

# **Biofísica i immunoteràpia contra el càncer**

Interacció entre cèl·lules T i hidrogels sintetitzats amb heparina

Maximilià Soler i Subirana  
1497326

Treball supervisat per Jordi Faraudo, científic titular a l'ICMAB

Treball de Fi de Grau en Física  
Convocatòria pel Juliol de 2023

**UAB**  
**Universitat Autònoma de Barcelona**  
**Facultat de Ciències**  
Departament de Física  
Barcelona, Espanya  
15 de Setembre de 2022

# **Biofísica i immunoteràpia contra el càncer**

Interacció entre cèl·lules T i hidrogels sintetitzats amb heparina

**Maximilià Soler i Subirana**

## **Resum**

# Sumari

<b>1</b>	<b>Introducció i objectius</b>	<b>3</b>
1.1	Motius . . . . .	3
1.2	Procediments . . . . .	3
1.3	Electrostàtica . . . . .	3
1.4	Hipòtesis i objectiu final . . . . .	4
<b>2</b>	<b>Mètode</b>	<b>5</b>
2.1	Equació de Poisson-Boltzmann . . . . .	5
2.2	PB per a medis salins . . . . .	7
<b>3</b>	<b>Resultats</b>	<b>9</b>
3.1	Sel·lecció del model proteic . . . . .	9
3.2	Càlcul electrostàtic superficial de la proteïna . . . . .	9
3.3	Heparina . . . . .	11
3.4	Encaix de molècules (molecular docking) . . . . .	12
<b>4</b>	<b>Conclusions i comentari final</b>	<b>13</b>
<b>A</b>	<b>Apèndix</b>	<b>14</b>
A.1	Bibliografia . . . . .	14
A.2	Annexos . . . . .	15
A.2.1	Estudi i sel·lecció del millor model virtual . . . . .	15
A.2.2	Obtenció dels arxius per l'APBS . . . . .	18
A.3	Índex alfàbetic . . . . .	20
A.4	Declaracions . . . . .	21
A.4.1	Declaració d'autoria del Treball de Fi de Grau . . . . .	21
A.4.2	Declaració d'Extensió del Treball de Grau . . . . .	22
A.4.3	Declaració Supervisió Treball de Grau . . . . .	23

# 1 Introducció i objectius

## 1.1 Motius

En el present treball s'estudia la interacció, suposadament electroestàtica, entre dues molècules, la proteïna CCL21 i un hidrogel compost principalment d'heparina, per a la millora d'una de les noves teràpies contra el càncer, la immunoteràpia.

Aquesta consisteix, a gran escala, en millorar l'efectivitat de les cèl·lules T del mateix pacient que pateix un càncer i les defenses del qual no poden lluitar adequadament contra el tumor. Per a poder millorar la condició d'aquestes cèl·lules, és necessari extreure-les i tractar-les al laboratori. Això implica haver de trobar un entorn apte per a les cèl·lules T, que ha de ser similar al de l'interior del cos humà.

Per al correcte desenvolupament d'aquestes, necessitarem descobrir i estudiar les principals interaccions entre les seves membranes, que és per on poden unir-se i interactuar amb el medi, i els diferents teixits del cos humà. Aquí és on indaga aquest treball: l'estudi de la interacció electroestàtica d'una proteïna de la membrana cel·lular altament electropositiva amb un entorn similar al d'un gangli limfàtic, altament electronegatiu.

## 1.2 Procediments

Mitjançant extraccions de sang i després els corresponents mètodes separadors de tipus de cèl·lules (afèresi), és possible aïllar una col·lecció de cèl·lules T del pacient el qual volem tractar. En comptes d'utilitzar recipients coneguts com la placa de Petri (considerat un mètode bidimensional), el que es vol és cultivar aquestes cèl·lules de forma tridimensional. Aquests necessiten un entorn similar al d'un gangli per a poder multiplicar-se, que anomenem matriu extracel·lular. Un cop s'han seleccionat adequadament, es comprova que responguin adequadament a les proteïnes del càncer i que siguin prou efectives. Finalment, aquestes cèl·lules són retornades al pacient per a lluitar contra el tumor en qüestió.

En tractar-se de cèl·lules del propi pacient, la taxa de rebuig és mínima, pel que normalment té una eficiència molt superior a la dels altres mètodes més agressius [1].

## 1.3 Electroestàtica

Dins les cèl·lules hi ha una quantitat inimaginable de macromolècules i altres components repartides pel medi intracel·lular. Aquestes, formades per diferents parts unides per diversos tipus d'enllaços, poden perdre algunes zones connectades feblement al cos principal. Això és a causa de l'agitació tèrmica: al món microscòpic, tots els components estan en moviment a causa de la temperatura. Aquests, generalment àtoms o petites molècules, solen abandonar electrons a l'estructura principal d'on s'han separat, que transforma aquesta en una molècula carregada negativament, anomenada macroió. És per això que diem que aquesta molècula queda carregada. Els àtoms que han marxat s'anomenen contraions, ja que la suma de totes les seves càrregues cancel·la la del macroió.

Un exemple interessant aplicable al nostre camp seria el de l'ADN, que en els medis intracel·lulars, els aminoàcids que el componen acostumen a ionitzar-se en interactuar amb aquest entorn salí. Això és possible perquè no presenten enllaços covalents, és a dir que l'energia que comparteixen junts és inferior a la que tenen per separat, i, normalment l'energia sempre tendeix a romandre a l'estat més baix i estable possible.

Aquests ions ara confronten una dicotomia: si s'allunyen generen entropia, però per a dispersar-se necessiten energia ja que provenen d'una molècula que han deixat carregada. El resultat final, és doncs, que adquireixen la màxima entropia que els és possible amb la mínima energia, com hem vist a l'exemple anterior. Dins els medis cel·lulars que es volen estudiar, acaba passant un fenomen conegut com a capa difusa de càrrega. En altres paraules, els àtoms formen una espècie de núvol que rodeja el macroió, generalment a una distància de pocs nanòmetres, que cancel·la la càrrega més enllà d'aquests, creant una càrrega total neutra vista des de fora [2]. És a dir, un altre macroió proper no interaccionarà fins a atravesar aquesta capa que contrarresta l'energia del macroió en qüestió. Aquests àtoms del núvol juntament amb les càrregues a la superfície del macroió es coneix com doble capa elèctrica.

Considerant el solvent tant a temperatura ambient (293.15K) com corporal (309.65K), el sistema presenta certa energia tèrmica. Ara prenem en compte  $k_B T^\circ$ , factor àmpliament utilitzat per a representar l'energia

tèrmica d'un sistema: quan tenim una energia total igual o superior a aquesta, les vibracions s'elevan i permeten la separació d'un element respecte la molècula principal. En altres paraules, si tenim  $E_{total} \geq k_B T^o$ , diem que tenim un sistema ionitzable. A més, els graus de llibertat per component es defineix com  $1/2k_B T$ , que ens diu de quines formes es pot moure aquest en un sistema.

És interessant fixar-se que, a causa de les condicions esmentades anteriorment, quan dos macroions siguin suficientment a prop un de l'altre, seran crucials les interaccions amb cada component pertanyent al nívol i la geometria de la superfície de cada macroió, ja que són els principals responsables de la forma en què encaixaran aquestes dues macromolècules. Es diu doncs que es tracten d'estructures estereoespecífiques. Per a solucionar aquest problema, és a dir, determinar el potencial a la superfície d'aquests macroions, es fa servir l'equació de Poisson-Boltzmann, explicada més endavant a la secció de Mètodes.

## 1.4 Hipòtesis i objectiu final

Finalment doncs, aquest treball rau en la creació d'un àmbit quasi-perfecte per a aquestes cèl·lules i la seva reproducció. És necessari generar un entorn gairebé idèntic al de l'interior d'un gangli limfàtic. Per a aconseguir-ho, s'han d'analitzar les proteïnes que s'encarreguen de vincular la cèl·lula a les *parets* del gangli, per a així determinar-ne la interacció electroestàtica i poder decidir el millor mètode per a produir-lo.

En el cas del present treball, es fa servir un hidrogel creat principalment amb heparina i un químic anomenat polietilenglicol per a simular aquest entorn del gangli limfàtic. Als resultats experimentals duts a terme al laboratori, s'ha descobert que certa proteïna localitzada dins els gànglis limfàtics, la CCL21, amb una càrrega de 17.0eV a pH 7, té la taxa de reproducció de cèl·lules T més elevada en aquest hidrogel. Això condueix a pensar que aquesta proteïna té una gran interacció electroestàtica amb l/heparina, que serà un dels principals punts d'aquest estudi.

La hipòtesi que es planteja és, doncs, comprovar si la interacció entre la proteïna CCL21 i l/heparina és purament elecetroestàtica, o si n'hi ha d'altres rellevants. En cas afirmatiu, caldrà comprovar, a més, quins *binding sites* utilitzen cadascuna i la geometria assolida. Finalment, serà necessari validar o refusar si es tracta d'una interacció suficientment efectiva per a l'estudi en immunoteràpia.

S'investigaran doncs, les principals interaccions entre aquests dos elements, els càlculs electroestàtics que les regeixen i la seva resolució, tant analítica com numèrica.

En resum, els objectius punt per punt seran:

- Determinar la millor estructura digital de la molècula CCL21.
- Estudiar-la.
- Trobar-ne el potencial electroestàtic a la superfície.
- Realitzar els mateixos passos per a una estructura fiable de l/heparina.
- Comprovar la interacció i geometria assolida (el punt i orientació més probables d'unió) d'aquestes dues molècules.
- Verificar si es tracta d'una interacció purament electroestàtica o hi intervenen d'altres.
- Determinar finalment si es prou bona per a l'estudi en immunoteràpia.

## 2 Mètode

### 2.1 Equació de Poisson-Boltzmann

A partir de la definició dels macroions esmentada anteriorment, s'estudia el cas simple del macroió carregat negativament envoltat de càrregues positives a pocs nanòmetres de la superfície. En aquest cas, es pot aproximar un camp elèctric  $E$  amb direcció  $-\hat{x}$ , a causa de la curta distància. Si es consideren distribucions prou uniformes, es pot definir  $-\sigma_q$  com a distribució superficial de càrrega del macroió ( $Cm^{-2}$ ), mentre  $\rho_q$  ( $Cm^{-3}$ ) defineix la distribució volumètrica del nívol de càrregues positives que l'envolta. És a dir, aplicant la llei de Gauss, es té que el camp elèctric a la superfície és:

$$\mathcal{E}_{sup} = \frac{-\sigma_q}{\varepsilon} \quad (1)$$

On  $\varepsilon$  representa la permetibilitat rlativa del medi aquós (també anomenada constant dielèctrica de l'aigua) que se sol establir a 78.4 a temperatura ambient. Pel que fa a les càrregues del nívol, altre cop per la llei de Gauss es pot deduir que:

$$\frac{d\mathcal{E}}{dx} = \frac{\rho_q}{\varepsilon} \quad (2)$$

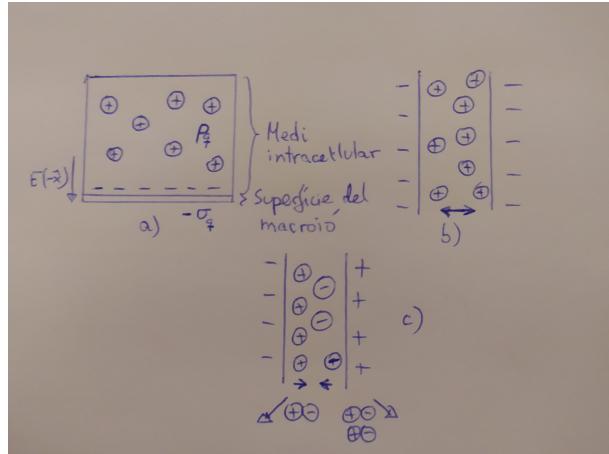


Figure 1: (a) Esquema simplificat del macroió rodejat pel nívol carregat positivament. Podem observar les distribucions de càrrega esmentades anteriorment, recalcant que només hi ha pocs nm de separació entre superfície i càrregues lliures. (b) Dos macroions de símbol igual interactuant: hi ha repulsió entre ells, però els contraions es mantenen pressionats per ambdues bandes. (c) Quan dos macroions de símbol diferent s'acosten, els contraions s'ajunten i surten dispersats, guanyant entropia.

Un cop determinades aquestes condicions, es pot començar a plantejar com determinar els potencials de cada banda. És interessant començar pel cas més simple: el nívol està carregat de ions monovalents positius. Sorgeix però un problema: no es pot determinar el camp elèctric sense saber la distribució de càrregues, i per tant s'han de determinar aquestes. A la vegada, hom pot veure que en estar carregades aquestes partícules, elles mateixes afecten la distribució de les altres. Per tant, hem de determinar dos aspectes a la vegada: la distribució de càrregues que afecta el potencial, i el potencial que afecta la distribució de càrregues.

Recordant que el potencial pot ser expressat en funció de la distribució de càrregues, es pot definir una distribució mitjana de càrregues  $\langle \rho_q \rangle$ , que representa la mitjana d'aquestes partícules pertanyents al nívol, i és possible en tenir innumerables contraions repartits al voltant del macroió en estudi, ja que es tracta d'un element biològic de gran mida. És a dir, cada ió notarà l'efecte de molts altres, i les fluctuacions en el potencial seran més aviat mínimes. El potencial que aquesta crea s'anomena camp mitjà (mean field), i s'escriu com a  $V(x)$ : s'anomena aproximació de camp mig. Per simplicitat és millor seguir escrivint  $\rho_q$ .

S'obté doncs, que necessitem determinar la manera de trobar el potencial elèctric a partir de la densitat de càrrega i viceversa. Començant per la concentració de contraions, i.e.  $c_+(x)$  (cal recordar que seguim en el cas simplificat monovalent), suposant un medi aquós no salí, la concentració lluny de la superfície tendeix a zero. Per altra banda, el potencial electroestàtic de qualsevol contraïó és  $eV(x)$ , ja que s'ha considerat un camp constant a x. Suposant que cada ió es mou de manera independent en un potencial fix  $V(X)$  (com hem mencionat abans), es pot veure que es tracta d'una distribució de Boltzmann, que podem escriure tal que  $c_+(x) = c_0 e^{-eV(x)/k_B T}$ , amb  $c_0 = c_+(0)$  tal que  $V(0) = 0$  (és a dir, la concentració a la superfície), i suposant que  $q = +e$ , ja que es tracta del cas monovalent sense sals al medi.

En aquest senzill cas, podem definir  $\rho_q$  com la concentració  $c_+(x)$  per la càrrega  $q = +e$ , i recordant que  $\mathcal{E} = -dV/dx$  per al camp establert anteriorment, i recordant ??, s'arriba a l'equació de Poisson  $\frac{d^2V}{dx^2} = -\rho_q/\varepsilon$ . Cal remarcar, a més, que  $-\sigma_q$  és representada pels aminoàcids carregats negativament presents a la superfície de la macromolècula, i  $\rho_q$  correspon a la distribució dels contraions més les sals presents en aquest medi, com soLEN ser-ho Na<sup>+</sup> i Cl<sup>-</sup>. A més, la càrrega de tots aquests components dependrà del pH del medi.

Aquí entra en dubte la unió d'aquestes dues equacions. Per una banda, tenim les càrregues fixes a la superfície, que vénen determinades per la relació de Poisson. Per altra però, tenim les lliures que són determinades a partir de la distribució de Boltzmann. És ràpid de veure que, tant una com l'altra, s'afechten mútuament: el canvi de potencial modifica la concentració d'ions, i el canvi de l'última en modifica el primer. Per tant, s'han de determinar a la vegada, fet que acostuma a ser complex ja que no sol existir una solució analítica en ser no-lineal, llevat d'alguns casos molt simples amb geometries molt senzilles. Hom podria pensar de fer una aproximació lineal, però aquesta només seria vàlida per a un potencial inferior a  $k_B T$ , que com bé s'ha mencionat prèviament, no succeeix en l'àmbit d'estudi. En unir les dues relacions, s'obté:

$$\frac{d^2V}{dx^2} = -\frac{ec_0 e^{-eV(x)/k_B T}}{\varepsilon} \quad (3)$$

Aquesta equació pot millorar-se per a entendre'n millor el comportament. Un canvi de terme interessant és el del potencial adimensional reescalat, que permet comparar la relació directa, l'ordre de magnitud, entre l'energia tèrmica de les molècules a temperatura ambient,  $k_B T$  amb el potencial de la superfície en el medi. Dit altres, representa el potencial electroestàtic que ve alliberat a causa de l'agitació tèrmica a certa temperatura, com l'ambient, o en aquest cas, la corporal. Es pot reescriure com:

$$\bar{V} = \frac{eV(x)}{k_B T} \quad (4)$$

Això modifica l'equació anterior tal que:

$$\frac{d^2\bar{V}}{dx^2} = -\frac{e^2 c_0 e^{-\bar{V}}}{\varepsilon k_B T} \quad (5)$$

On hem reescrit  $c_+(x)$  com  $c_0 e^{-\bar{V}}$ .

Aquest raonament fa que Niels Bjerrum (1879-1958) trobi una altra magnitud d'ampli interès d'estudi: la longitud de Bjerrum. Aquesta indica a partir de quina distància la interacció electroestàtica entre dues partícules de càrrega e és equivalent a l'energia tèrmica  $kT$ , o en altres paraules, quant es poden acostar dues molècules de càrrega de mateix signe sense que es repel·leixin, considerant que es disposa d'una energia de valor  $K_B T$ . En aigua a temperatura ambient amb ions monovalents, es dóna a 0.71nm. La fórmula que estableix és:

$$l_B = \frac{e^2}{4\pi\varepsilon k_B T} \quad (6)$$

Que correspon a equiparar la fórmula de l'energia elèctrica,  $E = kq_1 q_2 / r$ , amb l'energia tèrmica,  $E = K_B T$ , amb  $\varepsilon$  la constant dielèctrica de l'aigua i tenint dues càrregues idèntiques d'igual signe. Aillant  $r$ , es troba el valor mencionat. Introduïda a l'equació anterior, acaba proporcionant la coneguda equació de Poisson-Boltzmann sense sals en el medi i amb ions monovalents:

$$\frac{d^2\bar{V}}{dx^2} = -4\pi l_B c_0 e^{-\bar{V}} \quad (7)$$

Utilitzada per a la superfície que interessa analitzar [3]. Aquesta equació, com que acostuma a ser no-linial i per tant irresoluble manualment, es calcula de forma numèrica amb l'aplicació APBS (Automatic Poisson-Boltzmann Solver) al VMD, programa que és utilitzat en aquest treball per a estudiar les estructures en detall. El procés seguit és descrit amb més detall a l'annex.

Com a tota equació diferencial, cal afegir condicions inicials i de contorn per a trobar la solució desitjada. A partir de la llei de Gauss mencionada a ??, aplicant el potencial adimensional trobem que:

$$\frac{d\bar{V}}{dx_{sup}} = 4\pi l_B \frac{\sigma_q}{e} \quad (8)$$

Amb aquesta condició, i establint que  $\bar{V}(0) = 0$  i que:

$$\frac{d\bar{V}}{dx}_{\infty} = 0 \quad (9)$$

Es poden determinar les solucions per a aquest cas simplificat. Per a una solució analítica, necessitaríem una funció que en segona derivada fos equivalent a la seva exponencial. El llibre de P. Nelson [2], pàgina 236, en cita les més conegudes. Aquest concepte serà important per a entendre el següent apartat.

## 2.2 PB per a medis salins

El nombre d'aminoàcids acídics i bàsics determina la càrrega d'una proteïna. Ja siguin a la pròpia proteïna o al medi, aquests interactuen entre ells deixant carregada positiva o negativament la molècula estudiada. Això conduceix a pensar en el comportament les càrregues móbils de la superfície mencionada anteriorment i els ions presents a la dissolució: hi ha càrregues positives i negatives, sobretot en medis aquosos com és el cos humà. Per tant, és necessari reescriure l'equació de Poisson-Boltzmann per a diferents tipus de càrregues, en medis salins.

En els casos realistes, on normalment no tenim solució analítica, s'han de considerar tots els factors: el macroió serà segurament heterogeni, carregat de forma positiva i negativa a diferents zones de la molècula; els ions presentaran una situació similar, i, a més, seran presents en un medi salí, que provoca que la concentració a l'infinít  $c_{\infty}$  de les sals mai sigui exactament zero. A causa d'això, serà millor canviar la condició inicial del potencial: se n'ha d'establir un tal que  $V(x)$  tendeixi a zero a grans distàncies. Això duu a:  $c_+(x) = c_{\infty} e^{-eV(x)/k_B T}$  i  $c_-(x) = c_{\infty} e^{-(e)V(x)/k_B T}$ , suposant que ens trobem en un medi salí on tenim components de càrregues + i -.

Unit aquestes dues concentracions a l'equació de Poisson-Boltzmann ja esmentada, s'obté:

$$\frac{d^2\bar{V}}{dx^2} = -\frac{1}{2} 8\pi l_B c_{\infty} \left( e^{-\bar{V}} - e^{\bar{V}} \right) \quad (10)$$

On escribim  $-1/2 \cdot 8$  per a la següent simplificació: Peter Debye (1884-1966), de manera similar a Bjerrum, estipula una magnitud interessant en aquesta expressió: la longitud de Debye, que en aquests casos es representa com  $\lambda_D = (8\pi l_B c_{\infty})^{-1/2}$ . Aquesta magnitud ens indica fins a on persisteixen els efectes electroestàtics d'una solució. En altres paraules, a distàncies més grans que la longitud de Debye, els camps elèctrics comencen a patir un efecte d'apantallament que decau de forma exponencial. Per a l'aigua a temperatura ambient, és aproximadament 0.7nm, suposant 100mM per a cada ió.

Aquesta longitud pot ser implementada dins l'equació de Poisson-Boltzmann com:

$$\frac{d^2\bar{V}}{dx^2} = -\frac{1}{2} \lambda_D^{-2} \left( e^{-\bar{V}} - e^{\bar{V}} \right) \quad (11)$$

Finalment cal remarcar doncs, que s'utilitza el programari APBS [12] per a la resolució numèrica ja que l'analítica en aquest cas és impossible en ser geomètricament molt complexa.

No caldria ni mencionar l'ampli rang d'usos que proveeix aquesta equació en tots els diferents camps de la física: electroestàtica, sistemes biològics, sistemes electroquímics, sistemes biofísics, ciència de materials, entre molts altres. Això és degut a la gran quantitat de situacions realistes on intervé la presència de diferents càrregues en un medi determinat, com és el medi aquós salí del cos humà en aquest cas.

En altres paraules: en aquest treball es farà ús de l'equació de Poisson-Boltzmann per a trobar el potencial a la superfície de la proteïna CCL21 i a l/heparina, mitjançant el programa APBS, que proporciona la solució

numèrica de l'equació amb unes condicions prèviament estipulades, com a qualsevol equació diferencial, mencionades a l'apèndix. Per a preparar les coordenades d'aquesta proteïna, a més, farem ús del programari VMD, present a l'apèndix.

### 3 Resultats

#### 3.1 Sel·lecció del model proteic

La proteïna desitjada per a l'estudi és la CCL21. Al *Protein Databank*, lloc on s'emmagatzem les estructures moleculars de forma digital, s'hi troben dos models: 2L4N i 5EKI. Aquestes estructures de proteïnes han estat obtingudes de forma experimental, detallades als annexos. En comparar les estructures, es determina que l'estructura més adient per a aquest estudi és la 5EKI, més concretament la cadena E (ja que el model al PDB presenta sis cadenes similars de la mateixa proteïna). Aquesta sel·lecció és a causa de la important presència d'uns aminoàcids a la terminal C (a partir del 70) de l'estructura 5EKI per a la interacció amb altres proteïnes com la CCR7, que no són presents a la 2L4N. El protocol seguit a la comparació s'explica en detall als annexos.

Cal remarcar però, que cap de les dues estructures es considera complerta. La cadena E de la 5EKI manca els 5 primers aminoàcids, però no s'han considerat crucials per a aquest estudi. L'aparença d'aquesta molècula és la següent:

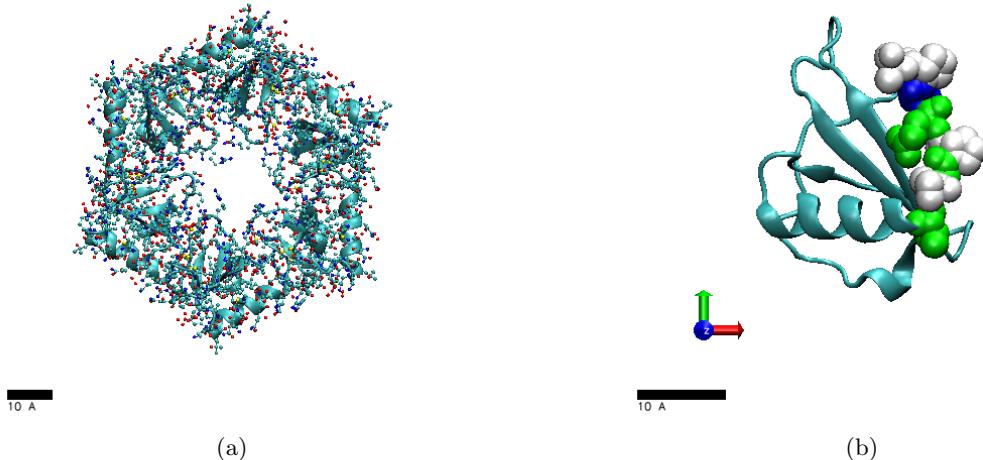


Figure 2: (a) Estructura de codi 5EKI, que presenta sis cadenes similars a causa del mètode d'obtenció de cristal·lografia de rajos X. Presenta 3639 àtoms, sense hidrògens. (b) Cadena E de l'estructura acabada de mencionar, que es considera el més bo (veure Annex) amb la terminal C d'aquest model, lloc crucial per a la interacció amb la proteïna CCR7. Presenta 583 àtoms, sense hidrògens. Els gràfics utilitzats per a l'estructura base són NewCartoon. A la primera fotografia utilitzem CPK per a veure els residus més clarament, i a la segona indiquem la terminal amb VDW colorats segons el tipus de residu.

Es calcula, a més, la càrrega total de la cadena E, que proporciona un resultat de 7e-. Per a la interpretació dels arxius pdb, es poden consultar els annexos.

#### 3.2 Càlcul electrostàtic superficial de la proteïna

Mitjançant l'APBS es resolen les equacions de Poisson-Boltzmann mencionades a la secció anterior. És interessant recalcar les condicions prèvies establertes. Establint unes concentracions de 100mM NaCl. Col·locant una temperatura de 298.15K (T ambient) i comparant-la amb 309.65 (T mitjana del cos humà), una permeabilitat del medi de 78.54 i un pH de 7.0, es troben les següents figures:

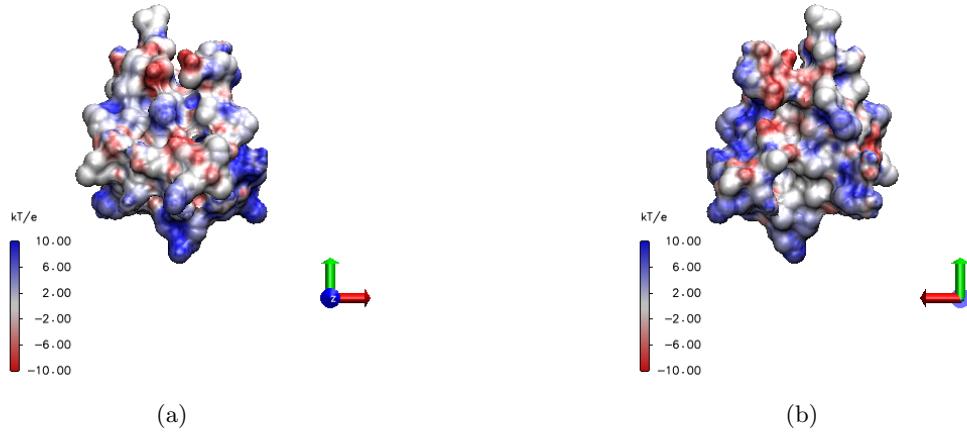


Figure 3: (a) Representació del potencial electroestàtic a la superfície de la proteïna CCL21 (5EKI) a una temperatura corporal de 309.65K. (b) Mateixa figura, rotada  $180^{\circ}$  a l'eix y. Els càlculs han estat realitzats amb APBS (consultar Annex) i les representacions amb VMD (Surf + Volume).

Cal recalcar que s'ha obtingut aquesta imatge fent servir un rang de potencials de -10 a +10 unitats de  $kT/e$  que representa el potencial adimensional. Recordant l'equació ?? es pot veure que  $V(x)$  s'expressa en unitats de  $kT/e$  multiplicats per aquest potencial adimensional. A més, a temperatura corporal,  $kT/e$  té un valor aproximat de 26.68mV, pel que es té una rang de 266mV a -266mV segons els valors trobats amb més freqüència a l'arxiu dx. Com que les figures obtingudes amb la temperatura ambient són molt similars a les del cos, s'han mostrat a l'annex.

Es pot veure que la zona blava, sobretot a la imatge a  $180^{\circ}$ , correspon aproximadament a la terminal que interessa estudiar, vista abans. Es pot deduir que es tracta d'una estructura molt inhomogènia, que de fet en part pot semblar útil a l'hora d'addherir-se a l/heparina, de càrrega negativa, ja que les zones positives que veiem al gràfic són relativament lluny de les negatives, que afavoreix la unió sense repel·lir-se.

### 3.3 Heparina

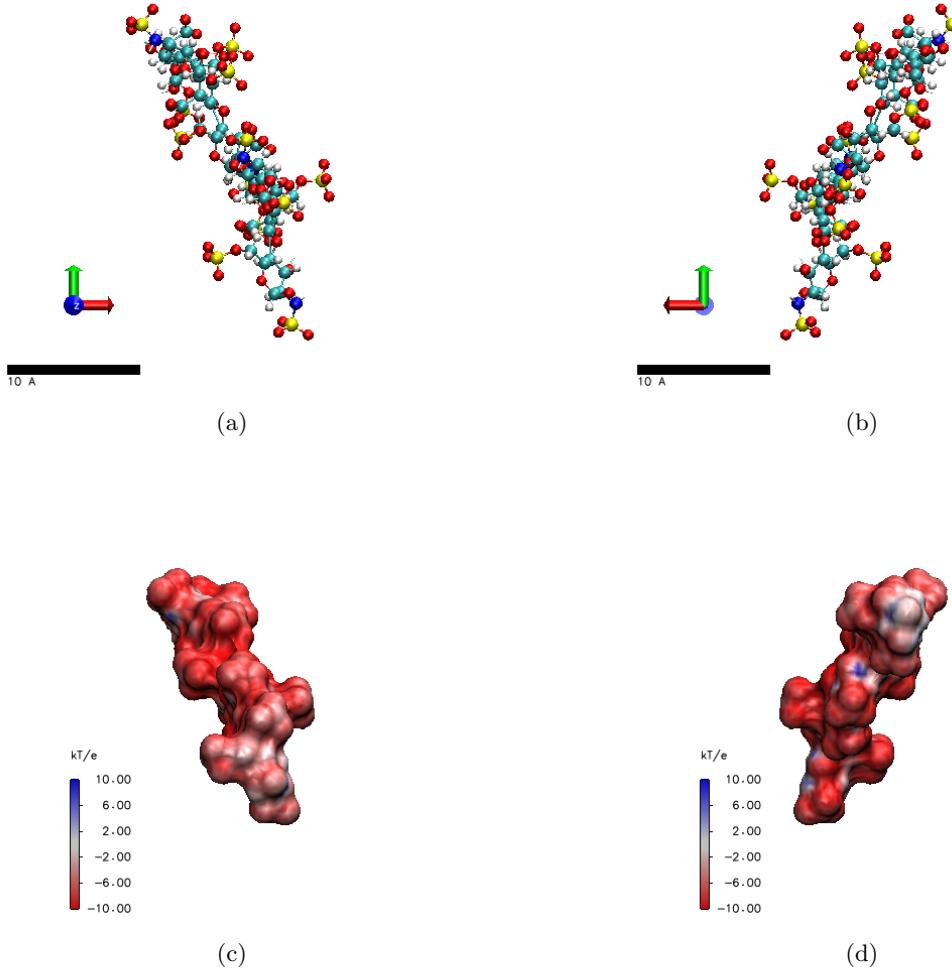


Figure 4: (a) Estructura general de l/heparina [14], amb 203 àtoms, hidrògens inclosos. (b) Estructura girada a  $180^\circ$  sobre l'eix y. (Tant (a) com (b) estan representades en format CPK, amb el programa VMD) (c) Potencial a la superfície de l/heparina, a temperatura corporal (veure resta de condicions a l'apèndix). (d) Potencial superficial de l/heparina, a  $180^\circ$  en l'eix y. El color blau representa la tendència cap a càrregues positives i, el vermell, negatives.(Tant (c) com (d) estan representades en format Surf-Volume, al VMD).

A les dues primeres imatges s'observa l'estruatura general de l/heparina amb escala [14]. Al segon parell, es pot observar el comportament del potencial a la superfície mitjançant l/APBS. Cal fixar-se, com destaca, la distribució gairebé homogènia de càrregues a la superfície, on gairebé no hi és present cap càrrega positiva ni neutra, llevat d'alguns punts molt discrets. Comparant ?? amb ??, podem veure que la zona electropositiva més destacada és la cantonada inferior dreta, on hi tenim un intens color blau. A més, correspon justament amb una part dels residus 70-77 de la cadena E, que segons els documents respecte al model 5EKI (mencionats a l'apèndix), són primordials en les interaccions d'aquesta proteïna. Per altra banda, l/heparina, tot i tenir càrrega negativa a tota la superfície, sembla ser més intensa a la part superior. Això condueix a pensar que, en una primera aproximació, les dues estructures s'unirien per aquella zona.

El que pertoca és doncs calcular, de forma analítica i aproximada, la interacció entre aquestes dues molècules. Segurament apareixerà un resultat força diferent al numèric, però almenys s'ha d'intentar que

quadri amb l'ordre de magnitud.

### **3.4 Encaix de molècules (molecular docking)**

## **4 Conclusions i comentari final**

## A Apèndix

### A.1 Bibliografia

- [1] <https://www.youtube.com/watch?v=eckgmqiBBRo> “Avancem en la immunoteràpia contra el càncer amb materials 3D amb Ignacio Crespo” per ICMAB-CSIC
- [2] Nelson, P., “Biological Physics: Energy, Information, Life,” *Entropic forces at work*, pp. 230-231, First Edition (2003).
- [3] Nelson, P., “Biological Physics: Energy, Information, Life,” *Entropic forces at work*, pp. 232-235, First Edition (2003).
- [4] <https://www.ks.uiuc.edu/Research/vmd/> Pàgina principal del programari VMD.
- [5] <https://www.rcsb.org/structure/2l4n> Primer model considerat, 2L4N.
- [6] [https://files.rcsb.org/pub/pdb/validation\\_reports/l4/2l4n/2l4n\\_full\\_validation.pdf](https://files.rcsb.org/pub/pdb/validation_reports/l4/2l4n/2l4n_full_validation.pdf) Informe sobre l’obtenció del model mencionat i els procediments aplicats.
- [7] <https://pubs.acs.org/doi/10.1021/bi201601k> Document informatiu sobre el model 2L4N de la proteïna CCL21.
- [8] <https://www.ks.uiuc.edu/Training/Tutorials/vmd/tutorial-html/> Tutorial utilitzat en aquest treball.
- [9] <https://www.rcsb.org/structure/5EKI> Segon model utilitzat.
- [10] <https://pubs.acs.org/doi/10.1021/acs.biochem.6b00304> Document informatiu del model 5EKI.
- [11] [https://files.rcsb.org/pub/pdb/validation\\_reports/ek/5eki/5eki\\_full\\_validation.pdf](https://files.rcsb.org/pub/pdb/validation_reports/ek/5eki/5eki_full_validation.pdf) Informe d’obtenció del model 5EKI. pp. 4, 9-10.
- [12] <https://server.poissonboltzmann.org/pdb2pqr> Convertidor en línia d’arxius PDB/PSF a PQR i APBS.
- [13] <https://github.com/MaxTFG/TFG> Arxius treballats per a consultar.
- [14] <https://amarolab.ucsd.edu/covid19.php> Arxius generats per a l’estudi de l’heparina.

## A.2 Annexos

### A.2.1 Estudi i selecció del millor model virtual

En fer ús del programa VMD, haurem d'utilitzar models virtuals de proteïnes principalment del *Protein Databank*, on hi ha emmagatzemades les que ens interessa estudiar. El primer model sel·leccionat per a estudiar és l'anomenat 2L4N, de la proteïna CCL21 com s'ha esmentat abans [4] [5]. Per a saber fer anar el programa, la pàgina conté alguns tutorials essencials. Els utilitzats en aquest treball han estat els mencionats a la bibliografia [8].

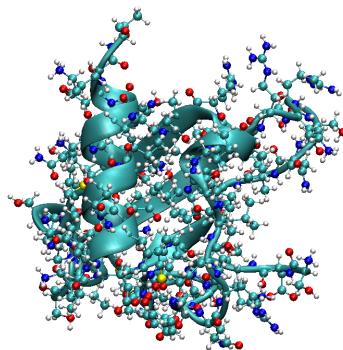


Figure 5: Estructura general del model 2L4N

La determinació de l'estructura d'aquesta proteïna s'ha realitzat mitjançant el mètode NMR (Nuclear Magnetic Resonance), que consisteix en col·locar la proteïna en un camp magnètic potent i constant mentre els seus àtoms són afectats per un altre camp magnètic feble i oscil·lant. En llegir la freqüència obtinguda d'aquest procés, se'n poden determinar els elements i molècules (aminoàcids) d'aquesta proteïna de manera força eficient [6].

En analitzar aquesta proteïna, descobrim que el model virtual només inclou els 70 primers aminoàcids, mentre menciona que en falten uns 40 a la terminal C.

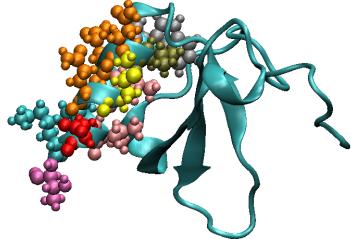


Figure 6: Terminal C del model 24LN (59 a final)

És interessant revisar el document adjunt ja que ens menciona els llocs d'unió (*binding sites*) de la proteïna CCL21 amb la CCR7, que és la proteïna localitzada a la capa més exterior de les cèl·lules T que facilita la interacció electrostàtica i és d'estudi primordial en aquest treball, ja que d'elles depèn el possible cultiu en 3D al laboratori [7]. En són els següents:

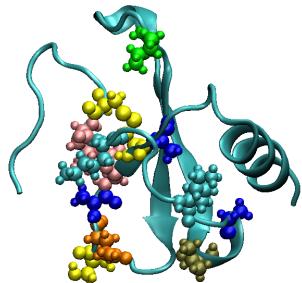


Figure 7: Residus 9-11, 16, 19, 21, 36, 47-50, 52 i 53 del model 2L4N

En descobrir la manca d'aminoàcids a la terminal C, decidim cercar si existeix un model virtual amb més aminoàcids determinats a la terminal C ja que són crucials a l'hora de la interacció amb l/heparina, perquè s'ha demostrat que el motiu pel qual la proteïna CCL21 atrau tan eficientment les cèl·lules T és gràcies a aquesta terminal que altres similars menys eficients no contenen. Trobem un model força interessant, de la mateixa proteïna CCL21, anomenat 5EKI, que s'ha determinat mitjançant cristal·lografia r�os X [9].

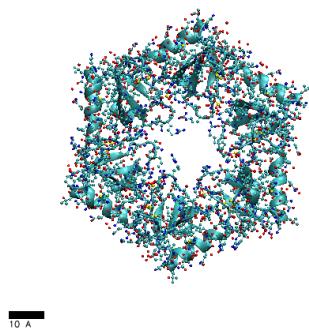


Figure 8: Estructura general del model 5EKI

Trobem que el model conté 6 formes de la mateixa proteïna (a causa de la cristal·lografia), pel que decidim buscar quina és la més adient. Finalment veiem que la cadena E no conté cap col·lisió entre àtoms [10] [11] (el model no té àtoms massa a prop entre ells que puguin causar problemes a l'hora de calcular interaccions) i a més és el que menys residus mancants proporciona. Decidim doncs utilitzar aquesta cadena per a l'estudi.

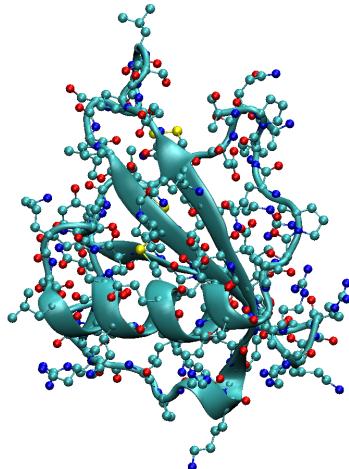


Figure 9: Cadena E del model 5EKI

Fent un cop d'ull al document adjunt d'aquest model veiem que menciona que els aminoàcids de la terminal C a partir del 70 fins el 79, són de gran importància en la interacció d'aquesta proteïna, els quals el model 2L4N no conté. Finalment, decidim doncs que el model 5EKI és el que ens interessa utilitzar per a l'estudi, tot i no contenir els 5 primers aminoàcids de la terminal N, que de moment, considerem no-crucials per als càculs que ens interessen realitzar.

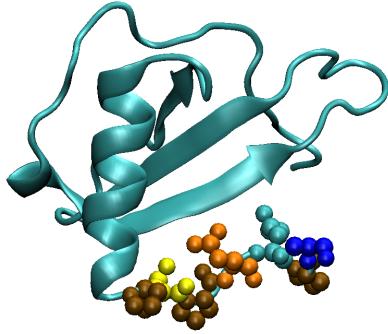


Figure 10: Terminal C (70-79) de la cadena E (5EKI)

Per a acabar de comparar i descartar el millor model, fem servir l'eina *MultiSeq* al VMD que ens compara l'estructura dels dos models de la proteïna. Veiem que l'única discrepància es troba als primers aminoàcids com mencionat anteriorment.

### A.2.2 Obtenció dels arxius per l'APBS

Per al correcte estudi i càlculs amb el model escollit, hem de perfeccionar la nostra proteïna, és a dir, afegir-hi els elements que puguin faltar-hi. En general solen faltar només àtoms d'hidrogen, però decidim fer-ho amb el VMD mitjançant uns arxius de topologia que l'ajuda a interpretar cada aminoàcid de la proteïna i determinar quins elements falten. Podem utilitzar els convertidors en línia mencionats a la bibliografia [12], que afegeixen pel seu compte els hidrògens i altres àtoms que falten a la proteïna.

A més, un cop obtinguda aquesta nova estructura, utilitzarem el programa *PDB2PQR* que ens permet proporcionar una càrrega i un radi a cada element de la proteïna amb el pH desitjat. Amb tot això podem procedir a utilitzar el programa APBS, que ens proporcionarà una imatge tridimensional del potencial a la superfície d'aquesta proteïna.

A part, convindria calcular la càrrega total de la proteïna, que ho podem fer sumant la segona columna de l'arxiu PQR amb Excel o similar. Necessitarem aquest resultat per a establir quins límits, en termes de  $kT/e$ , desitgem al gràfic del potencial de superfície.

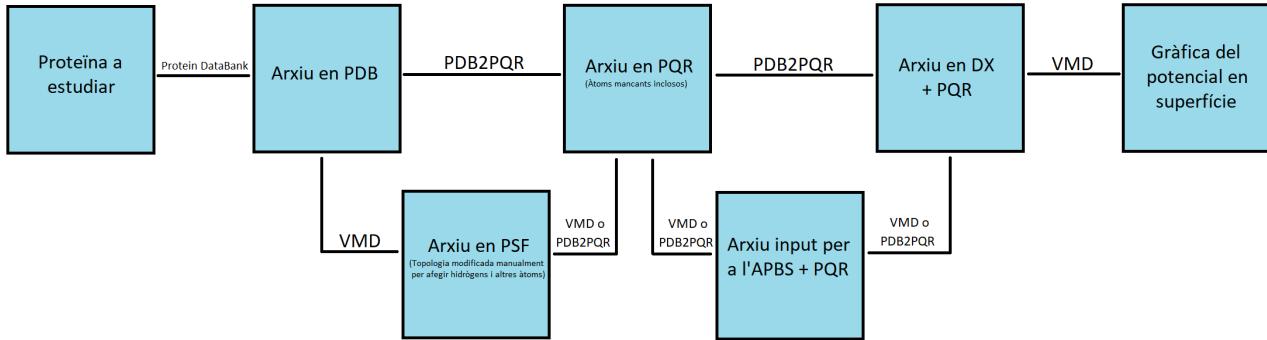


Figure 11: Procés seguit a la primera part.

La disposició dels arxius PQR és la següent: la primera columna especifica el tipus d'element (ATOM / HETATM), el segon l'ordre, el tercer la posició dins l'aminoàcid, i el quart l'aminoàcid al qual pertany.

La cinquena columna menciona el chainID, a quina cadena pertany l'àtom. Les tres següents especificuen les posicions X Y Z, i la següent és la càrrega, que en sumar tota la columna obtenim la càrrega total de la proteïna. Finalment l'última columna és el radi de cada àtom. Per a consultar els arxius utilitzats, es pot veure el GitHub a [13].

Els resultats per a la temperatura ambient del potencial a la superfície són:

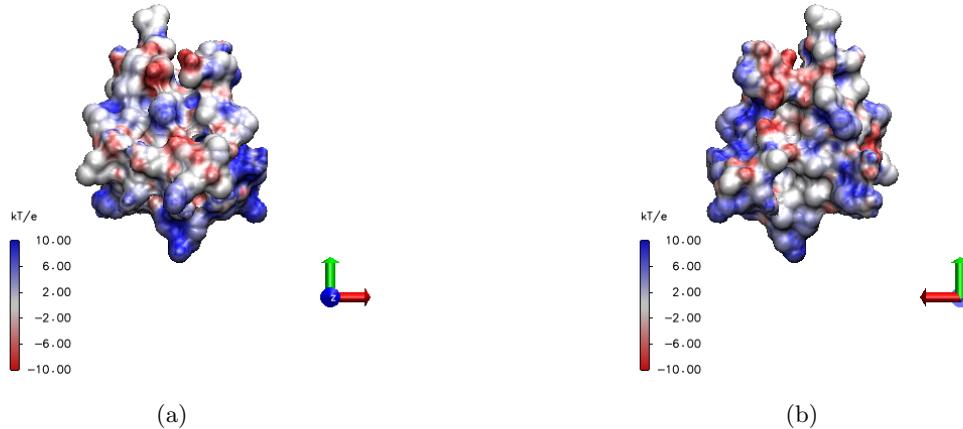


Figure 12: (a) Representació del potencial electroestàtic a la superfície de la proteïna CCL21 (5EKI) a una temperatura corporal de 298.65K. (b) Mateixa figura, rotada 180º a l'eix y. Els càlculs han estat realitzats amb APBS (consultar Annex) i les representacions amb VMD (Surf + Volume).

Pel que fa a l/heparina, repetim el mateix procés esmentat abans amb els arxius trobats a [14] dins la carpeta ”*spike\_apbs\_marolab.tar.gz*”, amb l/arxiu ”*complex.pqr*” trobat a ”*pqr*s” dins *rbshep* i ”*complex.in*” trobat a ”*ins*”. Tot això al convertidor en línia ens proporciona les imatges mencionades en l/apartat de l/heparina.

### A.3 Índex alfabètic

## A.4 Declaracions

### A.4.1 Declaració d'autoria del Treball de Fi de Grau

Jo, Maximilià Soler Subirana, amb Document Nacional d'Identitat 21770721W, i estudiant del Grau en Física de la Universitat Autònoma de Barcelona, en relació amb la memòria del treball de final de Grau presentada per a la seva defensa i avaluació durant la convocatòria de Juliol del curs 2022-2023, declaro que

- El document presentat es original i ha estat realitzat per la meva persona.
- El treball s'ha dut a terme principalment amb l'objectiu d'avaluar l'assignatura de treball de grau en física en la UAB, i no s'ha presentat prèviament per a ser qualificat en l'avaluació de cap altra assignatura d'aquesta universitat ni de cap altra.
- En el cas de continguts de treballs publicats per terceres persones, l'autoria està clarament atribuïda, citant les fonts degudament.
- En els casos en els que el meu treball s'hagi realitzat en col·laboració amb altres investigadors i/o estudiants, es declara amb exactitud quines contribucions es deriven del treball de tercers i quines es deriven de la meva contribució.
- A l'excepció dels punts esmentats anteriorment, el treball presentat és de la meva autoria.

Signat:

#### **A.4.2 Declaració d'Extensió del Treball de Grau**

Jo, Maximilià Soler Subirana, amb Document Nacional d'Identitat 21770721W, i estudiant del Grau en Física de la Universitat Autònoma de Barcelona, en relació amb la memòria del treball de final de Grau presentada per a la seva defensa i avaluació durant la convocatòria de Juliol del curs 2022-2023, declaro que:

- El nombre total de paraules (segons el comptatge proposat) incloses en les seccions des de la introducció fins a les conclusions és de X paraules.
- El nombre total de figures es de X.
- En total el document, comptabilitza: paraules + x 200 paraules/figura = Que compleix amb la normativa en ser inferior a 10000.

Signat:

#### **A.4.3 Declaració Supervisió Treball de Grau**

Aquest document té com a objectiu certificar que l'alumne Maximilià Soler Subirana, amb DNI 21770721W, ha realitzat el seu treball titulat X sota la meva direcció, i el presentarà per tal de defensar l'assignatura al Treball final de grau en el Grau de física. L'alumne m'ha fet arribar el formulari word d'avaluació de l'activitat de l'alumne per part del director. (Si us plau, feu-lo arribar omplert a tfg.física@uab.cat abans del lliurament de les memòries)

## List of Figures

1	(a) Esquema simplificat del macroió rodejat pel núvol carregat positivament. Podem observar les distribucions de càrrega esmentades anteriorment, recalcant que només hi ha pocs nm de separació entre superfície i càrregues lliures. (b) Dos macroions de símbol igual interactuant: hi ha repulsió entre ells, però els contraions es mantenen pressionats per ambdues bandes. (c) Quan dos macroions de símbol diferent s'acosten, els contraions s'ajunten i surten dispersats, guanyant entropia. . . . .	5
2	(a) Estructura de codi 5EKI, que presenta sis cadenes similars a causa del mètode d'obtenció de cristal·lografia de rajos X. Presenta 3639 àtoms, sense hidrògens. (b) Cadena E de l'estructura acabada de mencionar, que es considera el més bo (veure Annex) amb la terminal C d'aquest model, lloc crucial per a la interacció amb la proteïna CCR7. Presenta 583 àtoms, sense hidrògens. Els gràfics utilitzats per a l'estructura base són NewCartoon. A la primera fotografia utilitzem CPK per a veure els residus més clarament, i a la segona indiquem la terminal amb VDW colorats segons el tipus de residu. . . . .	9
3	(a) Representació del potencial electroestàtic a la superfície de la proteïna CCL21 (5EKI) a una temperatura corporal de 309.65K. (b) Mateixa figura, rotada 180º a l'eix y. Els càlculs han estat realitzats amb APBS (consultar Annex) i les representacions amb VMD (Surf + Volume). . . . .	10
4	(a) Estructura general de l/heparina [14], amb 203 àtoms, hidrògens inclosos. (b) Estructura girada a 180º sobre l'eix y. (Tant (a) com (v) estan representades en format CPK, amb el programa VMD) (c) Potencial a la superfície de l/heparina, a temperatura corporal (veure resta de condicions a l'apèndix). (d) Potencial superficial de l/heparina, a 180º en l'eix y. El color blau representa la tendència cap a càrregues positives i, el vermell, negatives.(Tant (c) com (d) estan representades en format Surf-Volume, al VMD). . . . .	11
5	Estructura general del model 2L4N . . . . .	15
6	Terminal C del model 24LN (59 a final) . . . . .	16
7	Residus 9-11, 16, 19, 21, 36, 47-50, 52 i 53 del model 2L4N . . . . .	16
8	Estructura general del model 5EKI . . . . .	17
9	Cadena E del model 5EKI . . . . .	17
10	Terminal C (70-79) de la cadena E (5EKI) . . . . .	18
11	Procés seguit a la primera part. . . . .	18
12	(a) Representació del potencial electroestàtic a la superfície de la proteïna CCL21 (5EKI) a una temperatura corporal de 298.65K. (b) Mateixa figura, rotada 180º a l'eix y. Els càlculs han estat realitzats amb APBS (consultar Annex) i les representacions amb VMD (Surf + Volume). . . . .	19

## **List of Tables**