the carrier gas collisions.

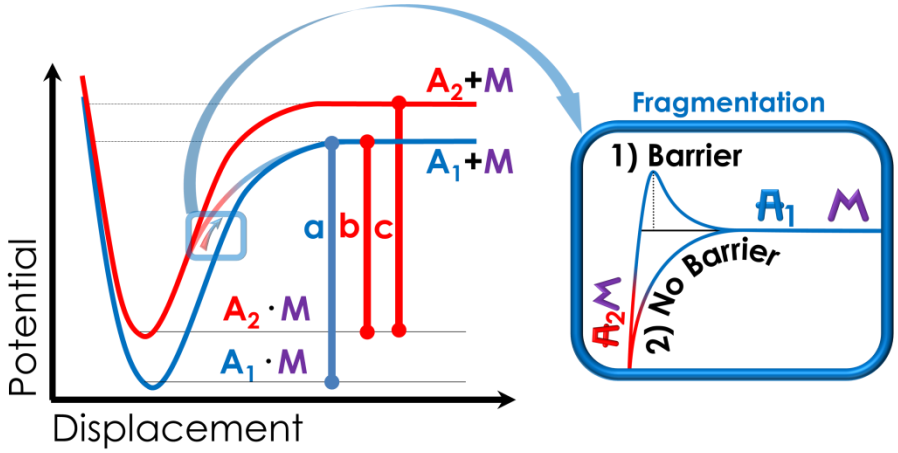


Figure 2.3.3.2.2: Potential energy profile for the fragmentation process. Two main pathways are possible: through the **Global Minimum (GM):** for $A_1M \rightarrow A_1$ + M the (a) difference is required, and for $A_2M \rightarrow A_1$ + M the (b) one. In this latter case, a further barrier can also be present or not (see the cases 1 and 2 inside the blue box); in general it is necessary to assume that there is barrier and to unravel its height, an extra calculation is required. The other case regards the fragmentation through the **Most Similar** monomer conformation **(MS):** for $A_1M \rightarrow A_1$ + M the same (a) path is required as before, while for $A_2M \rightarrow A_2$ + M the (c) difference is necessary, which is higher in energy.

Based on these diagrams (Figures 2.3.3.2.1 and 2.3.3.2.2.), it is possible to perform some energy calculations. The first step regards the Gibbs free energy calculations for the complex; about this another parameter is required, namely the Counter Poise correction ($\Delta E_{BSSE_{A_iM}}$).[11] In equation 3 the Gibbs free energy is reported for a

temperature T for the complex AM, where the species A lies in its *i*-th conformation, while the M species presents only one possible arrangement ($\Delta G_{A_iM,T}$).

3) $$\Delta G_{A_iM,T} = \Delta E_{Electronic\ A_iM} + \Delta E_{ZPE\ A_iM} + \Delta E_{BSSE\ A_iM} + \Delta E_{ddH\ A_iM,T} - T \cdot \Delta S_{dS\ A_iM,T}$$

Using equation (3), it is possible to calculate the binding Gibbs free energy. In order to obtain these results, different ways can be used and each of them can reveal different type of information. The MS methodology uses the most similar conformation in the monomer ($\Delta G_{A_i,T}$) to calculate the binding energy ($\Delta G_{(MS)A_iM,T}$), see equation (4). In this kind of diagram, the ΔG=0 point identifies the maximum temperature at which the complex is stable. Above this temperature ($T_{\Delta G_{(MS)A_iM}=0}$) the complex fragments. Equation (5) is specific for the global minimum and matches that used for the calculations of the GM relative energy in equation (8), since the global minimum is the same in both cases ($\Delta G_{(MS)A_{GM}M,T}$=$\Delta G_{(GM)A_{GM}M,T}$). The MS relative binding Gibbs free energy reveals information about the binding energy and the height of the potential energy profile without any isomerization, see equation (6).

4) $$\Delta G_{(MS)A_iM,T} = \underset{Equation\ 3}{\Delta G_{A_iM,T}} - (\underset{Equation\ 1}{\Delta \boldsymbol{G_{A_i,T}}} + \Delta G_{M,T})$$
$$\boldsymbol{T_{\Delta G_{(MS)A_iM}=0}} \equiv Equilibrium\ Temperature\ (MS)$$

5) $$\Delta G_{(MS)A_{GM}M,T} = \Delta G_{A_{GM}M,T} - (\Delta G_{A_{GM},T} + \Delta G_{M,T})$$

6) $$\boldsymbol{\Delta G_{Relat.(MS)A_iM,T} = \Delta G_{(MS)A_iM,T} - \Delta G_{(MS)A_{GM}M,T}}$$
$$\boldsymbol{Relative\ Binding\ Gibbs\ Free\ Energy\ (MS)}$$

The other calculation methodology uses the GM conformation to estimate the relative binding Gibbs free energy, see equation (7). The equilibrium temperature obtained in this way it is not the real one because in many cases, when it is not based on the most stable conformation of the constituent monomers, the separation in energy between the different isomers of the complex implicates an isomerization and in most of the cases this is not allowed. With respect to the GM binding energy, this energy does not produce very useful information, but the GM relative binding Gibbs free energy is equal to the relative Gibbs free energy of equation (11), and it is useful for the understanding of the relative stability of the complexes.

$$7)\quad \Delta G_{(GM)A_iM,T} = \underset{Equation\ 3}{\Delta G_{A_iM,T}} - \underset{Equation\ 1}{(\Delta G_{A_{GM},T} + \Delta G_{M,T})}$$

$$\boldsymbol{T}_{\Delta \boldsymbol{G}_{(\boldsymbol{GM})\boldsymbol{A_iM}}=\boldsymbol{0}} \equiv Equilibrium\ Temperature\ (GM)$$

$$8)\quad \Delta G_{(GM)A_{GM}M,T} = \Delta G_{A_{GM}M,T} - (\Delta G_{A_{GM},T} + \Delta G_{M,T})$$

$$9)\quad \Delta G_{Relat.(GM)A_iM,T} = \Delta G_{(GM)A_iM,T} - \Delta G_{(GM)A_{GM}M,T}$$

$$10)\quad \Delta G_{Relat.(GM)A_iM,T} = \underset{Equation\ 7}{\left(\Delta G_{A_iM,T} - (\Delta G_{A_{GM},T} + \Delta G_{M,T})\right)} - \underset{Equation\ 8}{\left(\Delta G_{A_{GM}M,T} - (\Delta G_{A_{GM},T} + \Delta G_{M,T})\right)}$$

$$Substitution\ of\ the\ equations\ 7\ and\ 8\ in\ equation\ 9$$

$$\boldsymbol{\Delta G_{Relat.(GM)A_iM,T} = \Delta G_{A_iM,T} - \Delta G_{A_{GM}M,T}}$$

$$\boldsymbol{11)\quad Simplified\ equation\ 10}$$

$$\boldsymbol{Relative\ Gibbs\ Free\ Energy\ (GM)}$$

The isomerization barrier that separates two conformations of the monomer can change due to the intermolecular interactions, which can stabilize the transition state. This effect is limited by many factors but the main one to consider is related to the position adopted during complexation, and if it is the same involved during the interconversion process. In fact, in this situation the isomerization barrier can change significantly, even it can decrease to the point where the isomerization can be achieved by collisions with the carrier gas, see Figure 2.3.3.2.3. This type of phenomenon is considered in Chapter 4 for the case of the β-β dimer (see Figure A4.4).

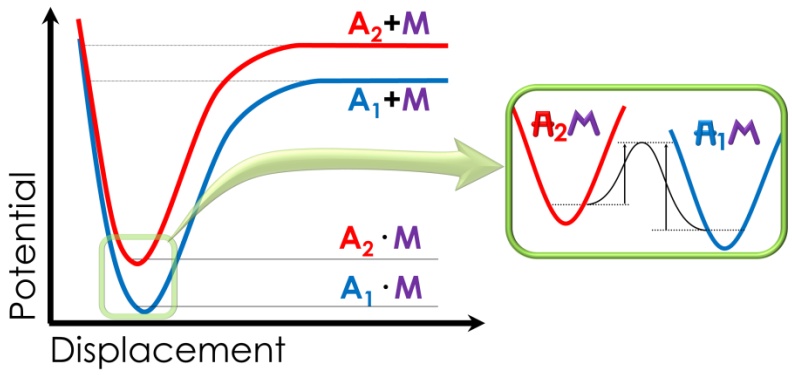


Figure 2.3.3.2.3: Interconversion pathways occurring after that the complex is formed.

2.3.3.2.1. Effect of the Temperature

The last point to analyze is the temperature. In Figure 2.3.3.2.1.1, the energy relative to the sum of the free monomers with respect to the complexes is represented. At high temperatures, the energy of the complexes is higher and the equilibria are shifted to the dimer fragmentation. Close to ΔG=0, the dimer starts forming, and below this point the formation of the complex is the most favored process. The supersonic expansion produces an adiabatic cooling, but the calculated Gibbs free energy data at the equilibrium do not represent this specific cooling process that occurs inside the expansion.[28] The Gibbs free energy data show which the most stable species at each temperature are, helping to understand which situation was frozen by the sudden cooling.

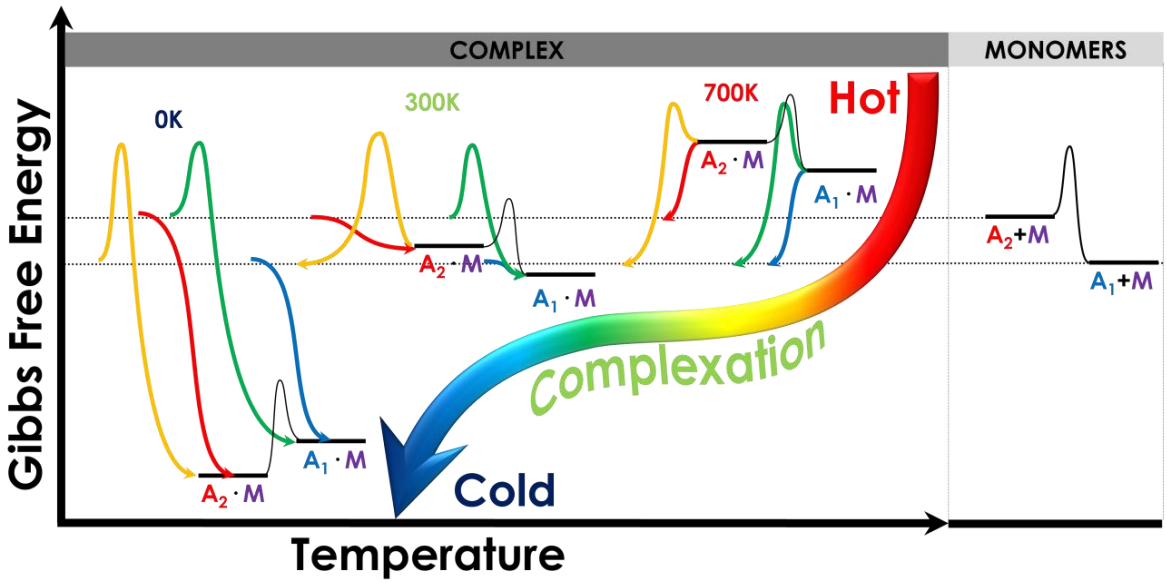


Figure 2.3.3.2.1.1: Effect of the temperature considering that the monomers remain energetically constant between them in all the temperature range. Diagram of the cooling process in the expansion. A2+M and A1+M represent the sum of the energies for the separated monomers. A2M and A1M represent the energies of the complex. The diagram compares the monomers with the dimers at several temperatures.

By using the equations (4) and (7), it is possible to represent the binding energy (Figure 2.3.3.2.1.2: a_1 and b_1). This is useful mainly for the MS calculation, and by exploiting this representation it is possible to compare the stability of different complexes. In order to obtain a better overview of the relative stability between complexes, it is necessary to use equations (6) and (11), and so calculate the relative data. The resulted diagrams show the relative binding Gibbs free energy for the MS methodology and the relative Gibbs free energy for the GB methodology, as reported in Figure 2.3.3.2.1.2: a_2 and b_2.

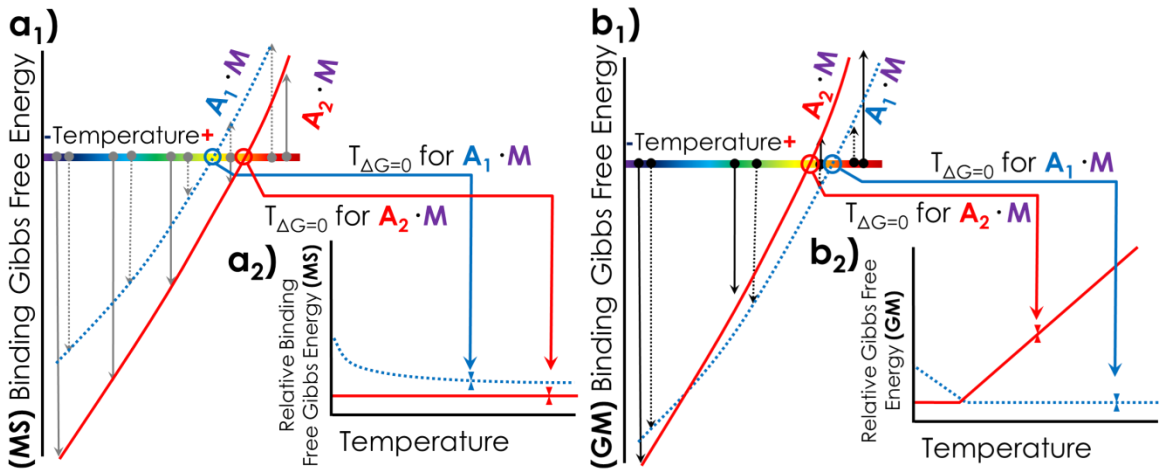


Figure 2.3.3.2.1.2: Two model diagram for the representation of the MS and GM binding Gibbs free energy data. **a_1)** MS Binding Gibbs Free energy diagram build by equation 4; **a_2)** MS Relative Binding Gibbs free energy diagram build by equation 6; **b_1)** GM Binding Gibbs free energy diagram build by equation 7; **b_2)** GM Relative Gibbs free energy build by equation 11. The usual temperature range is from 0K to 700K, when the organic matter decomposes. The diagram can be relative to all the isomers or a selection of the most representative ones.

2.4 References

(1) Demtröder, W. *Laser Spectroscopy. Basic Concepts and Instrumentation*; Springer, 2003.

(2) Levy, D. H. Laser Spectroscopy of Cold Gas-Phase Molecules. *Annu. Rev. Phys. Chem* **1980**, *31*, 197–225.

(3) De Vries, M. S.; Hobza, P. Gas-Phase Spectroscopy of Biomolecular Building Blocks. *Annu. Rev. Phys. Chem* **2007**, *58*, 585–612.

(4) Scoles, G. Atomic and Molecular Beam Methods. In *Atomic and Molecular Beam Methods*; Oxford University Press, 1988; Vol. 1, pp 14–53.

(5) Camiruaga, A.; Usabiaga, I.; Insausti, A.; León, I.; Fernández, J. A. Sugar–peptidic Bond Interactions: Spectroscopic Characterization of a Model System. *Phys. Chem. Chem. Phys.* **2017**.

(6) Leon, I.; Millan, J.; Cocinero, E. J.; Lesarri, A.; Castano, F.; Fernandez, J. A. Mimicking Anaesthetic-Receptor Interaction: A Combined Spectroscopic and Computational Study of Propofol···phenol. *Phys. Chem. Chem. Phys.* **2012**, *14* (25), 8956–8963.

(7) Vanommeslaeghe, K.; Guvench, O.; MacKerell, A. D. Molecular Mechanics. *Curr. Pharm. Des.* **2014**, *20* (20), 3281–3292.

(8) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120* (1–3), 215–241.

(9) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 Energy Evaluation by Direct Methods. *Chem. Phys. Lett.* **1988**, *153* (6), 503–506.

(10) Gaussian 09 Revision D0.1 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, "; M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. M.; G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. H.; A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. H.; M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. N.; Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J.; J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. B.; K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. N.; K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. T.; M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. C.; et al. Gaussian 09 Revision D0.1. *Gaussian, Inc., Wallingford CT* **2013**.

(11) Boys, S.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19* (4), 553–566.
(12) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. Macromodel—an Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics. *J. Comput. Chem.* **1990**, *11* (4), 440–467.
(13) Leon, I.; Montero, R.; Longarte, A.; Fernandez, J. A. Influence of Dispersive Forces on the Final Shape of a Reverse Micelle. *Phys. Chem. Chem. Phys.* **2015**, *17* (3), 2241–2245.
(14) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* **2005**, *26* (16), 1752–1780.
(15) Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17* (5–6), 490–519.
(16) CASE, D. A.; CHEATHAM, T. E.; DARDEN, T. O. M.; GOHLKE, H.; LUO, R. A. Y.; MERZ, K. M.; ONUFRIEV, A.; SIMMERLING, C.; WANG, B.; WOODS, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688.
(17) KIRSCHNER, K. N.; YONGYE, A. B.; TSCHAMPEL, S. M.; GONZÁLEZ-OUTEIRIÑO, J.; DANIELS, C. R.; FOLEY, B. L.; WOODS, R. J. GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates. *J. Comput. Chem.* **2008**, *29* (4), 622–655.
(18) Grimme, S. Density Functional Theory with London Dispersion Corrections. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1* (2), 211–228.
(19) Campargue, R. Progress in Overexpanded Supersonic Jets and Skimmed Molecular Beams in Free-Jet Zones of Silence. *J. Phys. Chem.* **1984**, *88* (20), 4466–4474.
(20) Patterson, D.; Doyle, J. M. Cooling Molecules in a Cell for FTMW Spectroscopy. *Mol. Phys.* **2012**, *110* (15–16), 1757–1766.
(21) Deckers, J.; Fenn, J. B. High Intensity Molecular Beam Apparatus. *Rev. Sci. Instrum.* **1963**, *34* (1), 96–100.
(22) Smalley, R. E.; Wharton, L.; Levy, D. H. Molecular Optical Spectroscopy with Supersonic Beams and Jets. *Acc. Chem. Res.* **1977**, *10* (4), 139–145.
(23) Kudoh, S.; Takayanagi, M.; Nakata, M. Conformational Cooling in a Supersonic Jet of 1,2-Dichloroethanestudied by Matrix Isolation Infrared Spectroscopy. *Chem. Phys. Lett.* **1998**, *296* (November), 329–334.
(24) Ruoff, R. S.; Klots, T. D.; Emilsson, T.; Gutowsky, H. S. Relaxation of Conformers and Isomers in Seeded Supersonic Jets of Inert Gases. *J. Chem. Phys.* **1990**, *93* (5), 3142–3150.
(25) Srinivasan, R.; Braren, B. Ultraviolet Laser Ablation of Organic Polymers. *Chem. Rev.* **1989**, *89* (6), 1303–1316.
(26) Lesarri, A.; Mata, S.; López, J. C.; Alonso, J. L. A Laser-Ablation Molecular-Beam Fourier-Transform Microwave Spectrometer: The Rotational Spectrum of Organic Solids. *Rev. Sci. Instrum.* **2003**, *74* (11), 4799–4804.
(27) Desyatnyk, O.; Pszczolkowski, L.; Thorwirth, S.; Krygowski, T. M.; Kisiel, Z. The Rotational Spectra, Electric Dipole Moments and Molecular Structures of Anisole and Benzaldehyde. *Phys. Chem. Chem. Phys.* **2005**, *7* (8), 1708–1715.
(28) Godfrey, P. D.; Brown, R. D. Proportions of Species Observed in Jet Spectroscopy−Vibrational Energy Effects: Histamine Tautomers and Conformers. *J. Am. Chem. Soc.* **1998**, *120* (41), 10724–10732.
(29) Çarçabal, P.; Cocinero, E. J.; Simons, J. P. Binding Energies of Micro-Hydrated Carbohydrates: Measurements and Interpretation. *Chem. Sci.* **2013**, *4* (4), 1830–1836.

3. Intermolecular Interactions in Glucose · Phenol Complexes: the "*Proven Survivors*"

3.1. Introduction and Objectives

The biological world is a metastable equilibrium of a complex network of interactions and interconnections at molecular level.[1] One of the consequences of this interconnection is that obtaining a complete picture of a living system is a huge and challenging task. The more recent studies use complex mathematical algorithms combined with the most advanced imaging spectroscopic methods and other state-of-art techniques, in order to find out a general scheme of the behavior of some biochemical systems.[2] The inferred probabilistic connections obtained in this way, are compared with the well-known biochemical reactions and with the concentration of metabolites in the cell at certain instants in order to validate the analysis.[3] Nowadays, a full and reliable computational simulation of such systems is not possible, since they are constituted of multiple dynamic phenomena, many of them still unknown.[4] In this intricate environment, the molecular complexes are forming and breaking constantly.

Certainly, nature relies in the non-covalent interactions between the biological molecules to build the intricate mechanism that results in what we call life.[5] Such weak interactions allow living organism to maintain their metabolism working with a minimum energy exchange. For example, many processes are triggered when two molecules, a signaling molecule and a receptor, interact and form a complex that induces a conformational change in the receptor, starting a signaling cascade. As the two molecules are held together by weak interactions (van der Waals and H-bonds), at some point the complex dissociates again, terminating the process and leaving the system ready after a short recovering time to start the process again when necessary. If the signaling molecule and the receptor were bonded by covalent interactions, such recovery would be significantly more difficult or even impossible. Actually, there are many toxic compounds whose side effects derive precisely from irreversible reactions. The formation of complexes with a high chemical affinity can often interfere on the physical chemical balance, producing in many cases a dramatic collapse.[6]

As stated above, generally, biological reactions are produced by transient interactions. An example is represented by the anesthetic effects caused by propofol: when it attaches the $GABA_A$ channel in the brain leaving it open, it blocks the pain signal allowing Cl^- to enter the post-synapses and compensating the aperture of the Na^+ channels; as soon as its concentration in the blood decreases and the drug leaves the active center, the effect disappears.[7] Administering the correct amount of propofol to a patient results in sedation.

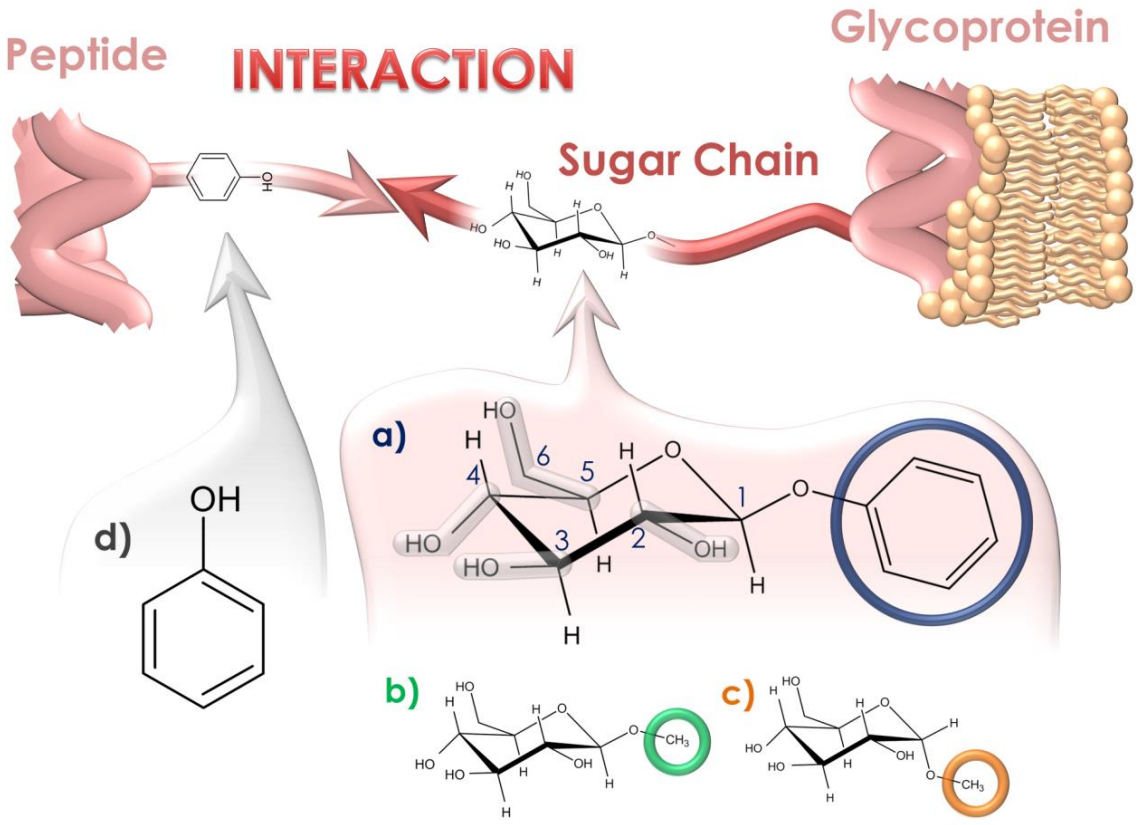
However, an overdose of propofol causes death, due to its (lower) affinity for the adrenergic receptors. Put in other words, the first molecules of propofol in the blood are able to find their way to the $GABA_A$ channels; too many molecule of propofol start attaching to other receptors, in this case, to the adrenergic receptors. Actually, the most difficult part in the intelligent design of new drugs is not the design of a ligand that targets the desired protein, but to avoid any other interaction with the other +100,000 different proteins also present and will presumably result in (undesired) side effects. The ability to give rise to a subtle extra stabilization with respect to any other interaction, could represent the determinant factor that confers to the molecule an active biological function.[8] Actually, the full understanding of the mechanisms of these reactions is fundamental in order to influence or even control them, as it happens, for example, in the repair of a disequilibrium in the human body caused by the deficit of a hormone, or when the docking of a virus to its target cells is avoided.

Several techniques are able to disentangle specific information regarding these interactions. Among them, crystallography is probably the most broadly used. It enables to determine the structures of the molecules or even of molecules catch in the interaction.[9,10] However, it requires formation of crystals, solid state highly-ordered structures in which the molecular conformation is the result of a collapse into a solid. The data obtained with this technique are very useful, but the connection with the real biological environment is not straightforward, as the interaction distances and conformations may not correlate. Thus, it would be more convenient to study these systems in solution. However, the intrinsic dynamics of solution complicates the determination of a single structure and only an averaged conformation can be obtained.

In principle, it could seem that study of biological systems in the isolated phase is simply not possible, because it imposes radically different conditions to those in biological media. However, it also presents some very interesting characteristics.[11] Using supersonic expansions, it is possible to create the conditions to form isolated molecular aggregates whose structure is not perturbed by any additional interaction and therefore, the true conformational preferences of the molecules may be unraveled, using a combination of powerful spectroscopic techniques and quantum-mechanical calculations. With these tools, the true nature of the interactions can be revealed. Furthermore, the supersonic jet expansion presents similar balances during the cooling with respect to other phases. In fact, the molecules are heated (either thermally or by laser desorption processes) and transferred to gas phase. In this way, the energy carried to the system allows the

molecules to populate and so map the entire conformational landscape up to very high temperatures. The subsequent sudden cooling produced by the collisions in the nozzle freezes the molecules in a variety of situations that reflect the initial ones and may be far away from equilibrium.[12] Thus, the conformational isomers that survive the cooling stage always contain the most stable structure, but may also include others relatively high in energy, depending on the barriers connecting them with other more stable minima. Each of those structures is an invaluable piece of information that allows one not only to understand the forces at play but also to discover the competition that rules and shapes the aggregate. Each structure is like a high-definition snapshot of the multiple conformations that the system may adopt in solution. Interestingly, some of those (in principle) less stable structures, represent more tightly-bonded conformations, although the extra interaction energy comes to the expenses of adopting less stable conformations of the interacting molecules. This phenomenon can be the consequence of strong intermolecular bonds, which increase the probability of these clusters to survive to the collisions in the expansion, and this can be consider a benefit in other highly competitive environments, as it happens in the biological systems, resulting in a small increase of the aggregation lifetime. The change on the strength of the interactions proves to be a key factor in the biological balances, as the kinetics of the processes play an important role, and the possibility, under the appropriate conditions, of isolating one of these types of conformations can shed light on the mechanisms underlying the process.

Following the above considerations, we deal in this chapter with the interactions between glucose and phenol as a model system of the sugar-tyrosine interaction.[13,14] Although it is a simplified molecular model, the interaction results in a complex potential energy surface with numerous relative minima of similar stability that will challenge both the experimental methodology and computational models. Glucose is one of the most abundant carbohydrates and actually, it presents a rich conformational landscape with several conformers, differing in the orientation of the hydroxyl and hydroxymethyl groups.[15] Scheme 3.1 shows the structure of phenol and the sugars used in this experiment. Both anomers of methyl-D-glucopyranose (α/β-MeGlc) and the β anomer of phenyl-D-glucopyranose (β-PhGlc) were complexed with phenol (P). The former two systems will serve to study the differences introduced by the orientation of the anomeric conformation, while the latter interacting pair will allow us to evaluate the influence of a more voluminous substituent in the intermolecular interactions, resembling what happens in the interaction of glucose oligomers with other molecules.



Scheme 3.1: Phenol and sugar interaction: **a)** phenyl-β-D-Glucopyranose (β-PhGlc), in blue the numbering of the glucoses atoms. Grey bars identify the flexible parts of D-Glucoses. **b)** Methyl-β-D-Glucopyranose (β-MeGlc). **c)** Methyl-α-D-Glucopyranose (α-MeGlc). **d)** Phenol (P).

3.2. Experimental Results

Figure 3.2.1 shows the REMPI spectra of the two monomers that present optically-accessible transitions, P and β-MeGlc, and of the heterodimers formed in this work, obtained using the experimental conditions described previously in Chapter 2. Since there are two chromophores, the spectral region can be divided into two sections, one for each chromophore (the phenyl group of β-PhGlc and P). Although in principle one could think that P is a better chromophore, the stronger intensity of the transitions of β-PhGlc···P dimer in the 36700-37100 cm^{-1} region, which most likely corresponds to the excitation of the 0_0^0 transition of β-PhGlc, seem to indicate that this is not the case. The relative intensity of the signals in the two chromophores' spectral regions demonstrates that the signal produced by the phenyl chromophore of β-PhGlc is stronger than that from P. The peak observed around 36800 cm^{-1} in methyl-β-D-glucopyranose · phenol (β-MeGlc·P) and methyl-α-D-glucopyranose · phenol (α-MeGlc·P) spectra arises mainly from the fragmentation of higher mass complexes (Figure A3.1-b) and Figure A3.2-b)). A

more detailed explanation of these spectra can be found in the already published paper.[16]

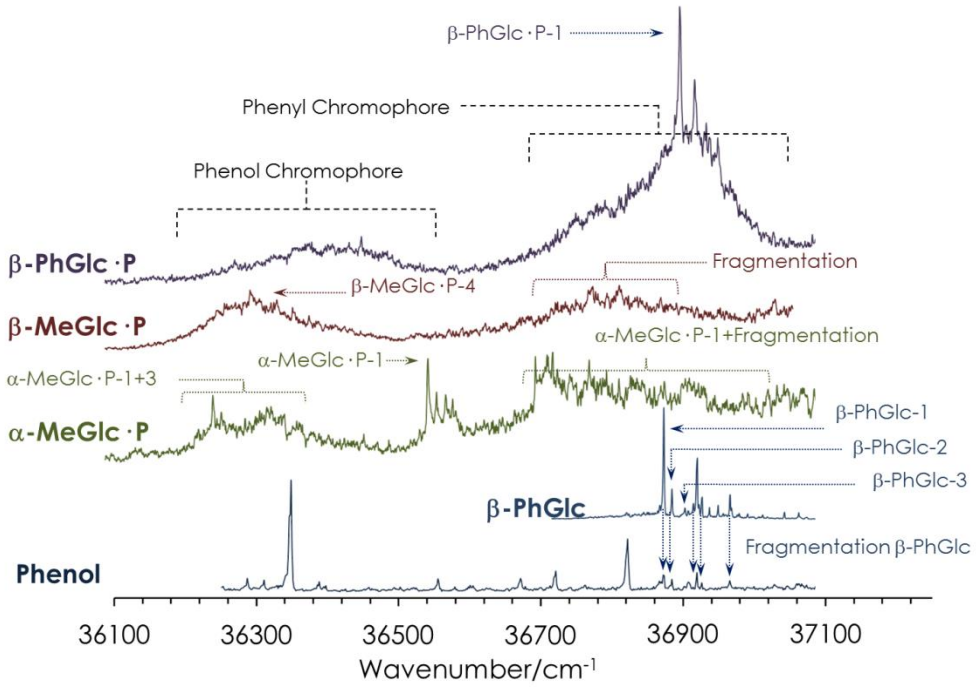


Figure 3.2.1: REMPI spectra of Phenol (P), Phenyl-β-D-Glucopyranose (β-PhGlc), Methyl-α-D-Glucopyranose · Phenol (α-MeGlc·P), Methyl-β-D-Glucopyranose · Phenol (β-MeGlc·P), and Phenyl-β-D-Glucopyranose · Phenol (β-PhGlc·P). The spectra of P and β-PhGlc monomers are in good agreement with previous publications.[17,18]

The mass-resolved IR spectra obtained tuning the UV laser at certain wavenumbers of the REMPI spectra of Figure 3.2.1, may be found in Figure 3.2.2. The spectra of the monomers are used to calibrate the correction factors to account for the anharmonicity. Certainly, the simulated spectra are constructed using the normal mode analysis and therefore they are based on harmonic calculations. However, different groups present a different level of anharmonicity and it changes depending on the environment in which the chemical group is found. For example, the formation of H-bonds results in an increase of the anharmonicity, qualitatively proportional to the strength of the H-bond. The region scanned in Figure 3.2.2 contains CH and OH groups and they present different anharmonicity, making necessary the use of two different correction factors: 0.9530 for CH stretch vibrations and 0.9385 for OH vibrations. These factors were calibrated using the spectra of the monomers and on several additional molecules of unequivocal assignment.

Analysis of the monomer's spectra is particularly useful for the next assignment of the complexes. Actually, the spectrum of a molecular aggregate can be considered as the spectrum of the chromophore perturbed by the interacting partner in the case of REMPI spectroscopy, or as the sum of the spectra of the individual molecules perturbed by the interaction in the case of the IR spectra.

Two free OH regions may be distinguished in the spectra of Figure 3.2.2: the 3610-3640 cm^{-1} region corresponds to the glucose OH stretches. These groups are not hydrogen bonded due to structural constrains, but they present noticeable interactions and are forming a cooperative interaction network. For the sake of brevity, we will refer to this network of interactions as a hydrogen bond network.

On the other hand, the single OH stretch of P appears at ~3650 cm^{-1}, in good agreement with previous publications.[17] Any interaction that these OH moieties may present in the complex, will result in shifts in their position that will correlate with the strength of the interactions. Certainly, the band of the phenolic OH appears shifted to the red in the spectra of the dimers, (Figure 3.2.2-e) -h)), indicating that this moiety is involved in moderate/strong interactions.

Regarding the rest of the bands, all the spectra of the complexes in Figure 3.2.2 present the same number of OH groups and therefore, they should contain the same number of bands. The position of the bands varies between aggregates, depending on the interactions that take place in each system. Such variation is a valuable source of information to assign the spectra to the computed structures.

A large number of calculations were carried out to locate all the relevant minima in the potential energy surface (PES). Presenting all the theoretical results along the text would be impractical, and therefore they are summarized in the Appendix (Figures: A3.1-d), A3.2-c), A3.3-d)). Moreover, in the Appendix (Figures: A3.1-c), A3.2-b), and A3.3-c)), the comparison between the experimental spectra and those predicted for the most relevant structures of each family of computed structures may be found. For the sake of brevity, we will present here only the comparison between the experimental and computed spectra for those structures to which they were assigned.

Two different IR/UV spectra of α-MeGlc·P complex were recorded, depending on the wavenumber probed with the UV laser (Figure 3.2.2-f), -g)). Comparison of those two spectra with the predicted ones (see Figure A3.1) allowed us to assign the experimental spectra to isomers α-MeGlc·P-1 and -3. The two families differ in

the relative position of the two interacting molecules: phenol interacts with the oxygen 6 ($O_{PhH}\cdot O6$) of glucose in the former, while it inserts its OH group between O4 and O6 in the latter ($O4H\cdot O_{PhH}\cdot O6$). In addition, the two different intermolecular interactions result in very different hydrogen bond networks, as the OH groups orient clockwise in the former (α-MeGlc·P-1) and anti-clockwise in the latter (α-MeGlc·P-3). Thus, the cooperative network acts as an amplifier of the intermolecular interactions, increasing its effect in the final shape of the aggregate.

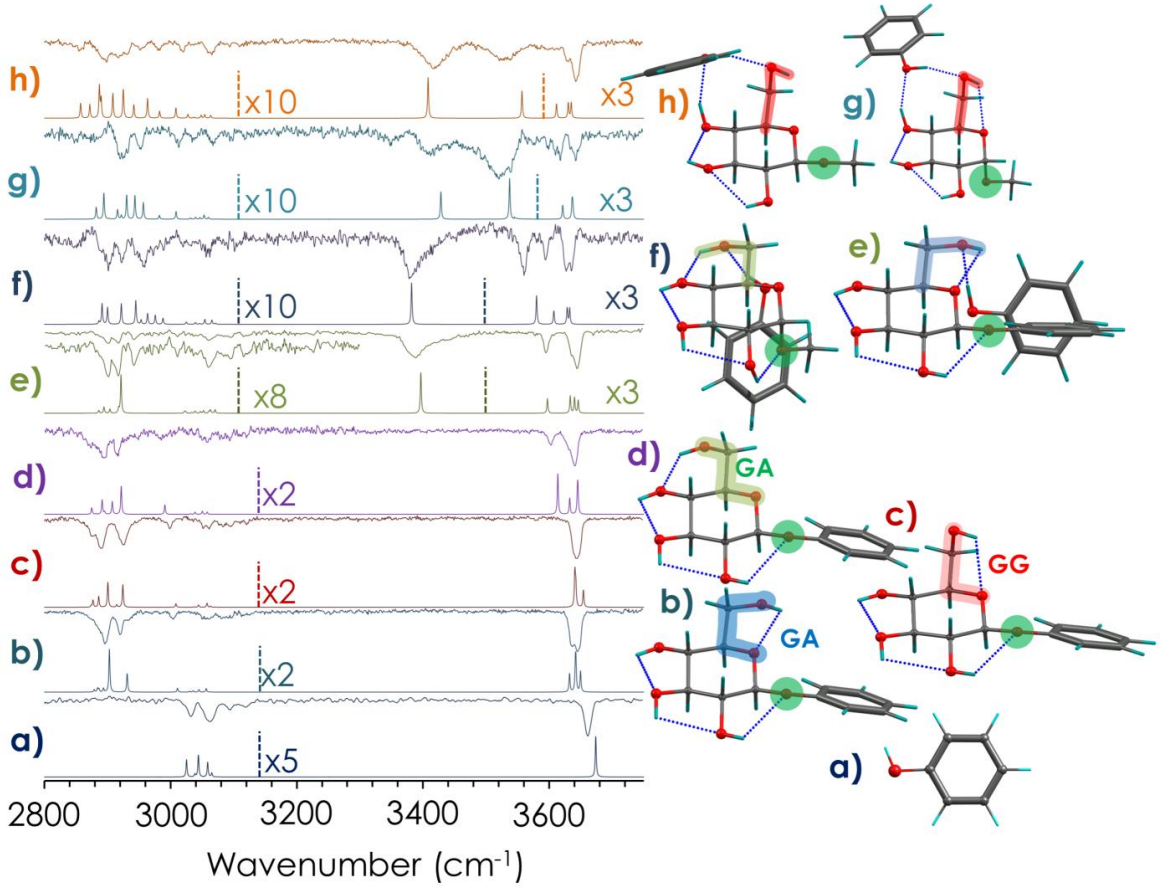


Figure 3.2.2: IDIRS spectra of: **a)** phenol **b)** β-PhGlc-1 **c)** β-PhGlc-2 **d)** β-PhGlc-3 **e)** β-PhGlc·P-1 **f)** α-MeGlc·P-1 **g)** α-MeGlc·P-3 **h)** β-MeGlc·P-4.

Two systems containing a β anomer were studied: β-PhGlc·P (Figure 3.2.2-e)) and β-MeGlc·P (Figure 3.2.2-h)). No significant changes in the shape of the spectra were observed probing different parts of the REMPI spectra with the UV laser. Furthermore, a single isomer is sufficient to explain each of the two spectra (see Figures A3.2 and A3.3 for more detailed information). In the case of β-PhGlc·P, the experimental spectrum matches very well that predicted for the global minimum, in which P interacts with the hydroxymethyl group of glucose ($O_{PhH}\cdot O6$) and at the

same time both aromatic rings interact. Probably, this latter C-H···π interaction is decisive to give to the complex its final shape.

Regarding β-MeGlc·P, the assigned family is not the global minimum, but matches with the second most stable family of structures observed for α-MeGlc·P (O4H· O_{Ph}H· O6). All the data for these two complexes are reported in the Appendix for Chapter 3, in Figures A3.2 and A3.3 for β-PhGlc·P and β-MeGlc·P, respectively.

3.3. Discussion

As described above, for what concerns the assigned structures of β-MeGlc·P complex, it is worth noting that they correspond to the second family (O4H· O_{Ph}H· O6) and thus do not represent the global minimum. This result could be explained with a missing detection of the global minimum caused by a change in the excited state conformation, with the subsequent reduction in the Franck-Condon window for detection. But on the other hand, the first family (O_{Ph}H· O6) was detected for the other type of aggregates analyzed. Since they present similar structures to that of β-MeGlc·P, one should in principle expect them to have similar Franck-Condon windows and therefore it should allow their detection. Therefore, there must be other reasons for the absence of the detection of the global minimum.

In Figure A3.4, the diagrams of the relative Gibbs Free energy are shown. As mentioned in Chapter 2, the data can be analyzed in two ways: one can represent the Gibbs relative free energy, taking the ΔG of the most stable species as the minimum, and subtracting it from the rest of the isomers (labeled as GM in Figure A3.4). The second method uses the ΔG of the binding of the most tightly bonded isomer and subtract such value from the rest (labeled as MS in Figure A3.4). While the former diagrams show the variation of the relative stability of the different families with the temperature, the latter give the dependence of the interaction energy with the temperature across those families and it is useful in a situation where many species present similar stability. In fact, in such cases, the most tightly bonded species may represent the preferred conformation.

Thus, Figure 3.3 shows the analysis of ΔG for the four aggregates studied in this section, while the complete study may be found in Figure A3.4. According to the data in Figure 3.3, two different situations may exist. For β-PhGlc·P, the global minimum is somehow more stable than the rest of the species (Figure 3.3-1, dark blue section), although it presents a similar dissociation energy than β-PhGlc· P-3

(see MS diagram in Figure A3.4-c$_2$)). Since only the isomer 1 was detected, it seems that formation of this cluster is governed by the relative stability and not by the binding energy.

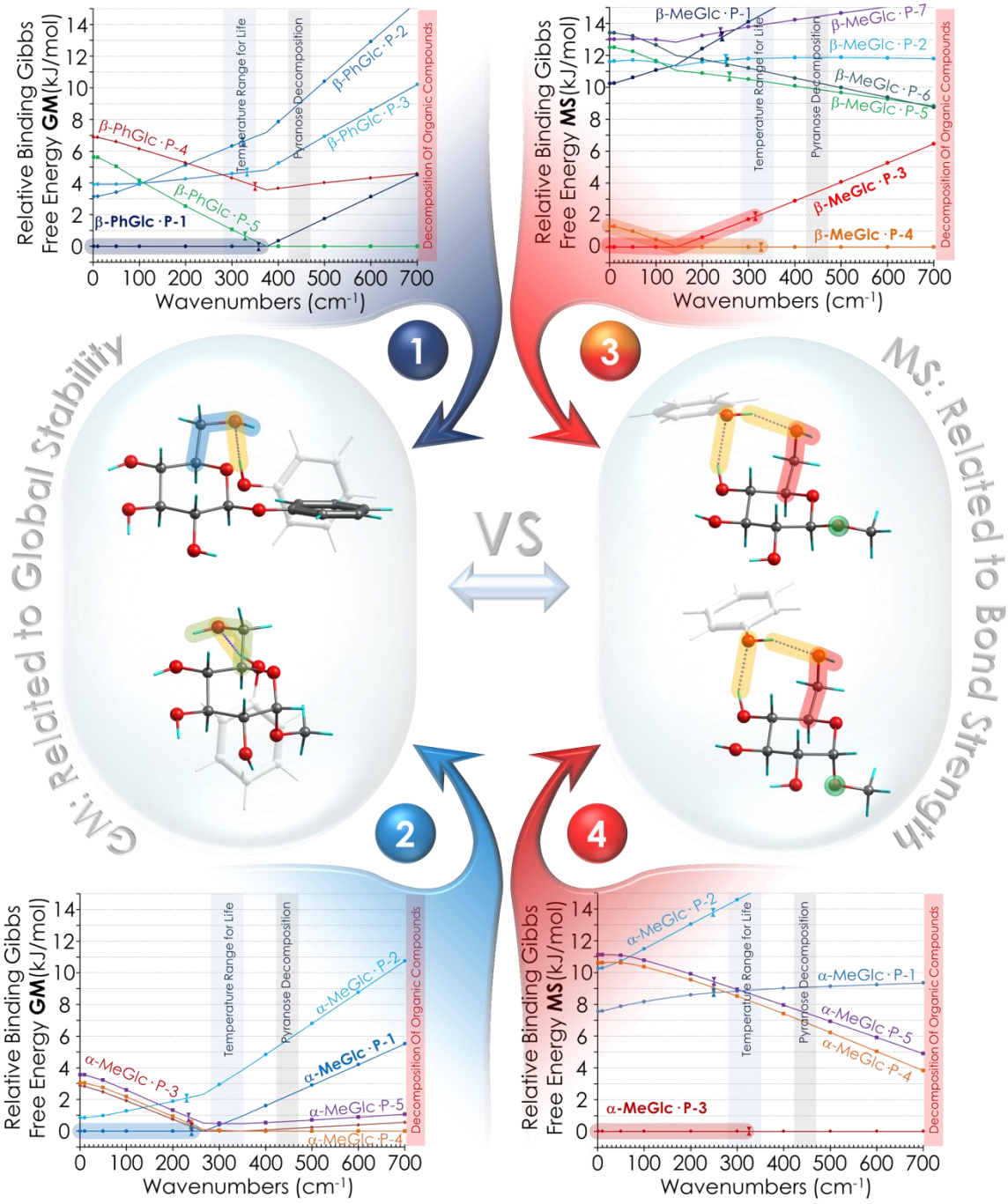


Figure 3.3: First Family 1&2) β-PhGlc·P and α-MeGlc·P dimers and the Gibbs Free Energy diagrams calculated with respect to the Global Minimum monomers (GM); **Second Family 3&4)** β-MeGlc·P and α-MeGlc·P dimers and the Gibbs Free Energy diagrams calculated with the Most Similar monomers with the geometry adopted in the complex (MS).

In the case of β-MeGlc·P dimer, the situation is reversed: the GM diagram (Figure A3.4-b_2)) reveals a competition between several candidates, as observed also in the MS diagram of β-MeGlc·P (Figure A3.4-b_1)). Nevertheless, in this case the MS diagram highlights a relevant energy difference of ~10 kJ/mol between the representative structures of the second family β-MeGlc·P-3 & -4 and the rest of conformations in the whole interval of temperature (see Figure 3.3-3 and -4, orange and red sections). This energy difference is larger than the expected errors (~4 kJ/mol) usually achieved by these type of calculations.[19] Moving to the comparison between experimental and computed spectra, it can be noted that the β-MeGlc·P-3 & -4 are the most probable candidates for the assignment. Furthermore, the only difference between both structures lies in the rotation of the phenyl group with respect to the sugar molecule, and therefore it will relax in the expansion, giving rise to only one of the two isomers. Thus, both the comparison with the computed spectra and the analysis of the relative stability and dissociation Gibbs free energy agree with this assignment. For all these reasons, it can be stated that in this case, the decisive factor in the selection of species during the aggregation process is the binding energy, and not the relative stability. In a situation where several structures have similar stability, those conformers that present greater binding energy seem to be favored during the cooling process. Under equal stability conditions, a larger binding energy appears to be an "evolutionary advantage" to survive the "natural selection" of the cooling process.

Regarding α-MeGlc·P, a more complex situation was observed: the GM diagram shows that the assigned α-MeGlc·P-1 conformer is slightly more stable, although several other structures are close in energy in the temperature gap of our expansion (Figure 3.3-2, light blue section). However, the MS diagram shows a clear dominance of the assigned α-MeGlc·P-3 conformer (Figure 3.3-4, red section) with a relative ΔG of binding significantly larger, > 7 kJ/mol, than the next most tightly-bonded conformer. In this system, both the most stable and the most tightly bonded conformers were found in the expansion, and they correspond to different families and therefore, high potential energy barriers are expected for the isomerization process.

Summarizing, the observed trends are compared in the same Figure 3.3, in which the β-XxGlc·P dimers are related to the two assigned conformations of the α-MeGlc·P cluster. Basically, the assigned structures belong to two families: the first one ($O_{Ph}H \cdot O6$) is built on the sugar's most stable conformer, and the second one ($O4H \cdot O_{Ph}H \cdot O6$) correspond to the preferred interaction site. Therefore, the first family is the product of a global energy stabilization (both of the interacting

molecules and the cluster), while the second family comes from the higher binding energy that, through a selection mechanism at the intermolecular bond level, produced an extra stabilization enabling the formation and detection of these conformers in the CI-phase.

3.4. Conclusions

The relative stabilization of a complex is one of the most important parameters to take into consideration in gas phase, in order to achieve a correct assignment. In this study, a new and interesting criterion has been revealed, which is the competition between the most stable and most tightly-bonded structures in the formation of the clusters. Actually, inside a molecule all the bonds are formed by covalent, strong interactions, but in a complex the interactions are due to a combination of some weak non-covalent bonds. The maximization of these secondary interactions and the correlated stabilization of the complex, confer to the latter the ability to survive in the CI-phase through the expansion, and sometimes also to coexist with more stable structural arrangements (as demonstrated for the clusters analyzed in this section). This interacting mechanism, which gives rise to a stabilizing balance of the chemical forces, could be one of the parameters that play a determinant role in the biological activity, and in the responses produced by drugs. In fact, a stronger binding interaction can provide to the complexes the ability to survive in many different environments. These "*proven survivors*" complexes, in highly competitive environments like the intra- and intercellular media, can sometimes overcome the global minimum conformation, becoming the real rulers of the biochemical reactions. Through the systematic study here reported this class of "*proven survivors*" clusters was isolated and detected.

3.5. References

(1) Dill, K.; Bromberg, S. *Molecular Driving Forces. Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience*, 2nd ed.; Taylor & Francis, 2010.

(2) Gamez-Pozo, A.; Berges-Soria, J.; Arevalillo, J. M.; Nanni, P.; Lopez-Vacas, R.; Navarro, H.; Grossmann, J.; Castaneda, C. A.; Main, P.; Diaz-Almiron, M.; Espinosa, E.; Ciruelos, E.; Fresno Vara, J. A. Combined Label-Free Quantitative Proteomics and microRNA Expression Analysis of Breast Cancer Unravel Molecular Differences with Clinical Implications. *Cancer Res.* **2015**, *75* (11), 2243–2253.

(3) Dweep, H.; Sticht, C.; Pandey, P.; Gretz, N. MiRWalk - Database: Prediction of Possible miRNA Binding Sites by " Walking" the Genes of Three Genomes. *J. Biomed. Inform.* **2011**, *44* (5), 839–847.

(4) Young, D. C. *COMPUTATIONAL DRUG DESIGN A Guide for Computational and Medicinal Chemists*; WILEY, 2009.

(5) Riley, K. E.; Hobza, P. Noncovalent Interactions in Biochemistry. *Wiley Interdiscip. Rev. Comput.*

Mol. Sci. **2011**, *1* (1), 3–17.
(6) Nelson, D. L.; Nelson, D. L.; Lehninger, A. L.; Cox, M. M. *Lehninger Principles of Biochemistry*; W.H. Freeman: New York, 2008.
(7) Trapani, G. M.; Altomare, C.; Sanna, E.; Liso, G. B. and G. Propofol in Anesthesia. Mechanism of Action, Structure-Activity Relationships, and Drug Delivery. *Current Medicinal Chemistry*. 2000, pp 249–271.
(8) Jorgensen, W. L. The Many Roles of Computation in Drug Discovery. *Science* **2004**, *303* (5665), 1813–1818.
(9) Garman, E. F. Developments in X-Ray Crystallographic Structure Determination of Biological Macromolecules. *Science (80-.).* **2014**, *343* (6175), 1102 LP-1108.
(10) Glaeser, R. M. Review: Electron Crystallography: Present Excitement, a Nod to the Past, Anticipating the Future. *J Struct Biol* **1999**, *128* (1), 3–14.
(11) De Vries, M. S.; Hobza, P. Gas-Phase Spectroscopy of Biomolecular Building Blocks. *Annu. Rev. Phys. Chem* **2007**, *58*, 585–612.
(12) *Jet Spectroscopy and Molecular Dynamics*; Hollas, J. M., Phillips, D., Eds.; 1995.
(13) León, I.; Millán, J.; Cocinero, E. J.; Lesarri, A.; Castaño, F.; Fernández, J. A. Mimicking Anaesthetic-Receptor Interaction: A Combined Spectroscopic and Computational Study of Propofol···phenol. *Phys. Chem. Chem. Phys.* **2012**, *14* (25), 8956–8963.
(14) Usabiaga, I.; Gonzalez, J.; Arnaiz, P. F.; Leon, I.; Cocinero, E. J.; Fernandez, J. A. Modeling the Tyrosine-Sugar Interactions in Supersonic Expansions: Glucopyranose-Phenol Clusters. *Phys. Chem. Chem. Phys.* **2016**, *18* (18), 12457–12465.
(15) Alonso, J. L.; Lozoya, M. a.; Peña, I.; López, J. C.; Cabezas, C.; Mata, S.; Blanco, S. The Conformational Behaviour of Free D-Glucose—at Last. *Chem. Sci.* **2014**, *5* (2), 515.
(16) Usabiaga, I.; González, J.; Arnáiz, P. F.; León, I.; Cocinero, E. J.; Fernández, J. A. Modeling the Tyrosine-Sugar Interactions in Supersonic Expansions: Glucopyranose-Phenol Clusters. *Phys. Chem. Chem. Phys.* **2016**, *18* (18), 12457–12465.
(17) Billes, F.; Mohammed- Ziegler, I. Vibrational Spectroscopy of Phenols and Phenolic Polymers. Theory, Experiment, and Applications. *Appl. Spectrosc. Rev.* **2007**, *42* (4), 369–441.
(18) Talbot, F. O.; Simons, J. P. Sugars in the Gas Phase: The Spectroscopy and Structure of Jet-Cooled Phenyl β-D-Glucopyranoside. *Phys. Chem. Chem. Phys.* **2002**, *4* (15), 3562–3565.
(19) Marianski, M.; Supady, A.; Ingram, T.; Schneider, M.; Baldauf, C. Assessing the Accuracy of Across-the-Scale Methods for Predicting Carbohydrate Conformational Energies for the Examples of Glucose and α-Maltose. *J. Chem. Theory Comput.* **2016**, *12*, 6157–6168.

4. Equatorial β-D-glucopyranose: Structural Properties in One of the Most Abundant Carbohydrates

4.1. Introduction and Objectives

Sugars are very versatile molecules: they can act as structural support in cell walls[1] or in DNA,[2,3] or they may serve to label cells in molecular recognition processes. Certainly, they are commonly found bound to proteins and/or lipids outside the plasmatic membranes.[4,5,6] Like in DNA or proteins, where the code is composed of different letters of an alphabet of DNA bases or amino acids, in the case of carbohydrates, the "message" is codified in the orientation of the hydroxyl groups of five- or six-membered rings. The different types of rings in which a given sugar may be found are also classified as *epimers*, according to the equatorial/axial sequence of their hydroxyl substituents. The epimers also admit many modifications, such as sialylation or acetylations, increasing even further the number of possible letters. One of the most important epimer is β-D-glucopyranose, as it is one of the most abundant sugars in nature, where it plays several relevant functions.

The objective of the study here reported is the evaluation of the properties that make β-D-glucopyranose different from the rest of pyranosyl carbohydrates. Indeed, it is the only sugar that presents all the hydroxyl groups of the ring in equatorial position.

By using an experimentally-validated theoretical methodology, its potential energy surface (PES) was analyzed. The very high flexibility of the molecule makes not possible a simple 3D representation. For this reason the interconversion pathways between the most stable structures of β-D-glucopyranose were investigated. This type of analysis is far from the full interpretation of its PES, but it is very useful in order to understand the conformational behavior of the molecule.

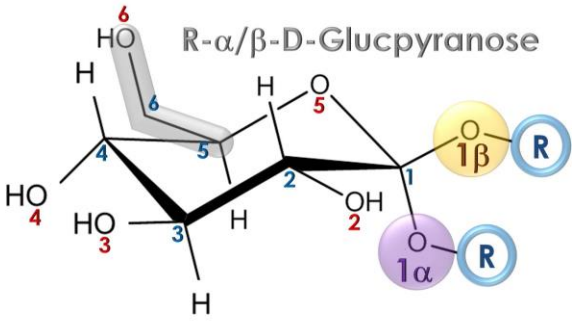
Once the monomer's structural properties were determined, and based on the information obtained, the study was directed to deepen into the aggregation process. In order to evaluate the behavior of β-D-glucopyranose, several dimers mixing β- and α-anomer were considered, paying special attention to the effect of the equatorial hydroxyls position on the complexation process. For this analysis, the Gibbs free energy of binding diagrams were used in order to elucidate the existence of particular energetically determinant trends for the dimerization processes.

Finally, many crystallographic matrices (intermolecular and intramolecular properties of the crystal were taken into account) of glucose were decomposed:

monosaccharides, disaccharides and polysaccharides. A special effort was done to understand the structure of cellulose, and the interactions that form the crystal. This polysaccharide is very abundant in nature and it usually plays a structural function in vegetables, where the cellulose crystals are organized in stacked layers, formed by cooperative hydrogen bond networks, able to add intra/intermolecular stabilization to the cellulose chains. From this whole picture, arises the interest to investigate if a relation between these equatorial hydroxyl groups and the macromolecular properties of cellulose exists. The final purpose of this chapter is to connect the orientation of the hydroxyl groups and the hydrogen bond networks in the crystal with the aggregation process and with the intrinsic properties of β-D-glucopyranose.

4.2. β-D-glucopyranose Monomers: Conformational Analysis

The PES analysis is a very common procedure,[7] that usually helps to extract information on the chemical properties of molecules. However, due to their complexity, this type of calculations is limited to very simple systems, where the conformational flexibility of the molecule is defined by a few degrees of freedom (DF). Some examples are propofol[8,9] and the diol 1,2-butanediol,[10] in which there are more than two DF describing their conformational landscapes, complicating the study and producing multidimensional PES. The more complicated the molecular system is, the more simplifications are required in order to limit the number of possible structural representations and focus on the most important information.



Scheme 4.1: Structure and labeling of glucopyranose. The blue numbers refer to the carbon atoms, while the red ones to the oxygen atoms, except for the anomeric position that is expressly highlighted.

Regarding glucopyranose, it has four hydroxyl and one hydroxymethyl substituents, each of which has several DFs (see Scheme 4.1 and Scheme A4.1 in the Appendix for Chapter 4 for more details) resulting in an overwhelming number of possible combinations. In order to overcome this problem, the molecule can be divided in different sections, the stability of each of which can be analyzed separately. In this way, the number of calculations needed is significantly reduced.

The first parameter to be taken into account is the heteroatomic six-membered ring, which can adopt either chair or twist (boat) conformations. However, the large energy difference between these two conformations allows one to consider only chair conformers, simplifying the situation (see Figure A2.3 in the Appendix for Chapter 2 for more details). On the other hand, two are the possible simplifications with regard the hydroxyl groups bound to the ring: the first one is to consider the isomerization pathways between the full interconversion of all hydroxyls structures from one direction (β-Glc-03,04 conformations) to the other (β-Glc-09,10 conformations), crossing through the intermediate local minima (β-Glc-05,06,07,08 conformations). The corresponding isomerization barriers are significantly reduced if the intermediate geometry presents an eclipsed structure with respect to a hydrogen atom (Diagram 4.1, orange bar), rather than with a carbon of the aliphatic ring (Diagram 4.1, blue bars). The second simplification is related to the aliphatic carbons of the pyranose ring. All of them have a tetrahedral structure, which splits the conformational minima into three possibilities: Gauche-Anti 1 (GA1), Gauche-Anti 2 (GA2), and Gauche-Gauche (GG), however the latter does not present true minima (Diagram 4.1, red bar), and therefore, only two possible conformations (GA1 and GA2 in Diagram 4.1, green bars) have to be taken into account, with a single preferential interconversion pathway: eclipsed-H, the orange bar in Diagram 4.1.

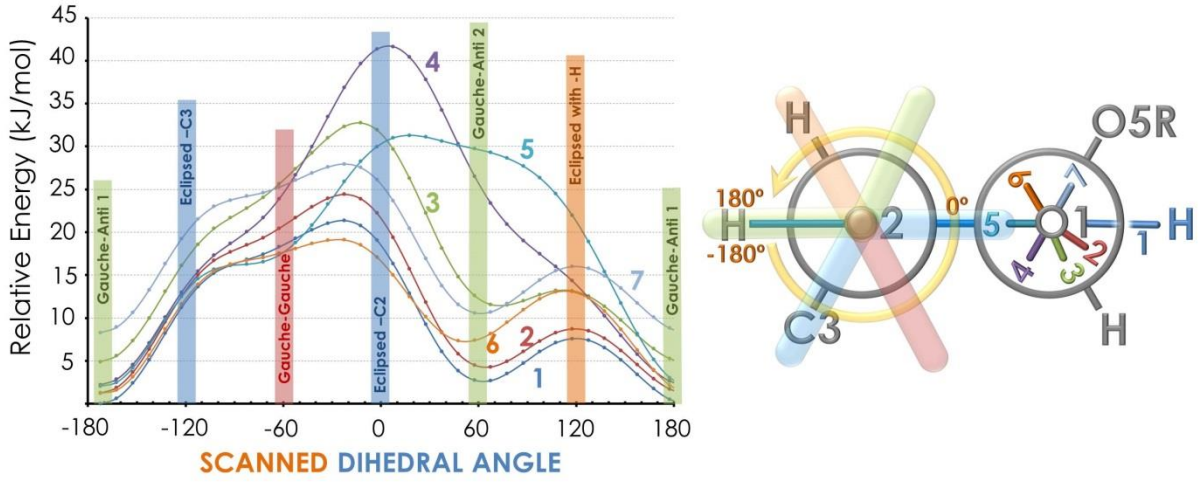


Diagram 4.1: Newman projection of O2-C2 and O1-C1 of β-D-glucopyranose. The first dihedral angle (O5-C1-O1-H) was maintained fixed while the second OH dihedral angle (C1-C2-O2-H) was scanned. The scan was performed in seven different positions relative to the first dihedral angle. The M06-2X/6-311++G(d,p) calculation level was used. More details relative to these scans and other calculations run with other methodologies can be found in the Appendix for Chapter 4, more precisely in Figure A4.2.

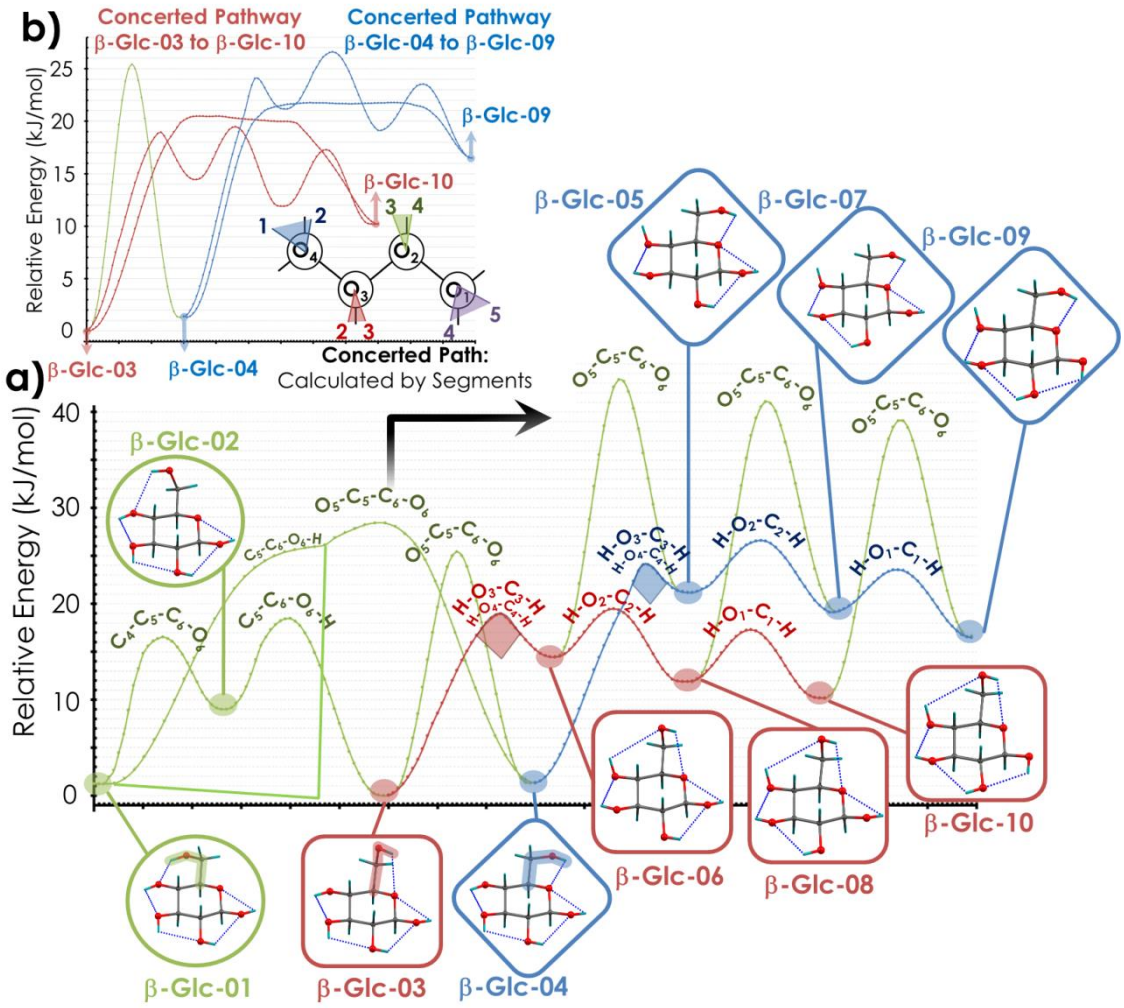


Figure 4.2.1: SPES diagram for β-D-glucopyranose molecule. a) Isomerization pathways connecting the closest conformations; b) concerted interconversion pathways (red and blue barriers with the plateau) compared to those in a). It can be noted the concerted movement of all the hydroxyl groups to isomerize from β-Glc-03&04 to β-Glc-10&09. Calculation level: M06-2X/6-311++G(d,p). The complete data are reported in Table A4.1.

Based on these preliminary results, a new method for the analysis of the PES is proposed. Once the most relevant minima are identified, it is possible replace the PES with a *simplified* PES (SPES), which is an estimation of the minimum energy isomerization paths between those minima. The SPES for the present case of β-glucose may be found in Figure 4.2.1. This diagram presents three lower energy minima and several other intermediate structures, at least 10 kJ/mol higher in energy. Moreover, two types of barriers can be identified: ≲10 kJ/mol and ≳20 kJ/mol. The three main conformers: β-Glc-01, 03 and 04 are connected by ≳20 kJ/mol barriers, while the rest of the structures are connected by ≲10 kJ/mol barriers. Thus, the relative energies, the isomerization pathways and the values of

the barriers predict that only three structures will survive the cooling process of the supersonic expansion. According to our calculations, the collision of an Ar atom (our carrier gas) at the normal speed of the expansion (Match 2, Figure A4.4-c)), provides enough energy to the molecule to overcome barriers of ≲10 kJ/mol, but not barriers ≳ 20 kJ/mol, and therefore such collisions would be able to move all the population into the three most stable conformers. The data here displayed are in agreement with previously published results: the three structures predicted here match perfectly well those detected by microwave spectroscopy[11] and also correspond to the equivalent structures of phenyl-β-D-glucopyranose.[12] In this work, these three structures were re-investigated by laser spectroscopy, extending the IR spectra down to the CH region and confirming the already published results (Figure A4.1).

The interactions between the hydroxyl groups of this molecule further complicates its conformational behavior, making also possible a concerted isomerization mechanism, jumping from β-Glc-03&04 to β-Glc-10&09 (see Figure 4.2.1). Such mechanism results in the observed plateau in the interconversion curve (Figure 4.2.1-b)). This is produced when the first DF, or scanned coordinate, arrives to the maximum of the curve and the geometry changes while the energy remains constant, because the scanned coordinate was replaced by a new one along the barrier (the coordinates shown in Figure 4.2.1-b)): the selected dihedral angle 1-2 (blue trace) is replaced by the 2-3 (red trace), and later by the 3-4 (green trace) and finally by the 4-5 dihedral angle (purple trace), resulting in several isoenergetic intermediate geometries. The energy difference observed between red and blue isomerization pathways is smaller when the process is concerted (plateau of the blue and red traces in Figure 4.2.1-b)) compared to the sequential mechanism (peaks of the blue and red traces in Figure 4.2.1-b)).

Taking into account all the data obtained in this section, it is possible to summarize the conformational behavior of β-D-glucopyranose as follows: from one side, the hydroxymethyl substituents can adopt three possible conformations: GA β-Glc-01; GG β-Glc-03 and AG β-Glc-04; on the other side, the hydroxyl groups bound to the ring operate in general as a single entity, interconverting and influencing each other through a concerted isomerization pathway, despite that some stable intermediate structures exist, such as β-Glc-05, 06, 07 and 08.

4.3. β-D-glucopyranose Dimerization Process in the Cold-Isolated (CI) Phase

Several factors were taken into account during the study of the dimerization process. On the one hand, the influence of the anomeric effect in the aggregation: the small changes in the hydrogen bond network introduced by the anomeric carbon may be amplified by the aggregation process. The difference in stability between anomers is small but usually α anomers are slightly favored. As the molecule can convert from one anomer to the other going through its linear form, such difference in stability results in an accumulation of the population in the α conformation. The conversion may be prevented introducing a substituent in that carbon, such as in methyl-/phenyl-α/β-D-glucopyranose. On the other hand, the inclusion of a substituent in the anomeric carbon may also modulate the intermolecular interactions and thus the three systems: D-glucopyranose, methyl-α/β-glucopyranose and phenyl-α/β-glucopyranose were considered, together with their mixed dimers, to study the influence in the molecular behavior of the symmetry breaking produced by the mixture of different anomers. However, as no commercial source of phenyl-α-D-glucopyranose was found, the study of α-α dimers was limited to a computational experiment.

Finally, the study of the phenyl-β-D-glucopyranose · D-glucopyranose dimer was hampered by the existence of a mixture of β/α-D-glucopyranose anomers in gas phase, resulting in a large number of possible structures to assign to the experimental spectrum. Nevertheless, this also gave us the chance to study the effect produced by the competition between both anomers during the complexation process.

Firstly, the experimental results and their interpretation on the light of quantum chemistry calculations will be presented, to continue with the results from pure computational experiments in an attempt to generalize the observed trends.

The REMPI spectra are reported in Figure 4.3.1, from which the determination of the excited state and the selection of the appropriate ionization wavelength for the IDIRS experiment was done. Although the structural information contained in the REMPI spectrum requires of a complicate subtraction procedure and its accuracy is difficult to evaluate, some related data can help to understand better the conclusions extracted from the IR spectra. The reference spectrum of phenyl-β-D-glucopyranose is reported in Figure 4.3.1: the molecule was found adopting three different conformations that interact to a greater or lesser degree with the

chromophore, producing a shift in the 0_0^0 transition. By exploiting this shift as reference for the 0_0^0 transition of all the others complexes, it is possible to obtain information on the role of the chromophore as interaction site. The phenyl-β-D-glucopyranose · D-glucopyranose dimer, trace (b) in the figure, presents several well-resolved band built on top of a broad absorption. The strongest transition lies at 36900 cm^{-1}. If we take this band as the 0_0^0 transition, it would mean that interaction between the two molecules takes place away from the aromatic ring, as a minimum shift in the origin band from the monomer to the dimer was observed.

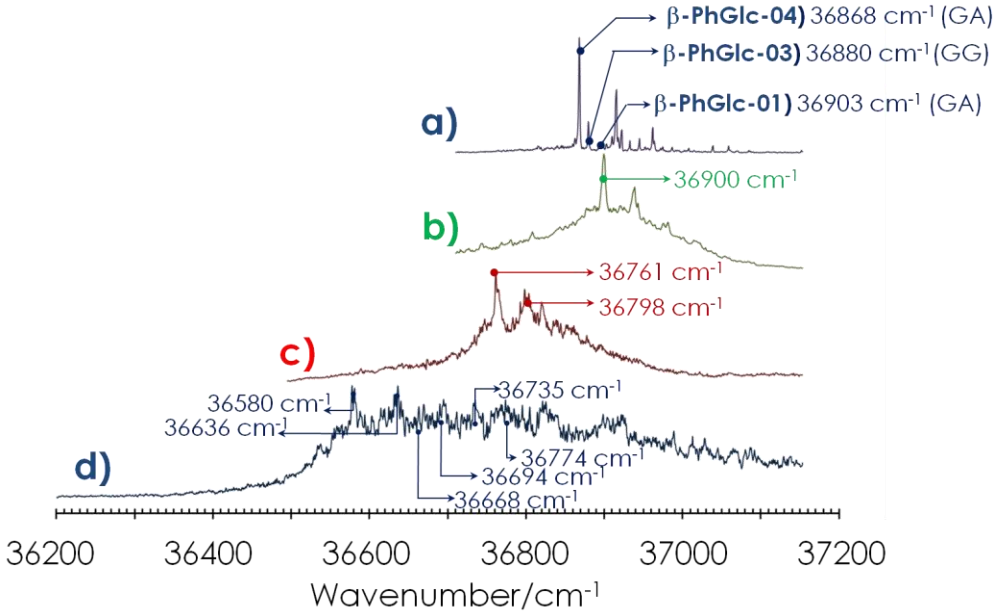


Figure 4.3.1: REMPI spectra obtained for: a) β-D-glucopyranose b) phenyl-β-D-glucopyranose · D-glucopyranose c) phenyl-β-D-glucopyranose · methyl-β-D-glucopyranose d) phenyl-β-D-glucopyranose · methyl-α-D-glucopyranose.

Regarding the phenyl-β-D-glucopyranose · methyl-β-D-glucopyranose, trace (c), the spectrum shows a red shift and a significant broadening, probably coming from a direct interaction with the chromophore. In fact, such interactions usually result in a strong coupling between the new lower energy vibrational modes produced during the dimerization couple with the electronic transition, producing a broadening of the peaks in the spectrum.[13] This phenomenon is mainly observed in complexes where the interaction with the chromophore is strong.

The last spectrum, trace (d), corresponds to the phenyl-β-D-glucopyranose · methyl-α-D-glucopyranose dimer. The shape is totally different from that of the monomer: it is a broad, unresolved trace, shifted ~ 300 cm^{-1} from that of the bare

molecule. These differences can be attributed to the presence of more than one conformer, and at least one of which presents a strong interaction with the chromophore.

With the preliminary information extracted from the REMPI spectrum, one can proceed with the recording of the IR spectra. Important structural information was obtained by comparing the IR spectra with the theoretical simulations for each calculated conformer. In order to get the best assignment, the information extracted from the REMPI experiments and from the IR data was evaluated in the light of computations on the relative Gibbs free energy. For clarity, the IR assignment of the complexes (see Figure 4.3.2) is individually treated.

The IDIRS trace obtained for the β-β dimer phenyl-β-D-glucopyranose···methyl-β-D-glucopyranose may be found in Figure 4.3.2-a), together with the proposed assignment. Additional information may be found in Figure A4.3 and Table A4.2, including IDIRS spectra recorded probing two different wavenumbers (panel a), a comparison between the experimental spectrum and the simulated ones for the five most stable calculated conformers (panel b), the relative Gibbs free energy diagram (panel c), and ball&stick representation of the five most stable conformers (panel d).

The comparison between experimental and calculated traces in Figure 4.3.2-a) shows an excellent agreement. The simulated spectrum of the global minimum faithfully reproduces the experimental spectrum, while the computed Gibbs free energy reinforces such assignment, as there is an energy gap of about 10 kJ/mol with respect to the other possible conformations. The isomerization paths from each structure falling onto the assigned conformer point to shallow barriers for such processes. Indeed, the energy barriers related to these isomerization pathways show that while the three detected conformers of β-PhGlc are isolated from each other by ≳ 25 kJ/mol barriers (Figure A4.4-a)), the aggregation process significantly reduces such barriers to ≲10 kJ/mol) in the direction of formation of the global minimum, while remain ≳ 20 kJ/mol in the opposite direction. When such values are compared with the kinetic energies for helium and argon carrier gases at the speed of our supersonic expansion (Figure A4.4-c)), it is clear that the energy of the collisions with the buffer gas allows the β-PhGlc core to relax into de global minimum when it is part of a dimer, but not when it is isolated, explaining the existence of three conformers of the monomer, but a single conformer of the dimer.

Furthermore, the analysis of the influence of the temperature in the height of the barriers for isomerization of the dimer (Figure A4.4-d)), shows that at the temperature of the beam, the molecules have enough energy available to relax into the dimer's global minimum. All these calculated data further support the assignment in favor of the 01 conformation of the phenyl-β-D-glucopyranose· methyl-β-D-glucopyranose dimer (β-PhGlc-2 · β-MeGlc-2, see section 4.3 in the Appendix for Chapter 4).

The situation is somehow different in the mixed α-/β–phenyl-β-D-glucopyranose· methyl-α-D-glucopyranose dimer. The analysis of the relative stability of the computed structures (see Table A4.3), predicts very small energy differences between the global minimum and the second most stable conformer. In good agreement, the experimental spectrum in Figure 4.3.2-b), cannot be only reproduced by the simulated spectrum of the global minimum and a contribution of at least another conformer is required. Surprisingly, the spectrum that fits better is not that of the second most stable species from an enthalpic point of view but that of a conformer that lies relatively high in energy. However, when ΔG is taken into account instead, such conformer becomes the second most stable structure in the interval of temperatures of the beam (100-200 K, Figure A4.5). As an additional test, the IDIRS spectra probing different wavenumbers of the REMPI spectrum were recorded, and an evident change in appearance was detected, depending on the wavenumber of the probe laser. This can be taken as a further evidence of the existence of at least two isomers.

The structures of the main conformers of this β-α dimer are shown in Figure A4.5-d). It is worthy to note, that conformer 01 presents a direct interaction with the chromophore, contrary to isomer 08, and therefore, not only a shift in the 0_0^0 transition from the bare molecule to the dimer is expected but also between the two isomers.

The investigation of the anomers' competition in forming aggregates can be evaluated by studying the complex phenyl-β-D-glucopyranose⋯D-glucopyranose (β/β-α). As explained above, the latter is present in both anomeric conformations as it can isomerize through its linear form. Nevertheless, the population of the α anomer is expected to be higher due to its relatively higher stability.[14] The results presented here for this system are complemented by the data In Tables A4.4 and A4.5 and in Figures A4.6 and A4.7.

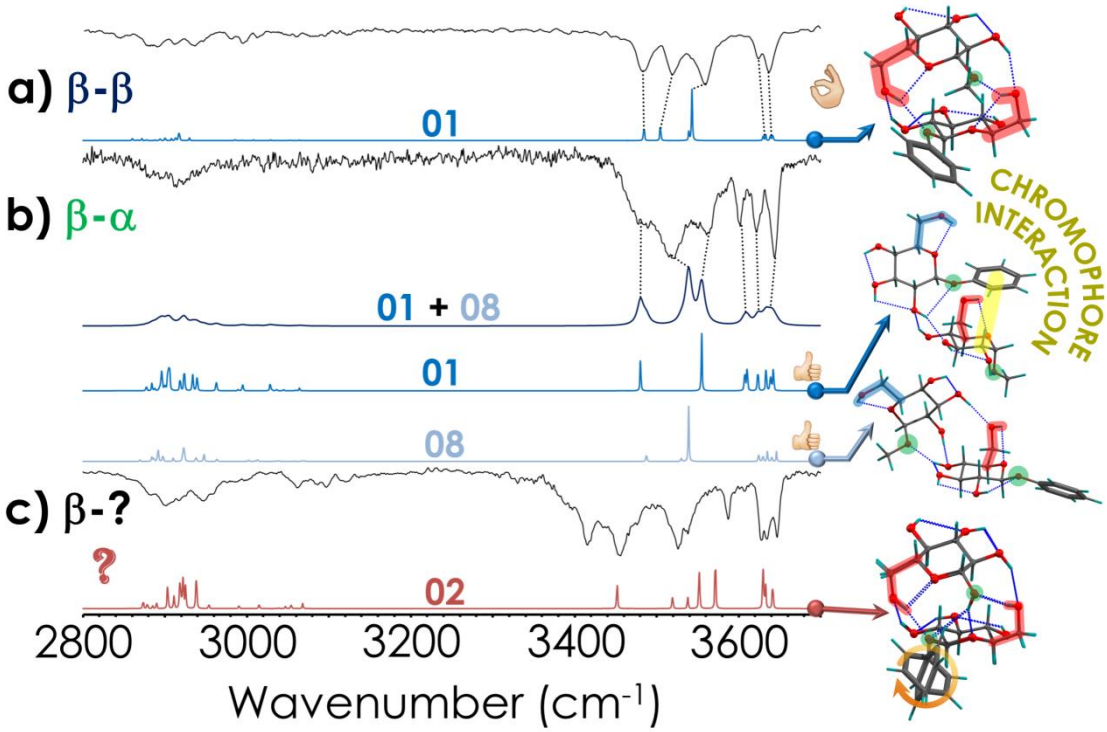


Figure 4.3.2: IDIRS spectra for: a) phenyl-β-D-glucopyranose · methyl-β-D-glucopyranose; b) phenyl-β-D-glucopyranose · methyl-α-D-glucopyranose; c) phenyl-β-D-glucopyranose · D-glucopyranose. The experimental spectra are reported in black, while the vibrational simulations at the M06-2x/6-311++G(d,p) level of calculation are displayed as colored traces.

The IR Hole-Burning experiment points to the presence of a single conformer, or if more than one isomer was formed, all the species present similar spectra. Other ionization wavelengths were also tested, but the spectrum presented the same features. Still, the experimental spectrum may be due to different species with overlapping REMPI spectra and therefore the IR obtained is invariant of the wavenumber probed by the UV laser. Furthermore, the assignment is complicated by the fact that the α-anomer is the most stable one, so it is the most abundant species in the expansion, but the β-anomer produces more stable complexes, and the anomers distribution in the solid powder of D-glucopyranose is not precisely known. Nevertheless, taking into account the Hole-Burning results, if all the IR peaks come from the same conformer, the structure here observed could be β-PhGlc · β-Glc 02, see Figure A4.6, which is equivalent to the one detected for the β-β dimer discussed previously. In fact, it is the only conformation that is able to produce five H-bonds, reproducing the five peaks observed in the experimental spectrum, corresponding to the bound hydroxyl groups.

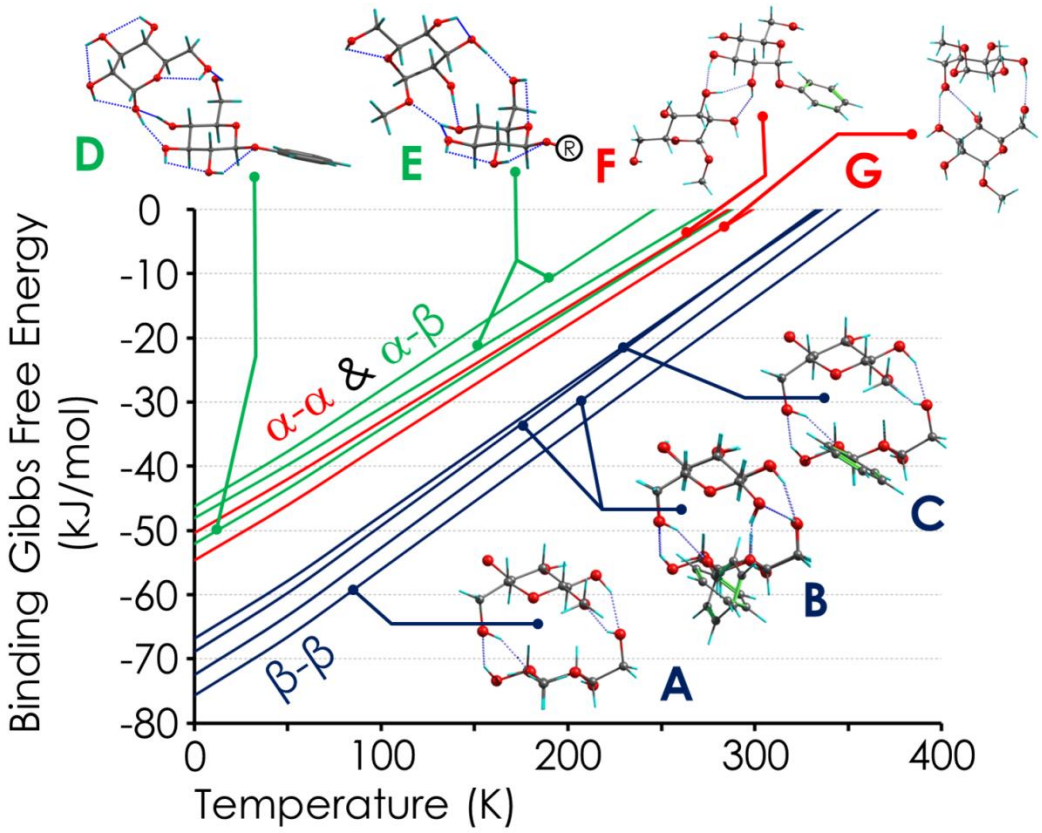


Figure 4.3.3: Most Similar (MS) binding energies for: **A)** β-MeGlc · β-MeGlc (theoretical methylated version of the structure 01 of β-PhGlc · β-MeGlc) **B)** β-PhGlc · β-Glc (experimental candidates: 01 and/or 02) **C)** β-PhGlc · β-MeGlc (experimentally determined: 01) **D)** β-PhGlc · α-Glc (conformer 04 without the chromophore interaction) **E)** β-Ph/MeGlc · α-MeGlc (08 assigned geometry and the equivalent methylated version) **F)** α-PhGlc · α-MeGlc (theoretical structure 48 without the chromophore interaction) **G)** α-MeGlc · α-MeGlc (01 theoretical structure).

The theoretical study for the α-α dimer phenyl-α-D-glucopyranose · methyl-α-D-glucopyranose was carried out employing the same level of calculation that was able to reproduced the experimental results of the previous systems (Tables A4.6 and A4.7). In most of the conformers of this dimer there is a direct interaction with the chromophore, highlighting the stabilization produced by the phenolic ring but, on the other hand, the proper interactions between sugars result masked by this collateral effect (Figure A4.8-a)). For this reason, the Gibbs free energy diagrams of the complexes without the interaction with the chromophore are reported apart, together with the corresponding structures (see Figure A4.8-b) and -c)). In order to get a more accurate view of the α-α interaction, the (methyl-α-D-glucopyranose)$_2$ homodimer was also calculated. Table A4.8 summarizes the energy results obtained. The Gibbs free energy diagrams and the structures may be found in Figures A4.9-a) and –b), while in Figure A4.9-c), the conformers of the two types of α-α dimer (phenyl-α-D-glucopyranose·methyl-α-D-glucopyranose) and (methyl-α-

D-glucopyranose)$_2$ are grouped in families, depending on the type of interaction displayed. In this latter section, the α-α interaction on those structures in which no interaction with the chromophore exists is analyzed, and the good correlation between the predicted structures for both systems is emphasized.

In addition to all these results, the structures of the assigned complexes in their corresponding methylated form (e.g. phenyl-β-D-glucopyranose·methyl-β/α-D-glucopyranose → methyl-β-D-glucopyranose· methyl-β/α-D-glucopyranose) were also calculated. The data are compared in Figure 4.3.3, where the structures and energies are reported and represented as Gibbs free energy of binding. The diagram reveals that, in general, all the observed dimers are stabilized by the formation of up to three H-bonds. The exception is the most stable conformer of the β-β family, which presents a network of four H-bonds (or five, when the dimer is formed by a β-D-glucopyranose molecule and the anomeric oxygen is free to interact).

The ability of the β-β complexes to generate a wide network of coordinated H-bonds is related to the position of the hydroxyl groups in the equatorial plane. Actually, this arrangement allows each hydroxymethyl group to form two H-bonds. This gives rise to a four H-bond compact network that presents a high stability, although there are other two arrangements that exhibit a similar symmetric interaction (with four H-Bonds), and other two more with a non-symmetric interaction but highly coordinated. It is interesting to note that only the first type of structures was observed (see Figure 4.3.4-a)). In order to evaluate why the other structures are not as stable as the one detected, a further M06-2X/6-311++G(d,p) geometry optimization was carried out on the β-β complexes. Figure 4.3.4 shows the results of the relative stabilities of these structures: the central-blue section contains the assigned conformation (a) and the other two symmetric conformers (c) and (e), which are more than 20 kJ/mol above the global minimum. In fact, these two structures are only held together by two stable H-bonds, while the rest of the intermolecular interactions are weaker, as their long distances indicate. Regarding the remaining structures (b) and (d), their asymmetry allows them to reach closer interaction distances, but without achieving the formation of four H-bonds, as in the global minimum.

Three main conclusions may be extracted from this information, regarding the aggregation and complexation processes: (*i*) the orientation of the hydroxyl groups (equatorial or axial) has a strong impact in the hydrogen bond network; (*ii*) collateral attractive weak H-bonds between the CH and the lone pairs of the

oxygen atoms are formed, and (*iii*) the sterically repulsive CH-HC and lone pair-lone pair interactions determine the final shape of the structure. In fact, the lone pairs and the CHs affect the aggregation process as much as the position of the hydroxyl groups, by blocking the possibility to reach the four stabilizing H-bonds network.

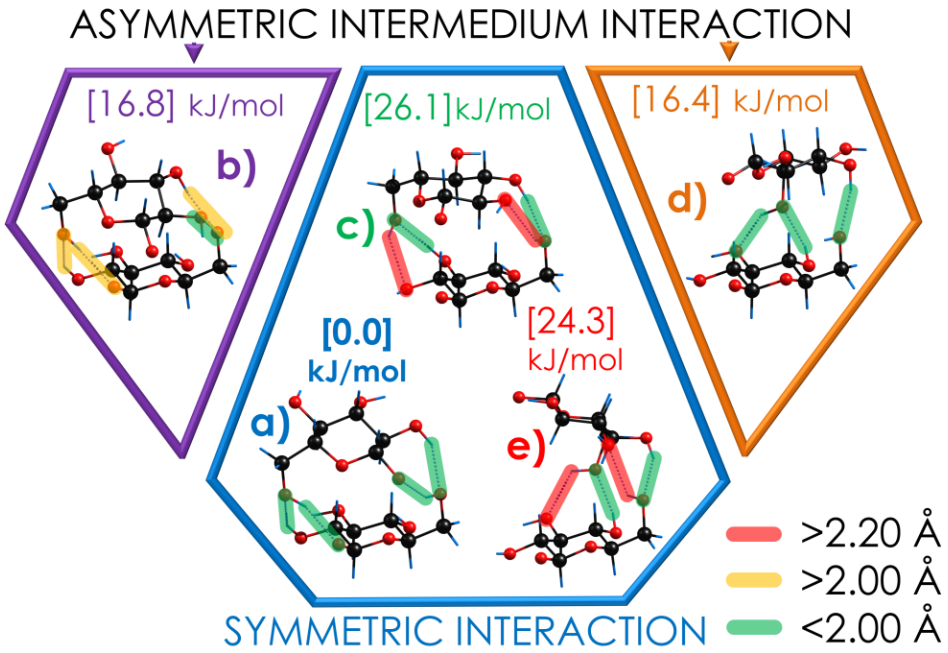


Figure 4.3.4: The assigned minima and the remaining equivalent hydroxymethyl bond types. Structures optimized at the M06-2X/6-311++G(d,p) calculation level: a) The bond type of the assigned minimum: P2M6P1-M2P6M1 b) Bond type: P2M6P1-M3P6M2 c) Bond type: P3M6P2-M3P6M2 d) Bond type: P4M6P3-M3P6M2 e) Bond type: P4M6P3-M4P6M3. The nomenclature tutorial is in Scheme A4.1.

4.4. Formation of Poly-1,4-β-D-glucopyranose (Cellulose I) Crystal

Analysis of many crystallographic structures[15–26] (see Figure A4.10), uncovered a particular property of the cellulose I crystal.[27–29] In the crystal, the equatorial hydroxyl groups interact with those from neighboring molecules, producing a layer in which the interactions extend along the equatorial plane of the whole crystal (see Figure A4.11). Thus, in addition to result in very strong intermolecular interactions in the dimers, the equatorial position of all the hydroxyl groups, produces linear chains organized in stacked layers in the cellulose crystal. The ability to form stable linear chains is one of the key factors for the generation of layers, and also maximizes the intermolecular interactions. In order to better

understand what happens in the crystal, it is necessary to divide it in sections, and analyze these sections separately (see Figure 4.4.1).

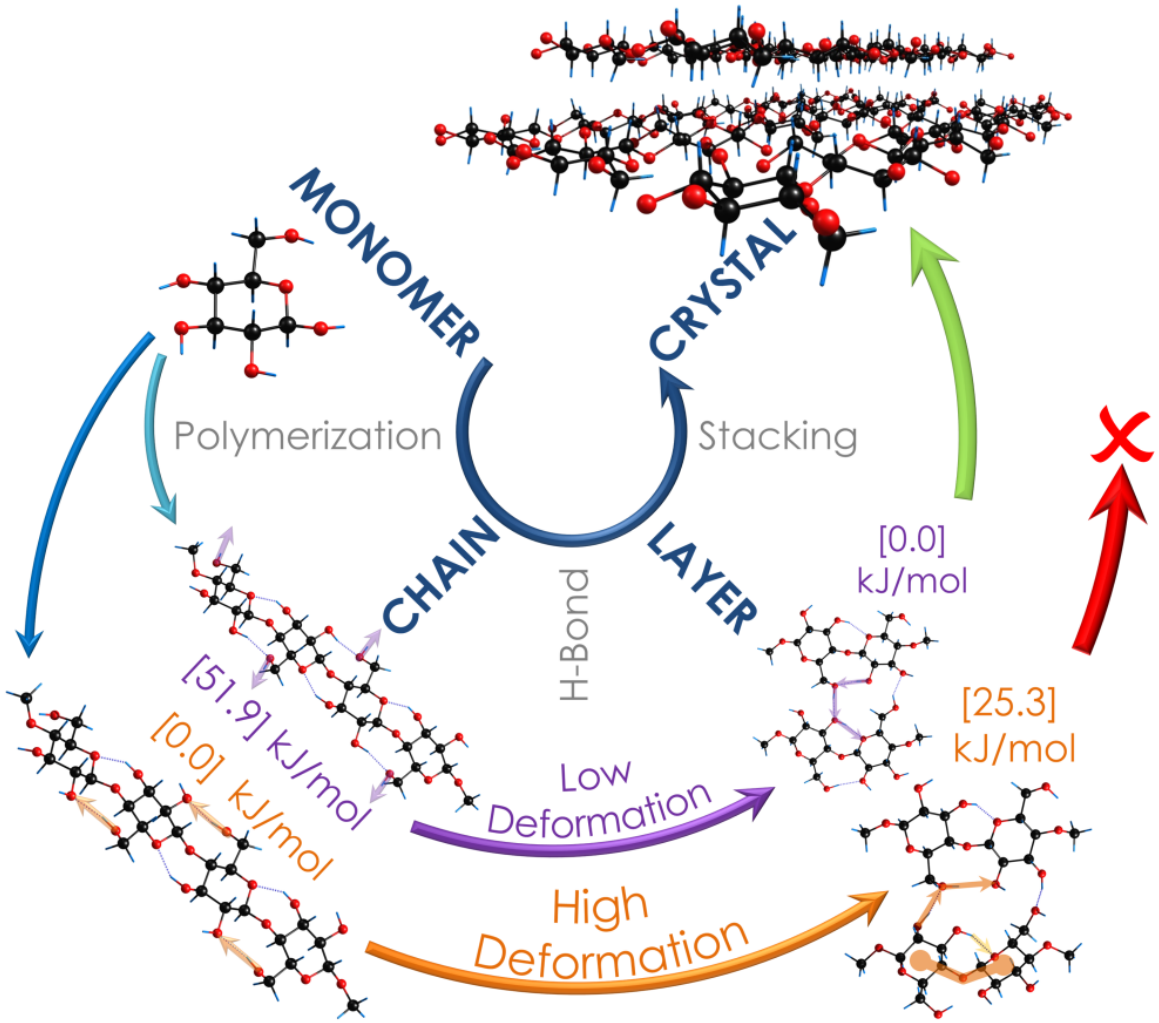


Figure 4.4.1: Unraveling the structure of the Cellulose I^{27} crystal. Level of calculation: M06-2X/6-31+G(d).

The smallest unit is the monomer, but its properties were already analyzed in section 4.2. Then, to subtract the effect produced by the chains, some calculations were done. In Figure A4.12, the optimized geometries of the possible conformations of the glucoses chains are reported. It can be noted that there are two types of chains: one with all the glucose molecules on the same side (see Figure A4.12-2)), that can be folded in a full negative bending conformation (Figure A4.12-2d)), a full positive bending conformation (Figure A4.12-2f)), or a mixture of both positive and negative bending conformations, producing and undulated chain (Figure A4.12-2e)). Although it could be possible to crystalize cellulose in any of these conformations, these geometries do not fulfill the property

highlighted before, that the equatorial plane propagates the stabilizing intermolecular interactions along the entire crystal. Moreover, they do not correlate with the crystallographic geometry in Figure A4.11, and therefore these types of structures were rejected. The other three linear geometries (see Figure A4.12-1)) present the monomeric units alternated, like in the crystal. A close view at these structures shows that the one in c) should be excluded because of three reasons: (*i*) its relative stability is too low; (*ii*) its linearity is not high enough; and the most important one, (*iii*) its steric hindrance. In fact, if a hypothetic crystallographic structure is built using this geometry, the hydrogen atoms bound to O6 and to O3 would occupy the same space, which is not possible. For all these reasons this structure has to be rejected, and at most it may act as an intermediate in hydrated cellulose, where interactions with other chains do not occur. Thus only two candidates remain: the first one with all the OH forming an intramolecular hydrogen bond network and producing the most stable arrangement (Figure A4.12-a), and the orange chain depicted at the bottom of Figure 4.4.1), and an additional structure that is less stable but presents some free OH and therefore, available to stabilize the structure through intermolecular interactions (Figure A4.12-b), and the purple chain in Figure 4.4.1).

Next, the determinant parameters for the generation of layers must be analyzed. For this calculation, the smallest portion of the crystal was taken into account in order to fully reproduce the interaction, and for the same geometry three different arrangements of hydrogen bond networks were tested. In a preliminary step, the proposed hydrogen bond network was optimized at the M06-2X/6-31+G(d) level by fixing the carbon and oxygen atoms in the geometry of the crystallographic structure (see Figure A4.13-1)). The color code used is preserved in all the figures of this section (e.g. the purple chain correlates with the purple layers in Figures A4.12, A4.13 and 4.4.1). It is worth noting how the stability is inverted moving from the chains to the layers, where the purple option becomes now the most stable arrangement. Next, the structures were re-calculated but imposing no constrains in the atoms during optimization. This second simulation is useful to understand the relative stability of the crystallographic structures and the proposed arrangements, checking also how far these structures are from the equivalent isolated minimum. Actually, this can be evaluated by looking at the structural deformation: if it is large, this means that it is far from representing the real minimum. Instead, if it is small, this reveals that the proposed geometry resembles that in the crystallographic matrix (see Figure A4.13-2)). The crystallographic option 2 (the purple structure) represents the most probable configuration in order to reproduce the crystal structure. In particular, the layers can stack each one on top of the

other in two different orientations: either avoiding the repulsive interactions produced between the lone pairs of the oxygen atoms (O···O) and the hydrogen atoms of the CH groups (CH···HC), or maximizing the attractive interactions produced by the weak H-bonds (CH···lone pair).

In conclusion, it is worth noting that the existence of different types of crystals (cellulose I α or β) with a wide variety of stacking interactions holding together the layers,[27] makes their complete description a huge and complicate task, way beyond the scope of this work in which we exclusively attempt to give a simple and general overview of this topic.

4.5. Conclusions

The synergy between the CI-phase experiments, the crystallographic data and the quantum mechanical simulations led us to reveal the behaviors of β-D-glucopyranose during formation of aggregates and crystals. Three important properties have been highlighted: (*i*) the most stable conformations of the hydroxymethyl substituents were newly confirmed: GG, GA 1 and GA 2, and they were located with accuracy in the corresponding PES; (*ii*) the ability of the hydroxyl groups of the pyranosyl ring to move either independently one from the others (producing intermedia geometries β-Glc-05,06,07 and 08), or, the most favored option, in a synchronous way to produce global conformational changes through concerted pathways (geometric variations produced from 03 or 04 to 09 or10); (*iii*) the attractive/repulsive CH and lone pairs secondary but fundamental interactions. Moreover, the main property of this sugar appears to be the equatorial arrangement of all the hydroxyl groups that arises an extra stabilization.[30] Regarding the crystallographic structure, the observed linear chains are sustained by the equatorial arrangement of the hydroxyl groups; this linearity maximizes the number of interactions between chains and, also thanks to the better puckering, allows them to form higher density polymers.[31,32] This property is also observed in the dimer: the equatorial hydroxyl groups permits the simultaneous interaction of both hydroxymethyl groups maximizing the number of strong H-bonds. In summary, through this study the extraordinary stabilizing ability of β-D-glucopyranose epimer has been pointed out, making it suitable for the large amount of structural functions that is able to accomplish in nature.

4.6. References

(1) Carpita, N. C.; Gibeaut, D. M. Structural Models of Primary Cell Walls in Flowering Plants: Consistency of Molecular Structure with the Physical Properties of the Walls during Growth. *Plant J.* **1993**, *3* (1), 1–30.

(2) Wang, a H.; Nathans, J.; van der Marel, G.; van Boom, J. H.; Rich, a. Molecular Structure of a Double Helical DNA Fragment Intercalator Complex between Deoxy CpG and a Terpyridine Platinum Compound. *Nature* **1978**, *276*, 471–474.

(3) Wang, A. H.; Fujii, S.; van Boom, J. H.; van der Marel, G. A.; van Boeckel, S. A.; Rich, A. Molecular Structure of r(GCG)d(TATACGC): A DNA--RNA Hybrid Helix Joined to Double Helical DNA. *Nature* **1982**, *299* (5884), 601–4.

(4) Varki, A. Biological Roles of Oligosaccharides: All of the Theories Are Correct. *Glycobiology* **1993**, *3* (2), 97–130.

(5) Dwek, R. A. Glycobiology: Towards Understanding the Function of Sugars. *Chem. Rev.* **1996**, 96, 683–720.

(6) *CARBOHYDRATE RECOGNITION Biological Problems, Methods, and Applications*; Boons, B. W. and G.-J., Ed.; WILEY, 2001.

(7) Marquardt, R.; Quack, M. Handbook of High-Resolution Spectroscopy; 2011; Vol. 1, pp 539–549.

(8) León, I.; Millán, J.; Cocinero, E. J.; Lesarri, A.; Castaño, F.; Fernández, J. A. Mimicking Anaesthetic-Receptor Interaction: A Combined Spectroscopic and Computational Study of Propofol···phenol. *Phys. Chem. Chem. Phys.* **2012**, *14* (25), 8956–8963.

(9) Leon, I.; Cocinero, E. J.; Millán, J.; Jaeqx, S.; Rijs, A. M.; Lesarri, A.; Castaño, F.; Fernández, J. A. Exploring Microsolvation of the Anesthetic Propofol. *Phys. Chem. Chem. Phys.* **2012**, *14* (13), 4398–4409.

(10) Calabrese C., Maris A., Marcelino N., Vigorito A., M. S. High Resolution Free Jet Millimeter Wave Absorption Spectroscopy: A Bridge to Astrochemistry. In *The 24th International Conference on High Resolution Molecular Spectroscopy PRAHA*; 2016.

(11) Alonso, E. R.; Pen, I.; Cabezas, C.; Alonso, J. L. Structural Expression of Exo-Anomeric Effect. *J. Phys. Chem. Lett.* **2016**, *7*, 845–850.

(12) Talbot, F. O.; Simons, J. P. Sugars in the Gas Phase: The Spectroscopy and Structure of Jet-Cooled Phenyl β-D-Glucopyranoside. *Phys. Chem. Chem. Phys.* **2002**, *4* (15), 3562–3565.

(13) Nibbering, E.; Dreyer, J.; Kühn, O. Chapter 7. Vibrational Dynamics of Hydrogen Bonds. *Anal. Control Ultrafast Photoinduced React. Chem. Phys.* **2007**, 619–687.

(14) Alonso, J. L.; Lozoya, M. a.; Peña, I.; López, J. C.; Cabezas, C.; Mata, S.; Blanco, S. The Conformational Behaviour of Free D-Glucose—at Last. *Chem. Sci.* **2014**, *5* (2), 515.

(15) Brown, G. M.; Levy, H. A. A-D-Glucose : Precise Determination of Crystal and Molecular Structure by Neutron-Diffraction Analysis. *Science (80-.).* **1995**, *147* (9), 1038–1039.

(16) Jefrey, S. S. C. C. G. A. The Refinament of the Crystal Structures of B-D-Glucose and Cellobiose. *Acta Crystallogr. Sect. B* **1967**, *B24*, 830–838.

(17) Ferrier, W. G. The Crystal and Molecular Structure of B-D-Glucose. *Acta Crystallogr.* **1963**, *16*, 1023.

(18) Kimura, F.; Oshima, W.; Matsumoto, H.; Uekusa, H.; Aburaya, K.; Maeyama, M.; Kimura, T. Single Crystal Structure Analysis via Magnetically Oriented Microcrystal Arrays. *CrystEngComm* **2014**, *16* (29), 6630.

(19) Prusov, A. N.; Prusova, S. M.; Zakharov, A. G. Interaction of Cellulose and Lignocellulosic Polymers with Water and Aqueous Systems *. *Russ. Chem. Bull. Int. Ed.* **2014**, *63* (9), 1926–1945.

(20) Melekh, N. V; Frolova, S. V; Aleshina, L. A. POLYMERS X Ray Analysis of Powdered Celluloses Obtained with the Use of Lewis Acids. **2014**, *56* (2), 134–141.

(21) Moncrief, J. W.; Sims, S. P. Absolute Configuration Determination Using the Anomalous Scattering of Cu-K, X-Rays by Oxygen Atoms: Cellobiose. *Chem. Commun.* **1969**, 914–915.

(22) Kalenius, E.; Kekäläinen, T.; Neitola, R.; Beyeh, K.; Rissanen, K.; Vainiotalo, P. Size- and Structure-Selective Noncovalent Recognition of Saccharides by Tetraethyl and Tetraphenyl Resorcinarenes in the Gas Phase. *Chem. - A Eur. J.* **2008**, *14* (17), 5220–5228.

(23) Popov, D.; Burghammer, M.; Bule, A.; Montesanti, N.; Putaux, J. L.; Riekel, C.; Synchrotron, E.; Facility, R.; Cedex, G. A-Amylose Single Crystals: Unit Cell Refinement from Synchrotron Radiation Microdiffraction Data. **2006**, 3704–3706.

(24) Popov, D.; Bule, A.; Burghammer, M.; Chanzy, H.; Montesanti, N.; Putaux, J. Crystal Structure of A-Amylose: A Revisit from Synchrotron Microdiffraction Analysis of Single Crystals. *Macromolecules* **2009**, *42* (4), 1167–1174.
(25) Nishiyama, Y.; Mazeau, K.; Morin, M.; Cardoso, M. B.; Chanzy, H.; Putaux, J. Molecular and Crystal Structure of 7-Fold V-Amylose Complexed with 2-Propanol. *Macromolecules* **2010**, *43* (20), 8628–8636.
(26) Sarko, A.; Biloski, A. CRYSTAL STRUCTURE OF THE KOH-AMYLOSE COMPLEX*. *Carbohydr. Res.* **1980**, *79*, 11–21.
(27) O'Sullivan, A. C. Cellulose: The Structure Slowly Unravels. *Cellulose* **1997**, *4* (3), 173–207.
(28) VanderHart, D. L.; Atalla, R. H. Native Cellulose: A Composite of Two Distinct Crystalline Forms. *Science (80-.).* **1984**, *223*, 283–284.
(29) Nishiyama, Y.; Langan, P.; Chanzy, H. Crystal Structure and Hydrogen-Bonding System in Cellulose I from Synchrotron X-Ray and Neutron Fiber Diffraction. **2002**, 9074–9082.
(30) Usabiaga, I.; Gonzalez, J.; Leon, I.; Arnaiz, P. F.; Cocinero, E. J.; Fernandez, J. A. Influence of the Anomeric Conformation in the Intermolecular Interactions of Glucose. *J. Phys. Chem. Lett.* **2017**, *8* (6), 1147–1151.
(31) Sun, C. True Density of Microcrystalline Cellulose. *J. Pharm. Sci.* **2005**, *94* (10), 2132–2134.
(32) Sarko, A.; Wu, H.-C. H. The Crystal Structures of A-, B- and C-Polymorphs of Amylose and Starch. *Starch/Stärke* **1978**, *30* (3), 73–78.

5. Analysis of Folding Structures

5.1. Introduction and Objectives

Foldamers are discrete linear molecules that fold into conformational-ordered states in solution. These molecules mimic the ability of proteins, nucleic acids and polysaccharides to fold into well-defined conformations, such as helices and sheets.[1] This chapter will deal with a peptide sequence considering it as a foldamer, since the methodology of the study would be identical.

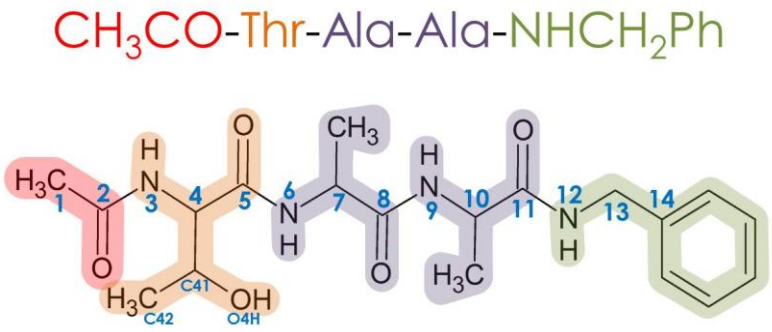
Depending of the presence and number of versatile elements, foldarmers can vary in size, and this confers to these molecules their ability to adopt many different structures by exploiting the flexibility of the monomeric units. Their structures are stabilized by non-covalent interactions (mainly H-bonds and dispersive forces) between non-adjacent monomeric units. Hence, the importance of the role played by these intramolecular interactions links the structure of the foldamers to the function that they accomplish in nature.[2]

In this chapter, the gas phase structure of different foldamer-like systems is unraveled. The conformation adopted in the gas phase may not correspond to the same structure that the polymer assumes in solution or crystal. However, the information obtained from the analysis of the gas phase data is related to the intrinsic conformation of the molecular target, in absence of any interference with the environment. These kinds of studies help to identify the predominant intramolecular interactions and other types of effects involved. The focus of the chapter will be the investigation of the folding process determining the properties of two foldamers. It can be noted that this type of study is significantly more complex than others treated previously due to the high conformational flexibility of the molecules to be studied.

The two investigated compounds belong to two important and representative molecular groups: a tripeptide, and a polymer based on anthracene units. The first one is composed of a threonine and two alanine units (Ac-Thr-Ala-Ala-NHBn), capped by acetyl and NH-benzyl groups.[3] These groups preserve the peptide bonds which appear in polypeptides or proteins. Furthermore, this peptide sequence usually appears as part of the backbone of O-linked glycoproteins with antifreeze activity.[4] The goal of this analysis is to determine the predominant conformation(s) adopted by this tripeptide in the CI-phase and evaluating the role of the intramolecular interactions which favor this structure.

The other macromolecule investigated is a polymer, which is composed by repetitive units of two and three anthracene-containing monomers. This polymer, similar to what occurs between layers of graphite,[5] has an exceptionally high conductance. In this case, the objective is to determine the arrangement and interactions of this macromolecule that give rise to this specific conductance.

5.2. Analysis of the Folding Process in a Highly Flexible Tripeptide



Scheme 5.2.1: Structure of Threonine-Alanine-Alanine tripeptide.

As already explain before, this tripeptide (TAA hereafter, and shown in Scheme 5.2.1) is formed by three monomeric units: one threonine and two alanine. The chain is capped in N3 by an acetyl (-$COCH_3$) group, and in N12 by an NH-benzyl (NHBn) group (for more details see Scheme A5.1). The Bn substituent was inserted as the chromophore required to carry out the spectroscopic study. The sample is not commercial, it was synthesized by Prof. Ben Davis' group (Oxford University, United Kingdom) and this allows us to have tailor-made samples to focus on specific aspects. Only few milligrams of TAA were synthetized, thus the efforts were focused mainly on the measurement of the IDIRS spectrum. The REMPI spectrum obtained for this molecule presents two sharp well-resolved peaks: one at 37530 cm^{-1} and the more intense at 38059 cm^{-1}. Both peaks generated the same IDIRS spectrum (see Figure 5.2.1).

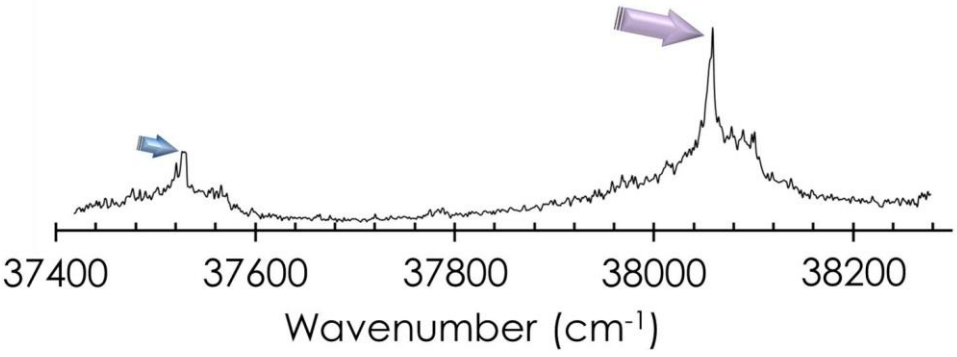


Figure 5.2.1: One-color REMPI spectrum of TAA. The bands selected to record the IDIRS spectrum are highlighted.

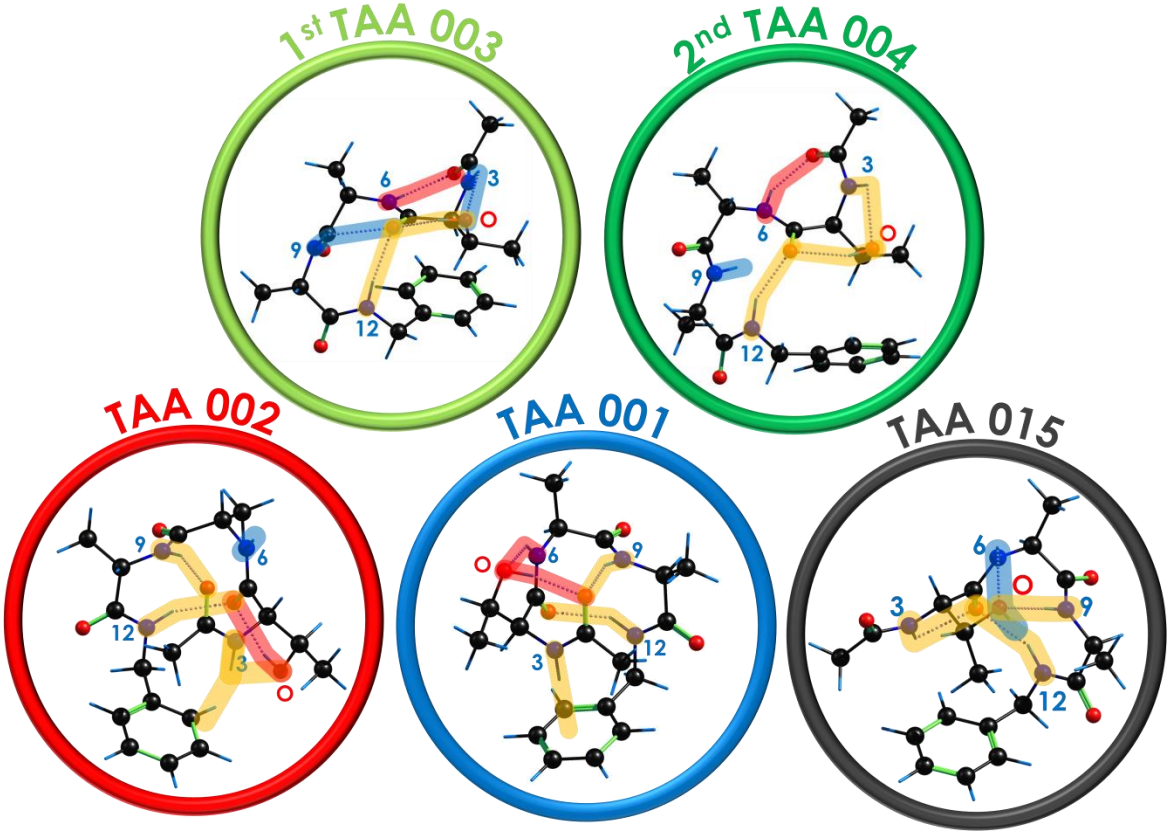


Figure 5.2.2: Five representative conformations of TAA. The blue numbers correspond to nitrogen atoms and the red "O" letter marks the OH group of threonine. TAA 003 and TAA 004 conformations exhibit very similar geometries; TAA 001 and TAA 002 are folded in opposite directions, and TAA 015 presents the same folding direction than TAA 001 but with a different network of intramolecular interactions.

A huge computational effort was performed in this study: 130 conformations were fully optimized with a DFT method sensitive to dispersive forces (M06-2X/6-31+G*). In addition, several force fields were used in the initial conformational search using molecular mechanics. The computational requirements for this work were very demanding, and it could be accomplished only thanks to the supercomputing facility at the University of the Basque Country. In Appendix of this Chapter 5, all the structural parameters and theoretical results can be found. In particular, Figures A5.1, A5.2, A5.3 highlight the main intramolecular interactions observed, and Tables A5.1.1, A5.1.2, A5.1.3, A5.1.4 report the relative energy and Gibbs free energy derived from the force field used for the conformational search, while the Tables A5.2.1, A5.2.2, A5.2.3, A5.2.4 show the dihedral angles of the tripeptide's backbone for all the structures. Finally, Tables A5.3.1, A5.3.2, A5.3.3, A5.3.4, A5.3.5 summarize the main intramolecular interactions for each conformer. These tables are very rich in information but at the same time they are not very helpful to have a general view of the system. For this reason, Figure A5.4 collects all the calculated structures, grouped into families and in Figure A5.5 ball&stick representations of

selected conformers of these families are displayed. Once the families are identified, their relative stability in a broad range of temperatures was analyzed using the relative Gibbs free energy (Figure A5.6). A detailed view of the , most stable candidates for each family can be found in Figure A5.6-d).

It is worthy to note that a thorough exploration of the conformational landscape for this system could only be achieved by employing several force fields. In fact, many conformers were detected only in a particular force field. Hence, an initial screening using a single force field would be incomplete, missing many low energy conformers, as can be seen in Table A5.1.1.

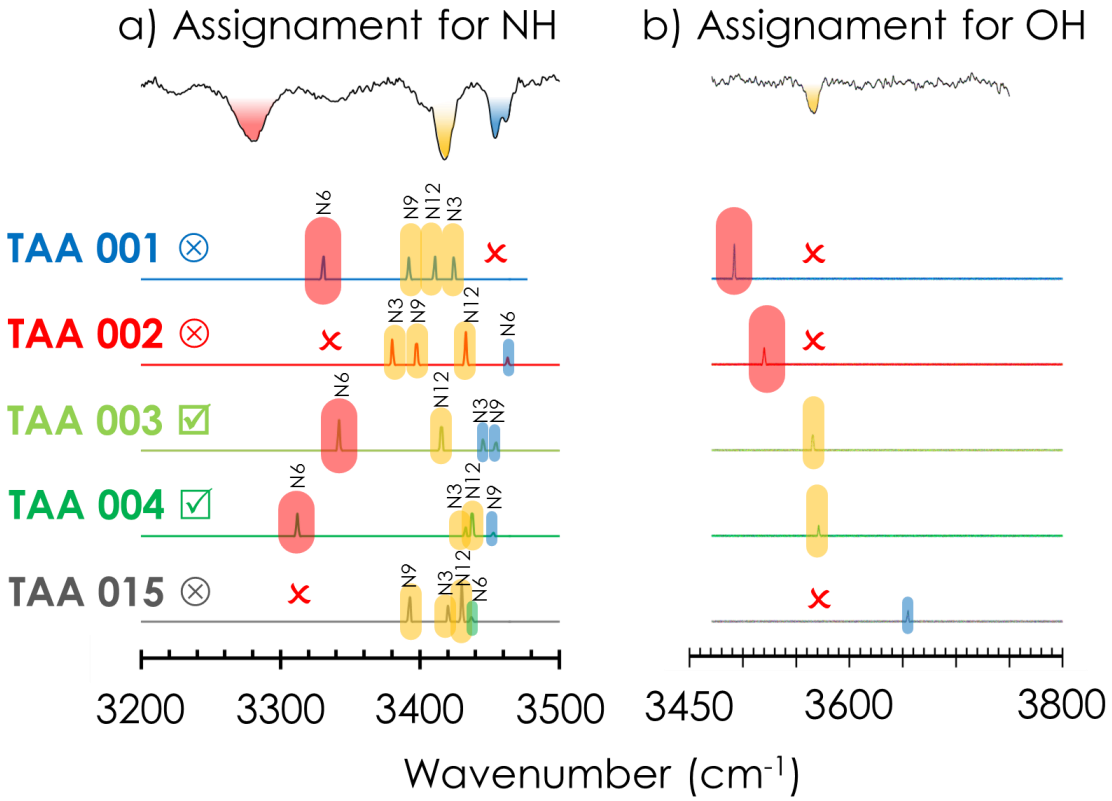


Figure 5.2.3: Comparison of the experimental IDIRS spectrum (upper trace in black) and the predicted spectra (color traces). The simulations were calculated at M06-2X/6-31+G(d) level. In order to take into account the anharmonicity, a factor of 0.953 was used, which was determined using already assigned molecules. TAA presents four NHs and one OH belonging to threonine: a) assignment for the NH groups; b) assignment for threonine´s OH.

The theoretical conformations selected (see Figure 5.2.2) were compared with the experimental IDIRS spectrum. In order to reach a reliable assignment , and since the structures of these conformers are complicate to fully understand in a 2D-view, all the geometries were decomposed in *fundamental components* in order to facilitate the procedure (see Figure A5.8). Figure 5.2.3 summarizes the assignment.

Here, the spectrum is separated into two regions: one for the stretch of the NH groups (section a) and the other for the OH band (section b). These OH and NH bands are very sensitive to the environment, in particular to the strength of the H-bonds. In fact, a strong H-bond produces a red shift in the band proportional to its strength and in some cases also a broadening of the band (exception blue shifts cases)[6], while weak interactions or non-bonded modes present very small alterations. In the spectrum reported in Figure 5.2.3, the peaks are highlighted using colors: the blue and narrow peaks represent the free NHs and OH vibrational modes, while the red and broad one represents the strong H-bond interactions, which appears to the red of the rest of transitions. Finally, the yellow peak represents a intermediate situation where a weak interaction is present. The comparison between the experimental bands and the theoretical predictions attests that the simulated spectrum of TAA 003 conformation faithfully reproduces the experimental spectrum, followed by TAA 004. The structures TAA 001, 002 and 015 are directly rejected because their peaks do not match with the experimental ones. In particular, TAA 001 does not present weakly bound OH and free NHs; TAA 002 does not have a weakly bound OH and also a strongly bound NH is missing; finally several bands of TAA 15 does not match the bands in the experimental spectrum, for example the free OH of threonine is predicted at ~3650 cm^{-1}. A complete comparison with the theoretical predictions is reported in Figure A5.7.

The isomerization energy between conformers is very high in most cases, since several H-bonds are involved in the stabilization of each structure. The interconversion pathway of TAA 003/TAA 004 pair is shown in Figure 5.2.4. Both structures TAA 003 and TAA 004 are superposed in order to better appreciate their structural differences. The profile of the interconversion pathway between these two conformers was calculated and its estimation is reported in Figure A5.9. It reveals a barrier of approximately 40 kJ/mol. This high barrier blocks any possibility of isomerization between the two conformers under the experimental conditions.

The assigned structure corresponds to the most stable conformational family described by the Gibbs free energy at relatively high temperatures (above 100 K). In particular, the best agreement is showed by TAA 003 (ΔG_{0K}~10kJ/mol, $\Delta G_{80\text{-}120K}$~0kJ/mol, $\Delta G_{120\text{-}400K}$<5kJ/mol), but an absolute rejection of TAA 004 (ΔG_{0K}~10kJ/mol, $\Delta G_{120\text{-}600K}$~0kJ/mol) cannot be done. Actually, the experimental spectrum could be the result of the contribution of both conformations TAA 003 and TAA 004. Finally, it can be noted that the folded structure adopted by this tripeptide corresponds to the semi-folded family. The experimental spectrum leaves out the other theoretically stable structures predicted as the most stable

ones at low temperatures, like TAA 002 (ΔG_{0K}~10kJ/mol, ΔG_{80K}~0kJ/mol, ΔG_{350K}>10kJ/mol) and TAA 001 ($\Delta G_{0\text{-}80K}$~0kJ/mol, $\Delta G_{120\text{-}700K}$>15kJ/mol).

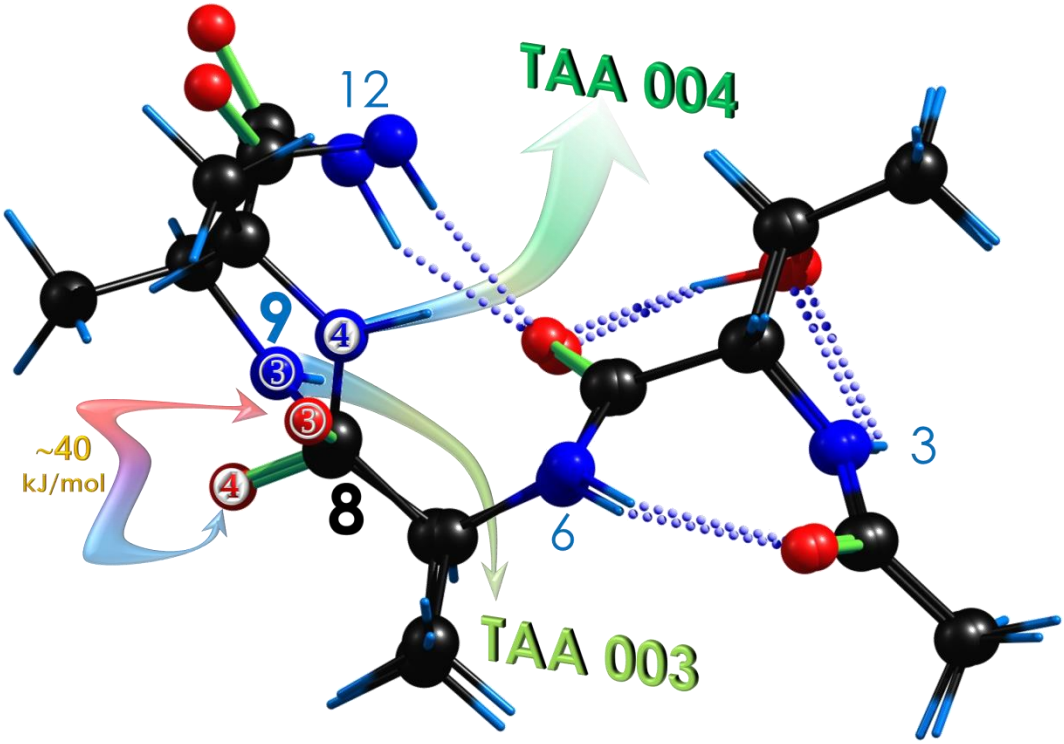


Figure 5.2.4: Comparison between two sections of the structures TAA 003 and TAA 004. The main difference consists of the value of the dihedral angle between C8-O and N9-H groups. The isomerization pathway selected and the corresponding energy estimation are reported in Figure A5.9.

Many factors affect the experimental findings and in turn the observed structure. A typical overview of the reactions involved in the process of folding was already displayed in Chapter 2, in Figure 2.3.1.1.3. The usual definition of the activation energy is the energy required to start a reaction.[7] In this complex experiment, the activation energy can be associated to the conformational distribution before the expansion.[8,9] However, before the expansion a full unfoldment of the tripeptide is allowed, thus all the conformations can be considered available. Additionally, the laser vaporization processes delivers a high amount of energy that remains available for the molecules to move freely and explore the conformational hypersurface. This is in strong contrast with the detection in our jet of a limited number of species. Somehow, the initial instants of the cooling process act also as a filter that only a limited number of species can overcome. In this case, only those structures folded in a certain, like TAA 003, survived the dynamic process of the expansion .

5.3. Analysis of the Folding Process in a Polymer Based on Anthracene Units

In this section, we tackled the spectroscopy and structure determination of a new macromolecule. The molecular target was synthesized by Prof. A. Mateo's group (POLYMAT, San Sebastian, Spain) and it consists of 9,10-dimethyleneanthracene units stitched together by 2,6-pyridinedicarboxamides. Two tailor-made samples were provided: a dimer and a trimer (see Figure 5.3.1). This polymer presents an exceptionally high conductance in molecular junctions. The objective of the study is to unravel and determine the folded structure responsible of this specific ability: are the stacked anthracene units (π···π interactions) that produce this high conductance? Or is it due to the CH···π interactions between the anthracenes?

For such large molecules (dimer m/z = 984.33 Da and trimer m/z = 1532.10 Da), it is really complicated to obtain a well-resolved REMPI spectrum. The anthracene monomers that form the oligomer present intramolecular interactions that usually couple with low-frequency modes, producing a broadening of the bands.[10] The size of this macromolecule exceeds the typical size investigated with this technique, and the setup has shown a limited cooling ability for what concern the low frequencies. Another complication regards the sample amount: only a few milligrams (~30 mg) were available for the whole experiment, transforming the study on a real challenge (around 40 hours of measuring non-stop).

With these premises, all the experimental efforts were put in obtaining good IDIRS spectra. The structural assignment comes from the comparison between the experimental IDIRS spectrum and the calculated structures. The calculations, performed by M. Melle's group (Centro ALGORITMI, Braga, Portugal), give rise to different possible structures (see the Appendix of this chapter for more details and information on the theoretical procedure). As regards the dimer, the most stable minima are represented by four different conformations, attending to the nature of the intermolecular interactions: π···π 1, π···π 2, CH···π 1 and CH···π 2, while for the trimer two stable arrangements were found: π···π 1 and CH···π 1, see Figures A5.10 and A5.11. The strategy adopted was: (*i*) determination of the mass spectrum (see Figure 5.3.1); (*ii*) anthracene 0_0^0 transitions bibliographic identification.[10]

Several wavenumbers (36363, 34343 and 34013 cm^{-1}) were probed to the blue of anthracene's 0_0^0 transitions, although a REMPI spectrum was not recorded in order to save sample. Nevertheless, no remarkable differences were observed between

the spectra recorded at the above-mentioned positions. Strong-enough signal was obtained from the dimer, while it was close to the limit of detection for the trimer. Figure 5.3.1 shows the mass spectra for both species (dimer and trimer) obtained by accumulating 500 spectra. Both spectra show complicated fragmentation patterns: this is the reason why a lot of strong peaks appear at lower masses values. Finally, the signal was divided to obtain the IDIRS spectra using active subtraction.

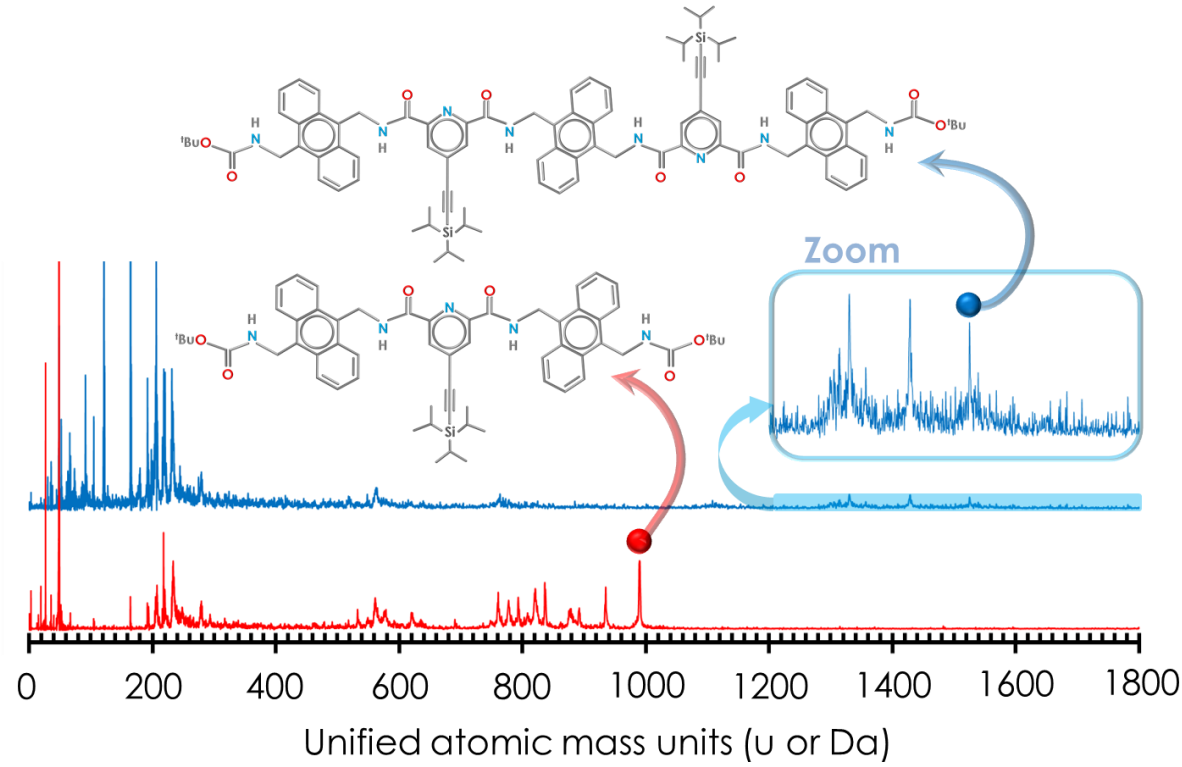


Figure 5.3.1: Mass spectra corresponding to the dimer (red) and the trimer (blue) of the anthracene polymer. The corresponding molecular schematic structures are also reported. The x-axis corresponds to m/z ratio, but it was assumed that most of the ions present a single positive charge.

The conformational-specific vibrational spectra were recorded in the C-H and N-H stretch regions, which are extremely sensitive to the local hydrogen bonded conformational environment. The IR experimental spectra of dimer and trimer are very similar (see Figure 5.3.2), with a set of bands in the 2850-3100 cm^{-1} region, assigned to the C-H stretch vibrational modes. As regards the N-H modes, the band around 3430 cm^{-1} contains the symmetric and antisymmetric coupled modes of the amide moieties. The presence of a single red-shifted band points out that the amide N-H bonds are equivalent, and that they are weakly hydrogen bonded to the nitrogen atoms of pyridine rings. A second set of N-H bands should

be observed, which would correspond to the carbamate nitrogen atoms. However, these modes are not observed in the experimental spectra. This finding is not unusual, since these atoms are not forming H-bonds and hence, these bands should be weaker. The spectra of several minima conformations of dimer and trimer were calculated and compared with the experimental spectra in order to assign the folded conformations. The agreement between experimental and calculated spectra (Figure 5.3.2) is almost perfect for the conformers stabilized by π···π 1 interactions between the anthracene moieties. The simulated spectra show that a CH···π interaction is not possible in the case of the dimer, and this univocally confirms that the observed structure is the π···π one.

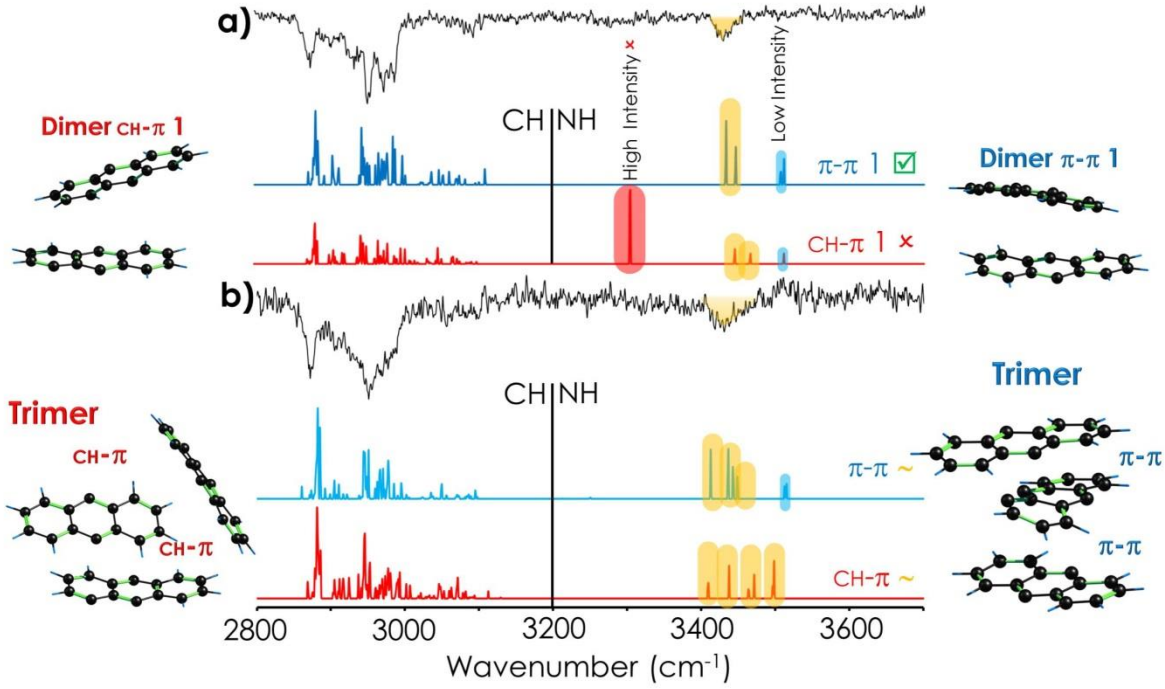


Figure 5.3.2: Comparison between the experimental IDIRS spectrum (above in black) and the theoretical predictions (below in colors). a) Dimer assignment. b) Trimer assignment. The level of calculation used was B3LYP/6-31g(d,p). Two factors were necessary to correct the anharmonicity: one for the CH region (0.980) and another one for the NH region (0.995).

On the other hand, in the case of the trimer, both simulations could correspond to the experimental spectrum. However, the difference in energy (~8.5 kJ/mol), the slightly better correlation with the experimental spectrum and the observation of the π···π interaction for the dimer, safely point out that the detected structure is the π···π one also in the trimer.

In summary, π-clouds seem to play a key role in the exceptionally high conductance of these foldamers, although the mechanism is not entirely clear yet.

The full structures of the assigned isomers are reported in Figures A5.10 and A5.11, where the main polymeric structures are shaded in grey in order to highlight the conformation adopted by the anthracene units. More details on this study can be found in the already published article.[11]

5.4. Conclusions

In conclusion, in this chapter the determination of the structure of two very large molecules were analyzed: a tripeptide with a high conformational flexibility, and an oligomer which represents the upper limit of mass detection for this kind of experiment, by means of mass resolved laser spectroscopy. In both cases, this spectroscopic technique has proved to be able to solve two scientific issues of high interest and complexity, and was able to give conclusive answers by exploiting the most of its capabilities.

In particular, the investigation of the tripeptide was supported by and compared to an exhaustive theoretical study, which led to a specific analysis of the results (conformation by conformation). Moreover, the examination of the Gibbs free energy in a broad temperature ranges enabled to put in context the experimental findings, and this was fundamental in order to determine the range where the conformational distribution fits better with the experimental observations. The good agreement between the experimental results and the predictions has revealed the properties of this tripeptide in the CI-phase and the preferred conformational arrangement.

For what concerns the anthracene polymer, the experimental technique has been pushed to its limit of mass detection measuring the dimer (984.33 Da) and the trimer (1532.10 Da), which roughly duplicate and triplicate, respectively, the typical molecular weights for this kind of measurements.

Summing-up, this section has revealed the true potential of the mass resolved laser spectroscopy technique, and highlighted new applications, that can represent only a small part of the whole spectrum of possibilities which can be explored with this technique.

5.5. References

(1) Le Bailly, B. A. F.; Clayden, J. Dynamic Foldamer Chemistry. *Chem. Commun.* **2016**, *52* (27), 4852–4863.

(2) Brioche, J.; Pike, S. J.; Tshepelevitsh, S.; Leito, I.; Morris, G. A.; Webb, S. J.; Clayden, J. Conformational Switching of a Foldamer in a Multicomponent System by pH-Filtered Selection between Competing Noncovalent Interactions. *J. Am. Chem. Soc.* **2015**, *137* (20), 6680–6691.

(3) DeVries, A.; Vandenheede, J.; Feeney, R. E. Primary Structure of Freezing Point-Depressing Glycoproteins. *J. Biol. Chem.* **1971**, *246* (2), 305–308.

(4) Jia, Z.; Davies, P. L. Antifreeze Proteins: An Unusual Receptor-Ligand Interaction. *Trends Biochem. Sci.* **2002**, *27* (2), 101–106.

(5) Sengupta, R.; Bhattacharya, M.; Bandyopadhyay, S.; Bhowmick, A. K. A Review on the Mechanical and Electrical Properties of Graphite and Modified Graphite Reinforced Polymer Composites. *Prog. Polym. Sci.* **2011**, *36* (5), 638–670.

(6) Hobza, P.; Havlas, Z. Blue-Shifting Hydrogen Bonds. *Chem. Rev.* **2000**, *100* (11), 4253–4264.

(7) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The Energy Landscape and Motions of Proteins. *Science (80-.).* **1991**, *254*, 1598.

(8) Florio, G. M.; Christie, R. A.; Jordan, K. D.; Zwier, T. S. Conformational Preferences of Jet-Cooled Melatonin: Probing Trans- and Cis-Amide Regions of the Potential Energy Surface. *J. Am. Chem. Soc.* **2002**, *124* (34), 10236–10247.

(9) Godfrey, P. D.; Brown, R. D. Proportions of Species Observed in Jet Spectroscopy-Vibrational-Energy Effects: Histamine Tautomers and Conformers. *J. Am. Chem. Soc.* **1998**, *120* (41), 10724–10732.

(10) Cockett, M.; Kimura, K. A Study of anthracene–Arn (n=0–5) in the Ground Cationic State by Laser Threshold Photoelectron Spectroscopy: Selective Ionization of Complex Isomers Formed in the Free Jet Expansion. *J. Chem. Phys.* **1994**, *100* (5), 3429.

(11) Carini, M.; Ruiz, M. P.; Usabiaga, I.; Fernández, J. A.; Cocinero, E. J.; Melle-Franco, M.; Diez-Perez, I.; Mateo-Alonso, A. High Conductance Values in π-Folded Molecular Junctions. *Nat. Commun.* **2017**, *8* (May), 15195.

6. Analysis of the Nucleation Process in Theobromine

6.1. Introduction and Objectives

Xanthines are a family of molecules whose structure contains the purine base. Actually, they are also known as purine derivatives as they are structurally related with the purine nucleobases adenine and guanine. One of the most famous xanthines is caffeine, which is one of the most popular drugs among humans. It is present in coffee, tea and many other drinks. Its structural similarity to the purine base is probably behind its ability to irreversibly block the receptor of adenosine, preventing the onset of drowsiness induced by adenosine.[1,2] Caffeine is metabolized into other xanthines, like theobromine (see Figure 6.1), and finally all these molecules are converted to uric acid or into its methylated derivative.[3]

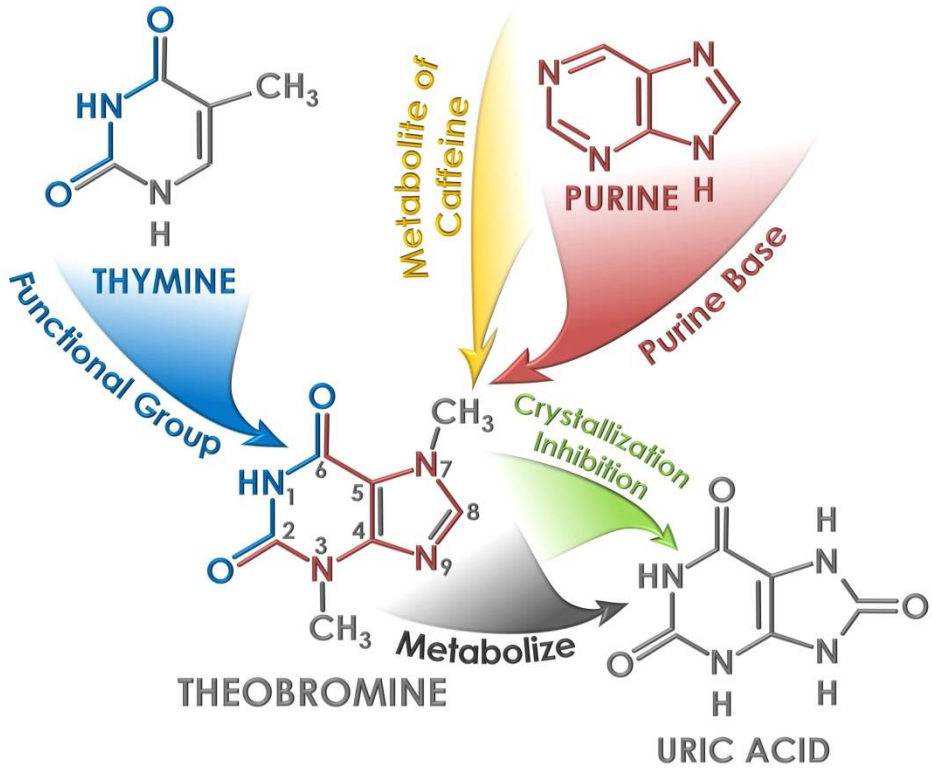


Figure 6.1: Structure and labeling of theobromine and its derivatives.

Regarding theobromine, it is present in some very common foods like chocolate and it can also take part in some metabolic processes.[4] From a structural point of view, this molecule is a purine base derivative, like adenine and guanine, which presents the same combination of -CO-NH-CO- functional groups also found in thymine, a pyrimidine nucleobase. Thus, in principle, this molecule should produce the same kind of interactions, and in particular a double H-bond interaction, that

takes place in the canonical thymine-adenine DNA base pair.[5–7] Interestingly, despite that for the homodimer of theobromine one should expect H-bonds of similar strength as for theobromine-adenine, it shows a preference for stacking interactions, as it will be shown below. Such behavior is also in contrast with that observed for guanine, cytosine, xanthine and theophylline (see Chapter 7).

Some studies suggest that theobromine inhibits the nucleation process of uric acid in the human body,[8] but specific information regarding the first stages of its nucleation are still unknown. The objective of this chapter is to unravel how the first molecules of theobromine aggregate into amorphous clusters, and the interactions which drive the process.[9] Despite that the measures are taken in supersonic expansions, the first instants of the aggregation process may be similar to the first instants of the formation of a crystal, and therefore, the conclusions extracted may be also valid for the latter case. Thus, the data obtained in CI-phase will be compared to crystallographic data and theoretical calculations, in order to build a global picture of the nucleation of theobromine.[10,11] This picture will be complemented with data on the influence that water has in the process, obtained both from experimental results and from the SPES (Simplified Potential Energy Surface) of the interaction between theobromine and water. By exploiting the results obtained, it is possible to correlate the affinity of the water for each binding site, and then, by comparison with the results from theobromine dimer, an estimation of the importance of the direct interference produced by a single water molecule in the dimerization process can be done.

Before starting with the analysis of the intermolecular interactions, a complete characterization of the theobromine molecule is required. Indeed, it presents two tautomeric structures (enol and keto forms) that can affect the full nucleation process. In order to analyze this aspect, the energy barriers for the tautomerization process were determined, also for taking into account the possibility that only one tautomer is stable at the conditions of the expansion, and therefore to explore if the whole subsequent study may be carried out considering only one tautomer.

6.2. Tautomerization of Theobromine

As regards the tautomerization process, it is strongly perturbed by the interaction with other molecules, specially by formation of H-bonds, which may facilitate the process. In order to estimate such effects, the relative tautomeric equilibrium was considered for different complexes: in particular, theobromine monomer (1THEOB), theobromine with water (THEOB+W), and theobromine dimer (2THEOB).

Examination of Figure 6.2 clearly shows that the keto form is the most stable tautomer in all cases. Such energy difference decreases with complexation, as well as the barrier for the tautomerization process. The main consequence of this change in the morphology of the tautomerization energy pathway, is the efficiency of the cooling process to trap population in the keto and enol forms during the first instants of the expansion. The barrier to move from the enol (the less stable tautomer) to the keto tautomer is ~140 kJ/mol in the monomer; a single water molecule is able to reduce it to ~40 kJ/mol, but the reduction is even more pronounced in the dimer, where its value is ~6 kJ/mol. Such small barrier is most likely not able to prevent the evolution of the enol tautomer into the keto form.

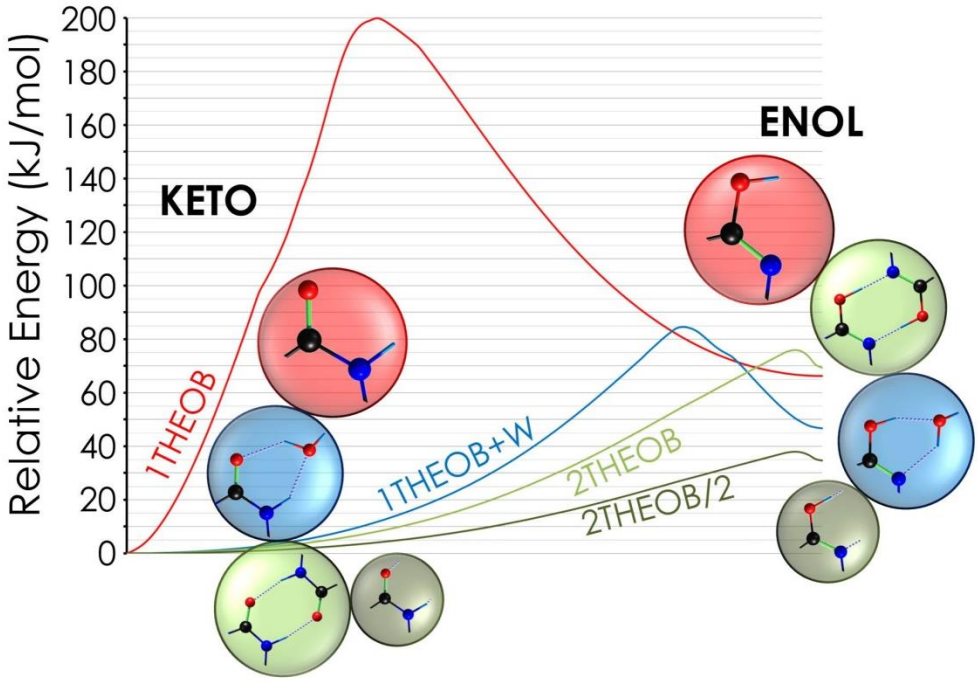


Figure 6.2: Averaged interconversion curves for the tautomerization of theobromine molecule. In order to obtain a general overview, since all the curves are very similar, they were averaged and for this purpose the x-axis was normalized maintaining the maxima in the same x-position. 1THEOB: isolated theobromine molecule; 1THEOB+W: theobromine · water complex; 2THEOB: theobromine dimer. In the latter case the result was divided by two producing the 2THEOB/2 curve, in order to obtain the energy related to the contribution of a single molecule. All the calculations were performed at the M06-2X/6-311++G(d,p) level. Full information is reported in Figures: A6.2, A6.6 and A6.9 of the Appendix.

In summary, the possibility to isolate enol molecules decreases during condensation, and therefore no molecular aggregates containing the enol form are expected to be detected (and they were not, see section 6.4). For this reason, these results allow us to simplify the analysis considering only the keto form for the nucleation process.

6.3. The Water Effect

The characterization of theobromine-water interaction is a way to map the possible intermolecular interactions that theobromine establishes. Water is an excellent proton-donor, but it is also a good proton-acceptor and can even interact with the electronic cloud of an aromatic ring.[12,13] In some way, the water acts as a probe that allows one to determine the interaction strength of the functional groups of theobromine. Furthermore, any biologically-relevant effect of theobromine will take place necessarily in solution.

Figure 6.3.1 shows the SPES for theobromine-water interaction. It is clear from the diagram that the structures bound by a H-bond are significantly more stable than the single O-H$\cdots\pi$ conformer, 1THEOB+W 07, which lies 20 kJ/mol above the global minimum and that represents the less stable conformational isomer found.

Regarding the aggregates formed by H-bonds, two types may be distinguished depending on the type of H-bond formed: N···HOH, 1THEOB+W 05 and 1THEOB+W 06; O···HOH where the main interaction is between a carbonyl oxygen of theobromine and a hydrogen atom of the water unit,1THEOB+W 04 and 1THEOB+W 03; and the most stable family that encloses the complexes where two H-bonds are present: O···HOH//NH···OH_2, 1THEOB+W 01 and 1THEOB+W 02. For these isomers, a second interaction where the nitrogen of theobromine acts as proton-donor with respect to the oxygen of the water adds an extra stabilization that makes them the most stable structures.

Regarding the isomerization between structures, the SPES in Figure 6.3.1 show direct connections between some of the isomers. At first glance, 1THEOB+W 01 and 1THEOB+W 02, the two most stable isomers, are separated by a high-enough barrier to expect both of them to be formed and detected in the expansion. On the other hand, the N···HOH complexes (1THEOB+W 05 and 1THEOB+W 06), are at least 12 kJ/mol less stable with respect to the global minimum, and therefore, even if they survive the cooling stage, a very small population is expected to get trapped in those minima, resulting in very weak signals, probably below the detection threshold. As regards 1THEOB+W 04 and 1THEOB+W 03, some considerations on the relaxation barriers should be done. In fact, as depicted in Figure 6.3.1, some energy barriers connecting the different minima can be overcome by collisions with argon during the expansion, like from 1THEOB+W 07 to 06 and then to 05, as well as from 1THEOB+W 09 to 05 or 04 and from 1THEOB+W 08 to 03. In particular, there are two barriers that lie on the limit of the energy

delivered by the collisions with argon: from 1THEOB+W 04 to 01 and from 1THEOB+W 03 to 02, therefore these two complexes may not be able to survive in the expansion.

Finally, the analysis of the stability of each isomer in Figure 6.3.1 gives also hints on the strength of the interaction of the different functional groups of theobromine with the water molecule. In particular, the binding energy shows that the less bound conformations are characterized by weak interactions involving the methyl groups of theobromine (~10-15 kJ/mol), followed by the ring interaction which presents a stabilization energy around 25 kJ/mol, to end with the most tightly bound conformations, which present H-bond interactions above 30 kJ/mol. The total dissociation energy of the most stable isomer is ~45 kJ/mol.

In conclusion, for all these reasons the structures that are expected to be found in the experimental spectrum are 1THEOB+W 01 and 1THEOB+W 02.

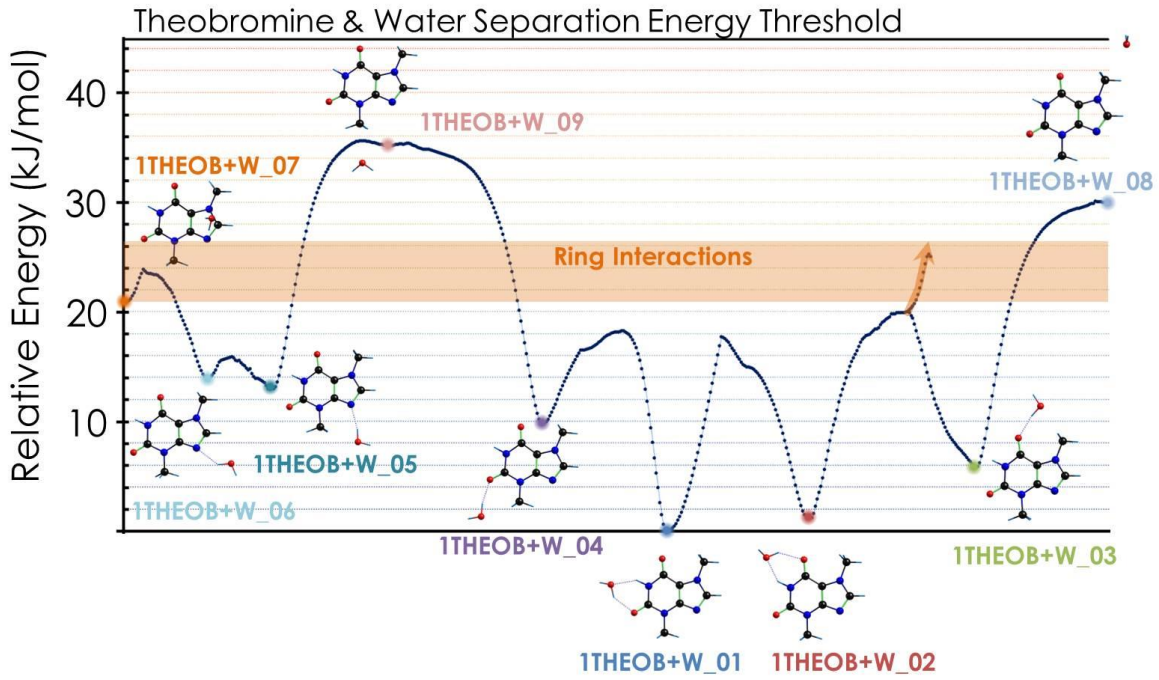


Figure 6.3.1: SPES of theobromine-water. The main stable geometries are reported and they are connected to the most similar structures, through a simulated path that shows the barriers found for each isomerization process. The orange window corresponds to the energy difference between the isomer with the HO-H···π interaction and an intermediate geometry connected with a transition state. In the Appendix (Figure A6.4), some examples of different geometries were optimized starting from a ring interaction, but except for 1THEOB+W_07, all the other trials fall in the H-bond structures (see an example of the trial displayed as an orange arrow in the diagram). All the details of the interconversion barriers and structural parameters scanned are reported in Figure A6.5.2. All the calculations were done at the M06-2X/6-311++G(d,p) level.

The REMPI spectrum of this system (Figure A6.1-b)) is not very informative and does not enable the identification and isolation of the different species that may be present in the beam. Certainly, it is a broad absorption that extends through the whole measured spectral region, from the red to the blue of the monomer's 0_0^0 transition. The spectral region where the monomer's spectrum appears could not be scanned, due to the strong signals it produces that blind the detector.

For what concerns the IDIRS spectrum, Figure 6.3.2, it was recorded ionizing at ~36095 cm^{-1}, and presents three strong bands, which are well reproduced in the calculated spectra of 1THEOB+W 01 and 02 conformers. In this experiment, these two clusters are indistinguishable and the relaxation process of 1THEOB+W 02 into 01 is not possible under our experimental conditions, therefore the spectrum is very likely the result of the contributions of both conformers. The full analysis for this assignment can be found in Figure A6.7.

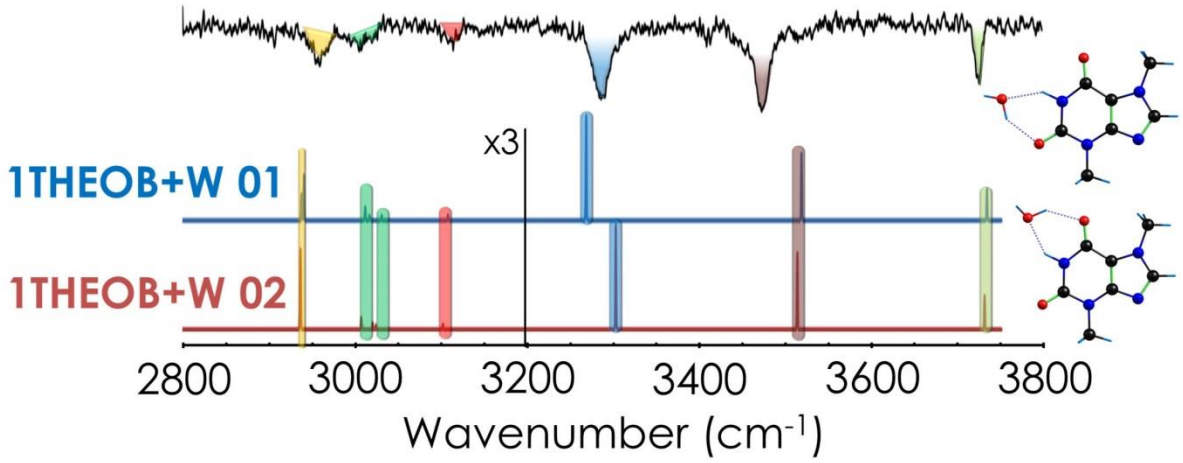


Figure 6.3.2: Experimental IDIRS spectrum (black trace) obtained for the theobromine-water complex and the corresponding theoretical predictions (blue and red traces). All the calculations performed at M06-2X/6-311++G(d,p) level, correction factor of 0.9530 for NH and CH, and of 0.9385 for OH stretching vibrations.

The comparison between the most strongly bound dimers of theobromine (which will be discussed in the section below) with the theobromine-water complexes highlights that, the competition between self- and water-aggregation may explain specific effects observed in this system, such as its solubility and the first instants of crystal formation. More precisely, a close look to the relative Gibbs free energy of binding (MS) of 1THEOB+W 01, 2THEOB 10, 1THEOB+W 07 and 2THEOB 01 (Figure 6.3.3) shows that, up to the standard temperature of 298 K, both homodimers are the most stable conformers, and that the complex with water characterized by a H-bond interaction (1THEOB+W 01) is the only one who can compete with the

formation of homodimers, taking into account its higher affinity. For this reason, this H-bond geometry can interfere during the nucleation process of theobromine (see Figure A6.8 for a complete description of the dimer geometries). This hypotheses is reinforced by the low energy difference between the H-bond complexes (1THEOB+W 01 vs 2THEOB 10), while the relative stability of the other geometries (1THEOB+W 07 and 2THEOB 01) is much higher (> 40 kJ/mol).

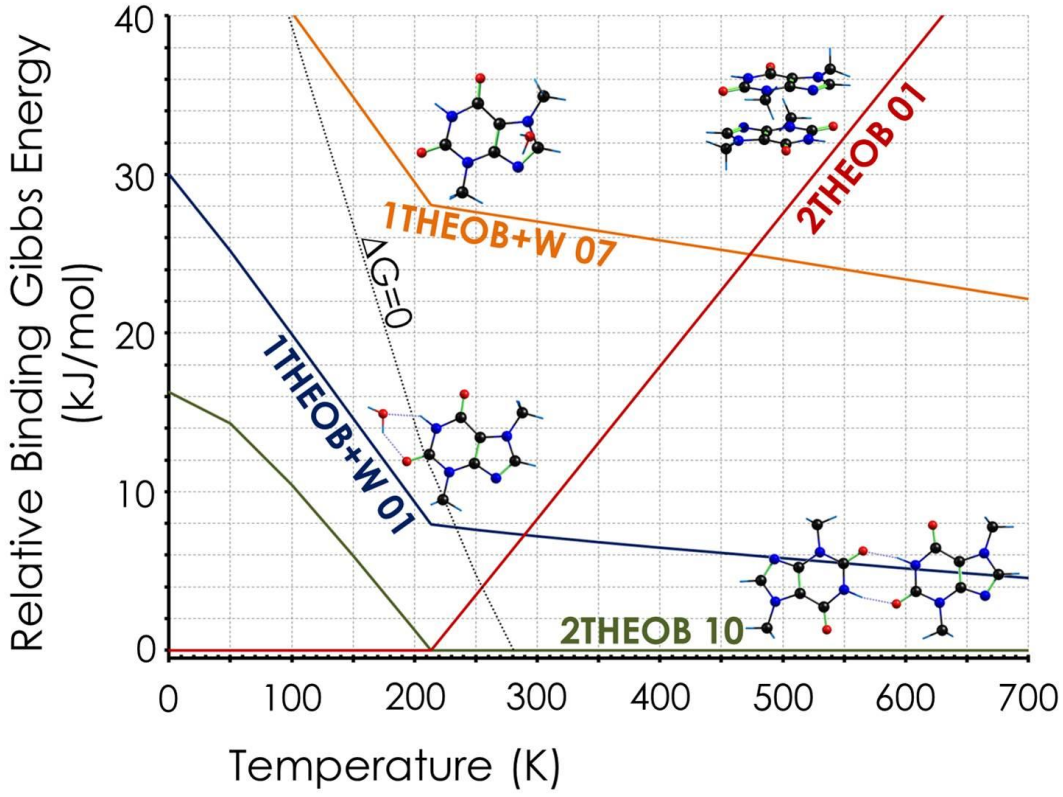


Figure 6.3.3: Binding Gibbs free energy diagram comparing two theobromine dimers (2THEOB 01 and 2THEOB 10, stacked and hydrogen bonded ones respectively) with equivalent theobromine-water complexes (1THEOB+W 01 and 1THEOB+W 07, which correspond to the conformations interacting towards the ring and the observed H-bonded one, respectively). Additional information are in Figure A6.12.

A further conclusion can be inferred looking at Figure 6.3.3 regarding the behavior of theobromine in solution, and in particular considering the weakness of the ring-water interaction. This may be one of the factors that makes this molecule relatively insoluble.[14,15] Actually, the self-aggregation of water is substantially more stabilizing than the interaction with the aromatic ring in theobromine and therefore, introducing a theobromine molecule in a water solution implies the rupture of several H-bonds that the water-theobromine interaction cannot compensate. Regarding the H-bond conformations, they can compete favorably

with water self-aggregation since their stabilization energy is significantly higher and this effect can play a role in helping the solubilization of theobromine.

6.4. Theobromine Nucleation: the First Two Steps

The experimental measurements of theobromine homoclusters were done at several aggregation levels: monomer, dimer and trimer. The REMPI spectroscopy (see Figure A6.1), produced broad absorptions, except for the monomer. The already-cited problem related to the strong signal produced by the monomer remains also in this case. Nevertheless, the spectra of these homoclusters start to the red of the monomer's 0_0^0 transition and therefore a significant part of their spectra could be recorded without damaging the detector.

The IDIRS spectra of the monomer and dimer were reported by M. P. Callahan *et al.*: in the case of monomer, they used the band at 35320 cm^{-1} in the REMPI spectrum.[16] In order to test if there are interferences due to fragmentation and to explore the spectroscopy of the monomer in a different region, the IDIRS spectrum of theobromine was recorded probing the same band used for the complexes: ~34843 cm^{-1}. However, the spectrum recorded does not exhibit fragmentation from higher mass channels and it is similar to the one reported in the literature.[16] The assignment was carried out by comparison with the theoretical prediction for the corresponding tautomer (see Figure 6.4, the full data are reported in Figure A6.3-b)), and confirms that the observed conformer corresponds to the keto form (1THEOB NH 01), although the presence of a small contribution from the most stable enol forms (1THEOB OH_C6_T and 1THEOB OH_C2_T) cannot be excluded (as depicted in Figure A6.3-b)).

The IDIRS spectrum of the dimer (see Figure 6.4), reveals that only the stacked conformer was detected. This assignment is deduced from the absence of a binding NH···O interaction that usually produces an intense and broad band at ~3200 cm^{-1}, and from the presence of the characteristic free NH stretch band, exactly at the same position where it appears in the spectrum of the monomer. For this dimer, the IDIRS spectrum was also recorded at 36420 cm^{-1}, which is to the blue of the monomer's 0_0^0 transitions in order to avoid the damage of the detector (see Figure A6.11). A broadening of the free NH stretch peak was observed, probably due to the contribution from other stacked conformers. All the possible structures and the complete dataset are reported in the Appendix, in Figures: A6.8, A6.10 and A6.11 and in Table A6.1.

The assignment of theobromine dimer shows that the most stable conformers are 2THEOB 01 and 2THEOB 02, characterized by stacked interactions (pairing or not the same side: A-A→2THEOB 01 and A-B→2THEOB 02 , see Figure A6.8). According to the calculations, they represent the lowest energy conformations at the temperature of our expansion (< 200 K). Nevertheless, other stable conformations also can contribute to the detected peaks (see Figure A6.11 for a complete description of the dimer assignment).

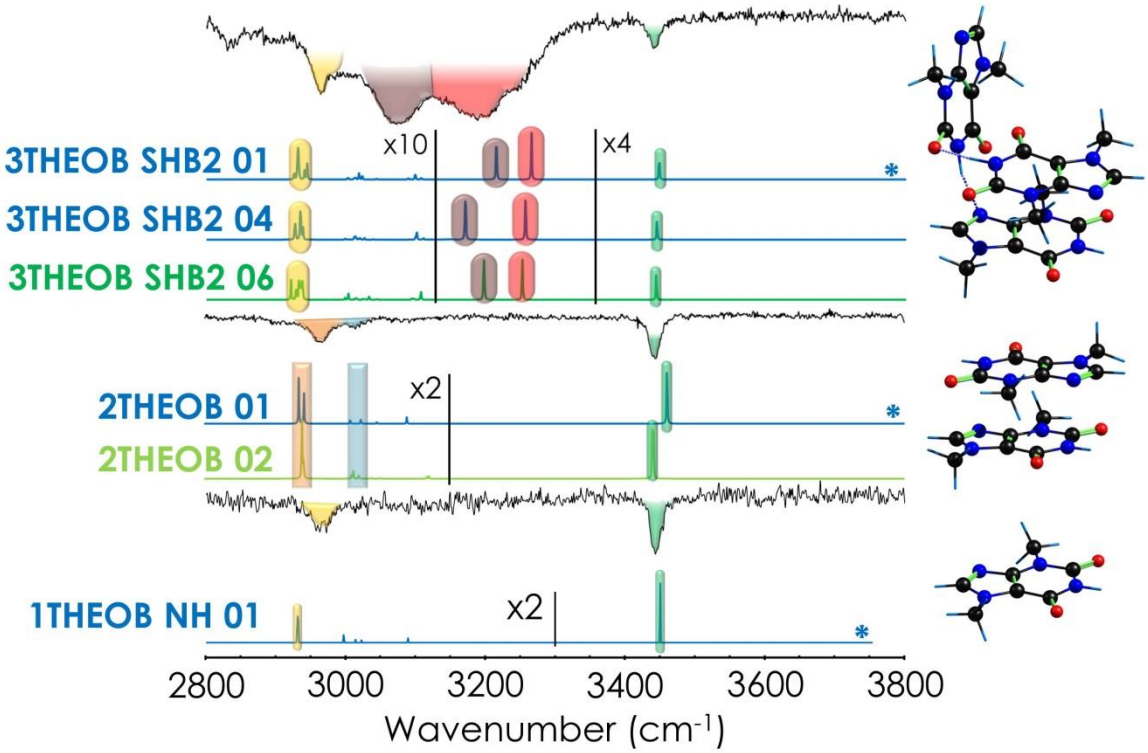
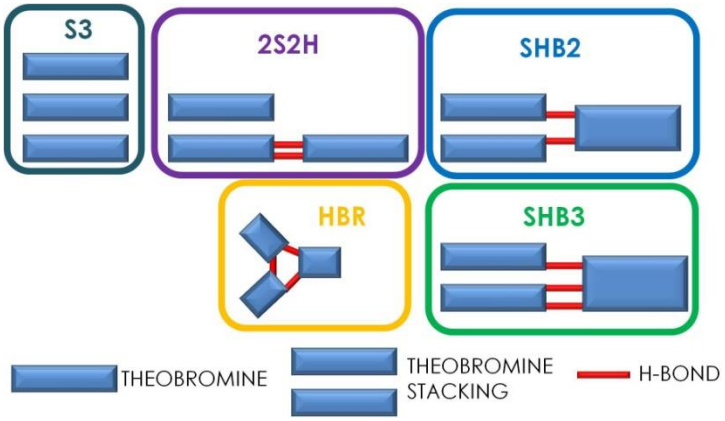


Figure 6.4: Summary of the assignment of theobromine monomer, dimer and trimer. The experimental results and theoretical predictions are compared: the experimental spectra appear as black traces, while the color traces correspond to the predictions. All the calculations were done at M06-2X/6-311++G(d,p) level, and a correction factor of 0.953 was used both for NH and CH. The depicted structures correspond to the prediction trace marked with an asterisk.

The addition of a third molecule significantly complicates the conformational landscape and therefore, the assignment. Moreover, a laborious conformational search was necessary, with the subsequently optimization and organization of the resulting structures into families. This latter step was performed manually, by looking at the main interactions characterizing each conformer (see Table A6.2 and Figure A6.13 for a general view). In particular, five families were identified: HBR; 2S2H; SHB2; SHB3; S3, in which the labeling used for their description reflects the type of intermolecular interactions that hold the molecules together (see Scheme 6.4). The HBR family (see Figure A6.14) groups those conformers based exclusively on H-

bond networks. 2S2H family (see Table A6.3 and Figures A6.15 and A6.16) is characterized by the stacking of two molecules and formation of two H-bonds with the third one. The SHB2 family (see Table A6.4 and Figures A6.17 and A6.18) is formed by structures that present one stacked interaction and an H-bond between each molecule in the stack and the third one. The SHB3 family, (see Table A6.5 and Figures A6.19 and A6.20) resembles the previous one, but in this case the H-bonds acquire more relevance, as three H-bonds are present. The last family is the S3 (see Figure A6.21), where the monomers are held together by stacking interactions alone. All the most stable conformations of each family were considered for the assignment of the experimental results.



Scheme 6.4: Sketch of the classification of theobromine trimer into families.

Based on stability considerations, the three most likely candidates for the assignment are SHB2, SHB3 and 2S2H families, as the components of S3 and HBR families are too high in energy, see Figure A6.13.

A close examination of the most stable families highlights that the 2S2H family is the combination of the most stable families of stacking and H-bonded dimers (2THEOB 10, 11 and 12). Dimer's formation seems to be governed by stacking interactions rather than by formation of H-bonds , as also observed by M. P. Callahan *et al*.[16] The two most stable members of this family are also the ones that better reproduce the experimental spectrum (Figure 6.4 and A6.22), in which the signature of the stacking interactions, the band at 3500 cm^{-1}, is also present. The existence of this family of structures in the beam would also fit very well with the commonly accepted growing scheme of the molecular aggregates, in which each aggregate is built adding a new molecule to one of the most stable structures of a pre-existing species. Only in cases where a very high energy difference between species exists and the binding energy is high enough to climb the barriers for

isomerization, different species that break the rule may be found. The other two families represent, in fact, two new ways of assembling (SHB2 and SHB3 families) characterized by different intermolecular H-bonds which play a fundamental role in their stabilization, and that were not found in the dimers. In particular, it has been estimated that the relative stability of SHB3 and 2S2H families of conformers are quite similar (from about 5 to 14 kJ/mol above SHB2 family, see Table A6.2 for the complete dataset).

In summary, the experimental spectrum (see Figure 6.4) is assigned to the most stable conformations of the SHB2 family, which are the only geometries able to explain all the peaks of the spectrum: free-NH in green, and bound-NH in red, although for the latter the prediction is not as good as for the free-NH stretches, because of the anharmonicity. However, it is not possible to discard contributions from other species belonging to the other two families, SHB3 and 2S2H (see Figure 6.4 and A6.22), because their spectra could be also hidden in the experimental trace.

6.5. The Aggregation Process of Theobromine

In order to get insight to the nucleation process of theobromine, it is interesting to compare our results with the crystal structure. Actually, the aggregates of limited size studied here may be considered as the species formed in the first stages of any nucleation process of theobromine, independently of where it takes place: in a saturated solution or in gas phase. However, the final structure of the solid resulting from the aggregation may differ significantly from those initial stages or perhaps it may inherit part of their characteristics.

The literature on the crystal structure of theobromine is very limited. To the best of our knowledge, there is a single report on crystals of anhydrous theobromine obtained using vacuum sublimation.[17] Such procedure somehow resembles the supersonic expansion's environment, since in both cases the molecules are in gas phase, although in the latter, the aggregation takes place also in gas phase and in a colder environment, while during the former, the molecules in gas phase condense into a solid (crystal). Furthermore, in the expansion all the molecules are cooled suddenly by collisions with the buffer gas, limiting the available energy for isomerization and re-arrangement processes just to the energy released by the aggregation itself. Conversely, sublimation is a slow process in which the molecules carry a significant amount of ro-vibrational energy (K. Ford *et. al.* performed the sublimation at 498 K) and condense on a cold surface, which is usually around 273

K. Therefore, they can use the excess thermal energy to overcome potential energy barriers, and to form a more ordered crystal structure.

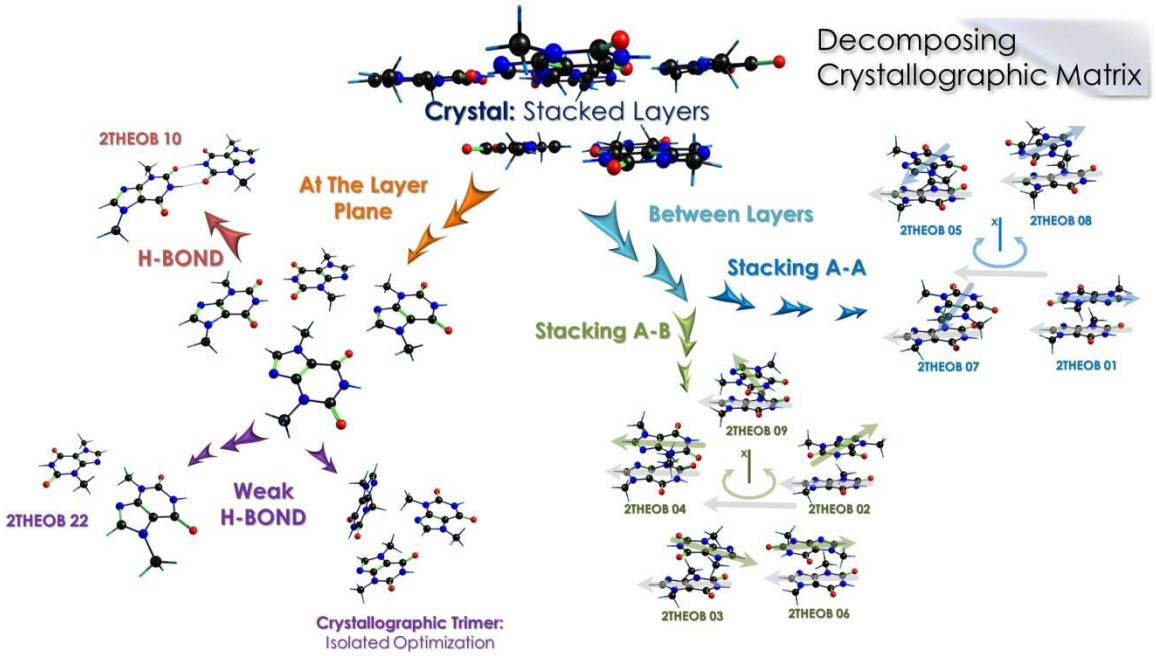


Figure 6.5.1: The crystallographic structure of theobromine is decomposed down to its possible precursor dimers.

The crystallographic structure of theobromine[17] was obtained from the Cambridge Crystallographic Data Centre (CCDC). It may be decomposed in pairs and trimers of interacting molecules, so they can be compared with the aggregates found in our expansion. As depicted in Figure 6.5.1, the crystal is organized in stacked layers, formed by H-bonded theobromine molecules. Although stacking is the main interaction between layers, some C-H···N H-bonds are also present. Interestingly, the stacking seems to be more or less insensitive to the orientation of the molecules, and both A-A and A-B orientations were found in the crystal. For that reason, none of the two families can be rejected in the forthcoming analysis. Moreover, comparison between the pairs and trimers of interacting molecules in the crystal with those found in our expansion and during the conformational search, shows that the most stable trimers formed in the beam are readily recognized also in the crystal. More precisely, they correspond to the 2S2H_A-A & A-B families. The only difference with the structures in the crystal is a slight shift of the stacked molecules and the non-planarity of the H-bond unit in the molecular aggregate formed in the expansion (see Figure A6.16). Nevertheless, the latter structural difference is simply a consequence of the absence of interactions between the trimer and a surrounding medium. Certainly, the unit that composes the crystal is the H-bonded dimer 2THEOB 10. Apart from this and the stacked

dimers, no other arrangements with a relevant stability seem to take part as precursors of the crystal structure.

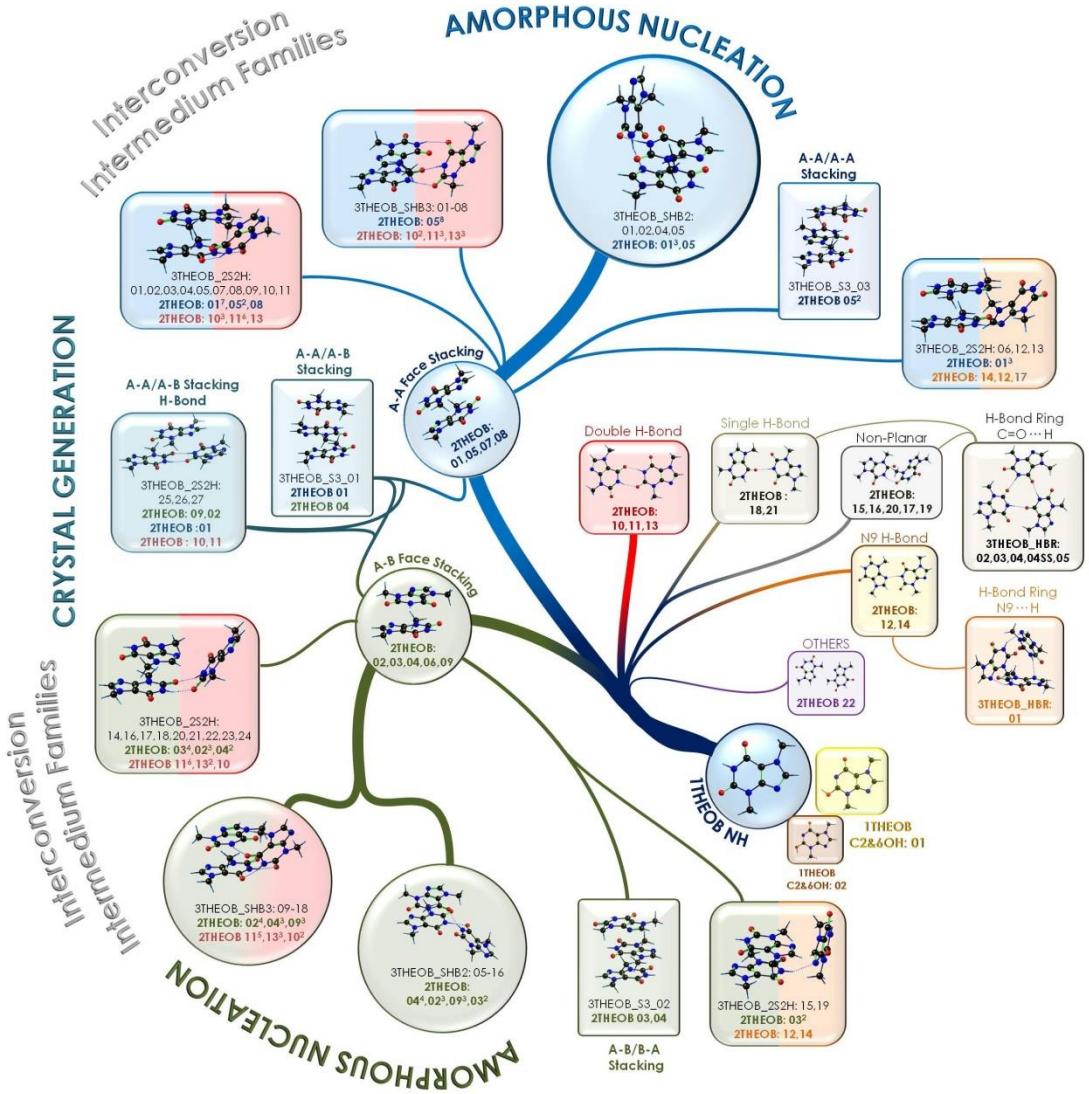


Figure 6.5.2: *Family tree* of the nucleation of theobromine: from the monomer to the trimer. The circles show the families isolated in supersonic expansion. The dimensional limit of the image makes it necessary to reduce the structures represented: only the main conformer of each family for each branch of the tree is depicted.

The final challenge is the creation of a *genealogical tree* illustrating the nucleation process of theobromine, able to unify all the information reported in this chapter.[9] Certainly, that the aggregates formed in the beam may be also found forming part of the crystal, seems to indicate that there is a connection between nucleation in both types of environments, and that the clusters formed in the

beam are also the structures formed in the first instants of the formation of the crystal. In Figure 6.5.2, a scheme to connect the two worlds is presented. The branches of the tree represent mechanisms of growth and isomerization pathways, connecting the structures of the same level, which cannot be shown for the sake of clarity. The structures reported inside circles correspond to the ones observed in the supersonic expansion, and represent the most probable arrangements for an amorphous nucleation. The main origin of the crystal is represented by the 2S2H family of trimers, which can be found in most of the branches. The tree is ordered by similarity, so closer structures are structurally related: the assigned SHB2 family can easily interconvert to the SHB3 family in a warmer environment. From SHB3, just by breaking a single H-bond, the 2S2H family is obtained. As the scheme shows, the 2THEOB 10 dimer is required in order to form the crystal.

As regards the influence of water, it mainly affects the H-bond interactions, as it attaches to the side of the molecule that renders stronger H-bonds. The preliminary analysis of the experimental data obtained in this work on monohydrated theobromine dimer, show that water can be enclosed between the two theobromine molecules. This may indicate that water strongly perturbs crystal formation: the preference of theobromine to form stacked layers would not be affected by the presence of water, but the disposition of the theobromine molecules in the layers would be strongly altered by the presence of water molecules. This can be represent one of the reasons why theobromine crystal had to be formed by sublimation.

6.6. Conclusions

The combined analysis of experimental results, theoretical calculations and crystallographic data enabled to advance in the understanding of the competing forces involved in the complicate process of nucleation of theobromine. In particular, starting from the monomeric unit, moving to its water complex, and investigating the homo-dimer and trimer, the structure of the different aggregates of this molecule were deeply analyzed. The most stable families of conformations for each kind of molecular aggregate were identified, concluding that the dimer presents a stacking structure, while the trimer is composed of the stacked dimer plus the third molecule hydrogen bonded to one of the other two units. The same spot is also chosen by water to attach to theobromine. All these structures are strongly connected with those found in the crystal and may even explain how water can cause interferences in the process of crystal formation.

Finally, it is worth noting that the study reported in this chapter represents only the beginning of the analysis of the experimental data obtained in this work. In fact, other theobromine clusters have also been experimentally analyzed, but the enormous amount of calculations required to understand the experimental results and the corresponding computational time-costs, did not allow us to present them in this context. Nevertheless, the experimental results already obtained are reported in the following chapter. The aim of the future research is to give a detailed panorama of the structures characterizing the nucleation process of theobromine, exploiting the CI-phase data in order to extrapolate to the behavior in the crystalline state, and to infer the underlying mechanisms using a synergic approach.

6.7. References

(1) Nehlig, A.; Daval, J.-L.; Debry, G. Caffeine and the Central Nervous System: Mechanisms of Action, Biochemical, Metabolic and Psychostimulant Effects. *Brain Res. Rev.* **1992**, *17* (2), 139–170.

(2) Ribeiro, J. A.; Sebastio, A. M. Caffeine and Adenosine. *J. Alzheimer's Dis.* **2010**, *20* (SUPPL.1).

(3) Tang-Liu, D. D.; Williams, R. L.; Riegelman, S. Disposition of Caffeine and Its Metabolites in Man. *J. Pharmacol. Exp. Ther.* **1983**, *224* (1), 180 LP-185.

(4) Smit, H. J. Theobromine and the Pharmacology of Cocoa. In *Methylxanthines, Handbook of Experimental Pharmacology*; Springer, 2011; Vol. 200, pp 201–234.

(5) Johnson, I. M.; Prakash, H.; Prathiba, J.; Raghunathan, R.; Malathi, R. Spectral Analysis of Naturally Occurring Methylxanthines (Theophylline, Theobromine and Caffeine) Binding with DNA. *PLoS One* **2012**, *7* (12), 1–11.

(6) Kim, S.; Lee, W.; Herschbach, D. Cluster Beam Chemistry: Hydration of Nucleic Acid Bases; Ionization Potentials of Hydrated Adenine and Thymine. *J. Phys. Chem.* **1996**, *3654* (96), 7933–7937.

(7) Sugawara, M.; Mochizuki, T.; Takekuma, Y.; Miyazaki, K. Structure-Affinity Relationship in the Interactions of Human Organic Anion Transporter 1 with Caffeine, Theophylline, Theobromine and Their Metabolites. *Biochim. Biophys. Acta - Biomembr.* **2005**, *1714* (2), 85–92.

(8) Grases, F.; Rodriguez, A.; Costa-Bauza, A. Theobromine Inhibits Uric Acid Crystallization. A Potential Application in the Treatment of Uric Acid Nephrolithiasis. *PLoS One* **2014**, *9* (10), 1–6.

(9) Byrn, S.; Pfeiffer, R.; Ganey, M.; Hoiberg, C.; Poochikian, G. Pharmaceutical Solids: A Strategic Approach to Regulatory Considerations. *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists*. 1995, pp 945–954.

(10) Chung, J.; Kim, I. W. Oriented Crystallization of Xanthine Derivatives Sublimated on Self-Assembled Monolayers. *Korean J. Chem. Eng.* **2011**, *28* (1), 232–238.

(11) Cesaro, A.; Starec, G. Thermodynamic Properties of Caffeine Crystal Forms. *J. Phys. Chem.* **1980**, *84* (11), 1345–1346.

(12) Calabrese, C.; Gou, Q.; Maris, A.; Caminati, W.; Melandri, S. Probing the Lone Pair···π-Hole Interaction in Perfluorinated Heteroaromatic Rings: The Rotational Spectrum of Pentafluoropyridine·Water. *J. Phys. Chem. Lett.* **2016**, *7* (8), 1513–1517.

(13) Pribble, R. N.; Zwier, T. S. Probing Hydrogen Bonding in Benzene-(Water) N Clusters Using Resonant Ion-Dip IR Spectroscopy. *Faraday Discuss.* **1994**, *97*, 229–241.

(14) Johannsen, M.; Brunner, G. Solubilities of the Xanthines Caffeine, Theophylline and Theobromine in Supercritical Carbon Dioxide. *Fluid Phase Equilib.* **1994**, *95*, 215–226.

(15) Li, S.; Varadarajan, G. S.; Hartland, S. Solubilities of Theobromine and Caffeine in Supercritical Carbon Dioxide: Correlation with Density-Based Models. *Fluid Phase Equilib.* **1991**, *68*, 263–280.

(16) Callahan, M. P.; Gengeliczki, Z.; Svadlenak, N.; Valdes, H.; Hobza, P.; de Vries, M. S. Non-

Standard Base Pairing and Stacked Structures in Methyl Xanthine Clusters. *Phys. Chem. Chem. Phys.* **2008**, *10* (19), 2819–2826.

(17) Ford, K. A.; Ebisuzaki, Y.; Boyle, P. D. Methylxanthines. II. Anhydrous Theobromine. *Acta Crystallogr. Sect. C Cryst. Struct. Commun.* **1998**, *54* (12), 1980–1983.

7. Additional Experimental Results and Conclusions

7.1. Introduction and Objectives

The molecular systems investigated in the previous chapters of this thesis represent examples of very different situations: starting from semi-flexible systems and their interactions (sugars), continuing with the folding process of some macromolecules (peptide and polymer), and ending with the analysis of the nucleation process in biomolecules and the connection with the formation of their crystal.

Each system presents its own challenges, which required of different theoretical approaches to better understand the experimental results and to put them in a broader context. Hence, the time devoted to computations was similar or even larger than that required to obtain the experimental results. In all cases, successive iterations were necessary in order to obtain a full description of the conformational panorama. In fact, once that the conformational search was successfully complete, all the structures optimized and the frequencies calculated, a global analysis was carried out to generate a general view that guided us in a second cycle of calculations, aiming at explaining some of the observations or to support preliminary interpretations.[1] For example, if a stable species was absent from the spectrum, the determination of isomerization barriers between that species and the detected ones was carried out. The theoretical calculations were time-consuming: the averaged CPU-time for each optimized geometry was estimated to be around three hundred hours. Certainly, a result was always obtained once that the calculation was completed. In this aspect, calculations differ from the experiment, for which an "output" is never certain. On the other hand, the answer obtained from a calculation may or may not be right and very often it is difficult to assess the computation error.

The experimental techniques used in this work, based on laser spectroscopy, entails a very complicated methodology related to lasers alignment, optimization of experimental conditions, and even to the maintenance of the specific electronic devices. Actually, the experimental setups employed are not commercial and for this reason required of continuous and precise maintenance and of double check of the whole operation steps. In particular, laser experiments are quite laborious since they depend on the laser status. Usually, the best measurements are reached after long operating sessions and in a stable environment; in particular the laser stability and therefore the quality of the signal improves at the end of the session of measures. Many factors limit the ability to get signal and/or a valid experimental measure: the type of molecular system and its size are, for example, aspects which affect the results and the time required for the

experiment. In the most demanding system investigated in this thesis, the foldamer of 1500 Da and with only 30 mg of sample available, more than 35 hours were required, and it is important to highlight that they were non-stop. The experimental part is on the one hand, the core of the investigation, but on the other hand represents the limiting part of this type of studies. In fact, the results are not given for granted, they usually take more time than calculations (except for the largest systems) and with the restriction that they require of continuous attention.

During this thesis, the main efforts were directed to the experimental part, and always aiming at taking the technique to its limits. For this reason, some calculations were also very time-consuming: because the size of the systems experimentally analyzed, and also the corresponding measurements, required of more time than normal. Thus, most of the experimental results obtained could not be completely analyzed, simply because the calculations requested may still take several years to be completed, despite that the University of the Basque Country counts with exceptional computational resources. The objective of this last chapter is to propose a preliminary interpretation of the remaining experimental results which form part of this project, making use of some already finished calculations or just by simply extrapolation of the results from other systems, when no calculations are still available.

Finally, some general conclusions will be offered about all the contents presented in this dissertation, and about the degree of accomplishment of the objectives proposed at the beginning.

7.2. Nucleation and Water Effects

The results already reported for theobromine in Chapter 6, represent a complete description of the initial stages of the nucleation process that allowed us to establish a connection between nucleation in jets and during crystal formation. Nevertheless, as mentioned before, the IDIRS experiments were performed up to the pentamer of theobromine. A close look at the spectrum in Figure 7.2.1, shows that the spectra of the tetramer and pentamer are very similar to that of the trimer. The only noticeable difference is the relative intensity of the free-NH band (3450 cm^{-1}) that decreases in the tetramer. This effect can be related to the number of possible H-bonds which are present at each stage of nucleation or, in other words, to the number of unpaired N-H moieties.

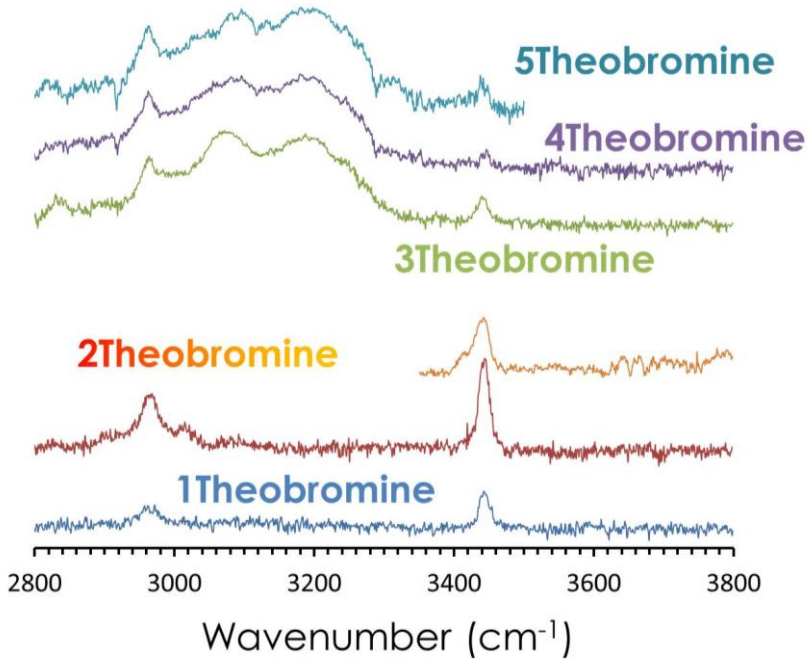


Figure 7.2.1: IDIRS spectra of theobromine homoclusters from monomer to pentamer.

Thanks to a fine tune of the experimental conditions, it was also possible to record the IR spectra of theobromine dimer and trimer with a water molecule (see Figure 7.2.2). Analysis of these systems is expected to shed some light on the influence of water in the nucleation of theobromine and more precisely, on the possible perturbations that water may introduce in crystal formation. This additional data will enable to test if water can be captured inside a theobromine complex, despite that theobromine is a very hydrophobic molecule.[2,3] Furthermore, the presence of water is helpful to confirm the assignment proposed for dimer and trimer.

The reported experimental spectra (Figure 7.2.2) for the water complexes are very complicated, in part due to the two OH stretch vibrations introduced by the water molecule, and their analysis will require of extensive calculations. Nevertheless, the number of bands due to the free-NH stretching may give hints on the structure of the aggregates: in the 1:1 cluster only one peak was present, while in the 2:1 species the spectrum shows three peaks, and the only possible explanation is the presence of at least two conformations.[4] It is also interesting to note that the free-OH band disappears in the IR spectrum of the 1:2 species, and it is a weak and slightly shifted band in the spectrum of the 1:3. This may be the signature of a water molecule trapped within the theobromine molecules, but further confirmation is required before making a final assignment.

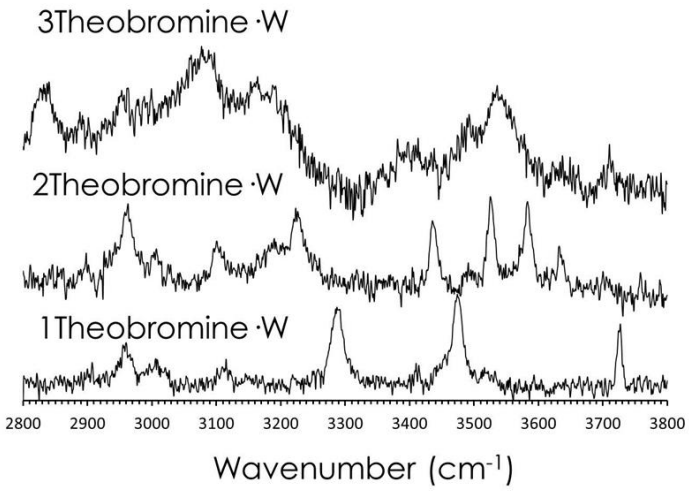


Figure 7.2.2: IDIRS spectra of complexes between a single water molecule and different homoclusters of theobromine (1 Theobromine, 2 Theobromine, 3 Theobromine).

One of the most interesting feature of theobromine is its propensity to form stacking dimers rather than H-bonded ones.[5] Such behavior is very unusual, compared with the aggregation preferences of similar molecules, such as DNA bases or other xanthines. For example, Figure 7.2.3 shows the IDIRS spectra obtained from different theophylline homoclusters.[6,7] The mass spectrum is also shown for comparison, demonstrating the existence of the so-called "magic numbers", aggregates of higher-than-expected stability. In this case, theophylline tetramer produced a significantly stronger signal than the trimer, and the peak of the hexamer is also stronger than expected, compared to the pentamer. Preliminary calculations (Figure 7.2.4), show that the predicted spectrum for the H-bonded dimer fits very well with the experimental observation. Furthermore, it is significantly more stable than the stacked structure. The formation of the trimer from the dimer is complicate and not as favored as the tetramer. This latter one results to be the stacking of two H-bonded dimers, recovering its high stability. Very probably, a similar situation takes place in the pentamer and hexamer (see Figure 7.2.4).

Finally, the interaction with water was also investigated for theophylline, by recording the spectra of the water complex for the monomer and dimer (the signals for the trimer and tetramer were insufficient in order to record a spectrum).[2] In Figure 7.2.5, the experimental spectra are reported and, as it happened with theobromine, the assignment is not straightforward. In fact, several isomers are probably present, with very different geometries. Once the calculations are finished, comparison between the results from all the systems analyzed may allow

us to understand the complex relationships of these molecules with water, during the nucleation process.

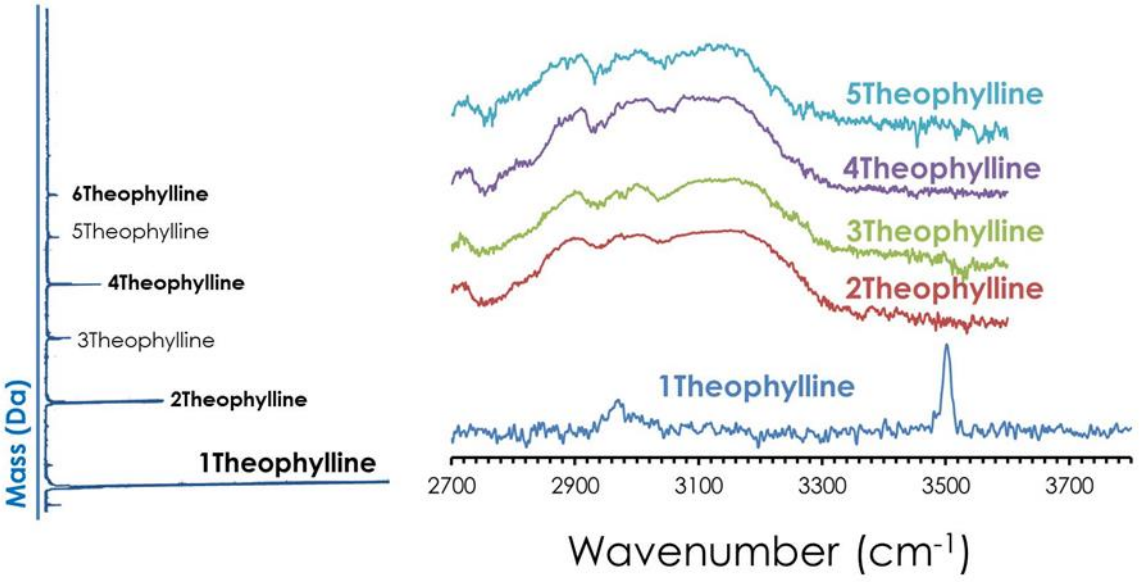


Figure 7.2.3: Theophylline mass spectrum (left panel) and IDIRS spectra (right panel), from the monomer to the pentamer homoclusters.

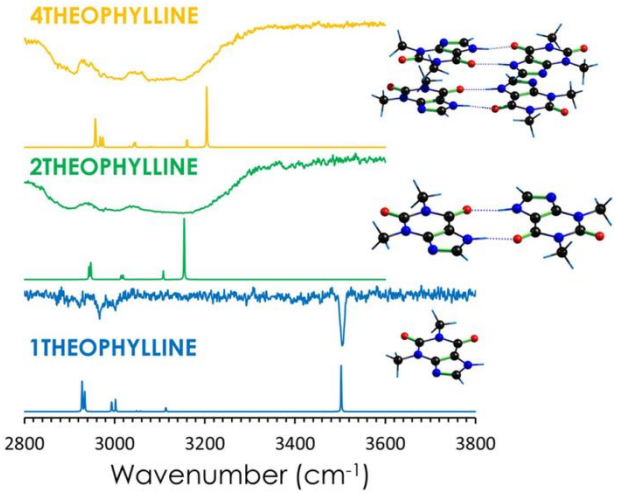


Figure 7.2.4: Preliminary assignment of theophylline. The predictions were obtained at M06-2X/6-311++G(d,p) level of calculation for the monomer and the dimer, while M06-2X/6-31+G(d) was used for the tetramer.

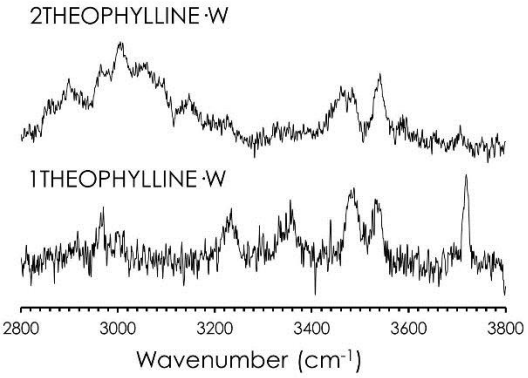


Figure 7.2.5: IDIRS spectra of the water complexes of theophylline monomer and dimer (1 Theophylline, 2 Theophylline, respectively).

7.3. Heteronucleation

The nucleation process which involves different types of molecules is very often known as heteronucleation. The study of such processes yields specially relevant information when different molecules belonging to the same family, are mixed in order to evaluate their interaction preferences. For instance, if the molecules are similar but blocked in different stable positions, due for example to the presence of methyl substituents, the study can help in the understanding of the affinity for the interaction sites and the role that they play in the aggregation process.[8]

In this regards, two systems were investigated. The first one is the complex of uric acid with theobromine. The study aims at investigating the ability of theobromine to avoid the crystallization of the uric acid reported in the literature, and that may be related to the interference of theobromine during the first instants of the nucleation process of uric acid.[9,10,11] We analyzed precisely those first stages, finding very interesting results.

The second heteronucleation system analyzed was the mixture of theobromine and theophylline with xanthine and uric acid.[12] This study aims at revealing detailed information on the binding interactions of this family of molecules and the relevant structural parameters to describe the interaction. Actually, all of them are important biochemical compounds, with pharmacologic effects, and also are structurally related to DNA and RNA bases and consequently, to many other nucleotide-derivatives present in the human body. Determination of their interaction preferences can be useful in order to understand particular biochemical mechanisms and to produce important information for drug design, by predicting the chemical behavior of a molecule or functional group with a specific surrounding.[7,10,11,13–15]

Despite that the calculations are not completed yet, preliminary assignments will be offered, by comparison with the spectra of the free monomers: theobromine, theophylline, xanthine and uric acid (see Figure 7.3.1). In fact, looking at their IDIRS spectra, the bands corresponding to the free-NH stretches correlate very well with the corresponding NH groups of the molecule. In particular, since theobromine and theophylline present only a single NH bond, their spectra present a single NH stretch. For the sake of clarity, the NH groups are labeled always in the same way in all the xanthines: theobromine presents only the free N1H; theophylline the N7H, while xanthine does not present methyl groups and therefore, three well-resolved peaks appear in its IR spectrum, corresponding to free N1H, N7H, and N3H

stretches. Regarding uric acid, it presents a C=O in the position of the xanthine CH, and an additional NH group: the N9H.[16,17] The four peaks relative to the four free NH groups can be readily identified in the spectrum of uric acid. Despite that the calculations of these systems are not completed yet, the presence or absence of the NH stretches will give useful hints on the real structures of the complexes: when an H-bonded aggregate is formed, the corresponding stretches of the interacting NH groups shift to the red, highlighting which the interaction site is. If no NH band shifts, the aggregate is formed by stacking interactions.

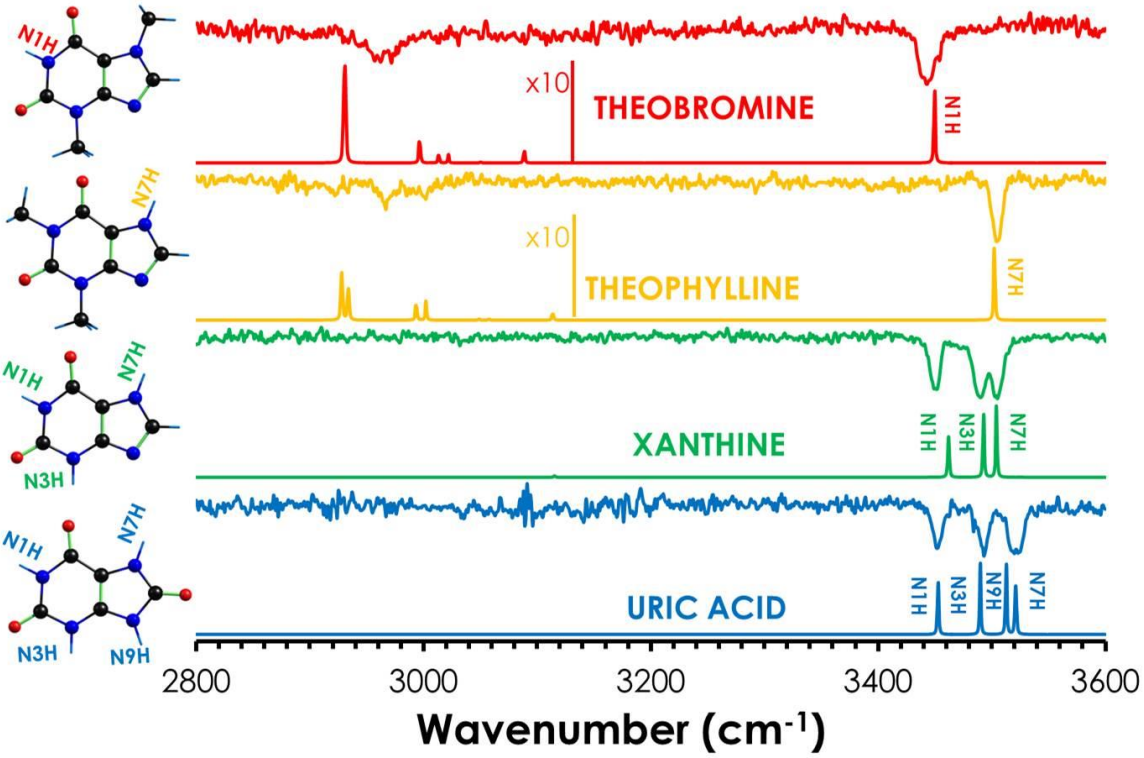


Figure 7.3.1: IDIRS spectra of the monomer of theobromine, theophylline, xanthine and uric acid compared to the relative theoretical predictions calculated at the M06-2X/6-311++G(d,p) level.

Figure 7.3.2 shows the results for the uric acid-theobromine study. Uric acid dimer, presents three conformations, all of which exhibit similar spectra, although only the most stable one is shown in the figure. Several isomers are also expected for the trimer, as the most stable structure, the one shown in the figure, would result in a different pattern regarding the free-NH stretches.

When mixed with theobromine, the number of possible conformations seems somehow reduced. For example, two conformers are expected for the 1:1 heterocluster (1THEOB:1UA), both based on formation of H-bonds, despite the

strong tendency of theobromine to form stacking structures. It seems that the ability of uric acid to form H-bonds, governs the aggregation of the two molecules. The situation becomes very complex in the case of the 1:2 (1THEOB:2UA) heterocluster, with several families contributing to the spectrum, or at least, with similar predicted spectra. Therefore, the calculations will be required to reach a conclusion, even if it can be already assert that the stacking contribution to the 1:2 heterocluster is not significant. The assignment is done by comparison of the different spectra recorded, paying special attention to the missing free-NH peaks, like theobromines N1H, uric acid N3H and uric acid N9H. The assignment of 2:1 (1THEOB:2UA) is straightforward, because only one family can reproduce these two peaks. Furthermore, this 2:1 heterocluster results from the combination of two structures proposed for the dimer. In the dimer of theobromine, the peaks relative to N3H and N9H reduced their intensity until disappearing, so the experimental spectrum recorded can be contemplated as the combination of the two corresponding conformers, in which the N3H or the N9H are interacting.

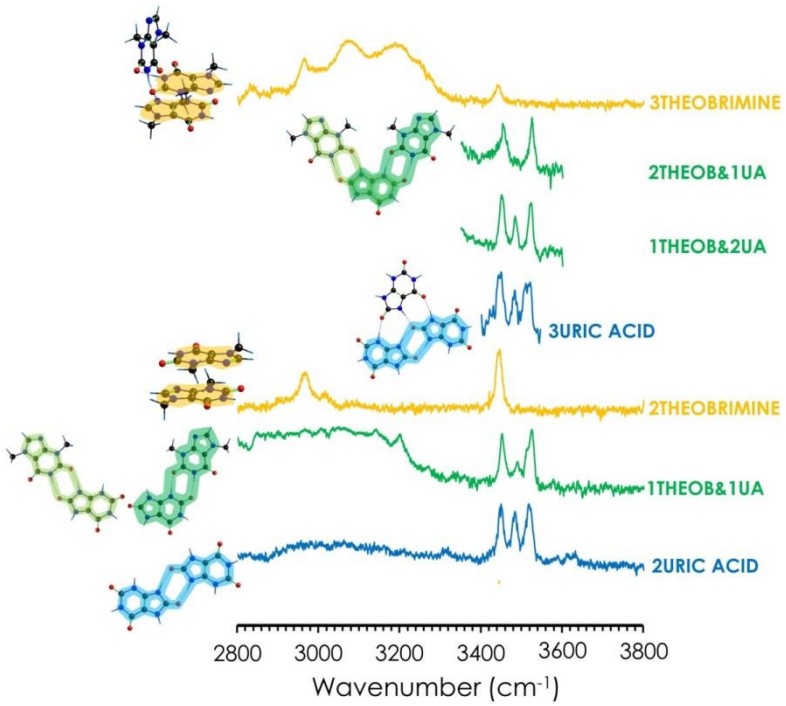


Figure 7.3.2: IDIRS spectra of theobromine and uric acid homo- and heteroclusters.

The calculations on the heterodimer (1THEOB:1UA) also confirm that the two structures displayed in Figure 7.3.2 are the two most stable ones, even if they are not the global minima in all the temperature range, but only at room temperature (in fact, at low temperature the most stable conformations are the stacking ones).

The effects that can be deduced from preliminarily calculations (since they are not finished yet), reveal a competition between the interaction of theobromine and uric acid: the crystal phase of the uric acid[18] is formed by H-bond interactions, and the crystallographic structure of theobromine[19] is organized in layers of stacking molecules. The interesting results obtained for theobromine and uric acid shed light on the preferences which rule their aggregation process.

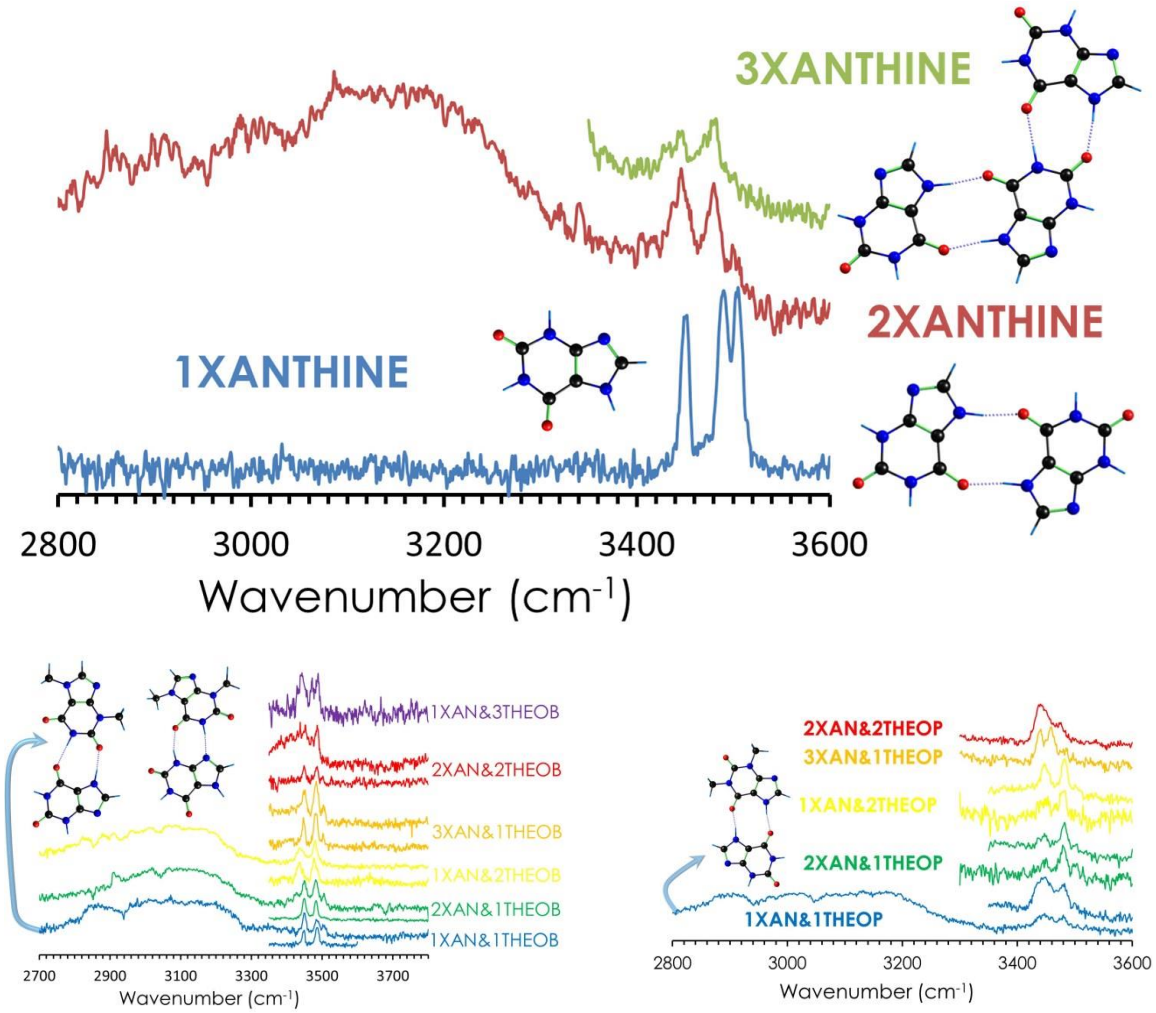


Figure 7.3.3: Xanthine nucleation and its heteronucleation with theobromine and theophylline.

Next, the spectroscopy of xanthine was also tackled, in order to add, this molecule to the study of heteronucleation, in a second stage.[12] For some reasons, probably a complicated excited state dynamics, this molecule rendered very weak signals, hindering the experimental part of the study. In order to determine the relevance of the different functional groups, the heteronucleation of xanthine-theobromine and xanthine-theophylline was analyzed, from dimers to hetero-tetramers. All the data are reported in Figure 7.3.3.

Once more, one must keep in mind that only preliminary assignments can be proposed. Xanthine dimer resembles very closely that of theophylline, being the H-bond geometries the preferred ones. As regards the trimer, it shows a completely different behavior with respect to theophylline, where the signal in the mass spectrum was weaker than for the tetramer. In the case of xanthine, it does not present "magic numbers", and the spectrum of the trimer seems to correspond to an H-bonded species, according to the relative intensities of the NH peak. One of the best candidates for the assignment of the spectrum is shown in Figure 7.3.3.

The analysis of the heteronucleation of theobromine and theophylline, unraveled a different panorama with respect to the previous case. In Figure 7.3.3, the IDIRS spectra of the detected heteroclusters are also displayed. In many cases the spectrum was recorded ionizing at two different wavenumbers: at lower frequencies using the 0_0^0 transition of theobromine or theophylline or at higher frequencies when the 0_0^0 transition of xanthine was selected. Depending on the ionization region, the recorded spectrum may be different, pointing to the existence of several species, as it happened for the 1:1 heterocluster xanthine-theobromine: the IR spectrum revealed two conformations. This was also confirmed by preliminary calculations that predicted a stacking structure for the most stable species. The spectrum of such structure is not able to reproduce the experimental traces. In fact, the presence of H-bonds is clear in the corresponding spectrum (see 1XAN&1THEOB spectrum in the bottom left box of Figure 7.3.3). Also in this case, like in the theobromine with uric acid heterocluster, the species that better reproduces the experimental spectrum is the most stable one at temperatures above 0 K, or put in other words, the entropic contribution is important for this system, even at the temperature of the beam. For the rest of the xanthine-theobromine larger clusters, it is difficult to make predictions, but the sequence of spectra does not seem to follow the same pattern as in other heteroclusters.

Regarding xanthine-theophylline, the 1:1 heterocluster shows the same interaction that theophylline and xanthine homodimers. The three of them start the aggregation in a similar way, but then they continue the nucleation with a different mechanism: forming stacking dimers in theophylline, H-bonds in xanthine, and following a mixed and more complicated mechanism in the case of the heteronucleation process.

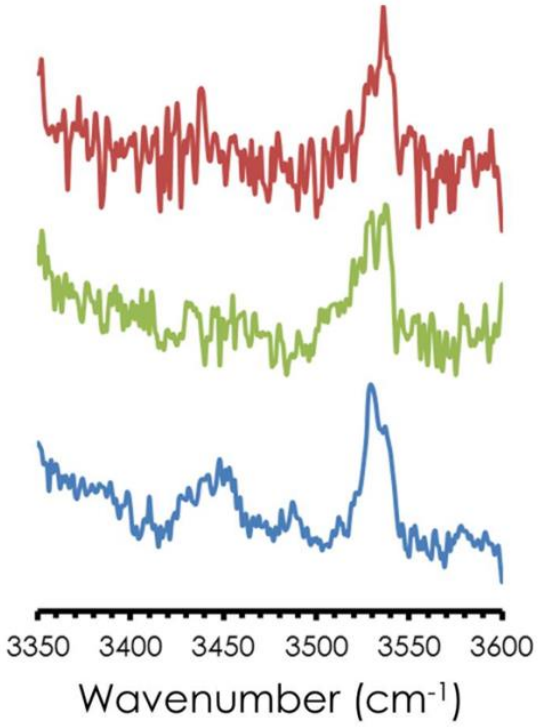


Figure 7.3.4: IDIRS spectra of the 1:1 cluster between cytosine and theophylline, ionizing at three different wavenumbers.

Heterocluster of xanthines with DNA bases were also investigated. In particular, the complex between theophylline and cytosine was analyzed. In Figure 7.3.4 the IDIRS spectra of this dimer are presented, ionizing at three different wavenumbers.[5,14,20] However, no conclusions can be reached without the calculations.

7.4. Flexible Nucleation

The combination of flexible molecules, like sugars, and rigid molecules, like xanthines, was also investigated.[21] In Figure 7.4 the complexes between phenylglucose (PhGlc) and phenylgalactose (PhGal) with theobromine and theophylline are shown. As can be seen, the IR spectra of some heterodimers and heterotrimers of one or two units of xanthine with a molecule of a sugar derivative were obtained. The assignment of these complexes is rather complicate, specially without calculations. At first glance, the two theobromine molecules seem to interact with the sugar, instead of forming the stacked structure already described for the theobromine dimer.

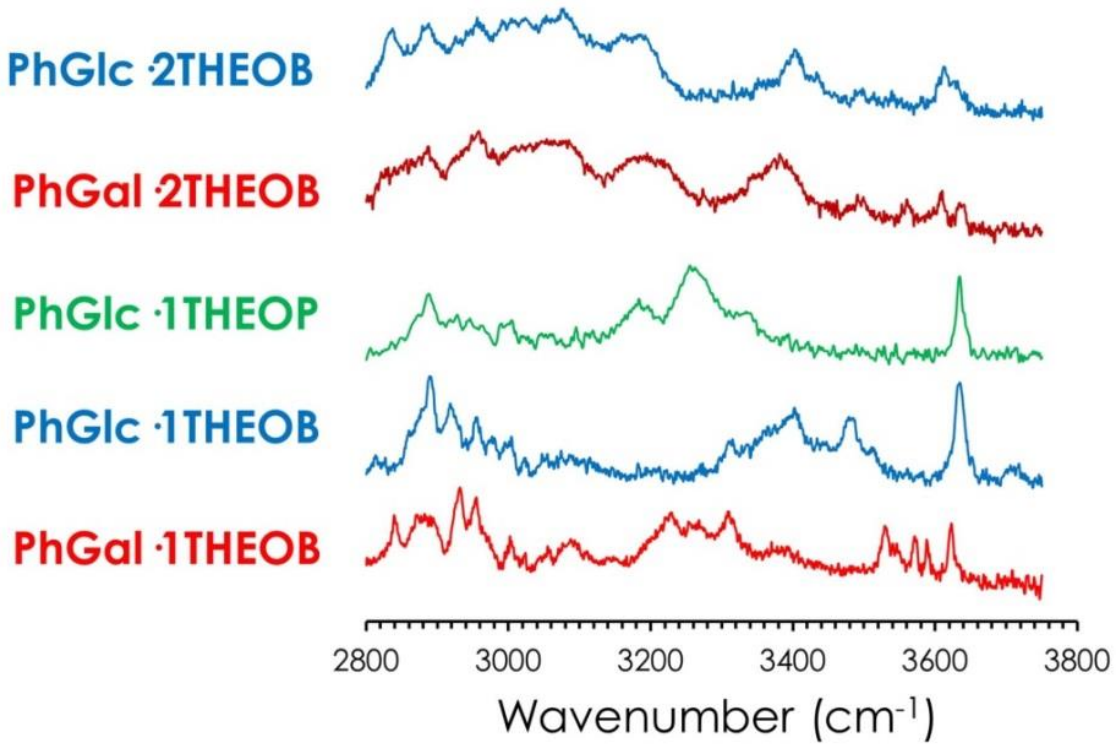


Figure 7.4: IDIRS spectra of heteroclusters of theobromine and theophylline with PhGlc/PhGal sugars.

7.5. Other Complexes

We also extended the studies about sugars aggregation to phenyl-β-D-galactopyranose (PhGal), to evaluate the influence of the phenyl substituent in the aggregation mechanism taking place during the formation of several dimers.[22]

A first set of investigated aggregates were the homo-dimers of phenyl-β-D-glucopyranose (PhGlc) and PhGal. To add information of the interactions which characterize these systems, also the complexes of PhGal with phenol and methyl-β-D-glucopyranose were analyzed. All the spectra obtained are collected in Figure 7.5.1.

The aggregation preferences of caffeine, the fully-methylated isomer of xanthine, were also investigated.[23–26] This molecule is one of most popular legal drugs, and here the complexes with phenol and PhGlc were investigated in order to increase the information about its binding and interaction preferences.[27] In Figure 7.5.2 the spectra recorded are reported. As caffeine does not present free-NH stretches, only the CH stretches appear in this region of the spectrum. Regarding the

complex with phenol, three new bands appeared that can only be due to the phenolic OH stretch and therefore indicate that at least three different isomers are formed. Unfortunately, the spectrum of the complex with PhGlc requires of calculations for its interpretation.

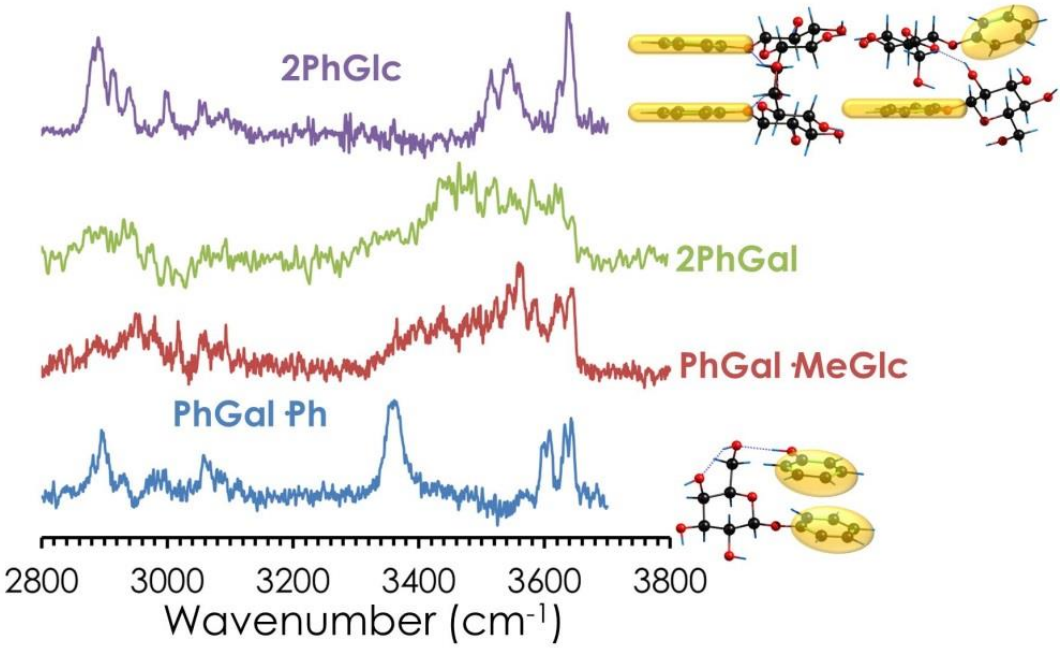


Figure 7.5.1: IDIRS spectra of different phenyl sugars dimers: homodimers of phenyl-β-D-glucopyranose (2PhGlc) and phenyl-β-D-galactopyranose (2PhGal), heterodimer of phenyl-β-D-galactopyranose with methyl-β-D-glucopyranose (PhGal·MeGlc), and heterodimer of PhGal with phenol (PhGal·Ph).

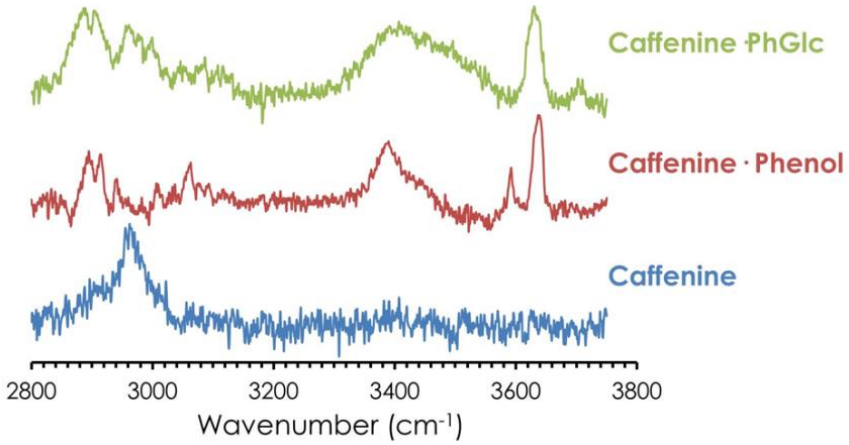


Figure 7.5.2: IDIRS spectra of caffeine, caffeine···phenol and caffeine···phenyl-β-D-glucopyranose (PhGlc).

7.6. Flexible Folding

The mixture of flexible molecules and foldamers was analyzed by studying sections of glycoproteins. The glycoproteins are involved in a variety of molecular docking mechanisms and their activity is related to the strong intermolecular interactions that they produced in a very selective way.[28]

The experimental results reported in this section can be divided in two parts. The first one refers to the study of the functional part of the antigen TN. This molecule is found in the surface of cancer cells but not in normal cells and therefore may be used as a target to train the immune system to destroy tumor cells.[29] The antigen is composed of an N-acetylgalactosamine (GalNac) linked through a glycosidic bond to either a threonine or a serine that form part of a longer chain of amino acids.[30,31] In this work, the functional part of the antigen formed by GalNac-Ser or GalNac-Thr were studied. To simulate the peptidic chain, the amino acid was terminated in a peptidic bond and a benzyl substituent was also added as chromophore.

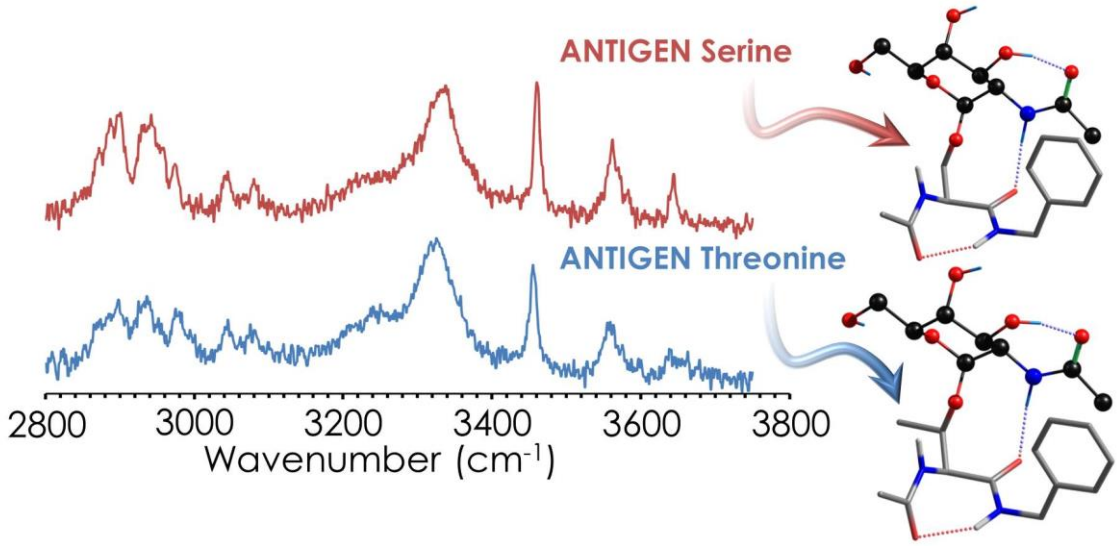


Figure 7.6.1: IDIRS spectra of GalNac-Threonine and GalNac-Serine

Threonine and serine have a very similar structure, which only differs in the position of the OH substituent.[32,33] Although it is certainly a modest difference, it results in different intermolecular interactions that can differentiate and limit the activity of these amino acids.[34] The preliminary results obtained in CI-phase revealed that both antigens adopt similar conformations (see Figure 7.6.1). The difference in their activity may be caused by the presence of the water solvent, which produces

significant changes in their structural arrangements through intermolecular interactions. This is an excellent example of the importance of the molecular flexibility and how it can influence the chemical properties.

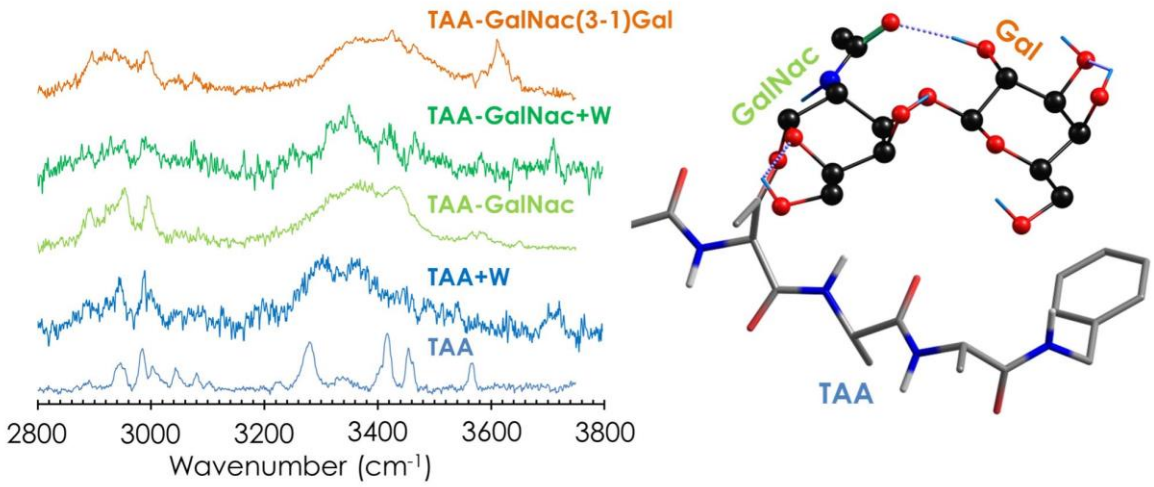


Figure 7.6.2: IDIRS spectra of TAA, TAA+W, TAA-GalNac, TAA-GalNac+W and TAA-GalNac-Gal.

The second example is based on the tripeptide already introduced in Chapter 5, composed by threonine, alanine and alanine (TAA). In this case the interaction of the TAA tripeptide with GalNac and galactopyranose (Gal), and also the corresponding water complexes, were investigated, exploring different types of macromolecules. In particular, the studied systems were: TAA, TAA+W, TAA-GalNac and TAA-GalNac, TAA-GalNac-W, and TAA-GalNac-Gal. The structure of the largest one, TAA-GalNac-Gal, is shown in Figure 7.6.2. Analyzing the results from the simplest to the most complicated one, the effects produced on the TAA folding are probed through the addition of the sugar molecules. Moreover, the water moiety, which mimics a possible chemical surrounding, helps to understand the structural investigation of TAA more in detail. Actually, its geometry adopted in solution is unfolded, and the addition of water units allows to unravel its effective arrangements also in the CI-phase.[35] The molecules like TAA are classified as antifreeze, and the final purpose related to this wide investigation is to determine how this type of molecules develop this function.[36,37]

7.7. Conclusions

As stated in the introduction, the smart experimental design is sustained by a deep knowledge of the mechanisms that govern the chemical properties. The isolation and characterization of these properties are powerful tools to gain insight into their nature, and this in turn can originate new ideas.[38] At the same time, the unraveling of particular chemical properties is fundamental for the design of new projects aimed at predicting further molecular systems on which to base the discovery of, for example, new active molecules.[8]

This thesis aims at exploring the structure of aggregates, the aggregation mechanisms and the forces that govern these processes. To achieve this goal, some relevant molecules containing biologically representative functional groups as well as their aggregates were fully characterized using an experimental-computation combined approach. The intensive experimental effort carried out allowed us to obtain a substantial amount of results, which combined with suited theoretical calculations gave rise to a new and synergistic approach where the theoretical methodology not only matches with the experiment, but contributes to outline a general view of the chemical properties of the molecules. Finally, the relation and combination with literature and scientific databases are fundamental in order to confirm the representativeness of the study performed and to increase the knowledge on the topics that this thesis deals with.

The results obtained are varied in nature and therefore, they were grouped in different chapters in order to better understand the conclusions extracted from their global analysis. In Figure 7.7, the interactions studied are classified in three general groups: flexibility, folding and nucleation. Actually, these groups are linked, since all these chemical properties are interconnected. Also, the subgroups *in progress*, will require of the combination of all the computational strategies already developed in the study of the other systems, to fully understand their behavior. Since they deal with more complex topics, development of specific approaches to fully comprehend the experimental results will be needed.

The first molecular property analyzed is flexibility, which refers to the ability of a molecule to switch between conformations to suite to the intermolecular interactions. Among the flexible molecules, sugars were chosen, in part due to their relevance in Nature. This class of compounds is very interesting also because their structure presents several degrees of freedom that allow them to adapt to

the surrounding: the ring, the hydroxyl groups, and the hydroxymethyl group that, with its flexibility represents in many cases the principal binding site.

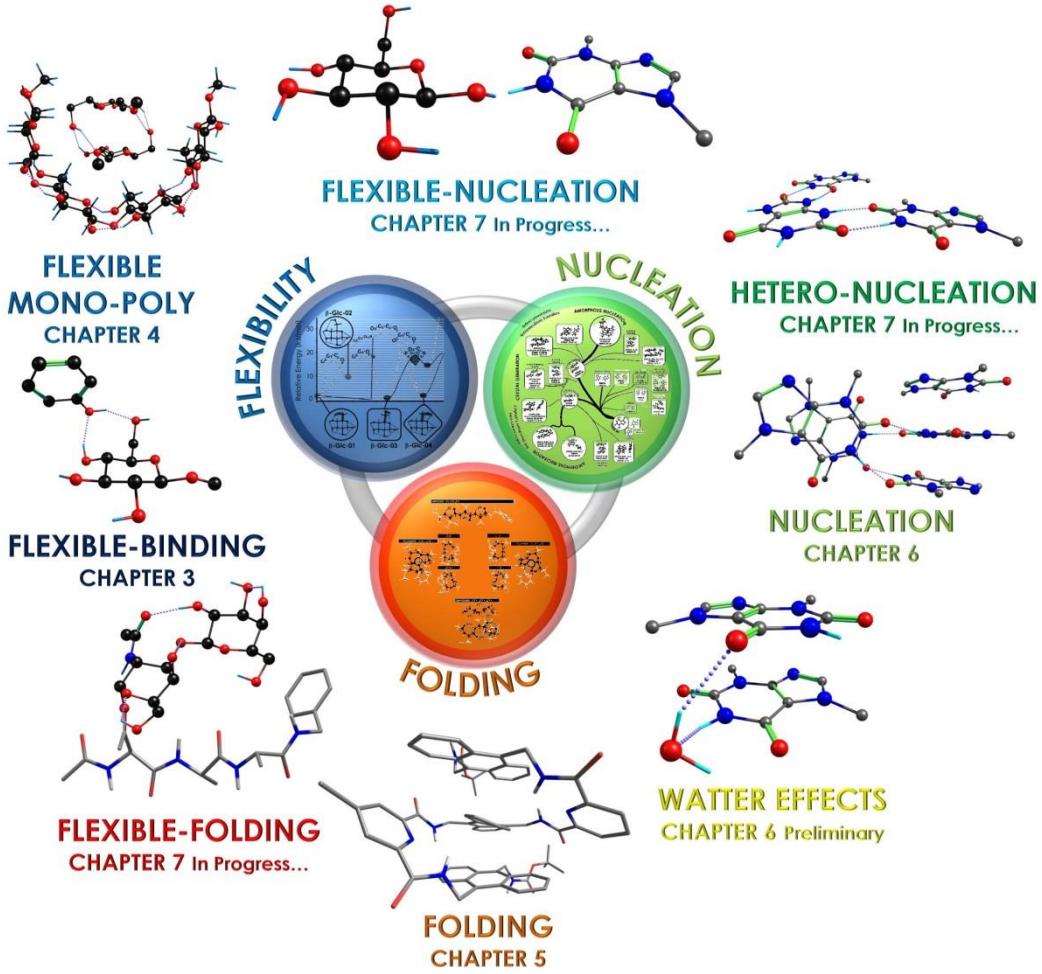


Figure 7.7: summary and classification of the studied interactions.

The effects of the flexibility in the stabilization of a complex was the main subject of Chapter 3, revealing the ability of a flexible molecule to modify the orientation of its different groups to maximize the interaction energy, moving from an orientation that stabilizes intramolecular interactions, to a new one that optimizes the intermolecular ones. Two types of energetic analysis were compared: the relative stability and the relative binding energy. Surprisingly, it was concluded that when there are several possible minima of similar relative stability, but one of them presents a stronger binding energy, the latter is favored. Thus, in specific situations the binding energy reveals as the key element to discriminate between stable geometries.[39]

Certainly, many are the aspects involved in the stabilization of the macroscopic systems, but it is useful and interesting to determine the molecular mechanisms underneath the macroscopic behavior, and to establish a link between the two worlds. In Chapter 4, the influence of the equatorial position of all the hydroxyl groups in glucose in the formation of the dimer or in the generation of cellulose crystal was demonstrated. In conclusion, the extra stabilization in β-β dimers, also observed for the corresponding linear chains and layers in the crystal, is in part due to this equatorial feature. Furthermore, the results of that study somehow explain why glucose is selected by Nature to accomplish structural functions.[40]

Flexible systems and foldamers present a very rich conformational landscape where the ability to interconvert between minima is restricted by the intramolecular bonds and the energy required to surmount potential energy barriers. Foldamers express their structural diversity through an equilibrium between a huge amount of potential structures connected by barriers of different height: the challenge is to know which the preferred one will be in each situation. Actually, the mechanism that defines the geometry of a macromolecule is very complicate and includes numerous parameters. In Chapter 5, we showed how a relatively small tripeptide folds in the CI-phase and in absence of any solvent. The results and the structure adopted in the jet cannot be directly extrapolated to its structure in solution, because in the latter case the environment is markedly different and strongly influences the final shape of the molecule, but the results from the jet reveal its intrinsic trend to fold. This information is very interesting also in order to evaluate the influence of the environment (water) and may, for example, serve to predict which structure this molecular system could adopt if it is placed in a hydrophobic environment. Continuing with the exploration of the folding process, the conformational preferences of a polymer based on anthracene units were also studied and the structural parameters that conferee the compound its high conductance were determined, attributing this specific property to the stacked interactions of the anthracene units, as it happens for graphite. The importance of this topic is validated by the excellent publication resulting from this study.[41]

Regarding the analysis of the intermolecular interactions, the first stage of the nucleation process is somehow frozen in the expansion. The complexity of this process, in which many variables are at play, required of a simplification of the system, and therefore a family of rigid molecules was chosen. In this way, the number of variables is significantly reduced, as the molecules cannot jump between different conformations, simplifying the conformational search, and

allowing us to examine aggregates with higher stoichiometry. Thus, we presented in Chapter 6 the analysis of the nucleation of theobromine. This molecule is composed by the thymine's functional group with the same basic purine skeleton. The study presents an added interest since theobromine exhibits well-known pharmacological effects. It is well known that theobromine interferes in the nucleation process of uric acid inside the human body, thanks to a yet-to-be-determined mechanism. The investigation aimed at unraveling the first instants of its nucleation process, starting from the monomer up to the trimer. Indeed, as mentioned in this chapter, the complete analysis includes up to the pentamer and also some complexes with water, but the calculations are not completed yet. Nevertheless, the results already obtained for the nucleation process point out that, out of the equilibrium in supersonic expansions, formation of an amorphous phase is achieved. Thanks to the theoretical data, the conformational panorama was analyzed, producing a *genealogical tree* (see Figure 6.4.2 in Chapter 6). The experimental geometries can be located inside this diagram, deducing the paths followed by the molecules to form the crystal and the amorphous phases, in order to obtain an overview of the basic mechanism.

Continuing with the investigation on the nucleation process, several other homo- and hetero-aggregates were isolated and their spectroscopy tackled, as reported in the previous part of this chapter. The results from those studies, that include a collection of different types of aggregates, represent another step forward not only for the understanding of the nucleation process, but also for the analysis of the interactions describing the flexible molecules, and in turn the folding process. In fact, these complexes represent the combination between the three molecular properties highlighted in Figure 7.7: flexibility, folding and nucleation, whose investigation is the main purpose of this thesis. The complexity of these systems will require of extensive calculations and probably, of the refinement of the computational strategies, and for these reasons their study is still in progress.

Finally, one would like to remark that the purpose of this work was not only the characterization of some molecular systems, but also to create a basic reference manuscript, which treats different kinds of relevant molecules (sugars, foldamers and some puric base derivatives), able to introduce any researcher interested in these topics, in the search for the molecular mechanisms underlying fundamental chemical properties.

The basic scientific research is usually considered far from any kind of applications, but this thesis project tries, with the figures, the summaries offered and the

extremely processed data, to facilitate the understanding of the results and observed trends, in an attempt to inspire the improved design of new experiments.

7.8. References

(1) Henninger, H. B.; Reese, S. P.; Anderson, A. E.; Weiss, J. A. Validation of Computational Models in Biomechanics. *Proc. Inst. Mech. Eng. H.* **2010**, *224* (7), 801–812.

(2) Johannsen, M.; Brunner, G. Solubilities of the Xanthines Caffeine, Theophylline and Theobromine in Supercritical Carbon Dioxide. *Fluid Phase Equilib.* **1994**, *95*, 215–226.

(3) Li, S.; Varadarajan, G. S.; Hartland, S. Solubilities of Theobromine and Caffeine in Supercritical Carbon Dioxide: Correlation with Density-Based Models. *Fluid Phase Equilib.* **1991**, *68*, 263–280.

(4) Danilov, V. I.; Slyusarchuk, O. N.; Poltev, V. I.; Alderfer, J. L.; Wollman, R. M.; Brickmann, J. A. W.; Lautenschlager, P. A Monte Carlo Simulation of Hydration of Xanthine-Derivatives and Their Stacked Forms. *J. Biomol. Struct. Dyn.* **1992**, *9* (6), 1239–1252.

(5) Callahan, M. P.; Gengeliczki, Z.; Svadlenak, N.; Valdes, H.; Hobza, P.; de Vries, M. S. Non-Standard Base Pairing and Stacked Structures in Methyl Xanthine Clusters. *Phys. Chem. Chem. Phys.* **2008**, *10* (19), 2819–2826.

(6) Rodrigues, M. A.; Padrela, L.; Geraldes, V.; Santos, J.; Matos, H. A.; Azevedo, E. G. Theophylline Polymorphs by Atomization of Supercritical Antisolvent Induced Suspensions. *J. Supercrit. Fluids* **2011**, *58* (2), 303–312.

(7) Nafisi, S.; Manouchehri, F.; Tajmir-Riahi, H.-A.; Varavipour, M. Structural Features of DNA Interaction with Caffeine and Theophylline. *J. Mol. Struct.* **2008**, *875* (1), 392–399.

(8) Byrn, S.; Pfeiffer, R.; Ganey, M.; Hoiberg, C.; Poochikian, G. Pharmaceutical Solids: A Strategic Approach to Regulatory Considerations. *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists*. 1995, pp 945–954.

(9) Grases, F.; Rodriguez, A.; Costa-Bauza, A. Theobromine Inhibits Uric Acid Crystallization. A Potential Application in the Treatment of Uric Acid Nephrolithiasis. *PLoS One* **2014**, *9* (10), 1–6.

(10) Sugawara, M.; Mochizuki, T.; Takekuma, Y.; Miyazaki, K. Structure-Affinity Relationship in the Interactions of Human Organic Anion Transporter 1 with Caffeine, Theophylline, Theobromine and Their Metabolites. *Biochim. Biophys. Acta - Biomembr.* **2005**, *1714* (2), 85–92.

(11) Gouda, H.; Kuntz, I. D.; Case, D. A.; Kollman, P. A. Free Energy Calculations for Theophylline Binding to an RNA Aptamer: Comparison of MM-PBSA and Thermodynamic Integration Methods. *Biopolymers* **2003**, *68* (1), 16–34.

(12) Ucun, F.; Saglam, A.; Guçlu, V. Molecular Structures and Vibrational Frequencies of Xanthine and Its Methyl Derivatives (Caffeine and Theobromine) by Ab Initio Hartree-Fock and Density Functional Theory Calculations. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2007**, *67* (2), 342–349.

(13) Doré, A. S.; Robertson, N.; Errey, J. C.; Ng, I.; Hollenstein, K.; Tehan, B.; Hurrell, E.; Bennett, K.; Congreve, M.; Magnani, F.; et al. Structure of the Adenosine A 2A Receptor in Complex with ZM241385 and the Xanthines XAC and Caffeine. *Structure* **2011**, *19* (9), 1283–1293.

(14) Zimmermann, G. R.; Jenison, R. D.; Wick, C. L.; Simorre, J.-P.; Pardi, A. Interlocking Structural Motifs Mediate Molecular Discrimination by a Theophylline-Binding RNA. *Nature* **1997**, *4*, 686–689.

(15) Johnson, I. M.; Prakash, H.; Prathiba, J.; Raghunathan, R.; Malathi, R. Spectral Analysis of Naturally Occurring Methylxanthines (Theophylline, Theobromine and Caffeine) Binding with DNA. *PLoS One* **2012**, *7* (12), 1–11.

(16) Asami, H.; Urashima, S.; Saigusa, H. Structural Identification of Uric Acid and Its Monohydrates by IR-UV Double Resonance Spectroscopy. *Phys. Chem. Chem. Phys.* **2011**, *13* (45), 20476–20480.

(17) Jiménez, V.; Alderete, J. B. Theoretical Calculations on the Tautomerism of Uric Acid in Gas Phase and Aqueous Solution. *J. Mol. Struct. THEOCHEM* **2005**, *755* (1), 209–214.

(18) Ringertz, H. The Molecular and Crystal Structure of Uric Acid. *Acta Crystallogr.* **1966**, *20* (2), 397–403.

(19) Ford, K. A.; Ebisuzaki, Y.; Boyle, P. D. Methylxanthines. II. Anhydrous Theobromine. *Acta Crystallogr. Sect. C Cryst. Struct. Commun.* **1998**, *54* (12), 1980–1983.

(20) Nir, E.; Hunig, I.; Kleinermanns, K.; de Vries, M. S. The Nucleobase Cytosine and the Cytosine Dimer Investigated by Double Resonance Laser Spectroscopy and Ab Initio Calculations. *Phys. Chem. Chem. Phys.* **2003**, *5* (21), 4780.

(21) Maresca, M.; Derghal, A.; Carravagna, C.; Dudin, S.; Fantini, J. Controlled Aggregation of Adenine by Sugars: Physicochemical Studies, Molecular Modelling Simulations of Sugar-Aromatic CH-P Stacking Interactions, and Biological Significance. *Phys. Chem. Chem. Phys.* **2008**, *10* (19), 2792–2800.

(22) Usabiaga, I.; González, J.; Arnáiz, P. F.; León, I.; Cocinero, E. J.; Fernández, J. A. Modeling the Tyrosine-Sugar Interactions in Supersonic Expansions: Glucopyranose-Phenol Clusters. *Phys. Chem. Chem. Phys.* **2016**, *18* (18), 12457–12465.

(23) Cesaro, A.; Starec, G. Thermodynamic Properties of Caffeine Crystal Forms. *J. Phys. Chem.* **1980**, *84* (11), 1345–1346.

(24) Weber Brun, G.; Martín, Á.; Cassel, E.; Vargas, R. M. F.; Cocero, M. J. Crystallization of Caffeine by Supercritical Antisolvent (SAS) Process: Analysis of Process Parameters and Control of Polymorphism. *Cryst. Growth Des.* **2012**, *12* (4), 1943–1951.

(25) Ribeiro, J. A.; Sebastio, A. M. Caffeine and Adenosine. *J. Alzheimer's Dis.* **2010**, *20* (SUPPL.1).

(26) Tang-Liu, D. D.; Williams, R. L.; Riegelman, S. Disposition of Caffeine and Its Metabolites in Man. *J. Pharmacol. Exp. Ther.* **1983**, *224* (1), 180 LP-185.

(27) Larsen, R. W.; Jasuja, R.; Hetzler, R. K.; Muraoka, P. T.; Andrada, V. G.; Jameson, D. M. Spectroscopic and Molecular Modeling Studies of Caffeine Complexes with DNA Intercalators. *Biophys. J.* **1996**, *70* (1), 443–452.

(28) Cole, C. R.; Smith, C. A. Glycoprotein Biochemistry (Structure and Function) — a Vehicle for Teaching Many Aspects of Biochemistry and Molecular Biology. *Biochem. Educ.* **1989**, *17* (4), 179–189.

(29) Brockhausen, I.; Schachter, H.; Stanley, P. Chapter 9 O-GalNAc Glycans. In *Essentials of Glycobiology. 2nd edition.*; Varki, A., Cummings, R., Esko, J., Eds.; Cold Spring Harbor (NY), 2009.

(30) Singhal, A.; Fohn, M.; Hakomori, S. Induction of a - N -Acetylgalactosamine- O -Serine / Threonine (Tn) Antigen-Mediated Cellular Immune Response for Active Immunotherapy in Mice Induction of a-W-Acetylgalactosamine-O-Serine / Threonine (Tn) Antigen-Mediated Cellular Immune Response for. *Cancer Res.* **1991**, *51*, 1406–1411.

(31) Wu, A. M. Expression of Binding Properties of Gal/GalNAc Reactive Lectins by Mammalian Glycotopes. In *The Molecular Immunology of Complex Carbohydrates ---2*; Wu, A. M., Ed.; Springer US: Boston, MA, 2001; pp 55–64.

(32) Alonso, J. L.; Perez, C.; Eugenia Sanz, M.; Lopez, J. C.; Blanco, S. Seven Conformers of L-Threonine in the Gas Phase: A LA-MB-FTMW Study. *Phys. Chem. Chem. Phys.* **2009**, *11* (4), 617–627.

(33) Blanco, S.; Sanz, M. E.; López, J. C.; Alonso, J. L. Revealing the Multiple Structures of Serine. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (51), 20183–20188.

(34) Zhang, W.; Laursen, R. A. Structure-Function Relationships in a Type I Antifreeze Polypeptide: THE ROLE OF THREONINE METHYL AND HYDROXYL GROUPS IN ANTIFREEZE ACTIVITY . *J. Biol. Chem.* **1998**, *273* (52), 34806–34812.

(35) Narayanan Krishnamoorthy, A.; Holm, C.; Smiatek, J. Local Water Dynamics around Antifreeze Protein Residues in the Presence of Osmolytes: The Importance of Hydroxyl and Disaccharide Groups. *J. Phys. Chem. B* **2014**, *118* (40), 11613–11621.

(36) Jia, Z.; Davies, P. L. Antifreeze Proteins: An Unusual Receptor-Ligand Interaction. *Trends Biochem. Sci.* **2002**, *27* (2), 101–106.

(37) Haymet, A. D. J.; Ward, L. G.; Harding, M. M. Winter Flounder "Antifreeze" Proteins: Synthesis and Ice Growth Inhibition of Analogues That Probe the Relative Importance of Hydrophobic and Hydrogen-Bonding Interactions. *J. Am. Chem. Soc.* **1999**, *121* (5), 941–948.

(38) De Vries, M. S.; Hobza, P. Gas-Phase Spectroscopy of Biomolecular Building Blocks. *Annu. Rev. Phys. Chem* **2007**, *58*, 585–612.

(39) Usabiaga, I.; Gonzalez, J.; Arnaiz, P. F.; Leon, I.; Cocinero, E. J.; Fernandez, J. A. Modeling the Tyrosine-Sugar Interactions in Supersonic Expansions: Glucopyranose-Phenol Clusters. *Phys. Chem. Chem. Phys.* **2016**, *18* (18), 12457–12465.

(40) Usabiaga, I.; Gonzalez, J.; Leon, I.; Arnaiz, P. F.; Cocinero, E. J.; Fernandez, J. A. Influence of the Anomeric Conformation in the Intermolecular Interactions of Glucose. *J. Phys. Chem. Lett.* **2017**, *8* (6), 1147–1151.

(41) Carini, M.; Ruiz, M. P.; Usabiaga, I.; Fernández, J. A.; Cocinero, E. J.; Melle-Franco, M.; Diez-Perez, I.; Mateo-Alonso, A. High Conductance Values in π-Folded Molecular Junctions. *Nat. Commun.* **2017**, *8* (May), 15195.

Appendix for Chapter 2

Experimental Setup and Methodology

2.1. Experimental Setup

2.3. Theoretical Methodology

Table A2.1: Delays Configuration

DELAY&CHANNEL	CONFIGURATION	DESCRIPTION
Delay 1 (D1)	T_0=Int Rate= 10Hz	Operating in internal mode; it is the principal delay that controls the valve pulse and the ablation laser
A (D1/A)	A=T_0+ **(500µs adj.)**	Position of the valve pulse with respect to all the other events **(adjustable in A)** equivalent to **(C)**
B (D1/B)	B=A+ **(~240µs adj.)**	Duration of valve opening: **(adjustable in B)**
C (D1/C)	C= T_0+ **(~482µs adj.)**	IOTA ONE configuration: External cycle Connection: A ⎍ B → input TTL* Flash lamp (FL) Connection: C → Ext Trig* Q-Switch (QS) Connection: D → Ext Q-Sw* Connected to Quantel ULTRA; Configuration: HV Test **O**; Q-Sw Trig **I**; Frontal Display freq: 00
D (D1/D)	D=C+230µs	It adjusts the moment of the ablation laser shot during the valve opening: **(adjustable in C)**
Delay 2 (D2)	T_0=Ext	Connection: (D1)C ⎍ D →(D2)EXT TRIG* Control of 2 UV/Laser
A (D2/A)	A=B-220µs	Flash lamp UV Laser 1 & 2 Connection: A→ Flashlamp Synchro in* (Laser 1) Connection: A ⎍ B → Flashlamp Synchro in* (Laser 2)
B (D2/B)	B=T_0+ **(~380µs adj.)**	Q-Switch UV Laser 1 Connection: B→ Q-Switch Synchro in* It adjusts the UV ionization Laser with respect to the ablation and the valve **(adjustable in B)**
C (D2/C)	C=B- **215,05µs**	**Fixed IR beam 50 ns before UV beam** T_0 OPO/IR Laser; Connection: (D2)C →(D3)EXT TRIG* T_0 FA/IR Laser; Connection: (D2)C →(D4)EXT TRIG* SCOPE&PULSE DIVISOR; Connection: C ⎍ D → Trig. In* Scope & Pulse Divisor configuration: ⎿ ↓
D (D2/D)	D=B± (**Variable**) ns	Q-Switch UV Laser 2 Connection: D→ Q-Swich Synchro in* The second UV laser; timing varies depending on the experiment that is carrying out **(adjustable in D)**
Delay 3 (D3)	T_0=Ext	Connection: (D2)C ⎍ D →(D3)EXT TRIG* Control of OPO IR/Laser
A (D3/A)	A=T_0+ 25µs	Flash lamp IR Laser
B (D3/B)	B=A+ 10µs	Connection: A ⎍ B → FL in*
C (D3/C)	C=A+ 190µs	Q-Switch IR Laser
D (D3/D)	D=C+ 10µs	Connection: A ⎍ B → FL in* or 5Hz frequency divisor
Delay 4 (D4)	T_0=Ext	Connection: (D2)C ⎍ D →(D4)EXT TRIG* Control of Dye IR/Laser
A (D4/A)	A=T_0+ 0µs	Flash lamp IR Laser Connection: A → FL TRIG.*(in the power supply) Connection, signal active subtraction: (D2)C ⎍ D → FL TRIG.*(in the power supply)
B (D4/B)	B=A+ 215µs	Q-Switch IR Laser Connection: back delay out B → PC_Tgrigger* in the Laser back delay out B configuration: 50Ω; VAR: Amplitud 1V; offset: 0V
C (D4/C)	C=A+ 0000µs	Operating at 10 Hz
C (D4/C)	C=A+ 150000µs	Operating at 5 Hz signal active subtraction

*Name of the output, →(connected to) name of the other instrument input

Figure A2.1: General view of the layout of the laboratory devices

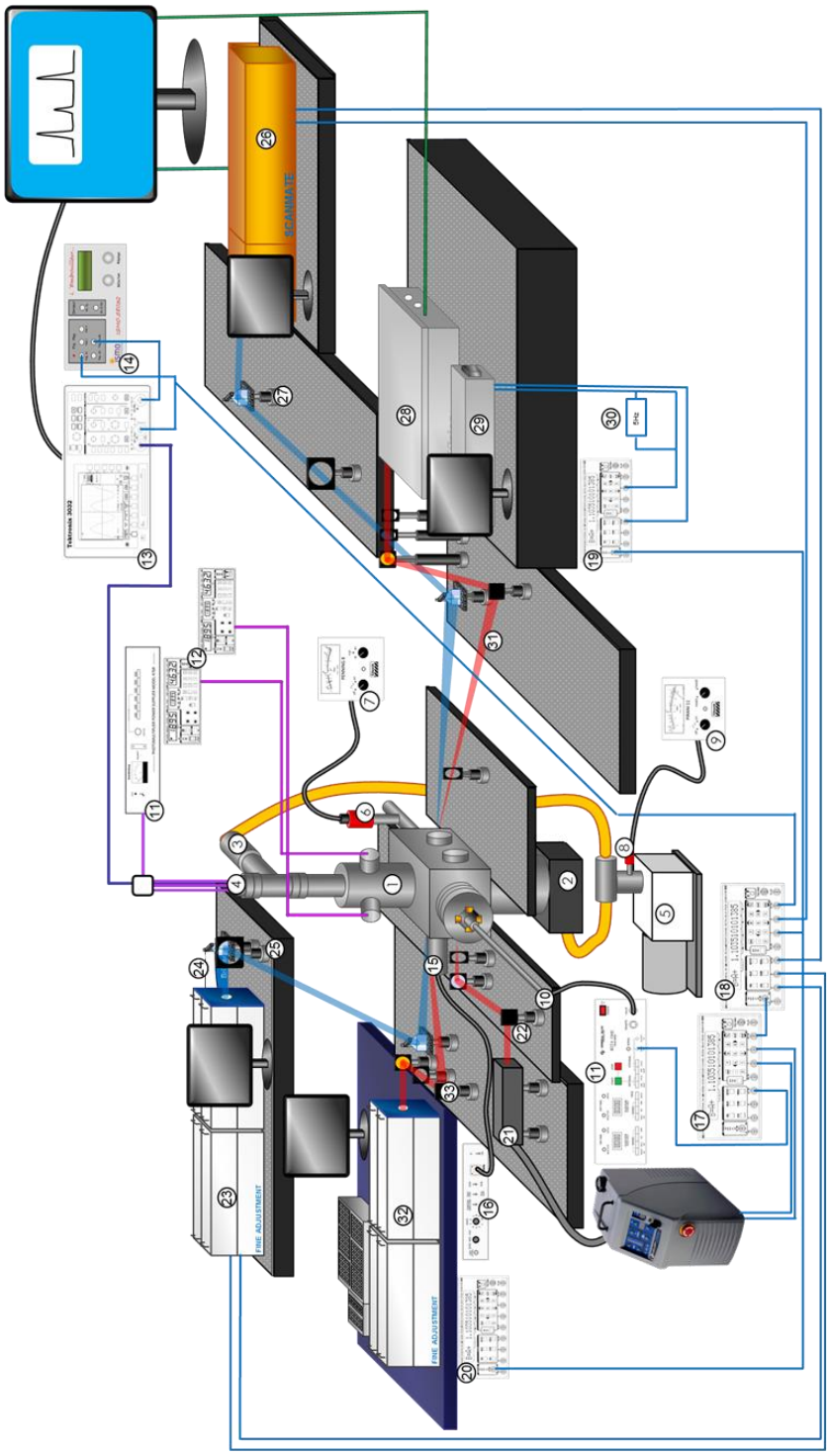


Figure A2.2: Example of the flexibility in a four-membered ring: cyclobutanol. By adding a substituent like the OH, part of the symmetry of the molecular system is broken. The diagram shows the interconversion pathway between two twisted geometries of the ring: the most stable one has the OH group in equatorial position (conformation 1 and 2 in the diagram), while the less stable geometry presents the OH group in axial position (conformation 3 and 4 in the diagram). Eq=equatorial; Ax=Axial; G=Gauche; A=Anti.

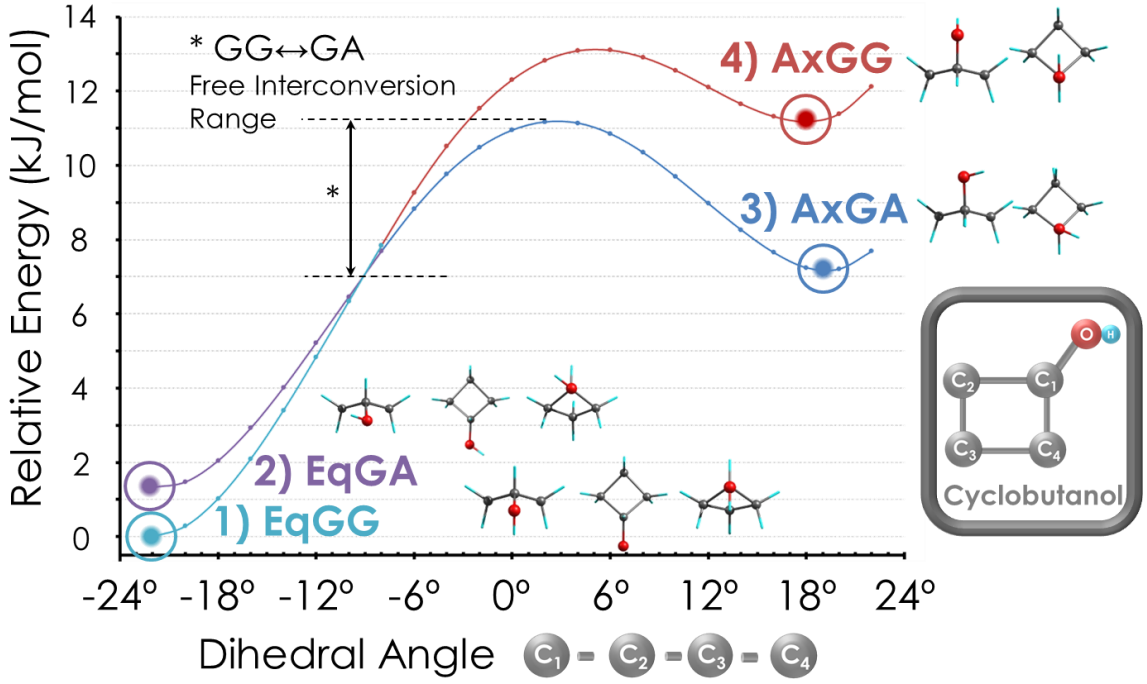


Figure A2.3: The case of cyclopentanol as an example of the flexibility in a five-membered ring: cyclopentanol. **a)** Scan of one dihedral angle, the second one is free to optimize and reach the most stable position at each point. The data obtained were tested using the Cremer-Pople parameters in order to evaluate the accuracy of the scan and determine if it is useful to get the interconversion path between the axial (Ax) and the equatorial (Eq) position of the OH substituent. **b)** The interconversion pathway between Eq- and Ax-hydroxyl groups: the Gauche-Gauche (GG) position of the OH group with respect to the ring, produces only one interconversion path (red line). The Gauche-Anti (GA) geometry breaks the symmetry producing two possible interconversion pathways (blue lines).

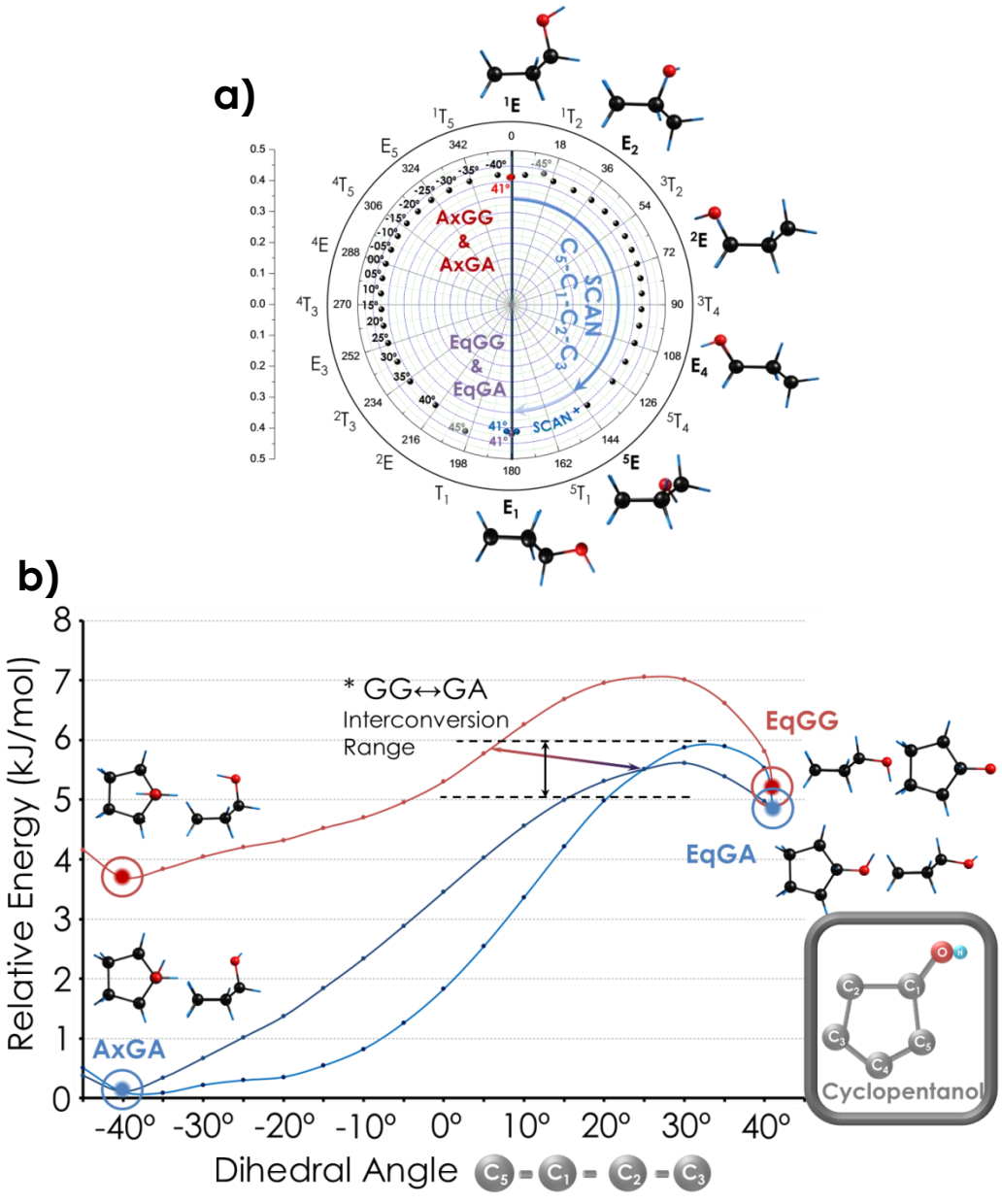


Figure A2.4: Interconversion path for the cyclohexanol molecule, from the chair to the twist geometry. This interconversion path is very complex and in some points there is a gap produced by a second interconversion coordinate. The main coordinate along the x axis is $C_6C_1C_2C_3$, but this presents a jump from -18° to -20°. In order to fill this gap, the corresponding coordinate was fixed at -19° approximately, and the secondary coordinate $C_1C_2C_3C_4$ was scanned from the beginning of the jump (corresponding to an angle of 5°) to a final angle of 63°. For the second jump that takes place between -52° and -54°, the same approach was used.

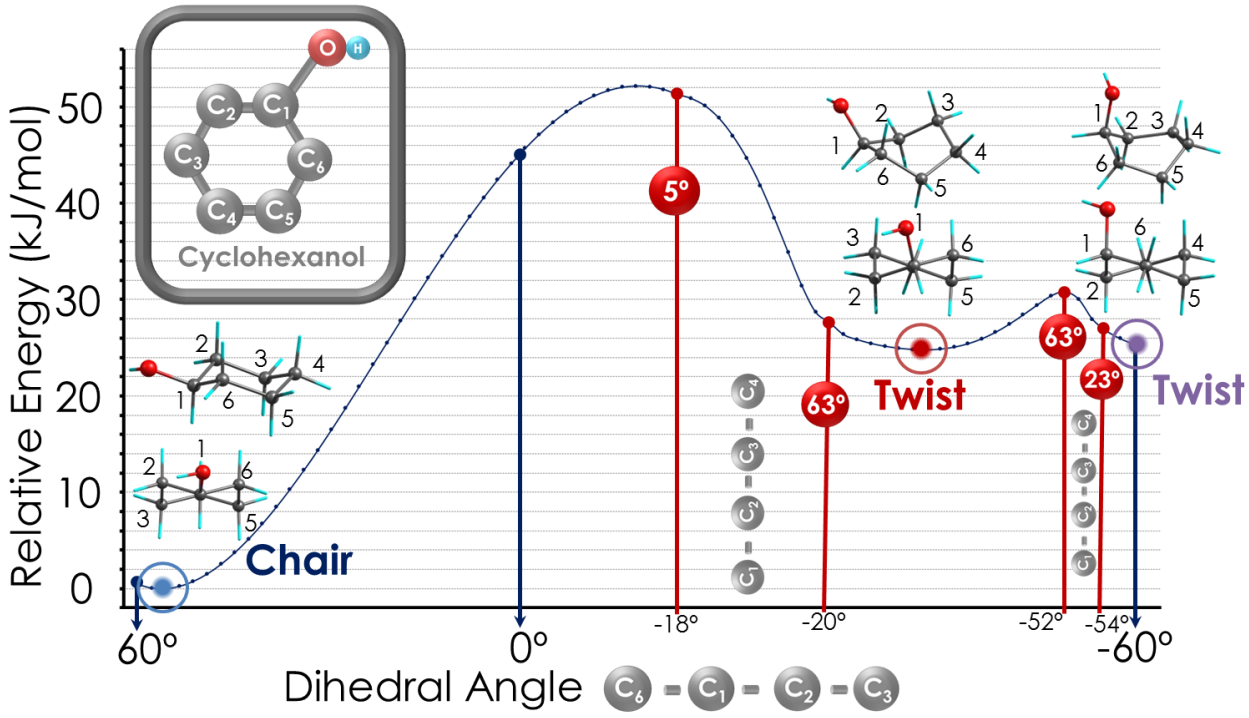
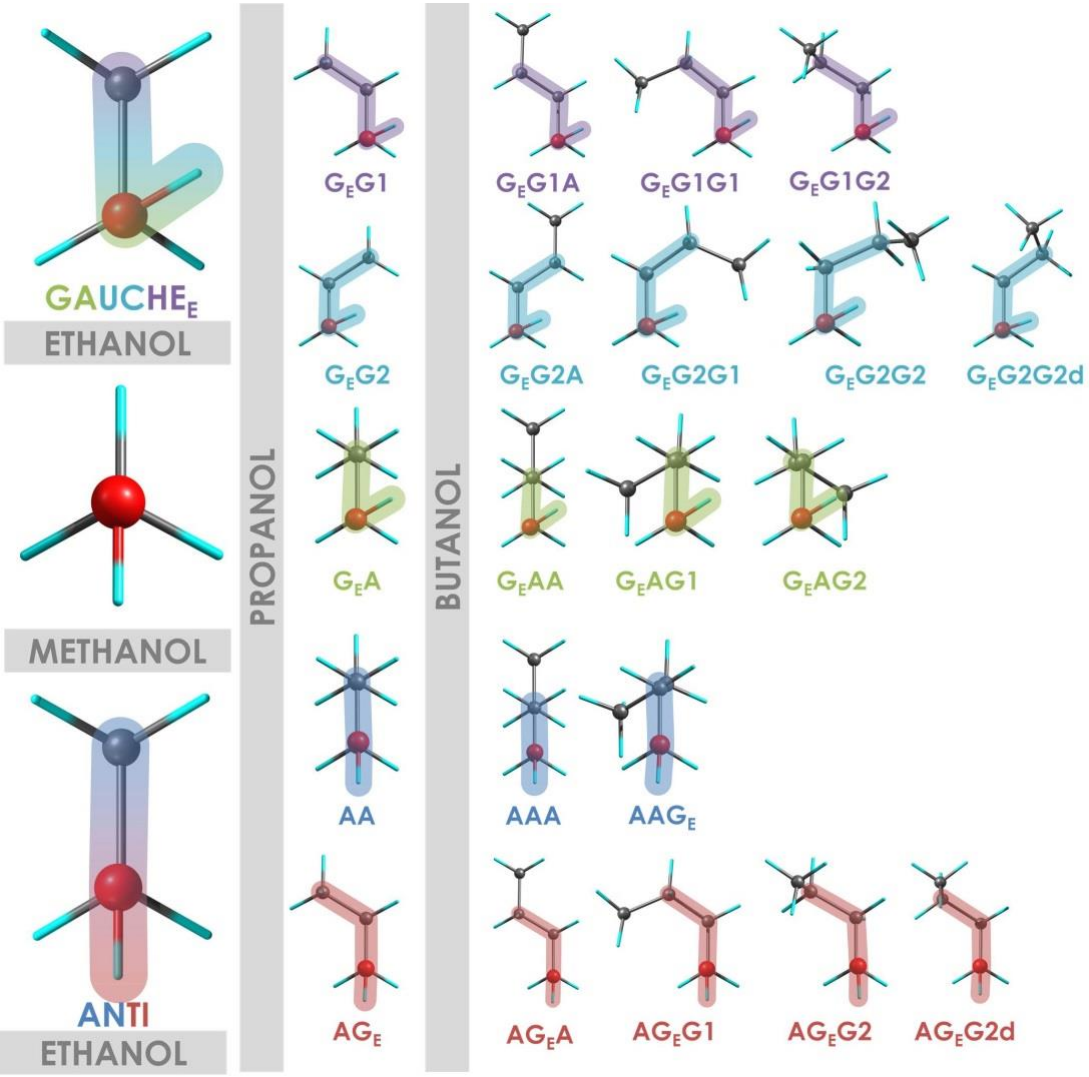


Figure A2.5: Conformational flexibility in linear alcohols as an example of the isomeric richness in linear molecules. The terminal -CH_3 produces three equivalent positions or equivalent isomers so this contribution is rejected by symmetry. The alcohols have also symmetrical properties that simplify their conformational landscape: indeed in this figure the difference between enantiomers, dextrorotation (D) and levorotation (L), was omitted and it was considered only one enantiomer (as indicated by the subscript E in the figure). Despite all these simplifications, it is obvious that the isomeric richness increases with the elongation of the chain. Starting from the first dihedral angle, (H-O-C-C) the case of ethanol presents two isomers in anti (A) and gauche (G) positions. Propanol with two dihedral angles produced by the addition of one methyl to the ethanol chain (H-O-C-C-C) presents five isomers. Finally, butanol (H-O-C-C-C-C) with three dihedral angles should present fourteen isomers but it presents at least sixteen isomers, because some folded structures where the gauche minimum, that should be close to 60°, split in two minima: one approximately at 30° and the other one at 90° (labeled with d in the figure).



Appendix for Chapter 3

Intermolecular Interactions in Glucose · Phenol Complexes: the "*Proven Survivors*"

3.2. Experimental Results

3.3. Discussion

Figure A3.1: Methyl-α-D-Glucopyranose·phenol spectroscopic results: **a)** REMPI spectrum of α-MeGlc·P complex and IDIRS spectra tuning the UV laser at several ionization wavelengths: the shapes of the spectra change with the ionization wavelength used; **b)** IR/UV Hole Burning experiment; **c_1)** IDIRS spectrum obtained tuning the UV laser at 36539 and 36714 cm^{-1}; **c_2)** IDIRS spectrum obtained tuning the UV laser at 36318 cm^{-1}; **c_3)** spectrum obtained subtracting trace (c_1) from (c_2); **c)** the experimental spectra described above (c_1, c_2, c_3 traces) are reported in black, while the corresponding simulated ones of the more stable isomers are displayed in different colors. M06-2X/6-311++G(d,p) calculation level was used, with correction factors of 0.953 for CH, and of 0.939 for OH; **d)** structures of the α-MeGlc·P most stable conformers.

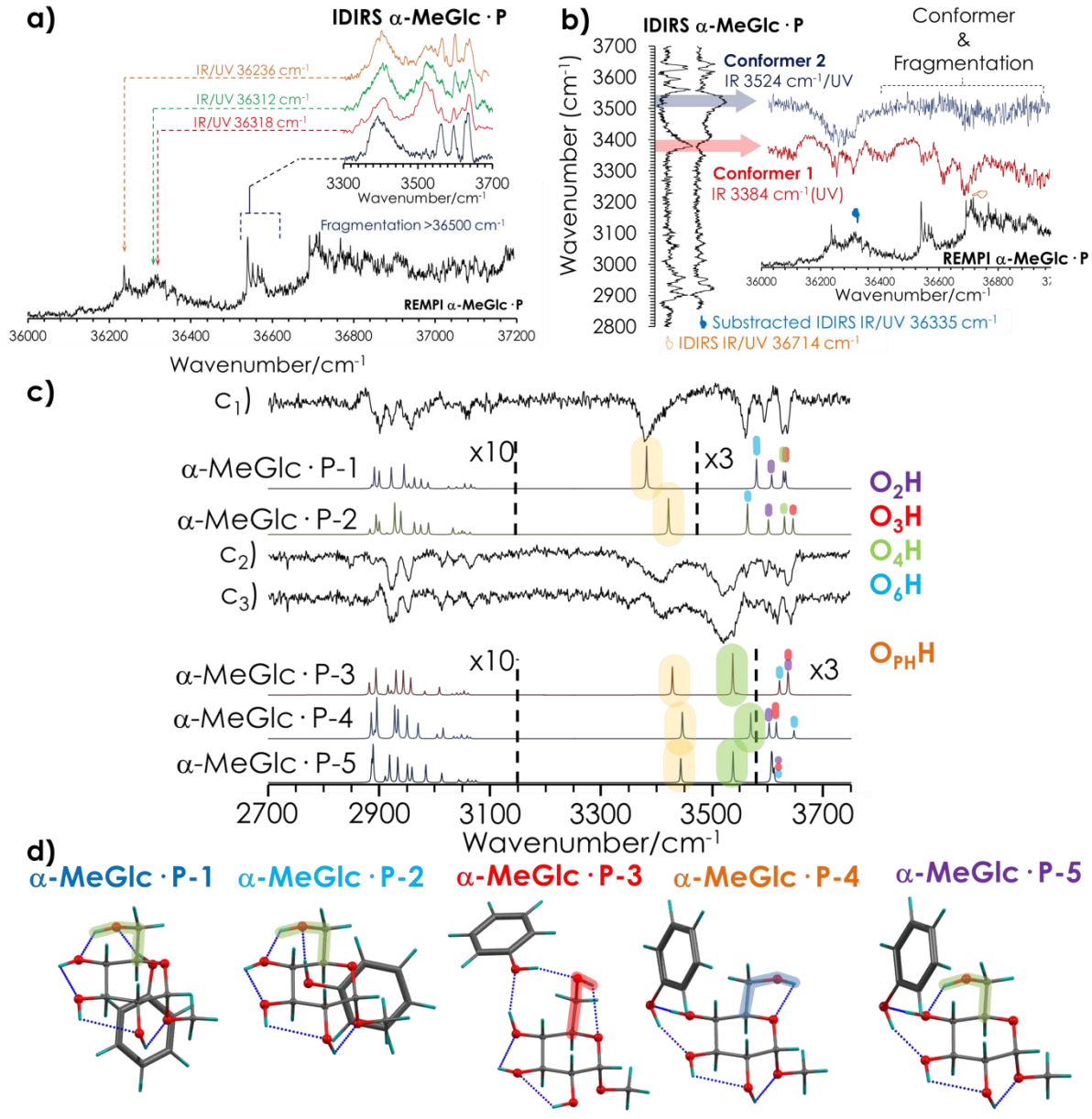


Figure A3.2: Phenyl-β-D-Glucopyranose·phenol spectroscopic results: **a)** REMPI spectrum of β-PhGlc·P complex and UV/UV Hole Burning experiment; **b)** comparison between IDIRS recorded tuning the UV laser at 36892 cm^{-1} and simulated spectra for some selected structures, computed at M06-2X/6-311++G(d,p) level. Correction factors used of 0.953 and 0.939 for CH and OH, respectively; **c)** structures of the β-MeGlc·P more stable conformers.

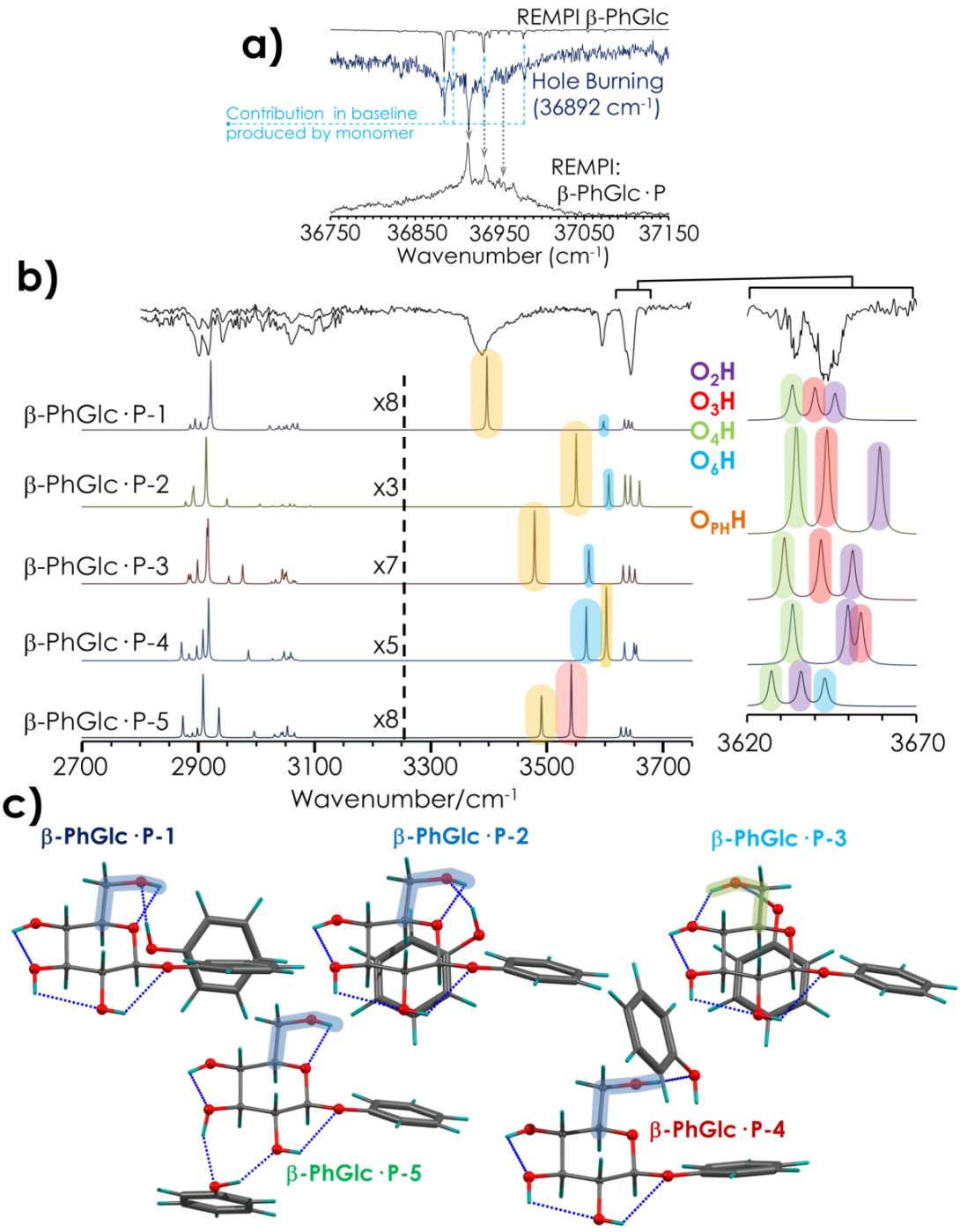


Figure A3.3: Methyl-β-D-Glucopyranose·phenol spectroscopic results: **a)** REMPI spectrum of β-MeGlc·P complex and IDIRS spectra at several ionization wavelengths: the shapes of the spectra do not change with the ionization wavelength used, apart from the signal coming from fragmentation; **b)** IR/UV Hole Burning experiment; **b$_1$)** confirmation of the contribution from high mass channel, the positive peaks in the blue trace come from the complex and the negative peaks are due to fragmentation; **c)** the experimental spectra are reported in black, while the corresponding simulated ones of the more stable isomers computed at M06-2X/6-311++G(d,p) level are displayed in colors. Correction factors used of 0.953 and 0.939 for CH and OH, respectively; **d)** structures of β-MeGlc · P most stable conformers.

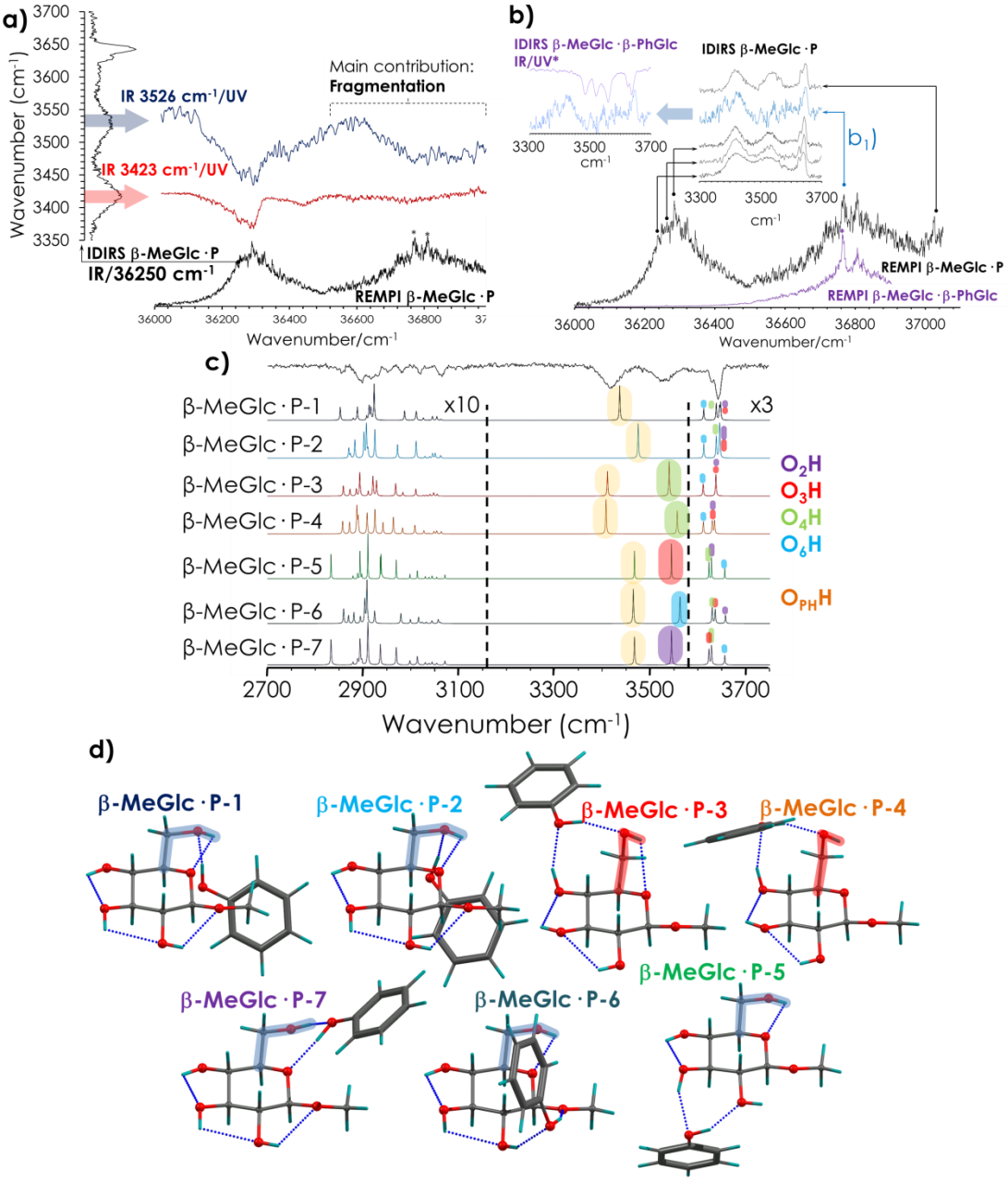
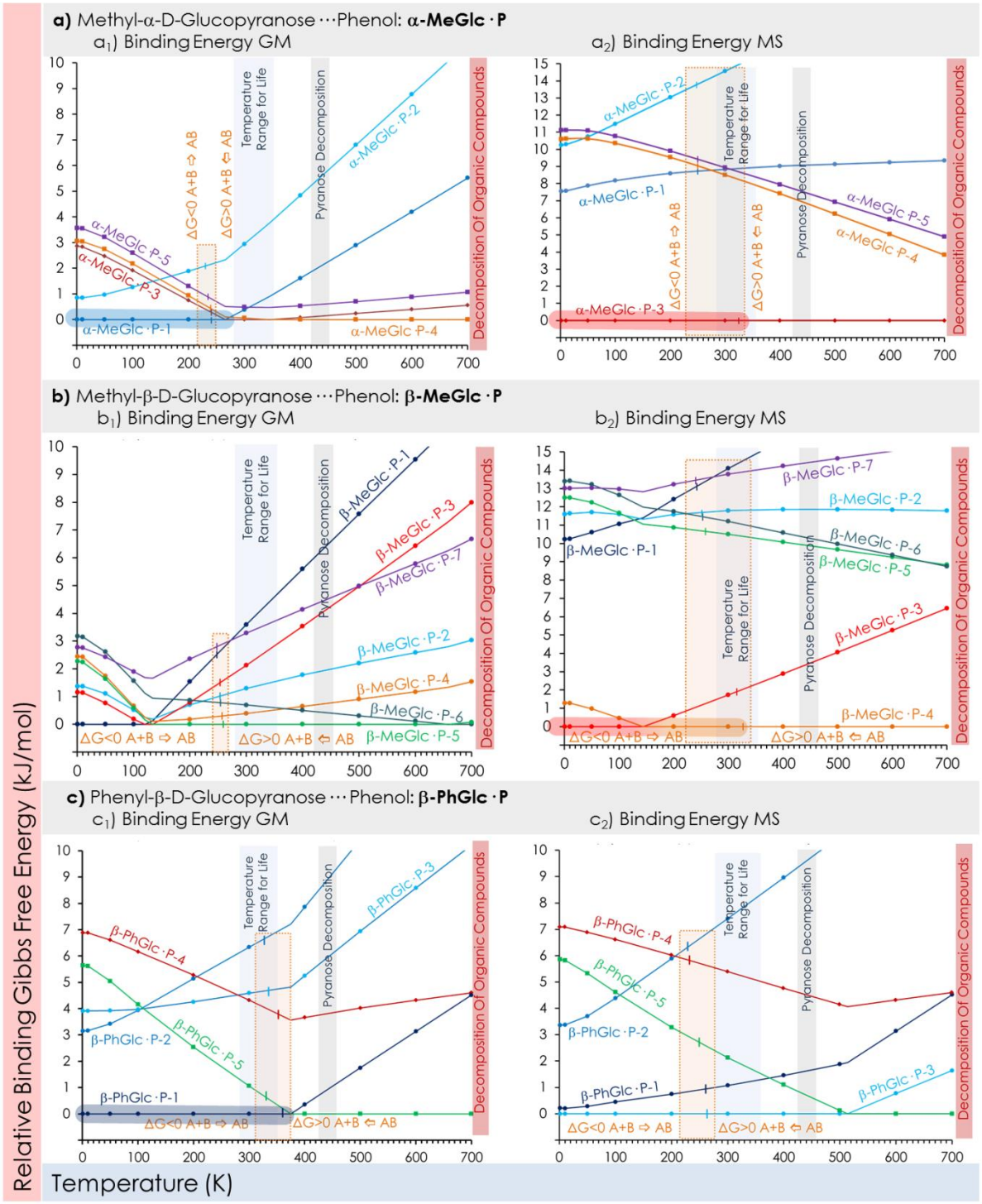


Figure A3.4: Relative energy data for: **a)** methyl-α-D-Glucopyranose · phenol; **b)** methyl-β-D-Glucopyranose · phenol;**c)** phenyl-β-D-Glucopyranose · phenol.



Appendix for Chapter 4

Equatorial β-D-glucopyranose: Structural Properties in One of the Most Abundant Carbohydrates

4.2. β-D-glucopyranose Monomers: Conformational Analysis

4.3. β-D-glucopyranose Dimerization Process in the Cold-Isolated (CI) Phase

4.4. Formation of Poly-1,4-β-D-glucopyranose (Cellulose I) Crystal

Scheme A4.1: Labeling of the sugars atoms. Oxygen atoms are in red and carbon atoms in blue. The anomeric oxygen (**O1**) is shaded in green and the hydroxymethyl substituent (**C5-C6-O6H**) in grey or in different colors, depending on the conformation adopted.

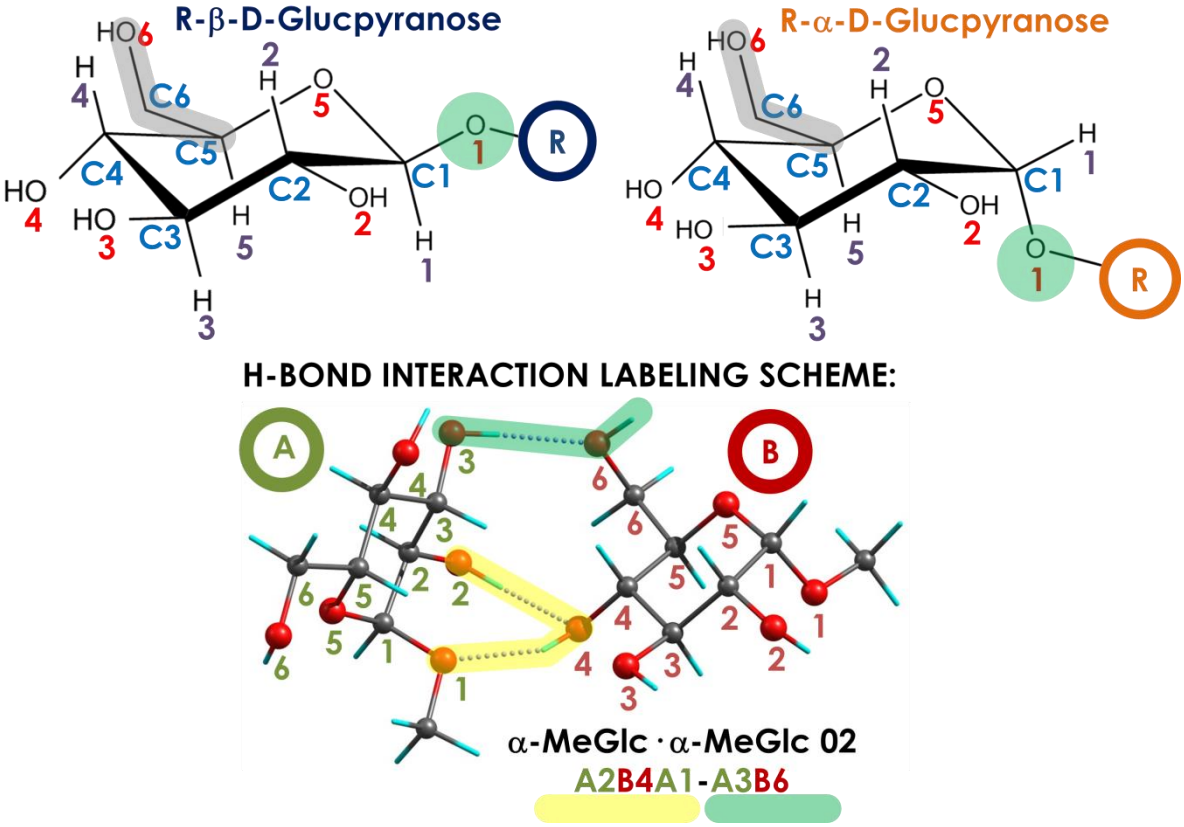


Figure A4.1: REMPI spectroscopy and corresponding IDIRS spectra for phenyl-β-D-glucopyranose (β-PhGlc): three conformers were observed.

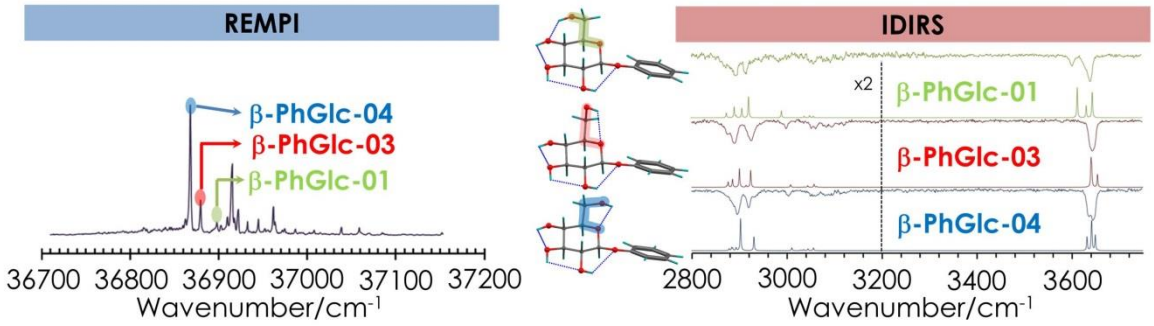


Table A4.1: β-D-glucopyranose (β-Glc): **a)** Table referred to the conformers in Figure 4.2.1-a), reporting their barrier for the interconversion and the scanned dihedral angles. **b)** Table referred to Figure 4.2.1-b) with the concerted interconversion pathways and the geometrical combination of the dominant dihedral angles.

a) Calculation Method & Level: M06-2X/6-311++G(d,p)

Initial Geometry	Final Geometry	Forward Barrier (kJ/mol)	Back Barrier (kJ/mol)	Geometrical properties		
				Dihedral Angle	Initial	Final
β-Glc-01	β-Glc-02	**15.3**	7.5	C_4-C_5-C_6-O_6	-75	39
β-Glc-02	β-Glc-03	9.4	**18.4**	C_5-C_6-O_6-H	-76	55
β-Glc-03	β-Glc-04	**25.4**	**24.1**	O_5-C_5-C_6-O_6	-57	59
				O_5-C_5-C_6-O_6(162)*	165	59
β-Glc-01	β-Glc-04	**27.1**	**28.4**	C_5-C_6-O_6-H	50	-54
β-Glc-05	β-Glc-06	**28.9**	**22.2**	O_5-C_5-C_6-O_6	-64	58
β-Glc-07	β-Glc-08	**29.1**	**21.9**	O_5-C_5-C_6-O_6	-64	59
β-Glc-09	β-Glc-10	**28.9**	**22.7**	O_5-C_5-C_6-O_6	-63	59
				H-C_3-O_3-H (-25)*	66	-64
β-Glc-03	β-Glc-05	**18.9**	4.5	H-C_4-O_4-H	17	-36
β-Glc-05	β-Glc-07	5.0	7.6	H-C_2-O_2-H	-57	70
β-Glc-07	β-Glc-09	5.4	7.1	H-C_1-O_1-H	54	-78
				H-C_3-O_3-H (-28)*	65	-61
β-Glc-04	β-Glc-06	**22.7**	2.9	H-C_4-O_4-H	22	-43
β-Glc-06	β-Glc-08	5.4	7.5	H-C_2-O_2-H	-58	70
β-Glc-08	β-Glc-10	4.4	7.2	H-C_1-O_1-H	50	-77

* Calculation of the minimum energy pathway for the interconversion with two involved coordinates: when the first dihedral angle is fixed (labeled with an asterisk), the second one is scanned.

b) Calculation Method & Level: M06-2X/6-311++G(d,p)

Concerted Path β-Glc-04 to β-Glc-09

Dominant Segment	Dihedral Angle	1	2	3	4	5
1 (1 - 2)	**H-C_4-O_4-H**	**42°**	**-55°**	-68°	-69	-70
2 (2 - 3)	**H-C_3-O_3-H**	-63°	**-4°**	**43°**	62	65
3 (3 - 4)	**H-C_2-O_2-H**	69°	58°	**22°**	**-27**	-60
4 (4 - 5)	**H-C_1-O_1-H**	-77°	-75°	-67°	**-41**	**55**
Barrier from β-Glc-04 to β-Glc-09:		**Forward: 20.5 kJ/mol Back: 5.3 kJ/mol**				

Concerted Path β-Glc-03 to β-Glc-10

Dominant Segment	Dihedral Angle	1	2	3	4	5
1 (1 - 2)	**H-C_4-O_4-H**	**36°**	**-55°**	-65°	-66°	-67°
2 (2 - 3)	**H-C_3-O_3-H**	-65°	**9°**	**44°**	63°	65°
3 (3 - 4)	**H-C_2-O_2-H**	70°	55°	**23°**	**-31°**	-59°
4 (4 - 5)	**H-C_1-O_1-H**	-78°	-77°	-69°	**-36°**	**59°**
Barrier from β-Glc-03 to β-Glc-10:		**Forward: 20.5 kJ/mol Back: 10.3 kJ/mol**				

Figure A4.2: Computational evaluation of the energy barriers related to rotation of the hydroxyl groups in the pyranose ring (β-Glc). The simulations were determined with two different theoretical methods: M06-2X and MP2 with the 6-311++G(d,p) basis set, and correspond to the scan of the C1-C2-O2-H dihedral angle, while O5-C1-O1-H is fixed at different angles. Thus, the curves in the two graphics represent cuts into the bidimensional PES of those two dyhedrals.

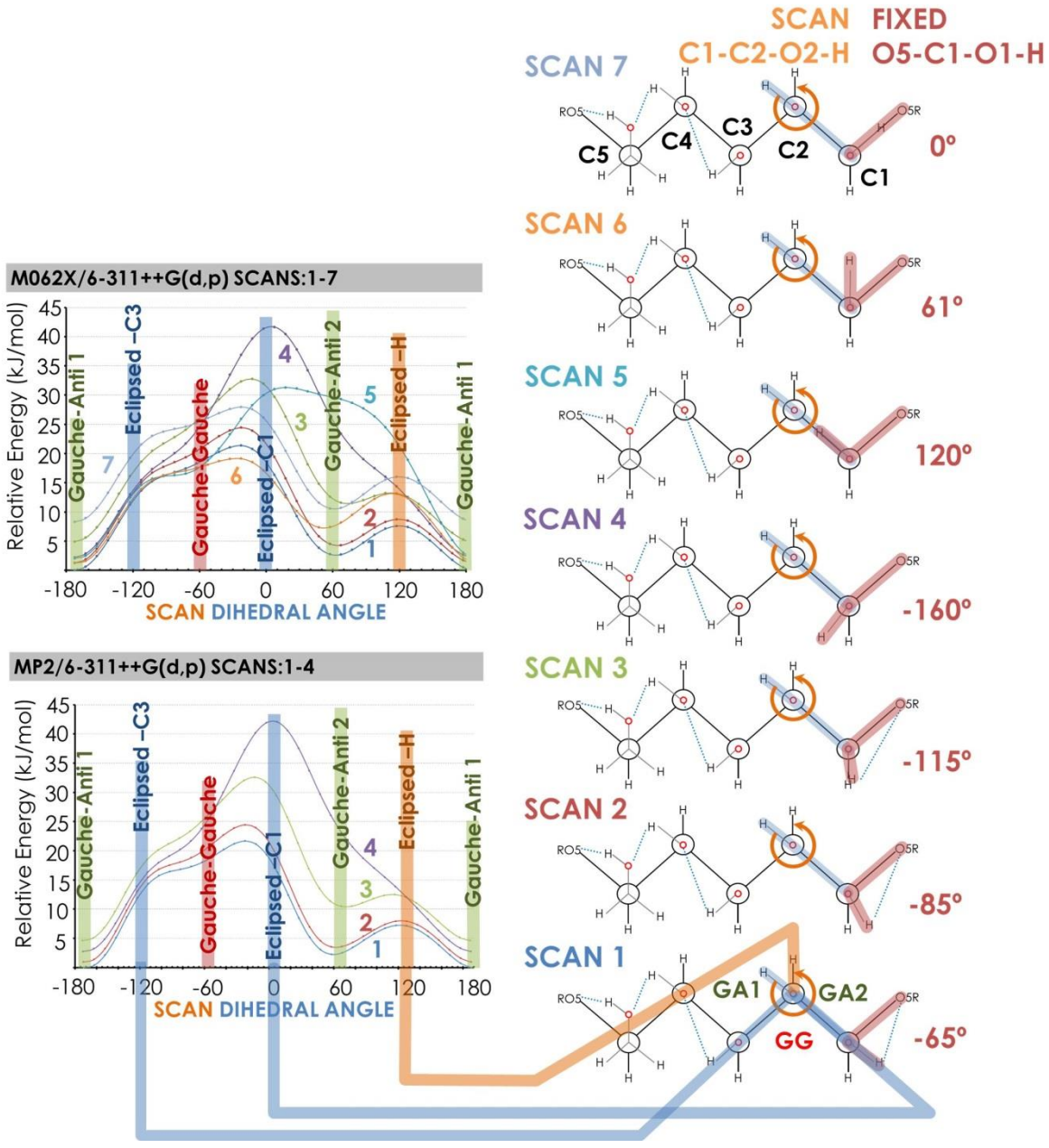
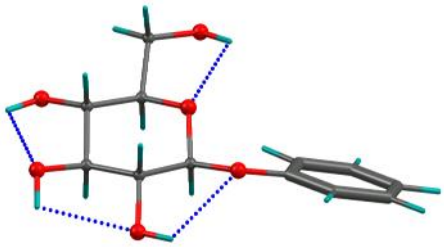
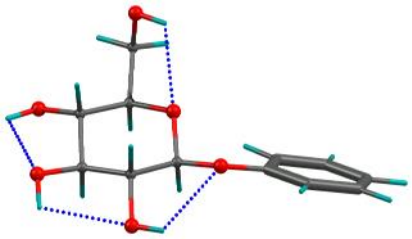


Table A4.2: Conformational isomers of phenyl-β-D-glucopyranose · methyl-β-D-glucopyranose (β-PhGlc · β-MeGlc) found in a 20 kJ/mol energy window, and optimized at the M06-2X/6-311++G(d,p) level, together with their relative energy (RE). RE ZPE: ZPE-corrected relative energy; RE GIBBS: relative ΔG at 298 K; β-PhGlc-n/β-MeGlc-n corresponds to the isomer of the monomer present in the complex (see the structures at the bottom of the table); H BOND describes the hydrogen bond network of the complex, using the nomenclature reported in the Scheme A4.1.

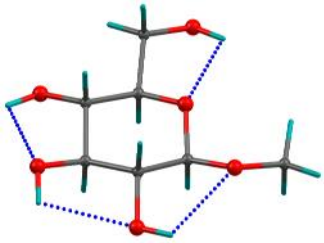
LABEL	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)	β-PhGlc (P)	β-MeGlc (M)	H BOND
01	0.00	0.00	β-PhGlc-2	β-MeGlc-2	P2M6P1-M2P6M1
02	10.27	15.49	β-PhGlc-1	β-MeGlc-2	P2M6P1-M2P6M1
03	12.46	12.51	β-PhGlc-2	β-MeGlc-2	P4M6P3-M2P6
04	12.96	15.12	β-PhGlc-2	β-MeGlc-1	P2M6P1-M2P6M1
05	19.01	18.81	β-PhGlc-1	β-MeGlc-1	P2M6P1
06	20.49	19.98	β-PhGlc-1	β-MeGlc-1	P6M6P5
07	21.02	18.00	β-PhGlc-2	β-MeGlc-1	P6M5-M6P1
08	21.94	21.24	β-PhGlc-1	β-MeGlc-3	M2P2M1
09	22.67	26.04	β-PhGlc-1	β-MeGlc-1	M6P6M5
10	22.74	22.70	β-PhGlc-1	β-MeGlc-2	M2P6-M6P2
11	23.47	20.98	β-PhGlc-1	β-MeGlc-1	M3P6M2



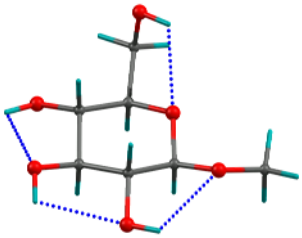
β-PhGlc-1



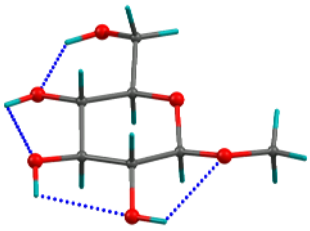
β-PhGlc-2



β-MeGlc-1



β-MeGlc-2



β-MeGlc-3

Figure A4.3: phenyl-β-D-glucopyranose · methyl-β-D-glucopyranose (β-PhGlc · β-MeGlc): **a)** IDIRS spectra at different ionization wavelengths. **b)** Comparison between the experimental spectrum (black line) and the simulated ones (colored lines) for the more stable isomers, predicted at the M06-2X/6-311++G(d,p) calculation level (correction factor for CH: 0.953 and for OH: 0.939). **c)** Gibbs free energy diagrams. **d)** Structures of the most stable conformers.

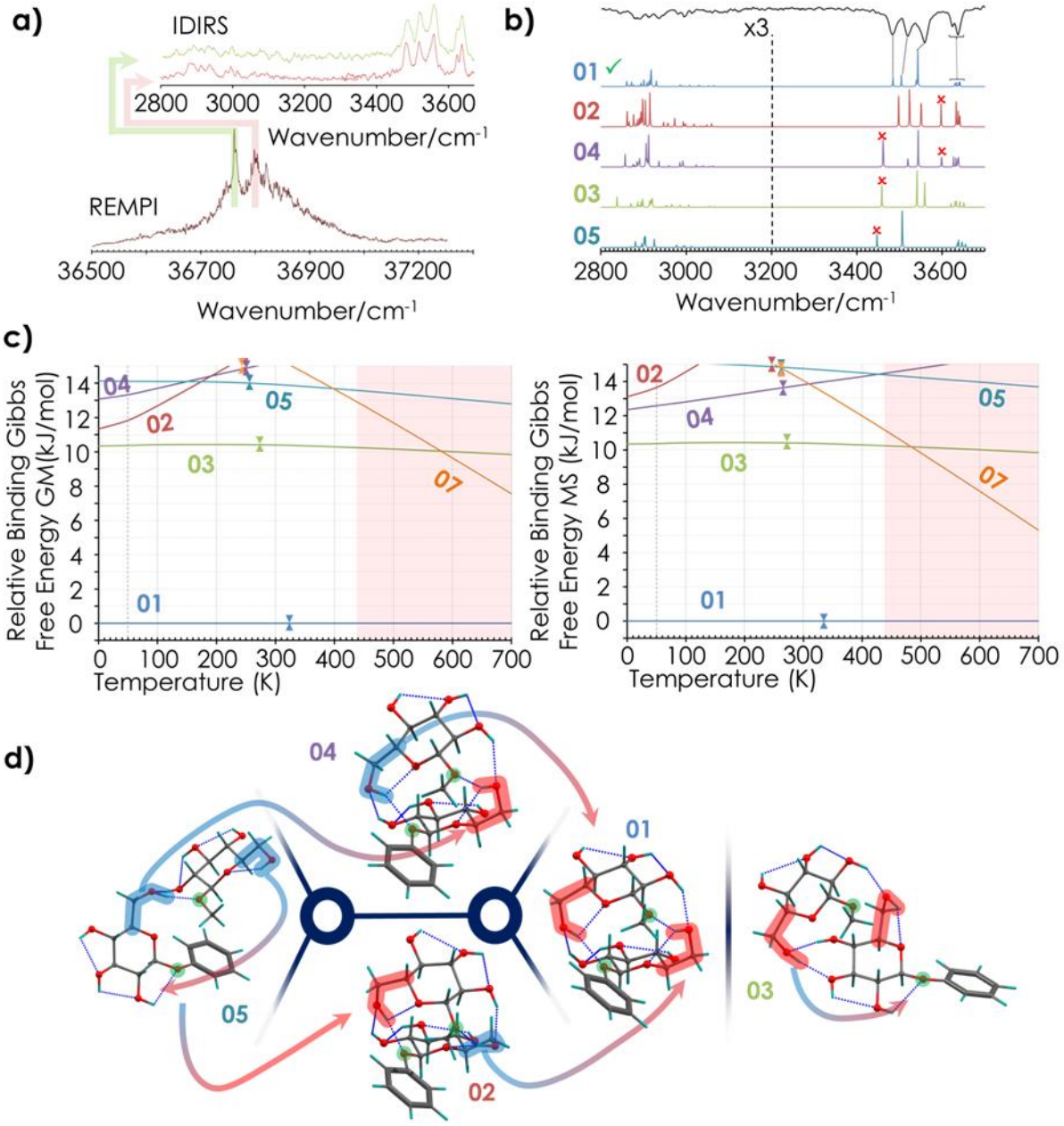
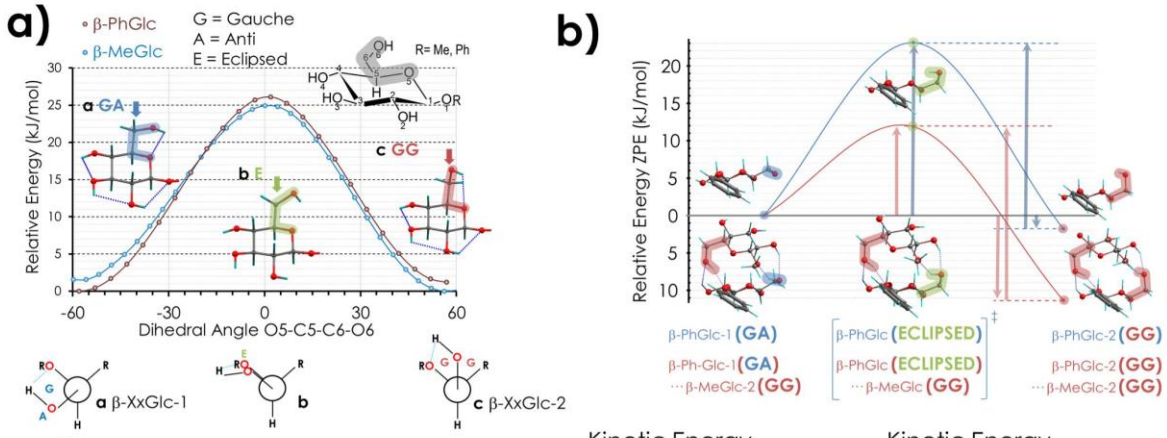


Figure A4.4: **a)** Energy barriers calculated at the M06-2X/6-311++G(d,p)level for the monomers phenyl-β-D-glucopyranose (β-PhGlc) and methyl-β-D-glucopyranose (β-MeGlc). **b)** Comparison of the energy barriers calculated at the M06-2X/6-311++G(d,p)level between the monomer phenyl-β-D-glucopyranose (β-PhGlc) and the dimer phenyl-β-D-glucopyranose · methyl-β-D-glucopyranose (β-PhGlc · β-MeGlc). **c)** Kinetic energy released by the collision of several carrier gases at different Mach speeds. **d)** Multidimensional diagram of the Gibbs free energy (M06-2X/6-311++G(d,p)) at different temperatures, with the relative possible effects on the dimerization process and the relative interconversion barriers.



c)

Mach Number	Carrier Gas	Kinetic Energy (kJ/mol)	Kinetic Energy (cm^{-1})
1	Ar	2.35	196.44
2	Ar	9.40*	785.75
1	He	0.24	19.67
2	He	0.94	78.68

Calculated speed of the jet used in the experiments: 1-2 Mach (Speed=Distance/($T_{Ablation\ Laser}$-$T_{Ionization\ Laser}$))

*These collisions presents enough energy to cross the barrier (ref. S1 and S2) from the GA conformation to the GG one in the cluster; but in the monomer, where both conformers were observed, this doesn't happen.

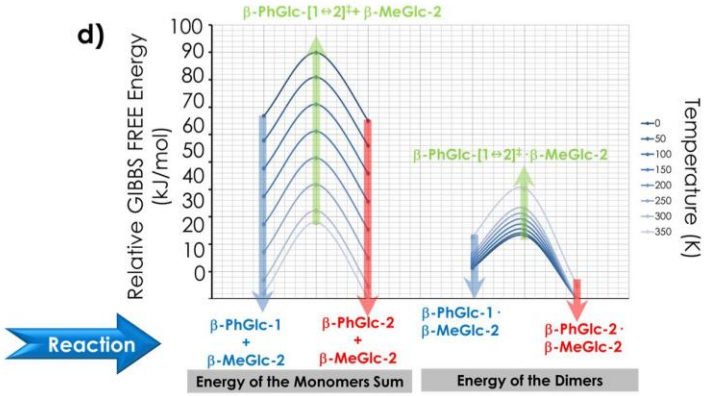
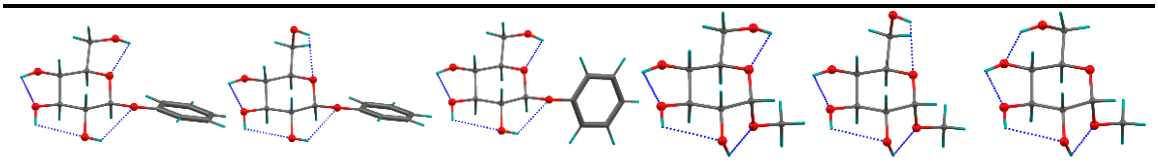


Table A4.3: Conformational isomers of phenyl-β-D-glucopyranose · methyl-α-D-glucopyranose (β-PhGlc · α-MeGlc) found in a 15 kJ/mol energy window, computed at the M06-2X/6-311++G(d,p) level, together with their relative energy (RE). RE ZPE: ZPE-corrected relative energy; RE GIBBS: relative ΔG at 298 K; β-PhGlc-n/α-MeGlc-n corresponds to the isomer of the monomer present in the complex (see the structures at the bottom of the table); H BOND describes the hydrogen bond network of the complex, using the nomenclature reported in the Scheme A4.1.

LABEL	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)	β-PhGlc (P)	α-MeGlc (M)	H BOND
01	**0.00**	**0.00**	**β-PhGlc-1**	**α-MeGlc-2**	**M4P2M3**
02	**1.79**	**7.50**	**β-PhGlc-1**	**α-MeGlc-1**	**M2P6M1**
03	**4.77**	**4.92**	**β-PhGlc-1**	**α-MeGlc-1**	**P6M6P5**
04	**4.68**	**3.43**	**β-PhGlc-1**	**α-MeGlc-1**	**P6M2P5**
05	**6.18**	**7.49**	**β-PhGlc-2**	**α-MeGlc-1**	**M6P6M5**
06	7.03	9.59	β-PhGlc-1	α-MeGlc-1	M6P1-P6M5
07	**3.86**	**7.33**	**β-PhGlc-1**	**α-MeGlc-1**	**P2M6P1**
08	**5.08**	**3.54**	**β-PhGlc-2**	**α-MeGlc-1**	**M3P6-M2P4M1**
09	7.32	10.30	β-PhGlc-1	α-MeGlc-2	M6P6M5
10	8.37	8.14	β-PhGlc-1	α-MeGlc-1	P6M3P5
11	8.57	9.80	β-PhGlc-1	α-MeGlc-1	M4P2M3
12	8.60	7.99	β-PhGlc-1	α-MeGlc-1	M3P5-P6M2
13	9.89	11.64	β-PhGlc-1	α-MeGlc-1	M6P6M5
14	8.84	7.83	β-PhGlc-1	α-MeGlc-1	M3P2M2
15	**7.22**	**7.38**	**β-PhGlc-1**	**α-MeGlc-1**	**M3P3M2P2**
16	10.84	8.29	β-PhGlc-2	α-MeGlc-1	M6P6M5
17	11.43	12.25	β-PhGlc-1	α-MeGlc-3	M4P6M3
18	**12.27**	**9.77**	**β-PhGlc-4**	**α-MeGlc-1**	**P2M6**
19	13.04	14.27	β-PhGlc-2	α-MeGlc-2	P2M6P1-M2P6
20	15.12	16.79	β-PhGlc-1	α-MeGlc-1	P2M4
21	15.70	15.61	β-PhGlc-4	α-MeGlc-1	M2P6M3P5



β-PhGlc-1 β-PhGlc-2 β-PhGlc-4 α-MeGlc-1 α-MeGlc-2 α-MeGlc-3

Figure A4.5: phenyl-β-D-glucopyranose · Methyl-α-D-glucopyranose (β-PhGlc · α-MeGlc): **a)** IDIRS spectra at different ionization wavelengths. **b)** Comparison between the experimental spectrum (black line) and the simulated ones (colored lines) for the more stable isomers, predicted at the M06-2X/6-311++G(d,p) calculation level (correction factor for CH: 0.953 and for OH: 0.9385). **c)** Gibbs free energy diagrams. **d)** Structures of the most stable conformers.

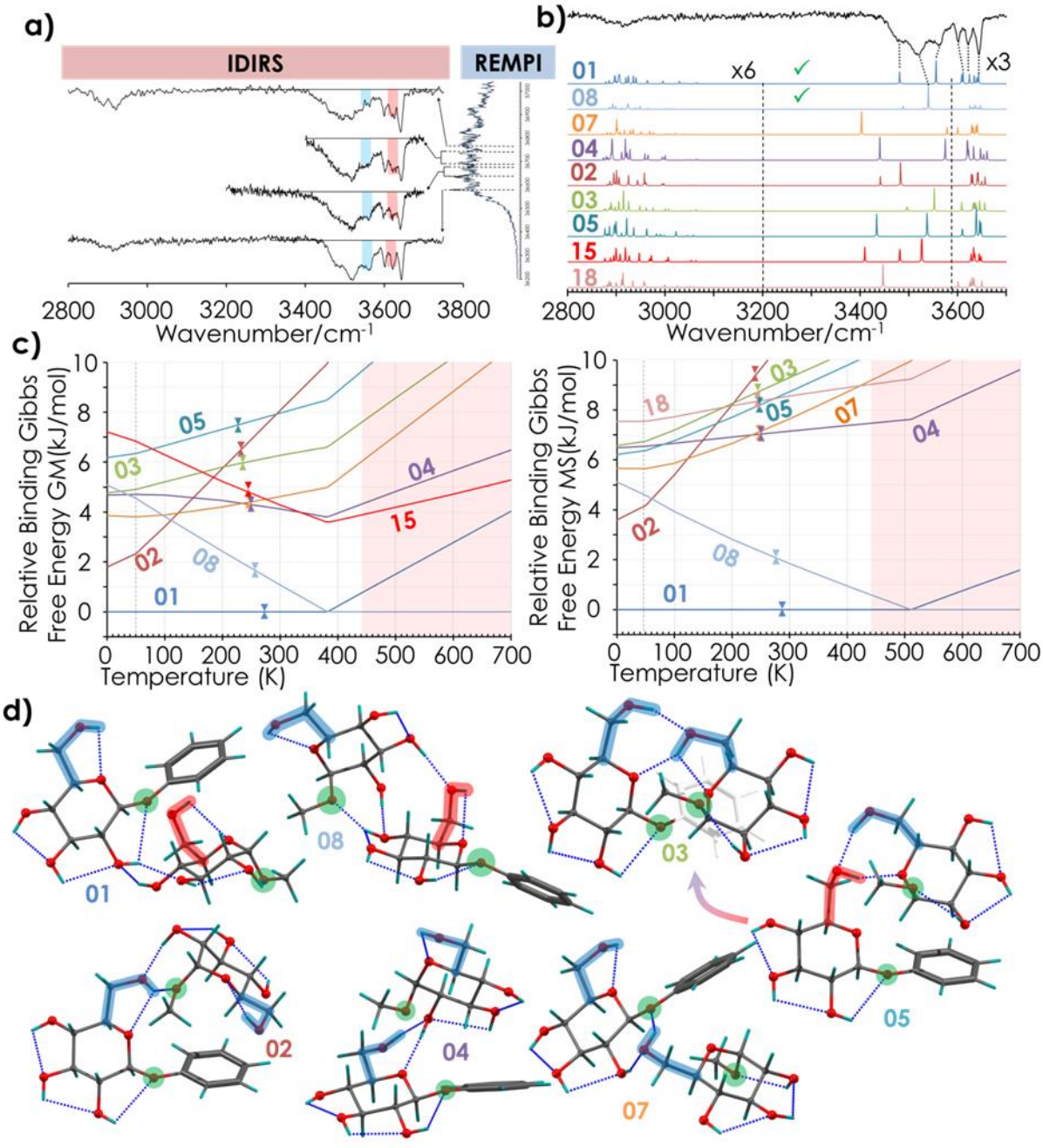
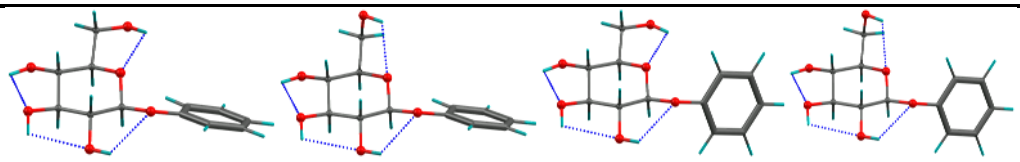
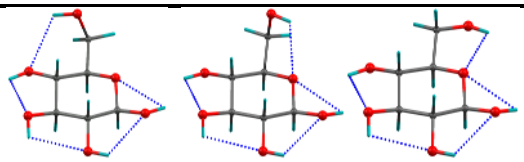


Table A4.4: Conformational isomers of phenyl-β-D-glucopyranose· β-D-glucopyranose (β-PhGlc· β-Glc) found in a 35 kJ/mol energy window, computed at the M06-2X/6-311++G(d,p) level, together with their relative energy (RE). RE ZPE: ZPE-corrected relative energy; RE GIBBS: relative ΔG at 298 K; β-PhGlc-n/β-Glc-n corresponds to the isomer of the monomer present in the complex (see the structures at the bottom of the table); H BOND describes the hydrogen bond network of the complex, using the nomenclature reported in the Scheme A4.1.

LABEL	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)	β-PhGlc (P)	β-Glc (B)	H BOND
01	**0.00**	**0.00**	**β-PhGlc-2**	**β-Glc-3**	**B2P6B1P5-P2B6P1**
02	**0.79**	**0.38**	**β-PhGlc-5**	**β-Glc-3**	**B2P6B1P5-P2B6P1**
03	**12.10**	**10.43**	**β-PhGlc-2**	**β-Glc-4**	**P6B1P5-P2B6P1**
04	13.24	10.46	β-PhGlc-1	β-Glc-3	B2P6B1-P2B6P1
05	**14.36**	**6.26**	**β-PhGlc-1**	**β-Glc-4**	**B6P6-B1P4**
06	**15.29**	**10.56**	**β-PhGlc-1**	**β-Glc-2**	**B1P4-P3B6P2**
07	16.90	15.45	β-PhGlc-2	β-Glc-4	B2P6B1P5-P2B6P1
08	18.21	10.98	β-PhGlc-2	β-Glc-3	B4P2B3
09	**20.75**	**6.50**	**β-PhGlc-1**	**β-Glc-4**	**B6P4B5-P1B3**
10	22.14	16.85	β-PhGlc-5	β-Glc-4	P6B1P5-P2B6
11	23.41	15.49	β-PhGlc-1	β-Glc-4	B6P4B5-P1B3
12	24.64	14.95	β-PhGlc-4	β-Glc-4	B1P2B5-B6P1
13	24.77	19.59	β-PhGlc-1	β-Glc-4	B2P2B1
14	26.14	21.03	β-PhGlc-2	β-Glc-4	B3P6B2P5
15	26.39	21.11	β-PhGlc-2	β-Glc-4	B3P6B2
16	26.45	15.89	β-PhGlc-2	β-Glc-4	P4B6P3-B1P6
17	26.88	19.22	β-PhGlc-1	β-Glc-4	B3P5-P6B2
18	34.77	30.93	β-PhGlc-1	β-Glc-4	B2P6B3
19	36.69	33.91	β-PhGlc-1-4	β-Glc-4	B2P6B3P5



β-PhGlc-1 β-PhGlc-2 β-PhGlc-4 β-PhGlc-5



β-Glc-2 β-Glc-3 β-Glc-4

Figure A4.6: phenyl-β-D-glucopyranose···β-D-glucopyranose (β-PhGlc···β-Glc): **a)** IR-Hole Burning spectra at different IR-Burning wavelengths; **b)** comparison between the experimental spectrum, upper black trace, and simulated spectra of the more stable isomers of β-PhGlc···β-Glc, the rest of the traces in different colors. Calculations carried out at M06-2X/6-311++G(d,p) level. Correction factors of 0.953 and 0.939 were applied respectively, to CH and OH normal modes to account for the anharmonicity; **c)** Gibbs free energy diagram; **d)** structure of the most stable β-PhGlc···β-Glc conformers.

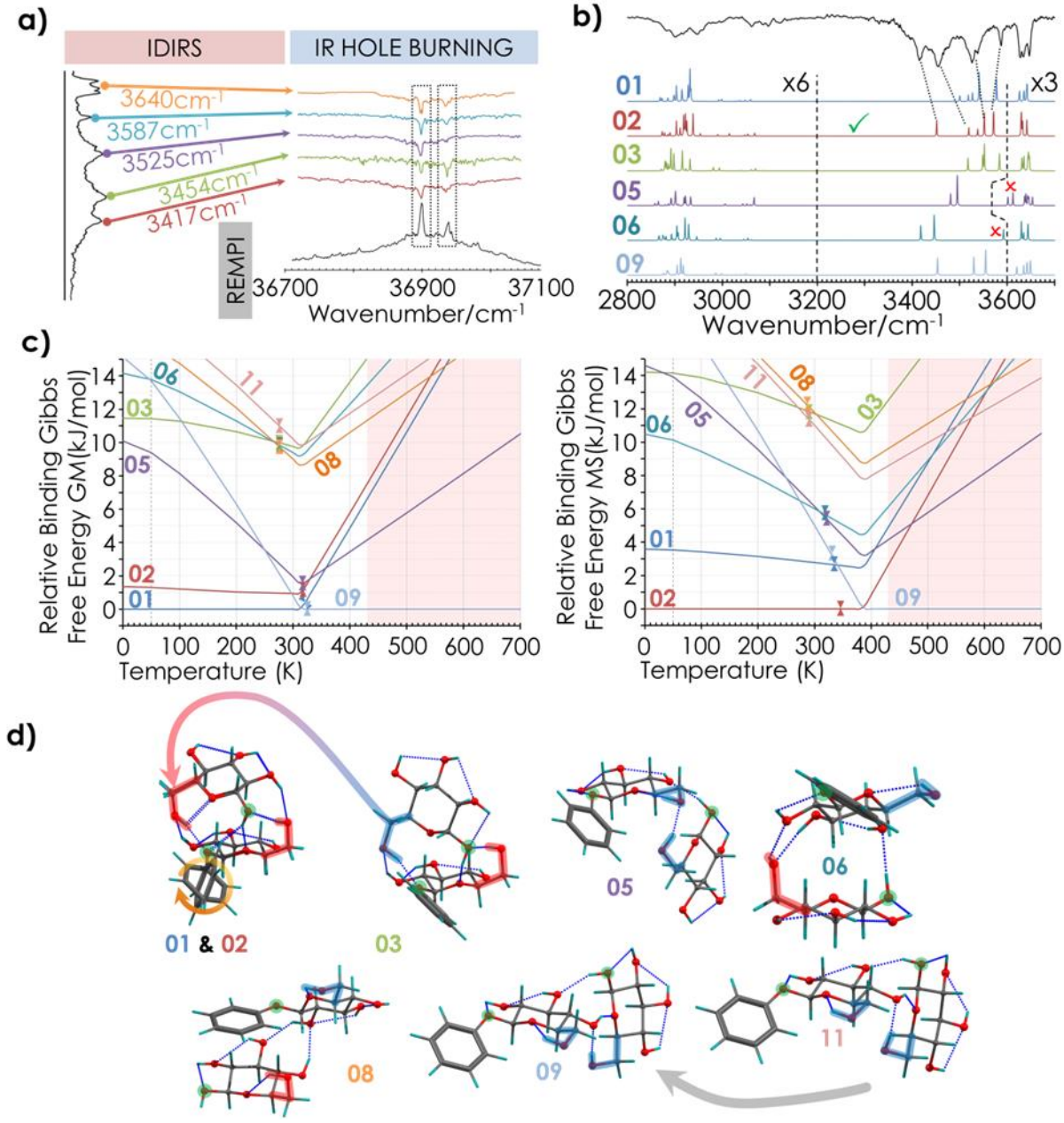
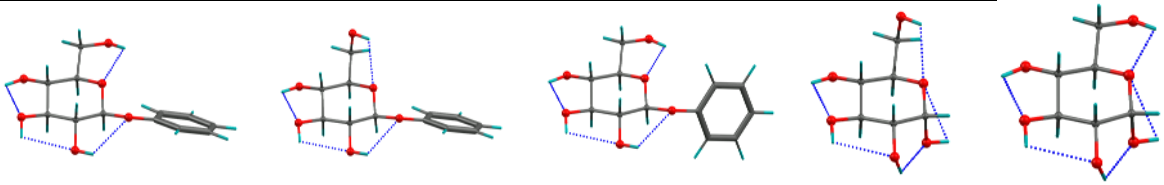


Table A4.5: Conformational isomers of phenyl-β-D-glucopyranose · α-D-glucopyranose (β-PhGlc · α-Glc) found in a 40 kJ/mol energy window, computed at the M06-2X/6-311++G(d,p) level, together with their relative energy (RE). RE ZPE: ZPE-corrected relative energy; RE GIBBS: relative ΔG at 298 K; β-PhGlc-n/α-Glc-n corresponds to the isomer of the monomer present in the complex (see the structures at the bottom of the table); H BOND describes the hydrogen bond network of the complex, using the nomenclature reported in the Scheme A4.1.

LABEL	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)	β-PhGlc (P)	α-Glc (A)	H BOND
01	0.00	0.00	β-PhGlc-1	α-Glc-4	A6P6A5-A1P5
02	7.49	6.72	β-PhGlc-1	α-Glc-4	A6P4-A1P6
03	8.49	8.35	β-PhGlc-2	α-Glc-4	A3P6A2
04	8.58	6.46	β-PhGlc-2	α-Glc-4	P6A6-P4A1P3
05	8.81	11.85	β-PhGlc-1	α-Glc-3	A6P6A5
06	12.07	9.50	β-PhGlc-2	α-Glc-4	P6A2P5
07	12.33	14.11	β-PhGlc-1	α-Glc-4	A6P6A5-A1P5
08	18.60	14.86	β-PhGlc-1	α-Glc-4	P6A2
09	18.69	18.21	β-PhGlc-1	α-Glc-4	P6A2
10	18.77	18.08	β-PhGlc-1	α-Glc-3	A6P5
11	20.01	20.39	β-PhGlc-2	α-Glc-4	A2P6A1
12	20.35	21.30	β-PhGlc-1	Boat Structure	A6P6A5-A1P5
13	20.36	21.31	β-PhGlc-1	Boat Structure	A6P6A5-A1P5
14	20.56	20.54	β-PhGlc-1	α-Glc-4	P6A6P5
15	21.56	21.43	β-PhGlc-2	α-Glc-3	A2P6A1
16	22.17	24.64	β-PhGlc-2	α-Glc-4	A2P5-P6A1
17	26.56	23.67	β-PhGlc-4	α-Glc-4	P6A2
18	35.63	36.91	β-PhGlc-1	Boat Structure	A6P6A5
19	36.35	36.72	β-PhGlc-4	Boat Structure	A1P6A6
20	38.20	40.87	β-PhGlc-2	Boat Structure	A6P5-P6A4
21	38.34	38.93	β-PhGlc-2	Boat Structure	P6A5-A1P5



β-PhGlc-1 β-PhGlc-2 β-PhGlc-4 α-Glc-3 α-Glc-4

Figure A4.7: phenyl-β-D-glucopyranose···α-D-glucopyranose (β-PhGlc···α-Glc): **a)** IR-Hole Burning spectra at different IR-Burning wavelengths; **b)** comparison between the experimental spectrum, upper black trace, and simulated spectra of the more stable isomers of β-PhGlc···α-Glc, the rest of the traces in different colors. Calculations carried out at M06-2X/6-311++G(d,p) level. Correction factors of 0.953 and 0.939 were applied respectively to CH and OH normal modes to account for the anharmonicity; **c)** Gibbs free energy diagram; **d)** structure of the most stable β-PhGlc···α-Glc conformers.

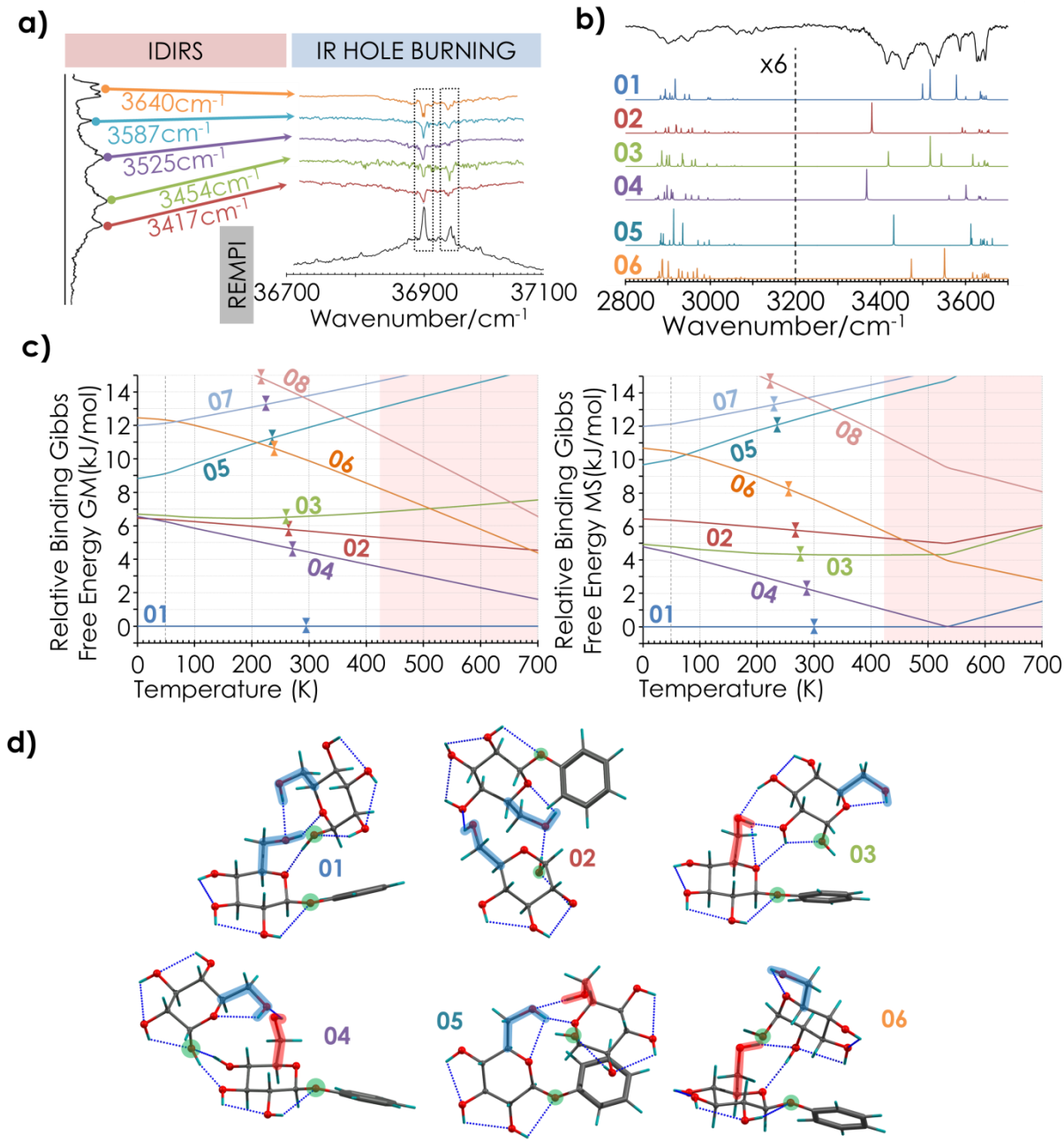


Table A4.6: Conformational isomers of phenyl-α-D-glucopyranose· methyl-α-D-glucopyranose (α-PhGlc· α-MeGlc) in which there is a direct interaction with the chromophore, found in a 25 kJ/mol energy window, computed at the M06-2X/6-311++G(d,p) level, together with their relative energy (RE). RE ZPE: ZPE-corrected relative energy; RE GIBBS: relative ΔG at 298 K; α-PhGlc-n/α-MeGlc-n corresponds to the isomer of the monomer present in the complex (see the structures at the bottom of the table); H BOND describes the hydrogen bond network of the complex, using the nomenclature reported in the Scheme A4.1. Ph Interaction: interaction with the chromophore.

LABEL	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)	α-PhGlc (P)	Dihedral Angle C1-O1-Cph1-Cph2	α-MeGlc (M)	H BOND	Ph Interaction
01	0.0	0.0	α-PhGlc-3	-13.678	α-MeGlc-03	M2P6	O1Me · Ph
02	2.8	0.7	α-PhGlc-6	-12.676	α-MeGlc-01	M2P6	O1Me · Ph
03	3.5	5.7	α-PhGlc-1	-141.919	α-MeGlc-01	M3P6M2P5	H5 · Ph
04	3.7	1.8	α-PhGlc-6	-21.297	α-MeGlc-01	P6M6	H3+O2H · Ph
05	4.0	2.4	α-PhGlc-6	-23.722	α-MeGlc-01	P6M6	H3+O2H · Ph
06	4.4	2.9	α-PhGlc-2	-11.861	α-MeGlc-01	M4P2M3	O2H · Ph
07	5.1	3.0	α-PhGlc-6	-23.754	α-MeGlc-01	P6M4	O2H · Ph
08	5.4	8.1	α-PhGlc-6	-66.588	α-MeGlc-03	P6M6	H3 · Ph
09	5.4	2.7	α-PhGlc-1	-5.257	α-MeGlc-01	M4P2M3	O2H · Ph
10	5.5	7.5	α-PhGlc-1	-7.539	α-MeGlc-03	M3P4-M2P6	O1Me+H5 · Ph
11	5.8	8.6	α-PhGlc-1	-73.649	α-MeGlc-01	P2M6P1	H1 · Ph
12	5.9	1.5	α-PhGlc-3	-10.677	α-MeGlc-01	M4P2M3	O2H · Ph
13	6.0	8.2	α-PhGlc-1	-141.286	α-MeGlc-03	M3P6M2P5	H5 · Ph
14	6.2	7.2	α-PhGlc-6	-49.998	α-MeGlc-01	M6P4-P6M4	O2H+H3 · Ph
15	6.3	4.9	α-PhGlc-1	-6.199	α-MeGlc-02	M4P2M3	O2H · Ph
16	6.3	6.9	α-PhGlc-6	-21.892	α-MeGlc-03	P6M6	O4H · Ph
17	6.4	2.9	α-PhGlc-1	-10.828	α-MeGlc-03	M4P2M3	O2H · Ph
19	6.7	4.0	α-PhGlc-1	-58.724	α-MeGlc-02	P2M6P5	H2 · Ph
20	6.8	6.8	α-PhGlc-1	-30.226	α-MeGlc-01	M6P1	H2 · Ph
21	6.9	8.3	α-PhGlc-1	-102.674	α-MeGlc-03	M2P6M1	O1Me · Ph
22	7.1	4.5	α-PhGlc-6	-46.294	α-MeGlc-02	P6M4	O2H · Ph
23	7.2	7.8	α-PhGlc-6	-19.67	α-MeGlc-01	P6M5-P6M6P4	H4 · Ph
24	7.7	7.0	α-PhGlc-1	-102.971	α-MeGlc-01	M2P6M1	O1Me · Ph
25	8.0	11.6	α-PhGlc-1	-53.61	α-MeGlc-01	P4M6-P6M4	H3 · Ph
26	8.4	7.5	α-PhGlc-3	-22.446	α-MeGlc-01	P4M6P6	H2 · Ph
27	8.5	6.7	α-PhGlc-5	-68.62	α-MeGlc-01	P4M6	H3 · Ph
28	8.7	7.8	α-PhGlc-4	-74.439	α-MeGlc-01	P4M6	H3 · Ph
29	8.8	10.3	α-PhGlc-4	-73.636	α-MeGlc-03	P4M6	H3 · Ph

LABEL	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)	α-PhGlc (P)	Dihedral Angle	α-MeGlc (M)	H BOND	Ph Interaction

				C1-O1-Cph1-Cph2			
30	9.0	8.6	α-PhGlc-1	-73.649	α-MeGlc-01	P2M6P1	H1 · Ph
34	10.1	9.3	α-PhGlc-1	-10.632	α-MeGlc-01	M6P4	H4 · Ph
37	11.1	10.5	α-PhGlc-1	4.15	α-MeGlc-01	P2M6	Weak H5+Me · Ph
38	11.1	11.7	α-PhGlc-1	18.197	α-MeGlc-01	M3P2M2	O1Me · Ph
42	11.4	9.3	α-PhGlc-1	-3.459	α-MeGlc-02	M4P2M3	H2 · Ph
44	12.0	11.8	α-PhGlc-1	-20.203	α-MeGlc-01	M6P4	H4 · Ph
45	12.3	11.4	α-PhGlc-1	-1.988	α-MeGlc-01	M2P4	O1Me · Ph
46	12.4	12.0	α-PhGlc-1	-31.622	α-MeGlc-01	M2P6	H5 · Ph
47	12.6	12.0	α-PhGlc-1	-11.942	α-MeGlc-01	M3P6M2	Lateral O2H
49	12.7	12.9	α-PhGlc-3	-0.93	α-MeGlc-01	P2M6	Ph Lateral O1Me · Ph
52	13.1	9.9	α-PhGlc-1	-31.406	α-MeGlc-02	M2P6	H5 · Ph
54	13.7	12.0	α-PhGlc-3	-67.483	α-MeGlc-01	P2M6P1	H1 · Ph
55	13.7	14.0	α-PhGlc-1	-117.933	α-MeGlc-01	M3P6M2	O2H · Ph
56	14.0	10.9	α-PhGlc-1	-7.59	α-MeGlc-02	M6P4	H2 · Ph
57	14.0	16.3	α-PhGlc-1	-12.638	α-MeGlc-01	P6M1	O1Me · Ph
59	14.1	12.4	α-PhGlc-1	-8.02	α-MeGlc-03	M2P6M1	Ph Lateral H5 · Ph
60	14.5	16.6	α-PhGlc-4	-5.286	α-MeGlc-03	M4P6	O1Me · Ph
61	14.8	13.0	α-PhGlc-1	0.814	α-MeGlc-01	M2P2M1	H5 · Ph
62	14.8	16.5	α-PhGlc-4	-28.834	α-MeGlc-01	P4M6	O5+H1 · Ph
63	14.9	13.9	α-PhGlc-2	-8.174	α-MeGlc-01	M4P2M3	H2 · Ph
64	15.0	13.9	α-PhGlc-2	-12.651	α-MeGlc-01	M2P4	O1Me+H5 · Ph
65	15.6	14.0	α-PhGlc-1	-5.9	α-MeGlc-01	M4P2M3	H2 · Ph
68	15.7	12.2	α-PhGlc-4	-15.452	α-MeGlc-08	P4M6	H2 · Ph
69	15.8	13.6	α-PhGlc-1	-4.082	α-MeGlc-01	P2M6	H5 · Ph
70	15.9	14.9	α-PhGlc-1	-8.197	α-MeGlc-01	M3P6M2	H2 · Ph
71	16.1	13.8	α-PhGlc-4	-24.168	α-MeGlc-01	P4M6	H2 · Ph
72	16.5	18.7	α-PhGlc-1	-19.649	α-MeGlc-01	M2P6M1	O1Me · Ph
73	16.6	15.2	α-PhGlc-1	-0.786	α-MeGlc-07	P2M6	H2 · Ph
74	16.6	15.7	α-PhGlc-3	-7.624	α-MeGlc-01	M4P2M3	H2 · Ph
75	16.7	15.9	α-PhGlc-2	-10.065	α-MeGlc-02	P2M6	H5 · Ph
76	16.9	12.3	α-PhGlc-1	3.014	α-MeGlc-03	P2M6	O1Me · Ph
77	17.9	13.9	α-PhGlc-1	-11.114	α-MeGlc-01	M3P2M2	H2 · Ph
78	18.1	15.5	α-PhGlc-1	-0.352	α-MeGlc-01	P2M6&5	H2 · Ph

LABEL	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)	α-PhGlc (P)	Dihedral Angle C1-O1-Cph1-Cph2	α-MeGlc (M)	H BOND	Ph Interaction
81	19.0	17.1	α-PhGlc-4	-25.636	α-MeGlc-01	P4M6	H2 · Ph
82	19.1	16.7	α-PhGlc-1	-18.879	α-MeGlc-01	P2M6	O6H · Ph
83	19.1	15.4	α-PhGlc-1	17.299	α-MeGlc-02	P2M6	H2 · Ph
84	19.6	17.9	α-PhGlc-1	-132.484	α-MeGlc-01	P2M6	H5 · Ph
85	20.0	16.0	α-PhGlc-1	-15.917	α-MeGlc-01	P2M4	H3 · Ph
86	20.5	19.4	α-PhGlc-1	-21.339	α-MeGlc-06	M6P6	H3 · Ph
87	20.7	21.6	α-PhGlc-1	-18.523	α-MeGlc-03	M2P6	H5 · Ph
88	21.2	19.3	α-PhGlc-1	-16.977	α-MeGlc-01	-	H5 · Ph
89	21.2	18.0	α-PhGlc-1	2.564	α-MeGlc-02	M2P6	H2 · Ph
90	21.2	19.7	α-PhGlc-1	-4.533	α-MeGlc-01	P2M4	H5 · Ph
91	21.3	21.6	α-PhGlc-1	-4.337	α-MeGlc-04	M4P6	H3 · Ph
92	21.7	16.3	α-PhGlc-1	-9.045	α-MeGlc-01	P2M6	O5 · Ph
93	21.9	16.9	α-PhGlc-1	-8.978	α-MeGlc-01	P2M6	O5+O6H · Ph
94	22.3	16.7	α-PhGlc-4	-19.208	α-MeGlc-01	P4M6	Ph Lateral H3 · Ph

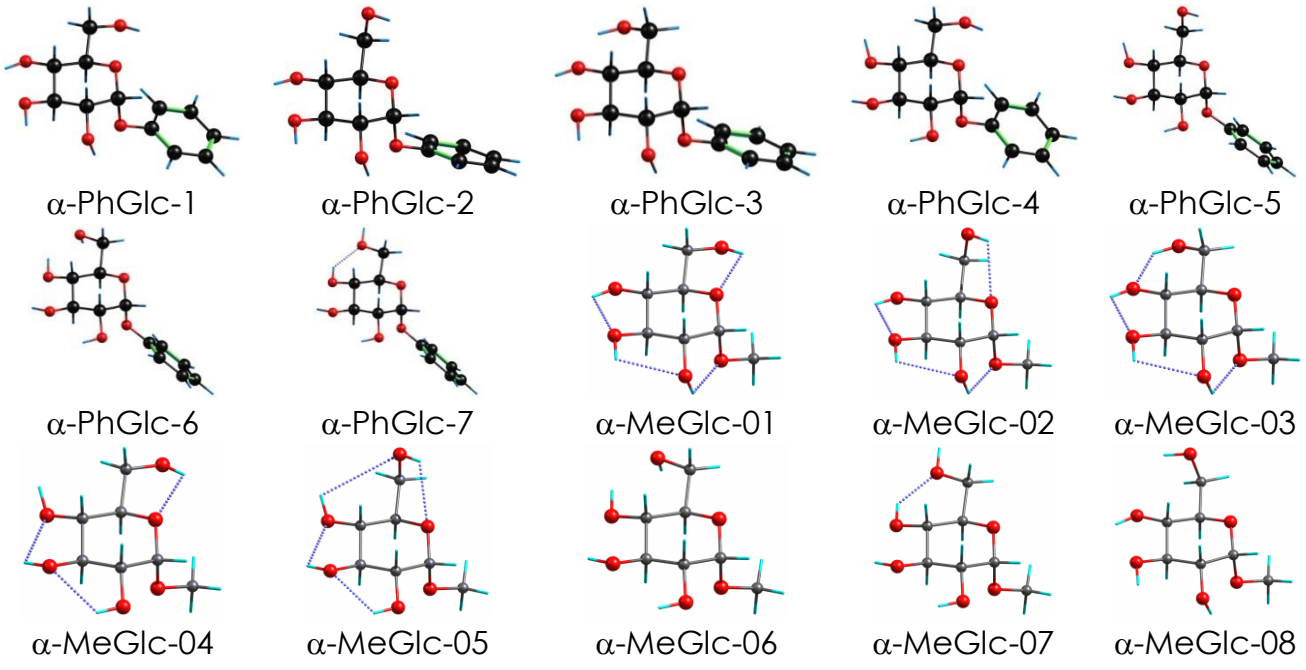


Table A4.7: Conformational isomers of phenyl-α-D-glucopyranose· methyl-α-D-glucopyranose (α-PhGlc· α-MeGlc) in which there is not a direct interaction with the chromophore, found in a 25 kJ/mol energy window, computed at the M06-2X/6-311++G(d,p) level, together with their relative energy (RE). RE ZPE: ZPE-corrected relative energy; RE GIBBS: relative ΔG at 298 K; α-PhGlc-n/α-MeGlc-n corresponds to the isomer of the monomer present in the complex (see the structures at the bottom of the table); H BOND describes the hydrogen bond network of the complex, using the nomenclature reported in the Scheme A4.1.

LABEL	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)	α-PhGlc (P)	Dihedral Angle C1-O1-Cph1-Cph2	α-MeGlc (M)	H-BOND
18	6.4	4.2	α-PhGlc-1	11.254	α-MeGlc-02	P3M6-P2M4
31	9.4	5.5	α-PhGlc-2	-14.689	α-MeGlc-01	P3M3P2M2
32	10.0	5.9	α-PhGlc-1	-10.38	α-MeGlc-01	P3M3P2M2
33	10.1	6.2	α-PhGlc-3	-15.445	α-MeGlc-01	P3M3P2M2
35	10.9	6.7	α-PhGlc-2	-26.933	α-MeGlc-01	M2P6M1
36	11.1	9.9	α-PhGlc-1	-12.107	α-MeGlc-02	P3M3P2M2
39	11.2	10.1	α-PhGlc-1	-62.189	α-MeGlc-03	P2M6
40	11.3	7.8	α-PhGlc-1	-9.962	α-MeGlc-03	P3M3P2M2
41	11.3	4.8	α-PhGlc-1	-2.845	α-MeGlc-01	M3P3M2P2
43	12.0	12.7	α-PhGlc-2	-0.577	α-MeGlc-01	P2M6
48	12.7	9.5	α-PhGlc-1	-89.803	α-MeGlc-01	P3M3P2M2
50	12.8	11.7	α-PhGlc-1	-9.83	α-MeGlc-01	M2P6M1
51	12.9	8.3	α-PhGlc-1	-10.199	α-MeGlc-01	P3M4P2M3
53	13.2	10.7	α-PhGlc-3	-19.578	α-MeGlc-01	P3M4-P2M6
58	14.1	12.0	α-PhGlc-7	-26.167	α-MeGlc-01	M3P2M2P3
66	15.6	12.3	α-PhGlc-1	-19.086	α-MeGlc-03	M2P6M1
67	15.6	13.2	α-PhGlc-1	-5.143	α-MeGlc-01	M3P4M2P3
79	18.2	16.2	α-PhGlc-4	-11.92	α-MeGlc-03	P4M6
96	23.9	22.9	α-PhGlc-2	-17.195	α-MeGlc-02	M2P6-P6M2-M4P3-P4M3

Figure A4.8: a) Relative Binding Gibbs Free Energy of all the calculated conformational isomers of phenyl-α-D-glucopyranose· methyl-α-D-glucopyranose (α-PhGlc· α-MeGlc) found in a ~25 kJ/mol energy window. **b)** Relative Binding Gibbs Free Energy for the most stable conformers of α-PhGlc· α-MeGlc without the interaction with the chromophore. **c)** Structure of the most stable conformers.

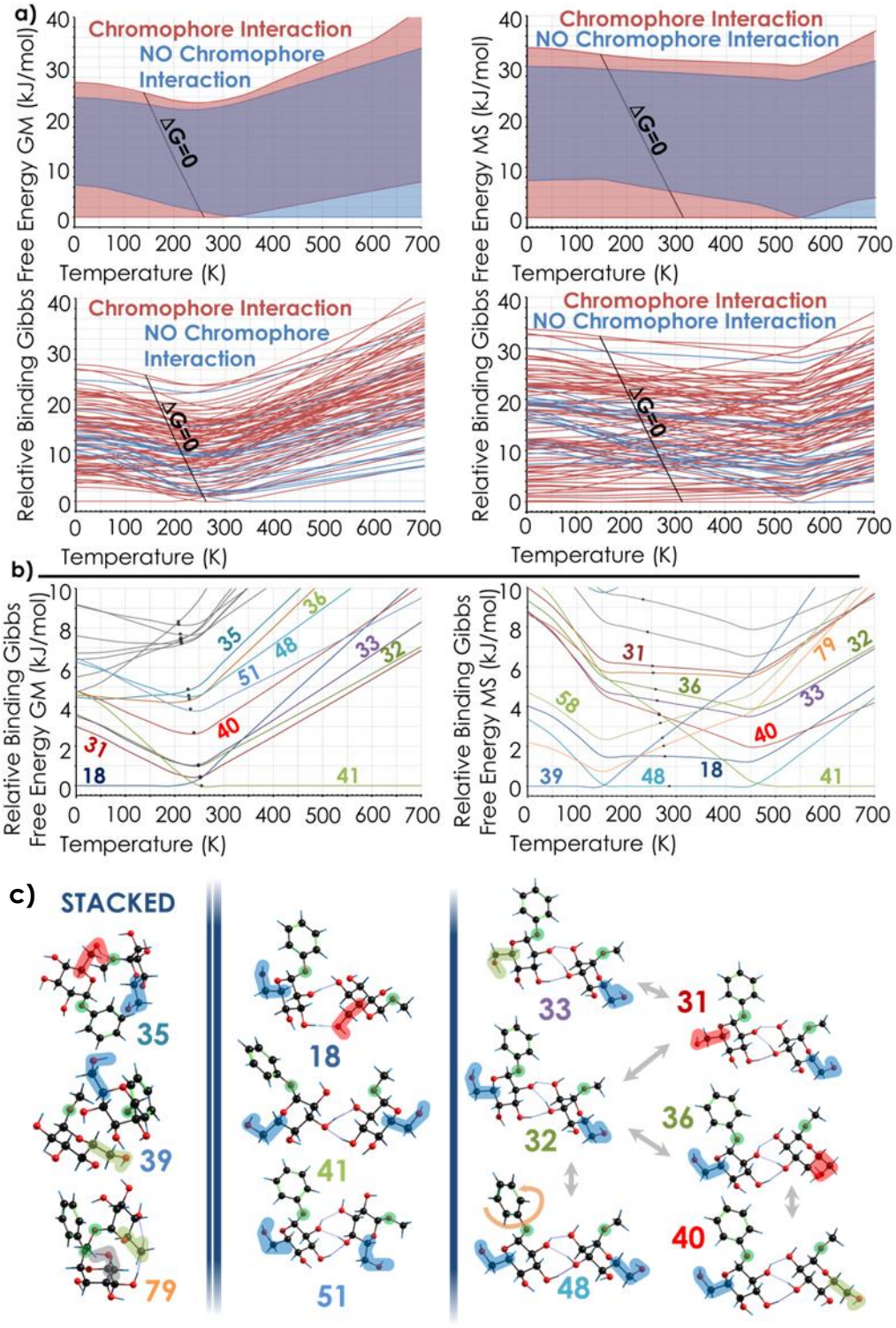
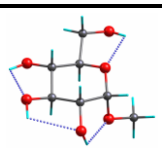
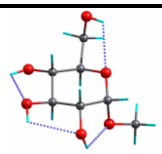


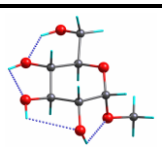
Table A4.8: Conformational isomers of α-MeGlc· α-MeGlc found in a 20 kJ/mol energy window, computed at the M06-2X/6-311++G(d,p) level, together with their relative energy (RE). RE ZPE: ZPE-corrected relative energy; RE GIBBS: relative ΔG at 298 K; α-PhGlc-n/α-MeGlc-n corresponds to the isomer of the monomer present in the complex (see the structures at the bottom of the table); H BOND describes the hydrogen bond network of the complex, using the nomenclature reported in the Scheme A4.1.

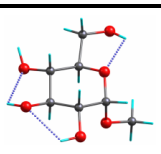
LABEL	E ZPE (kJ/mol)	E GIBBS (kJ/mol)	α-MeGlc-A (A)	α-MeGlc-B (B)	H BOND
01	**0.0**	**0.0**	**α-MeGlc-02**	**α-MeGlc-02**	**A4B6A3-B3A6**
02	**5.6**	**2.1**	**α-MeGlc-01**	**α-MeGlc-02**	**A2B4A1-A3B6**
03	**6.9**	**7.4**	**α-MeGlc-02**	**α-MeGlc-03**	**B2A4B1-B3A6**
04	8.2	9.2	α-MeGlc-01	α-MeGlc-01	A3B6A2B5
05	**9.7**	**8.6**	**α-MeGlc-01**	**α-MeGlc-01**	**A3B3A2B2**
06	10.3	8.8	α-MeGlc-01	α-MeGlc-02	B3A3B2A2
07	10.4	9.0	α-MeGlc-01	α-MeGlc-01	A3B2A2
08	**10.5**	**8.1**	**α-MeGlc-01**	**α-MeGlc-03**	**B3A3B2**
09	10.5	8.4	α-MeGlc-01	α-MeGlc-03	A3B3A2(B2)
10	**10.6**	**11.7**	**α-MeGlc-03**	**α-MeGlc-08**	**A2B6A1-B3A6**
11	10.9	12.9	α-MeGlc-01	α-MeGlc-03	B3A6B2
12	11.0	8.2	α-MeGlc-01	α-MeGlc-03	B3A3B2A2
13	11.0	10.7	α-MeGlc-01	α-MeGlc-02	A3B3A2B2
14	11.3	8.7	α-MeGlc-01	α-MeGlc-03	B3A2B2
15	11.3	11.6	α-MeGlc-01	α-MeGlc-01	A2B4
16	11.7	11.1	α-MeGlc-01	α-MeGlc-02	A3B2A2
17	12.0	10.3	α-MeGlc-01	α-MeGlc-01	A3B4A2
18	12.1	13.9	α-MeGlc-01	α-MeGlc-01	A2B6A1
19	12.1	11.2	α-MeGlc-01	α-MeGlc-02	A2B6A1
20	12.2	12.2	α-MeGlc-01	α-MeGlc-02	B3A2B2A1
21	12.4	14.3	α-MeGlc-01	α-MeGlc-03	A2B6
22	12.4	11.5	α-MeGlc-01	α-MeGlc-02	B6A4
23	12.5	14.7	α-MeGlc-01	α-MeGlc-03	A2B6
24	12.7	13.4	α-MeGlc-01	α-MeGlc-01	A3B5-A2B6
25	12.9	10.2	α-MeGlc-01	α-MeGlc-01	A3B2-B3A2
26	12.9	12.3	α-MeGlc-01	α-MeGlc-03	A3B2A2
27	13.0	13.1	α-MeGlc-03	α-MeGlc-03	A3B3A2B2
28	13.2	12.7	α-MeGlc-02	α-MeGlc-03	B3A6B1
29	**13.3**	**9.4**	**α-MeGlc-01**	**α-MeGlc-01**	**A3B2-B3A2**
30	13.4	14.1	α-MeGlc-01	α-MeGlc-01	A3B4-A2B6
31	13.7	10.6	α-MeGlc-01	α-MeGlc-01	-
32	13.7	14.0	α-MeGlc-01	α-MeGlc-03	A3B2A2

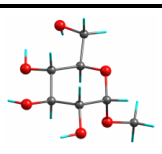
LABEL	E ZPE (kJ/mol)	E GIBBS (kJ/mol)	α-MeGlc-A (A)	α-MeGlc-B (B)	H BOND
33	**14.0**	**10.1**	**α-MeGlc-01**	**α-MeGlc-03**	**A3B2-B3A2**
34	14.1	16.1	α-MeGlc-01	α-MeGlc-03	B2A6B1
35	**14.2**	**12.3**	**α-MeGlc-02**	**α-MeGlc-08**	**A2B6A1**
36	14.3	15.0	α-MeGlc-01	α-MeGlc-02	B2A6B1
37	14.3	14.4	α-MeGlc-01	α-MeGlc-03	B3A4B2
38	14.3	14.7	α-MeGlc-01	α-MeGlc-01	A4B6A3
39	14.6	12.3	α-MeGlc-01	α-MeGlc-08	A2B6A1
40	14.9	14.8	α-MeGlc-03	α-MeGlc-03	A2B6
41	15.0	12.4	α-MeGlc-01	α-MeGlc-01	A3B3-B4A2
42	15.4	12.3	α-MeGlc-01	α-MeGlc-01	A4B3A3
43	15.7	13.8	α-MeGlc-01	α-MeGlc-01	A6B4
44	16.1	15.9	α-MeGlc-02	α-MeGlc-03	A6B4
45	16.2	18.2	α-MeGlc-01	α-MeGlc-01	A3B6A2
46	16.9	17.6	α-MeGlc-01	α-MeGlc-01	A6B6
47	17.2	14.8	α-MeGlc-02	α-MeGlc-03	A2B6
48	17.4	16.4	α-MeGlc-01	α-MeGlc-01	A4B3-B4A3
49	17.7	16.9	α-MeGlc-01	α-MeGlc-01	A4B2A3
50	17.8	20.0	α-MeGlc-01	α-MeGlc-03	A4B6
51	18.9	17.7	α-MeGlc-01	α-MeGlc-01	A6B4
52	19.1	19.3	α-MeGlc-01	α-MeGlc-01	A2B6
53	20.3	18.3	α-MeGlc-01	α-MeGlc-01	A2B4A1
54	20.4	20.9	α-MeGlc-02	α-MeGlc-03	A2B6
55	20.6	16.3	α-MeGlc-01	α-MeGlc-01	A4B2A3
56	20.8	21.0	α-MeGlc-01	α-MeGlc-01	A6B6
57	21.5	18.3	α-MeGlc-01	α-MeGlc-01	A4B2A3
58	**22.0**	**21.6**	**α-MeGlc-01**	**α-MeGlc-04**	**B4A6**
59	22.0	25.6	α-MeGlc-01	α-MeGlc-01	A6B4-B6A4
60	22.9	19.5	α-MeGlc-01	α-MeGlc-01	A2B4A1
61	23.4	19.4	α-MeGlc-01	α-MeGlc-08	A2B6
62	24.7	26.0	α-MeGlc-01	α-MeGlc-06	B6A6
63	25.9	23.2	α-MeGlc-01	α-MeGlc-01	A2B6


α-MeGlc-01


α-MeGlc-02


α-MeGlc-03


α-MeGlc-04


α-MeGlc-06

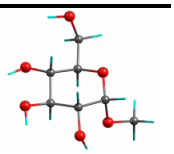

α-MeGlc-08

Figure A4.9: a) Relative Binding Gibbs Free Energy for the most stable conformers of methyl-α-D-glucopyranose· methyl-α-D-glucopyranose (α-MeGlc· α-MeGlc). **b)** Structure of the most stable conformers of α-MeGlc· α-MeGlc. **c)** Comparison between the most stable conformers of α-MeGlc· α-MeGlc and α-PhGlc· α-MeGlc without interaction with the chromophore.

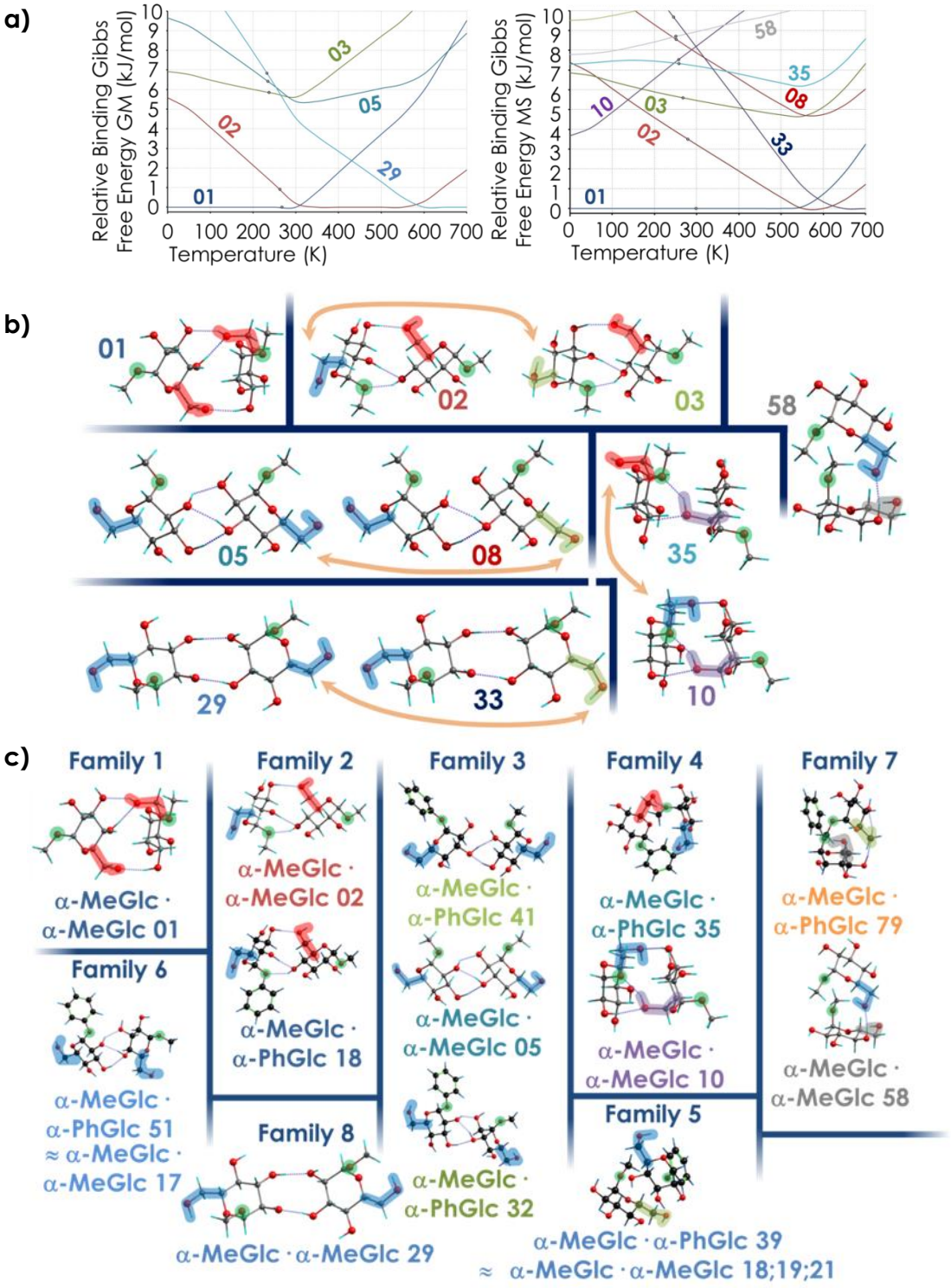


Figure A4.10: Crystallographic structures of some monosaccharides (glucopyranose), disaccharides (cellobiose), and polysaccharides (starch). Decomposition of the crystallographic structure and highlight of the main possible H-bond interactions.

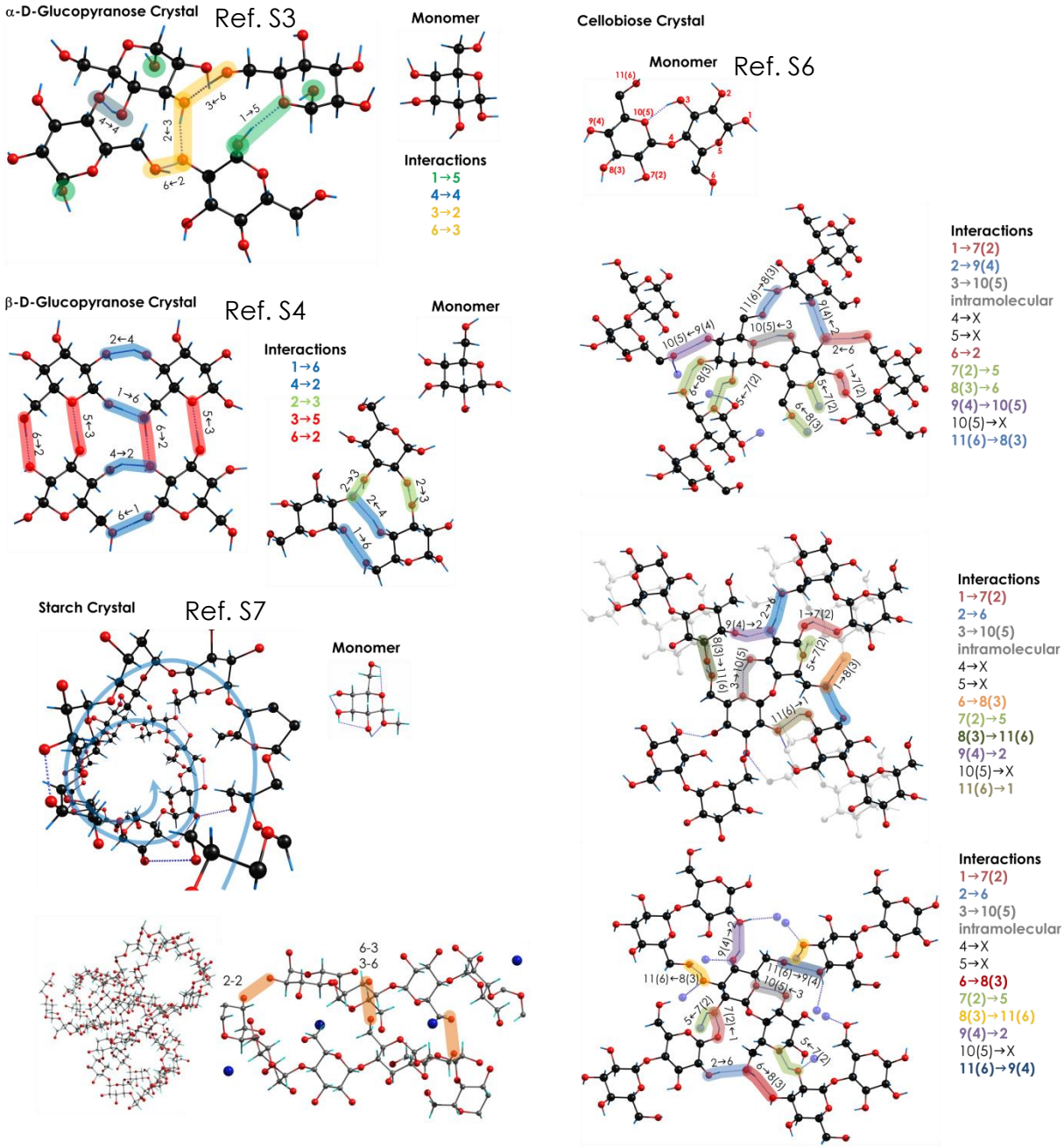


Figure A4.11: Crystallographic structures of the Cellulose I[S5] crystal, decomposition in layers and preliminary scheme for its H-bond arrangement.

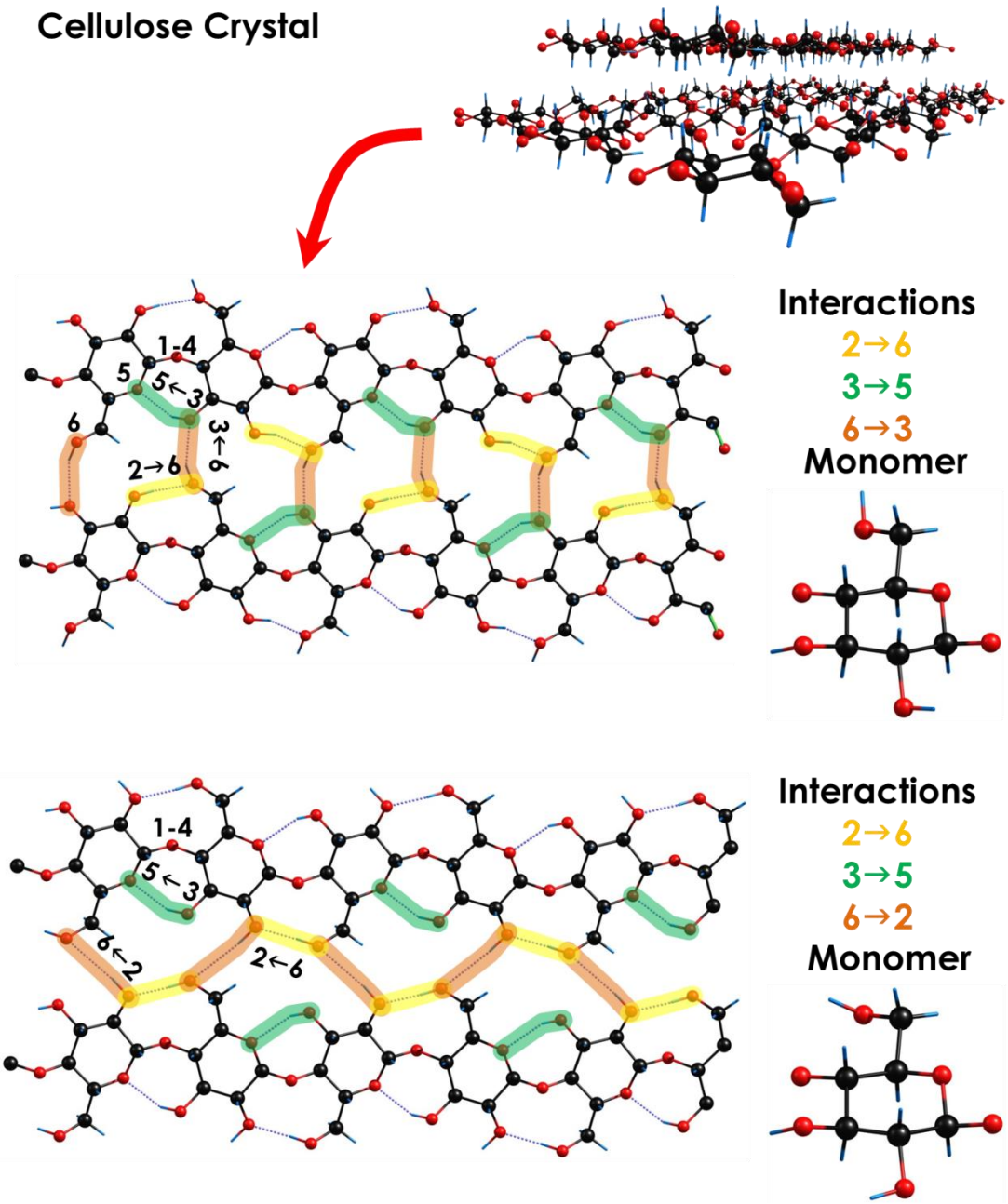


Figure A4.12: Optimization of the isolated Cellulose I chains with M06-2X/6-31+G(d). **1)** Show the a, b and c alternated glucoses chains arrangements. **2)** c, e and f show all the glucoses chains in the same direction. In particular: **a)** equivalent to the crystallographic chain with the hydrogens in one possible arrangement; **b)** equivalent to the crystallographic chain but with the hydrogens in another possible arrangement; **c)** optimized arrangement: intramolecular instability and incompatible with the crystallographic matrix because of the superposition of the hydrogens; **d)** all the glucoses are in a negative bending trend; **e)** alternated chain of negative and positive bending; **f)** all the glucoses are in a positive bending.

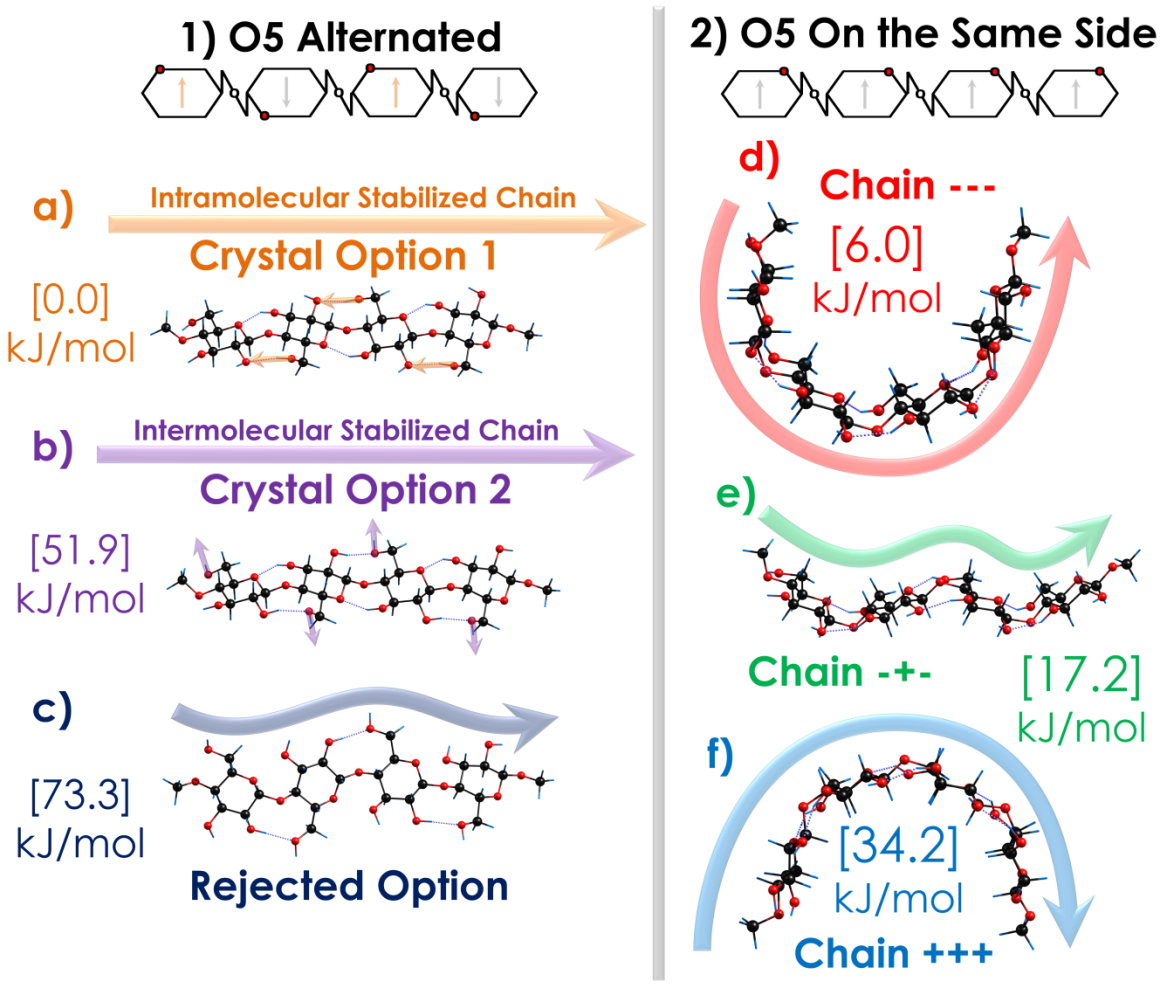
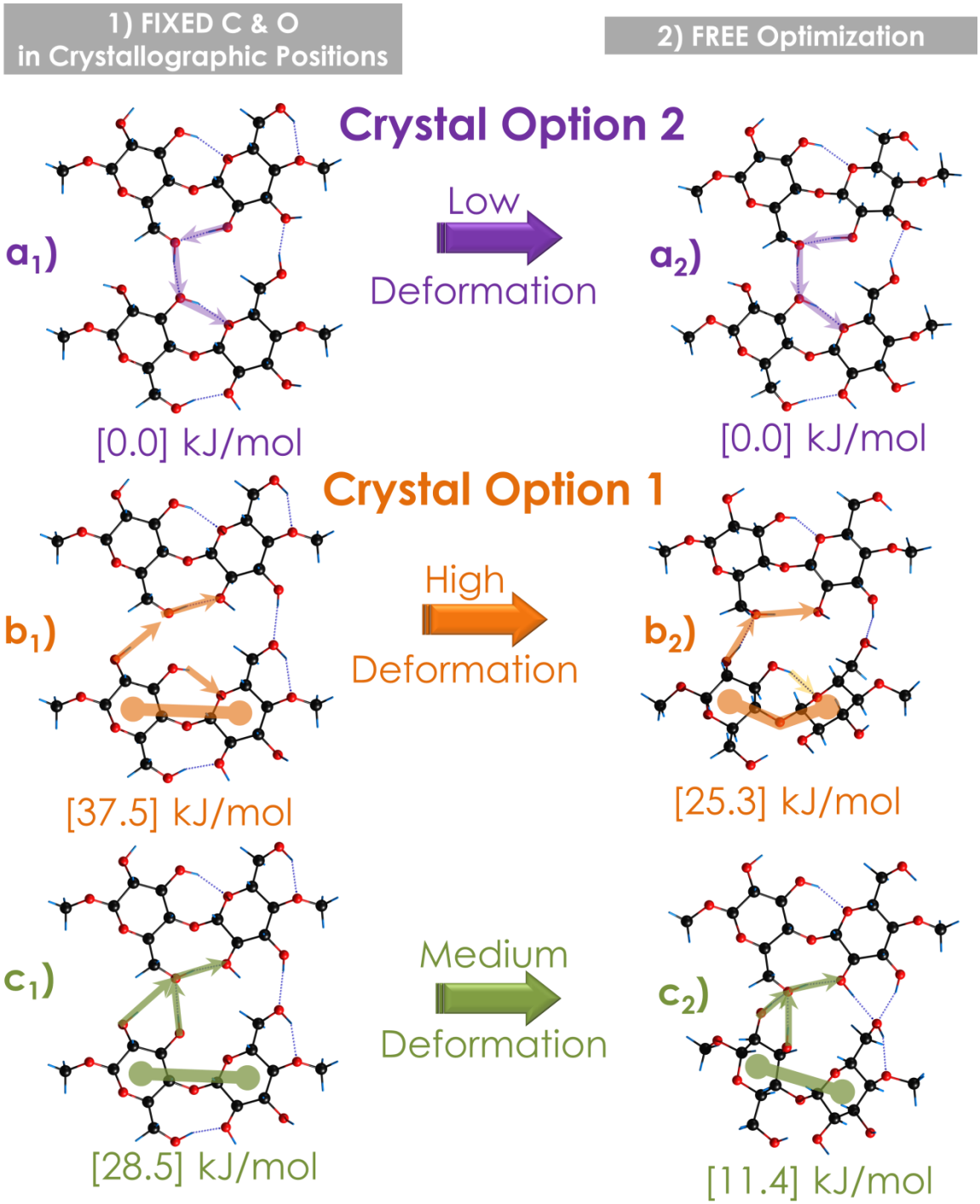


Figure A4.13: Optimization of isolated cellulose layers in the crystal using M06-2X/6-31+G(d). **1)** Relative stability of the tree possible arrangements for interacting hydrogen when the carbon and oxygen are fixed in the position of the crystallographic structure. **2)** Optimization of the structures reported in column 1, with all the atoms free.



REFERENCES

S1. R. S. Ruoff, T. D. Klots, H. S. Gutowsky, *J. Chem. Phys.* **1990,** *93*, 3142-3150.

S2. G. M. Florio, R. A. Christie, K. D. Jordan, T. S. Zwier. *J. Am. Chem. Soc.*, **2002**, *124*, 10236-10247.

S3. G.M.Brown, H.A.Levy; Science **1965**, 147, 1038.

S4. W.G.Ferrier; Acta Crystallogr. **1963**, 16, 1023.
S.S.C.Chu, G.A.Jeffrey; Acta Crystallogr.,Sect.B:Struct.Crystallogr.Cryst.Chem. **1968**, 24, 830,
M.L.C.E.Kouwijzer, B.P.van Eijck, H.Kooijman, J.Kroon; Acta Crystallogr.,Sect.B:Struct.Sci. **1995**, 51, 209.

S5. Y.Nishiyama, P.Langan, H.Chanzy; J.Am.Chem.Soc. **2002**, 124, 9074.

S6. R.A.Jacobson, J.A.Wunderlich, W.N.Lipscomb; *Acta Crystallogr.* **1961**, 14, 598
L.A.Aleshina, N.V.Melekh; Unknown ,**2012**, 6, 101.
F.Kimura, W.Oshima, H.Matsumoto, H.Uekusa, K.Aburaya, M.Maeyama, T.Kimura; CrystEngComm , **2014**, 16, 6630.

S7. A. Sarko, A. Biloski; Carbohydrate Research, **1980** 79,11-2.
Y. Nishiyama, K. Mazeau, M. Morin, M. B. Cardoso, H. Chanzy, and J. Putaux, Macromolecules, **2010**, 43, 8628.

Appendix for Chapter 5

Analysis of Folding Structures

Tripeptide:

Polimer:

Scheme A5.1: Labelling, numbering and sections of 1-methylated and 13-phenylated Threonine-Alanine-Alanine tripeptide (TAA).

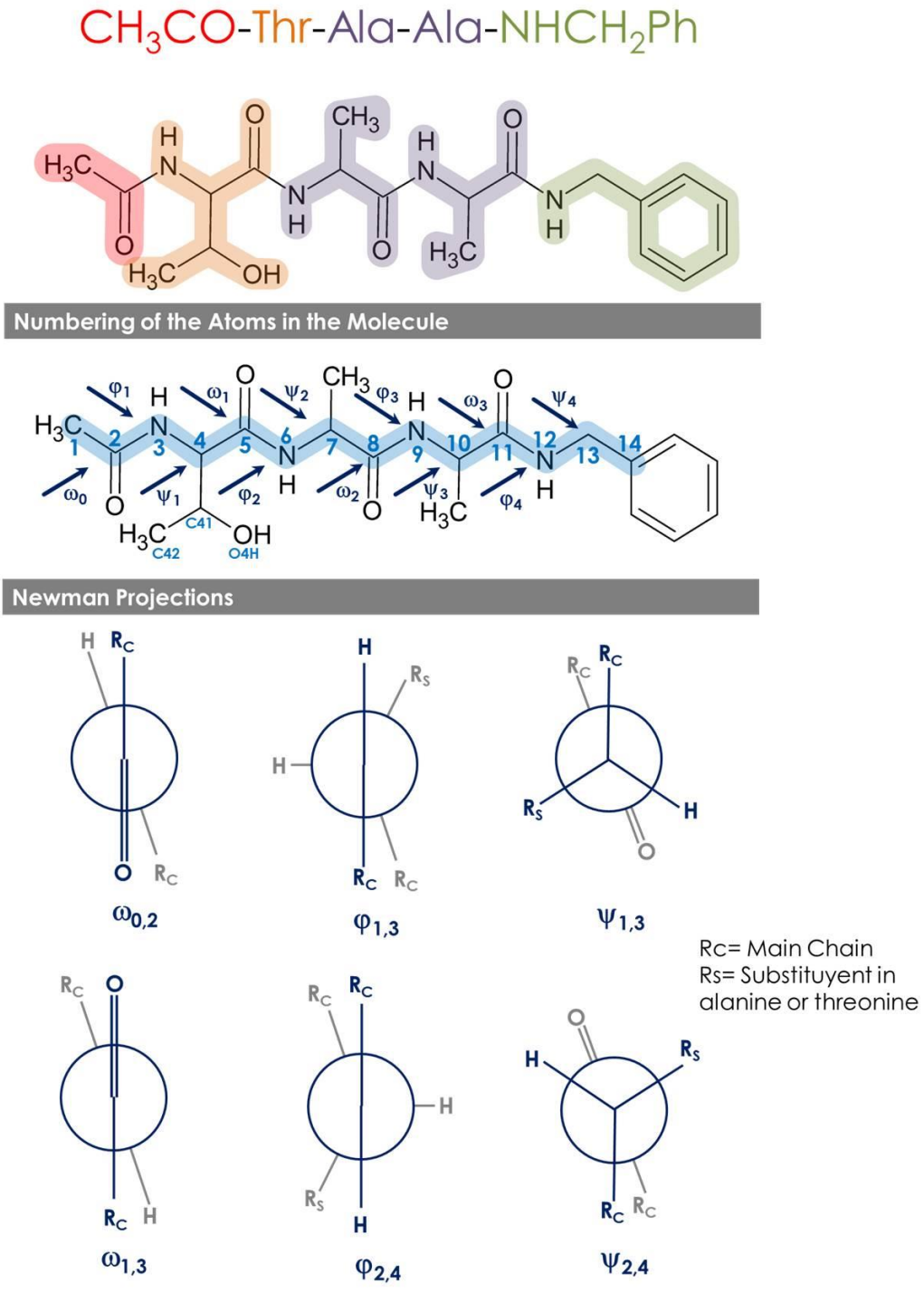


Figure A5.1: Description of the intramolecular interactions in the main chain of TAA.

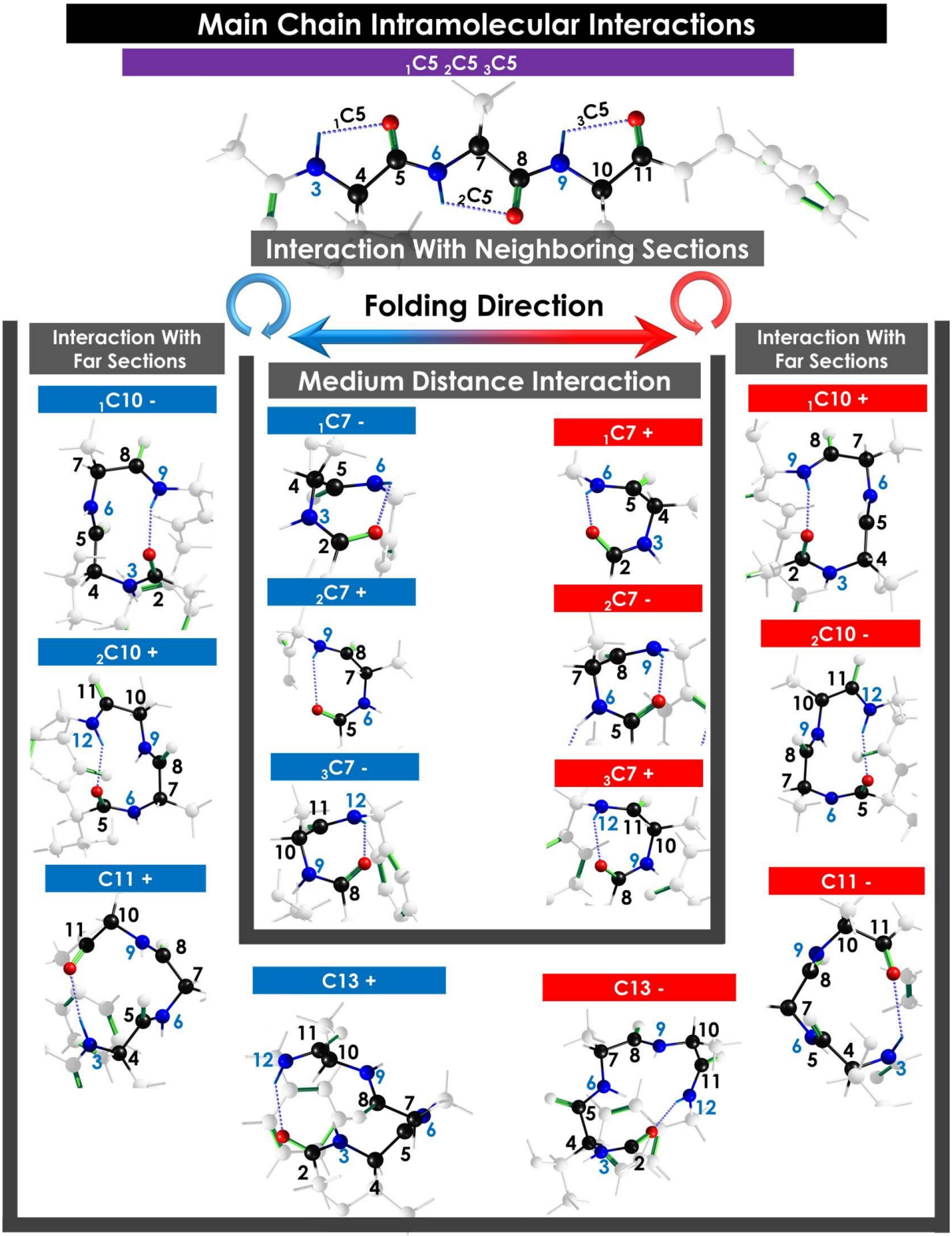


Figure A5.2: Main intramolecular interactions of the threonine unit.

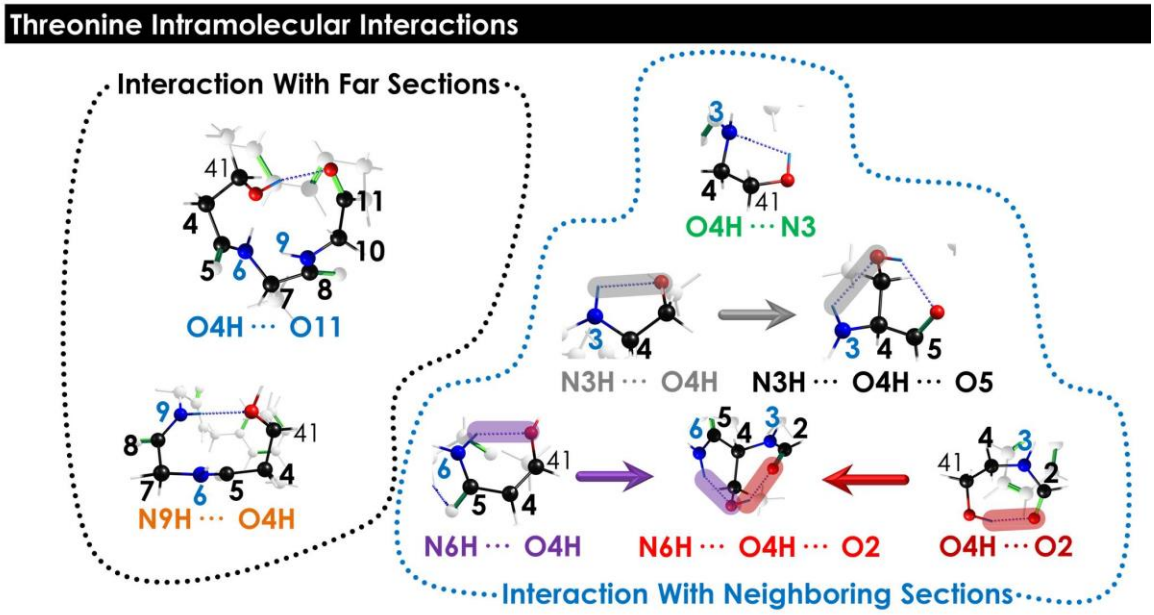


Figure A5.3: Main intramolecular interactions of the phenyl chromophore.

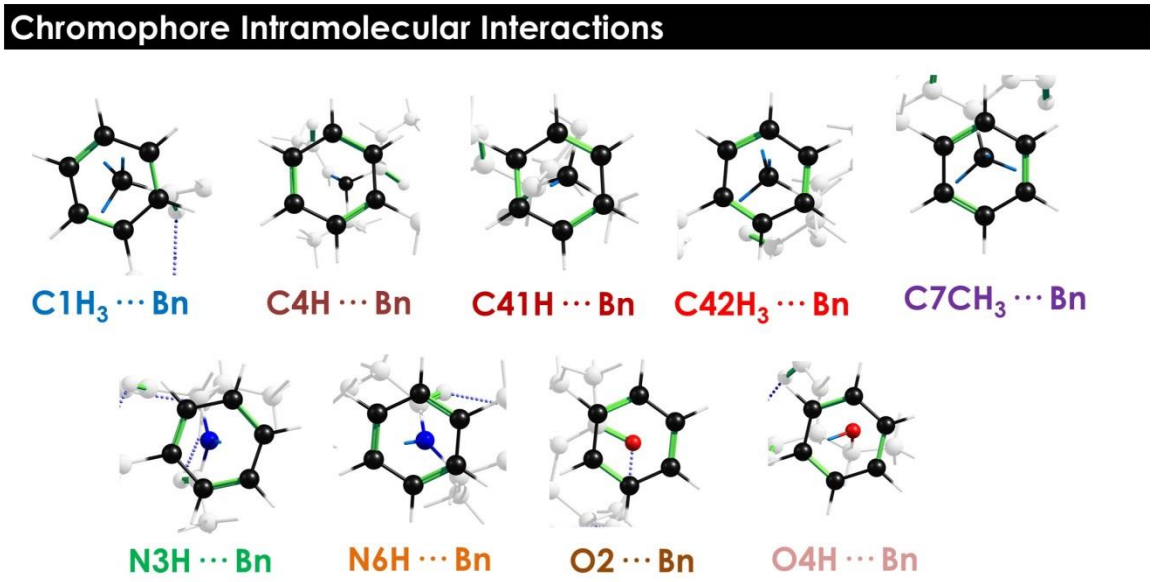


Table A5.1.1: Relative Energies (RE) of TAA computed conformers. The force fields used for the conformational search are also indicated.

LABEL	AMBER Glycam	OPLS 2005	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)
TAA_001	✓	✓	**0.00**	**0.00**
TAA_002		✓	**9.21**	**7.95**
TAA_003		✓	**9.72**	**3.47**
TAA_004		✓	**11.46**	**0.05**
TAA_005	✓		**12.45**	**7.77**
TAA_006	✓		**12.81**	**10.72**
TAA_007	✓		**14.14**	**12.01**
TAA_008			**14.45**	**8.07**
TAA_009	✓		**15.43**	**11.25**
TAA_010	✓	✓	**15.45**	**13.47**
TAA_011	✓	✓	**16.98**	**16.34**
TAA_012	✓		**17.09**	**12.56**
TAA_013	✓		**17.15**	**14.74**
TAA_014	✓		**17.31**	**14.12**
TAA_015		✓	**17.57**	**10.30**
TAA_016		✓	**17.86**	**11.26**
TAA_017		✓	**18.26**	**8.31**
TAA_018	✓		**18.72**	**15.42**
TAA_019	✓		**18.76**	**15.08**
TAA_020	✓	✓	**19.11**	**15.57**
TAA_021		✓	**19.27**	**14.72**
TAA_022	✓	✓	**19.37**	**12.53**
TAA_023		✓	**19.53**	**9.18**
TAA_024		✓	**19.60**	**4.59**
TAA_025	✓		**20.08**	**16.74**
TAA_026	✓		**20.14**	**9.82**
TAA_027	✓		**21.27**	**18.60**
TAA_028	✓		**21.43**	**8.11**
TAA_029	✓	✓	**21.58**	**12.15**
TAA_030		✓	**22.57**	**13.84**
TAA_031	✓		**22.66**	**12.20**
TAA_032	✓		**22.68**	**15.16**
TAA_033	✓		**23.44**	**15.10**
TAA_034	✓		**24.44**	**21.47**

Table A5.1.2: Relative Energies (RE) of TAA computed conformers. The force fields used for the conformational search are also indicated.

LABEL	AMBER Glycam	OPLS 2005	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)
TAA_035	✓		**24.49**	**21.59**
TAA_036	✓		**24.62**	**14.65**
TAA_037		✓	**24.77**	**17.70**
TAA_038	✓		**25.07**	**22.72**
TAA_039	✓		**25.16**	**24.61**
TAA_040	✓	✓	**25.49**	**20.85**
TAA_041	✓		**26.54**	**20.68**
TAA_042	✓		**27.26**	**16.32**
TAA_043	✓		**27.28**	**23.99**
TAA_044	✓		**27.33**	**21.65**
TAA_045	✓		**27.62**	**23.61**
TAA_046	✓		**28.07**	**15.43**
TAA_047	✓		**28.22**	**26.24**
TAA_048	✓		**28.54**	**22.84**
TAA_049	✓		**29.06**	**15.03**
TAA_050	✓		**29.74**	**24.64**
TAA_051	✓		**30.06**	**22.19**
TAA_052	✓		**30.80**	**28.13**
TAA_053	✓		**30.95**	**21.59**
TAA_054		✓	**30.99**	**17.71**
TAA_055	✓		**31.42**	**17.23**
TAA_056	✓		**31.46**	**30.12**
TAA_057	✓		**31.66**	**15.65**
TAA_058	✓		**31.79**	**25.74**
TAA_059	✓		**32.52**	**29.60**
TAA_060	✓		**33.02**	**31.26**
TAA_061	✓		**33.25**	**26.74**
TAA_062	✓		**33.92**	**19.65**
TAA_063		✓	**34.10**	**22.88**
TAA_064	✓		**34.43**	**22.81**
TAA_065	✓		**34.50**	**35.20**
TAA_066	✓		**34.83**	**30.35**

Table A5.1.3: Relative Energies (RE) of TAA computed conformers. The force fields used for the conformational search are also indicated.

LABEL	AMBER Glycam	OPLS 2005	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)
TAA_067	✓		**34.85**	**27.50**
TAA_068	✓		**34.92**	**29.93**
TAA_069	✓		**35.11**	**25.96**
TAA_070	✓		**35.19**	**25.14**
TAA_071	✓		**35.30**	**23.53**
TAA_072	✓		**35.32**	**26.28**
TAA_073	✓		**35.49**	**24.39**
TAA_074	✓		**35.65**	**27.89**
TAA_075	✓		**35.87**	**32.49**
TAA_076	✓		**35.94**	**27.31**
TAA_077	✓		**36.48**	**32.60**
TAA_078	✓		**36.64**	**27.10**
TAA_079	✓		**36.87**	**31.64**
TAA_080	✓		**37.59**	**24.29**
TAA_081	✓		**37.78**	**31.73**
TAA_082	✓		**37.83**	**27.07**
TAA_083	✓		**37.95**	**32.19**
TAA_084	✓		**38.04**	**36.62**
TAA_085	✓		**38.06**	**31.07**
TAA_086	✓		**38.10**	**29.77**
TAA_087	✓		**38.52**	**28.72**
TAA_088	✓		**39.30**	**32.92**
TAA_089	✓		**39.32**	**40.79**
TAA_090	✓		**39.43**	**29.82**
TAA_091	✓		**39.44**	**26.29**
TAA_092		✓	**39.44**	**27.08**
TAA_093	✓		**39.55**	**36.47**
TAA_094	✓		**39.66**	**33.58**
TAA_095	✓		**40.22**	**31.30**
TAA_096	✓		**40.42**	**16.99**
TAA_097	✓		**40.50**	**34.58**

Table A5.1.4: Relative Energies (RE) of TAA computed conformers. The force fields used for the conformational search are also indicated.

LABEL	AMBER Glycam	OPLS 2005	RE ZPE (kJ/mol)	RE GIBBS (kJ/mol)
TAA_098	✓		**40.93**	**32.40**
TAA_099	✓		**41.56**	**35.10**
TAA_100	✓		**41.72**	**27.75**
TAA_101	✓		**41.90**	**31.50**
TAA_102		✓	**42.74**	**29.47**
TAA_103	✓		**43.01**	**19.26**
TAA_104	✓		**43.65**	**39.43**
TAA_105	✓		**43.93**	**36.15**
TAA_106	✓		**44.26**	**41.81**
TAA_107	✓		**44.67**	**33.67**
TAA_108	✓		**44.78**	**41.55**
TAA_109	✓		**46.47**	**41.33**
TAA_110	✓		**47.18**	**32.82**
TAA_111	✓		**49.13**	**42.94**
TAA_112	✓		**49.38**	**38.62**
TAA_113	✓		**49.66**	**38.87**
TAA_114	✓		**49.68**	**40.71**
TAA_115	✓		**52.35**	**40.68**
TAA_116	✓		**52.61**	**46.20**
TAA_117	✓		**53.93**	**48.35**
TAA_118	✓		**53.94**	**31.04**
TAA_119	✓		**54.29**	**45.85**
TAA_120	✓		**54.61**	**43.55**
TAA_121	✓		**55.80**	**38.64**
TAA_122	✓		**55.90**	**30.75**
TAA_123	✓		**56.04**	**32.61**
TAA_124	✓		**56.13**	**34.98**
TAA_125	✓		**56.19**	**30.85**
TAA_126	✓		**57.73**	**38.27**
TAA_127	✓		**58.95**	**32.30**
TAA_128	✓		**61.25**	**57.21**
TAA_129	✓		**61.78**	**48.34**
TAA_130	✓		**63.78**	**50.36**

Table A5.2.1: Dihedral angles of TAA computed conformers.

	ω0	φ1	ψ1	ω1	φ2	ψ2	ω2	φ3	ψ3	ω3	φ4	ψ4
TAA_001	**-175**	69	-108	**170**	-60	-26	**175**	-96	26	**172**	105	-91
TAA_002	**174**	-67	110	**-170**	54	39	**180**	79	-5	**-179**	-86	-93
TAA_003	**-176**	-83	66	**-176**	-73	101	**-170**	58	37	**-177**	110	-37
TAA_004	**-175**	-83	74	**-168**	-69	-21	**176**	-100	21	**-179**	101	125
TAA_005	**170**	-79	81	**-158**	53	-128	**-177**	-72	-3	**179**	-89	56
TAA_006	**174**	-140	165	**162**	-91	57	**-167**	68	-79	**166**	-70	121
TAA_007	**-177**	-90	73	**-165**	69	3	**163**	-81	79	**-153**	85	-80
TAA_008	**-170**	-100	-2	**176**	77	-64	**156**	-70	93	**-170**	89	-146
TAA_009	**171**	-80	81	**-154**	49	-137	**-179**	-72	-7	**177**	-92	49
TAA_010	**-174**	62	-96	**171**	-60	-29	**175**	-100	33	**173**	97	-87
TAA_011	**-163**	-85	44	**-179**	70	-81	**157**	-74	86	**-173**	152	-34
TAA_012	**172**	-79	80	**-157**	52	-131	**178**	-80	74	**-174**	-163	-134
TAA_013	**180**	-83	69	**-174**	69	-77	**-177**	-85	54	**172**	88	159
TAA_014	**-3**	-81	-9	**177**	77	-54	**155**	-68	94	**-164**	77	54
TAA_015	**-178**	-164	157	**171**	-62	-19	**175**	-86	13	**-178**	91	101
TAA_016	**-176**	-166	163	**164**	-68	-28	**-177**	-85	60	**-177**	76	72
TAA_017	**-171**	-66	-22	**174**	-62	-18	**175**	-115	29	**-175**	99	-54
TAA_018	**-175**	61	-125	**178**	-56	-30	**169**	-79	77	**-166**	166	120
TAA_019	**-170**	-119	80	**-4**	-107	132	**179**	73	-69	**169**	-77	106
TAA_020	**-178**	-83	71	**-171**	71	-72	**162**	-60	136	**180**	84	-7
TAA_021	**176**	-66	118	**-172**	53	40	**-178**	75	-50	**-180**	-84	-97
TAA_022	**180**	-78	85	**175**	-80	77	**-157**	64	-90	**174**	-78	-46
TAA_023	**-173**	-66	-19	**175**	-65	-18	**174**	-112	28	**-175**	98	-52
TAA_024	**-176**	-83	68	**-169**	-70	-21	**178**	-84	-1	**176**	180	82
TAA_025	**172**	-157	115	**-179**	75	-65	**154**	-62	95	**-164**	78	60
TAA_026	**-178**	-83	76	**-175**	-84	68	**-171**	70	-78	**164**	-73	110
TAA_027	**179**	67	-95	**169**	-60	-29	**175**	-100	35	**174**	94	-85
TAA_028	**179**	-83	77	**-171**	73	-46	**175**	-75	94	**-175**	102	139
TAA_029	**179**	-83	73	**-159**	72	-56	**167**	-144	165	**179**	77	-93
TAA_030	**-179**	-85	74	**-171**	73	-52	**170**	-71	99	**-175**	102	143
TAA_031	**-177**	-83	74	**-170**	-86	54	**170**	-83	77	**-150**	79	93
TAA_032	**-180**	66	-123	**175**	-55	-32	**165**	-78	80	**-165**	-103	-132
TAA_033	**-178**	-164	155	**171**	-62	-19	**177**	-79	-4	**178**	160	-82
TAA_034	**179**	67	-94	**168**	-60	-29	**175**	-101	35	**173**	94	-85
TAA_035	**170**	-77	75	**-158**	47	-140	**173**	-100	16	**178**	-102	46

Table A5.2.2: Dihedral angles of TAA computed conformers.

	ω0	φ1	ψ1	ω1	φ2	ψ2	ω2	φ3	ψ3	ω3	φ4	ψ4
TAA_036	**-176**	-83	66	**175**	-152	-62	**167**	-64	-40	**172**	-116	43
TAA_037	**-174**	-64	-27	**171**	-123	42	**-176**	74	-71	**159**	-66	121
TAA_038	**-175**	53	-95	**172**	-58	-30	**174**	-100	36	**174**	91	-86
TAA_039	**-179**	-146	98	**-178**	77	-58	**154**	-68	92	**-162**	82	60
TAA_040	**179**	-66	-14	**170**	-135	41	**-175**	74	-69	**162**	-65	132
TAA_041	**-169**	77	-72	**155**	-52	138	**-178**	62	20	**179**	90	123
TAA_042	**175**	-157	-179	**177**	-82	65	**-175**	70	-77	**174**	-93	143
TAA_043	**171**	-136	-86	**177**	-82	78	**-160**	60	-114	**179**	-67	-34
TAA_044	**175**	-158	173	**176**	-84	66	**-179**	71	-56	**-175**	-80	98
TAA_045	**-168**	-57	-38	**175**	-125	47	**-173**	73	-70	**158**	-70	111
TAA_046	**177**	-81	84	**-174**	-156	172	**170**	-77	83	**-164**	82	-97
TAA_047	**165**	-134	-85	**176**	-82	76	**-160**	60	-112	**178**	-69	-31
TAA_048	**178**	-77	84	**-172**	169	-35	**174**	-82	75	**-152**	73	81
TAA_049	**-177**	-83	73	**-171**	-86	67	**-173**	-86	64	**-179**	-96	-112
TAA_050	**-167**	-58	-37	**175**	-124	46	**-174**	73	-71	**158**	-69	111
TAA_051	**168**	-75	81	**-161**	50	-148	**165**	-107	56	**-160**	89	102
TAA_052	**-167**	76	-91	**166**	-128	6	**173**	-82	84	**-154**	80	-83
TAA_053	**-178**	-79	89	**-175**	77	-57	**168**	-76	82	**-169**	-160	59
TAA_054	**175**	-156	169	**-172**	-66	-27	**179**	-112	23	**-177**	-176	-138
TAA_055	**179**	-80	82	**-169**	75	-50	**177**	-81	76	**180**	-98	-118
TAA_056	**173**	139	-74	**178**	-81	72	**-158**	59	-105	**180**	-85	20
TAA_057	**176**	-82	70	**-167**	69	-73	**164**	-57	137	**-180**	-145	-127
TAA_058	**171**	85	-73	**159**	-54	-43	**168**	-90	68	**174**	83	-89
TAA_059	**168**	-79	85	**-159**	48	-140	**177**	-82	68	**-167**	-169	34
TAA_060	**172**	-150	-83	**179**	-88	67	**-155**	58	-97	**167**	-69	-72
TAA_061	**-176**	76	-45	**-175**	-83	89	**-171**	75	-44	**-180**	-91	-145
TAA_062	**179**	-83	77	**-167**	-171	167	**159**	-74	87	**-169**	165	141
TAA_063	**-172**	-84	9	**159**	-70	-17	**-180**	-121	-45	**179**	-105	-153
TAA_064	**-174**	-85	64	**176**	-156	-71	**160**	-81	63	**-171**	143	17
TAA_065	**170**	129	-87	**178**	-81	72	**-157**	63	-95	**178**	-91	-8
TAA_066	**-170**	-153	37	**178**	66	-107	**173**	-64	-26	**175**	-97	35
TAA_067	**169**	-151	121	**-176**	75	-65	**154**	-67	94	**-164**	79	-121

Table A5.2.3: Dihedral angles of TAA computed conformers.

	ω0	φ1	ψ1	ω1	φ2	ψ2	ω2	φ3	ψ3	ω3	φ4	ψ4
TAA_068	-177	-113	78	10	-123	126	180	69	-88	164	-91	73
TAA_069	-176	-84	64	-173	-85	64	-179	178	-28	-173	-175	71
TAA_070	171	-81	71	-163	68	-77	165	-60	130	-15	78	-150
TAA_071	-175	-108	91	-5	-117	127	173	-82	75	-164	86	-77
TAA_072	177	-86	70	-169	63	-91	-179	-83	56	171	95	137
TAA_073	170	-59	127	-177	-153	-148	177	-79	85	-160	95	115
TAA_074	-170	-139	122	175	-157	-151	174	-78	85	-157	87	102
TAA_075	-175	-83	67	-170	72	-73	155	-74	113	-6	-139	89
TAA_076	-170	-83	81	177	-82	73	-167	69	-77	170	-81	132
TAA_077	-176	-83	70	-169	72	-72	157	-74	114	-9	-137	-88
TAA_078	-171	-83	55	-178	73	-73	153	-67	94	-168	-102	-120
TAA_079	-170	77	-111	173	-79	72	-168	166	-32	178	162	-104
TAA_080	-177	-84	70	-174	-101	108	-11	-117	104	-174	-180	-128
TAA_081	7	-148	90	-10	-63	-35	166	-135	137	-172	82	87
TAA_082	179	-173	155	178	72	-87	168	-65	-29	172	-115	43
TAA_083	-173	-86	65	-177	68	-85	-180	-85	55	177	92	136
TAA_084	161	-168	-70	-177	-92	67	-152	59	-98	174	-67	-45
TAA_085	-179	81	-58	179	-75	95	-176	75	-47	179	-91	-145
TAA_086	-169	-84	79	-179	-83	71	-166	68	-80	168	-77	-57
TAA_087	169	-77	73	-153	45	-133	161	-109	131	178	95	116
TAA_088	174	-80	75	-167	74	-62	165	-104	153	-28	-84	87
TAA_089	-175	-114	93	-20	-95	129	-177	73	-62	174	-79	104
TAA_090	177	85	-66	169	-60	-27	168	-145	133	180	64	-125
TAA_091	-177	-86	69	-178	-160	161	-172	73	-75	155	-89	-115
TAA_092	-177	-76	84	-172	180	-48	-179	-124	-14	-177	-172	72
TAA_093	-159	-131	80	-15	-93	134	-178	72	-65	172	-79	98
TAA_094	-178	76	-33	-175	-116	85	-6	-109	131	175	90	111
TAA_095	-178	-150	139	174	-159	-145	171	-78	84	-160	89	-72
TAA_096	-179	-84	77	-176	-162	160	175	-162	162	169	-95	60
TAA_097	-170	-87	66	-178	68	-84	180	-85	54	177	91	137
TAA_098	179	-169	154	177	-138	-142	176	-79	83	-158	86	103
TAA_099	-170	79	-77	156	-97	74	27	-140	135	174	84	-94
TAA_100	-179	-87	73	-176	-162	139	-179	72	-77	162	-84	84
TAA_101	179	-170	149	174	-79	74	176	-82	77	-154	90	116
TAA_102	-171	-83	57	177	-154	17	178	-179	-34	-177	-149	69

Table A5.2.4: Dihedral angles of TAA computed conformers.

	ω0	φ1	ψ1	ω1	φ2	ψ2	ω2	φ3	ψ3	ω3	φ4	ψ4
TAA_103	-179	-84	79	-176	-161	158	175	-163	165	175	160	-71
TAA_104	-178	-172	130	-176	-97	57	-167	69	-76	159	-77	-79
TAA_105	-167	-149	88	-176	-89	63	-162	66	-81	163	-75	-74
TAA_106	176	-153	117	-172	74	-56	161	-104	93	-20	-94	147
TAA_107	-168	77	-76	159	-51	139	-172	75	-55	166	169	133
TAA_108	-177	-175	143	-178	-90	58	-171	69	-76	163	-74	-58
TAA_109	-180	83	-54	171	-101	92	-3	-118	128	178	88	102
TAA_110	179	-169	149	174	-158	164	169	-80	80	-166	85	-84
TAA_111	177	-166	138	-173	58	-135	-169	-84	69	177	88	174
TAA_112	-165	59	-148	180	-150	63	15	-157	148	178	149	104
TAA_113	172	-161	121	-173	72	-77	157	-72	90	-169	-92	-139
TAA_114	174	-153	150	163	-94	100	-3	-122	116	-176	174	-117
TAA_115	-171	-82	77	-173	168	-32	176	-82	76	-150	75	81
TAA_116	174	-153	150	164	-74	85	-170	163	-30	-171	137	-56
TAA_117	176	-151	151	173	-147	175	-170	74	-73	162	-97	41
TAA_118	178	-168	149	170	-156	170	177	-162	160	169	-100	-116
TAA_119	-13	-96	113	-146	65	-99	162	-142	137	-174	89	-81
TAA_120	179	-155	157	161	-66	123	-19	-120	113	-173	120	-62
TAA_121	178	-169	149	170	-156	169	177	-163	165	176	171	96
TAA_122	178	-168	148	170	-157	170	176	-161	167	-179	97	116
TAA_123	178	-169	149	170	-156	169	177	-162	170	-179	104	121
TAA_124	178	-169	149	170	-156	169	177	-162	166	178	-177	-96
TAA_125	179	-171	150	171	-156	164	179	-85	67	179	-93	72
TAA_126	175	-163	149	175	-81	79	179	-160	156	169	-97	61
TAA_127	179	-171	150	175	-81	81	178	-159	162	178	-158	58
TAA_128	-179	-176	151	174	-83	60	-174	169	-121	177	-84	-91
TAA_129	178	-171	152	164	-134	154	174	77	-59	-180	-97	46
TAA_130	-177	-156	121	-172	68	-87	166	-148	131	179	93	128

Table A5.3.1: Structural parameters of TAA computed conformers.

	$_1$C10	$_2$C10	$_1$C7	$_2$C7	$_3$C7	$_1$C5	$_2$C5	$_3$C5	Threonine substituent	Chromophore	Other Folding Parameters
TAA_001	-	+	-						N6H ··· O4H ··· O2	N3H ··· Bn	
TAA_002	+	-	+						N3H ··· O4H ··· O5	C1H3 ··· Bn	
TAA_003		+	+	+					N3H ··· O4H ··· O5	C4H ··· Bn	
TAA_004		+	+						N3H ··· O4H ··· O5	C41H ··· Bn	
TAA_005		-	+						N3H ··· O4H ··· O5	C1H3 ··· Bn	
TAA_006				+	-	✓			O4H ··· O11	C41H ··· Bn	
TAA_007			+		+				N3H ··· O4H ··· O5	C7CH3 ··· Bn	(N9H ··· N3)
TAA_008				-	+				O4H ··· O11	N3H ··· Bn	
TAA_009		-	+						O4H ··· N3	N3H ··· Bn	
TAA_010	-	+	-						O4H ··· N3	N3H ··· Bn	
TAA_011			+	-	+				O4H ··· O11	C1H3 ··· Bn	
TAA_012			+		+				N3H ··· O4H ··· O5	C1H3 ··· Bn	
TAA_013			+	-	+				N3H ··· O4H ··· O5	C1H3 ··· Bn	
TAA_014				-	+				O4H ··· O11	N6H ··· Bn	
TAA_015		+				✓			N9H ··· O4H	C42H3 ··· Bn	
TAA_016		+			+	✓			N9H ··· O4H	C42H3 ··· Bn	
TAA_017	-	+							N3H ··· O4H	C41H ··· Bn	
TAA_018	-				+				N6H ··· O4H ··· O2	-	
TAA_019	+				-				N3H ··· O4H ··· O5	N3H ··· Bn	
TAA_020			+	-					N3H ··· O4H ··· O5	C1H3 ··· Bn	N3H ··· O11 (C11)-
TAA_021	+	-			-				N3H ··· O4H ··· O5	O4H ··· Bn	
TAA_022			+	+	-				N3H ··· O4H ··· O5	C4H ··· Bn	
TAA_023	-	+							N3H ··· O4H	C4H ··· Bn	
TAA_024		+	+						N3H ··· O4H ··· O5	O5;O4H ··· Bn	
TAA_025				-	+				O4H ··· O2	N6H ··· Bn	N3H ··· O11 (C11)-
TAA_026			+	+	-				N3H ··· O4H ··· O5	N6H ··· Bn	
TAA_027	-	+	-						-	N3H ··· Bn	
TAA_028			+	-	+				N3H ··· O4H ··· O5	-	
TAA_029			+	-				✓	N3H ··· O4H ··· O5	C1H3 ··· Bn	
TAA_030			+	-	+				N3H ··· O4H ··· O5	-	

Table A5.3.2: Structural parameters of TAA computed conformers.

	$_1$C10	$_2$C10	$_1$C7	$_2$C7	$_3$C7	$_1$C5	$_2$C5	$_3$C5	Threonine substituent	Chromophore	Other Folding Parameters
TAA_031			+	+	+				**N3H ··· O4H ··· O5**	C7H ··· Bn	
TAA_032	-	+-			+				**N6H ··· O4H ··· O2**	-	
TAA_033		+				✓			**N9H ··· O4H**	N3H;O5 ··· Bn	
TAA_034	-	+	-						-	**N3H ··· Bn**	
TAA_035		-	+						**N3H ··· O4H ··· O5**	**O4H ··· Bn**	
TAA_036			+						**N3H ··· O4H ··· O5**	**C41H ··· Bn**	**N12H ··· O2 (C13)-**
TAA_037	+				-				**N3H ··· O4H**	**N6H ··· Bn**	
TAA_038	-	+	-						**N6H ··· O4H**	**N3H ··· Bn**	
TAA_039				-	+				**N3H ··· O4H ··· O5**	**N6H ··· Bn**	**N3H ··· O11 (C11)-**
TAA_040	+				-				**N3H ··· O4H**	**C1H3 ··· Bn**	
TAA_041		+	-						**O4H ··· O2**	**N3H ··· Bn**	
TAA_042				+	-	✓			**N6H ··· O4H ··· Ph**	**O4H ··· Bn**	
TAA_043				+	--				**N6H ··· O4H**	**O4H ··· Bn**	**N3H ··· O11 (C11)+**
TAA_044				+	-	✓			**N6H ··· O4H ··· Ph**	**C42H3 ··· Bn**	
TAA_045	+				-				**N6H ··· O4H**	N6H;N3;C1H3 ··· Bn	
TAA_046			+		+		✓		**N3H ··· O4H ··· O5**	**C7CH3 ··· Bn**	
TAA_047				+	--				**N6H ··· O4H**	O4H;N3H ··· Bn	**N3H ··· O11 (C11)+**
TAA_048	+		+		+				**N3H ··· O4H ··· O5**	**N6H ··· Bn**	
TAA_049			+	+	+				**N3H ··· O4H ··· O5**	-	
TAA_050	+				-				**N3H ··· O4H**	N6H;N3;C1H3 ··· Bn	
TAA_051		-	+						**N3H ··· O4H ··· O5**	**C7CH3 ··· Bn**	
TAA_052	-		-		+				**N6H ··· O4H ··· O2**	C10 ··· Bn	
TAA_053			+	-	+				**N3H ··· O4H ··· O5**	-	
TAA_054		+				✓			-	**N3H ··· Bn**	
TAA_055			+	-	+				**N3H ··· O4H ··· O5**	-	
TAA_056				+					**N6H ··· O4H ··· O2**	**C1H3 ··· Bn**	**N3H ··· O11 (C11)+**
TAA_057			+	-					**N3H ··· O4H ··· O5**	-	**N3H ··· O11 (C11)-**
TAA_058	--	+-	-		+				**O4H ··· O2**	N3H;C4H ··· Bn	
TAA_059			+		+				-	**N3H ··· Bn**	

Table A5.3.3: Structural parameters of TAA computed conformers.

	$_1$C10	$_2$C10	$_1$C7	$_2$C7	$_3$C7	$_1$C5	$_2$C5	$_3$C5	Threonine substituent	Chromophore	Other Folding Parameters
TAA_060				+	-				O4H ··· O2	N6H ··· Bn	N3H ··· O11 (C11)+
TAA_061			-	+	-				O4H ··· O2	N3H ··· Bn	
TAA_062			+		+		✓		N3H ··· O4H ··· O5	-	
TAA_063	+		--						N3H ··· O4H	C41H;C42H3 ··· Bn	N12H ··· O2 (C13)-
TAA_064			+		+				N3H ··· O4H ··· O5	C4H ··· Bn	
TAA_065				+	-				N6H ··· O4H ··· O2	O2 ··· Bn	N3H ··· O11 (C11)+
TAA_066		-		--					O4H ··· O2	N3H ··· Bn	
TAA_067				-	+	✓			-	N6H ··· Bn	N3H ··· O11 (C11)-
TAA_068	+				-		✓		N3H ··· O4H ··· O5	O11;N9H ··· Bn	
TAA_069		+-	+	+					N3H ··· O4H ··· O5	C41H ··· Bn	
TAA_070			+	-					N3H ··· O4H ··· O5	-	N3H ··· O11 (C11)-
TAA_071	+				+		✓		N3H ··· O4H ··· O5	C7CH3 ··· Bn	
TAA_072			+	-	+				O4H ··· N3;Ph	N3H ··· Bn	
TAA_073					+				N3H ··· O4H ··· O5	N6H ··· Bn	
TAA_074					+	✓			N3H ··· O4H ··· O5	N6H ··· Bn	
TAA_075			+	-					N3H ··· O4H ··· O5;O11	C1H3 ··· Bn	N3H ··· O11 (C11)-
TAA_076			+	+	-				N3H ··· O4H	C4H ··· Bn	
TAA_077			+	-					N3H ··· O4H ··· O5	C1H3 ··· Bn	N3H ··· O11 (C11)-
TAA_078			+	-	+				N3H ··· O4H ··· O5;O11	-	
TAA_079		+-	-	+					N3H ··· O4H ··· O5	N3H ··· Bn	
TAA_080		+	+						N3H ··· O4H ··· O5	C41H ··· Bn	
TAA_081								✓	-	N3H ··· Bn	N12H ··· O2 (C13)+
TAA_082		-		-		✓			O4H ··· O2	C4H ··· Bn	
TAA_083			+	-	+				-	N3H ··· Bn	
TAA_084				+	-				O4H ··· O2	N6H ··· Bn	N3H ··· O11 (C11)+
TAA_085			-	+	-				O4H ··· O2	N3H ··· Bn	
TAA_086			+	+	-				-	N3H ··· Bn	

Table A5.3.4: Structural parameters of TAA computed conformers.

	$_1$C10	$_2$C10	$_1$C7	$_2$C7	$_3$C7	$_1$C5	$_2$C5	$_3$C5	Threonine substituent	Chromophore	Other Folding Parameters
TAA_087			+					✓	N3H ⋯ O4H ⋯ O5	O4H ⋯ Bn	
TAA_088			+	-				✓	N3H ⋯ O4H ⋯ O5	C1H3 ⋯ Bn	N3H ⋯ O11 (C11)-
TAA_089	+				-				O4H ⋯ N3	N3H ⋯ Bn	
TAA_090			-					✓	O4H ⋯ O2	N3H ⋯ Bn	
TAA_091			+		-		✓		N3H ⋯ O4H ⋯ O5	N6 ⋯ Bn	
TAA_092	+		+						N3H ⋯ O4H ⋯ O5	O2 ⋯ Bn	N12H ⋯ O2 (C13)-
TAA_093	+				-		✓		-	N3H ⋯ Bn	
TAA_094		+	-					✓	O4H ⋯ O2	N3H ⋯ Bn	
TAA_095					+	✓			O4H ⋯ O2	N6H ⋯ Bn	
TAA_096			+				✓	✓	N3H ⋯ O4H ⋯ O5	-	
TAA_097			+	-	+				-	N3H ⋯ Bn	
TAA_098					+	✓			O4H ⋯ O2	N6H ⋯ Bn	
TAA_099		+	-					✓	O4H ⋯ O2	N3H ⋯ Bn	
TAA_100			+		-		✓		N3H ⋯ O4H ⋯ O5	N6 ⋯ Bn	
TAA_101				+	+	✓			O4H ⋯ O2	C7CH3 ⋯ Bn	
TAA_102	+		+						N3H ⋯ O4H ⋯ O5	O2 ⋯ Bn	N12H ⋯ O2 (C13)-
TAA_103			+				✓	✓	N3H ⋯ O4H ⋯ O5	-	
TAA_104				+	-	✓			O4H ⋯ O2	N6H ⋯ Bn	
TAA_105				+	-				O4H ⋯ O2	N6H ⋯ Bn	
TAA_106				-		✓			O4H ⋯ O2	N6H ⋯ Bn	N3H ⋯ O11 (C11)-
TAA_107			-		-				O4H ⋯ O2	N3H ⋯ Bn	
TAA_108				+	-	✓			O4H ⋯ O2	C42H3 ⋯ Bn	
TAA_109		+	-						O4H ⋯ O2	N3H ⋯ Bn	
TAA_110					+	✓	✓		O4H ⋯ O2	C7CH3 ⋯ Bn	
TAA_111					+	✓			O4H ⋯ O2	C1H3 ⋯ Bn	
TAA_112			-					✓	N6H ⋯ O4H ⋯ O2	O2 ⋯ Bn	N12H ⋯ O2 (C13)+
TAA_113				-	+	✓			O4H ⋯ O2	-	N3H ⋯ O11 (C11)-
TAA_114		+				✓			O4H ⋯ O2	N3H ⋯ Bn	
TAA_115	+		+		+				N3H ⋯ O4H	N6H ⋯ Bn	
TAA_116		+		+		✓			O4H ⋯ O2	N3H ⋯ Bn	

Table A5.3.5: Structural parameters of TAA computed conformers.

	$_1$C10	$_2$C10	$_1$C7	$_2$C7	$_3$C7	$_1$C5	$_2$C5	$_3$C5	Threonine substituent	Chromophore	Other Folding Parameters
TAA_117					-	✓	✓		**O4H ··· O2**	**C42H3 ··· Bn**	
TAA_118						✓	✓	✓	**O4H ··· O2**	-	
TAA_119								✓	**O4H ··· N3**	**N3H ··· Bn**	**N12H ··· O2 (C13)+**
TAA_120		+				✓			**O4H ··· O2**	**N3H ··· Bn**	
TAA_121						✓	✓	✓	**O4H ··· O2**	-	
TAA_122						✓	✓	✓	**O4H ··· O2**	-	
TAA_123						✓	✓	✓	**O4H ··· O2**	-	
TAA_124						✓	✓	✓	**O4H ··· O2**	-	
TAA_125					+	✓	✓		**O4H ··· O2**	-	
TAA_126				+		✓		✓	**O4H ··· O2**	-	
TAA_127				+		✓		✓	**O4H ··· O2**	-	
TAA_128				+		✓		✓	**O4H ··· O2**	**C41H ··· Bn**	
TAA_129					-	✓	✓		**O4H ··· O2**	**C42H3 ··· Bn**	
TAA_130				-		✓		✓	**O4H ··· O2**	**C1H3 ··· Bn**	

Figure A5.4: All the computed conformations were grouped into families. **Purple family:** unfolded structures, mainly constituted by neighbor interactions. **Blue family:** TAA folds in one direction, the interactions that characterize this family are mainly medium- and long-range interactions. **Red family:** the same situation observed in blue family, but the folding takes place on the opposite direction. **Green family:** composed by a combination of short, medium and long distance intramolecular interactions, with the particularity that the folding direction is a combination of both red and blue (see Figure A5.1), producing semi-folded isomers. **Grey and Black families:** the rest of the structures.

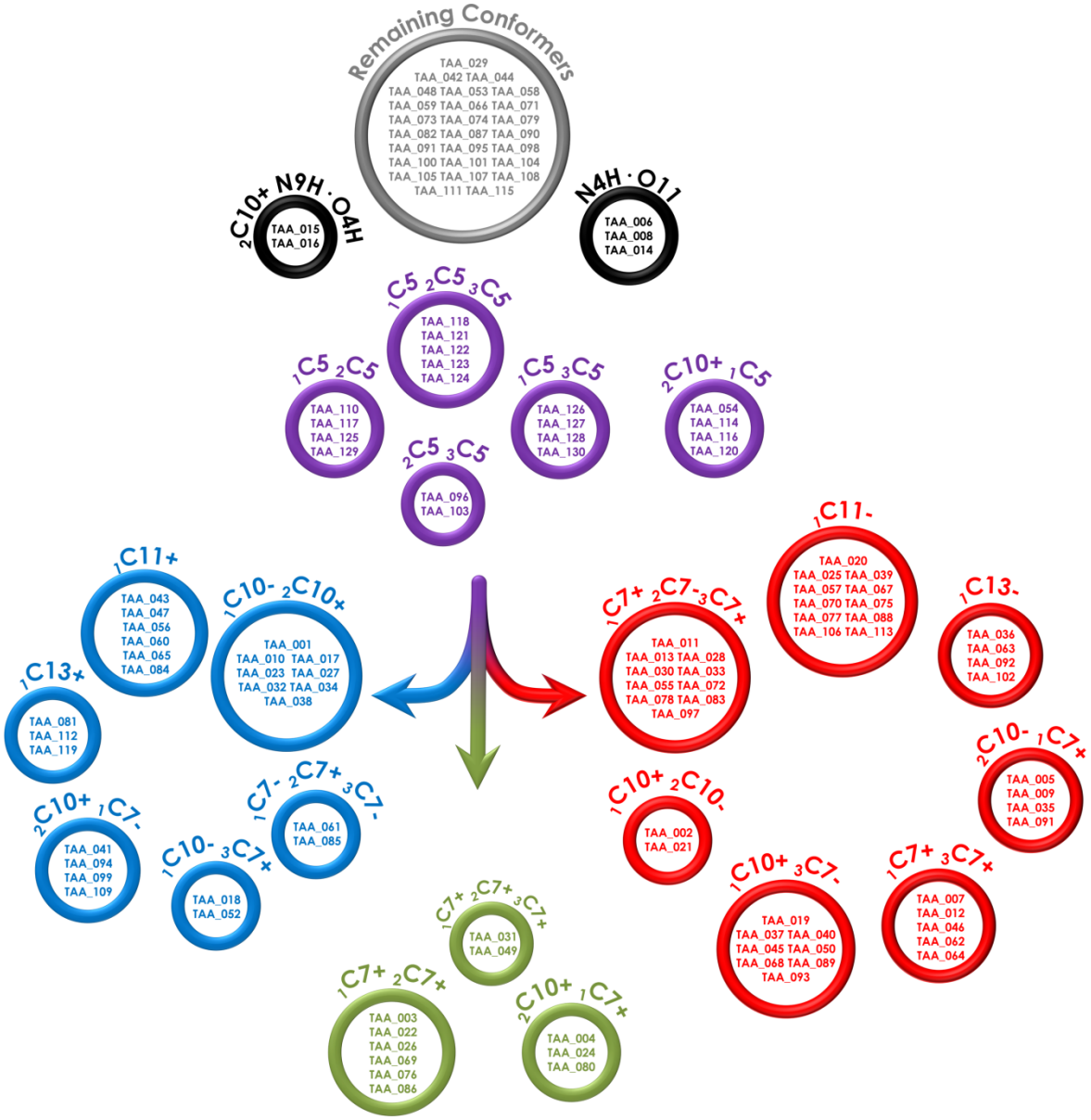


Figure A5.5: Representative conformers of each family.

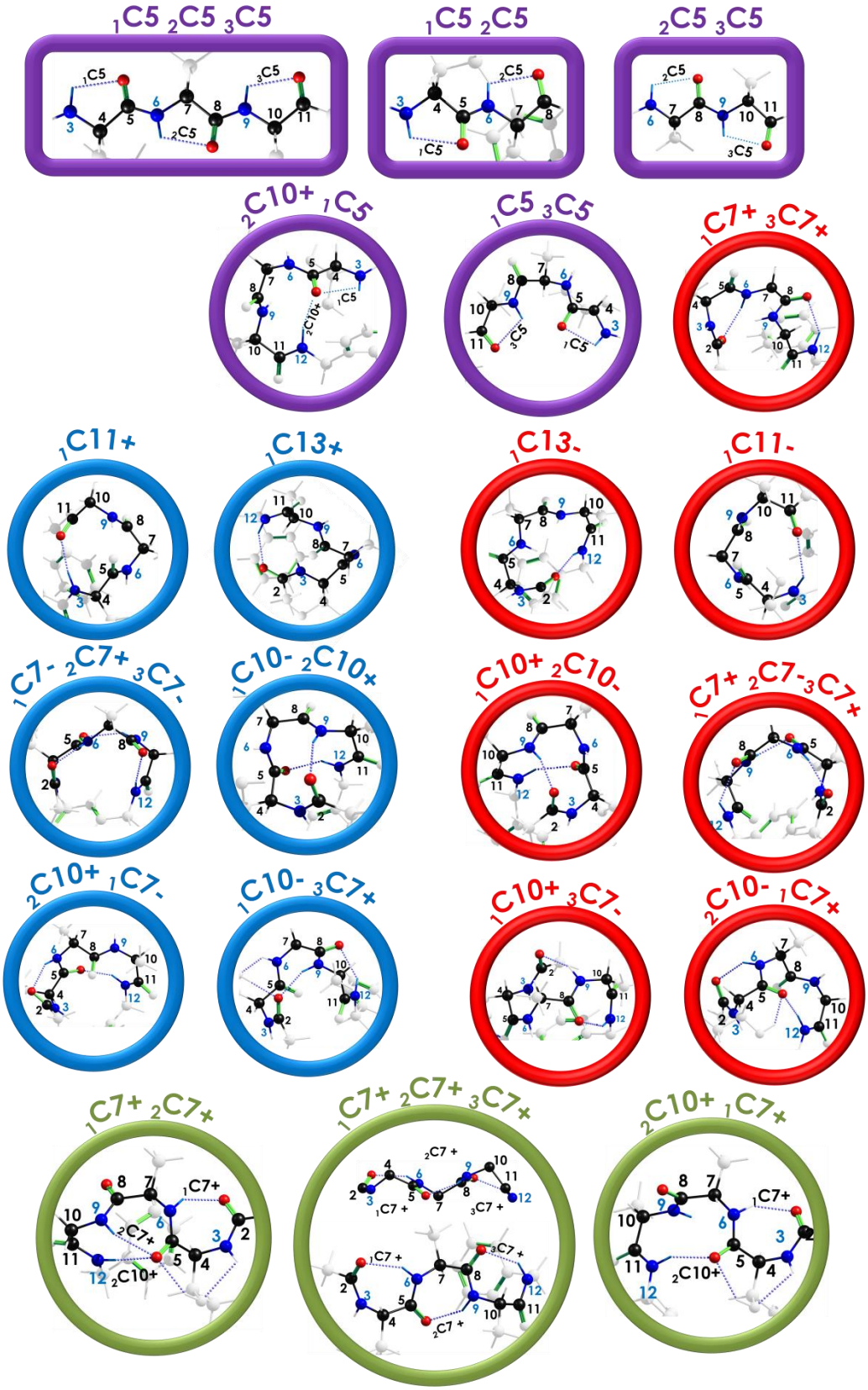


Figure A5.6: Relative Gibbs Free Energy Diagrams for **a)** black and grey families. **b)** purple and green families. **c)** blue and red families. **d)** Zoom of the most stable conformations of all families.

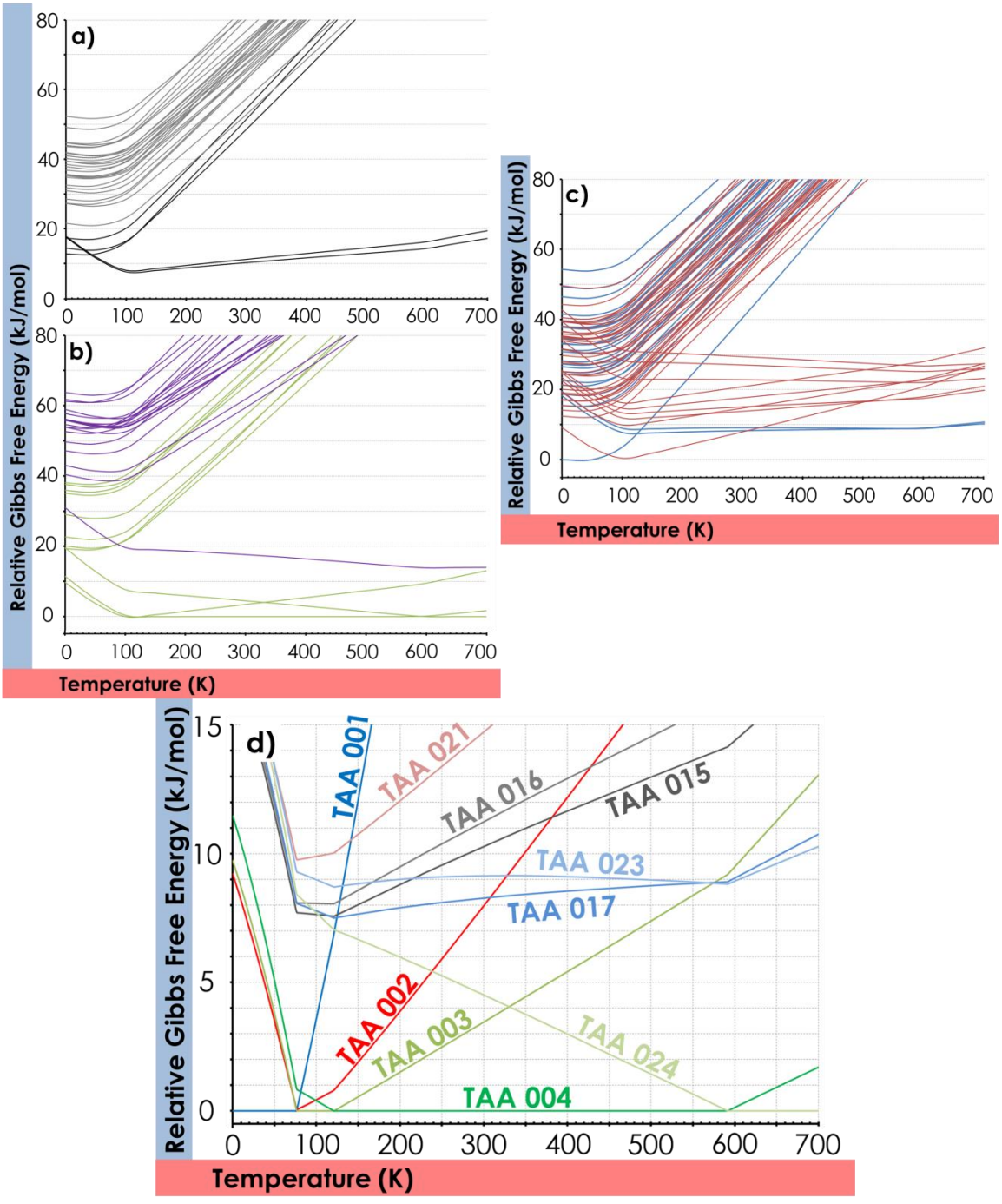


Figure A5.7: Comparison between the experimental spectrum (upper trace, in black) and the main theoretical predictions (traces in colors). The calculation level used was M06-2X/6-31+G(d). In order to account for the anharmonicity, a factor of 0.953 was used, which was determined using already assigned molecules. TAA presents four NHs and one OH belonging to threonine.

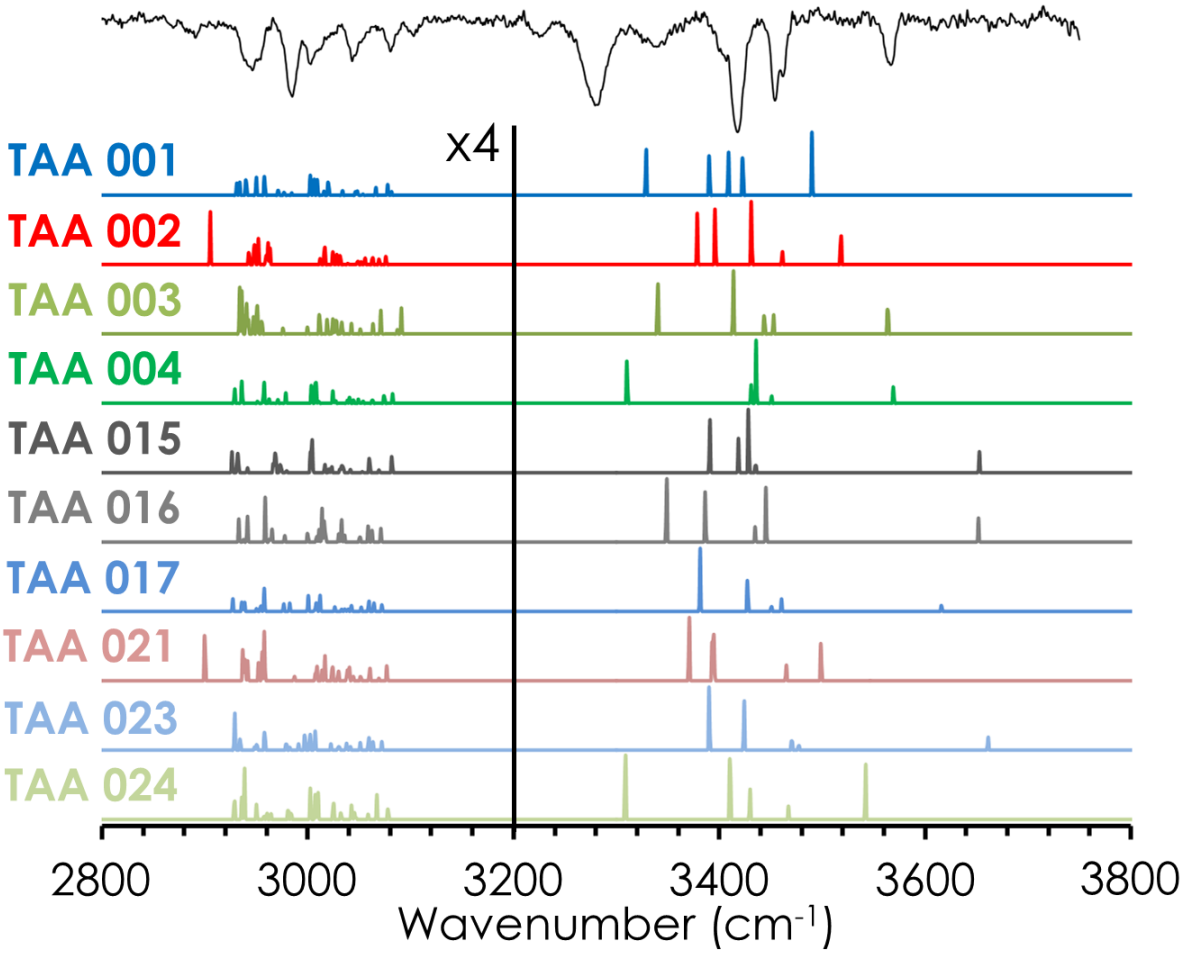


Figure A5.8: Structures of the five most stable conformations of TAA.

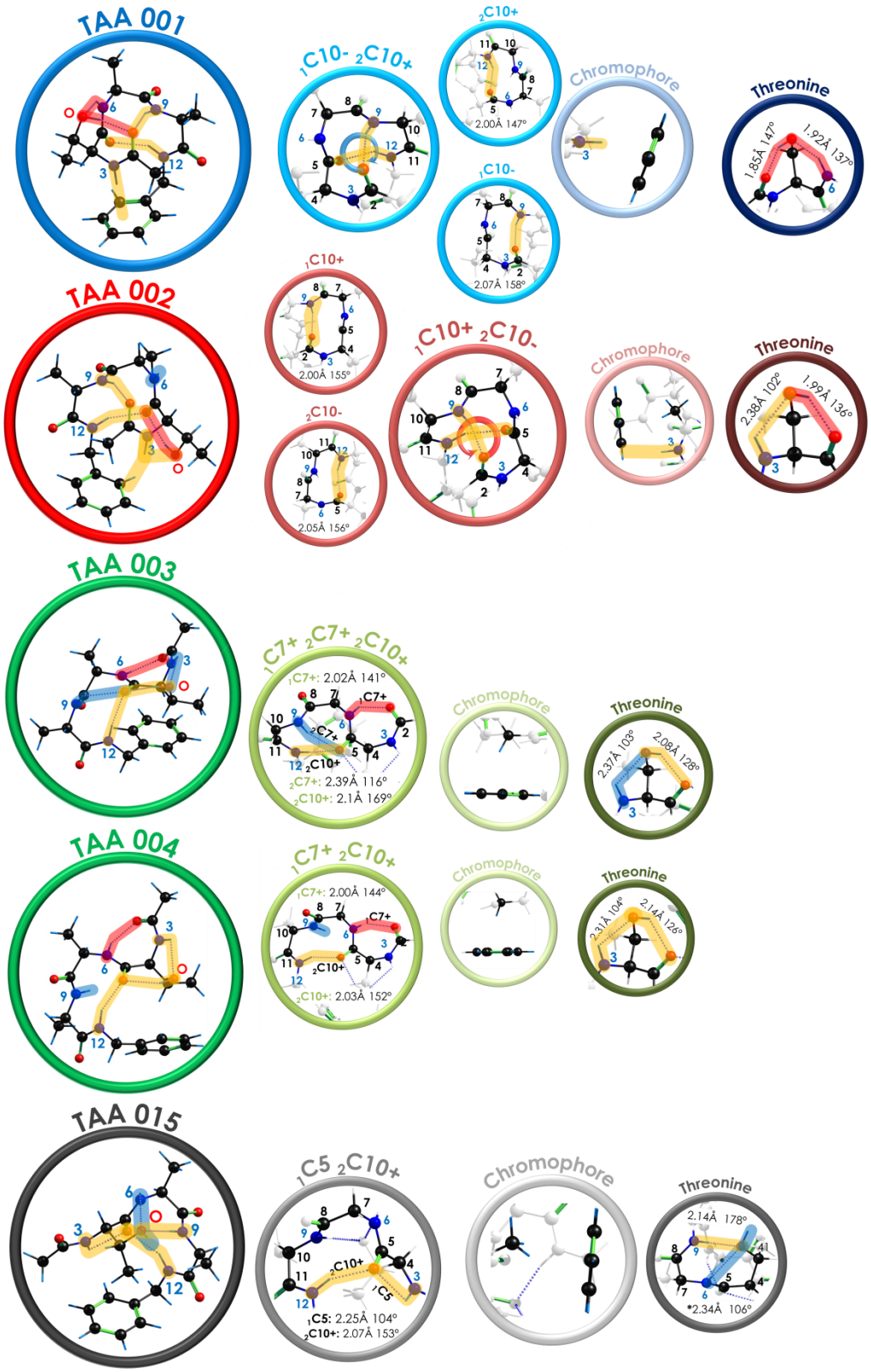


Figure A5.9: Estimation of the isomerization energy and pathway between TAA 003 and TAA 004. The traces highlighted in blue correspond to the scanned dihedral angle C11-C10-N9-H. Since the selected path does not have a clear and easy coordinate to scan, two jumps occurred in the scanned coordinate. The traces highlighted in yellow correspond to the energy optimization values coming from the same calculation and are reported in order to fill these gaps. Calculation level: M06-2X/6-31+G(d).

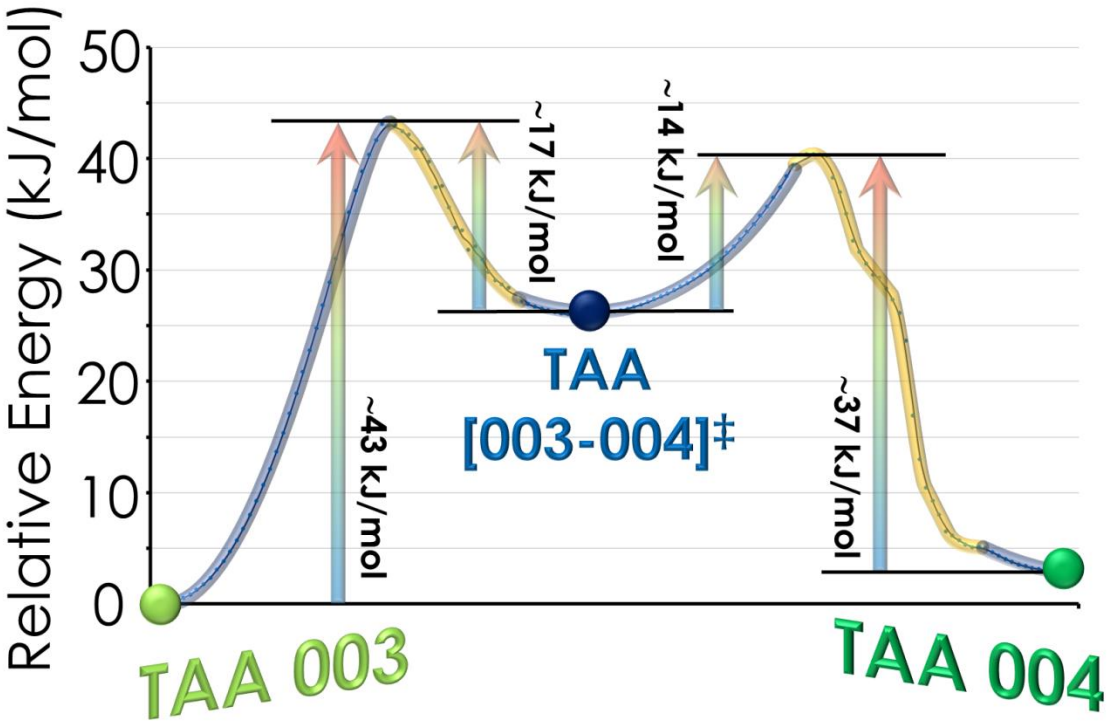


Figure A5.10: Most stable structures of the polymer based on π···π interactions . Calculation level: B3LYP/6-31g(d,p).

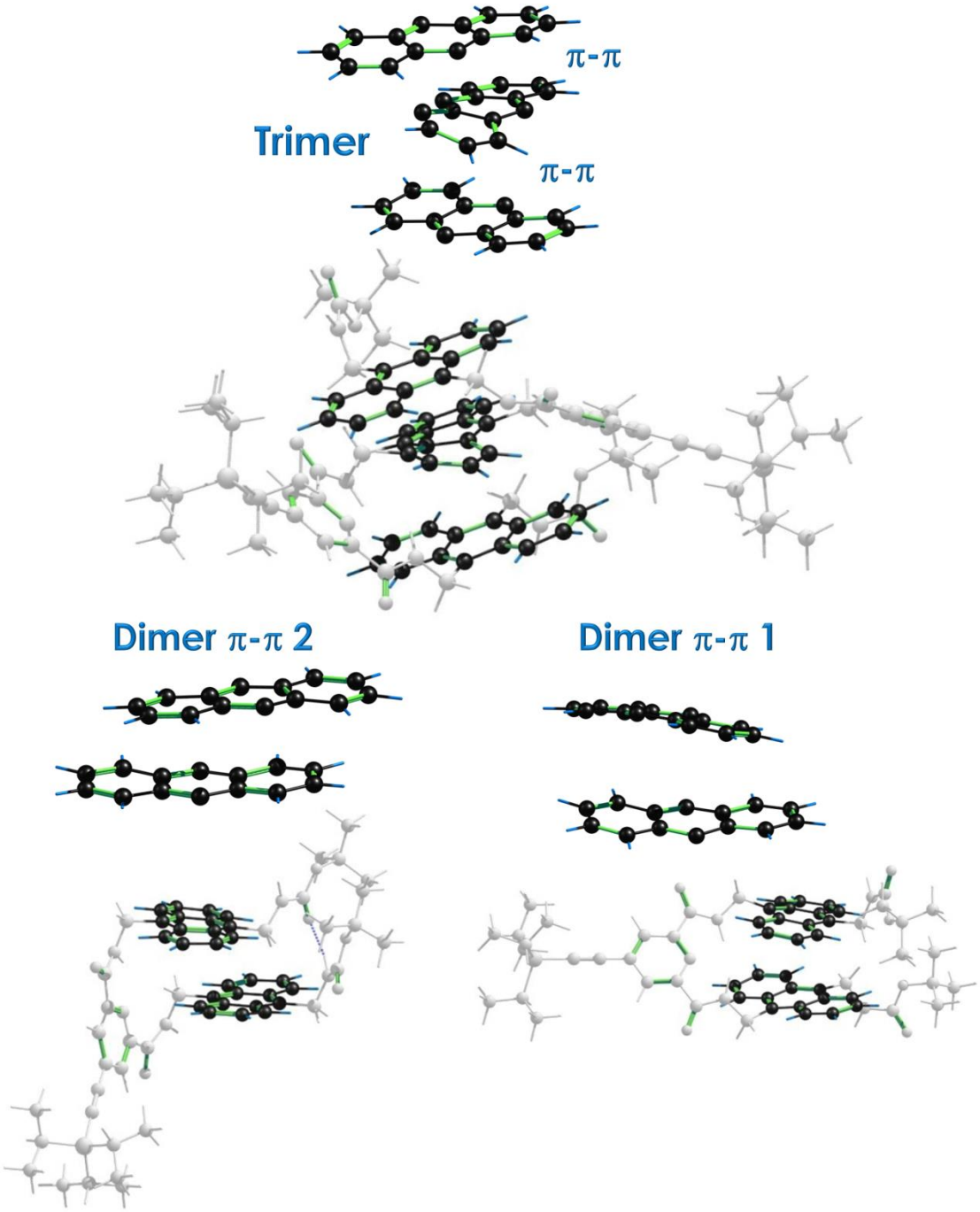
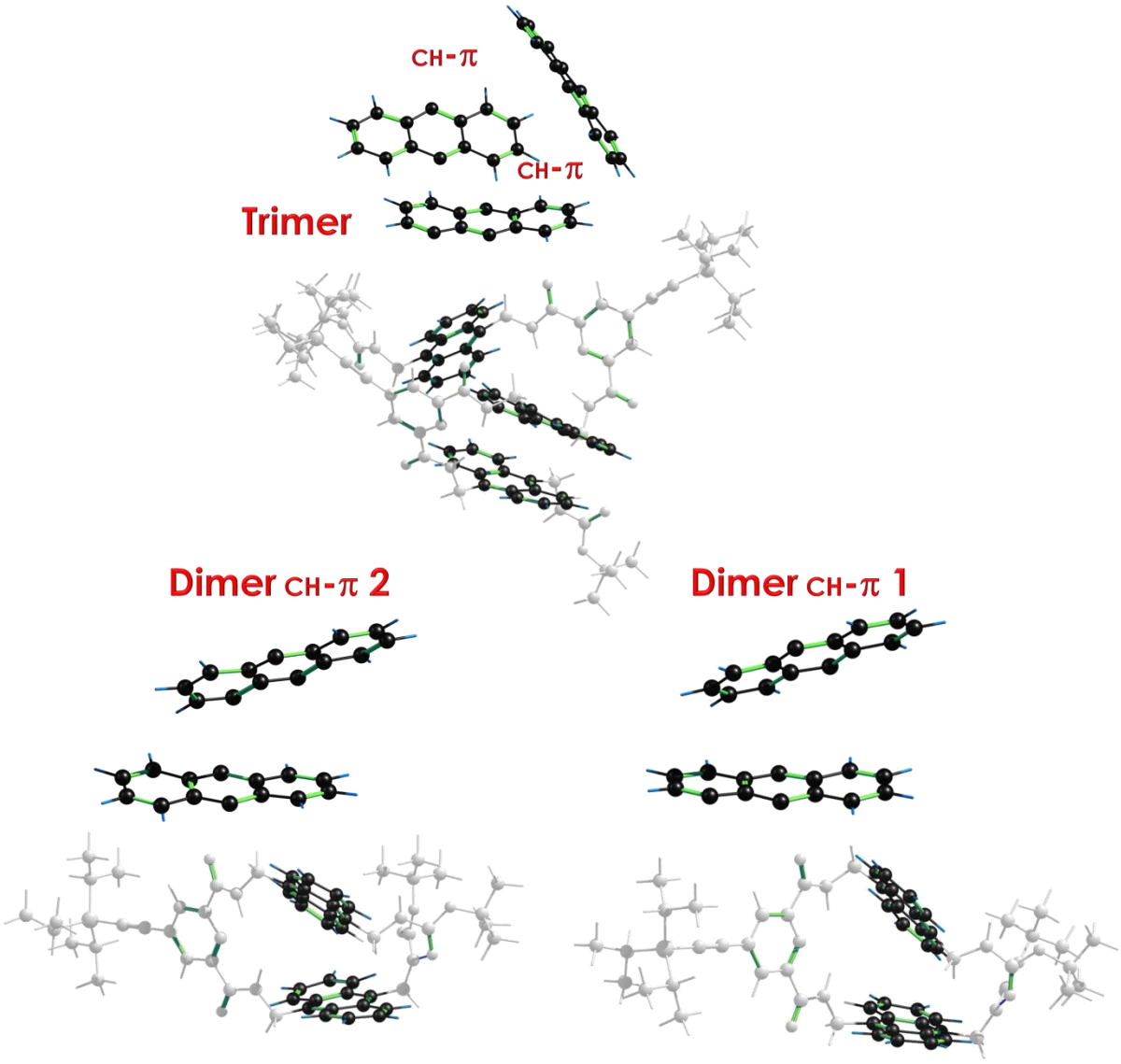


Figure A5.11: Most stable structures of the polymer based on CH···π interactions . Calculation level: B3LYP/6-31g(d,p).



Appendix for Chapter 6

Analysis of the Nucleation Process in Theobromine

Figure A6.1: REMPI spectra of: a) Theobromine b) Theobromine-water clusters.

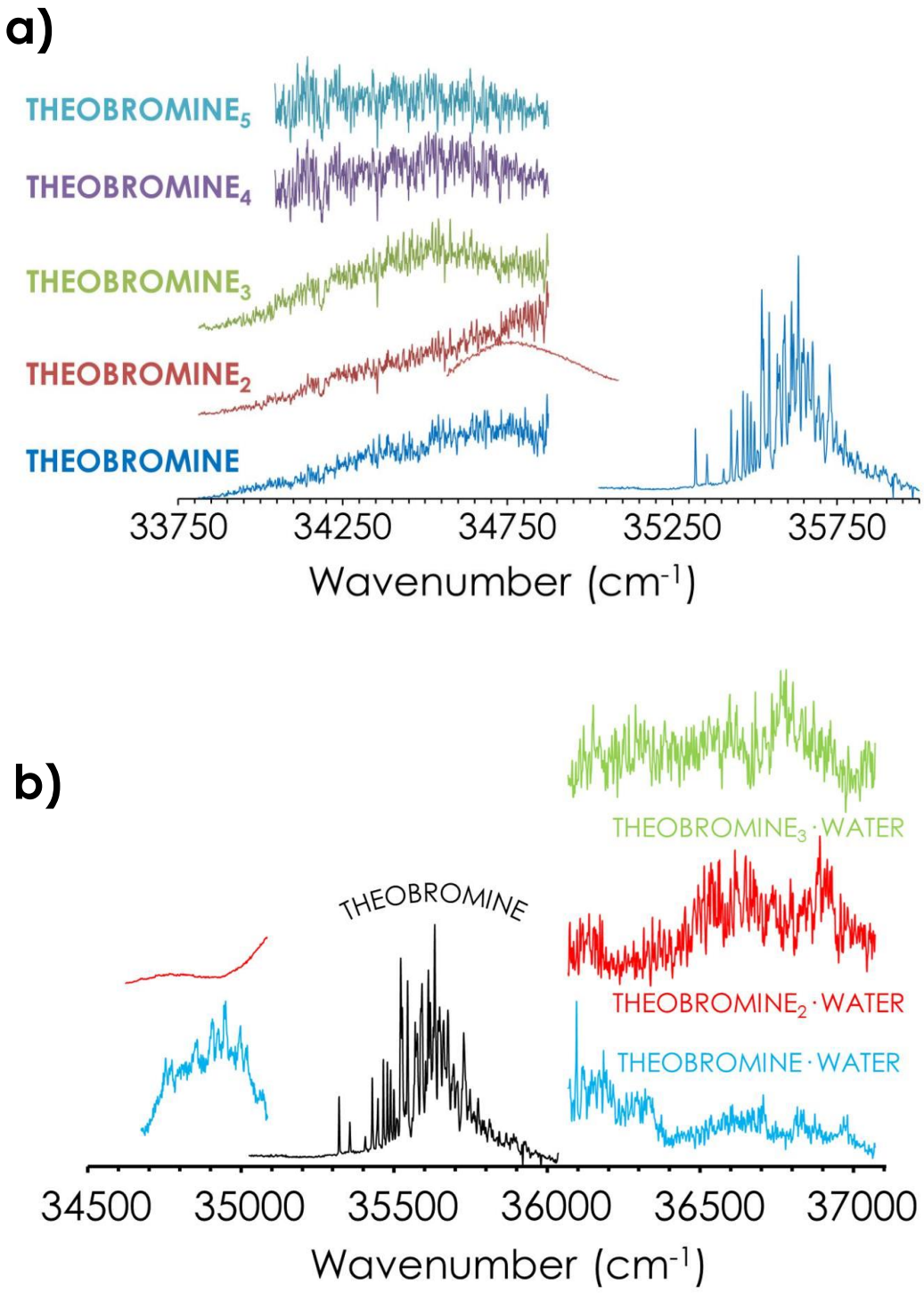


Figure A6.2: SPES of theobromine determined through the computation of the tautomerization energy in an isolated molecule. All the calculations are obtained using the M06-2X/6-311++G(d,p) level. a) ~6 kJ/mol (enol$_B$-enol$_T$); b) ~48 kJ/mol (enol$_T$-enol$_B$); c) ~129 kJ/mol (enol$_T$-keto); d) ~196 kJ/mol (keto-enol$_T$); e) ~204 kJ/mol (keto-enol$_T$); f) ~139 kJ/mol (enol$_T$-keto); g) ~52 kJ/mol (enol$_T$-enol$_B$); h) ~7 kJ/mol (enol$_B$-enol$_T$).

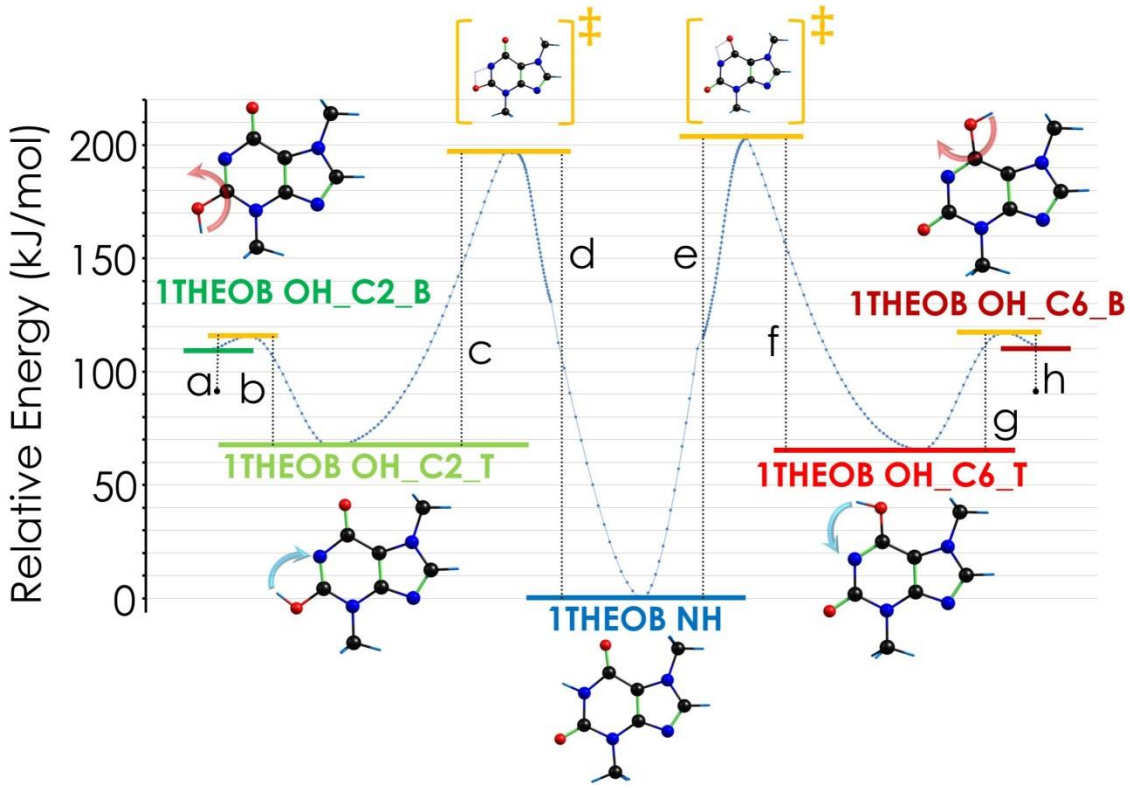


Table of the scanned coordinates:

OH Rotation	**Dihedral Angle**	**Initial**	**Final**		
a & b barrier	H-O-C2-N1	172°	362°		
Tautomerization	**Distance**	**Initial**	**Final**	**Fixed Coordinate**	
C2O-H to N1-H	N1-H	2.25 Å	1.01 Å		
c & d barrier	C2O-H	1.26 Å	2.56 Å	N1-H	F(1.31) Å
Tautomerization	**Distance**	**Initial**	**Final**	**Fixed Coordinate**	
C6O-H to N1-H	N1-H	2.25 Å	1.01 Å		
e & f barrier	C6O-H	1.260 Å	2.56 Å	N1-H	F(1.31) Å
OH Rotation	**Dihedral Angle**	**Initial**	**Final**		
g & h barrier	H-O-C6-N1	174°	364°		

Figure A6.3: a) Relative Gibbs free energy diagram of theobromine; b) experimental IDIRS spectrum of theobromine monomer, upper black trace, obtained ionizing at ~34843 cm^{-1}; color traces: theoretical predictions (correcting factor for OH 0.934, and for NH and CH: 0.953). All the calculations performed at M06-2X/6-311++G(d,p) level.

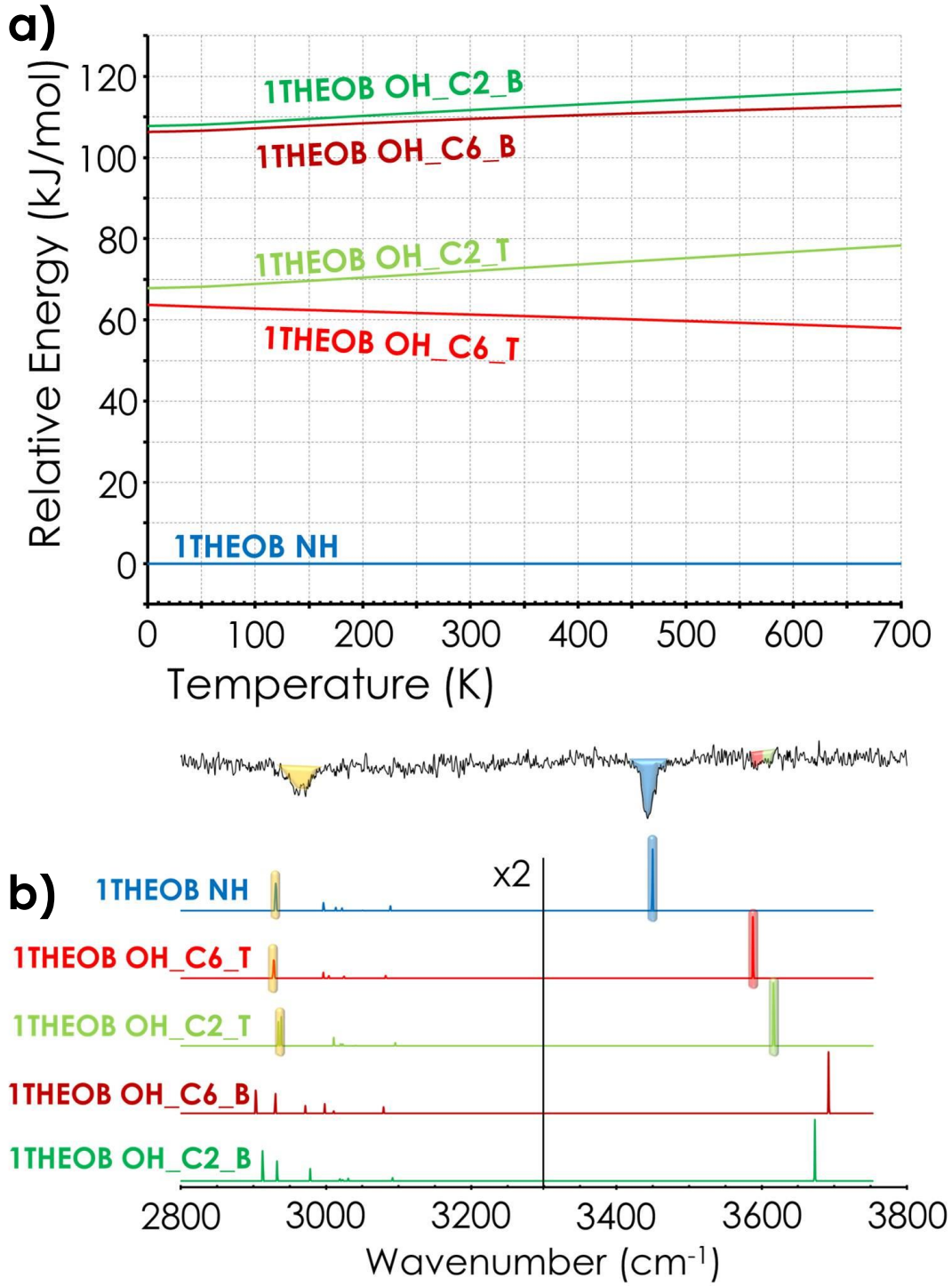


Figure A6.4: Theobromine-water complex: connections between the main conformations and other close conformers are connected by similarity. The conformational search was done using the MMFFs force field with Macromodel application, and completed using chemical intuition. Several structures were probed in order to obtain a conformer where the water interacts with the ring of theobromine (see the graphics below: energy optimization diagrams). Except for 1THEOB+W 07, all the other structures show the same trend to form H-bonds. Different initial positions were used as starting points for the optimization, but always the same four families were obtained. An example for each family is displayed below.

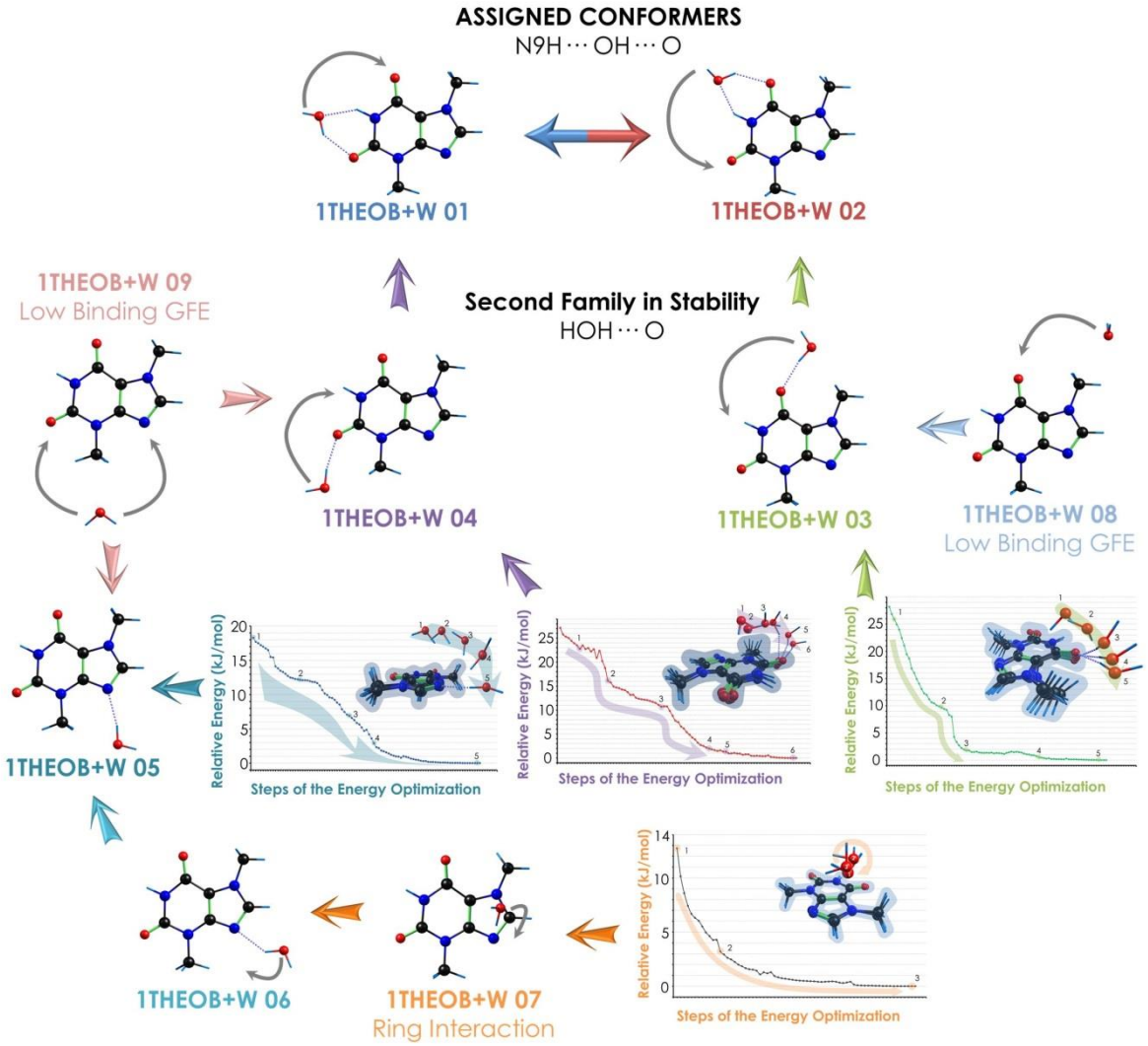
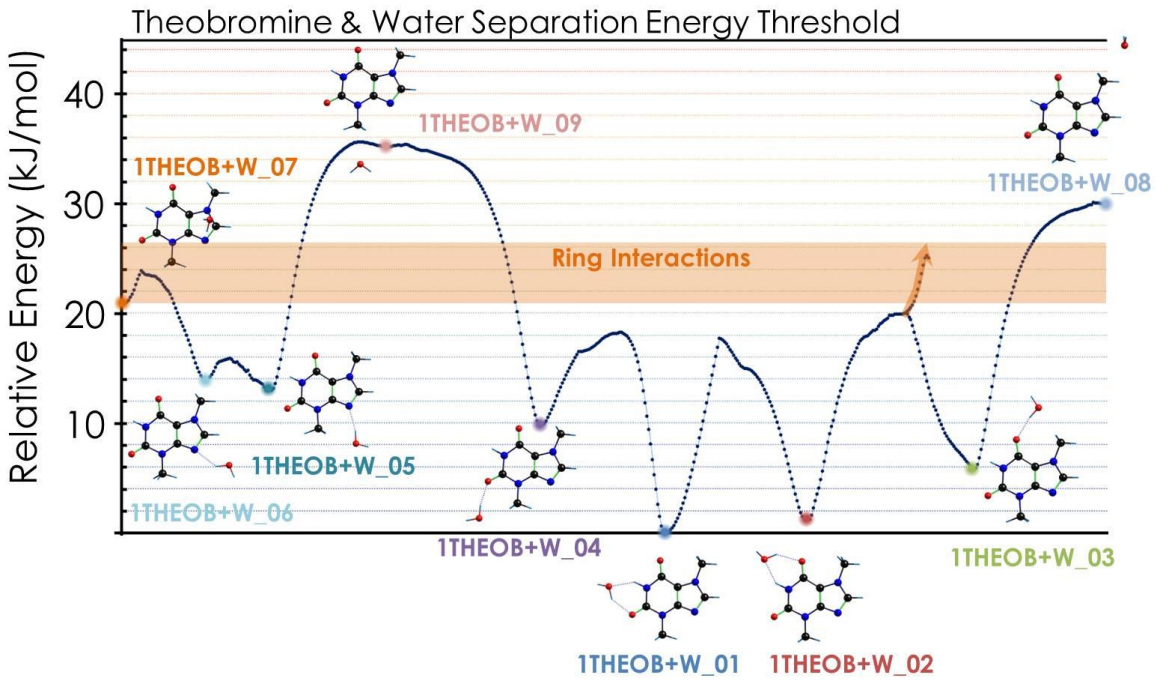


Figure A6.5.1: SPES for theobromine-water interaction. The main possible minima in energy are connected with the most similar geometry and the estimation of the barrier is also reported. All the calculations performed at the M06-2X/6-311++G(d,p) level.



Barriers Energy Table:

1THEOB+W_07	↓	~3 kJ/mol	1THEOB+W_04	↓	~8 kJ/mol
1THEOB+W_06	↑	~10 kJ/mol	1THEOB+W_01	↑	~18 kJ/mol
1THEOB+W_06	↓	~2 kJ/mol	1THEOB+W_01	↓	~18 kJ/mol
1THEOB+W_05	↑	~3 kJ/mol	1THEOB+W_02	↑	~17 kJ/mol
1THEOB+W_05	↓	~22 kJ/mol	1THEOB+W_02	↓	~19 kJ/mol
1THEOB+W_09	↑	~0.5 kJ/mol	1THEOB+W_03	↑	~14 kJ/mol
1THEOB+W_09	↓	~0.3 kJ/mol	1THEOB+W_03	↓	~24 kJ/mol
1THEOB+W_04	↑	~25 kJ/mol	1THEOB+W_08	↑	~0.1 kJ/mol

Figure A6.5.2: Table of the scanned coordinates for the calculations of Figure A6.5.1.

1THEOB·W_01 to 02	**Distance H-O**	**Initial**	**Final**		
	C2O···HO**H-O**C6	5.02 Å	1.92 Å		
	Angle N-H-O	**Initial**	**Final**	**Fixed Coordinate**	
	N1-H-OH_2	147°	177°	C2O···HO**H-O**C6	F(3.52) Å
1THEOB·W_03 to 02	**Distance H-O**	**Initial**	**Final**		
	N1**H-O**H_2	5.26 Å	1.96 Å		
	Angle C-O-O	**Initial**	**Final**	**Fixed Coordinate**	
	C6-O-OH_2	82°	132°	N1**H-O**H_2	F(3.40) Å
1THEOB·W_08 to 03	**Distance O-H**	**Initial**	**Final**		
	C6**O-H**OH	5.52 Å	1.86 Å		
1THEOB·W_01 to 04	**Distance O-C**	**Initial**	**Final**		
	H2**O-C**H_3N3	5.41 Å	3.05 Å		
	Angle C-O-O	**Initial**	**Final**	**Fixed Coordinate**	
	C2-O-OH_2	88°	132°	H2**O-C**H_3N3	F(4.11) Å
1THEOB·W_09 to 04	**Distance O-H**	**Initial**	**Final**		
	C2**O-H**OH	5.47 Å	1.87 Å		
	Dihedral angle	**Initial**	**Final**	**Fixed Coordinate**	
	C4-N3-C-H	50°	0°	C2**O-H**OH	F(5.03) Å
1THEOB·W_09 to 05	**Distance N-H**	**Initial**	**Final**		
	N9-HOH	5.29 Å	1.99 Å		
1THEOB·W_06 to 05	**Distance C-O**	**Initial**	**Final**		
	N3**C-O**H_2	5.01 Å	3.11 Å		
1THEOB·W_07 to 06	**Distance N-H**	**Initial**	**Final**		
	N9-HOH	2.59 Å	2.14 Å		
	Angle C-N-H	**Initial**	**Final**	**Fixed Coordinate**	
	C5-N9-HOH	79°	147°	**N9-H**OH	F(2.14) Å

Figure A6.6: Potential energy barriers between the most stable conformers of theobromine-water. a) ~37 kJ/mol (enol-keto) b) ~85 kJ/mol (keto-enol) c) ~85 kJ/mol (keto-enol) d) ~39 kJ/mol (enol-keto). All the calculations performed at M06-2X/6-311++G(d,p) level.

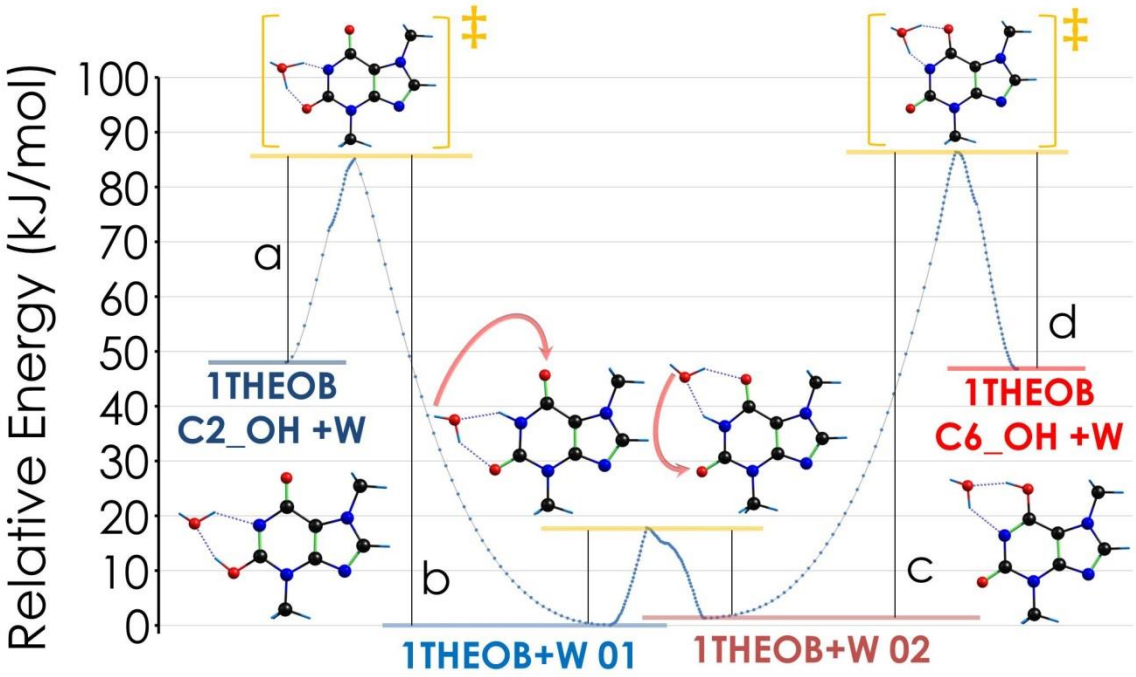


Table of the scanned coordinates:

NH to C2OH	**Distance H-O**	**Initial**	**Final**		
	N1**H-O**H	1.96 Å	0.98 Å		
	Distance H-O	**Initial**	**Final**	**Fixed Coordinate**	
	N1 ·H·(OH) ·**H-O**C2	1.06 Å	1.54 Å	N1**H-O**H	F(1.12) Å
NH to C6OH	**Distance H-O**	**initial**	**Final**		
	N1**H-O**H	1.97 Å	0.98 Å		
	Distance H-O	**Initial**	**Final**	**Fixed Coordinate**	
	N1 ·H·(OH) ·**H-O**C2	1.07 Å	1.55 Å	N1**H-O**H	F(1.13) Å

Figure A6.7: a) Relative Gibbs free energy diagram for theobromine; b) top black trace: experimental IDIRS spectrum of theobromine-water recorded ionizing at ~36095 cm^{-1}. Colored traces below: theoretical predictions for some of the computed structures (correcting factors: 0.939 for OH and 0.953 for NH and CH). All the calculation performed at M06-2X/6-311++G(d,p) level.

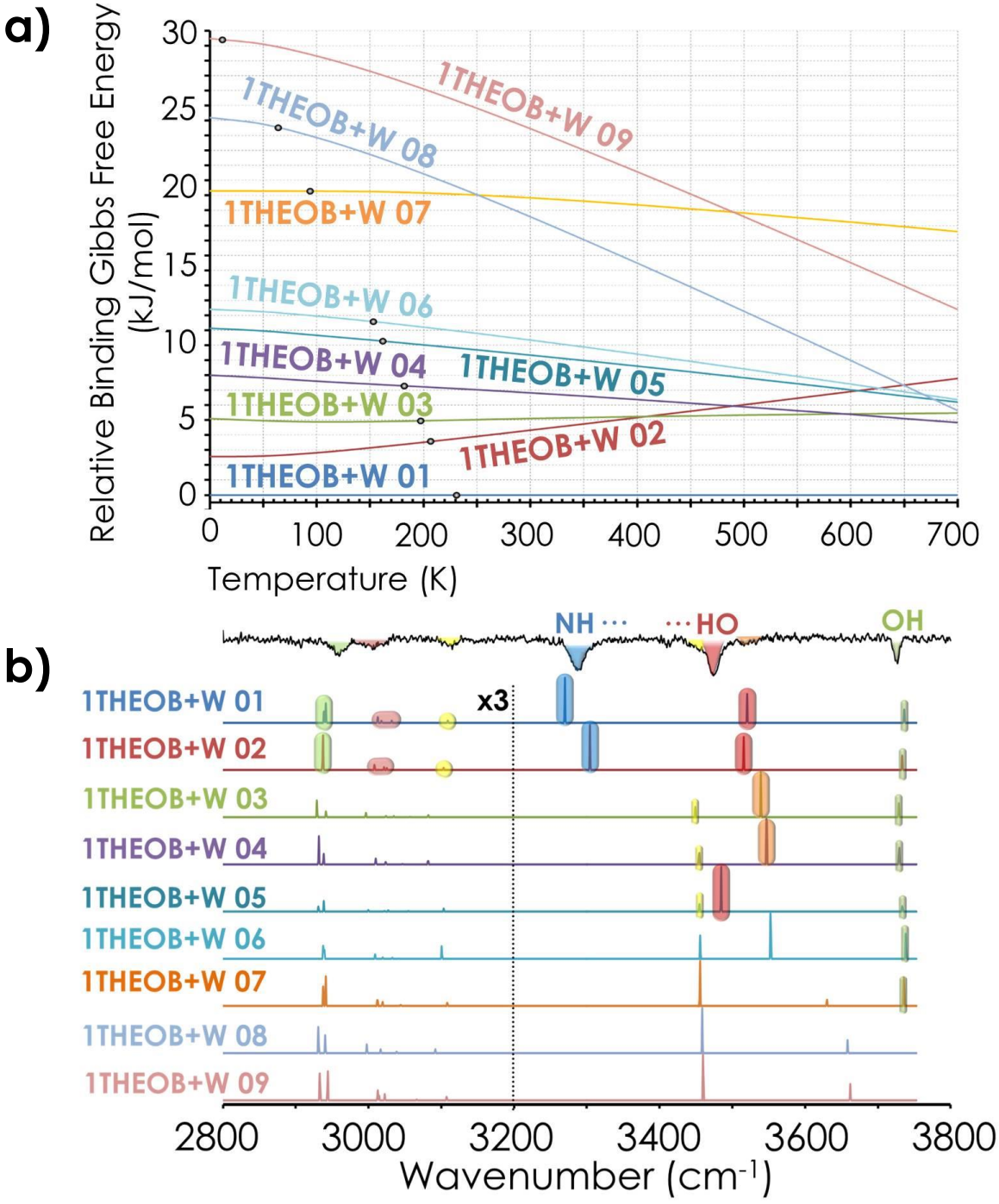


Figure A6.8: Structures and families for the theobromine dimer.

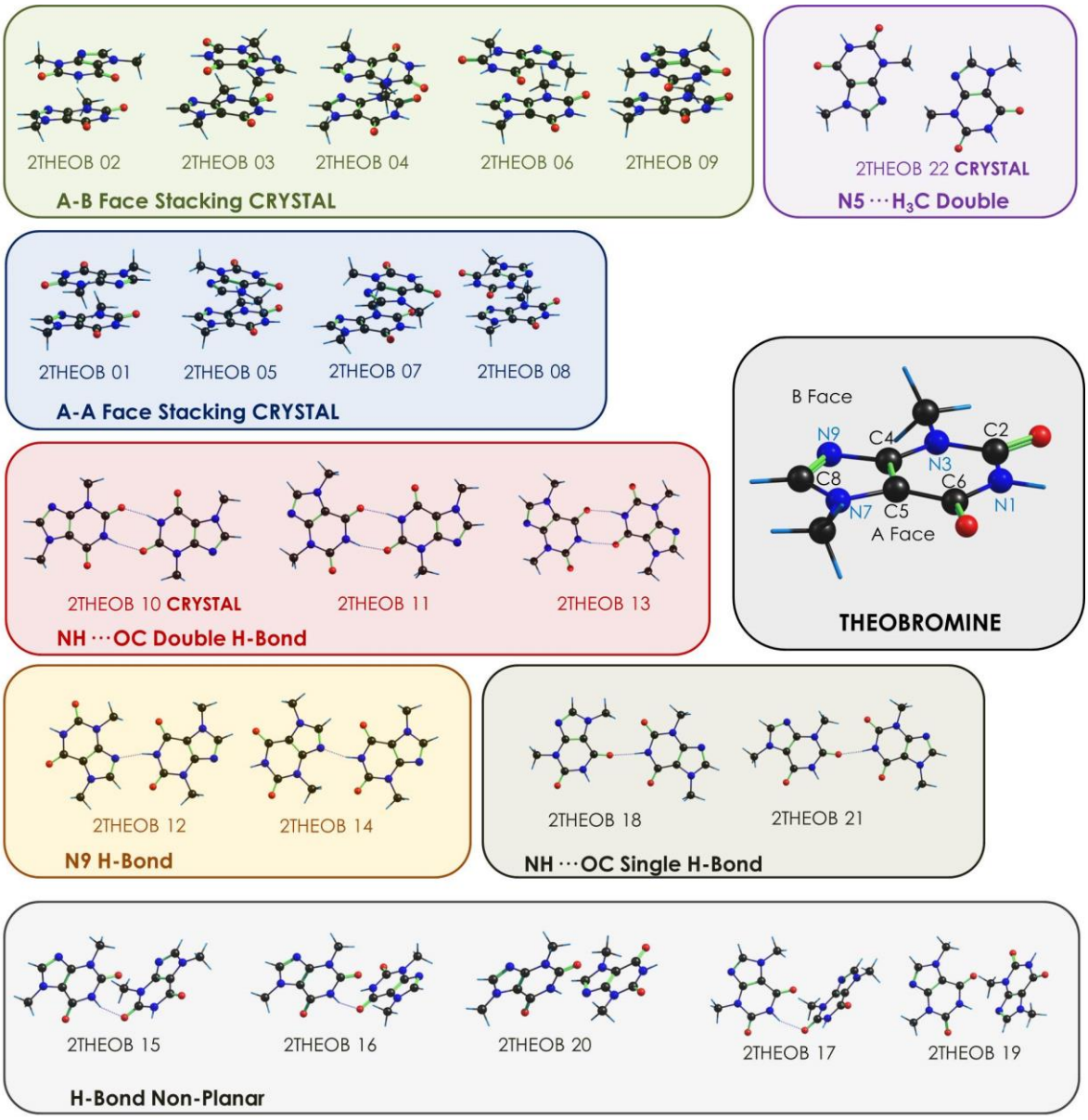


Table A6.1: Table of the energy and structural parameters relative of theobromine dimers. All the calculation done at the M06-2X/6-311++G(d,p) level.

LABEL	RE ZPE+BSSE (kJ/mol)	RE GIBBS (kJ/mol)	STACKING C4-C5-C´5-C´4	STACKING	H-BOND	Other Interactions
2THEOB_01	0.00	0.29	179.996	A-A	-	-
2THEOB_02	6.06	4.03	55.121	A-B	-	-
2THEOB_03	6.01	0.00	-39.42	A-B	-	-
2THEOB_04	9.12	7.38	177.798	A-B	-	-
2THEOB_05	8.55	4.04	64.913	A-A	-	-
2THEOB_06	10.39	2.50	17.756	A-B	-	-
2THEOB_07	10.58	8.46	70.383	A-A	-	-
2THEOB_08	13.99	11.70	-134.265	A-A	-	-
2THEOB_09	21.01	15.95	-115.401	A-B	-	-
2THEOB_10	16.29	1.32	-	-	N1H···OC2 + N1H···OC2	-
2THEOB_11	17.30	7.19	-	-	N1H···OC6 + N1H···OC2	-
2THEOB_12	19.30	10.99	-	-	N1H···N9	C8H···OC2
2THEOB_13	20.16	14.74	-	-	N1H···OC6 + N1H···OC6	-
2THEOB_14	20.01	10.11	-	-	N1H···N9	C8H···OC6
2THEOB_15	29.06	21.08	-	-	N1H···OC2	OC2···RING
2THEOB_16	32.59	23.98	-	-	N1H···OC2	OC2···RING
2THEOB_17	33.56	24.60	-	-	N1H···OC2	OC6···RING
2THEOB_18	31.42	18.75	-	-	N1H···OC6	$N7CH_3$···OC2
2THEOB_19	35.50	22.15	-	-	N1H···N9	OC6···RING
2THEOB_20	36.69	27.23	-	-	N1H···N9	OC2···RING
2THEOB_21	36.84	27.09	-	-	N1H···OC2	$N3CH_3$···OC2
2THEOB_22	48.82	28.75	-	-	-	$N3CH_3$···N5 + $N3CH_3$···N5

Figure A6.9: Tautomeric interconversion barriers for the most stable dimers of theobromine. The corresponding calculated barriers are: **Green)** ~7 kJ/mol (enol-keto); ~79 kJ/mol (keto-enol) **Red)** ~6 kJ/mol (enol-keto); ~76 kJ/mol (keto-enol) **Blue)** ~7 kJ/mol (enol-keto); ~73 kJ/mol (keto-enol). All the calculations performed at the M06-2X/6-311++G(d,p) level.

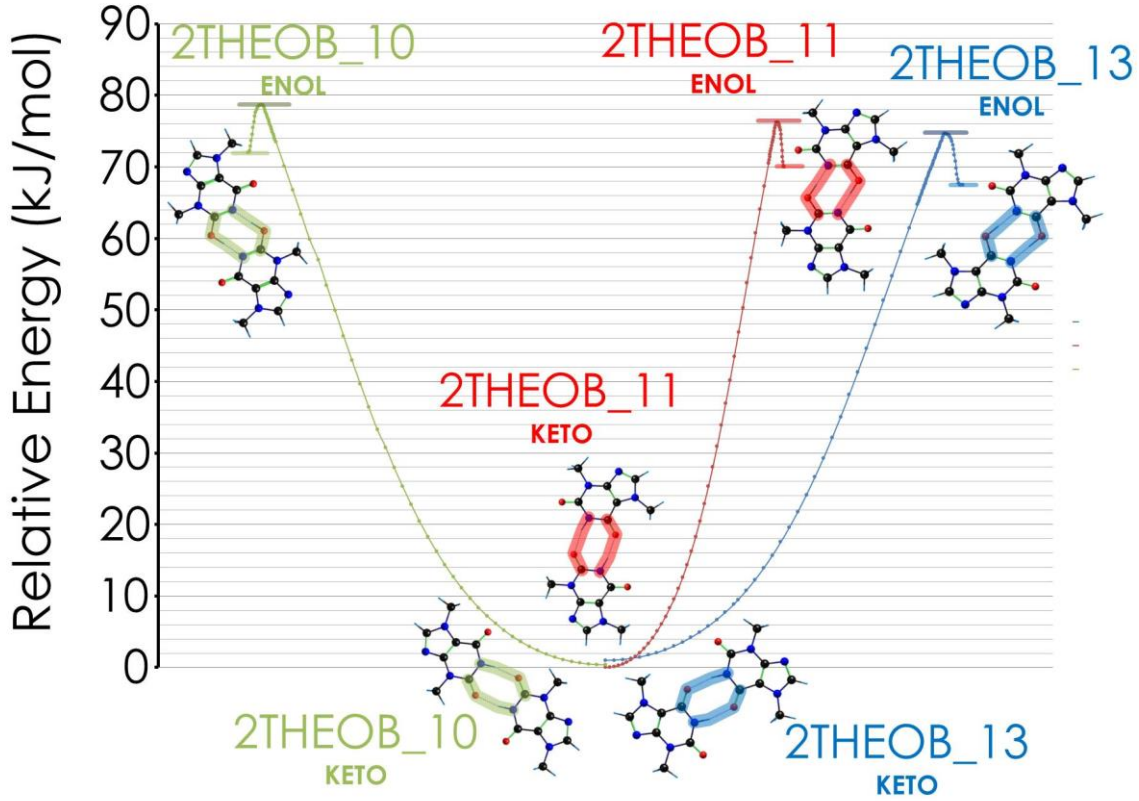


Table of the scanned coordinates:

2THOB_10	**Distance H-O´**	**Initial**	**Final**		
KETO-ENOL	N1**H-O´**C´2	1.85 Å	1.04 Å		
	Distance H-O´	**Initial**	**Final**	**Fixed Coordinate**	
	N´1**H´-O**C2	1.39 Å	1.05 Å	N1**H-O´**C´2	F(1.04 Å)
2THOB_11	**Distance H-O´**	**Initial**	**Final**		
KETO-ENOL	N1**H-O´**C´2	1.84 Å	1.05 Å		
	Distance H-O´	**Initial**	**Final**	**Fixed Coordinate**	
	N´1**H´-O**C2	1.39 Å	1.05 Å	N1**H-O´**C´2	F(1.05 Å)
2THOB_13	**Distance H-O´**	**Initial**	**Final**		
KETO-ENOL	N1**H-O´**C´2	1.87 Å	1.06 Å		
	Distance H-O´	**Initial**	**Final**	**Fixed Coordinate**	
	N´1**H´-O**C2	1.40 Å	1.04 Å	N1**H-O´**C´2	F(1.06 Å)

Figure A6.10: a) Relative binding Gibbs free energy for all the dimer conformational families showed in Figure A6.8; b) Relative binding Gibbs free energy of the most stable conformations.

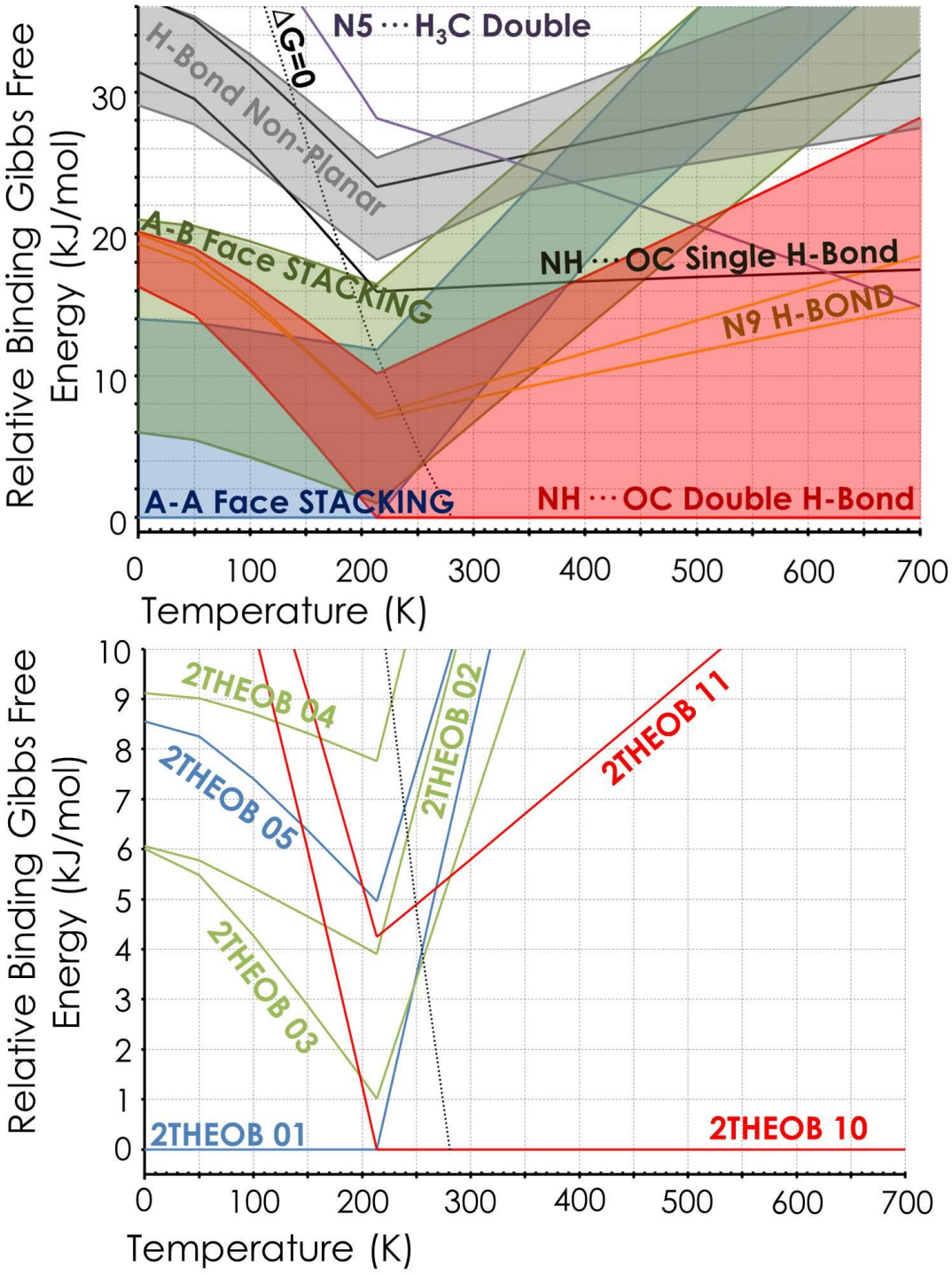


Figure A6.11: Upper black trace: experimental IDIRS spectrum of theobromine dimer. Rest of the traces: theoretical predictions for some computed structures (correction factors:

0.939 for OH and 0.953 for NH and CH). All the calculations performed at M06-2X/6-311++G(d,p) level.

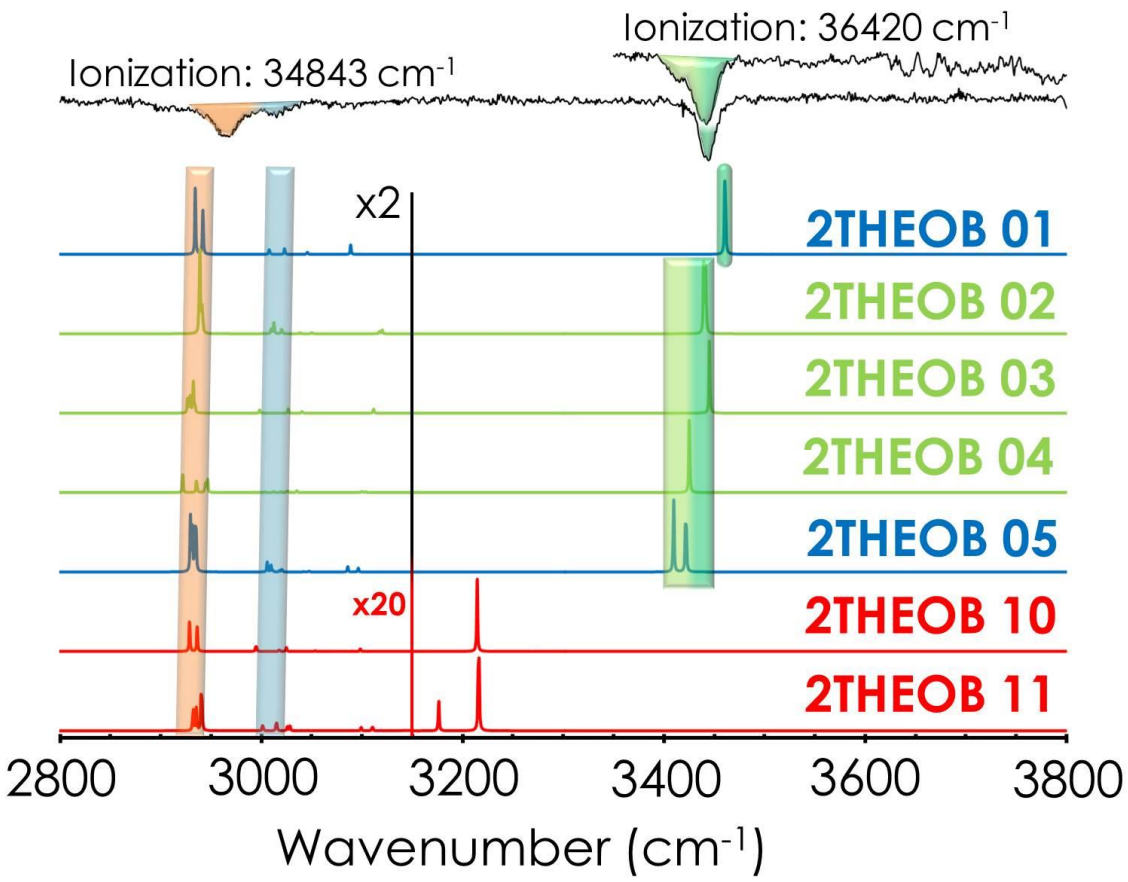
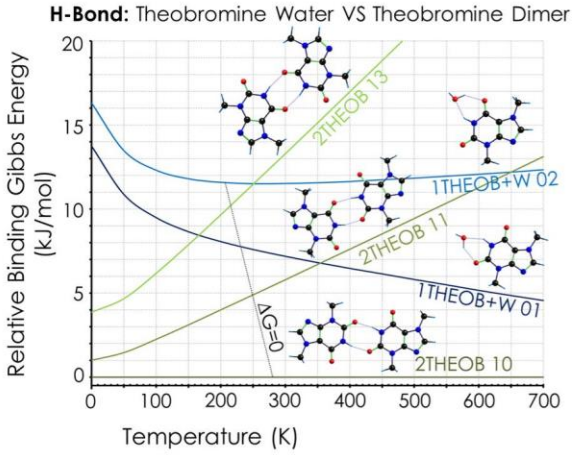
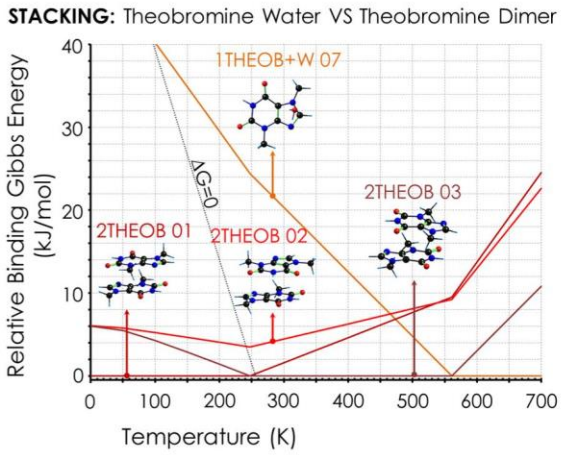


Figure A6.12: Binding Gibbs free energy diagrams comparing the Theobromine-water complex with the Theobromine dimer.


H-Bond: Theobromine Water VS Theobromine Dimer
Relative Binding Gibbs Energy (kJ/mol)
Temperature (K)
2THEOB 13
1THEOB+W 02
2THEOB 11
1THEOB+W 01
ΔG=0
2THEOB 10


STACKING: Theobromine Water VS Theobromine Dimer
Relative Binding Gibbs Energy (kJ/mol)
Temperature (K)
1THEOB+W 07
ΔG=0
2THEOB 03
2THEOB 01
2THEOB 02

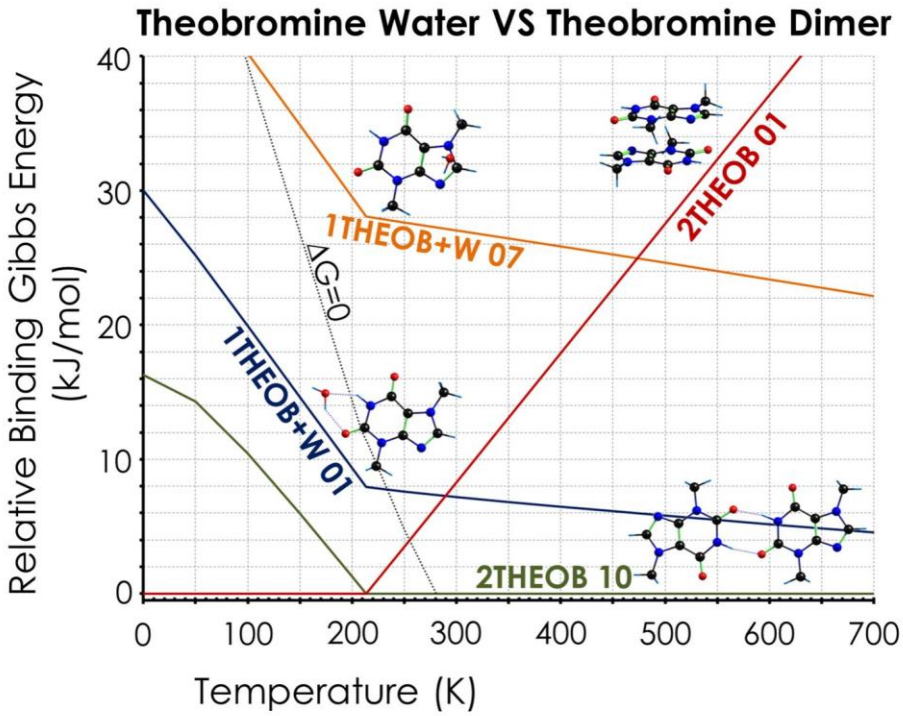

Theobromine Water VS Theobromine Dimer
Relative Binding Gibbs Energy (kJ/mol)
Temperature (K)
1THEOB+W 07
2THEOB 01
ΔG=0
1THEOB+W 01
2THEOB 10

Table A6.2: Relative stability of theobromine trimer conformers. All the calculations performed at M06-2X/6-311++G(d,p) level.

LABEL	RE ZPE+BSSE (kJ/mol)	RE GIBBS (kJ/mol)
3THEOB_SHB2_01	0.00	0.00
3THEOB_2S2H_02	5.73	6.81
3THEOB_2S2H_07	11.66	9.17
3THEOB_SHB3_10	4.56	9.21
3THEOB_SHB3_11	5.65	10.02
3THEOB_SHB3_09	4.20	10.39
3THEOB_SHB2_04	3.97	11.05
3THEOB_SHB2_06	4.67	11.15
3THEOB_SHB3_12	4.77	11.23
3THEOB_2S2H_16	10.66	12.05
3THEOB_2S2H_14	8.29	12.38
3THEOB_SHB2_07	9.93	12.52
3THEOB_2S2H_03	8.93	12.99
3THEOB_SHB3_13	6.76	13.24
3THEOB_2S2H_06	11.36	13.77
3THEOB_2S2H_17	9.97	13.92
3THEOB_2S2H_01	5.46	14.15
3THEOB_SHB3_01	8.08	14.39
3THEOB_2S2H_05	9.45	14.90
3THEOB_SHB3_05	11.30	15.12
3THEOB_SHB3_14	8.78	15.97
3THEOB_SHB3_03	9.90	16.10
3THEOB_SHB3_04	10.41	16.36
3THEOB_STACKING_01	11.69	16.59
3THEOB_2S2H_15	10.23	16.84
3THEOB_2S2H_08	13.86	17.23
3THEOB_2S2H_04	11.36	17.55
3THEOB_SHB3_02	9.99	17.57
3THEOB_2S2H_09	13.88	18.51
3THEOB_2S2H_21	13.55	18.61
3THEOB_2S2H_10	15.29	19.11
3THEOB_SHB3_15	12.81	19.18
3THEOB_SHB2_02	12.03	19.34
3THEOB_2S2H_22	14.77	19.80
3THEOB_SHB2_09	18.70	20.42
3THEOB_2S2H_24	17.00	20.86
3THEOB_SHB3_06	13.52	21.06
3THEOB_2S2H_23	16.47	22.00
3THEOB_2S2H_20	13.68	22.18
3THEOB_SHB2_05	18.40	22.38
3THEOB_2S2H_19	15.09	22.73
3THEOB_2S2H_18	12.12	22.92
3THEOB_2S2H_25	12.18	23.04
3THEOB_SHB2_08	18.51	23.39
3THEOB_SHB2_03	21.33	23.51
3THEOB_2S2H_26	19.97	23.92
3THEOB_STACKING_02	19.31	24.06
3THEOB_2S2H_11	17.61	24.06
3THEOB_SHB3_16	22.39	25.44
3THEOB_2S2H_12	20.50	25.99
3THEOB_SHB3_17	23.29	26.22
3THEOB_2S2H_27	22.38	27.10
3THEOB_SHB2_11	21.47	28.25
3THEOB_HBR_01	28.25	28.88
3THEOB_2S2H_13	25.14	29.42
3THEOB_HBR_02	31.74	29.73
3THEOB_SHB2_10	22.34	31.57
3THEOB_SHB2_12	25.74	32.61
3THEOB_SHB2_14	30.36	32.71
3THEOB_STACKING_03	31.83	33.25
3THEOB_SHB2_16	35.43	35.97
3THEOB_SHB2_13	31.96	37.37
3THEOB_SHB3_08	34.08	37.53
3THEOB_SHB3_07	33.36	38.14
3THEOB_HBR_04SS	39.45	38.39
3THEOB_SHB2_15	34.11	39.28
3THEOB_HBR_03	39.47	40.31
3THEOB_HBR_04	44.31	41.32
3THEOB_SHB3_18	39.93	41.67
3THEOB_HBR_05	46.94	41.72
3THEOB_CRYSTAL	69.76	65.43

Figure A6.13: a) Relative binding Gibbs free energy for all the families of Theobromine trimer; b) Relative binding Gibbs free energy of the most stable conformations.

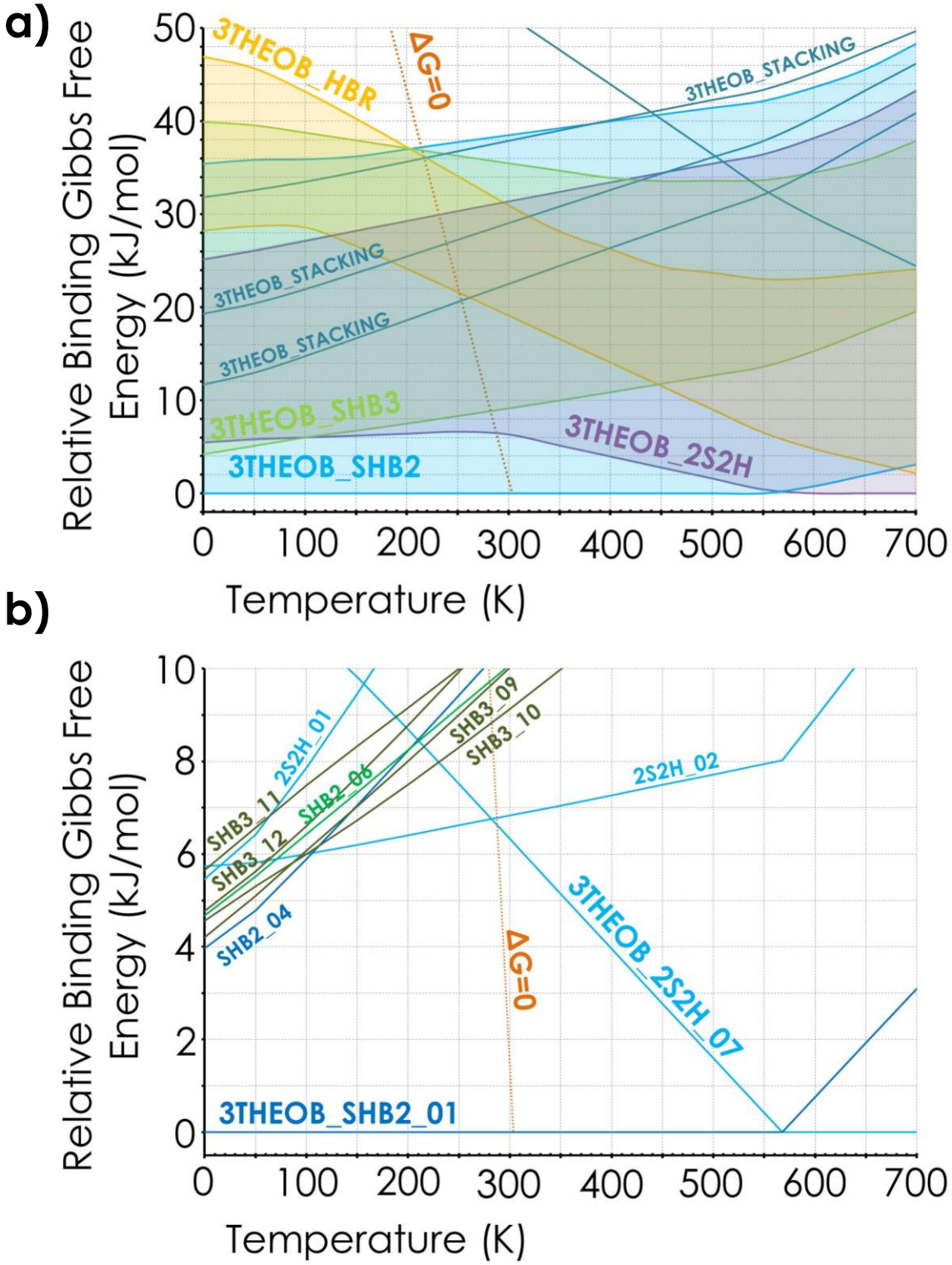
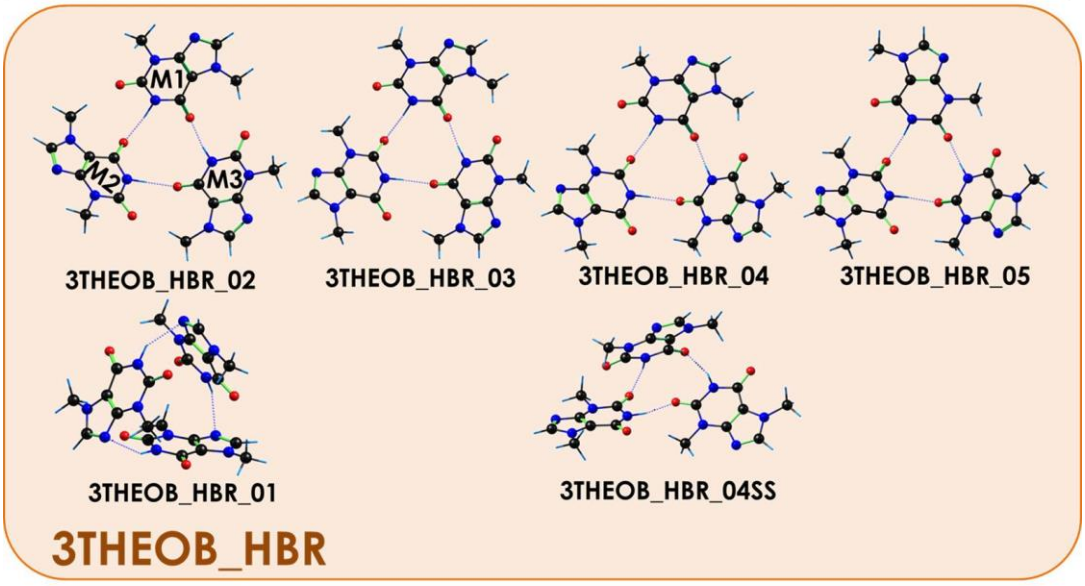


Figure A6.14: 3THEOB_HBR family: a) Energy and H-bond table; b) structures; c) relative binding Gibbs free energy diagram.

LABEL	RE ZPE+BSSE (kJ/mol)	RE GIBBS (kJ/mol)	H-BOMD (M1-M2)	H-BOMD (M2-M3)	H-BOMD (M3-M1)	Extra infor.
3THEOB_HBR_01	0.00	0.00	N1-H···N9	N1-H···N9	N1-H···N9	
3THEOB_HBR_02	3.48	0.85	N1-H···O=C6	N1-H···O=C6	N1-H···O=C6	
3THEOB_HBR_03	11.22	11.44	N1-H···O=C2	N1-H···O=C6	N1-H···O=C6	
3THEOB_HBR_04	16.06	12.44	N1-H···O=C2	N1-H···O=C2	N1-H···O=C6	
3THEOB_HBR_04SS	11.20	9.52	N1-H···O=C2	N1-H···O=C2	N1-H···O=C6	Semi-Stacked
3THEOB_HBR_05	18.68	12.85	N1-H···O=C2	N1-H···O=C2	N1-H···O=C2	



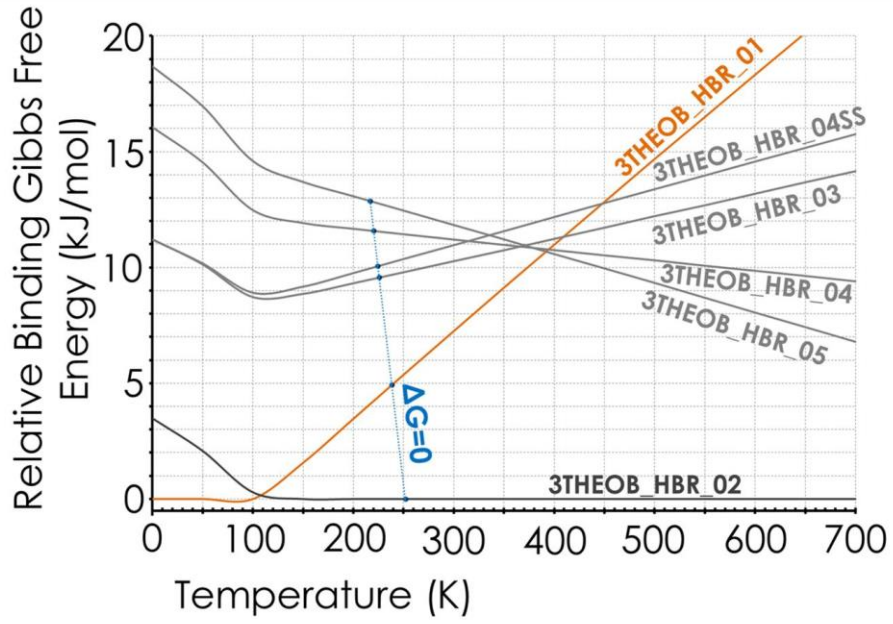
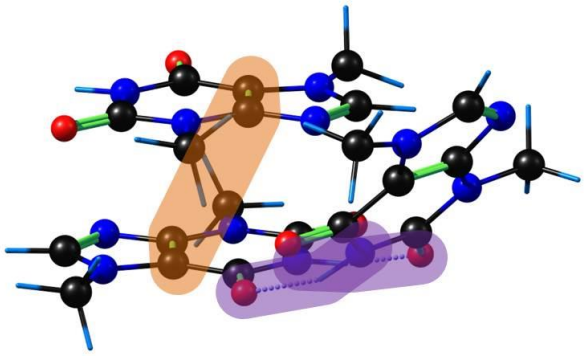


Table A6.3: 3THEOB_2S2H family: energy, H-bonds and structural parameters.

LABEL	RE ZPE+BSSE (kJ/mol)	RE GIBBS (kJ/mol)	STACKING C4-C5-C´5-C´4	STACKING	H-BOND Molecule 1	H-BOND Molecule 1
3THEOB_2S2H_01	0.00	7.35	**-179**	**A-A**	**C2=O···HN1**	**N1H···O=C6**
3THEOB_2S2H_02	0.27	0.00	-178	A-A	C6=O···HN1	N1H···O=C6
3THEOB_2S2H_03	3.47	6.18	178	A-A	C6=O···HN1	N1H···O=C2
3THEOB_2S2H_04	5.90	10.74	-125	A-A	C6=O···HN1	N1H···O=C2
3THEOB_2S2H_05	3.99	8.09	179	A-A	C2=O···HN1	N1H···O=C2
3THEOB_2S2H_06	5.90	6.97	180	A-A	N9···HN1	
3THEOB_2S2H_07	6.20	2.36	173	A-A	C2=O···HN1	N1H···O=C2
3THEOB_2S2H_08	8.40	10.43	-154	A-A	C2=O···HN1	N1H···O=C6
3THEOB_2S2H_09	8.42	11.70	153	A-A	C2=O···HN1	N1H···O=C6
3THEOB_2S2H_10	9.83	12.31	56	A-A	C2=O···HN1	N1H···O=C2
3THEOB_2S2H_11	12.15	17.26	63	A-A	C6=O···HN1	N1H···O=C2
3THEOB_2S2H_12	15.04	19.19	177	A-A	C2=O···HN1	
3THEOB_2S2H_13	19.68	22.62	-175	A-A	N9···HN1	
3THEOB_2S2H_14	2.82	5.57	51	A-B	C2=O···HN1	N1H···O=C2
3THEOB_2S2H_15	4.77	10.03	65	A-B	C9=O···HN1	
3THEOB_2S2H_16	5.20	5.25	41	A-B	C6=O···HN1	N1H···O=C2
3THEOB_2S2H_17	4.51	7.12	56	A-B	C6=O···HN1	N1H···O=C6
3THEOB_2S2H_18	6.66	16.12	-41	A-B	C6=O···HN1	N1H···O=C2
3THEOB_2S2H_19	9.63	15.92	175	A-B	N9···HN1	
3THEOB_2S2H_20	8.22	15.37	-175	A-B	C2=O···HN1	N1H···O=C6
3THEOB_2S2H_21	8.09	11.81	56	A-B	C2=O···HN1	N1H···O=C6
3THEOB_2S2H_22	9.31	12.99	-69	A-B	C2=O···HN1	N1H···O=C2
3THEOB_2S2H_23	11.01	15.19	67	A-B	C6=O···HN1	N1H···O=C2
3THEOB_2S2H_24	11.54	14.05	-179	A-B	C6=O···HN1	N1H···O=C6
3THEOB_2S2H_25	6.72	16.23	-41	A-B A-A	C6=O···HN1	N1H···O=C2
3THEOB_2S2H_26	14.50	17.11	-144	A-B A-A	C2=O···HN1	N1H···O=C6
3THEOB_2S2H_27	16.92	20.30	-150	A-B BOND	C2=O···HN1	N1H···O=C2



3THEOB_2S2H_01

Figure A6.15: a) Relative binding Gibbs free energy for all the conformers of 3THEOB_2S2H family; b) Relative binding Gibbs free energy of the most stable conformations.

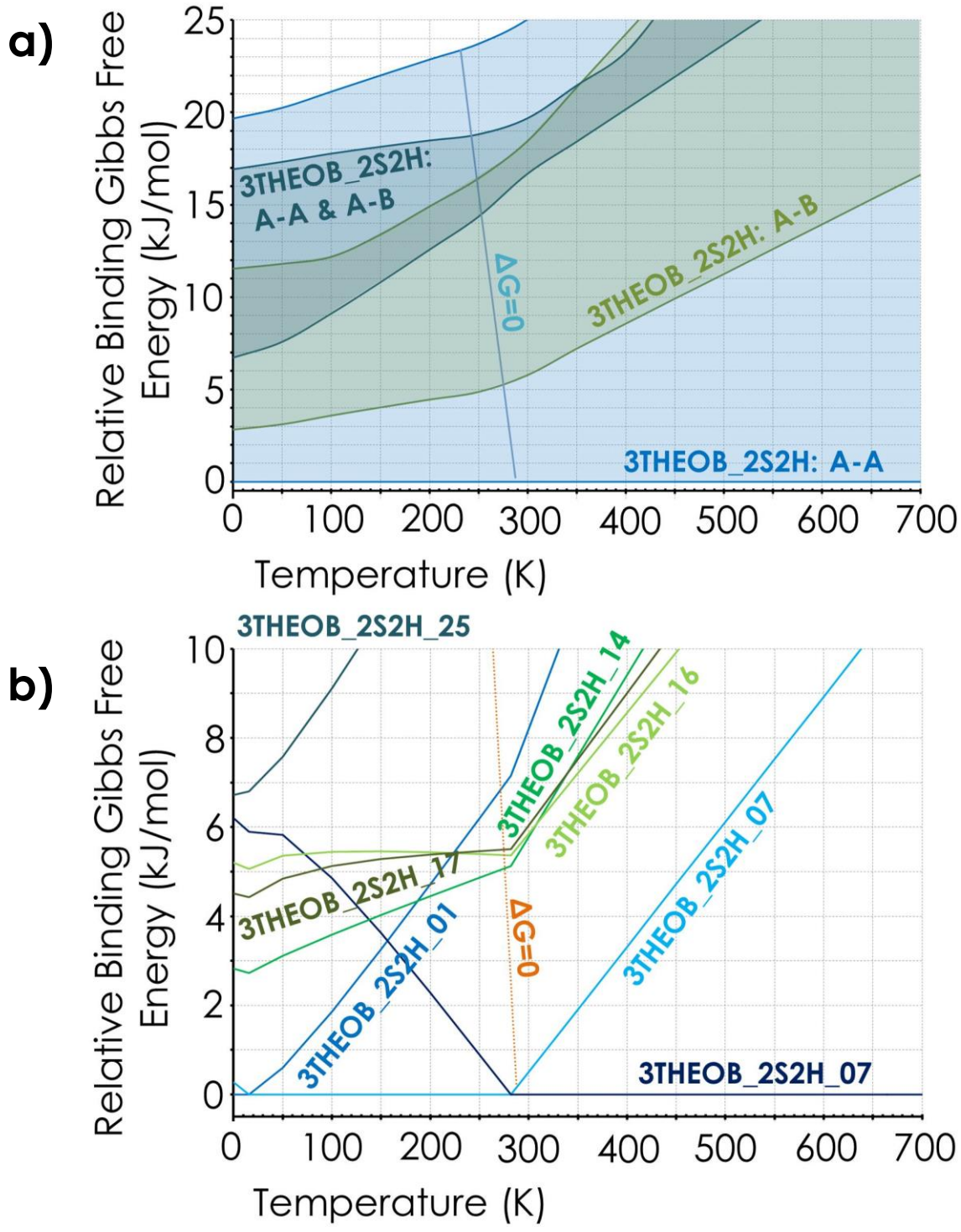


Figure A6.16: Structures of 3THEOB_2S2H family. Below each picture, the names of the corresponding precursor dimers are reported.

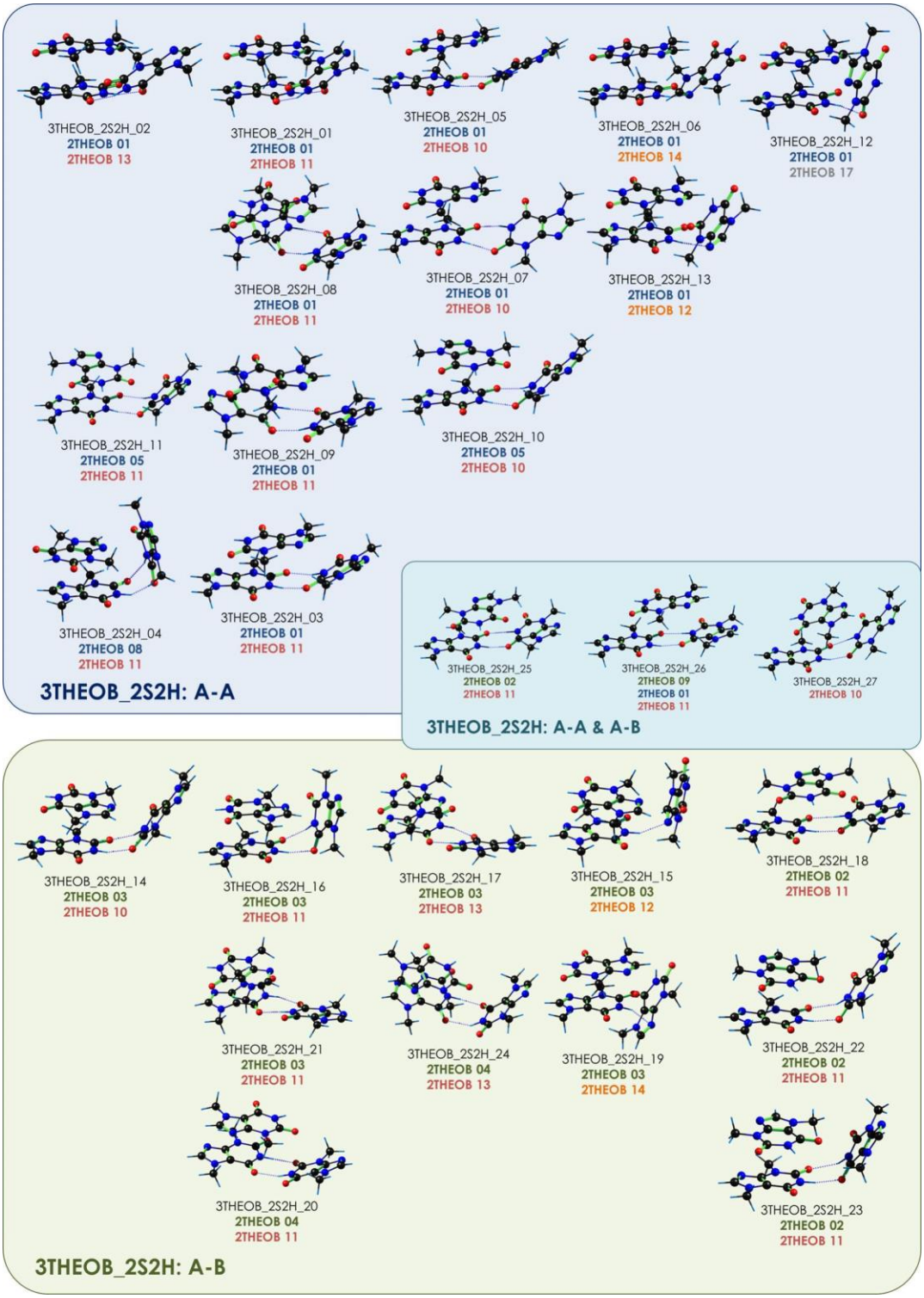
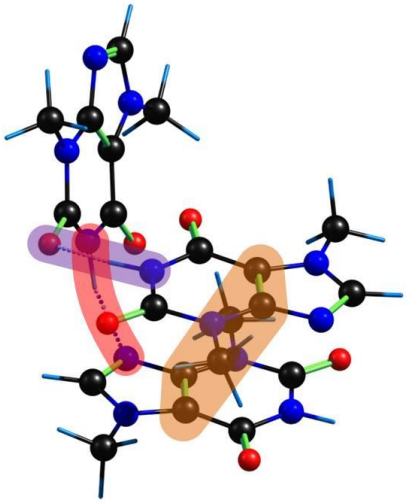


Table A6.4: 3THEOB_SHB2 family: energy, H-bonds and structural parameters.

LABEL	RE ZPE+BSSE (kJ/mol)	RE GIBBS (kJ/mol)	STACKING C4-C5-C´5-C´4	STACKING	H-BOND MOLECULE 1°	H-BOND MOLECULE 2°
3THEOB_SHB2_01	0.00	0.00	**-177.171**	**A-A**	**N1H···N9**	**C2=O···HN1**
3THEOB_SHB2_02	12.03	19.34	-111.581	A-A	N1H···N9	C2=O···HN1
3THEOB_SHB2_03	21.33	23.51	44.907	A-A	N1H···O=C2	C2=O···HN1
3THEOB_SHB2_04	3.97	11.05	178.137	A-A	N1H···N9	C6=O···HN1
3THEOB_SHB2_05	18.40	22.38	47.725	A-B	N1H···O=C6	C6=O···HN1
3THEOB_SHB2_06	4.67	11.15	38.641	A-B	N1H···N9	C2=O···HN1
3THEOB_SHB2_07	9.93	12.52	39.301	A-B	N1H···N9	C6=O···HN1
3THEOB_SHB2_08	18.51	23.39	-68.258	A-B	N1H···O=C6	C2=O···HN1
3THEOB_SHB2_09	18.70	20.42	-67.355	A-B	N1H···O=C6	C2=O···HN1
3THEOB_SHB2_10	22.34	31.57	177.042	A-B	N1H···O=C2	C2=O···HN1
3THEOB_SHB2_11	21.47	28.25	179.985	A-B	N1H···O=C2	C6=O···HN1
3THEOB_SHB2_12	25.74	32.61	-178.264	A-B	N1H···O=C2	C6=O···HN1
3THEOB_SHB2_13	31.96	37.37	-118.891	A-B	N9···H1N	C2=O···HN1
3THEOB_SHB2_14	30.36	32.71	-161.464	A-B	N1H···O=C2	C2=O···HN1
3THEOB_SHB2_15	34.11	39.28	-120.824	A-B	N9···H1N	C2=O···HN1
3THEOB_SHB2_16	35.43	35.97	-177.389	A-B	N9···H1N	C6=O···HN1



3THEOB_SHB2_01

Figure A6.17: a) Relative binding Gibbs free energy for all the families of 3THEOB_SHB2 conformers; b) relative binding Gibbs free energy of the most stable conformations.

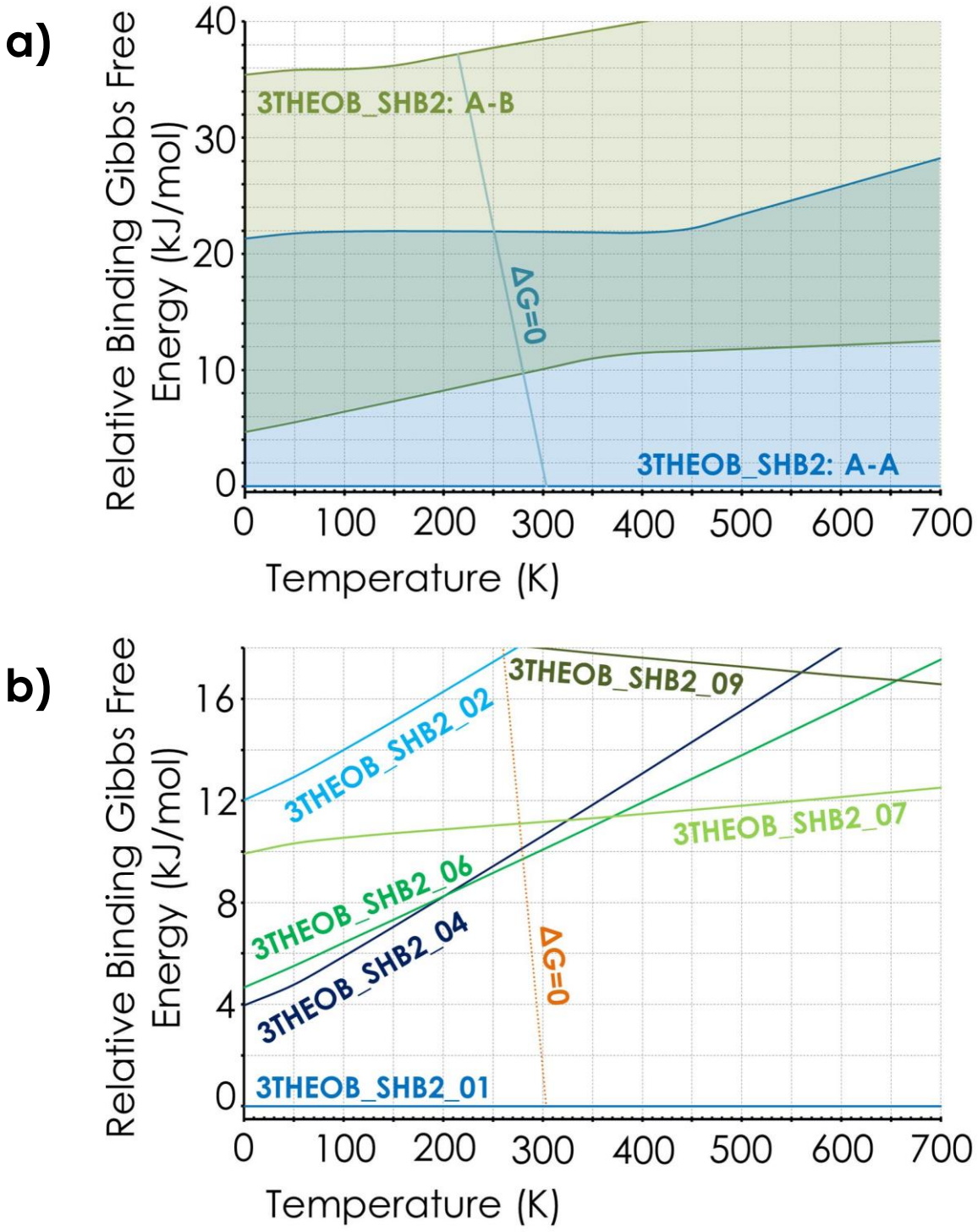
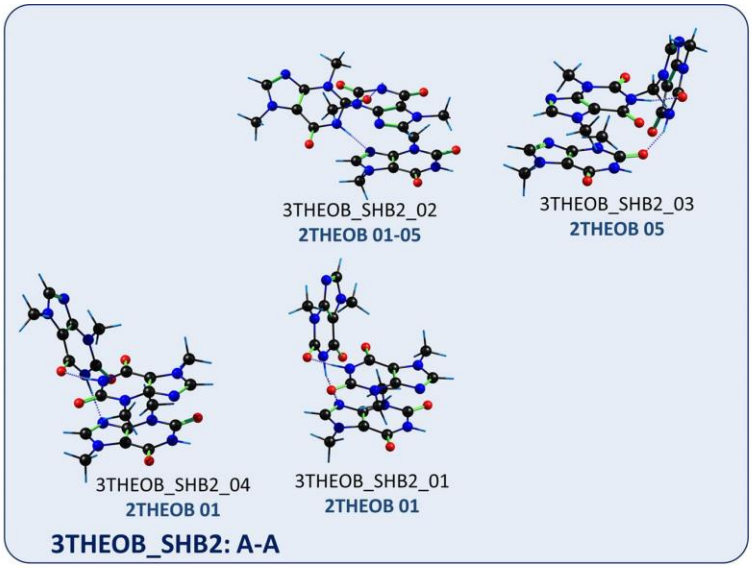


Figure A6.18: Structures of 3THEOB_SHB2 family. Below each picture, the names of the corresponding precursor dimers are reported.


3THEOB_SHB2_02
2THEOB 01-05
3THEOB_SHB2_03
2THEOB 05
3THEOB_SHB2_04
2THEOB 01
3THEOB_SHB2_01
2THEOB 01
3THEOB_SHB2: A-A

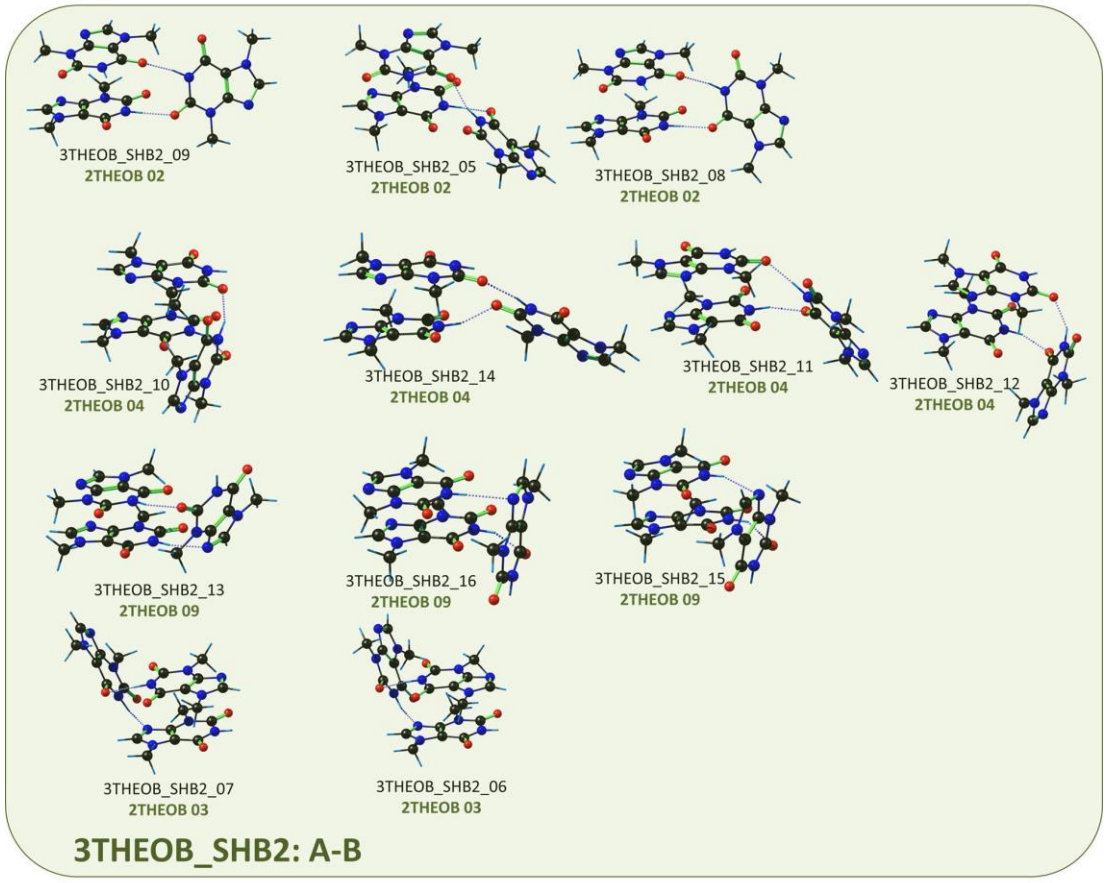
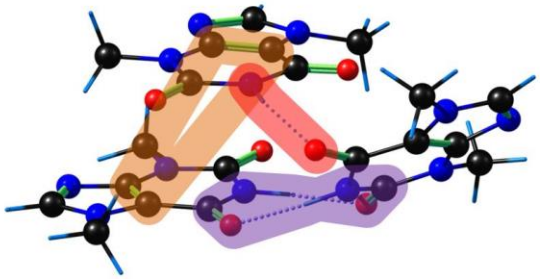

3THEOB_SHB2_09
2THEOB 02
3THEOB_SHB2_05
2THEOB 02
3THEOB_SHB2_08
2THEOB 02
3THEOB_SHB2_10
2THEOB 04
3THEOB_SHB2_14
2THEOB 04
3THEOB_SHB2_11
2THEOB 04
3THEOB_SHB2_12
2THEOB 04
3THEOB_SHB2_13
2THEOB 09
3THEOB_SHB2_16
2THEOB 09
3THEOB_SHB2_15
2THEOB 09
3THEOB_SHB2_07
2THEOB 03
3THEOB_SHB2_06
2THEOB 03
3THEOB_SHB2: A-B

Table A6.5: 3THEOB_SHB3 family: energy, H-bonds and structural parameters.

LABEL	RE ZPE+BSSE (kJ/mol)	RE GIBBS (kJ/mol)	STACKING C4-C5-C´5-C´4	STACKING	H-BOND MOLECULE 1°	H-BOND MOLECULE 1°	H-BOND MOLECULE 2°
3THEOB_SHB3_01	3.88	5.18	-70.527	A-A	N1H···O=C2	C2=O···HN1	C6=O···HN1
3THEOB_SHB3_02	5.79	8.35	-57.961	A-A	N1H···O=C6	C2=O···HN1	C6=O···HN1
3THEOB_SHB3_03	5.70	6.89	48.604	A-A	N1H···O=C6	C6=O···HN1	C2=O···HN1
3THEOB_SHB3_04	6.21	7.15	-51.261	A-A	N1H···O=C6	C6=O···HN1	C2=O···HN1
3THEOB_SHB3_05	7.10	5.90	49.377	A-A	N1H···O=C2	C6=O···HN1	C2=O···HN1
3THEOB_SHB3_06	9.32	11.85	61.711	A-A	N1H···O=C2	C6=O···HN1	C2=O···HN1
3THEOB_SHB3_07	29.16	28.93	15.479	A-A	N1H···O=C6	C6=O···HN1	C2=O···HN1
3THEOB_SHB3_08	29.88	28.32	-15.74	A-A	N1H···O=C2	C2=O···HN1	C6=O···HN1
3THEOB_SHB3_09	0.00	1.18	**-49.131**	**A-B**	**N1H···O=C6**	**C2=O···HN1**	**C6=O···HN1**
3THEOB_SHB3_10	0.36	0.00	-57.779	A-B	N1H···O=C6	C6=O···HN1	C2=O···HN1
3THEOB_SHB3_11	1.45	0.81	175.688	A-B	N1H···O=C2	C2=O···HN1	C6=O···HN1
3THEOB_SHB3_12	0.57	2.01	57.588	A-B	N1H···O=C6	C6=O···HN1	C6=O···HN1
3THEOB_SHB3_13	2.56	4.02	-55.272	A-B	N1H···O=C6	C2=O···HN1	C6=O···HN1
3THEOB_SHB3_14	4.57	6.76	177.724	A-B	N1H···O=C2	C6=O···HN1	C6=O···HN1
3THEOB_SHB3_15	8.61	9.97	177.882	A-B	N1H···O=C2	C2=O···HN1	C6=O···HN1
3THEOB_SHB3_16	18.19	16.22	104.99	A-B	N1H···O=C6	C2=O···HN1	C6=O···HN1
3THEOB_SHB3_17	19.09	17.00	-104.591	A-B	N1H···O=C6	C6=O···HN1	C2=O···HN1
3THEOB_SHB3_18	35.72	32.46	-115.583	A-B	N1H···O=C2	C6=O···HN1	C2=O···HN1



3THEOB_SHB3_09

Figure A6.19: a) Relative binding Gibbs free energy for all the conformers of 3THEOB_SHB3 family; b) relative binding Gibbs free energy of the most stable conformations.

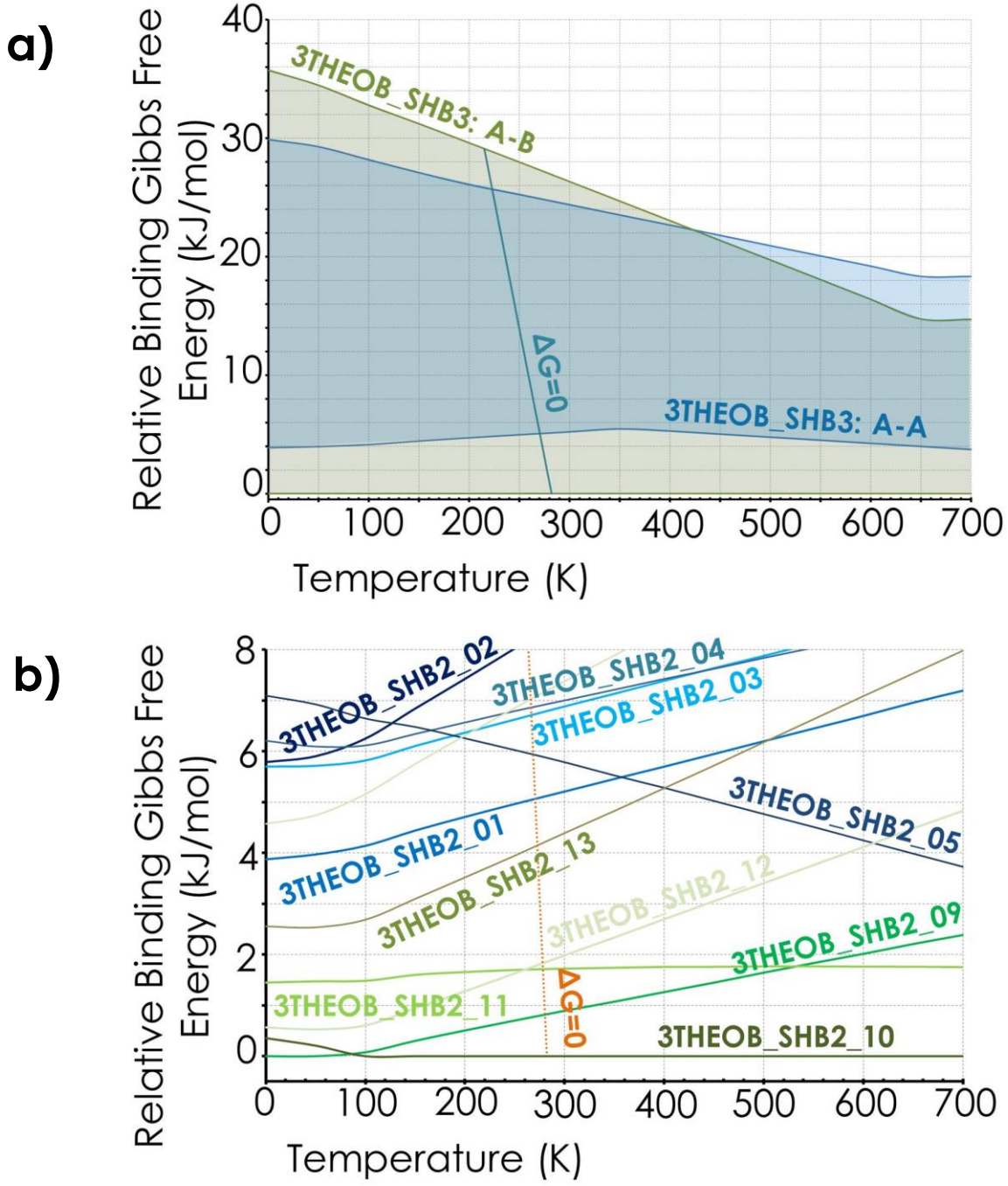


Figure A6.20: Structures of 3THEOB_SHB3 family. Below each picture, the names of the corresponding precursor dimers are reported.

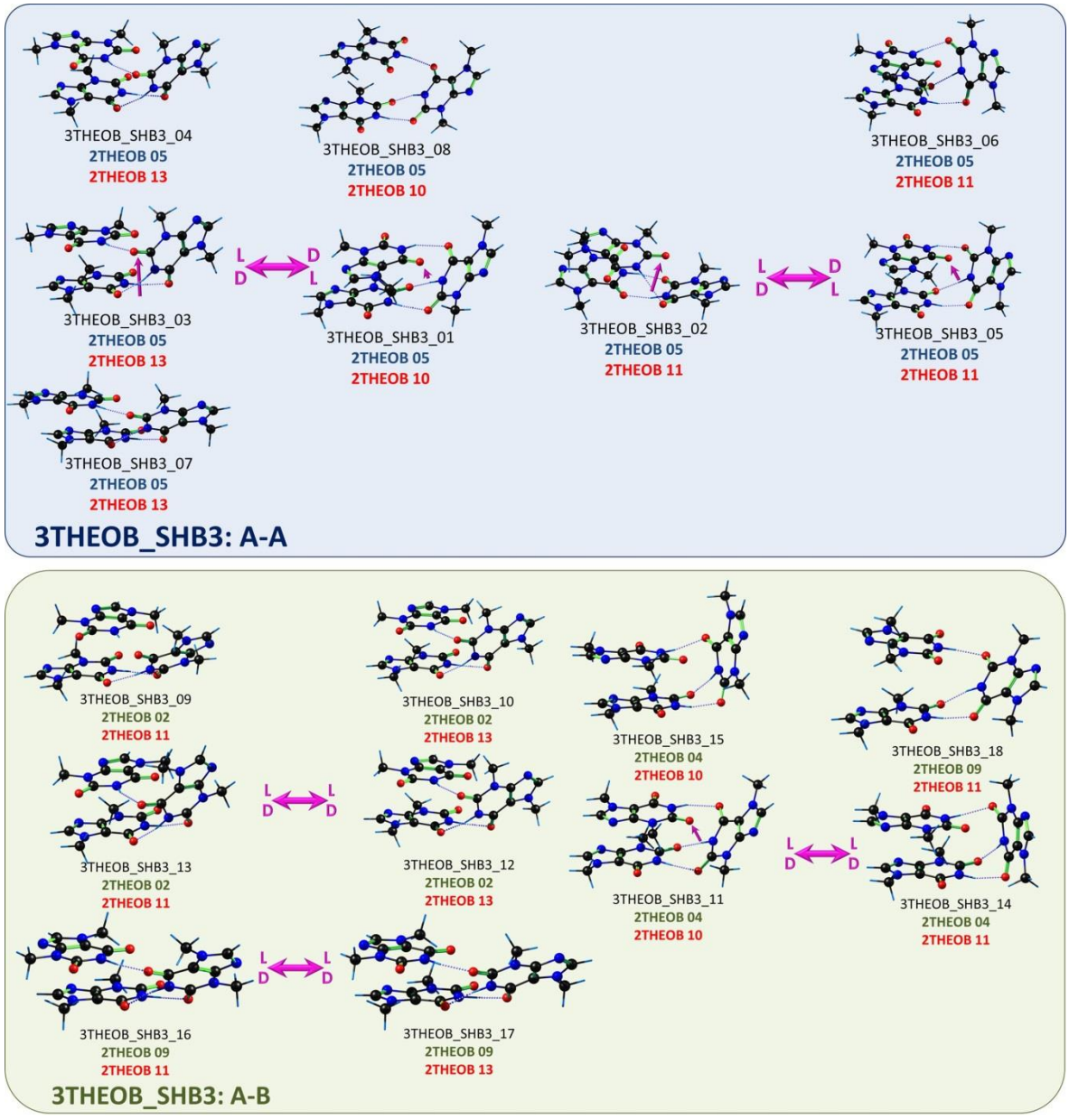
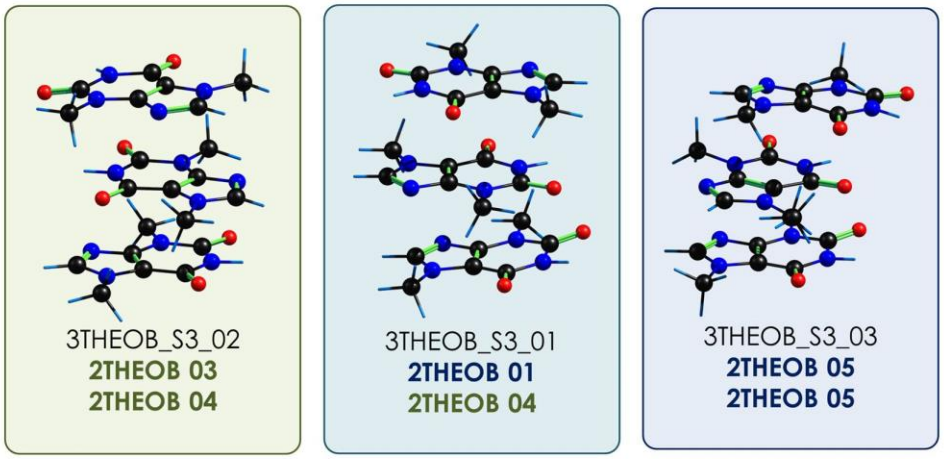


Figure A6.21: 3THEOB_S3 family: a) Energy and H-bonds data; b) structures; c) relative binding Gibbs free energy diagram.

LABEL	RE ZPE+BSSE (kJ/mol)	RE GIBBS (kJ/mol)	STACKING	STACKING M1 to M2 C4-C5-C´5-C´4	STACKING	STACKING M2 to M3 C´´4-C´´5-C´5-C´4
3THEOB_CRYSTAL	58.08	48.84	-	-	-	-
3THEOB_STACKING_01	0.00	0.00	A-A	-179.82	A-B	-171.934
3THEOB_STACKING_02	7.62	7.47	A-B	-50.461	A-B	176.691
3THEOB_STACKING_03	20.15	16.67	A-A	62.84	A-A	64.001



3THEOB_3S: A-B-A & A-A-B & A-A-A

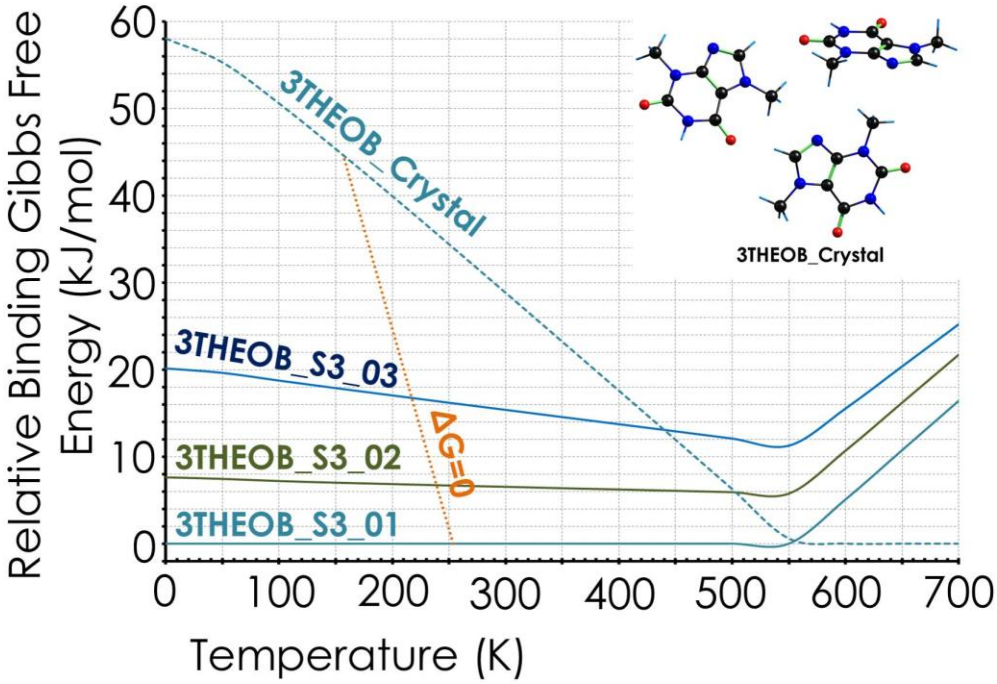
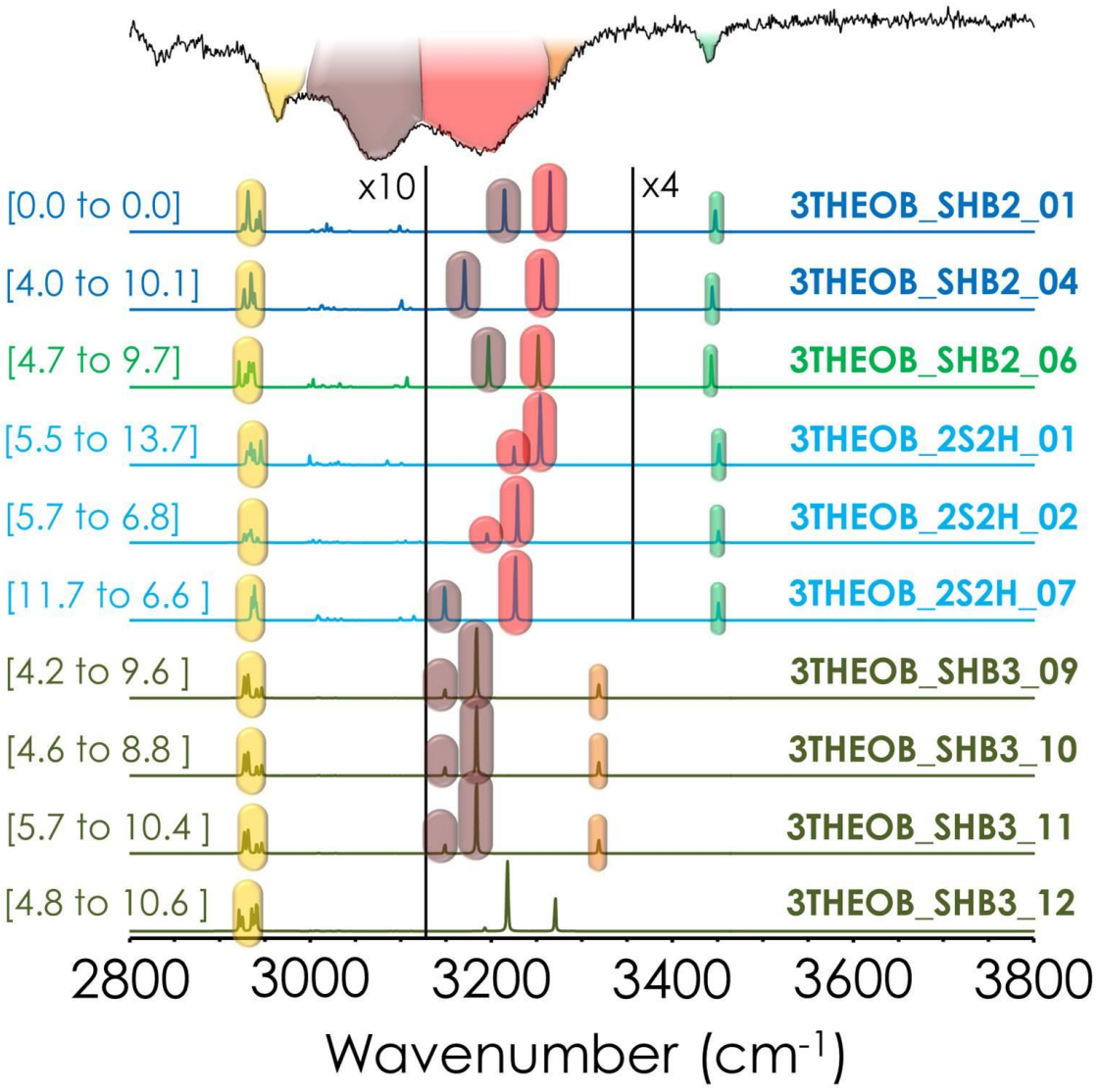


Figure A6.22: Above in black is reported the experimental IDIRS spectrum of Theobromine trimer, and below in colors the theoretical predictions (correction factors: 0.939 for OH and 0.953 for NH and CH). All the calculations performed at M06-2X/6-311++G(d,p) level. On the left side the energy value in kJ/mol at 0 K, and at the temperature where ΔG=0 is also displayed.



EXTERNAL EVALUATORS:

Dr. Luca Evangelisti – University of Bologna

Dr. Gustavo Garcia-Macias – SOLEIL Synchrotron

Prof. Zbigniew Kisiel - Institute of Physics, Polish Academy of Sciences