



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Escola d'Enginyeria de Barcelona Est

TREBALL FI DE GRAU

Grau en Enginyeria Química

SIMULACIÓ ATOMÍSTICA D'ANTICOSSOS SOBRE UNA SUPERFÍCIE D'OR PER A INTERFÍCIES DE SENSORS



Memòria i Annexos

Autor: Eduard Martín Martínez

Director: Joan Torras Costa

Convocatòria: Juny 2021



Resum

La modelització atomística computacional ha demostrat ser una eina de gran fiabilitat i s'ha consolidat com una mètode fonamental per l'anàlisi de l'estructura i les propietats de sistemes proteínics complexos. En l'estudi d'aquest projecte dos sistemes formats pels anticossos CR3022 i S309, específics per la detecció de la RBD del virus SARS-COV-2 han estat simulats.

L'objectiu principal amb el que s'ha desenvolupat el treball a estat l'anàlisi estructural d'aquests dos sistemes d'anticossos per comprovar la viabilitat de la seva futura aplicació en immunosensors capaços de detectar el virus SARS-COV-2. Per aquest motiu s'ha dut a terme una immobilització de la proteïna a través de la regió Fc, sobre una superfície d'or. Tot seguit s'han realitzat una sèrie de simulacions clàssiques i d'AMD per tal d'observar el comportament dels sistemes durant un determinat període de temps i explorar el major nombre de conformacions possibles.

Els posteriors resultats extrets mostren un comportament un comportament positiu i molt similar pels dos sistemes, que permeten portar l'objectiu inicial al següent nivell. Sent el sistema format per l'anticòs CR3022 aquell en que s'observa una major estabilitat en l'estructura final i funcionalitat en regions més específiques, malgrat estar inicialment sotmès a unes forces d'interacció amb la superfície majors.

Resumen

La modelización atomística computacional ha demostrado ser una herramienta de elevada fiabilidad y se ha consolidado como un método fundamental para el análisis estructural y de las propiedades de sistemas proteínicos complejos. En el estudio de este proyecto, dos sistemas formados por los anticuerpos CR3022 y S309, específicos para la detección de la región RBD del virus SARS-COV-2 han sido simulados.

El objetivo principal en el desarrollo del trabajo ha estado el análisis estructural de estos dos sistemas de anticuerpos para evaluar la viabilidad de su futura aplicación en inmunosensores capaces de detectar el virus SARS-COV-2. Para ello, se ha llevado a cabo un proceso de inmovilización de la proteína a través de la parte inferior de la región Fc, sobre una superficie de oro. Seguidamente se han realizado una serie de simulaciones clásicas y de aMD para poder observar la evolución en el comportamiento de los sistemas durante un periodo de tiempo determinado y explorar el mayor número de conformaciones posibles.

Los posteriores resultados muestran un comportamiento positivo y muy similar en los dos sistemas que permiten llevar el objetivo inicial al siguiente nivel. Siendo el sistema formado por el anticuerpo CR3022 aquel en que se observa una mayor estabilidad en la estructura final i manteniendo la funcionalidad en las regiones más específicas a pesar de estar inicialmente sometido a unas fuerzas de interacción con la superficie superiores.

Abstract

Computational atomistic modelling has proven to be a highly reliable tool and has established itself as a fundamental method for the structural analysis of complex protein systems. In the study of this project, two systems formed by the antibodies CR3022 and S309, specific for the detection of the RBD regions of the SARS-COV-2 virus has been simulated.

The main objective in the development of this work has been the structural analysis of these two antibody systems to evaluate the viability of their future application in immunosensors capable of detecting the SARS-COV-2 virus. For this, a protein immobilization process has been carried out through the lower part of the Fc region, on a gold surface. Subsequently, a series of classical and aMD simulations have been carried out to be able to observe the evolution in the behaviour of the systems during a determined period of time and to explore as many conformations as possible.

The subsequent results show a positive and very similar behaviour in the two systems that allow taking the initial objective to the next level. The system formed by the CR3022 antibody being the one in which greater stability is observed in the final structure and maintaining functionality in the more specific regions despite being initially subjected to higher interaction forces with the surface.

Agraïments

La realització d'aquest projecte no hagués estat possible sense l'ajuda i el suport de les persones que es mencionen en aquest apartat.

En primer lloc, expressar el meu immens agraïment al doctorant Dídac Martí, encarregat de la supervisió del projecte i sense el qual, aquests resultats no haguessin estat possibles. Gràcies per estar sempre predisposat a donar un cop de mà amb tot el que ha fet falta.

Al Dr Joan Torras, per la oportunitat descobrir i explorar aquest ampli món que és la química computacional.

Al meu company i amic Marc Alsina, qui em va incentivar a triar aquest projecte i m'ha brindat el seu suport durant tot el procés.

I per acabar, a totes aquelles persones, familiars i amics que d'una o altra manera han fet possible la finalització d'aquest treball.

Índex

Resum	3
Resumen.....	4
Abstract	5
Agraïments	6
Índex	7
1. Introducció	9
2. Descripció del sistema biològic d'estudi.....	11
2.1. Virus SARS-CoV-2.....	11
2.2. Proteïnes d'estudi.....	13
2.3. Superfície del sistema	15
2.4. Sistemes simulats	16
2.4.1. Sistema de l'anticòs CR3022 amb superfície d'Or	19
2.4.2. Sistema de l'anticòs S309 sobre la superfície d'Or.....	21
3. Mètodes de modelització molecular:	23
3.1. Mecànica molecular clàssica	23
3.1.1. Force Field i interaccions moleculars	23
3.1.2. Linker de proteïna-superfície	26
3.1.3. Condicions periòdiques de contorn	27
3.1.4. Dinàmica Molecular Clàssica.....	28
3.2. Accelerated Molecular Dynamics (aMD).....	33
4. Mètodes d'anàlisi Clàssica	35
4.1. RMSD	35
4.2. RMSF	36
4.3. Angles	37
4.4. Diedres.....	39
4.5. Distàncies	40
4.6. Radis de Gir	41
5. Mètodes d'anàlisi aMD.....	42
5.1. Potencial Mean Force.....	42
5.2. Free Binding Energy	43

6. Metodologia de Simulació	45
7. Resultats	47
7.1. Estudi estructural del sistema.....	47
7.2. Regions de major mobilitat del sistema	51
7.3. Estudi de les regions Fab del sistema	54
7.4. Estudi de la regió Fc del sistema.....	62
7.5. Anàlisi final i comparatiu dels resultats	69
8. Avaluació Econòmica	72
9. Conclusions	75
Bibliografia	77
Annex	83

1. Introducció

Els anticossos del tipus Immunoglobulina G (IgG) tenen un rol molt important en la regulació del sistema immunològic humà i es troben entre les molècules més ben dissenyades i construïdes en la natura [1]. Gràcies a la seva excepcional capacitat de bio-reconeixement, amb una alta especificitat i afinitat pel reconeixement d'antígens, els anticossos IgG són àmpliament utilitzats en dispositius de sensors serològics (els anomenats immunosensors), per la detecció de patògens i toxines [2]–[4]. En aquest context, la tecnologia de recombinació d'anticossos, és una eina molt robusta per la fabricació de nous immunosensors.

Els anticossos IgG presenten una estructura monomèrica formada per dues cadenes pesades (H) idèntiques i dues lleugeres (L) també idèntiques. Les dimensions de les cadenes pesades són aproximadament el doble que les lleugeres i estan lligades entre elles a través de ponts disulfurs. L'estructura quaternària està formada per dues branques idèntiques que resulten en una forma de Y, on la part superior, les dos branques, estan formades per mitja cadena pesada i una cadena lleugera cadascuna i són les encarregades de lligar-se al seu respectiu antígen, és per això que s'anomenen les regions *Fragment Antigen Binding* (Fab1 i Fab2) de la proteïna. Per altra banda, la part inferior està formada per les restants meitats de les dues cadenes pesades i formen la base de la proteïna, la regió cristal·lina (Fc).

El rendiment dels sensors basats en la detecció per anticossos depèn, entre d'altres coses, de la eficàcia de la immobilització de l'anticòs a una superfície mentre es conserva la seva activitat natural. Degut a això, és de vital importància que aquesta unió amb una superfície sòlida es doni mitjançant la regió Fc de la proteïna, deixant d'aquesta manera les regions Fab lliures i exposades per permetre el reconeixement de l'antigen i optimitzant així el límit de detecció i sensibilitat del sensor [5].

El coronavirus del tipus 2 causant del síndrome respiratori agut greu (SARS-COV-2) ha activat una crisi sanitària a nivell mundial amb un impacte social sense precedents a través de la malaltia COVID19. Una glicoproteïna en forma de espícula que es troba a la membrana del virus unida per una regió constant i formada per tres monòmers a la part més exposada, s'ha vist que té un paper molt important en la infecció viral i patògena [6]. La primera etapa d'aquest procés és la unió del *Receptor-Binding Domain* (RBD) al seu receptor, un enzim concret del cos humà (ACE2) per formar un complex que deriva en la fusió de la membrana, gràcies a altres parts de la espícula [7], [8] i que acaba derivant en la entrada viral a cèl·lula objectiu.

És per això, que el desenvolupament d'aquest tipus de immunosensors que permeten detectar de forma ràpida i eficaç el virus SARS-COV-2 en la etapa inicial del diagnòstic és de vital importància.

El disseny de neutralitzadors de la regió RBD del virus a partir d'anticossos del tipus IgG s'ha vist com un mètode prometedor contra la malaltia del COVID19 [9], [10]. Durant la pandèmia s'han anat determinant una sèrie d'anticossos efectius contra aquesta malaltia. En aquest projecte, ens centrarem en dos d'ells, els anticossos CR3022 i S309. Diferents estudis clínics i de dinàmica molecular s'han fet en els últims anys en referència a la unió d'aquests dos anticossos amb la regió RBD de la espícula del SARS-COV-2 [11]–[14].

D'aquesta manera, l'objectiu principal en la realització d'aquest projecte és l'estudi de dos anticossos del tipus IgG, que han estat dissenyats de manera específica per detectar el virus SARS-COV-2 per examinar l'estabilitat quan s'immobilitzen sobre una superfície sòlida d'or, que és habitualment utilitzada pels immunosensors [15]–[17]. L'estudi s'ha realitzat a partir de simulacions moleculars i models atomístics.

Així doncs, per tal de dividir el treball d'una forma clara i estructurada, en primer lloc es descriuran els diferents sistemes que s'han estudiat en aquest projecte. A continuació s'exposaran els fonaments teòrics de les diferents metodologies i models aplicats per la simulació dels sistemes i les posteriors tècniques analítiques emprades. I finalment es comentaran els resultats obtinguts, el seu anàlisi i les posteriors conclusions extretes.

2. Descripció del sistema biològic d'estudi

2.1. Virus SARS-CoV-2

El SARS-COV-2 (Coronavirus del tipus 2) és el virus principal de la malaltia COVID19. És un tipus de coronavirus altament infecciosos causant de la malaltia i que la seva fàcil expansió va provocar la pandèmia mundial actual. Per aquest motiu, entendre com aquest virus entra a les cèl·lules humanes és de vital importància per tal de neutralitzar aquells mecanismes que ho fan possible.

Per complir la seva funció el virus SARS-COV-2 s'uneix a l'enzim ACE2 (*Angiotensin-converting enzyme 2*) una metal·loproteïna que juga un paper important en el sistema cardiovascular i que es troba present en molts òrgans del cos humà. La infecció a la cèl·lula hoste, a la qual transmet el codi genètic RNA es dona mitjançant una glicoproteïna que forma part de la superfície del virus, anomenada *Spike Protein* (Proteïna S). Dins d'aquesta proteïna es troba una regió específica anomenada RBD (*Receptor-Binding Domain*) que té una major afinitat amb l'enzim ACE2 que qualsevol altre variant del virus SARS[18] i que a diferència del que seria més normal aquesta zona RBD es troba menys exposada que en la resta, facilitant d'aquesta manera la seva infectivitat i rang d'actuació de manera important i incrementant la dificultat de atacar aquesta zona [19].

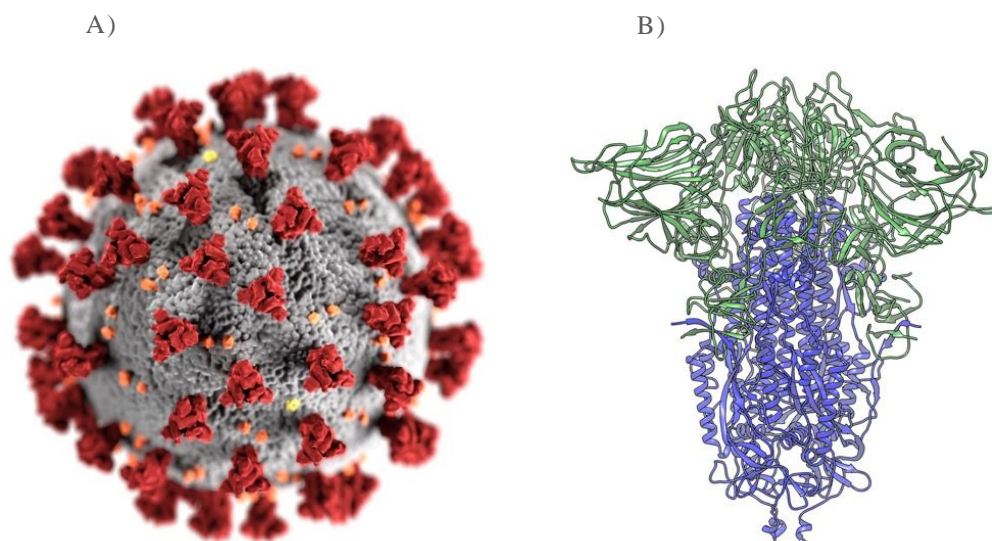


Figura 1: Imatges referents al virus SARS-COV-2. A) Il·lustració representativa de la morfologia ultraestructural del virus, creada pel CDC (Centres per al Control i Prevenció d'Enfermetats). Bb) Imatge obtinguda mitjançant criomicroscopia electrònica de la proteïna espícula mostrant en diferents colors les dues seccions principals (S1 verd i S2 blau). Font:[20].

Entendre el mecanisme d'entrada del Coronavirus a les cèl·lules hoste és important per la determinació de la infectivitat viral i patogènesis del virus, així com un objectiu per preveure la immunitat de l'hoste i les diferents estratègies d'intervenció[21]. Per entrar a la cèl·lula hoste, el virus primerament s'uneix a un receptor de la superfície de la cèl·lula objectiu, seguidament entra als endosomes i finalment es fusiona el virus amb la membrana del lisosoma per alliberar el genoma viral tal com es pot observar a la *Figura 2A*.

Una proteïna en forma d'espícula ancorada a la superfície és l'encarregada de controlar l'entrada del virus. Aquesta proteïna està formada per tres caps receptors-lligands S1 units a la part superior d'una zona més compacte anomenada S2 (*Figura 2B*). Aquesta zona està ben estudiada i es coneix que la zona S1 conté una sèrie d'aminoàcids anomenats *Receptor-Binding Domains* (RBD) que específicament reconeixen l'enzim ACE2. Aquest zona RBD varia la seva posició i passa d'estar en una posició elevada (*Standing-Up*) que li permet unir-se al seu receptor a una posició baixa que li concedeix una evasió important que dificulta la seva detecció[22] (*Figura 2B*). Per altra banda tal com es pot observar en la *Figura 2C-D* a diferència de la resta de virus SARS, la proteïna S del SARS-COV-2 conté una zona anomenada PPC que li concedeix un mecanisme de preactivació respecte a la resta de virus SARS que li doten d'una major efectivitat per entrar a la cèl·lula hoste.

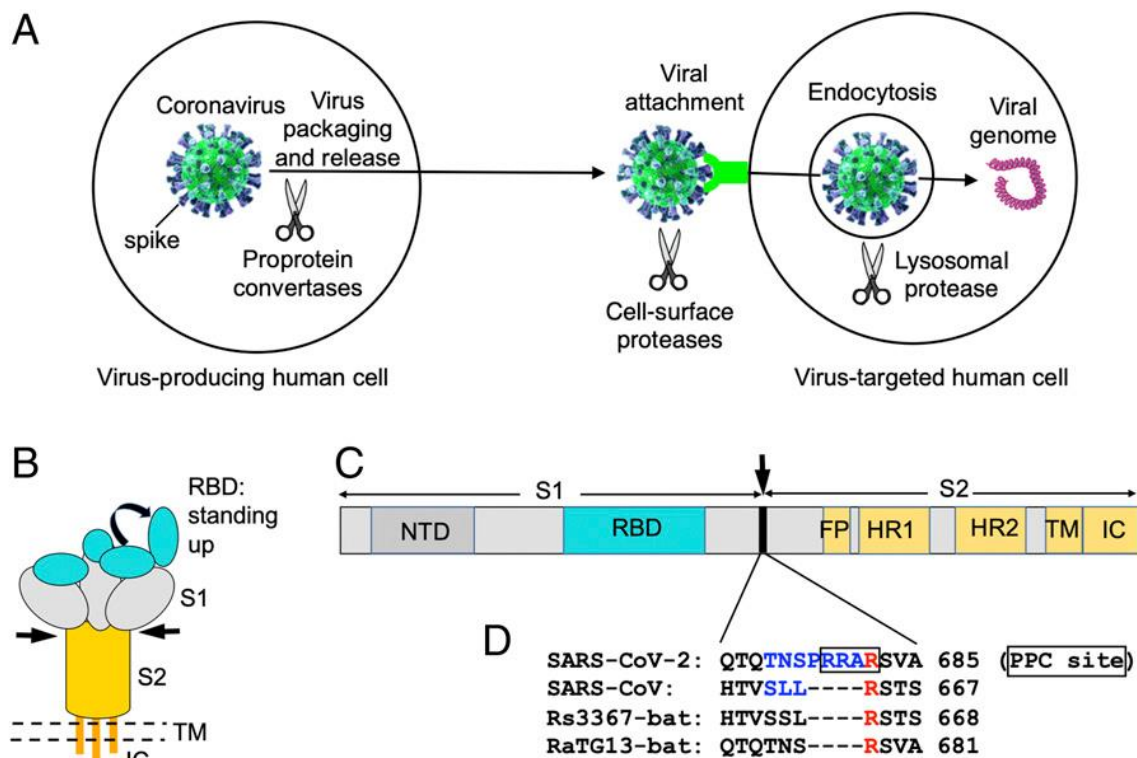


Figura 2 : Mecanisme d'actuació del Coronavirus. (A) Diferents etapes del mecanisme d'entrada del virus SARS-COV-2 a una cèl·lula humana. (B) Esquema de la proteïna S amb les seves parts. S1, la subunitat receptora d'unió; S2, subunitat unida a la membrana del

virus: RBD, en les posicions que ocupa segons la situació. (C) Dibuix esquemàtic de les parts de la espícula. (D) Part de la seqüència comparada de la proteïna S. Font: [19].

Per aquest motiu, en aquest projecte es realitza un estudi del comportament i estabilitat de dos anticossos que s'ha vist que ataquen i inutilitzen la proteïna S del virus SARS-COV-2 neutralitzant així el seu mecanisme d'actuació. Més concretament, aquests anticossos, actuen sobre la RBD (*Receptor-Binding Domain*) de l'espícula.

2.2. Proteïnes d'estudi

Com ja s'ha comentat, la proteïna a la que s'enllaça habitualment el virus és l'enzim ACE2, però en aquest cas l'estudi d'aquest projecte es centrarà en el comportament d'unes altres proteïnes que mostren certa afinitat cap a aquesta regió concreta del virus on es troben les anomenades RBDs de l'espícula del virus. Aquestes proteïnes són uns anticossos encarregats de bloquejar la RBD impedit d'aquesta manera que ataquin a l'enzim ACE2 i permetent que altres cèl·lules del sistema immunològic puguin neutralitzar l'antigen.

Aquest sistema proteínic està format per quatre cadenes, dues d'elles anomenades cadenes pesades i les altres dues, cadenes lleugeres. L'estructura global té una forma de Y on la part inferior està formada per la unió d'una part de les dues cadenes pesades i la part superior separada en dues branques on a cada cadena pesada se li complementa amb una de les cadenes lleugeres, tal i com es pot observar en la *Figura 3*. Aquestes unions entre cadenes són possibles gràcies a enllaços produïts mitjançant ponts di-sulfur.

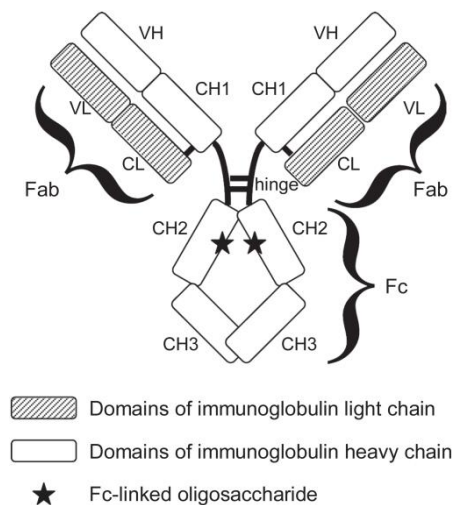


Figura 3: Representació de les diferents parts d'una proteïna corresponent a un anticòs IGG1 humà. Font: [23].

Aquest tipus de conformació mencionada correspon a un tipus concret de proteïnes, les immunoglobulines (Ig). Podem trobar cinc tipus de immunoglobulines principalment: IgG, IgA, IgM, IgD i IgE, on la més abundant en el cos humà i, per tant, amb la que s'ha treballat és la IgG[1], més concretament en la subclasse de la IgG1, que és la més abundant del seu tipus[24].

Aquesta proteïna es pot dividir principalment en dues regions: La Fab (*fragment Antigen binding*) que és la regió específica encarregada del reconeixement i d'interactuar amb el virus i que es troba dividida en dues branques a les que anomenarem Fab1 (La regió de la dreta) i Fab2 (Regió de l'esquerra); I la Fc (*Fragment Crystalline*) que és regió de l'anticòs que roman constant, determina el tipus de immunoglobulina en qüestió i que, a partir d'interaccions amb els seus lligands, intervé en varies funcions immune efectores i incrementa la vida de la molècula anticòs[25]. Finalment, aquestes dues parts estan unides per una regió anomenada *Hinge* que pren importància en la mediació de les dues zones. A la mateixa vegada, la regió Fab la trobem dividida en dos subregions: L'anomenada regió variable (Fv) i la regió continua que és comuna en totes les immunoglobulines i no te cap funció específica. La Fv en canvi, és la que aporta l'especificitat per unir-se i atrapar les proteïnes en qüestió i està formada per diferents subregions que li permeten realitzar aquesta funció, les CDRs (*Complementary-determining regions*). Aquestes regions CDR estan formades per un conjunt de residus específics i la funcionalitat i especificitat del l'anticòs depèn de que es conservi la seva estructura i es mantingui estable durant tot el procés de simulació.

Tenint en compte aquest tipus de proteïna, per adaptar el sistema al nostre objectiu que és la detecció del SARS-COV-2 s'ha substituït de manera computacional la part Fv de la regió Fab de la molècula per dos anticossos diferents específics per la interacció amb aquest virus, els anticossos CR3022 i S309. En el cas del CR3022, recentment s'ha confirmat la seva possibilitat d'unir-se a la proteïna S del SARS-COV-2 amb un cert grau d'activitat per desnaturalitzar-la. S'ha vist que l'anticòs CR3022 es pot lligar a l'epítip (Porció de l'antigen que reconeix el sistema immunitari, la seqüència concreta a la que s'uneix l'anticòs) del RBD del virus, el qual està format per un conjunt de 28 aminoàcids [11]. Per la seva banda, el S309 s'ha determinat com un possible anticòs per desnaturalitzar la proteïna S, gràcies a la seva capacitat per reconèixer un epítip que conté proteoglican diferent de la regió de la RBD a la que s'uneix l'enzim ACE2. Aquest epítip és accessible en els dos estats de la proteïna S, obert i tancat [26].

2.3. Superfície del sistema

Per tal de permetre la posterior aplicació d'aquests sistemes en aparells que permetin millorar la sensibilitat i rendiment d'immunosensors basats en mètodes mecànicoplàstics, s'ha escollit un tipus de material que simularà una superfície inorgànica on poder immobilitzar la proteïna. Aquesta superfície utilitzada és una capa composta únicament per àtoms d'Or (Au) formant una superfície uniforme i neutra.

L'or és un dels materials típicament més utilitzats en nanoestructures degut a les seves característiques que el fan un material bastant simple de sintetitzar en diferent tipus d'estructures i mides, es manté estable per llargs períodes de temps, té una funcionalitat com a superfície fàcil, és biocompatible, amb unes propietats òptiques úniques i una alta activitat catalítica que el fan un material ideal per al desenvolupament de sensors. Aquestes característiques el fan determinant en el desenvolupament de biosensors i nanosensors per la detecció de mostres, incloent contaminants inorgànics i orgànics, toxines, patògens, biomolècules, activitat enzimàtica i diagnòstic de diferents malalties. Les propietats òptiques de les partícules d'or combinades amb la seva capacitat de funcionar com a superfície, defineixen el seu ús efectiu en el desenvolupament d'aquests nano i biosensors. A més a més, l'ús de nanopartícules conjuntament amb aquests sensors permeten amplificar la senyal incrementant així la sensibilitat del dispositiu en qüestió[27].

Més concretament en aquest cas, la possibilitat d'immobilitzar anticossos en la superfície de nanopartícules d'or de manera fàcil i efectiva determina la seva efectivitat en el desenvolupament de nanoimmunosensors.

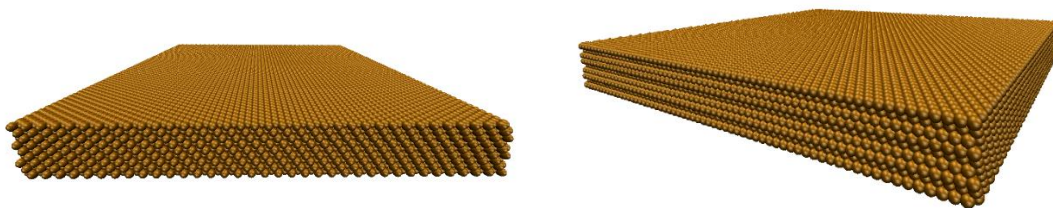


Figura 4: Representació obtinguda mitjançant VMD de la superfície d'or empleada en les simulacions des de diferents perspectives, formada completament per àtoms Au.

Un cop entesa la importància d'aquest material com a superfície del nostre sistema on s'immobilitzarà l'anticòs s'han analitzat les dimensions. És important tenir en compte que en cas real aquesta superfície s'hauria de considerar infinita, però degut a les limitacions dels mètodes computacionals i a l'elevat nombre d'àtoms que això suposaria pel sistema, s'ha determinat un rang de superfície lo suficientment gran per-què tot el conjunt de l'anticòs immobilitzat en la seva superfície es trobi sempre sobre aquesta superfície.

D'aquesta manera s'ha simulat una superfície d'or formada per una llosa de nou capes d'àtoms d'or amb unes dimensions de $179.55 \times 180.311 \times 18 \text{ \AA}^3$ representant l'estructura de més baixa energia de la cel·la unitària FCC[28], [29]. Més concretament cadascuna d'aquestes capes conté 4536 àtoms d'or formant un conjunt total de 40824 àtoms únicament per la superfície del sistema.

2.4. Sistemes simulats

Un cop vistes les principals parts del sistema, és moment de veure com es dona forma als sistemes finals que s'han utilitzat. S'ha simulat un total de quatre sistemes en funció de l'anticòs específic utilitzat i la superfície on s'ha immobilitzat la proteïna.

El sistema ideal a construir contindria tres parts: Una superfície on immobilitzar l'anticòs, que seria aquesta segona part i una tercera zona formada per la proteïna S del virus. Un sistema d'aquestes dimensions però, suposaria un cost computacional que a dia d'avui no es pot assumir. És per això, que en aquest projecte ens centrarem únicament en simular i analitzar el comportament del sistema compost per l'anticòs i la proteïna.

Per la preparació de la proteïna s'ha partit de dos sistemes inicials creats a partir de l'anticòs de la IGG1-B12 (PDB ID: 1HZH) utilitzada en la inactivació del HIV. S'han substituït les regions d'interacció amb el virus per les corresponents Fab de l'anticòs CR3022 (PDB ID: 6YLA) i Fv de l'anticòs S309 (PDB ID: 6WPT) (específics per la detecció del SARS-CoV-2).

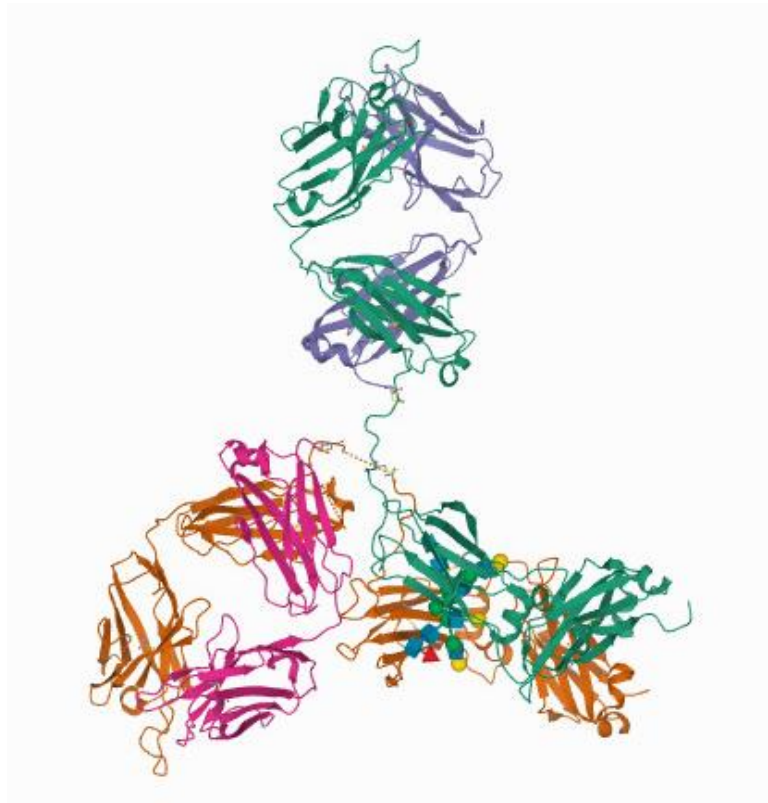


Figura 5: Representació de l'estructura cristal·lina de l'anticòs IgG1 B12 utilitzat en aquest projecte (PDB ID: 1HZH). Font: [30]

Per fer-ho, s'ha seguit el següent procediment:

Primerament utilitzant el *software* de UCSF Chimera[31] s'ha carregat el PDB corresponent a l'anticòs de la IGG1, (PDB Id: 1HZH). Seguidament es carrega el PDB de l'anticòs CR3022 (PDB Id: 6YLA) de tal forma que s'alineja amb una de les zones Fab del IGG1 per aconseguir que quedin correctament superposats. Aleshores es talla la cadena lleugera L de la Fab i el tros de la pesada H fins on acaba l'alineament de la proteïna original de la IGG1 i es substitueix pel CR3022. És repeteix el procediment per l'altre Fab de manera que es substitueixen les branques M (lleugera) i K (pesada) per acabar obtenint una proteïna amb la zona Fc original de la IGG1 però amb les zones Fab corresponents al CR3022.

En el cas del S309 el procediment és pràcticament el mateix, només varia que en comptes de tenir tota la Fab, del S309 únicament es té la part Fv i per tant de la IGG1 només es talla una quarta part de les cadenes pesades i la meitat de les lleugeres, deixant d'aquesta manera les regions Fc i contínues originals.

De manera que s'ha acabat generant els següents sistemes:



Figura 6: Representació de l'estructura cristal·lina de l'anticòs CR3022 juntament amb la regió RBD del SARS-COV-2, utilitzat en les regions Fab del sistema (PDB ID: 6YLA). Font: [32]



Figura 7: Representació de l'estructura cristal·logràfica de l'anticòs S309 (Regions en color groc i rosa) unit a l'epítoc corresponent de la proteïna S del SARS-COV-2 (PDB ID: 6WPT). Font: [33]

2.4.1. Sistema de l'anticòs CR3022 amb superfície d'Or

Aquest primer sistema està format per l'anticòs CR3022 (generat de manera que conté la Fc del IgG1 i els Fabs específics) immobilitzat sobre la superfície d'or vista també anteriorment. D'aquesta manera s'ha obtingut el sistema inicial que es pot observar en la Figura 5.

Aquest procés d'immobilització de la proteïna que és necessari en tots els sistemes és fa a partir de la substitució d'un residu de lisina (Lys448) situat a la part inferior de la cadena pesada CH3 de la regió Fc per un residu especial anomenat *Linker* (LNK) que permet unir la proteïna amb la superfície a partir d'un enllaç covalent molt estable, com ja es veurà més endavant.

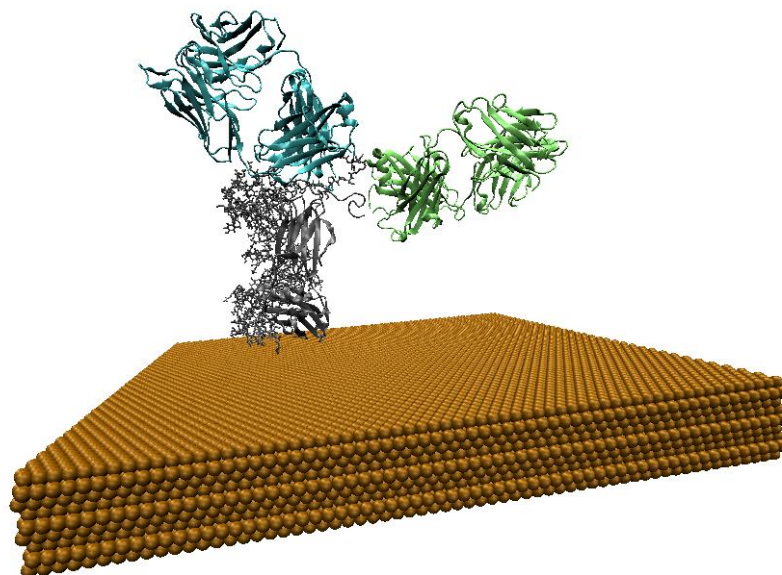


Figura 8: Representació del sistema CR3022_GOLD on es poden diferenciar les diferents parts principals: a) Molècula CR3022 on en verd es veu la regió Fab1, en blau la Fab2 i en gris la regió Fc, b) La superfície d'or representada per la massa d'àtoms de color ocre.

Aquest sistema es simula a un PH estable de 7 i s'ha solvatat en una caixa cúbica amb molècules d'aigua per simular així l'entorn aquós en què es solen trobar aquestes proteïnes, de manera que l'entorn de la caixa es troba en condicions periòdiques de contorn. D'aquesta manera s'obté un sistema format per 162597 molècules d'aigua explícita, una molècula formada per 1332 residus amb un total de 14607 àtoms i tenint en compte la superfície s'acaba obtenint un sistema format per 543221 partícules per simular.

Per tal de poder aplicar les condicions periòdiques de contorn necessàries per simular correctament el sistema és necessari establir els límits d'una caixa on incloure tot el conjunt de partícules del mateix. En aquest cas les dimensions de la caixa són $180 \times 180 \times 218 \text{ \AA}^3$.

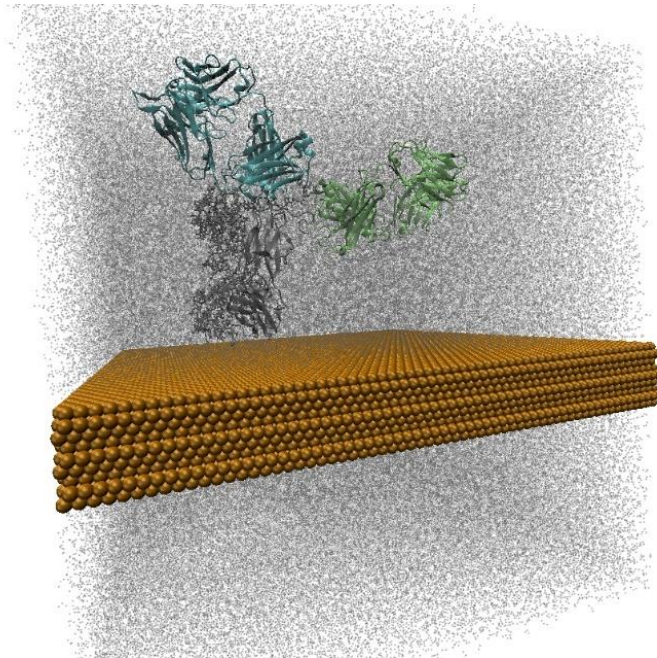


Figura 9: Representació del sistema CR3022_GOLD de la figura 5, incloent la caixa de solvatació d'aigua que cobreix el sistema (Imatge renderitzada amb VMD).

2.4.2. Sistema de l'anticòs S309 sobre la superfície d'Or

El sistema S309_GOLD està format per l'anticòs S309 immobilitzat en la superfície d'or de la mateixa forma que en la resta de sistemes, a través del residu LNK que en aquest cas substitueix el residu LYS456 situat en la mateixa zona inferior de la cadena pesada CH3 de la regió Fc.

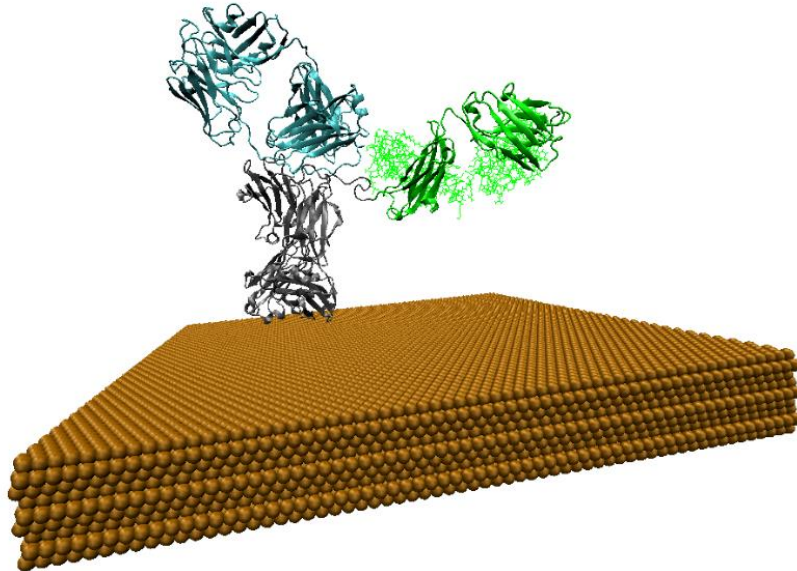


Figura 10: Representació del sistema S309_GOLD on es poden diferenciar les diferents parts principals: a) Molècula S309 on en verd es veu la regió Fab1, en blau la Fab2 i en gris la regió Fc, b) La superfície d'or representada per la massa d'àtoms de color ocre.

Com l'objectiu és obtenir els dos sistemes el més similars possibles per tal d'aconseguir uns resultats similars per comparar, el nombre de molècules d'aigua i d'àtoms de la superfície del sistema són iguals que en el cas del sistema CR3022 de manera que s'obté un sistema final format per 162597 molècules d'aigua solvatant una proteïna de 1327 residus i formant un sistema final de 523223 partícules controlat en condicions periòdiques de contorn a l'interior d'una caixa de $180 \times 180 \times 218 \text{ \AA}^3$.

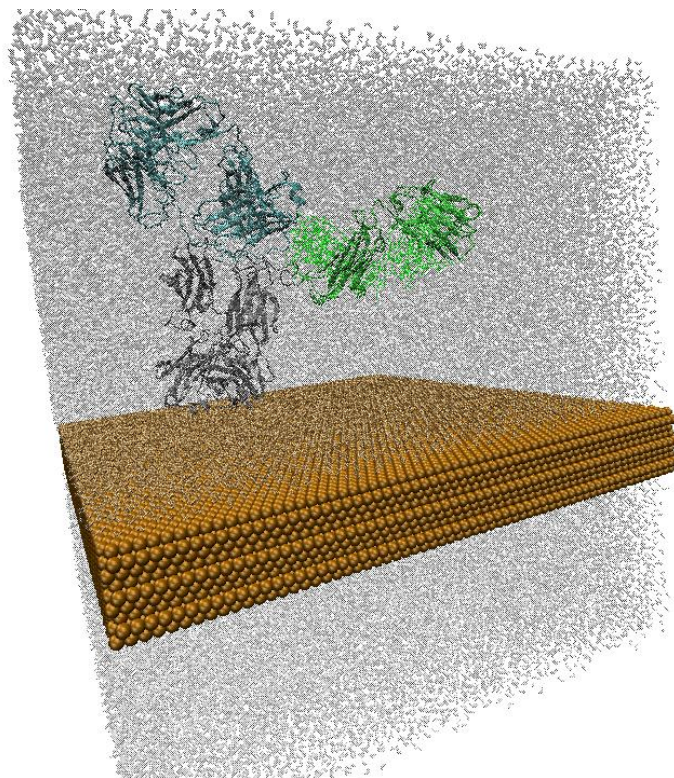


Figura 11: Representació del sistema S309_GOLD de la figura 7, incloent la caixa de solvatació d'aigua que cobreix el sistema.

3. Mètodes de modelització molecular:

3.1. Mecànica molecular clàssica

La mecànica clàssica va ser un dels primers models físics desenvolupats per modelitzar l'estructura i comportament de les molècules. Es basa en uns principis on els nuclis i electrons estan agrupats en partícules puntuals esfèriques on el radi i la càrrega venen determinades per mesures experimentals i dades teòriques. Les interaccions estan basades en potencials de la mecànica clàssica i determinen la distribució espacial de les partícules i les seves energies tenint en compte que el fonament principal d'aquest model és el fet que qualsevol àtom o molècula sempre es mou seguint un estat d'equilibri.

Les aplicacions de la mecànica clàssica molecular són molt variades gràcies al seu baix cost computacional, en comparació amb els mètodes quàntics però per d'altra banda, aquest sistema no contempla estats excitats o reactius i per tal d'observar altres propietats d'interès com la formació o ruptura d'enllaços s'hauria de tractar amb mètodes quàntics o híbrids.

3.1.1. Force Field i interaccions moleculars

En dinàmica molecular una molècula es descriu com una sèrie de punts amb càrrega (àtoms) enllaçades per molles (*bonds*). Per tal de descriure l'evolució al llarg del temps de la distància de l'enllaç, els angles, les torsions i les forces no enllaçants com les interaccions de van der Waals i electroestàtiques entre àtoms s'utilitza un *Force Field*[34].

El *Force Field* o camp de força del sistema és el conjunt d'equacions i paràmetres que descriuen les interaccions, tant intermoleculars com intramoleculars, d'una molècula, comptabilitzades en termes d'energia. Un cop generades les proteïnes, mitjançant el *software* TLeap de AMBER18 [35] s'han generat els diversos fitxers de paràmetres necessaris.

Abans d'iniciar qualsevol simulació és important tenir en compte i especificar tots aquells paràmetres necessaris per simular els diferents tipus d'àtoms que formen els sistemes i mantenir-los constants per tot el conjunt de simulacions. Per aquest motiu en el cas del conjunt aquós format únicament per molècules d'aigua que envolten tant la proteïna com la superfície d'or s'ha utilitzat el *Force Field water-TIP3P* [36], per la superfície formada únicament per àtoms d'or també s'han tingut en compte els paràmetres concrets al tractar-se d'àtoms metàl·lics [37] mentre que per la proteïna s'ha fet servir el *Force Field ff03ua* [38]. Finalment de manera més específica la parametrització i optimització del *linker* de la proteïna s'ha fet en *Gaussian*[39] i després s'ha passat a Amber mitjançant *tLeap*.

Els càlculs referents a les forces del sistema són els que tenen un cost computacional més elevat dins de la dinàmica molecular ja que s'obté del gradient negatiu d'una funció energètica potencial escalar i per sistemes biomoleculars implica el sumatori de les energies enllaçants i no enllaçants de manera que ve donat per un sumatori tal que [40]:

$$U(r) = \sum U_{\text{enllaçant}}(r) + \sum U_{\text{no-enllaçant}}(r) \quad (1)$$

Termes de potencial d'energia enllaçant:

Dins dels termes enllaçants trobem els següents potencials en referència al moviment harmònic vibracional, el moviment angular vibracional i el diedre (o també anomenat angle de torsió), obtinguts mitjançant les següents equacions:

$$U_{\text{bond}} = k(r_{ij} - r_0)^2 \quad (2)$$

$$U_{\text{angle}} = k(\theta - \theta_0)^2 + k_{ub}(r_{ik} - r_{ub})^2 \quad (3)$$

$$U_{\text{dihed}} = \begin{cases} k(1 + \cos(n\varphi + \phi)) & \text{if } n > 0 \\ k(\varphi - \phi)^2 & \text{if } n = 0 \end{cases} \quad (4)$$

On:

- $U(r)$, Energia potencial del sistema, depèn de r .
- r , Posició del sistema (en vectors) .
- k, k_{ub}, k_{ik} , Constants de cada sistema.
- θ, φ , Angle (en radiants) entre vectors de velocitat.
- r_0, θ_0 , Posicions d'equilibri de les variables.

El conjunt d'aquestes equacions permet el càlcul de les forces del sistema en funció de les dades geomètriques de les partícules i dels paràmetres de cada tipus d'àtom. Per cada pas, U_{bond} calcula l'energia de les oscil·lacions de la distància d'enllaç, U_{angle} les oscil·lacions de l'angle d'enllaç d'un conjunt de tres àtoms i finalment, U_{dihed} l'energia de rotació del diedre entre quatre àtoms.

Termes de potencial d'energia no enllaçant:

Els termes no enllaçants tenen en compte les interaccions entre aquells parells d'àtoms que, normalment, no es troben involucrats en termes d'energia enllaçants i té en compte el potencial de Lennard-Jones [41] que permet calcular les interaccions per forces de Van der Waals i el potencial electroestàtic de manera que s'obtenen dos termes a partir de les següents equacions:

$$U_{LJ} = -(E_{\min}) \left[\left(\frac{R_{\min}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\min}}{r_{ij}} \right)^6 \right] \quad (5)$$

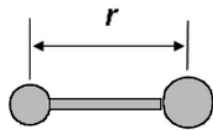
$$U_{elec} = \frac{\epsilon_{ij} C q_i q_j}{\epsilon_0 r_{ij}} \quad (6)$$

On:

- q_i, q_j , Càrregues de cada àtom.
- C , Constant de Columbus.
- ϵ_{ij} , Factor de *Lennard-Jones*.

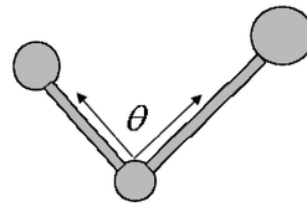
Physical model for the AMBER force field

$$V_{total} = V_{bond} + V_{angle} + V_{torsion} + V_{non-bond}$$



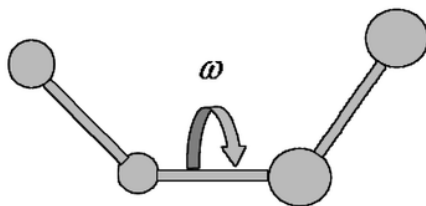
$$V_{bond} = k_{bond} (r - r_0)^2$$

(a)



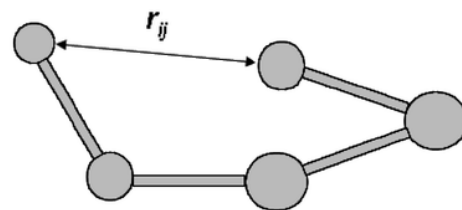
$$V_{angle} = k_{angle} (\theta - \theta_0)^2$$

(b)



$$V_{torsion} = \frac{1}{2} k_{torsion} \{1 + \cos(n\omega - \omega_0)\}$$

(c)



$$V_{non-bond} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\epsilon_{ij}}$$

(d)

Figura 12: Recull dels diferents paràmetres que es calculen per determinar el Force Field d'un sistema amb la representació gràfica dels àtoms que intervenen. Font: [42].

3.1.2. Linker de proteïna-superfície

Quan es parla del *linker* ens referim a aquell residu que permet immobilitzar la proteïna del sistema amb la superfície sobre la que es treballa, a partir de fixar-la sobre els àtoms d'or. En aquest cas s'ha optat per modificar la lisina C-terminal (LYS448 i LYS456) del final de la cadena pesada (CH) de la regió Fc. I, a continuació, mitjançant GAUSSIAN VIEW, es modifica la cadena lateral per canviar el grup NH₃ de la cadena lateral de l'aminoàcid final per enllaçant-lo amb un grup carboxil de la superfície d'or prèviament activada.

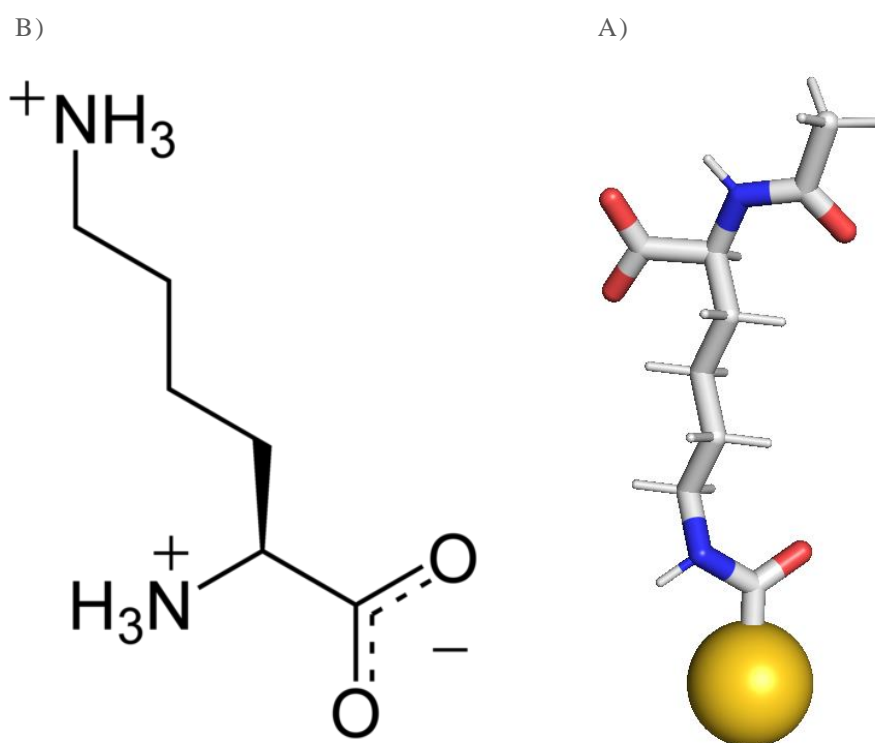


Figura 13: Figura A) Representació de l'estructura d'un aminoàcid de Lisina, Font: [43].
Figura B) Representació d'una lisina modificada i unida a un àtom d'or de la capa superior de la superfície.

Aquest sistema resultant es pot observar en la *Figura 13B* i, a partir d'aquest, s'ha treballat amb el *software* gaussian09 [44] i la metodologia R.E.D. [45]–[47] per calcular els paràmetres del *linker*. Primerament, s'ha utilitzat *Gaussian* per optimitzar la geometria del sistema i obtenir el paràmetre de la distància de l'enllaç Au-C ($d = 2.064 \text{ \AA}$) mitjançant el *Force Field* B3LYP [48]. Les funcions de base que s'han utilitzat en el cas dels àtoms de carboni, oxigen, nitrogen i hidrogen és la 6-31g* [49]. En aquest punt, s'han obtingut les càrregues de tot el sistema i s'ha transformat al *Force Field* ff03ua per tal de mantenir la uniformitat amb el conjunt total dels paràmetres del sistema.

3.1.3. Condicions periòdiques de contorn

Les condicions periòdiques de control (*Periodic Boundary Conditions*) permeten realitzar una simulació fent servir un nombre de partícules relativament petit en comparació al que s'haurien de fer servir per un sistema normal. D'alguna manera aquestes partícules en condicions periòdiques experimenten forces com si estiguessin en un fluid infinit. Imaginem una caixa cúbica plena de partícules que és replicada en totes direccions de manera que quedi completament envoltada d'altres caixes, donant així una matriu periòdica, com es pot observar en la *Figura 14*, pel cas d'un sistema en dues dimensions. En aquest sistema de la imatge per exemple, cada caixa estaria envoltada per altres 8 caixes; en un sistema en tres dimensions com el que simulem, serien un total de 26 caixes de control. Les coordenades de les partícules en les caixes envoltants que fan d'imatge, poden ser simulades simplement afegint o traient integrals múltiples dels costats de les caixes. D'aquesta manera, en el cas que una partícula abandonés la caixa per un costat, aquesta seria replicada per una partícula igual que entrés al sistema pel costat oposat, com es veu en el cas de les partícules ratllades del sistema de la *Figura 14*. Així s'aconsegueix que el nombre de partícules del sistema sempre es mantingui constant amb un cost computacional molt baix [50].

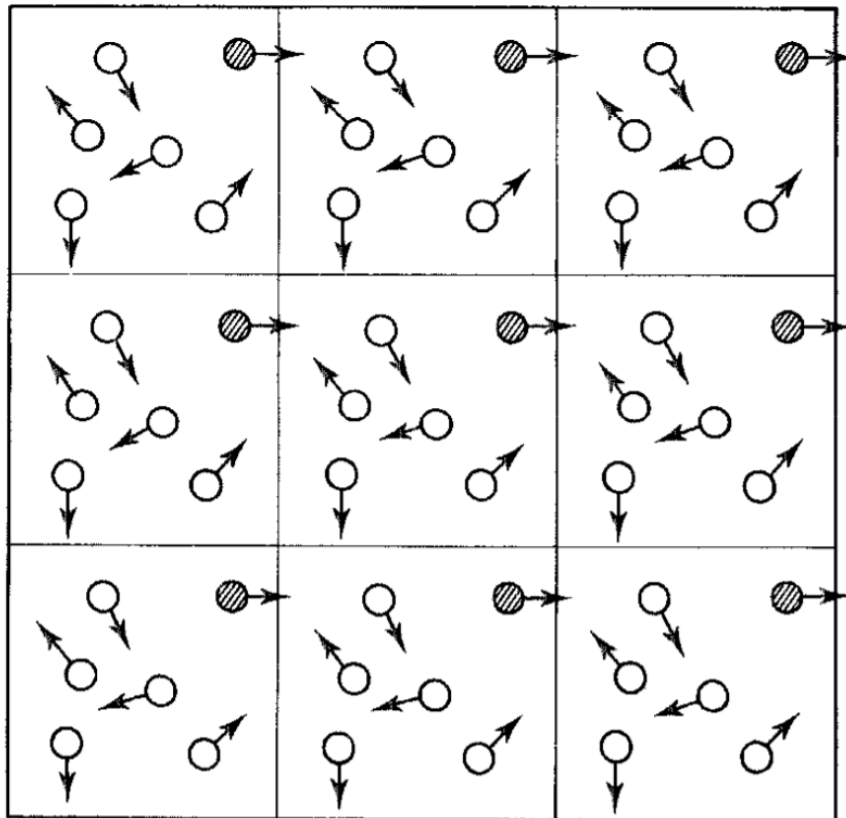


Figura 14: Imatge representativa de les condicions periòdiques de contorn per un sistema en dues dimensions. Font: [50]

3.1.4. Dinàmica Molecular Clàssica

Les tècniques de dinàmica molecular es poden utilitzar per explorar les diferents conformacions espacials i és un mètode altament utilitzat per simular molècules grans com les proteïnes amb les que es treballa en aquest projecte. En dinàmica molecular l'energia de superfície s'explora a partir de resoldre les lleis de Newton del moviment pel sistema [51]. En les simulacions clàssiques el sistema queda definit per dos factors: Les posicions que ocupen, r , i el moment conjugat, p , de les partícules de manera individual, on aquests factors estan representats per vectors en tres dimensions. D'aquesta manera en absència de limitacions i forces dependents de la velocitat el Hamiltonià del sistema es pot escriure de la manera següent:

$$H(r, p) = K(p) + U(r) \quad (7)$$

On $K(p)$ és l'energia cinètica que es pot calcular a partir de:

$$K(p) = \sum_i \frac{p_i^2}{2m_i} = \sum_i \frac{1}{2} m_i \cdot v_i^2 \quad (8)$$

El moment lineal p , segons la seva definició és la massa m de la partícula i per la seva velocitat v . D'aquest punt es pot extreure el temps a partir de la velocitat derivada de la posició de la partícula tal que, $v_i = dr_i/dt$.

Per altra banda, el terme $U(r)$ és la energia potencial obtinguda a partir de les interaccions entre partícules del sistema i queda recollida en els paràmetres de *Force Field* prèviament explicats. A partir de l'energia potencial és pot calcular la força f_i , d'una partícula i en respecte a la posició r .

$$f_i = -\frac{\partial}{\partial r_i} U(r) \quad (9)$$

Utilitzant aquesta força ja es pot aplicar l'equació de moviment que segons Newton [52] exposa, la segona derivada del temps o l'acceleració, de la posició d'una partícula es igual a la força exercida en ella entre la seva massa;

$$\frac{d^2}{dt^2} r_i = \frac{1}{m_i} f_i \quad (10)$$

La aplicació d'aquesta integració per totes les partícules del sistema de manera simultània permet seguir el moviment de les partícules al llarg del temps. Aquest procés és possible mitjançant diversos algorismes com els de LEAPFROG [53] o VERLET [54]. El *time*

step, o temps d'integració, escollit ha de ser prou petit per permetre la correcta integració dels paràmetres més ràpids del sistema. Si es tingués en compte el terme U_{bond} s'hauria de tenir un *time step* de 1,0 fs, però en les simulacions d'AMBER s'ha restringit els enllaços amb constricció d'àtoms d'hidrogen mitjançant l'algoritme de SHAKE[55]. D'aquesta manera es permet aplicar un *time step* en les simulacions de dinàmica molecular d'aquest projecte de 2,0 fs.

Finalment, a partir de la simulació de la molècula amb el seu entorn per un període de temps es pot obtenir les propietats aproximades d'aquella molècula que s'aproximarien a les obtingudes en mesures reals. Això implica que per la simulació del sistema en el temps, es poden extreure conformacions que corresponen a un conjunt termodinàmic o punt d'estat en equilibri. D'aquesta manera a partir de les equacions anteriors i segons l'algoritme emprat es pot obtenir un conjunt termodinàmic canònic (N, V, T), on la temperatura T , es manté constant; microcanònic (N, V, E), on es l'energia la que es conserva; o isotèrmic-isobàric (N, p, T), on la pressió es manté constant.

Per tal de poder aplicar cadascuna d'aquestes topologies a les simulacions dels sistemes proposats on moltes vegades és necessari mantenir certes variables constants en algunes etapes de la simulació, s'han de seguir una sèrie de protocols que queden descrits a continuació.

3.1.4.1. Protocol de Minimització

Dins de la bioquímica, el terme de conformació es refereix a una agrupació determinada d'àtoms en una molècula que pot variar per la rotació d'enllaços simples. Com sabem que els diferents tipus d'enllaços tenen associat uns valors energètics específics, la lliure rotació de l'enllaç sigma (σ), com la que es produeix entre els enllaços carboni-carboni, pot induir canvis energètics al conjunt total de la molècula, deixant-la en un estat energètic que pot ser més o menys favorable. L'energia requerida per passar d'una conformació a una altres és principalment l'energia de torsió (*Equació 4*), però també influeixen altres tipus de forces no enllaçants com les forces de Van der Waals, les interaccions dipol-dipol o els ponts d'hidrogen.

Sabem que les propietats químiques d'una molècula poden variar considerablement depenent de la distribució espacial 3D que ocupin. És degut a això, que la minimització energètica és essencial per determinar la distribució de la molècula en l'espai, ja que l'estructura cristal·logràfica de partida no sol ser energèticament favorable. Durant el procés de minimització, el programa busca el punt de mínima energia local i un cop assolida la conformació estructural que més s'aproxima a aquest punt, s'atura. Aquest punt trobat però, no és el punt de menys energia del sistema, és aquell mínim més proper que es pot assolir sense sobrepassar una barrera energètic tal com queda representat en la *Figura 15*.

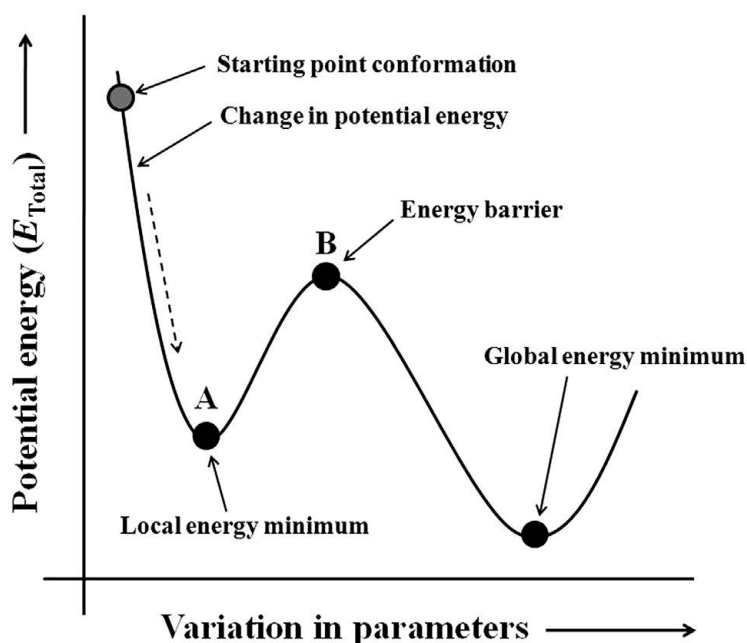


Figura 15: Diferents fases d'una molècula durant la minimització energètica. Font: [56]

D'aquesta manera s'ha aplicat el *software* de NAMD (V2.13) [57] que utilitza un mètode concret integrat per dur a terme aquest procés de minimització. Aquest mètode s'anomena *Conjugate gradient energy minimization* i utilitza un algoritme iteratiu de gradient descendent per trobar el valor mínim d'una funció. Aquest mètode permet resoldre de forma ràpida grans sistemes d'equacions lineals sense un mínim conegut utilitzant menys recursos computacionals i iteracions d'optimització que altres mètodes com el de *Steepest Descent* [58].

3.1.4.2. Protocol de Heating

El següent pas a seguir una vegada s'ha assolit una conformació de mínima energia, és portar el sistema a una temperatura determinada. Es parteix d'un estat inicial a 0K on les molècules estan estàtiques i es porta, en el nostre cas, a una temperatura de treball de 298K. Per realitzar aquests processos es poden fer servir diversos algoritmes que realitzen una funció de termòstats. En els dos sistemes que componen aquest projecte s'ha utilitzat l'algoritme de *Langevin Dynamics* [59].

Langevin introdueix un sistema amb una força fluctuant $\xi(t)$ que permet descriure el moviment erràtic de les partícules degut a l'atzar, a l'impacte descompensat de les molècules del solvent [60]. Langevin assumeix que aquest conjunt de forces fluctuants tenen una mitja de zero i que són independents de la posició de la partícula. D'aquesta manera es modifica l'equació de moviment de la partícula proposada per la dinàmica Newtoniana:

$$m \cdot a = -\gamma \cdot v \quad (11)$$

On:

- m , massa de la partícula.
- a , acceleració de la partícula.
- v , velocitat de la partícula.
- γ , coeficient de fricció.

I s'afegeix el terme proposat per Langevin com un sumatori, de forma que s'obté la següent expressió:

$$m \cdot a = -\gamma \cdot v + \xi(t) \quad (12)$$

A partir de la posterior correlació d'aquesta equació s'obté una relació en que per cada *time step* el termòstat de Langevin varia el moment tal que:

$$\Delta \vec{p}_i = \left(\frac{\partial V(r_i)}{\partial r_i} - \gamma \vec{p}_i + \delta p \right) \Delta t \quad (13)$$

Tal que \vec{p}_i correspon a moment de la partícula i δp és un nombre Gaussià aleatori obtingut a partir de la desviació estàndard del moment.

D'aquesta manera, a partir de l'algoritme de *Langevin Dynamics* NAMD permet l'equilibri del sistema a partir d'una reassignació de la temperatura, que modifica les velocitats de totes les partícules del sistema de manera periòdica fins que el conjunt del sistema assoleixi la temperatura desitjada. Per fer-ho, s'han d'especificar una sèrie de paràmetres en referència a la temperatura objectiu, l'increment per cada pas i la freqüència de *time steps* amb la que es realitza cada iteració [61].

3.1.4.3. Protocol NPT

Per l'equilibrat NPT del sistema s'ha utilitzat el procés de Nosé-Hoover Langevin *Piston Pressure Control* [62].

És un procés de NAMD que combina els mètodes de Nosé-Hoover *Constant Pressure Method* [63] i de *Langevin Dynamics* que permeten portar el sistema a la pressió desitjada mitjançant un sistema de pistó amb una temperatura i una oscil·lació determinades. Pel control d'aquest sistema, és important definir una sèrie de paràmetres. Primerament s'ha de fixar una temperatura (la mateixa que es vol mantenir a tot el sistema) per tal de

mantenir-la constant al llarg de tot l'equilibrat. A continuació s'han de definir dos paràmetres concrets que permetran simular aquest tipus d'equilibrat de pistó, un d'ells és habilitar la flexibilitat de la cel·la i l'altre és mantenir l'àrea constant. D'aquesta manera el que s'aconsegueix és mantenir constant el sistema en els eixos X i Y (les dimensions de l'àrea) i deixar amb llibertat únicament l'eix Z simulant així el comportament d'un sistema de pistó sobre la caixa de simulació.

Les equacions del moviment que intervenen en aquest cas són:

$$r' = \frac{p}{m} + e' \cdot r \quad (14)$$

$$p' = F - e'p - gp + R \quad (15)$$

$$V' = 3e' \quad (16)$$

$$e'' = \frac{3V}{W}(P - P_0) - g_e e' + \frac{R_e}{W} \quad (17)$$

On W és la massa del pistó, R és el soroll en els àtoms i R_e correspon al soroll sobre el pistó. De manera que s'obtenen a partir de les següents equacions:

$$W = 3N\tau^2kT \quad R^2 = \frac{2mgkT}{h} \quad R_e^2 = \frac{2Wg_e kT}{h} \quad (18)$$

Sent τ , el període d'oscil·lació.

Mitjançant aquest sistema s'ha d'especificar la pressió objectiu que es vol assolir, un període d'oscil·lació periòdic que de manera estàndard serà de 200fs i finalment la temperatura constant que es vol mantenir, on 298K serà la especificada en aquest projecte.

3.1.4.4. Protocol NVT

Quan s'intenta simular el comportament de determinats sistemes, com és el cas de les proteïnes on es molt important mantenir una temperatura amb variacions mínimes, de manera que es mantingui constant, en dinàmica molecular s'intenta reproduir un sistema isotèrmic. L'equilibrat NVT, com el seu nom indica es realitza mitjançant unes condicions de número de partícules, volum i temperatura constants i s'anomena ensamblatge canònic. Això no implica que per cada pas de l'equilibrat aquestes variables siguin constants, sinó que és el valor mig global al llarg de tot el procés, el que roman constant. D'aquesta forma es permet una fluctuació en l'energia total del sistema.

De la mateixa manera que en l'equilibrat NPT, s'utilitza el mètode de *Langevin Dynamics* per mantenir constant a la temperatura prèviament escalfada de 298K.

3.2. Accelerated Molecular Dynamics (aMD)

Després de les simulacions clàssiques una tècnica molt utilitzada en les simulacions MD és la *accelerated molecular dynamics* (aMD). Aquesta tècnica a diferència de les simulacions MD clàssiques permet sobrepassar els límits de les barreres energètiques que es troben al llarg del sistema de forma ràpida, de manera que es poden arribar a explorar conformacions i sistemes que mitjançant simulacions clàssiques no s'arriben a assolir o si ho fan, és amb un cost de temps molt més elevat.

De manera simplificada, les simulacions aMD són una tècnica de mostreig millorada que treballa implementant un increment no negatiu de potencial per reduir les barreres energètiques accelerant així la transició entre els estats de baixa energia del sistema.

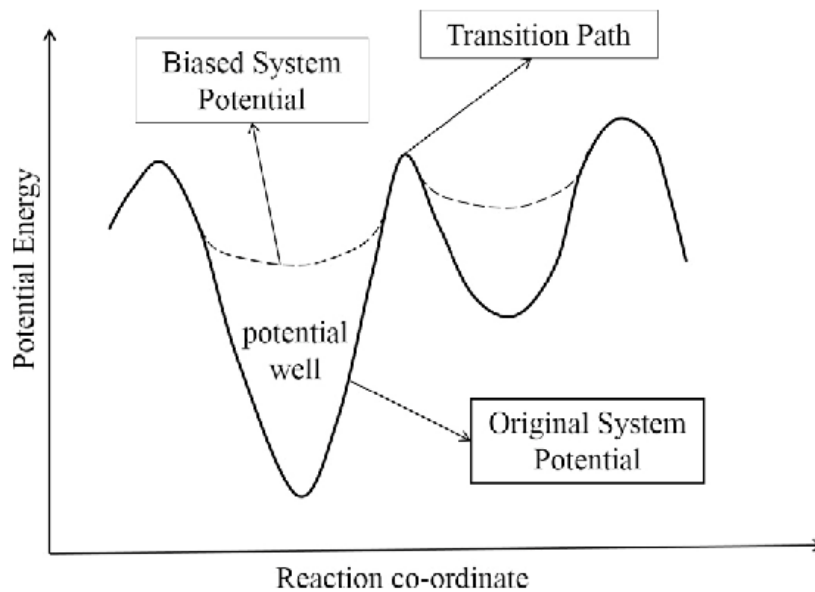


Figura 16: Representació del funcionament aMD. La línia sòlida representa el potencial original del sistema mentre que la discontinua mostra els punts que faciliten aquestes transicions energètiques entre les barreres. Font: [64]

El fonament darrera d'aquesta tècnica consisteix en modificar l'energia potencial del sistema de manera que quan aquesta energia es troba per sota d'un valor prèviament definit (E) s'aplica un factor per variar aquesta energia de tal manera que:

$$V'(r) = V(r), \quad V(r) \geq E \quad (19)$$

$$V'(r) = V(r) + \Delta V(r), \quad V(r) < E \quad (20)$$

On:

- $V(r)$ = Potencial original .
- $V'(r)$ = Potencial modificat.
- E = Energia de referència.
- $\Delta V(r)$ = Increment de potencial .

I aquest increment de potencial $\Delta V(r)$ donat per la següent fórmula:

$$\Delta V(r) = \frac{(E - V(r))^2}{\alpha + E - V(r)} \quad (21)$$

On α és el factor d'acceleració. A mida que aquest factor disminueix la facilitat per sobrepassar la barrera d'energia augmenten, i s'incrementa així, les transicions entre els diferents estats de baixa energia.

Finalment l'energia de referència i el factor d'acceleració són dos paràmetres que s'obtenen a partir de l'energia de diedre del sistema clàssic prèviament simulat i una sèrie de paràmetres determinats, de manera que:

$$E = V_{dihed_avg} + a_1 * N_{res} \quad (22)$$

$$\alpha = \frac{a_2 * N_{res}}{5} \quad (23)$$

On:

- N_{res} = Número de residus de la proteïna.
- V_{dihed_avg} = Energia mitja de diedre calculada a partir de les simulacions clàssiques.
- a_1, a_2 = Coeficients.

Els valors d'aquests coeficients a_1 i a_2 han estat proporcionats per Miao[24] després d'un estudi previ i s'han escollit els que proporcionaven millor resultats.

4. Mètodes d'anàlisi Clàssica

Un cop descrits els diferents processos emprats per simular els dos sistemes que formen part del projecte, és moment de veure i entendre els diferents mètodes que s'han fet servir per analitzar tots aquets resultats extrets per tal de poder observar amb informació concreta l'evolució dels dos sistemes d'anticossos al llarg del temps.

Per tal de poder realitzar un correcte anàlisis dels sistemes s'han utilitzat les diferents tècniques presentades a continuació. És important remarcar que per aplicar la majoria dels següents anàlisis s'ha empleat l'eina CPPTRAJ[65] del software d'Amber18. CPPTRAJ és el programa principal en Amber per processar trajectòries de coordenades i fitxers d'informació de simulacions.

4.1. RMSD

L'estadístic RMSD (*Root Mean Square Deviation*) és un indicador que permet mesurar les dades experimentals obtingudes d'un sistema amb els valors teòrics. En el cas de les simulacions MD ens mesura la suma de les diferències per a cada àtom del sistema, respecte a una estructura de referència, al llarg de tota la simulació i fent la mitjana pel número d'àtoms presents en el sistema.

Aquesta tècnica ens mostra la convergència i estabilitat del sistema que es pot apreciar quan el valor RMSD es manté constant i la tendència del mateix observant el seu creixement o disminució al llarg del temps. Per evitar les interferències de la superfície i el solvent, format per molècules aigua, del sistema s'ha d'especificar els residus sobre els que s'aplicaran aquests anàlisis de manera que només es calculin les interaccions a la proteïna, que al final és la part del sistema de la qual volem observar la seva evolució.

D'aquesta manera, el càlcul és realitza mitjançant la següent equació:

$$RMSD(t) = \sqrt{\frac{1}{N} \cdot \sum_{i=0}^N [(r_i(t) - r_i^{ref})^2]} \quad (24)$$

On:

- N , nombre d'àtoms del sistema.
- r_i , posició de la partícula i .
- r_i^{ref} , posició de referència de la partícula i .

4.2. RMSF

L'estadístic RMSF (Root Mean Square Fluctuation) és un indicador que mesura la desviació estàndard del RMSD d'una partícula. Amb aquesta tècnica s'analitzen les zones de l'estructura amb major i menor fluctuació durant una trajectòria i s'analitza en funció de cada residu de la proteïna anticòs.

Mitjançant aquesta tècnica es pot obtenir una anàlisi sobre l'estructura del sistema on s'observa, per residus, quines regions són més flexibles i amb més mobilitat, propietats necessàries per determinades regions amb una activitat específica de la proteïna i quines presenten major rigidesa i per tant major estabilitat.

Per aquesta anàlisi s'aplica la següent equació:

$$RMSF_i = \sqrt{\frac{1}{T} \cdot \sum_{t_j=1}^T [(r_i(t_j) - \bar{r}_i)^2]} \quad (25)$$

On:

- T , temps total de simulació.
- r_i , posició de l'àtom i .
- \bar{r}_i , posició mitja de l'àtom i en el temps T .

D'aquesta manera s'obté el valor RMSF per cada àtom, com el que interessa és el valor per residu s'aplica:

$$RMSF_{res} = \sum_{i=0}^A \frac{RMSF_i \cdot m_i}{M} \quad (26)$$

On:

- A , nombre d'àtoms corresponents a cada residu.
- $RMSF_i$, valor obtingut per l'àtom i .
- m_i , massa de l'àtom i .
- M , massa total del residu corresponent.

4.3. Angles

El càlcul d'angles (en graus [°]) s'obté a partir de tres punts del sistema escollits prèviament i ens permet veure com evoluciona l'angle que formen els àtoms escollits al llarg de cada *frame* de la trajectòria. És un tipus d'anàlisi que ens permet obtenir informació estructural de diferents regions d'un sistema durant tot el procés de simulació de manera que permet donar una idea clara sobre el comportament d'aquest.

Per als nostres sistemes s'han calculat mitjançant l'eina de CPPTRAJ, tres angles diferents amb tres punts determinats que queden recollits a continuació. Per entendre de forma més clara com es distribueixen els diferents angles i punts escollits en l'espai, en la *Figura 17*, es pot veure mitjançant un dibuix esquemàtic del sistema la localització d'aquests angles.

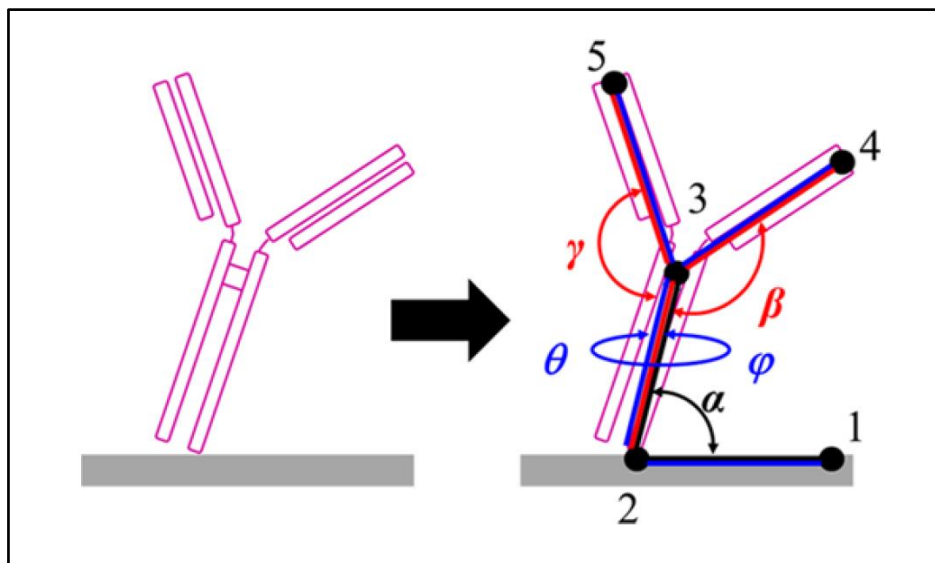


Figura 17: Esquema representatiu d'una estructura proteínica IgG sobre una superfície indicant els angles, diedres i punts per als càlculs. Font: [66]

En aquesta imatge es pot veure l'estructura esquemàtica d'un sistema IgG del mateix tipus que formen els anticossos del projecte. Aquest primer angle α és el que forma la superfície amb la base de la proteïna i s'anomena *Tilt Angle*. És una variable molt important a tenir en compte ja que proporciona informació referent a la facilitat de la proteïna per tombar-se i caure sobre la superfície. En la *Taula 1* podem veure quina són aquells àtoms escollits per cada sistema per aquesta anàlisi en concret. En referència a la nomenclatura d'aquestes taules, quan és parla de *CA*, ens referim sempre al carboni alfa del residu en qüestió. De la mateixa manera, en la *Figura 17*, es troben representats els angles que formen cadascuna de les branques Fab amb la regió constant de la proteïna; la Fc i la informació específica sobre els punts escollits en cada cas es troben en les *Taules 2* i *3*. Els residus concrets que formen els diferents CDR es troben en l'apartat *Annex.A*.

Taula 1: Angle Fc (α). Angle obtingut entre la superfície d'or i la regió Fc.

Punt	Tipus d'àtom	CR3022 ID residu	S309 ID residu	Info. Punt
1.	Au	Au@24959	Au@24959	Àtom d'or de la superfície
2.	CA	K457	K457	Residu que enllaça amb la superfície
3.	CA	P242	P228	Residu de la zona superior de la Fc

Taula 2: Angle Fab1 (β). Angle format entre la regió Fc i la regió Fab1.

Punt	Tipus d'àtom	CR3022 ID residu	S309 ID residu	Info. Punt
2.	CA	K457	K457	Residu que enllaça amb la superfície
3.	CA	P242	P228	Residu de la zona superior de la Fc
4.	Centre de masses	<i>CDR</i>	<i>CDR</i>	Centre de massa dels CDRs de la regió Fab1

Taula 3: Angle Fab2 (γ). Angle format entre la regió Fc i la regió Fab2.

Punt	Tipus d'àtom	CR3022 ID residu	S309 ID residu	Info. Punt
2.	CA	K457	K457	Residu que enllaça amb la superfície
3.	CA	P242	P228	Residu de la zona superior de la Fc
5.	Centre de masses	<i>CDR</i>	<i>CDR</i>	Centre de massa dels CDRs de la regió Fab2

4.4. Diedres

Igual que en el cas dels angles, el càlcul dels diedres (en graus [°]) permet realitzar una anàlisi sobre l'estructura principal del sistema i veure com varien al llarg de tota la simulació les diferents regions d'interès escollides. S'obté a partir de quatre punts diferents del sistema corresponents a uns àtoms escollits prèviament i ens permet veure com evoluciona l'angle de l'últim d'aquests punts respecte el pla que formen els tres primers al llarg de cada *frame* de la trajectòria.

En aquest cas s'han calculat dos diedres i cadascun d'ells està format per quatre punts recollits en les *Taules 4 i 5*, i representats gràficament en la *Figura 17*.

Taula 4: Diedre Fab1(ϕ). Diedre obtingut entre la superfície d'or, la regió Fc i el punt de centre de masses dels CDRs de la regió Fab1.

Punt	Tipus d'àtom	CR3022 ID residu	S309 ID residu	Info. Punt
1.	Au	Au@24959	Au@24959	Àtom d'or de la superfície
2.	CA	K457	K457	Residu que enllaça amb la superfície
3.	CA	P242	P228	Residu de la zona superior de la Fc
4.	Centre de masses	CDR	CDR	Centre de massa dels CDRs de la regió Fab1

Taula 5: Diedre Fab2(θ). Diedre obtingut entre la superfície d'or, la regió Fc i el punt de centre de masses dels CDRs de la regió Fab2.

Punt	Tipus d'àtom	CR3022 ID residu	S309 ID residu	Info. Punt
1.	Au	Au@24959	Au@24959	Àtom d'or de la superfície
2.	CA	K457	K457	Residu que enllaça amb la superfície
3.	CA	P242	P228	Residu de la zona superior de la Fc
5.	Centre de masses	CDR	CDR	Centre de massa dels CDRs de la regió Fab2

4.5. Distàncies

El càlcul de distàncies (en Àngstroms [Å]) ens permet obtenir informació sobre com s'aixafa la proteïna i cadascuna de les seves regions, al llarg del temps. Per això es calcula l'evolució de la distància entre dos punts. El primer s'ha agafat com el centre de masses de la regió en qüestió i el segon situat a la superfície d'or. En aquest cas en concret ens interessa observar l'evolució de la distància únicament en l'eix Z.

D'aquesta manera s'han obtingut tres càlculs diferents, cadascun amb els seus punts:

Taula 6: Distància Fc. Distància en l'eix Z, entre el centre de masses de la regió Fc del sistema i la superfície d'or.

Tipus d'àtom	CR3022 ID residu	S309 ID residu	Info. Punt
Centre de masses (CA)	C230-K457, C230-S451	C216-K443,G222-S430	Centre de massa de la Fc
Au	Au@19992	Au@19633	Àtom d'or de la superfície

Taula 7: Distància Fab1. Distància en l'eix Z, entre el centre de masses de la regió Fab1 del sistema i la superfície d'or.

Tipus d'àtom	CR3022 ID residu	S309 ID residu	Info. Punt
Centre de masses (CA)	T0-S219, D1-C220	V1-S215, I1-G215	Centre de massa de la Fab1
Au	Au@21660	Au@20875	Àtom d'or de la superfície

Taula 8: Distància Fab2. Distància en l'eix Z, entre el centre de masses de la regió Fab2 del sistema i la superfície d'or.

Tipus d'àtom	CR3022 ID residu	S309 ID residu	Info. Punt
Centre de masses	T0-S219, D1-C220	V1-L221, I1-G215	Centre de massa de la Fab2
Au	Au@19614	Au@20316	Àtom d'or de la superfície

4.6. Radis de Gir

En dinàmica molecular el radi de gir és un dels indicadors més rellevants per observar la compactació o esponjament de l'estructura de determinades zones de la molècula. Es calcula mesurant la distància entre el centre de masses de la proteïna (o la zona en qüestió) i els seus extrems de manera que:

$$R_g = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N |r_i - r_{center}|^2} \quad (27)$$

On:

- N , número d'àtoms de la proteïna.
- r_i , posició de l'àtom i .
- r_{center} , posició al centre de masses.

Si la proteïna es manté correctament plegada, el valor del radi de gir es mantindrà relativament estable. Mentre que si l'estructura es desordena el RoG fluctuarà molt més al llarg del temps.

Per aquest motiu, en aquesta anàlisi s'ha determinat el radi de gir de diferents zones conjuntes i per separat de la proteïna, de manera que es pugui explorar i analitzar cadascuna de les regions amb una funcionalitat rellevant de la proteïna. En la *Taula 9*, queda recollida la informació referent a la zona que s'ha analitzat amb cadascuna de les variables exposades.

Taula 9: Recull de les regions estudiades i nom de les variables donades.

Nº	Variable	Informació respecte la regió
1.	Rgyr Protein	Radi de gir de tota la proteïna.
2.	Rgyr Fab1	Conjunt de tota la branca Fab1
3.	Rgyr Fab1-H	Conjunt de CDRs de la cadena pesada de la Fab1
4.	Rgyr Fab1-L	Conjunt de CDRs de la cadena lleugera de la Fab1
5.	Rgyr Fab1-HL	CDRs en conjunt de les dues cadenes (H i L) de la Fab1
6.	Rgyr Fab2	Conjunt de tota la branca Fab2
7.	Rgyr Fab2-H	Conjunt de CDRs de la cadena pesada de la Fab2
8.	Rgyr Fab2-L	Conjunt de CDRs de la cadena lleugera de la Fab2
9.	Rgyr Fab2-HL	CDRs en conjunt de les dues cadenes (H i L) de la Fab2
10.	Rgyr Fc	Conjunt de tota la branca Fab2
11.	Rgyr Fc1-H	Conjunt de CDRs de la cadena pesada 1 (CH1) de la Fc
12.	Rgyr Fc2-H	Conjunt de CDRs de la cadena pesada 2 (CH2) de la Fc

5. Mètodes d'anàlisi aMD

Les simulacions aMD permeten explorar els sistemes de forma més amplia i ràpida que els mètodes clàssics, però aquest increment pot alterar de forma important les propietats estructurals de la proteïna dels sistemes, de manera que el primer conjunt d'anàlisis que s'aplicarà a aquestes simulacions serà el mateix que en el cas de les simulacions clàssiques. Per altre banda però, el fet d'aquesta anàlisi més exhaustiva del sistema ens permet obtenir informació més rellevant sobre els diferents mètodes energètics que es presenten a continuació.

5.1. Potencial Mean Force

Al examinar un sistema generat mitjançant eines computacionals, és interessant saber com varia l'energia lliure en funció d'alguna coordenada intermolecular o intramolecular (com la distància entre dos àtoms o un angle de torsió). La superfície d'energia lliure al llarg de la coordenada es denomina *Potencial Mean Force* (PMF). El PMF es pot obtenir en simulacions de Monte Carlo o de dinàmica molecular i, habitualment, aquest mètode es fa servir en conjunt amb *Umbrella Sampling* per poder mostrejar adequadament l'espai del sistema a mida que avança.

En aquest projecte, l'estudi de la PMF s'ha fet sobre el conjunt de simulacions aMD obtingudes de cada sistema. Al tenir en compte diverses produccions de cada sistema, la distribució de Boltzmann que s'obté amb l'ensamblatge canònic de la simulació no correspon a un únic valor. Per aquest motiu s'ha d'aplicar una modificació per tal de graduar la probabilitat en que es dona cadascuna d'aquestes produccions que conformen el sistema aMD total simulat.

A partir de la metodologia seguida per Miao [24] es defineix la distribució de probabilitat al llarg d'una coordenada de reacció seleccionada $A(r)$ escrita com a $p'(A)$, on r determina la posició dels àtoms $\{r_1, \dots, r_n\}$. D'aquesta manera donat un potencial $\Delta V(r)$ de cada *frame*, $p'(A)$ es pot modificar de manera que s'obtingui a la distribució canònica, $p(A)$, com:

$$p(A_j) = p'(A_j) \cdot \frac{\langle e^{\beta \Delta V(r)} \rangle_j}{\sum_{j=1}^M \langle e^{\beta \Delta V(r)} \rangle_j}, \quad j = 1, \dots, M \quad (28)$$

On:

- M , nombre de produccions.
- $\langle e^{\beta\Delta V(r)} \rangle_j$, factor mitjà de Boltzmann per $\Delta V(r)$ pels *frames* dins de j .

L'equació proposada proporciona un algoritme pels nous pesos de les simulacions aMD. El PMF calculat a partir d'aquesta fórmula s'obté tal que:

$$F(A_j) = -\frac{1}{\beta} \ln(p(A_j)) \quad (29)$$

Degut a que el factor de Boltzmann normalment s'obté a partir de zones pobrament mostrejades, el càlcul de mitjana potencial en el que es basa sol portar a grans fluctuacions d'energia. Per reduir aquest soroll s'ha aproximat el terme exponencial a partir del sumatori de Maclaurin de manera que:

$$\langle e^{\beta\Delta V} \rangle = \sum_{k=0}^{\infty} \frac{\beta^k}{k!} \langle \Delta V^k \rangle \quad (30)$$

D'aquesta manera es suprimeix el terme j però ens podem trobar en què en algun cas, la posició dels mínims d'energia del sistema es desplaça en comparació a la seva posició original. En el mateix article anteriorment mencionat, es proposen altres mètodes per suplir aquestes diferents aproximacions, però en aquestes simulacions no s'aprofundirà més.

5.2. Free Binding Energy

La *Free Binding Energy* és una funció que permet quantificar l'energia lliure associada a diferents estats de solvatació entre dues molècules. Els càlculs de l'energia lliure han demostrat ser molt útils per un gran nombre de aplicacions en química computacional com el disseny de medicaments i la determinació de l'estructura de proteïnes. La FBE és una funció que proporciona informació sobre la quantitat de superfície d'una proteïna, l'anticòs en el nostre cas, que es troba dins del complex i aïllat del solvent del sistema i en contacte amb la superfície.

Aquest estudi de la FBE s'ha calculat a partir del *software* MMPBSA que forma part d'Amber18. Aquest sistema calcula l'energia lliure d'enllaç del sistema a partir de les energies lliures del receptor i lligand sense enllaçar i del complexa enllaçat. Per al sistema del projecte, quan parlem del complexa, ens referim al conjunt global de la proteïna i la superfície d'or, mentre que el receptor i lligand corresponen respectivament a la proteïna per una banda i a la superfície per l'altra [67]. D'aquesta manera obtenim que:

$$\Delta G_{bind,solvated} = \Delta G_{complex,solvated} - (\Delta G_{receptor,solvated} + \Delta G_{ligand,solvated}) \quad (31)$$

On cadascuna de les energies lliures associades a les diferents parts del sistema que es calculen, $\Delta G_{solvated}$, es fa mitjançant el promig de les contribucions de l'energia en fase gas, l'energia lliure de solvatació i l'entropia del solut per cada pas del sistema.

$$\Delta G_{solvated} = E_{i,gas} + \Delta G_{i,solvation} - T \cdot S_{i,solute} \quad (32)$$

On i correspon a cadascun dels *frames* de les simulacions aMD emprats per càlcul. Finalment, per estimar el total de l'energia lliure de solvatació de la molècula, s'assumeix que aquesta es pot dividir en dues parts, les energies electrostàtiques i les no electrostàtiques, tal que:

$$\Delta G_{solvation} = \Delta G_{elec} + \Delta G_{surf} \quad (33)$$

On ΔG_{surf} correspon a l'energia lliure de solvatació de la molècula on s'han eliminat totes les càrregues dels àtoms i ΔG_{elec} és l'energia lliure obtinguda de treure prèviament totes les càrregues en el buit i després afegir-les en presència del solvent.

És important tenir en compte que el cost computacional per al càlcul de ΔG_{elec} és molt elevat i per això habitualment es simplifica mitjançant el mètode de càlcul numèric de *Poisson-Boltzmann* (PB). Però degut a les grans dimensions dels sistemes i als ordres de magnituds que això implica, s'ha optat per un mètode diferent, la aproximació *Generalized Born* (GB), pel càlcul de les energies electrostàtiques del sistema. Aquesta metodologia s'ha tornat cada vegada més utilitzada en simulacions MD degut a la seva senzillesa i eficiència computacional en comparació amb el mètode més estàndard de PB[68].

6. Metodologia de Simulació

A partir de les necessitats de cada sistema i tenint en compte els diversos protocols de simulació explicats amb anterioritat, en aquest punt es descriu el procediment seguit per realitzar tot el conjunt de simulacions necessàries per obtenir tot el conjunt de resultats per l'estudi proposat en el projecte.

El procediment seguit pels dos sistemes del projecte, CR3022 i S309 és exactament el mateix per tal d'obtenir uns resultats el més representatius possibles, tant per les simulacions cMD com les aMD.

6.1. Simulacions Clàssiques:

Un cop s'ha generat el sistema amb la proteïna, la superfície i el medi de solvatació amb el corresponent fitxer de paràmetres, el primer pas en les simulacions per tal d'aconseguir una estabilitat en el sistema és fixar tant la proteïna com la superfície i deixar únicament lliure les molècules d'aigua que envolten el sistema. D'aquesta manera s'aconsegueix portar el sistema a condicions estàndards de pressió i temperatura sense alterar la conformació inicial de la proteïna reduint així els riscos de que es produeixi una desnaturalització de la proteïna i eliminant interaccions no favorables que puguin aparèixer entre el solut i el solvent [69]. En aquesta etapa inicial, primerament s'ha fet una minimització per establir l'estabilitat del sistema i tot seguit s'ha realitzat un procés d'escalfament o *heating* per passar dels 0K inicials a una temperatura de 298K tal com s'especifica en el protocol corresponent. Aquesta primera etapa té una duració de 0.052ns i ve seguida d'un procés d'equilibrat NPT que es manté uns 20ns i permet estabilitzar la pressió del sistema a 1.013 bar.

Un cop acabada aquesta primera etapa d'equilibrat on s'ha portat l'aigua del sistema a condicions estàndards, es procedeix a alliberar la proteïna. A continuació, s'ha de tornar a fer un procés de minimització i escalfament del sistema per portar el sistema a les condicions òptimes per poder córrer l'equilibrat final de NVT de 150ns. Aquest últim procés de simulació ens proporcionarà tota la informació necessària per al posterior llançament d'aMD i a la posterior anàlisi clàssica del sistema.

6.2. Simulacions aMD:

Com ja vist en l'apartat 4.2, l'objectiu de l'anàlisi aMD és explorar diferents conformacions que en condicions estàndards no seria factible observar, ja sigui pel cost de temps o pel cost computacional. En aquest cas, seguint l'exemple proporcionat per Miao [24], [70] s'han realitzat quatre produccions de 90ns cadascuna que permeten explorar de forma molt més amplia el sistema; partint de diferents conformacions del tram final (Simulacions NVT dels 75ns als 150ns) d'aquesta primera part de simulacions en cMD obtenim en acabar, una producció total de 360ns de aMD per cada sistema. I per fer-ho, prèviament s'han hagut de calcular una sèrie de paràmetres a partir de les simulacions cMD que queden recollits en la *Taula 10*.

Taula 10: Paràmetres específics emprats en les simulacions aMD.

CR3022		S309	
N_{res}	1332	N_{res}	1327
V_{dihed_avg}	14.645,48	V_{dihed_avg}	14.540,55
a_1, a_2	4	a_1, a_2	4
E_D	19.973,48	E_D	19.848,55
α	1065.6	α	1061.6

Finalment a la *Taula 11*, es poden observar les dades referents al punt de partida de cada producció escollida pels diferents sistemes. Aquests valors han estat escollits a partir de la visualització gràfica del RMSD respecte el temps recollida en el *AnnexB*.

Taula 11: Time step i Frame de partida de cadascuna de les conformacions de partida de les simulacions aMD dels sistemes CR3022 i S309.

CR3022		S309	
Nº Prod.	Time_Step	Nº Prod.	Time_Step
1.	90ns	1.	90ns
2.	105ns	2.	100ns
3.	120ns	3.	110ns
4.	140ns	4.	120ns

7. Resultats

Un cop realitzades totes les simulacions, s'han tractat i analitzat tots els resultats amb l'objectiu principal de comprovar l'estabilitat estructural dels diferents sistemes dissenyats per així mantenir l'especificitat d'aquelles zones que hauran d'interactuar amb la regió concreta de la RBD de l'espícula del virus.

Per desenvolupar aquestes anàlisis de manera clara i estructurada, primerament comprovarem quina evolució ha seguit cadascun dels sistemes al llarg del temps durant les simulacions clàssiques. D'aquesta manera ens farem una idea inicial de que l'objectiu de mantenir una estructura proteínica damunt d'una superfície d'or en un sistema solvatat d'aigua es manté estable durant un cert període de temps determinat de simulació. Tot seguit, es veurà de forma més amplia que és el què succeeix en les diferents regions del sistema (Fab1, Fab2 i Fc) per veure amb més profunditat si es compleixen els requisits estructurals i de funcionalitats requerits per poder portar la idea d'aquest projecte al següent nivell. Finalment, s'han comparat els resultats obtinguts pels dos sistemes i s'ha fet una valoració dels comportaments observat en cadascun d'ells per tal de poder determinar quin es creu que podria aportar millors resultats.

7.1. Estudi estructural del sistema

En primer lloc, analitzarem l'evolució dels dos sistemes de manera visual a l'inici de la simulació i al final, per veure com s'ha transformat la proteïna al llarg d'aquest període i entendre l'evolució dels diferents paràmetres i variables estudiats més endavant. Els resultats obtinguts pel sistema CR3022 queden recollits en la *Figura 18*, mentre que pel sistema S309 es troba en la *Figura 19*. Aquestes primeres imatges ens permeten veure un comportament en la proteïna que ja preveïem. A mida que ha passat el temps, s'ha anat produint un procés d'aixafament on la proteïna ha anat perdent altura i l'estructura s'ha alterat mínimament.

En els dos sistemes es pot observar el mateix comportament respecte a la zona Fc de la proteïna. A mida que passa el temps de simulació s'aixafa de forma notable sobre la superfície. Concretament és una de les cadenes pesades la que entra en contacte amb la superfície mentre que l'altre es manté, però les regions CH2 i CH3 s'obren i formen un angle més agut. Per altra banda, les regions Fabs del sistema sembla que es mantenen més estables i encara que canviï la seva forma en l'espai, l'estructura principal sembla que no s'altera, al menys de manera que es pugi apreciar significativament de forma visual.

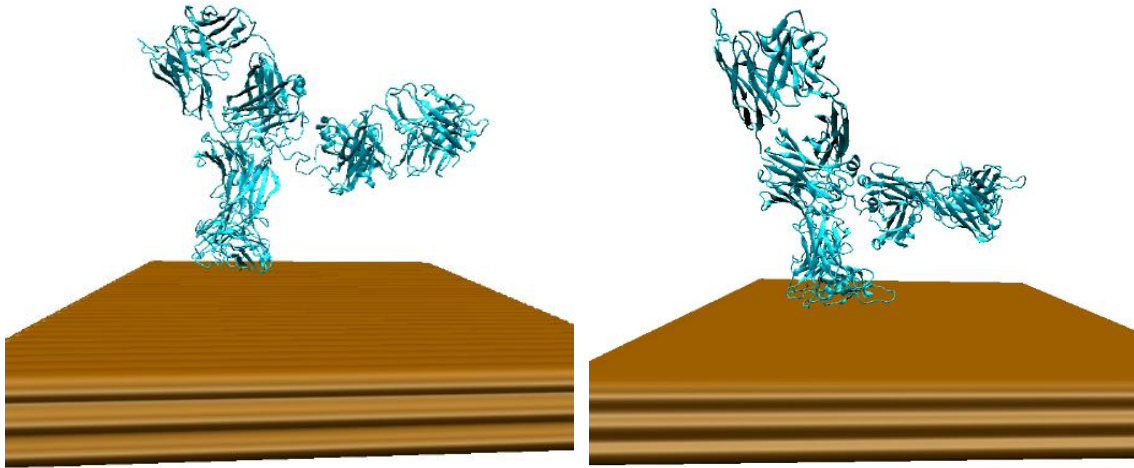


Figura 18: Representació mitjançant VMD del sistema CR3022. En la figura de l'esquerra es troba el sistema cristal·logràfic inicial i a la dreta el sistema al final del procés clàssic (després de 140ns de simulació).

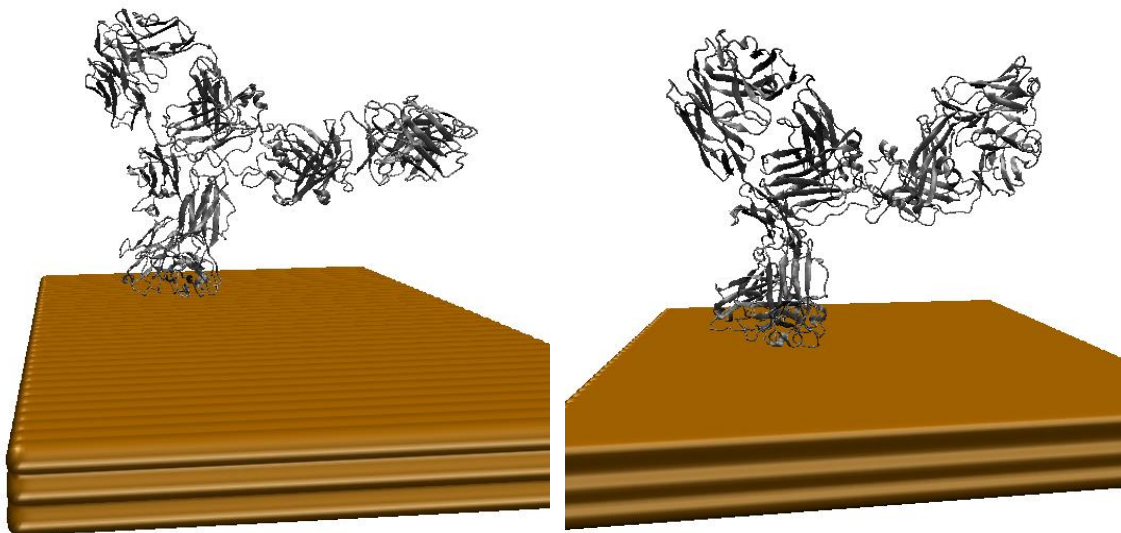


Figura 19: Representació mitjançant VMD del sistema S309. En la figura de l'esquerra es troba el sistema cristal·logràfic inicial i a la dreta el sistema al final del procés clàssic (després de 140ns de simulació).

A continuació, per entendre millor el comportament durant tota la simulació clàssica que dona lloc al sistema final observat, veurem una comparativa de les dades de forma i estructura que es van obtenir pel sistema cristal·logràfic inicial, és a dir, el sistema de partida, abans d'iniciar cap tipus de protocol de simulació que alterés cap variable d'aquest sistema i veurem com varien al llarg de la simulació clàssica, que és aquella realitzada de manera contínua i sense alterar dràsticament el sistema. La informació referent a les diferents variables i els punts emprats per al seu càlcul queden recollits en l'apartat 5 del treball.

Els resultats obtinguts queden recollits en les taules que es troben a continuació. En aquestes taules s'indiquen el valor cristal·logràfic de cadascuna de les variables, el valor mig dels últims 75ns de la simulació clàssica i l'increment absolut entre els dos valors. Si ens centrem en el sistema de l'anticòs CR3022 i analitzem les dades proporcionades per la *Taula 12*, podem veure que succeeix amb aquest sistema. Primerament, es pot veure com la distància global del centre de masses del sistema respecte la superfície d'or fixada (Distància Prot), varia i es redueix aproximadament 17\AA degut a les forces exercides pel sistema durant el procés de simulació, i de forma més concreta les diferents regions Fc, Fab1 i Fab2 es redueixen respectivament 11\AA , 13\AA i 1.6\AA . En aquest punt ens trobem amb un fenomen estrany que cal tenir en consideració. Les regions Fc i Fab1 del sistema disminueixen de manera constant i similar a la seva distància, que es la situació que cabria esperar degut a les forces a les que està sotmès el sistema durant la simulació, però en canvi, la branca de la Fab2 és manté pràcticament a la mateixa distància que en el seu estat inicial. El principal motiu d'aquest fet és que en la proteïna no es produeix únicament un procés en que s'aixafi, sinó que també, en estar lligada a la superfície per un únic punt, també gira i s'inclina tot el sistema en favor d'una de les branques, la Fab1.

Per justificar aquest fenomen, s'ha de veure que succeeix amb els angles i diedres del sistema. Si la proteïna es mantingues completament recte, l'angle que forma la regió lligada a la superfície, Fc, hauria de forma un angle estable de 90° , però com es pot observar aquest angle inicialment és de 71.6° i es redueix un total de 10° . Per altre banda, l'angle de la Fab1 també es redueix al voltant de 21.4° , fet que justifica aquesta reducció de la distància respecte a la superfície. L'angle que forma la branca Fab1 amb la Fc, en canvi, pateix un increment de 37° que fa que, malgrat que el conjunt global de la proteïna s'inclini i s'aixafi una mica, pel fet de compensar el sistema i mantenir una estructura estable amb la forma de Y original de la proteïna, aquesta regió no es vegi pràcticament alterada.

Finalment, és important comprovar que aquests canvis globals del sistema no alterin la conformació i especificitat de cadascuna de les regions mencionades ja que reduiria la funcionalitat del sistema. Més endavant s'estudiarà el comportament d'aquelles parts concretes que ens interessarà aprofundir més, però de forma més general ens centrarem en els radis de gir d'aquestes regions. Com es pot observar, el radi de gir de tota la proteïna es redueix al voltant de 4\AA , però es normal degut al moviment d'angles i distàncies que es produeix durant la simulació. El que si és rellevant i de vital importància es veure com en les tres regions del sistema es produeixen una variació del seu radi de gir menor de 0.5\AA . Una variació alta d'aquesta variable pot implicar que l'estructura es vegi alterada, per tant, el fet de que no es modifiqui al llarg de tot el període clàssic de simulació implica que les regions importants del sistema es mantenen estables i correctament plegades en el temps.

Taula 12: Variables més rellevants analitzades pel sistema inicial i clàssic del sistema CR3022.

Tècnica	Regió	Cristall	Final (75-150ns)	Increment Observat
Angles (°)	Fc (<i>Tilt Angle</i>)	71,6	61,6	10,0
	Fab1	113,9	92,5	21,4
	Fab2	121,4	159,7	38,3
Diedre (°)	Fab1	76,2	78,3	2,1
	Fab2	-119,5	-99,1	20,4
Distància (Å)	Prot.	77,595	60,387	17,208
	Fc	41,450	30,575	10,874
	Fab1	69,685	56,915	12,770
	Fab2	91,806	89,145	2,661
Radi de Gir (Å)	Prot.	48,523	44,205	4,318
	Fab1	24,696	24,284	0,412
	Fab2	24,696	24,279	0,417
	Fc	26,778	26,499	0,278

En el cas del sistema del S309 com ja es pot observar a partir de les representacions obtingudes mitjançant VMD, el comportament és una mica diferent ja que l'estructura global del sistema no varia de la mateixa manera que en el cas del CR3022. És per això que mitjançant les dades proporcionades per la *Taula 13* aprofundirem de forma més estructurada en l'evolució que ha portat aquest sistema. Començarem observant com s'ha aixafat aquesta proteïna en relació a la superfície de la mateixa forma que s'ha fet pel primer sistema. En aquest cas, el conjunt total del centre de masses de la proteïna s'ha reduït menys, aproximadament 9 Å. I si observem aquesta distància per cadascuna de les regions principals de la proteïna Fc, Fab1 i Fab2 obtenim que la distància s'ha reduït respectivament 11 Å, 1 Å i 5 Å. És a dir, que regió Fc, en contacte amb la superfície s'ha caigut de manera semblant al cas del CR3022, però en canvi les dues branques del sistema presenten un comportament una mica diferent. Les dues es mantenen molt més constants, si és cert que en aquest cas, en la Fab2 es redueix més la distància i és la branca Fab1 la que pràcticament no varia, però pel conjunt global de simulació aquesta variació es poc significativa.

Per això, tornarem a recórrer als angles i diedres per trobar una justificació a aquest comportament. Si ens fixem en el *Tilt Angle* veurem que igual que en el CR3022, apart d'aixafar-se, la proteïna també es tomba una mica, més concretament es passa d'un angle de 72° a un final de 52.6°. En canvi, si observem els angles que formen les regions Fab1 i Fab2, cap dels dos varia més de 5° en tot el procés, és a dir, que a diferència del sistema anterior, en l'estructura de forma de Y pràcticament no es produeix ninguna rotació. Ara bé, d'aquesta manera no podem acabar de justificar perquè les distàncies respecte a la superfície pràcticament no canvien. Per fer-ho, hem de veure com varien els diedres que es formen a partir de la superfície, la regió Fc i cadascuna de les regions Fab. L'increment de 35° en el cas de la Fab1 i de 20° en el cas de la Fab2 contraresten el procés d'aixafament

de la proteïna, de manera que quan més elevat sigui aquest diedre més petita serà la reducció de la distància respecte a la superfície d'or del sistema.

Finalment com calia esperar, el radi de gir del conjunt global de la proteïna varia menys pel sistema S309 que en el cas del CR3022, únicament 1 Å, degut a que l'estructura global ha patit menys canvis. Al contrari, els radis de girs per regions són tots una mica més alts, al voltant de 1 Å en les branques i de 2 Å per la zona Fc, de manera que s'haurà d'estudiar amb més profunditat si l'estructura secundària d'aquesta zona s'ha vist alterada en algun cas.

Taula 13: Variables més rellevants analitzades pel sistema inicial i clàssic del sistema S309.

Tècnica	Anàlisi	Cristall	Final (75-150ns)	Increment obtingut
Angles (°)	Surf_Fc	71,9	52,7	19,3
	Fc_Fab1	114,1	110,3	3,8
	Fc_Fab2	122,1	117,5	4,6
Diedre (°)	dihed_Fab1	74,9	109,2	34,3
	dihed_Fab2	-119,9	-140,013	20,1
Distància (Å)	Surf_Prot	75,563	66,889	14,467
	Surf_Fc	39,995	28,952	8,674
	Surf_Fab1	73,279	72,719	11,825
	Surf_Fab2	93,268	88,698	10,162
Radi de Gir (Å)	Rgyr_1	49,263	48,397	0,866
	Rgyr_Fab1	25,760	24,516	1,244
	Rgyr_Fab2	25,660	24,618	1,041
	Rgyr_Fc	26,375	28,607	5,232

7.2. Regions de major mobilitat del sistema

L'evolució general de l'estructura del sistema ens aporta informació sobre com es comporta el sistema de manera estructural. Ara bé, per veure la mobilitat de cada regió del sistema s'ha utilitzat l'anàlisi RMSF enfocat en cada residu durant els 150ns que s'han obtingut de simulació clàssica. Les gràfiques on es representen els RMSF de cada sistema es troben a l'apartat *Annex B*.

Per representar el RMSF s'han recollit els valors de cada residu i a partir d'aquests, s'ha generat un mapa de calor aplicat sobre la proteïna cristal·logràfica de partida de cada sistema de manera que es pot observar quines són aquelles regions amb una mobilitat més

elevada i més baixa durant les simulacions. Les regions amb menys mobilitat i, per tant, més rígides de la molècula són aquelles amb un valor RMSF més baix i es marquen en color blau. Per altra banda, les zones amb un valor RMSF elevat indiquen una alta mobilitat dels seus residus en comparació amb la resta del sistema i dins de la imatge estan representades de color vermell. Finalment, les zones de transició amb una mobilitat prou elevada per a tenir en compte, amb un valor de RMSF al voltant de 5 Å estan representades de color gris.

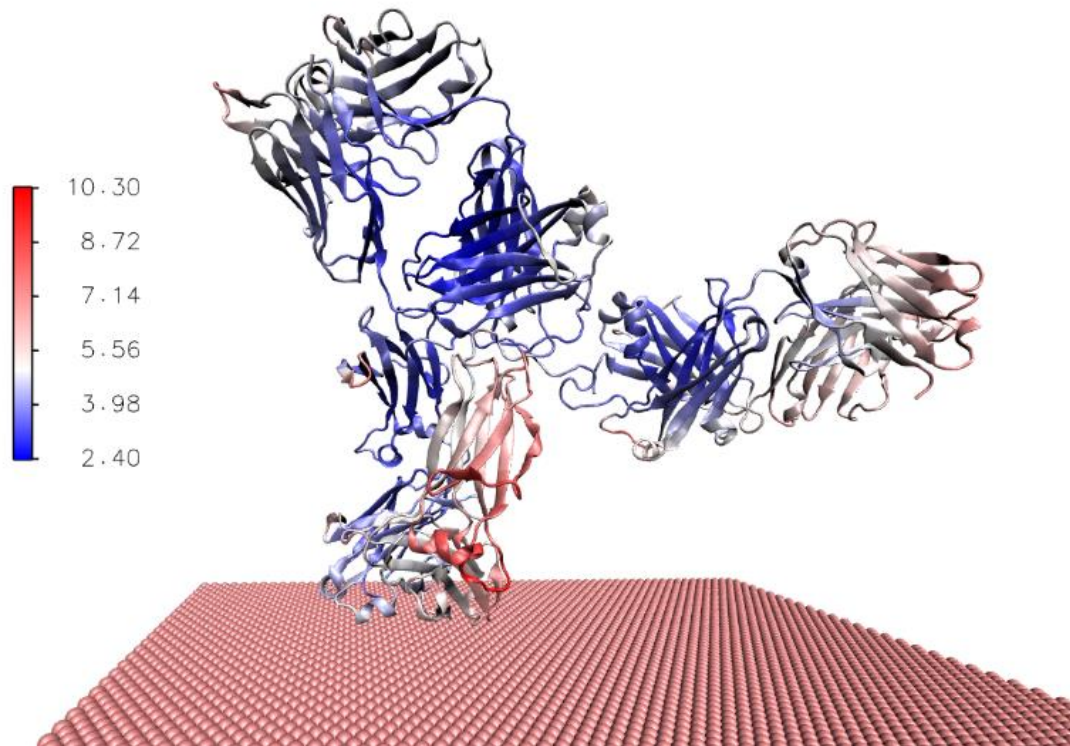


Figura 20: Representació mitjançant VMD d'un mapa de calor del RMSF del sistema CR3022 aplicat sobre l'estructura cristal·lina de la proteïna.

Gràcies a aquesta representació podem observar com pel sistema CR3022, obtingut en la *Figura 20*, les regions contínues de les branques Fab del sistema es mantenen molt rígides i estables, amb una mobilitat menor de 4 Å. Les regions Fv en canvi tenen una mobilitat bastant més notable amb zones on el RMSF pot arribar a assolir valors de fins a 8 Å. Aquesta mobilitat és esperada en aquestes regions degut a que en els extrems, l'exposició a les diferents forces externes i de solvatació genera una major interacció. Per altra banda, en aquestes regions més exposades de la proteïna és on es troben les anomenades CDR. Aquestes regions CDR són les zones concretes de l'anticòs que prèviament s'ha vist que interaccionen directament amb l'antigen. Per tal de permetre una correcta interacció amb regió RBD del virus, és necessari que existeixi aquest nivell de mobilitat, però quan més alt sigui aquest valor de RMSF més probabilitats hi ha de que l'estructura secundària necessària per realitzar la seva funció hagi patit algun tipus de desnaturalització.

Per altra banda, la mobilitat que s'observa en la regió Fc del sistema confirma el comportament observat en l'estudi estructural previ. Una de les dues cadenes pesades que la formen presenta un RMSF prou baix ja que excepte els extrems, que són les zones més exposades, està representada de color blau. Aquesta cadena pesada correspon a aquella en què al llarg de tota la trajectòria únicament variava la forma i l'angle que formaven les zones CH2 i CH3 de la cadena, però l'estructura general es conservava perfectament. La segona cadena pesada en canvi, presenta una mobilitat molt més elevada, amb zones on el valor del RMSF arriba als 10 Å, i amb una major part de la seva representació en la *Figura 20* sent de color vermell. Aquest comportament confirma que és aquesta regió de la Fc la que pateix més interaccions amb la superfície i en conseqüència, es produeix una variació significativa del *tilt angle* que fa que cada cop s'aixafi i es tombi més la proteïna. És interessant comentar que la cadena pesada del sistema que pateix més aquest procés d'aixafament sobre la superfície és la que lliga la proteïna amb la mateixa superfície mitjançant el residu *Linker* que s'ha modificat manualment seguint la metodologia prèviament explicada.

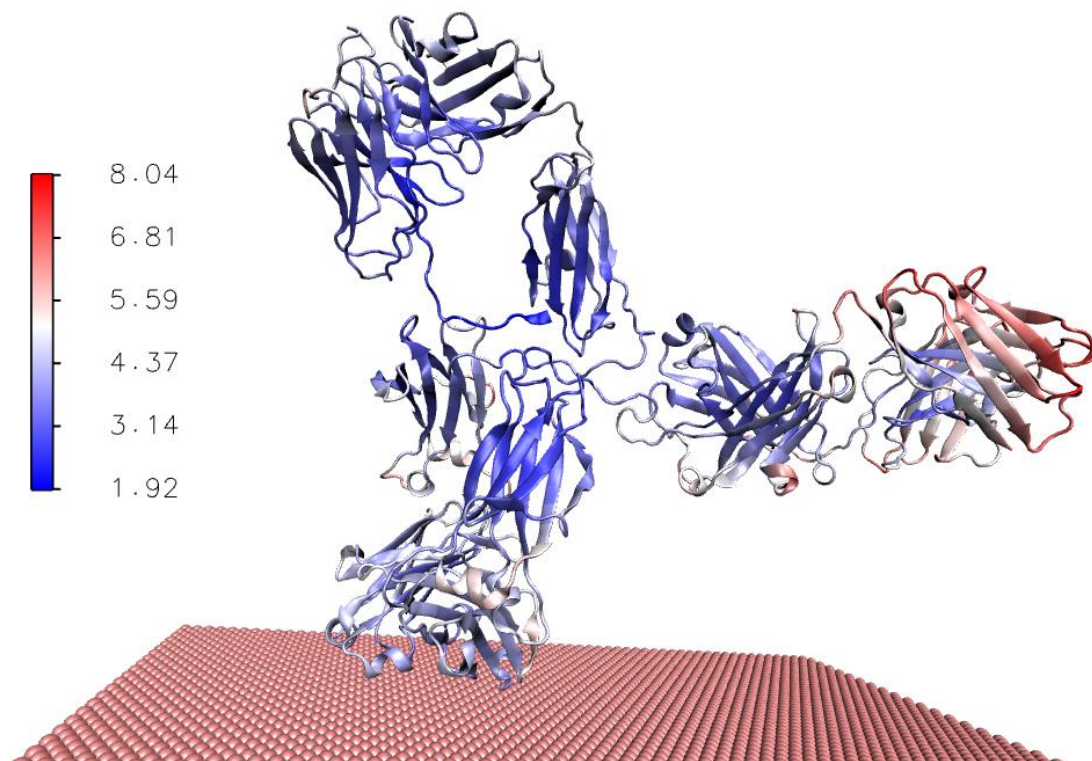


Figura 21: Representació mitjançant VMD d'un mapa de calor del RMSF del sistema S309 aplicat sobre l'estructura cristal·lina de la proteïna.

En el cas del sistema S309 el comportament que es pot observar en la *Figura 21* es similar al observat per al sistema CR3022, però amb unes variacions no tant elevades a simple

vista. En aquest cas tant les regions contínues de les Fabs com les zones CH2 de les cadenes pesades que formen la Fc presenten uns valors de RMSF menors de 4 Å. És a dir, que la estructura al voltant del *hinge* de la proteïna es manté rígida i estable, amb una mobilitat reduïda, fet que permet que els angles i distàncies entre les diferents zones del sistema es mantinguin molt més constants que en altres circumstàncies. Les regions Fv de les branques segueixen tenint una mobilitat elevada i similar al CR3022 amb uns valors de RMSF que van entre 5-8 Å. De la mateixa manera, en aquestes regions, és on es situen les CDR de l'anticòs encarregades de lligar-se i interaccionar amb el virus del SARS-COV-2 i aquesta mobilitat és d'esperar però s'ha d'analitzar, sobre tot en el cas de la regió Fab1 ja que és la que es troba més exposada i on les probabilitats de que l'estructura secundària es vegi alterada són més altes.

En definitiva, com ja hem vist, és en la Fc on els canvis respecte al CR3022 són més significatius. Si bé és cert que una de les cadenes pesades presenta una coloració més gris i amb alguns tons de vermell, el RMSF d'aquesta cadena no arriba a valors tan elevats com en el sistema anterior, aconseguint així que, malgrat que el *Tilt angle* també es vegi reduït, la rotació global de la proteïna i el canvi estructural sigui menys accentuat. D'igual manera que en el primer sistema, aquesta cadena que presenta uns valors de RMSF superiors, és la mateixa que lliga la proteïna amb la superfície a través del residu *Linker* de manera que el fet d'estar interaccionant amb la superfície d'una manera més propera pot ser un dels factors claus en el seu comportament respecte a la cadena pesada homònima del sistema.

7.3. Estudi de les regions Fab del sistema

Les regions Fab de les molècules són aquelles encarregades de detectar i lligar-se a les seves regions específiques del virus. Com ja hem vist, per tal que puguin complir la seva funció immunològica les branques han d'estar orientades de tal forma que les regions que contenen els CDRs específics puguin realitzar la seva funció. Això implica que aquestes regions són les que es trobem més exposades al solvent i al medi en general de tota la proteïna. Per aquest motiu, en aquest apartat dels resultats veurem si mantenen la seva estructura i funcionalitat correctament.

Observar l'evolució del radi de gir de l'estructura és una de les formes més utilitzades i efectives per determinar el comportament de l'estructura d'aquest tipus de regions. D'aquesta manera s'han analitzat cadascuna de les branques de manera més general i més concreta sobre les regions específiques a partir de graficar aquesta variable en funció del PMF del sistema. Aquest estudi s'ha fet a partir de les simulacions aMD ja que és un mètode que permet explorar el conjunt del sistema d'una manera més amplia amb un menor requeriment de temps de computació.

Començant pel sistema CR3022, a la *Figura 22* es pot veure representat el valor del radi de gir resultant de la branca Fab1 sobre l'eix x, en front del PMF en l'eix y. Mitjançant

aquesta anàlisi podem veure com evoluciona l'energia potencial en funció del radi de gir i determinar la regió de mínima energia del sistema. Marcat amb un rombe de color vermell trobem aquest mínim situat al voltant del valor de radi de gir de 24,3 Å. Sobre la gràfica també es troben tres línies de diferent color que aporten informació sobre els resultats obtinguts durant les simulacions de cMD. En verd s'indica el valor inicial de radi de gir de l'estat cristal·logràfic inicial de la proteïna, en lila el valor promig de tota l'anàlisi clàssica i en taronja el promig dels últims 75ns. El fet que aquests tres valors quedin compresos dins del rang explorat per les simulacions aMD i que a l'etapa final de la simulació s'apropi al mínim és un indicador molt bo de l'estabilitat de la regió Fab1 del sistema.

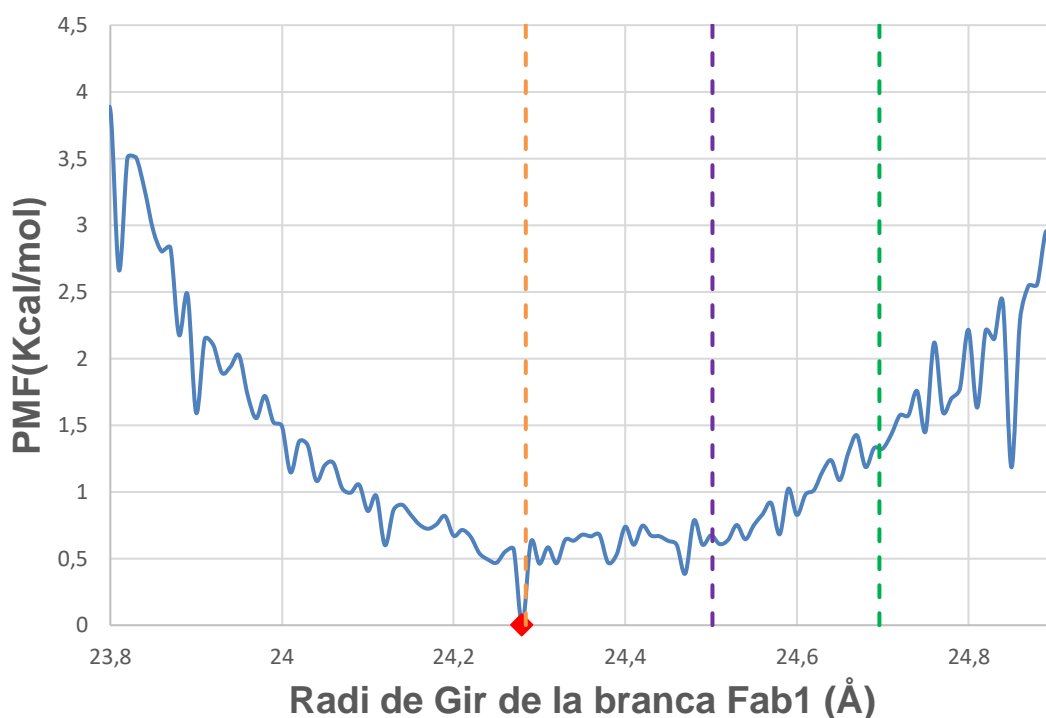


Figura 22: Gràfic representatiu del radi de gir Fab1 en funció al PMF del sistema CR3022. En un rombe vermell s'indica el mínim del sistema, i amb línies discontinuïques de color verd, lila i taronja és marca el valor corresponent al valor cristal·logràfic inicial, al promig global i al promig dels últims 75ns de clàssica respectivament.

Després de veure com globalment es conserva l'estructura i la fluctuació de la regió Fab1 és mínima, s'ha decidit analitzar el comportament específic de les CDR situades als extrems de la branca. Els residus que formen part d'aquesta regió estan dividits en dos parts, aquells que formen part de l'extrem de la cadena pesada de la branca CDR-H (THR32-GLY36, ILE51-GLY67, GLY100-VAL109) i els que pertanyen a la cadena lleugera CDR-L (LYS1135-ALA1151, TRP1167-SER1173, GLN1206-THR1214) i com ja s'ha vist, són aquells residus específics que detecten interaccions amb la regió epítop del virus SARS-COV-2. És per això, que en la *Figura 23* s'ha representat el radi de gir d'aquest conjunt de residus que conformen el CDR respecte el PMF obtingut. Com es pot veure, el comportament dels CDR no varia de l'observat en la branca Fab1 conjunta, els resultats obtinguts durant el procés clàssic corresponen al rang més explorat mitjançant

aMD i es manté constant. En aquest gràfic però, a part del mínim amb un valor de PMF de 0 Kcal/mol també hi apareixen un parell més de regions amb mínims energètics considerables on sembla que el sistema es podria establir en cas que el procés clàssic hagués sigut més extens. Si més no, el trobar aquests mínims no contradiu el fet que en la regió més important del sistema respecte la interacció amb l'antigen objectiu, es conserva l'estructura de partida.

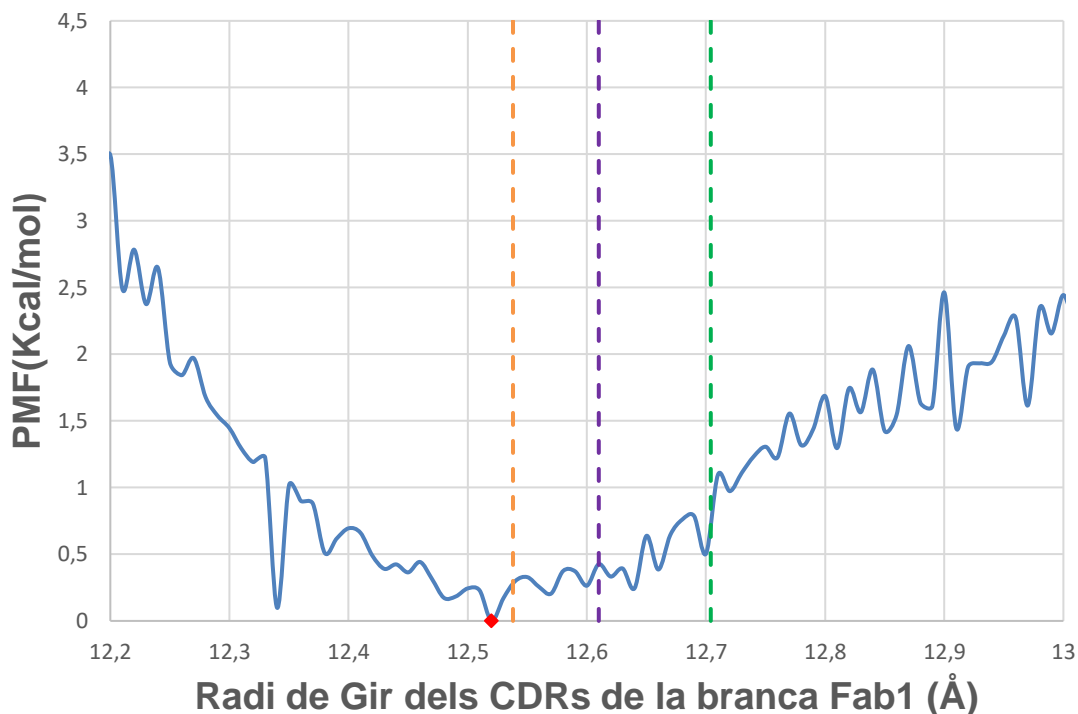


Figura 23: Evolució aMD del radi de gir de les regions CDR de la branca Fab1 del sistema CR3022 respecte el PMF.

Aquests resultats per tant, demostren la conservació de l'estructura global de la regió Fab1 i també de forma més localitzada, la conservació dels CDR concrets de la branca. Per tant, es pot afirmar que en aquesta regió es manté la funcionalitat de l'anticòs d'unir-se amb la RBD de la proteïna S del virus, sense que es vegi afectada per l'exposició durant un llarg període de temps a les interaccions resultants de la presència de la superfície d'or.

El comportament observat en la branca Fab2 de la proteïna del sistema CR3022 en la *Figura 24* és molt similar al prèviament vist per la seva branca homònima. És important tenir en compte que l'estructura i la distribució dels residus és idèntica en ambdós branques, de manera que un comportament similar és d'esperar. La distribució sobre la gràfica dels valors obtinguts de les simulacions clàssiques recauen sobre la regió explorada mitjançant el procés aMD de manera que el radi de gir de la regió Fab2 es manté constant i sense fluctuacions considerables en l'estructura. El fet que el mínim recaigui sobre el mateix valor de 24,3 Å que en la branca Fab1 ens indica que l'estabilitat es manté igual en tota la regió Fab del sistema CR3022.

Seguint el mateix procediment, en la *Figura 25* s'han graficat els resultats de radi de gir del CDR de la branca Fab1. Al ser branques idèntiques, els residus d'aquestes regions són els mateixos, però és important saber exactament la seva nomenclatura dins la molècula. Pel CDR-H de la Fab2 del sistema CR3022 trobem els residus: THR480-GLY484, ILE499-GLY515, GLY548-VAL557; i el CDR-L conté els residus: LYS914-ALA930, TRP946-SER952, GLN985-THR993. Com era d'esperar, la localització del mínim i els valors clàssics recauen sobre el mateix valor de radi de gir que en el cas del CDR de la Fab1 del sistema, de manera que es compleixen les mateixes condicions d'estabilitat del sistema prèviament raonades. Únicament afegir, que en aquest cas els mínims adjacents a la regió més estable del sistema, són més accentuats indicant altres possibles conformacions de baixa energia que podria adquirir el sistema en determinades condicions.

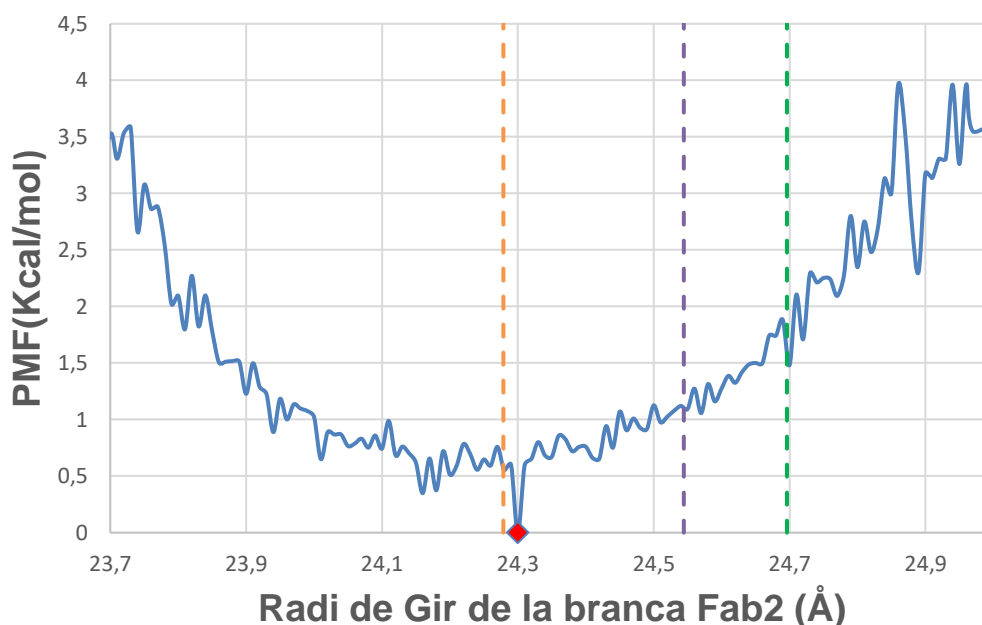


Figura 24: Gràfic representatiu del radi de gir Fab2 en funció al PMF del sistema CR3022. En un rombe vermell s'indica el mínim del sistema, i amb línies discontinues de color verd, lila i taronja és marca el valor corresponent al valor cristal·logràfic inicial, al promig global i al promig dels últims 75ns de clàssica respectivament.

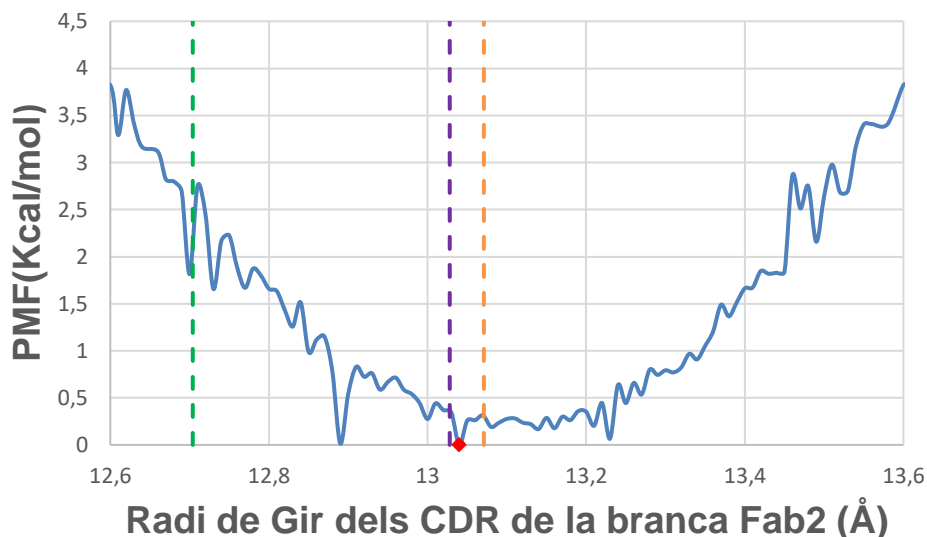


Figura 25: Evolució aMD del radi de gir de les regions CDR de la branca Fab1 del sistema CR3022 respecte el PMF.

De la mateixa manera que en el primer sistema, el procediment seguit pel sistema S309 amb l'objectiu d'analitzar l'estabilitat i funcionalitat de les regions Fab de la proteïna es basa en l'estudi del radi de gir global de la regió i de manera més específica, de la zona on es troben els CDRs de l'anticòs. Respecte la branca Fab1, en la *Figura 26* es distribueix el seu radi de gir global respecte al valor PMF prèviament calculat. És interessant remarcar que en aquest cas, els dos valors corresponents al promig global de les simulacions clàssiques i el corresponent als últims 75ns recau exactament sobre el punt amb un PMF mínim del sistema. Aquest fet ens indica que encara que el valor cristal·logràfic inicial no es trobi dins de la regió més explorada del sistema, aquest valor mínim i per tant de més estabilitat, s'assoleix ràpidament i s'aconsegueix mantenir constant durant tota la simulació. En definitiva, l'estabilitat que s'observa en aquesta regió del sistema ens confirma que després de tot temps de simulació tant de cMD com d'aMD, l'estructura secundària i la funcionalitat es conserven.

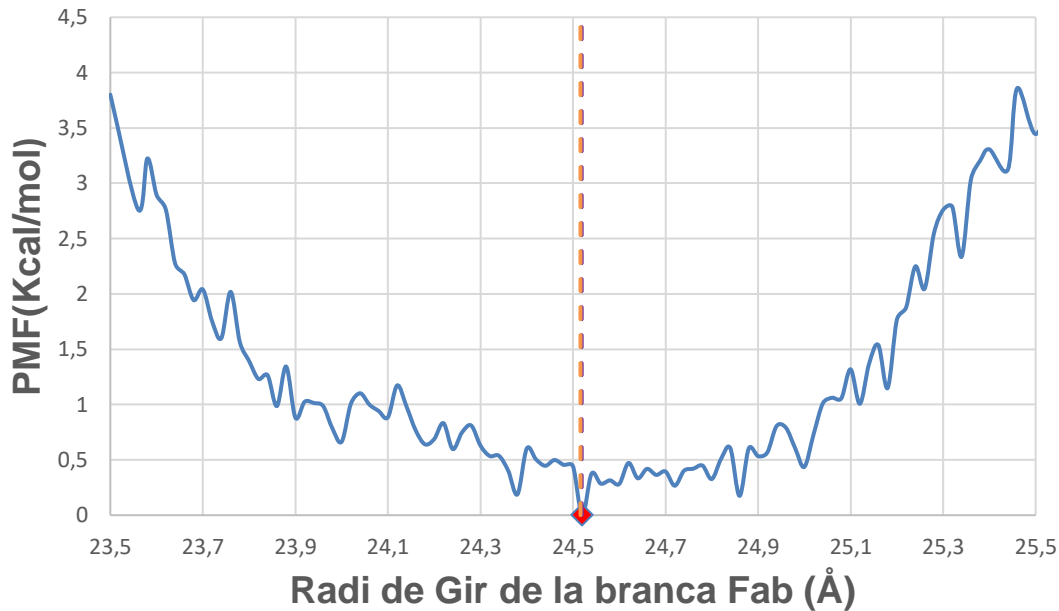


Figura 26: Gràfic representatiu del radi de gir Fab1 en funció al PMF del sistema S309. En un rombe vermell s'indica el mínim del sistema, i amb línies discontinues de color lila i taronja es marca el valor corresponent al promig global i al promig dels últims 75ns de clàssica respectivament.

Després de l'anàlisi global de la branca Fab1, passem a realitzar el mateix procediment per la regió de CDR específics del sistema S309. La distribució és la mateixa que pel sistema CR3022 i en la regió Fab1 trobem dos tipus de CDR en funció de la cadena a la que pertanyen: CDR-H (TYR31-TRP35, ILE50-ARG66, TYR99-TRP116) i CDR-L (ALA923-TRP934, ALA950-GLY956, GLN989-PHE996). El resultat obtingut d'aquestes regions queda recollit en la *Figura 27* i, per una banda, es pot veure com els valors tant dels promig de les simulacions clàssiques com el valor cristal·logràfic inicial, recauen a la zona central i de menor energia del sistema. En canvi, si ens centrem en la regió aMD explorada d'aquest radi de gir en específic, podem veure com el punt mínim energètic trobat recau en un valor de 12,8 Å i queda a la dreta de la zona més concorreguda de les simulacions cMD. El fet que aquest mínim energètic es trobi una mica desplaçat i que no sigui massa accentuat en referència als altres mínims observables en la gràfica, és un indicador de que, en aquest cas, el sistema no acaba d'assentar-se en un mínim energètic estable. Malgrat això, l'estructura inicial del sistema es manté constant, tant en el procés cMD com en el aMD.

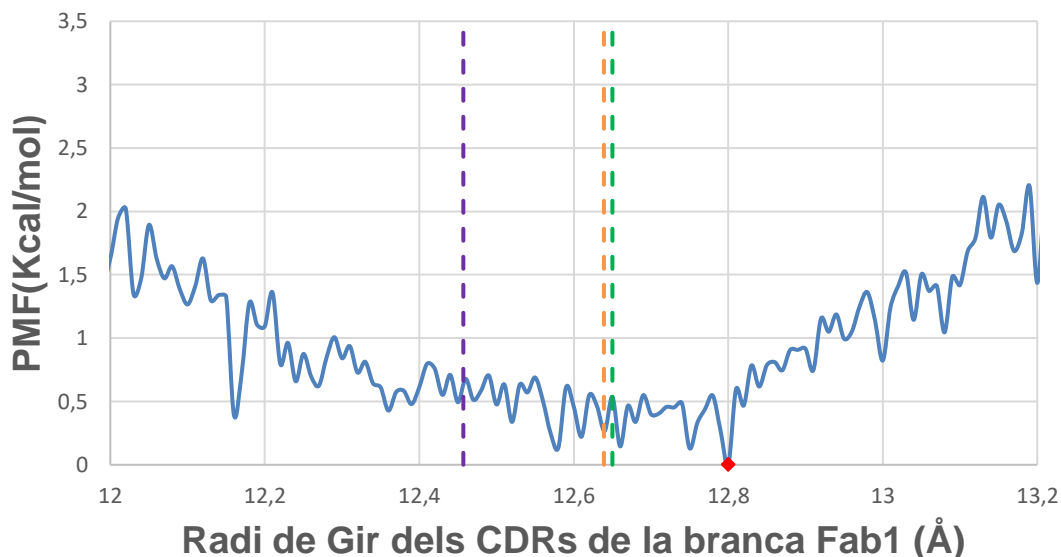


Figura 27: Evolució aMD del radi de gir de les regions CDR de la branca Fab1 del sistema S309 respecte el PMF.

Conseqüentment, en la *Figura 28* es pot veure com el comportament observat en la branca Fab2 del sistema S309 és el mateix que en el cas de la Fab1. En aquest cas, els valors promig del procés clàssic no recauen exactament sobre el mínim del sistema, però segueixen formant part de la zona més explorada del sistema. Igual que en el sistema CR3022 el mínim recau sobre el mateix valor que en la branca Fab1 del sistema S309 confirmant així un comportament idèntic a l'hora de conservar de forma constant l'estructura principal i secundària de la regió Fab2.

En aquest cas però, els resultats obtinguts per la regió CDR formada pels residus específics de les regions CDR-H (TYR487-TRP491, ILE506-ARG522, TYR555-TRP572) i CDR-L (ALA1137-TRP1148, ALA1164-GLY1170, GLN1203-PHE1210) difereixen una mica dels vistos en la branca Fab1 del S309. Tal com es pot observar en la *Figura 29*, el comportament clàssic és el mateix observat en l'anàlisi global de la branca Fab2, ja que malgrat trobar-se una mica allunyat del mínim energètic inicialment, aquest convergeix ràpidament cap la regió on el PMF adopta valors més baixos. Observant la regió aMD explorada, podem veure com en aquest cas, s'han trobat dos mínims ben definits amb un PMF de 0 Kcal/mol i el radi de gir observat fluctua entre aquests dos punts de mínima energia, però mantenint-se constant durant tot el període simulat.

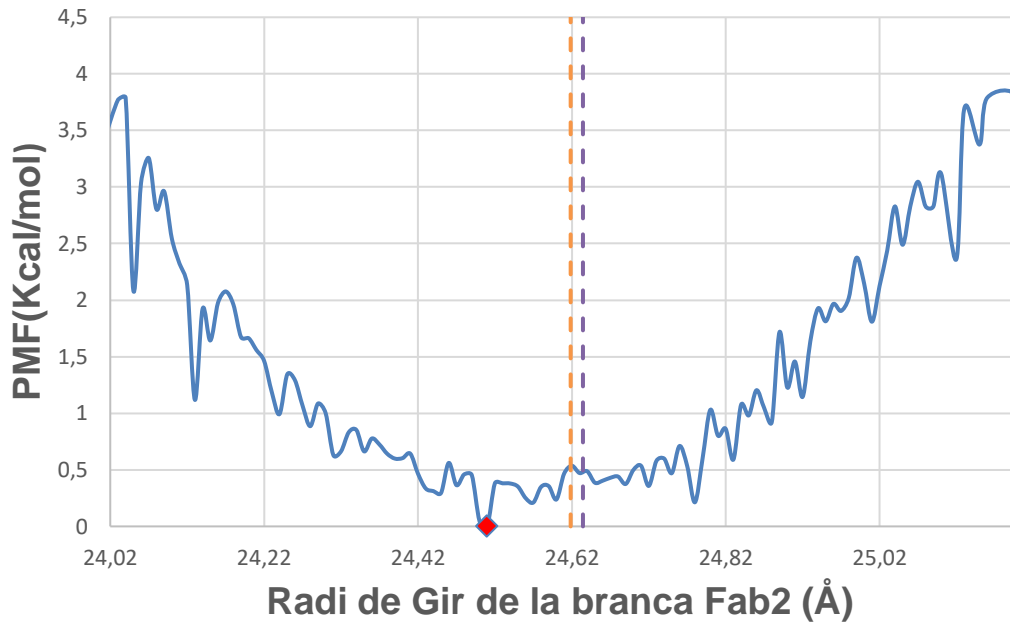


Figura 28: Gràfic representatiu del radi de gir Fab2 en funció al PMF del sistema S309. En un rombe vermell s'indica el mínim del sistema, i amb línies discontinúes de color lila i taronja es marca el valor corresponent al promig global i al promig dels últims 75ns de clàssica respectivament.

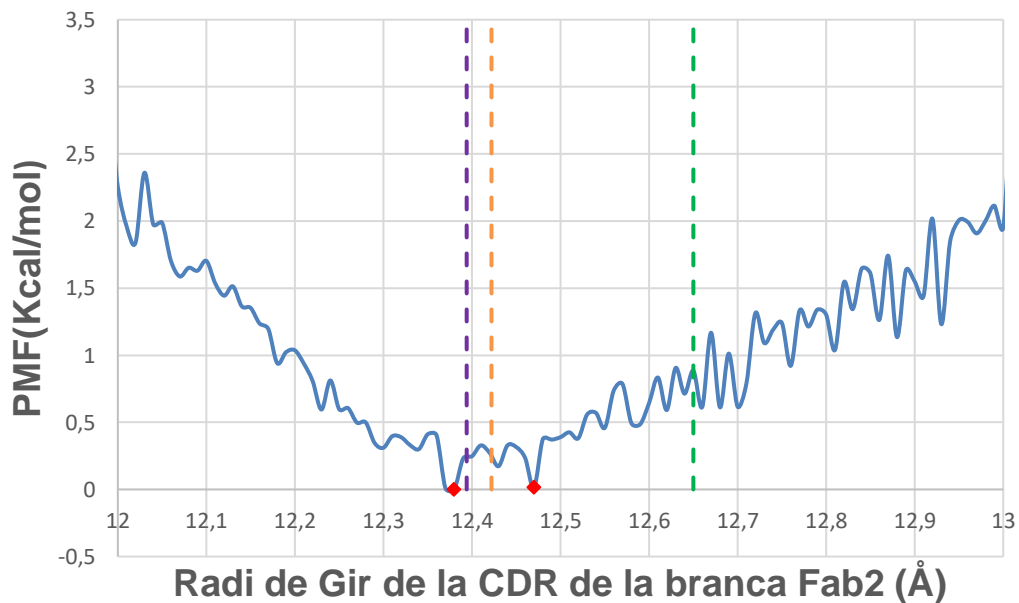


Figura 29: Evolució aMD del radi de gir de les regions CDR de la branca Fab2 del sistema S309 respecte el PMF.

Finalment, amb el que s'ha vist fins ara, s'ha pogut determinar tant en el cas del sistema CR3022 com en el del S309 el comportament de les regions Fab del sistema, a partir de l'anàlisi exhaustiva del radi de gir. El comportament que presenten les dues proteïnes sobre la superfície d'or és molt similar i el seu estudi demostra que les interaccions que aquesta superfície i el sistema de solvatació, a partir de molècules d'aigua que les envolten, no són prou rellevants per alterar l'estructura i funcionalitat d'aquestes branques més exposades del sistema. De la mateixa manera, l'anàlisi realitzada sobre les diferents CDRs de cada sistema ens ha aportat els resultats concrets esperats que indiquen com aquests residus més exposats de les branques Fab1 i Fab2 mantenen una estructura constant sobre el mínim energètic del sistema, garantint així l'especificitat i funcionalitat característiques d'aquest conjunt de residus que els permet detectar i interaccionar amb el virus SARS-COV-2 amb l'objectiu de desnaturalitzar la seva estructura i impedir d'aquesta manera la fusió amb la membrana de la cèl·lula hoste objectiu. Cal també destacar que mentre en el sistema CR3022 el comportament de la regió CDR de les dues branques segueix el mateix patró, en el cas del sistema S309 les regions CDR presenten un comportament un pèl diferent.

7.4. Estudi de la regió Fc del sistema

Gràcies a les anàlisis prèvies, s'ha pogut determinar l'evolució global dels dos sistemes amb els que s'ha treballat en aquest projecte i s'ha comprovat com efectivament es manté correctament l'estructura i funcionalitat de les regions més externes i importants a l'hora de captar i lligar-se a la regió RBD del virus. Ara bé, per tal que les futures aplicacions d'aquests sistemes d'anticossos per al desenvolupament dels immunosensors prèviament comentat és tant, o més important, entendre que succeeix en l'anomenada regió Fc del sistema. Com ja s'ha vist, és mitjançant aquesta part, que la proteïna es manté unida a la capa superior de la superfície d'or dels sistemes CR3022 i S309 a través del residu de lisina (LYS457) modificat i unit covalentment a un àtom d'or específic. De la mateixa manera, és la regió de la proteïna a través de la qual es transmet la informació des de les regions Fab1 i Fab2 a la capa d'or i d'aquí a la posterior interpretació del sensor.

Aleshores, per desenvolupar l'anàlisi exhaustiva d'aquesta regió Fc, seguirem el mateix procés aplicat en l'estudi de les Fabs per determinar les zones de mínima energia mitjançant el valor de PMF obtingut de cadascuna de les variables d'interès. Primerament, veurem el nivell d'aixafament de la molècula, seguirem amb l'anàlisi de l'estabilitat global i de la regió amb els CDRs de la zona més propera a la superfície i finalment es complementarà amb una anàlisi de l'energia lliure d'enllaç del sistema sobre la superfície.

Començant com sempre amb el sistema CR3022, en la *Figura 30* s'ha representat la distància del centre de masses de la regió Fc respecte el PMF obtingut. En aquest gràfic es segueix la metodologia emprada per als anteriors, de manera que en taronja es veu representat el valor dels últims 75ns de simulació clàssica de la distància. El valor inicial

es situaria a la dreta del rang que es mostra en la gràfica. Això és un indicador de que la distància s'ha reduït considerablement fins a arribar a un valor mínim d'energia com es pot observar a una distància de 30 Å. D'aquest gràfic podem extreure que malgrat el sistema estar avançant en una bona direcció cap a l'estabilitat energètica, amb els 150ns de clàssica simulat, no s'ha arribat a aquest punt, que sí s'ha pogut explorar de forma àmplia mitjançant les simulacions aMD. És a dir, aquest aixafament de la proteïna sobre la superfície es seguiria produint fins arribar al valor de mínima energia trobat. D'aquesta manera es pot determinar que un cop assolit aquest punt, el sistema s'estabilitzaria i les probabilitats de que es continués aixafant es reduïrien considerablement degut a què per fer-ho el sistema hauria de travessar salts energètics considerables.

Per altra banda, a part de l'aixafament de la proteïna, un altre procés que es produeix en la proteïna és el fet de tombar-se i caure sobre la superfície. Per tal de controlar aquest fenomen s'ha analitzat l'evolució del *Tilt Angle*, que és l'angle que forma la regió Fc del sistema respecte a la superfície d'or. En la *Figura 31*, podem veure com efectivament des del valor cristal·logràfic inicial, marcat amb la línia verda discontinua, fins a la regió de mínima energia del conjunt de simulacions aMD, hi ha un canvi significatiu en el valor de l'angle que indica aquesta tendència a tombar-se de la proteïna. Aquest mínim trobat però, és un molt bon indicador que ens permet veure com aquest sistema tendirà a estabilitzar-se a un angle final aproximat de 54,3° o com a molt arribant a una segona zona de mínima energia també bastant explorada, situada als 49,4°. De manera que un cop assolit aquest angle, les probabilitats de que la proteïna es segueixi tombant seran molt més baixes i tant la distància com l'angle mantindran uns valors constants que garantitzaran l'estabilitat estructural de la proteïna i sobre tot, de la regió Fc.

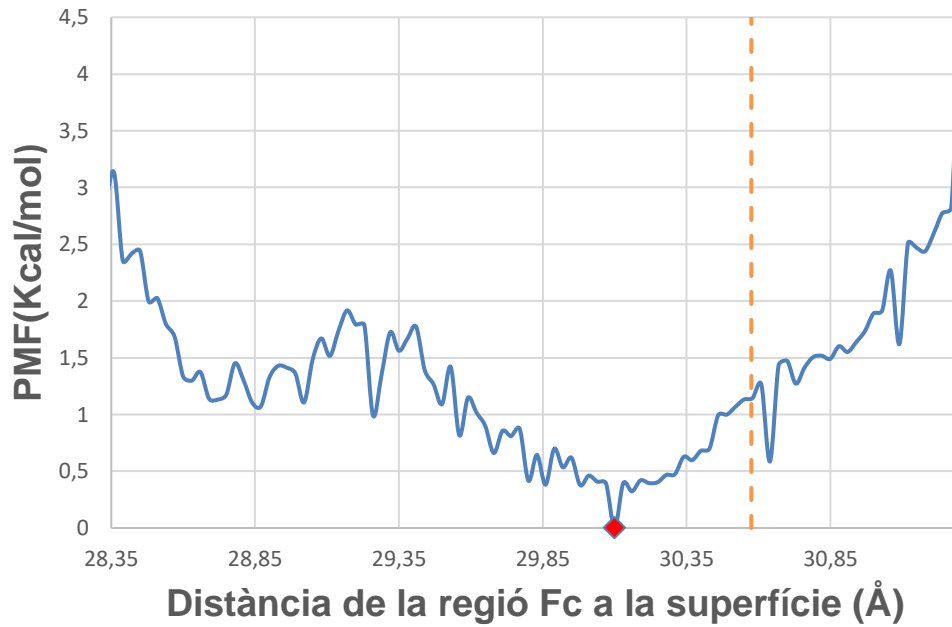


Figura 30: Gràfic representatiu de l'evolució de la distància de la regió Fc a la superfície del sistema CR3022 respecte el PMF. El rombe vermell indica el mínim del sistema i la línia taronja discontinua el valor promig dels últims 75ns de cMD.

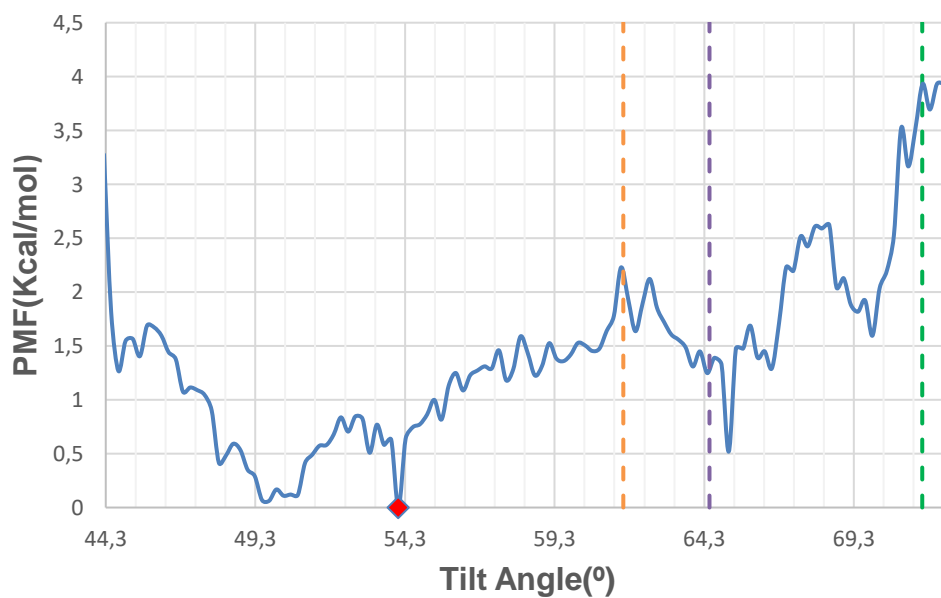


Figura 31: Gràfic representatiu del Tilt Angle del sistema CR3022 respecte el PMF.

Seguint el mateix procediment que en les anteriors anàlisis, s'ha estudiat el comportament estructural del sistema S309 de la mateixa manera. En la *Figura 32* es pot veure l'evolució de la distància respecte el PMF i de la mateixa manera que en el sistema CR3022 la distància de la proteïna cristal·lina inicial es veu reduïda fins a arribar a una regió on s'assoleix una estabilitat de baixa energia. En aquest cas però, malgrat que el promig dels últims 75ns de clàssica cauen sobre el mínim del sistema, no s'ha trobat un punt mínim prou accentuat on es mantingui el sistema, sinó que en el conjunt de distàncies explorades mitjançant les simulacions aMD s'observa un rang que va dels 27,5 als 29,8 Å on l'energia PMF és prou baixa. Per tant, és molt probable que la distància del centre de masses de la regió Fc respecte a la superfície no assolixi un valor final, sinó que fluctuarà de manera constant i estable dins del rang establert.

Pel que fa a l'angle que forma la superfície amb la Fc del sistema S309, observem un procés similar al del sistema CR3022, malgrat que el valor cristal·logràfic inicial de l'angle no queda recollit dins del rang aMD de simulacions, si es pot observar la tendència del sistema cap a la regió de menor energia. Com és d'esperar, igual que en el càlcul de distàncies no s'observa un únic mínim. En la regió de mínima energia trobem tres punts diferents amb mínims significatius que indiquen que un cop assolit, el sistema final no serà estàtic, sinó que fluctuarà dins d'un rang comprès entre els 48,4° i els 50,5°. Donant a un sistema resultant més mòbil que en cas del CR3022, però mantenint igualment una estructura estable i constant.

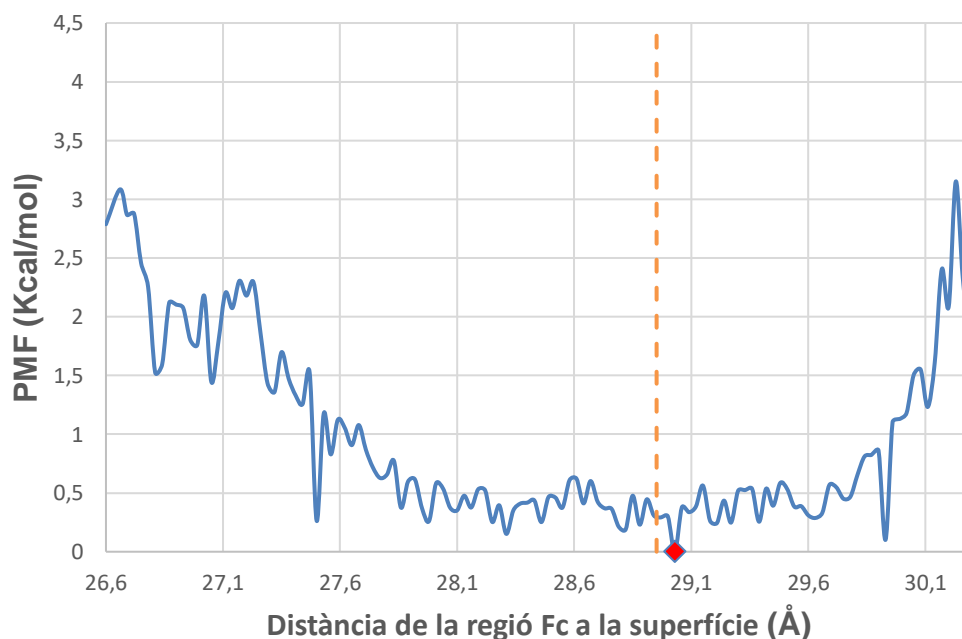


Figura 32: Evolució de la distància de la regió Fc a la superfície del sistema S309 respecte el PMF. El rombe vermell indica el mínim del sistema i la línia taronja discontinua el valor promig dels últims 75ns de cMD.

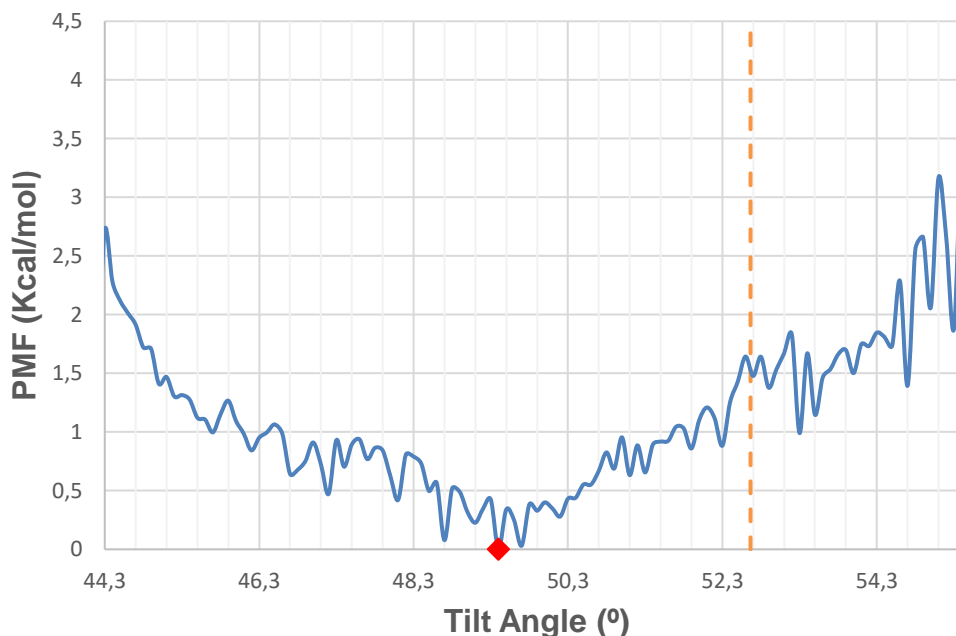


Figura 33: Gràfic representatiu del Tilt Angle del sistema S309 respecte el PMF.

Un cop analitzat l'aixafament i caiguda de la molècula sobre la superfície d'or del sistema és moment de comprovar de la mateixa manera que s'ha fet en les branques Fab1 i Fab2 de cada sistema, si durant aquest procés de simulació on s'ha vist alterat el sistema, l'estructura i funcionalitat de la regió i els CDRs específics es conserven o si pel contrari pateixen alguns canvis significatius.

En la *Figura 34* s'ha representat el radi de gir de la regió Fc respecte el PMF del sistema i seguint la mateixa nomenclatura que en les gràfiques anteriors, podem veure com en el cas del sistema CR3022 els valors clàssics obtinguts romanen pràcticament constants i sobre la regió de mínima energia observada sobre el valor de 26,45 Å i amb unes variacions del radi de gir menors de 0,5 Å per tot el conjunt de simulacions cMD. D'aquesta manera es confirma que malgrat el cert grau d'aixafament i caiguda de la proteïna sobre la superfície d'or, l'estructura i per tant, la funcionalitat d'aquesta regió del sistema per mantenir la proteïna estable no es veu prou afectada per les interaccions de la superfície ja que el volum de la regió s'adapta de manera que al caure sobre l'eix Z, s'expandeix de manera uniforme sobre els eixos X i Y. Aconseguint així una estabilitat en el radi de gir de la Fc.

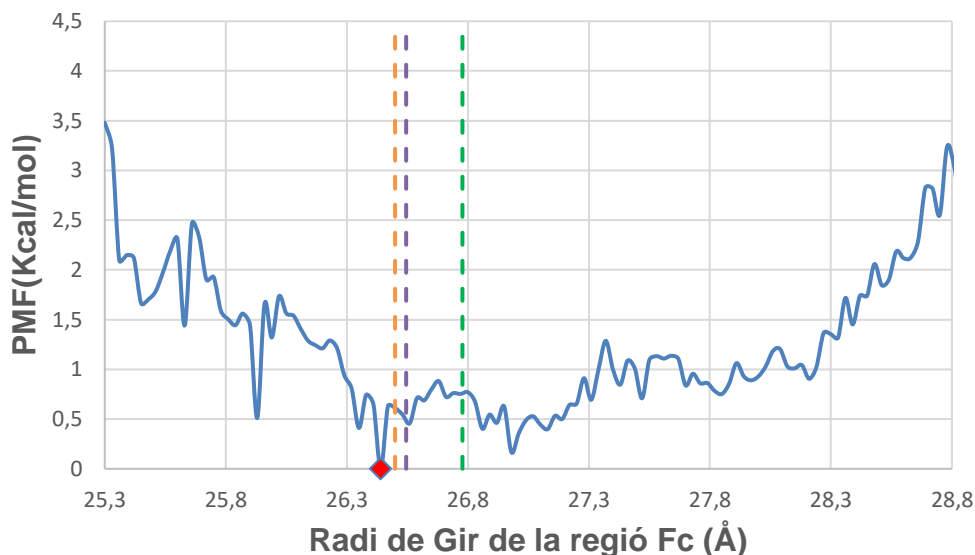


Figura 34: Radi de gir de la regió Fc del sistema CR3022 respecte el PMF.

En el cas del sistema S309, els resultats obtinguts corresponents al radi de gir de la regió Fc en funció del PMF denoten un comportament diferent. Com es pot veure en la *Figura 35* el valor cristal·logràfic inicial de la proteïna queda fora del rang observable en el gràfic, de manera que durant el procés de simulació ha patit un increment considerable. Per altra banda, els valors finals clàssics tampoc recauen sobre el mínim d'energia PMF trobat a partir de les simulacions aMD, sinó que tenen una tendència a incrementar el valor del radi de gir. Durant els 150ns de clàssica, l'increment ha estat de més de 5 Å, un valor molt superior a qualsevol altre radi de gir analitzat durant aquest projecte que confirma que l'aixafament que es produeix sobre la superfície no es veu prou compensat per l'expansió de la regió sobre els eixos X i Y fent que el volum es comprimeixi. És per això que d'aquesta anàlisi podem extreure que el fet que l'estabilitat del sistema no assoleixi un valor mínim energètic prou estable, sinó que aquesta fluctui com s'ha pogut veure en els diferents anàlisis, tant els globals de cada regió com els més específics, genera unes condicions d'instabilitat en el sistema que acaben afectant a l'estructura principal de la base de la proteïna, la regió Fc.

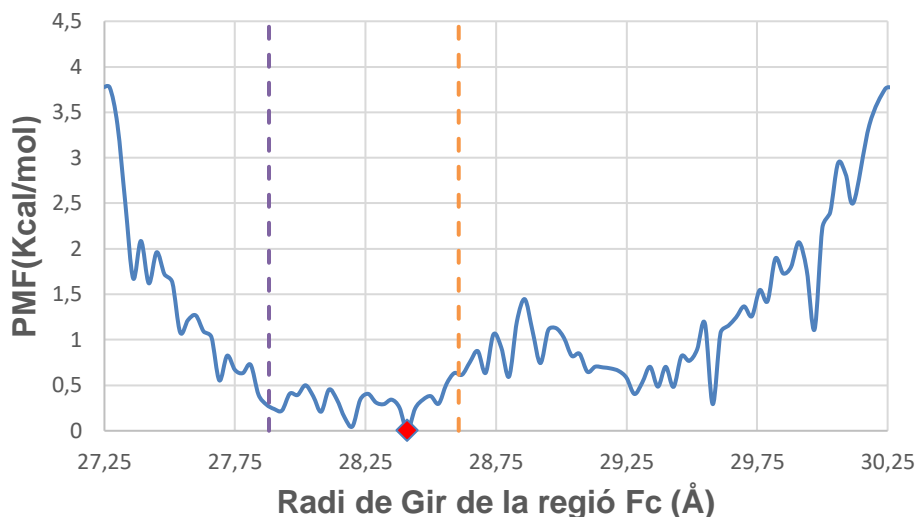


Figura 35: Radi de gir de la regió Fc del sistema S309 respecte el PMF.

Finalment, s'han determinat l'energia lliure d'enllaç (FBE) obtinguda de cada sistema a partir de les interaccions entre la proteïna i la superfície i s'ha correlacionat mitjançant una gràfica en dues dimensions amb l'angle que formen la superfície d'or amb la regió Fc dels sistemes. La FBE proporciona informació sobre la força d'atracció que es dona entre les dues parts del sistema. Per aquest motiu, el fet de graficar-la en funció del *Tilt Angle* ens permet veure el nivell de interacció d'aquestes conformacions amb valors mínims de PMF trobats que ja s'ha corroborat que impliquen una major estabilitat i convergència del sistema respecte a la superfície d'or dels sistemes.

D'aquesta manera, podem veure com en el cas del sistema CR3022 de la *Figura 36*, aquests mínims es troben sobre uns valors de FBE d'entre -6000 i -6500 Kcal/mol mentre que en el cas del sistema S309, aquests mateixos mínims recauen sobre un FBE més baix, entre -5600 i -5900 Kcal/mol. En canvi, l'angle que dona lloc a aquests mínims de PMF observats és més gran en el cas del CR3022 que pel sistema S309. El mínim més important es troba en un angle de 54° en el primer dels sistemes respecte a un angle de 49° en el segon. Amb aquests resultats podem concloure que malgrat la FBE del sistema CR3022 és més baixa, i per tant la força d'interacció de la proteïna amb la superfície és major que en cas del sistema S309 on la FBE és una mica més elevada, el primer sistema presenta una major estabilitat i un *Tilt angle* més gran i constant que el S309, on la proteïna presenta una tendència més accentuada a inclinar-se sobre a la superfície.

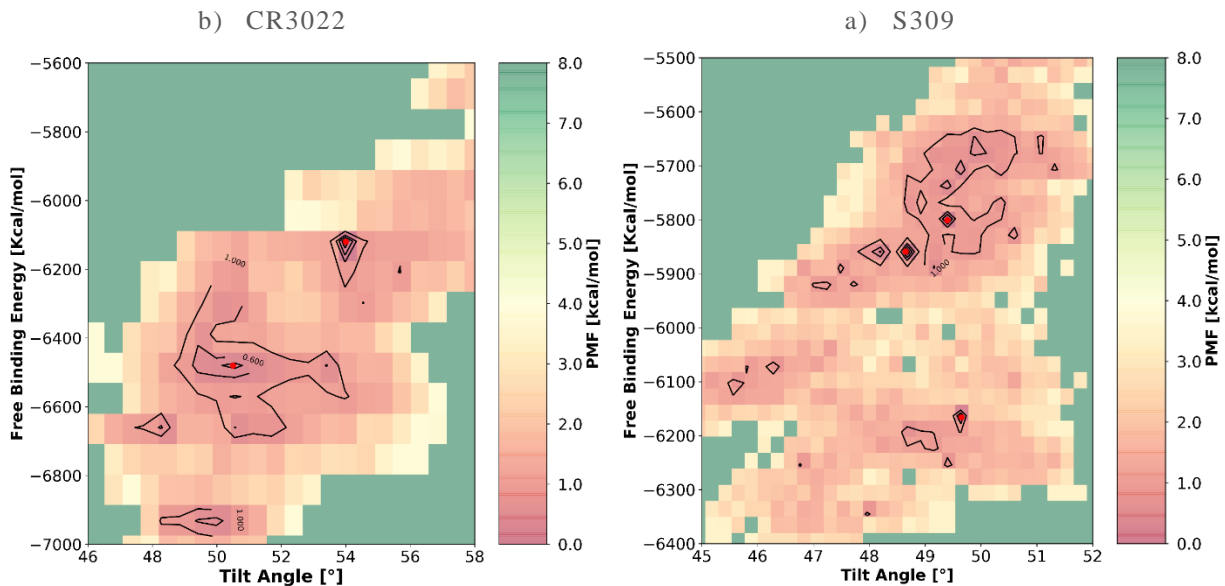


Figura 36: Gràfic 2D del Tilt Angle respecte la FBE del sistema en funció del PMF del sistema. Els punts vermells del sistema indiquen els mínims trobats. A) Gràfic del sistema CR3022. B) Gràfic del sistema S309.

7.5. Anàlisi final i comparatiu dels resultats

En aquest projecte s'han estudiat dos sistemes paral·lelament amb l'objectiu d'analitzar el comportament dels seus respectius anticossos sobre una superfície d'or per entendre la seva evolució en unes condicions de pressió, temperatura i forces d'interacció amb la superfície i el medi aquós que envolta el sistema generades computacionalment per tal de simular el comportament que tindrien aquestes proteïnes en aquestes condicions.

Recollint els resultats més importants del sistema CR3022, podem dir que durant les primeres etapes de la simulació el conjunt del sistema pateix una sèrie de canvis interessants. Per una banda, es produeix un aixafament de la proteïna que queda reflexada amb la reducció de les distàncies del centre de masses de cada regió respecte la superfície d'or i un procés d'inclinació que s'ha analitzat a partir de l'angle que formen la regió Fc i la superfície. Els efectes d'aquests dos processos respecte a l'estructura global de la proteïna s'han analitzat a partir dels angles i els diedres de les regions Fab1 i Fab2 de l'anticòs. D'aquesta forma, s'ha comprovat com encara que el conjunt global de la proteïna caigui cada cop més a prop de la superfície, les regions Fab roten de tal forma que l'estructura i orientació originals en que les zones CDR se situen en les parts més externes de les Fabs i en direcció oposades a la superfície es mantenen. Gràcies a això es garanteix la seva orientació òptima respecte a les futures interaccions amb l'antigen SARS-COV-2. Per altra banda, a partir del RMSF, s'ha vist com la regió central de la

proteïna tenia una mobilitat baixa que confirma que es mantingui l'estructura global de la proteïna, mentre que en les regions més externes la mobilitat era més elevada. En part, és degut a aquesta mobilitat en determinades regions del sistema, que s'ha realitzat un estudi més exhaustiu a partir dels resultats de les simulacions aMD i el PMF calculat. D'aquesta manera s'ha aportat llum sobre les dos incògnites que faltaven per resoldre. Primerament, s'ha determinat que malgrat la mobilitat en les regions externes de les branques Fab1 i Fab2, tant la seva estructura global com la funcionalitat dels residus que formen els CDRs específics de cada branca es mantenen i, per tant, tampoc es veuen afectades, per les interaccions amb la superfície i els processos que fan que s'aixafi i s'inclini la regió inferior de la proteïna. I en segon lloc, s'ha vist com els processos que aixafen el sistema acabaran estabilitzant-se i es mantindrà una certa estructura de la regió Fc del sistema que malgrat els canvis, conservarà la seva funcionalitat.

El comportament observat en el sistema S309 és molt similar. En un primer moment, l'estructura global del sistema també es veu alterada pel procés de simulació i els processos que afecten a l'aixafament i la inclinació de la proteïna sobre la superfície són els mateixos. A mida que avança el temps de simulació els valors, respecte a les distàncies i l'angle de la superfície respecte a la Fc es veuen reduïts i és a partir dels angles restants calculats i diedres del sistema que veiem com l'estructura de la proteïna varia i s'adapta per mantenir l'estabilitat. Igualment, les branques Fab1 i Fab2 del sistema mantenen l'orientació inicial que resulta en la correcta col·locació de les zones CDR. Per altra banda, la mobilitat de la proteïna observada a partir del RMSF és, en general, menor però seguint el mateix patró de regions més o menys mòbils. Per això, s'han analitzat els mateixos paràmetres a partir de les simulacions aMD. En aquest cas, s'ha observat que la proteïna tendeix a presentar un comportament més inestable on no s'estabilitza sobre una única conformació, sinó que el nivell d'aixafament i inclinació de la proteïna fluctua de tal forma que s'observa un comportament una mica menys estable en els CDRs de les branques Fab i que s'accentua més en l'estructura resultant de la Fc.

En el cas del procés d'aixafament i inclinació de la regió Fc de les proteïnes, s'ha observat un comportament molt similar en els dos sistemes. En tots dos casos ha estat la cadena pesada lligada a la superfície la que ha mig col·lapsat sobre la superfície. Això es deu a diversos factors. En primer lloc, és important tenir en compte que el sistema cristal·logràfic estava inicialment orientat per què, en el cas de tombar-se, ho fes de tal forma que tota la proteïna estigués sobre la superfície i així reduir el nombre d'àtoms d'or necessaris per la simulació, de manera que és probable que una de les dos cadenes quedés més exposada que l'altre. Un altre punt, és que la part inferior de les cadenes pesades, que correspon a la Fc és idèntica en els dos sistemes ja que aquesta part prové del mateix PDB original. En últim terme, el fet de que aquesta cadena estigues lligada a la superfície ha pogut donar lloc a un increment en la interacció respecte la seva cadena homònima i ha facilitat l'aixafament cap aquest costat determinat.

Per concloure, és important entendre que, encara que inicialment el sistema S309 presenta unes alteracions més baixes en l'estructura global de la proteïna respecte el CR3022, aquest últim acaba assolint un estat final més estable. És més, el fet que les interaccions

amb la superfície en funció del *Tilt angle* i el PMF del sistema, siguin més elevades en sistema CR3022 i que tot i així el punt de mínima energia permeti una inclinació menys accentuada que en el S309 ens dona més arguments per afirmar després d'un cert temps de simulació, que el sistema CR3022 serà el més probable de presentar una estabilitat en la proteïna que li permeti mantenir l'estructura i la funcionalitat específica i necessària de cadascuna de les regions estudiades en aquest projecte.

8. Avaluació Econòmica

A continuació es descriurà el pressupost final que s'ha calculat pel desenvolupament del projecte.

En la realització del mateix, s'han utilitzat únicament eines computacionals per tal de poder desenvolupar tot el conjunt de simulacions i els posteriors anàlisis necessaris. En conseqüència, s'ha dividit el preu total en funció del cost humà on s'inclouen aquelles persones que han participat en la d'alguna forma en el projecte, el cost total del *software* que s'ha utilitzat, el cost del *hardware* on s'han els diferents processos computacionals i per últim, el cost relacionat amb les eines pròpies necessitades i el proporcional al desplaçament.

Començant pel cost humà, s'ha considerat a l'estudiant com a un enginyer *junior*. També s'ha comptat amb el suport d'un enginyer tècnic amb més de tres anys d'experiència en el camp d'estudi i el respectiu tutor. Els salaris s'han determinat a partir del nivell acadèmic i de l'experiència de cada persona en base al Conveni d'Enginyeries i Oficines d'Estudi Tècnic XIX recollits en el BOE [71]. Les hores treballades s'han obtingut a partir del promig diari, tenint en compte que el projecte es va iniciar a principis de Febrer del 2021 i s'ha finalitzat a principis de Juny del mateix any i el desglossament és pot veure en la *Taula 14*. Amb la qual cosa, ha tingut una durada de quatre mesos.

Taula 14: Cost humà del projecte desglossat en funció del personal, el preu i el temps.

Persona	Salari (€/h)	Temps invertit (h)	Cost Total (€)
Enginyer Junior	9,62	640	6.156,8
Enginyer Tècnic	15,38	160	2.460,8
Tutor	27,88	32	892,16
TOTAL			9509.76 €

En l'apartat referent al costos de *software*, tenim la sort de que la majoria del programari utilitzat és o bé de codi obert i per tant gratuït o és pot fer servir mitjançant llicències acadèmiques que obté la universitat, amb la qual cosa el cost d'aquesta secció es veu dràsticament reduït en comparació amb el que s'obtidria en altres tipus de projectes experimentals. La distribució d'aquests costos queda recollida en la *Taula 15* i contempla que el període d'amortització dels programes informàtics és de 6 anys.

Taula 15: Cost del *Software* utilitzat en el projecte en funció de la seva amortització.

<i>Software</i>	Cost associat (€)	Durada de la llicència (dies)	Temps d'ús (dies)	TOTAL (€)
AMBER18	412,45	2190	120	22,6
NAMD2.13	0	-	120	0
NwChem	0	-	120	0
Gaussian09	4280	2190	120	234,52
Chimera	0	-	120	0
VMD1.9.3	0	-	120	0
			TOTAL	257,12 €

El cost del *hardware* físic utilitzat per les diferents simulacions i càlculs dels projecte s'ha determinat a partir de les hores totals de computació emprades. És important tenir en compte que els costos físics associats a l'ús dels diferents equips de computació utilitzats queden inclosos en el cost de càlcul de cada clúster. Per l'equilibrat inicial del sistema i els posteriors anàlisis de les simulacions s'ha utilitzat el clúster TOSCA de la UPC, per la resta de simulacions que inclouen els 150ns de cMD i els 360 de aMD de cadascun dels sistemes del projecte s'ha fet servir el clúster francès de IRENE i finalment, pel càlcul de les FBE que requeria un cost de memòria molt gran, s'ha fet servir el clúster de CSUC. És per això que com es pot observar en al *Taula 16*, només s'han comptabilitzat els costos referents al clúster propi de la universitat. Els costos de càlcul del clúster IRENE són zero ja que s'han aprofitat les hores concedides per investigació a la universitat, mentre que els costos del CSUC són els establerts pel propi clúster [72] i els de TOSCA s'han aproximat [14].

Taula 16: Cost associat al *Hardware extern* utilitzat per simular.

<i>Hardware</i>	Temps de computació (h)	Cost de càlcul (€/h)	TOTAL (€)
TOSCA UPC	15.000	0,033	495
IRENE	500.000	0	0
CSUC	2.464	0.017	41,065
		TOTAL	536.065€

Finalment, en referència al cost de l'equip personal és sumen els costos de 600€ de l'ordinador obtingut expressament per la realització d'aquest projecte i un total aproximat de 200€ en base al cost del transport públic utilitzat per desplaçar-se al centre universitari de recerca durant aquests quatre mesos de durada. Sumant un total de 800€ de costos propis.

El resum final amb la suma dels costos totals que han permet la realització d'aquest projecte queden recollits en la *Taula 17*.

Taula 17: Taula resum del cost total del projecte.

Tipus de cost	Cost (€)
Cost humà	9.509,76
<i>Software</i>	257,12
<i>Hardware</i>	536,07
Cost personal	800
TOTAL	11.102,19 €

9. Conclusions

Una vegada finalitzat l'estudi proposat en aquest projecte, és el moment d'analitzar els resultats obtinguts i extreure una sèrie de reflexions finals.

Sent el punt de partida del treball l'elaboració de dos anticossos, específics per la detecció del virus SARS-COV-2 causant de la malaltia COVID19 per la seva posterior aplicació en immunosensors, és van dissenyar dos sistemes que permetessin avaluar el seu comportament. Aquests sistemes es van formar a partir d'una proteïna IgG1 modificada per contenir les regions Fab específiques dels anticossos CR3022 i S309 en funció del sistema, immobilitzada sobre la superfície d'or i solvatats amb molècules d'aigua per simular les condicions més habituals del immunosensor.

Un cop obtinguda la cristal·lografia inicial dels dos sistemes es van seguir una sèrie de protocols estàndards de minimització energètica, escalfament i equilibrat de pressió per portar-los a unes condicions estables de 298K i 1.032 bar. A continuació tot el procés de simulació implicant les etapes de cMD i aMD es va dur a terme per cada sistema amb l'objectiu de simular el comportament que tindrien aquestes proteïnes en la realitat i explorar així tot el conjunt de conformacions possibles que podria assolir el sistema.

Arribats a aquest punt, es va desenvolupar un anàlisi exhaustiu que permetés comparar els resultats cristal·logràfics inicials de la proteïna, amb aquells obtinguts després de 150ns de simulació clàssica i 360ns de simulació aMD. Aquests anàlisis consistien en una primera part on es van comprovar els canvis produïts al final de la cMD en relació a l'estructura inicial del sistema seguit d'un segon anàlisi més exhaustiu de cada regió concreta dels sistemes aportant més informació sobre les zones amb mínims energètics del sistema.

En el cas del sistema simulat per l'anticòs CR3022 els resultats obtinguts indiquen que l'estructura principal del sistema es veu inicialment alterada a conseqüència de totes les forces a les que es troba sotmès el sistema durant el procés de simulació. Però un cop assolides les conformacions de mínima energia observades a partir dels PMF obtinguts, la proteïna tendeix a estabilitzar-se. Conjuntament, diferents resultats obtinguts, mostren com també es manté la orientació i funcionalitat de les regions CDR específiques per interaccionar amb la proteïna S del SARS-COV-2.

Per altra banda, els resultats obtinguts pel sistema de l'anticòs S309 mostren uns resultats similars. Aquest sistema s'ha vist que està sotmès a unes forces d'interacció amb la superfície menors que en el CR3022 i els canvis estructurals observats han sigut menors. En canvi, un cop s'ha assolit un temps de simulació prou elevat el sistema no aconsegueix estabilitzar-se sobre una conformació amb un mínim energètic important i es produeixen unes fluctuacions que malgrat no alterar el correcte funcionament i orientació de les CDR de les dues branques Fab1 i Fab2 del sistema si que s'aprecien més en la regió Fc.

Sobre els resultats obtinguts d'aquesta regió Fc, s'ha pogut extreure que es produeix un aixafament i una inclinació sobre la superfície d'or generada per les forces del sistema, la orientació inicial de la proteïna i el fet d'estar lligada mitjançant un residu determinat a la superfície. En el cas del sistema CR3022 s'ha vist com a pesar de tot això l'estructura i funcionalitat de la proteïna es mantenen. Paral·lelament, el resultat observat en el S309 mostra com les regions Fab1 i Fab2 mantenen la seva funcionalitat però els canvis produïts en la Fc són menys clars.

D'aquesta manera, podem concloure que els objectius generals de mantenir una estructura estable es compleix en els dos sistemes estudiats. En el cas del sistema CR3022 tant les regions Fabs com la Fc mostren uns resultats que confirmen els objectius de mantenir aquestes regions específiques funcionals malgrat estar sotmès a un conjunt de forces considerables mentre que en el sistema S309 la funcionalitat de les regions Fabs també es conserven però els canvis observats en l'estructura de la Fc són més elevats del podria interessar a l'hora d'elaborar els immunosensors a partir d'aquestes proteïnes.

Bibliografia

REFERÈNCIES:

- [1] G. Vidarsson, G. Dekkers, and T. Rispen, "IgG subclasses and allotypes: From structure to effector functions," *Front. Immunol.*, vol. 5, no. OCT, p. 520, Oct. 2014, doi: 10.3389/fimmu.2014.00520.
- [2] H. B. Habtamu, T. Not, L. De Leo, S. Longo, L. M. Moretto, and P. Ugo, "Electrochemical immunosensor based on nanoelectrode ensembles for the serological analysis of IgG-type tissue transglutaminase," *Sensors (Switzerland)*, vol. 19, no. 5, p. 1233, Mar. 2019, doi: 10.3390/s19051233.
- [3] W. Chaoharoen *et al.*, "Electrochemical detection of the disease marker human chitinase-3-like protein 1 by matching antibody-modified gold electrodes as label-free immunosensors," *Bioelectrochemistry*, vol. 101, pp. 106–113, Feb. 2015, doi: 10.1016/j.bioelechem.2014.07.006.
- [4] C. Zhu, G. Yang, H. Li, D. Du, and Y. Lin, "Electrochemical Sensors and Biosensors Based on Nanomaterials and Nanostructures," *Anal. Chem.*, vol. 87, p. 36, 2015, doi: 10.1021/ac5039863.
- [5] A. K. Trilling, J. Beekwilder, and H. Zuilhof, "Antibody orientation on biosensor surfaces: A minireview," *Analyst*, vol. 138, no. 6. Royal Society of Chemistry, pp. 1619–1627, Mar. 21, 2013, doi: 10.1039/c2an36787d.
- [6] N. Wang, J. Shang, S. Jiang, and L. Du, "Subunit Vaccines Against Emerging Pathogenic Human Coronaviruses," *Frontiers in Microbiology*, vol. 11. Frontiers Media S.A., p. 298, Feb. 28, 2020, doi: 10.3389/fmicb.2020.00298.
- [7] J. Lan *et al.*, "Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor," *Nature*, vol. 581, no. 7807, pp. 215–220, Mar. 2020, doi: 10.1038/s41586-020-2180-5.
- [8] L. Du, Y. Yang, Y. Zhou, L. Lu, F. Li, and S. Jiang, "MERS-CoV spike protein: a key target for antivirals," *Expert Opinion on Therapeutic Targets*, vol. 21, no. 2. Taylor and Francis Ltd, pp. 131–143, Feb. 01, 2017, doi: 10.1080/14728222.2017.1271415.
- [9] R. De Gasparo *et al.*, "Bispecific IgG neutralizes SARS-CoV-2 variants and prevents escape in mice," *Nature*, vol. 593, no. 7859, pp. 424–428, May 2021, doi: 10.1038/s41586-021-03461-y.
- [10] D. F. Robbiani *et al.*, "Convergent antibody responses to SARS-CoV-2 in convalescent individuals," *Nature*, vol. 584, no. 7821, p. 437, Jun. 2020, doi: 10.1038/s41586-020-2456-9.
- [11] W. Yu *et al.*, "Mechanistic Insights to the Binding of Antibody CR3022 Against RBD from SARS-CoV and HCoV-19/SARS-CoV-2: A Computational Study," *Comb. Chem. High Throughput Screen.*, vol. 23, Oct. 2020, doi: 10.2174/1386207323666201026160500.

- [12] M. Tuccori *et al.*, “Anti-SARS-CoV-2 neutralizing monoclonal antibodies: clinical pipeline,” *mAbs*, vol. 12, no. 1. Bellwether Publishing, Ltd., 2020, doi: 10.1080/19420862.2020.1854149.
- [13] M. A. Tortorici *et al.*, “Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms,” *Science (80-.)*, vol. 370, no. 6519, pp. 950–957, Nov. 2020, doi: 10.1126/science.abe3354.
- [14] M. A. Cowie, “Estudi conformacional del complexe anticòs-antigen format entre la glicoproteïna de l'espícula del virus SARS-CoV-2 i diferents anticossos,” 2021.
- [15] L. A. Layqah and S. Eissa, “An electrochemical immunosensor for the corona virus associated with the Middle East respiratory syndrome using an array of gold nanoparticle-modified carbon electrodes,” *Microchim. Acta*, vol. 186, no. 4, pp. 1–10, Apr. 2019, doi: 10.1007/s00604-019-3345-5.
- [16] F. Yang, Y. Li, X. Jin, Q. Xu, F. Cheng, and X. Wang, “Immunosensor-based rapid quantitative detection of Newcastle disease virus antibodies using innovative gold immunochromatographic assay,” *J. Appl. Microbiol.*, vol. 129, no. 6, pp. 1751–1757, Dec. 2020, doi: 10.1111/jam.14688.
- [17] H. Haji-Hashemi, P. Norouzi, M. R. Safarnejad, and M. R. Ganjali, “Label-free electrochemical immunosensor for direct detection of Citrus tristeza virus using modified gold electrode,” *Sensors Actuators, B Chem.*, vol. 244, pp. 211–216, Jun. 2017, doi: 10.1016/j.snb.2016.12.135.
- [18] J. Shang *et al.*, “Structural basis of receptor recognition by SARS-CoV-2,” *Nature*, vol. 581, 2020, doi: 10.1038/s41586-020-2179-y.
- [19] J. Shang *et al.*, “Cell entry mechanisms of SARS-CoV-2,” doi: 10.1073/pnas.2003138117/-/DCSupplemental.
- [20] “Details - Public Health Image Library(PHIL).” <https://phil.cdc.gov/Details.aspx?pid=23312> (accessed May 31, 2021).
- [21] F. Li, “Structure, Function, and Evolution of Coronavirus Spike Proteins,” *Annual Review of Virology*, vol. 3. Annual Reviews Inc., pp. 237–261, Sep. 29, 2016, doi: 10.1146/annurev-virology-110615-042301.
- [22] Y. Yuan *et al.*, “Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains,” *Nat. Commun.*, vol. 8, no. 1, pp. 1–9, Apr. 2017, doi: 10.1038/ncomms15092.
- [23] A. Natsume, R. Niwa, and M. Satoh, “Improving effector functions of antibodies for cancer treatment: Enhancing ADCC and CDC,” *Drug Design, Development and Therapy*, vol. 3, no. 3. Drug Des Devel Ther, pp. 7–16, 2009, doi: 10.2147/dddt.s4378.
- [24] Y. Miao, F. Feixas, C. Eun, and J. A. McCammon, “Accelerated molecular dynamics simulations of protein folding,” *J. Comput. Chem.*, vol. 36, no. 20, pp. 1536–1549, 2015, doi: 10.1002/jcc.23964.
- [25] P. M. Hogarth and G. A. Pietersz, “Fc receptor-targeted therapies for the

- treatment of inflammation, cancer and beyond,” *Nature Reviews Drug Discovery*, vol. 11, no. 4. Nature Publishing Group, pp. 311–331, Apr. 30, 2012, doi: 10.1038/nrd2909.
- [26] D. Pinto *et al.*, “Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody,” *Nature*, vol. 583, no. 7815, pp. 290–295, Jul. 2020, doi: 10.1038/s41586-020-2349-y.
- [27] L. Rieznichenko, “Features of gold nanoparticles application for bio- and nanosensors,” *Int. J. Biosens. Bioelectron.*, vol. 4, no. 2, pp. 39–40, 2018, doi: 10.15406/ijbsbe.2018.04.00095.
- [28] A. Kletsov, Y. Dahnovsky, and J. V. Ortiz, “Surface Green’s function calculations: A nonrecursive scheme with an infinite number of principal layers,” *J. Chem. Phys.*, vol. 126, no. 13, 2007, doi: 10.1063/1.2713743.
- [29] F. Hanke and J. Björk, “Structure and local reactivity of the Au(111) surface reconstruction,” *Phys. Rev. B. Condens. Matter Mater. Phys.*, vol. 87, no. 87, p. 235422, 2013, doi: 10.1103/PhysRevB.87.235422.
- [30] “RCSB PDB - 1HZH: CRYSTAL STRUCTURE OF THE INTACT HUMAN IGG B12 WITH BROAD AND POTENT ACTIVITY AGAINST PRIMARY HIV-1 ISOLATES: A TEMPLATE FOR HIV VACCINE DESIGN.” <https://www.rcsb.org/structure/1hzh> (accessed Jun. 06, 2021).
- [31] E. F. Pettersen *et al.*, “UCSF Chimera - A visualization system for exploratory research and analysis,” *J. Comput. Chem.*, vol. 25, no. 13, pp. 1605–1612, Oct. 2004, doi: 10.1002/jcc.20084.
- [32] “RCSB PDB - 6YLA: Crystal structure of the SARS-CoV-2 receptor binding domain in complex with CR3022 Fab.” <https://www.rcsb.org/structure/6yla> (accessed Jun. 06, 2021).
- [33] “RCSB PDB - 6WPT: Structure of the SARS-CoV-2 spike glycoprotein in complex with the S309 neutralizing antibody Fab fragment (open state).” <https://www.rcsb.org/structure/6WPT> (accessed Jun. 06, 2021).
- [34] “Force Fields for MD simulations.”
- [35] D. A. Case Ross C Walker and T. E. Darden Junmei Wang Robert Duke, “Amber 2018 Reference Manual Principal contributors to the current codes.” Accessed: May 24, 2021. [Online]. Available: <http://ambermd.org/contributors.html>.
- [36] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, “Comparison of simple potential functions for simulating liquid water,” *J. Chem. Phys.*, vol. 79, no. 2, pp. 926–935, Aug. 1983, doi: 10.1063/1.445869.
- [37] H. Heinz, R. A. Vaia, B. L. Farmer, and R. R. Naik, “Accurate simulation of surfaces and interfaces of face-centered cubic metals using 12-6 and 9-6 lennard-jones potentials,” *J. Phys. Chem. C*, vol. 112, no. 44, pp. 17281–17290, Nov. 2008, doi: 10.1021/jp801931d.
- [38] L. Yang *et al.*, “New-generation Amber united-atom force field,” *J. Phys. Chem. B*, vol. 110, no. 26, pp. 13166–13176, Jul. 2006, doi: 10.1021/jp060163v.

- [39] M. J. Frisch, "Gaussian~16 Revision C.01." Wallingford CT, 2016.
- [40] "Force Field Parameters."
<http://www.ks.uiuc.edu/Research/namd/2.9/ug/node21.html> (accessed May 31, 2021).
- [41] R. Rennie, *A Dictionary of Chemistry*. Oxford University Press, 2016.
- [42] "The physical models for the AMBER molecular mechanics force field.... | Download Scientific Diagram." https://www.researchgate.net/figure/The-physical-models-for-the-AMBER-molecular-mechanics-force-field-Atoms-and-bonds-are_fig1_5773728 (accessed May 31, 2021).
- [43] T. F. Koetzle, M. S. Lehmann, J. J. Verbist, and W. C. Hamilton, "Precision neutron diffraction structure determination of protein and nucleic acid components. VII. The crystal and molecular structure of the amino acid L-lysine monohydrochloride dihydrate," *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.*, vol. 28, no. 11, pp. 3207–3214, Nov. 1972, doi: 10.1107/s0567740872007757.
- [44] M. J. Frisch, "Gaussian 09, Revision A.02." Wallingford CT, 2009.
- [45] E. Vanquelef *et al.*, "R.E.D. Server: A web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments," *Nucleic Acids Res.*, vol. 39, no. SUPPL. 2, pp. W511–W517, Jul. 2011, doi: 10.1093/nar/gkr288.
- [46] F. Y. Dupradeau *et al.*, "The R.E.D. tools: Advances in RESP and ESP charge derivation and force field library building," *Phys. Chem. Chem. Phys.*, vol. 12, no. 28, pp. 7821–7839, Jul. 2010, doi: 10.1039/c0cp00111b.
- [47] C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman, "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model," *J. Phys. Chem.*, vol. 97, no. 40, pp. 10269–10280, 1993, doi: 10.1021/j100142a004.
- [48] K. Kim and K. D. Jordan, "Comparison of density functional and MP2 calculations on the water monomer and dimer," *Journal of Physical Chemistry*, vol. 98, no. 40. American Chemical Society, pp. 10089–10094, 1994, doi: 10.1021/j100091a024.
- [49] Y. Yang, M. N. Weaver, and K. M. Merz, "Assessment of the '6-31+Gt; + LANL2DZ' mixed basis set coupled with density functional theory methods and the effective core potential: Prediction of heats of formation and ionization potentials for first-row-transition-metal complexes," *J. Phys. Chem. A*, vol. 113, no. 36, pp. 9843–9851, Sep. 2009, doi: 10.1021/jp807643p.
- [50] L. B. Kier, "Molecular Modeling: Principles and Applications By Andrew Leach. Longman, Edinburgh. 1996. xvi + 595 pp. ISBN 0-582-23933-8. \$35.00.," *J. Med. Chem.*, vol. 40, no. 18, pp. 2969–2969, Aug. 1997, doi: 10.1021/jm970383k.
- [51] C. Oostenbrink, M. van L.-... M. C. II, and undefined 2007, "Applications of molecular dynamics simulations in drug design," *dlab.cl*, Accessed: May 31,

2021. [Online]. Available: http://dlab.cl/wp-content/uploads/2015/11/05_Oostenbrinnk_et_al_2007.pdf.
- [52] H. Goldstein, C. Poole, J. Safko, and S. R. Addison, “Classical Mechanics, 3rd ed.,” *Am. J. Phys.*, vol. 70, no. 7, pp. 782–783, Jul. 2002, doi: 10.1119/1.1484149.
- [53] R. H.-M. C. Phys. and undefined 1970, “The potential calculation and some applications,” *ci.nii.ac.jp*, Accessed: Jun. 01, 2021. [Online]. Available: <https://ci.nii.ac.jp/naid/10014832156/>.
- [54] L. Verlet, “Computer ‘experiments’ on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules,” *Phys. Rev.*, vol. 159, no. 1, pp. 98–103, Jul. 1967, doi: 10.1103/PhysRev.159.98.
- [55] J.-P. Ryckaert, G. Ciccotti+, and H. J. C. Berendsen, “Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes,” 1977. Accessed: Jun. 01, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0021999177900985>.
- [56] K. Roy, S. Kar, and R. N. Das, “Computational Chemistry,” in *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*, Elsevier, 2015, pp. 151–189.
- [57] J. C. Phillips *et al.*, “Scalable molecular dynamics on CPU and GPU architectures with NAMD,” *J. Chem. Phys.*, vol. 153, no. 4, Jul. 2020, doi: 10.1063/5.0014475.
- [58] “Conjugate Gradient Method and it’s Application in Energy Minimization | by Akshay Shah | Medium.” <https://akshayshah-96720.medium.com/conjugate-gradient-m-dfd9d2969e73> (accessed Jun. 02, 2021).
- [59] S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, “Constant pressure molecular dynamics simulation: The Langevin piston method,” *J. Chem. Phys.*, vol. 103, no. 11, pp. 4613–4621, Jun. 1995, doi: 10.1063/1.470648.
- [60] A. Rudin and P. Choi, “Diffusion in Polymers,” in *The Elements of Polymer Science & Engineering*, Elsevier, 2013, pp. 275–304.
- [61] “Temperature Control and Equilibration.” <http://www.ks.uiuc.edu/Research/namd/2.9/ug/node36.html> (accessed Jun. 02, 2021).
- [62] “NAMD 2.13 User’s Guide.” <https://www.ks.uiuc.edu/Research/namd/2.13/ug/> (accessed May 28, 2021).
- [63] G. J. Martyna, D. J. Tobias, and M. L. Klein, “Constant pressure molecular dynamics algorithms,” *J. Chem. Phys.*, vol. 101, no. 5, pp. 4177–4189, Aug. 1994, doi: 10.1063/1.467468.
- [64] S. Chakraborty, J. Zhang, and S. Ghosh, “Accelerated molecular dynamics simulations for characterizing plastic deformation in crystalline materials with cracks,” *Comput. Mater. Sci.*, vol. 121, pp. 23–34, Aug. 2016, doi: 10.1016/j.commatsci.2016.04.026.

- [65] D. R. Roe and T. E. Cheatham, "PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data," *J. Chem. Theory Comput.*, vol. 9, no. 7, pp. 3084–3095, Jul. 2013, doi: 10.1021/ct400341p.
- [66] D. Martí, J. Ainsley, O. Ahumada, C. Aleman, and J. Torras, "Tethering of the IgG1 antibody to amorphous silica for immunosensor development: A molecular dynamics study," *Langmuir*, vol. 36, no. 42, pp. 12658–12667, 2020, doi: 10.1021/acs.langmuir.0c02203.
- [67] B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, and A. E. Roitberg, "MMPBSA.py: An efficient program for end-state free energy calculations," *J. Chem. Theory Comput.*, vol. 8, no. 9, pp. 3314–3321, Sep. 2012, doi: 10.1021/ct300418h.
- [68] A. Onufriev, D. Bashford, and D. A. Case, "Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model," *Proteins Struct. Funct. Genet.*, vol. 55, no. 2, pp. 383–394, May 2004, doi: 10.1002/prot.20033.
- [69] Y. Takano and C. Stanton, "AMBER Tutorial DNA:DNA, polyA-polyT," 2003. Accessed: Jun. 02, 2021. [Online]. Available: <http://ndbserver.rutgers.edu/>.
- [70] Y. Miao, W. Sinko, L. Pierce, D. Bucher, R. C. Walker, and J. A. McCammon, "Improved reweighting of accelerated molecular dynamics simulations for free energy calculation," *J. Chem. Theory Comput.*, vol. 10, no. 7, pp. 2677–2689, Jul. 2014, doi: 10.1021/ct500090q.
- [71] M. De Trabajo, "Disposición 14977 del BOE núm. 251 de 2019," 2019. Accessed: Jun. 15, 2021. [Online]. Available: <http://www.boe.es>.
- [72] "(No Title)." https://www.csuc.cat/sites/default/files/2020-12/Tarifas_2021.pdf (accessed Jun. 15, 2021).

Annex

A. Residus CDR dels sistemes.

En les següents taules queda recollida la informació dels residus que formen part de les regions CDR del sistema. Inclou el número de residu corresponent en el sistema dissenyat, el tipus d'aminoàcid amb la nomenclatura estàndard de tres lletres i el ID del residu del PDB original.

Taula 18: Grups de residus que formen part del CDR de la cadena pesada de la regió FabI del sistema CR3022.

CDR-H.1			CDR-H.2			CDR-H.3		
Num	AA	6YLA H	Num	AA	6YLA H	Num	AA	6YLA H
32	THR	31	51	ILE	50	100	GLY	99
33	TYR	32	52	ILE	51	101	SER	100
34	TRP	33	53	TYR	52	102	GLY	101
35	ILE	34	54	PRO	53	103	ILE	102
36	GLY	35	55	GLY	54	104	SER	103
			56	ASP	55	105	THR	104
			57	SER	56	106	PRO	105
			58	GLU	57	107	MET	106
			59	THR	58	108	ASP	107
			60	ARG	59	109	VAL	108
			61	TYR	60			
			62	SER	61			
			63	PRO	62			
			64	SER	63			
			65	PHE	64			
			66	GLN	65			
			67	GLY	66			

Taula 19: Residus que formen part del CDR de la cadena lleugera de la regió FabI del sistema CR3022.

CDR-L.1			CDR-L.2			CDR-L.3		
Num	AA	6YLA L	Num	AA	6YLA L	Num	AA	6YLA L
1135	LYS	24	1167	TRP	56	1206	GLN	95
1136	SER	25	1168	ALA	57	1207	GLN	96
1137	SER	26	1169	SER	58	1208	TYR	97
1138	GLN	27	1170	THR	59	1209	TYR	98
1139	SER	28	1171	ARG	60	1210	SER	99
1140	VAL	29	1172	GLU	61	1211	THR	100
1141	LEU	30	1173	SER	62	1212	PRO	101

1142	TYR	31	1213	TYR	102
1143	SER	32	1214	THR	103
1144	SER	33			
1145	ILE	34			
1146	ASN	35			
1147	LYS	36			
1148	ASN	37			
1149	TYR	38			
1150	LEU	39			
1151	ALA	40			

Taula 20: Grups de residus que formen part del CDR de la cadena pesada de la regió Fab2 del sistema CR3022.

CDR-H.1			CDR-H.2			CDR-H.3		
Num	AA	6YLA H	Num	AA	6YLA H	Num	AA	6YLA H
480	THR	31	499	ILE	50	548	GLY	99
481	TYR	32	500	ILE	51	549	SER	100
482	TRP	33	501	TYR	52	550	GLY	101
483	ILE	34	502	PRO	53	551	ILE	102
484	GLY	35	503	GLY	54	552	SER	103
			504	ASP	55	553	THR	104
			505	SER	56	554	PRO	105
			506	GLU	57	555	MET	106
			507	THR	58	556	ASP	107
			508	ARG	59	557	VAL	108
			509	TYR	60			
			510	SER	61			
			511	PRO	62			
			512	SER	63			
			513	PHE	64			
			514	GLN	65			
			515	GLY	66			

Taula 21: Residus que formen part del CDR de la cadena lleugera de la regió Fab2 del sistema CR3022.

CDR-L.1			CDR-L.2			CDR-L.3		
Num	AA	6YLA L	Num	AA	6YLA L	Num	AA	6YLA L
914	LYS	24	946	TRP	56	985	GLN	95
915	SER	25	947	ALA	57	986	GLN	96
916	SER	26	948	SER	58	987	TYR	97
917	GLN	27	949	THR	59	988	TYR	98
918	SER	28	950	ARG	60	989	SER	99

919	VAL	29	951	GLU	61	990	THR	100
920	LEU	30	952	SER	62	991	PRO	101
921	TYR	31				992	TYR	102
922	SER	32				993	THR	103
923	SER	33						
924	ILE	34						
925	ASN	35						
926	LYS	36						
927	ASN	37						
928	TYR	38						
929	LEU	39						
930	ALA	40						

Taula 22: Residus que formen part del CDR de la cadena pesada 1 de la regió Fc del sistema CR3022.

CDR-H1.1			CDR-H1.2			CDR-H1.3		
Num	AA	1HZH H	Num	AA	1HZH H	Num	AA	1HZH H
359	LEU	368	384	SER	393	414	ASP	423
360	THR	369	385	ASN	394	415	LYS	424
361	LYS	370	386	GLY	395	416	SER	425
362	ASN	371	387	GLN	396	417	ARG	426
363	GLN	372	388	PRO	397	418	TRP	427
			389	GLU	398	419	GLN	428
			390	ASN	399	420	GLN	429
			391	ASN	400	421	GLY	430
			392	TYR	401	422	ASN	431
						423	VAL	432

Taula 23: Residus que formen part del CDR de la cadena pesada 2 de la regió Fc del sistema CR3022.

CDR-H2.1			CDR-H2.2			CDR-H2.3		
Num	AA	1HZH K	Num	AA	1HZH K	Num	AA	1HZH K
804	LEU	365	829	SER	390	859	ASP	420
805	THR	366	830	ASN	391	860	LYS	421
806	LYS	367	831	GLY	392	861	SER	422
807	ASN	368	832	GLN	393	862	ARG	423
808	GLN	369	833	PRO	394	863	TRP	424
			834	GLU	395	864	GLN	425
			835	ASN	396	865	GLN	426
			836	ASN	397	866	GLY	427
			837	TYR	398	867	ASN	428
						868	VAL	429

Taula 24: Grups de residus que formen part del CDR de la cadena pesada de la regió Fab1 del sistema S309.

CDR-H.1			CDR-H.2			CDR-H.3		
Num	AA	6WPT H	Num	AA	6WPT H	Num	AA	6WPT H
31	TYR	31	50	ILE	50	99	TYR	99
32	GLY	32	51	SER	51	100	THR	100
33	ILE	33	52	THR	52	101	ARG	101
34	SER	34	53	TYR	53	102	GLY	102
35	TRP	35	54	ASN	54	103	ALA	103
			55	GLY	55	104	TRP	104
			56	ASN	56	105	PHE	105
			57	THR	57	106	GLY	106
			58	ASN	58	107	GLU	107
			59	TYR	59	108	SER	108
			60	ALA	60	109	LEU	109
			61	GLN	61	110	ILE	110
			62	LYS	62	111	GLY	111
			63	PHE	63	112	GLY	112
			64	GLN	64	113	PHE	113
			65	GLY	65	114	ASP	114
			66	ARG	66	115	ASN	115
						116	TRP	116

Taula 25: Residus que formen part del CDR de la cadena lleugera de la regió Fab1 del sistema S309.

CDR-L.1			CDR-L.2			CDR-L.3		
Num	AA	6WPT L	Num	AA	6WPT L	Num	AA	6WPT L
923	ALA	24	950	ALA	51	989	GLN	90
924	SER	25	951	SER	52	990	HIE	91
925	GLN	26	952	SER	53	991	ASP	92
926	THR	27	953	ARG	54	992	THR	93
927	VAL	28	954	ALA	55	993	SER	94
928	SER	29	955	THR	56	994	LEU	95
929	SER	30	956	GLY	57	995	THR	96
930	THR	31				996	PHE	97
931	SER	32						
932	LEU	33						
933	ALA	34						
934	TRP	35						

Taula 26: Grups de residus que formen part del CDR de la cadena pesada de la regió Fab2 del sistema S309

CDR-H.1			CDR-H.2			CDR-H.3		
Num	AA	6WPT H	Num	AA	6WPT H	Num	AA	6WPT H
487	TYR	31	506	ILE	50	555	TYR	99
488	GLY	32	507	SER	51	556	THR	100
489	ILE	33	508	THR	52	557	ARG	101
490	SER	34	509	TYR	53	558	GLY	102
491	TRP	35	510	ASN	54	559	ALA	103
			511	GLY	55	560	TRP	104
			512	ASN	56	561	PHE	105
			513	THR	57	562	GLY	106
			514	ASN	58	563	GLU	107
			515	TYR	59	564	SER	108
			516	ALA	60	565	LEU	109
			517	GLN	61	566	ILE	110
			518	LYS	62	567	GLY	111
			519	PHE	63	568	GLY	112
			520	GLN	64	569	PHE	113
			521	GLY	65	570	ASP	114
			522	ARG	66	571	ASN	115
						572	TRP	116

Taula 27: Residus que formen part del CDR de la cadena lleugera de la regió Fab2 del sistema S309.

CDR-L.1			CDR-L.2			CDR-L.3		
Num	AA	6WPT L	Num	AA	6WPT L	Num	AA	6WPT L
1137	ALA	24	1164	ALA	51	1203	GLN	90
1138	SER	25	1165	SER	52	1204	HIE	91
1139	GLN	26	1166	SER	53	1205	ASP	92
1140	THR	27	1167	ARG	54	1206	THR	93
1141	VAL	28	1168	ALA	55	1207	SER	94
1142	SER	29	1169	THR	56	1208	LEU	95
1143	SER	30	1170	GLY	57	1209	THR	96
1144	THR	31				1210	PHE	97
1145	SER	32						
1146	LEU	33						
1147	ALA	34						
1148	TRP	35						

Taula 28: Residus que formen part del CDR de la cadena pesada 1 de la regió Fc del sistema S309.

CDR-H1.1			CDR-H1.2			CDR-H1.3		
Num	AA	1HZH H	Num	AA	1HZH H	Num	AA	1HZH H
367	LEU	354	392	SER	379	422	ASP	409
368	THR	355	393	ASN	380	423	LYS	410
369	LYS	356	394	GLY	381	424	SER	411
370	ASN	357	395	GLN	382	425	ARG	412
371	GLN	358	396	PRO	383	426	TRP	413
			397	GLU	384	427	GLN	414
			398	ASN	385	428	GLN	415
			399	ASN	386	429	GLY	416
			400	TYR	387	430	ASN	417
						431	VAL	418

Taula 29: Residus que formen part del CDR de la cadena pesada 2 de la regió Fc del sistema S309.

CDR-H2.1			CDR-H2.2			CDR-H2.3		
Num	AA	1HZH K	Num	AA	1HZH K	Num	AA	1HZH K
813	LEU	344	838	SER	369	868	ASP	399
814	THR	345	839	ASN	370	869	LYS	400
815	LYS	346	840	GLY	371	870	SER	401
816	ASN	347	841	GLN	372	871	ARG	402
817	GLN	348	842	PRO	373	872	TRP	403
			843	GLU	374	873	GLN	404
			844	ASN	375	874	GLN	405
			845	ASN	376	875	GLY	406
			846	TYR	377	876	ASN	407
						877	VAL	408

B. Gràfiques RMSD i RMSF del sistema clàssic

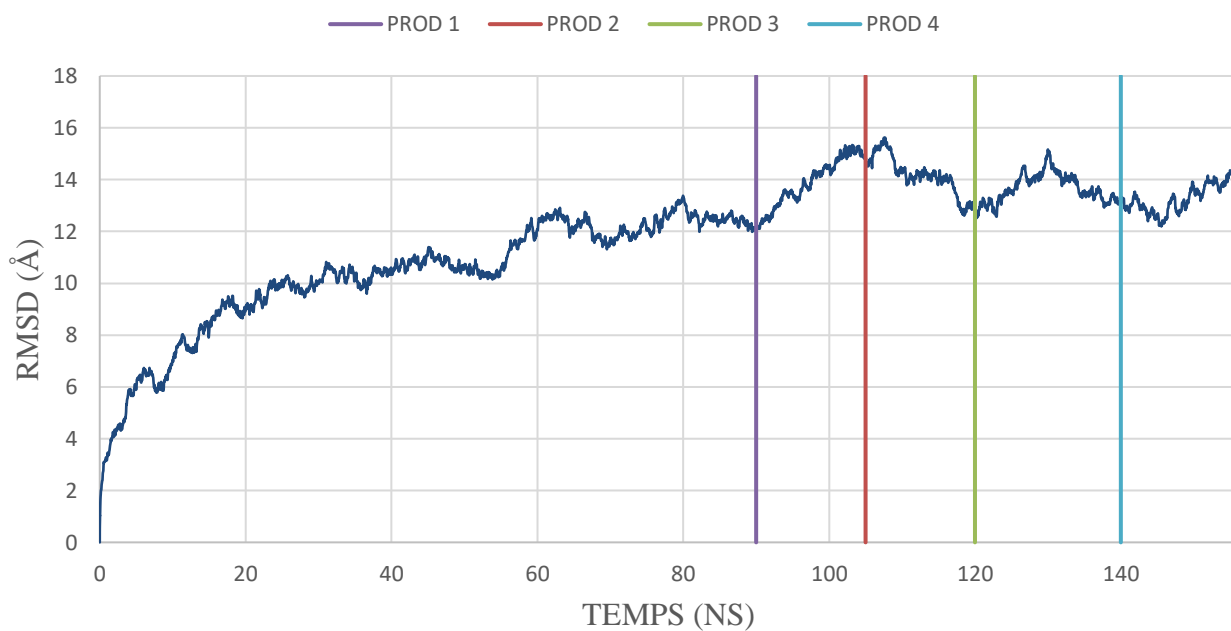


Figura 37: Gràfica del RMSD obtingut dels 150ns de simulació clàssica en front del temps del sistema CR3022. A partir de les línies verticals s'indiquen els punts del sistema escollits per les quatre produccions AMD.

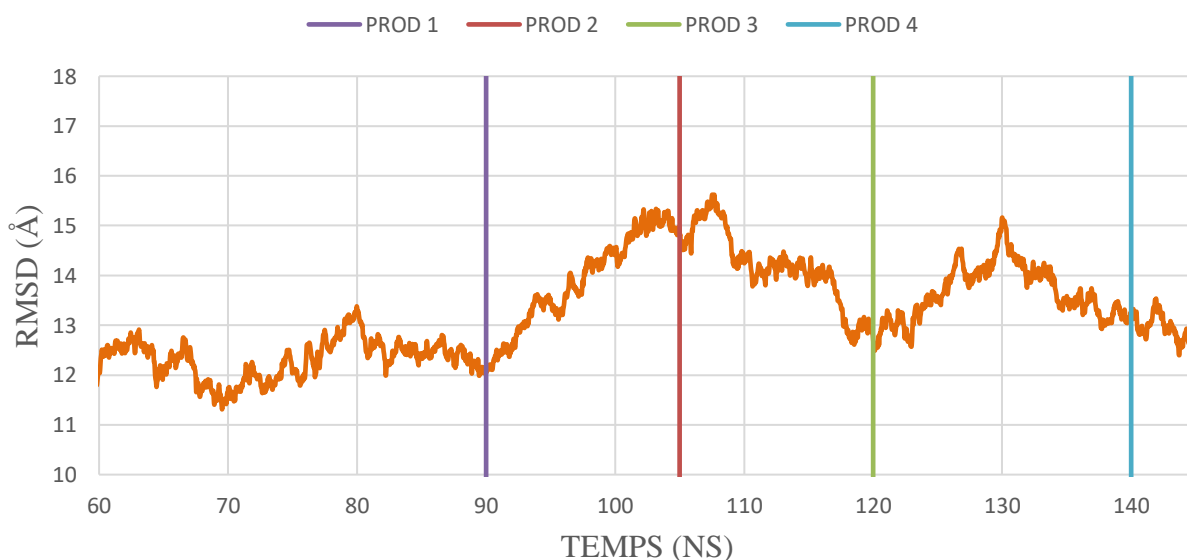


Figura 38: Gràfic ampliament al tram final de la trajectòria clàssica simulada del RMSD respecte al temps pel sistema CR3022 on s'observen de manera més ampla els valors que prenen les conformacions triades.

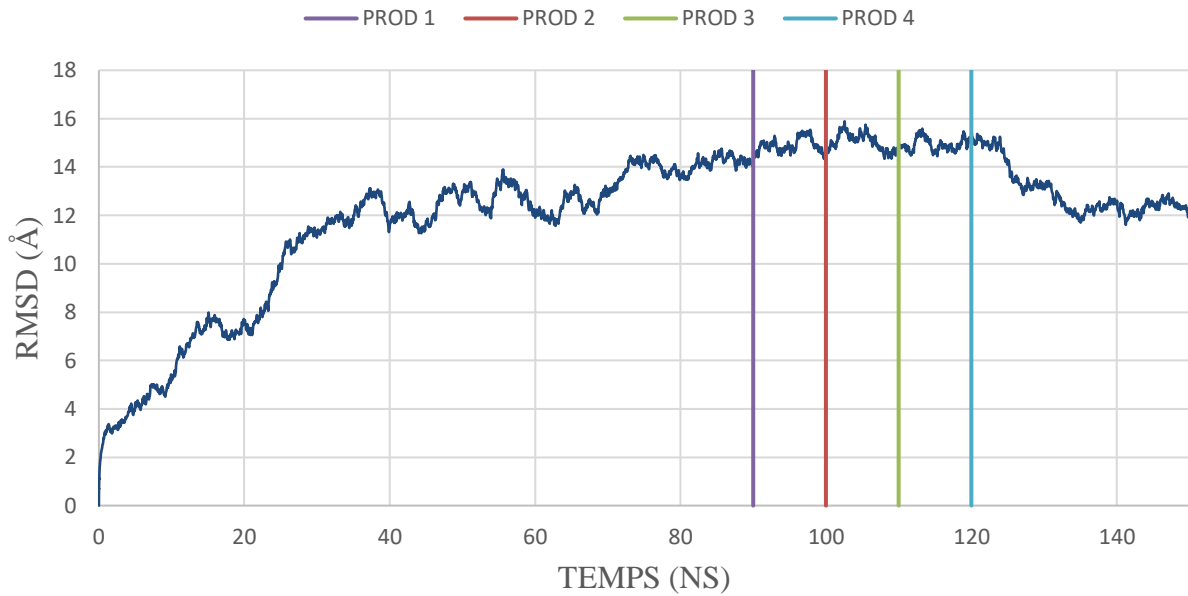


Figura 39: Gràfica del RMSD obtingut dels 150ns de simulació clàssica en front del temps del sistema S309. A partir de les línies verticals s'indiquen els punts del sistema escollits per les quatre produccions amd.

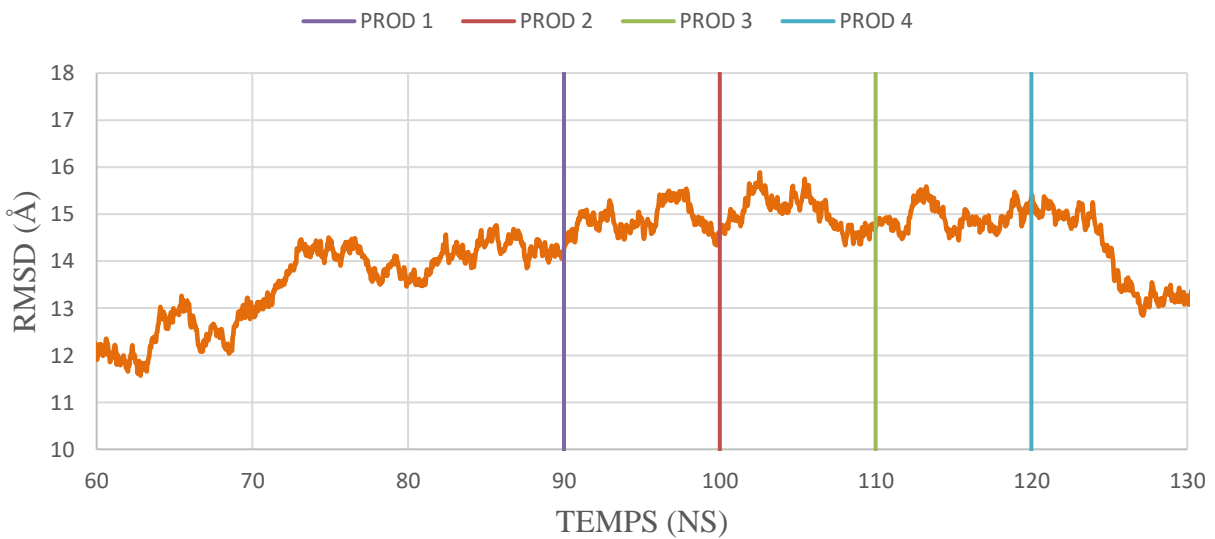


Figura 40: Gràfic ampli al tram final de la trajectòria clàssica simulada del RMSD respecte el temps pel sistema S309 on s'observen de manera més amplia els valors que prenen les conformacions triades.

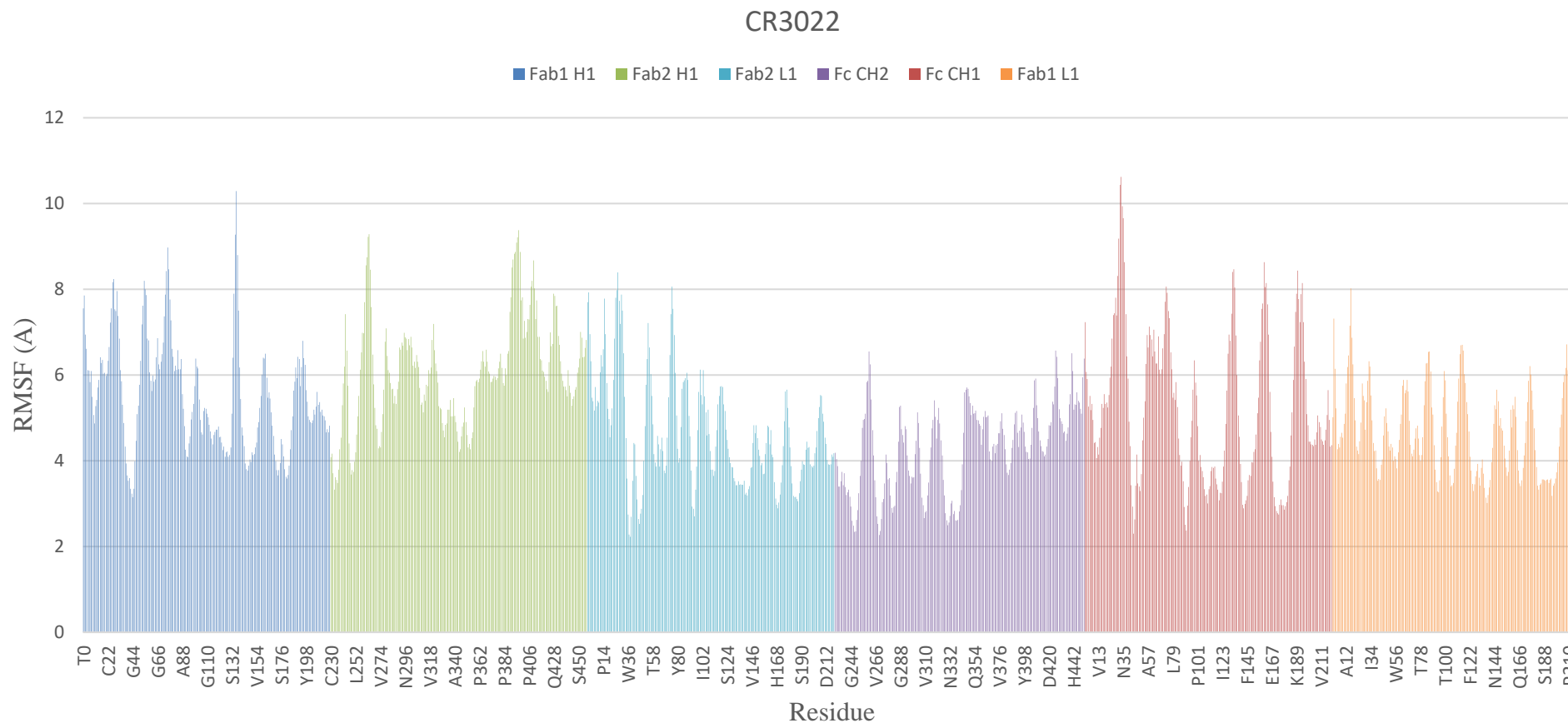


Figura 41: Gràfic sobre el valor del RMSF per cada residu del sistema CR3022. En diferents colors queden diferenciades les regions estudiades.

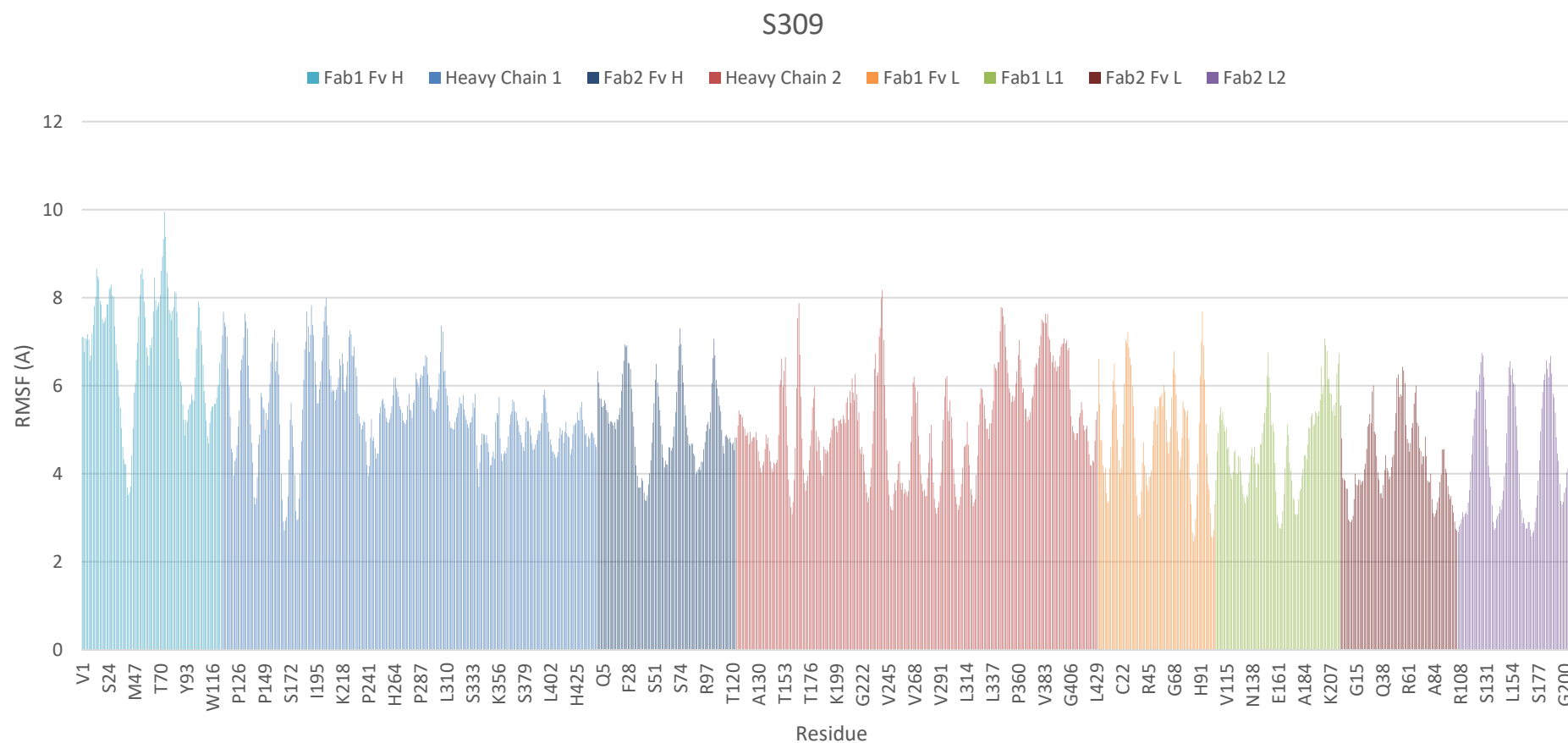


Figura 42: Gràfic sobre el valor del RMSF per cada residu del sistema S309. En diferents colors queden diferenciades les regions estudiades.