

**Molekulaszerkezeti és -dinamikai adatok rezgési spektrumokkal való
kapcsolatának analízise fehérjék esetén.**

TDK dolgozat

**Oláh Tamás István
Fizika Bsc. III évfolyam**

ELTE, Természettudományi kar

Témavezető:

**Dr. Smeller László
Semmelweis Egyetem, Biofizika és Sugárbiológia Intézet**

Budapest 2012

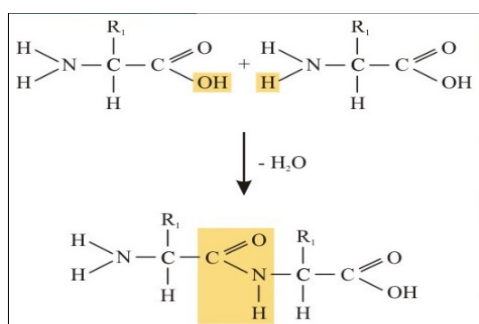
Tartalomjegyzék:

1 Bevezetés.....	2
1.1 A fehérjék.....	2
1.1.1 A fehérjék szerkezete.....	2
1.1.2 A másodlagos szerkezet.....	3
1.2 Hidrogénkötés.....	5
1.3 A fehérje térszerkezetének meghatározására szolgáló módszerek.....	6
1.3.1 Röntgendiffrakció.....	6
1.3.2 Mágneses magrezonancia.....	6
1.3.3 Infravörös spektroszkópia.....	8
1.3.4 Molekuladinamikai szimulációk.....	8
1.4 Molekularezgések.....	9
1.4.1 N-metilacetamid rezgései.....	9
2. Alkalmazott módszerek.....	10
2.1 Torziós szögek kiszámítása.....	10
2.2 Illesztés.....	11
3 Célkitűzések.....	11
4 Eredmények	12
4.1 C=O és O-N távolságok különböző fehérjék esetén.....	12
4.2 Infravörös spektrumok illesztése.....	20
4.3 Különböző fehérje szerkezetre jellemző amid I rezgések meghatározása.....	27
5 Összefoglalás.....	29
6 Köszönetnyilvánítás.....	29
7 Hivatkozások.....	30

1.Bevezetés

1.1 Fehérjék

A fehérjék az élő szervezeteket felépítő biomolekulák egyik fontos csoportját alkotják. Gerardus Johannes Mulder fedezte föl őket, majd 1838 Jöns Jakob Berzelius adta nekik a protein („elsődleges fontosságú”) nevet. A fehérjék húsz különböző α aminosav lineáris polimerjeiből felépülő szerves makromolekulák. A polipeptidláncot alkotó aminosavak peptidkötéssel kapcsolódnak egymáshoz. A peptidkötés két szomszédos aminosav között jön létre, az egyik aminosav karboxilcsoportja kapcsolódik a másik aminosav nitrogénjéhez víz molekula kilépésével, ahogy az 1.1 ábrán látható. A szénatom és nitrogén atom között kialakult kötés delokalizált, planáris és dipólusmomentuma van. A kötés körül rotáció nem lehetséges. A fehérje gerincét alkotó, de a peptidkötésben részt nem vevő atomok (C_α -C, N- C_α) között egyszeres kovalens kötések vannak, melyek körül szabad forgás jöhet létre, ami aminosavanként két rotációs lehetőséget rejt.



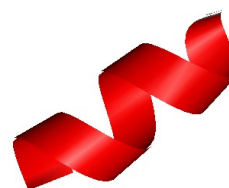
1.1. ábra, peptidkötés kialakulása

(forrás: tudásbázis.sulinet.hu)

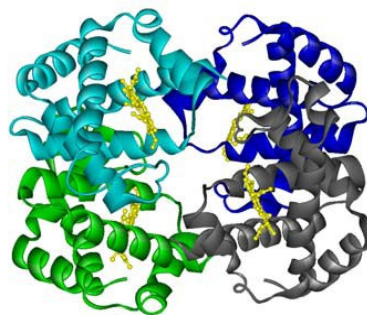
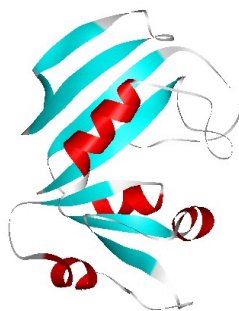
1.1.1 Fehérjék szerkezete

A fehérjék szerkezetének négy szintjét különböztetik meg (1.2 ábra):

- Elsődleges szerkezet az aminosavak kapcsolódási sorrendje (szekvenciája).
- Másodlagos szerkezet a fehérjék rövidtávú rendezettségét írja le, leggyakrabban helikális, vagy lemezszerű. Tipikusan 10-20 aminosavat tartalmaz.
- Harmadlagos szerkezetnek a másodlagos szerkezeti elemek által térben kialakított alakzatot nevezzük. Ezt a szerkezetet a Van der Waals kölcsönhatások, a hidrogénhidkötések, töltött aminosavak között ionos kötések stabilizálják.
- Negyedleges szerkezete a több különálló polipeptidláncból felépülő fehérjéknek van, a polipeptidláncok egymáshoz képesti elrendeződését nevezzük negyedleges szerkezetnek.

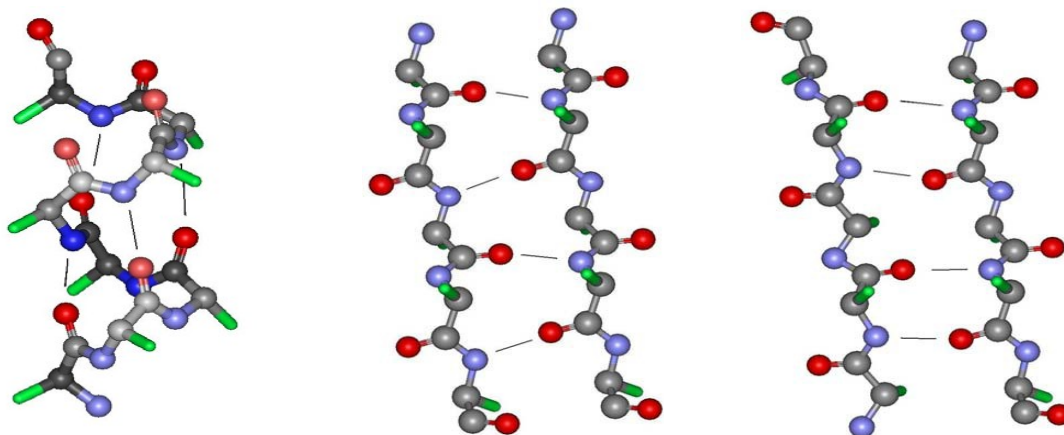


b)



d)

a)példa az elsődleges szerkezetre, b) másodlagos szerkezet egy tipikus képviselőjét az alfa hélixet mutatja be, c)DHFR enzim teljes harmadlagos szerkezete látható d) négy polipeptidláncból felépülő hemoglobin negyedleges szerkezete
Az ábrák a viewerlite és az open office programmal készültek.



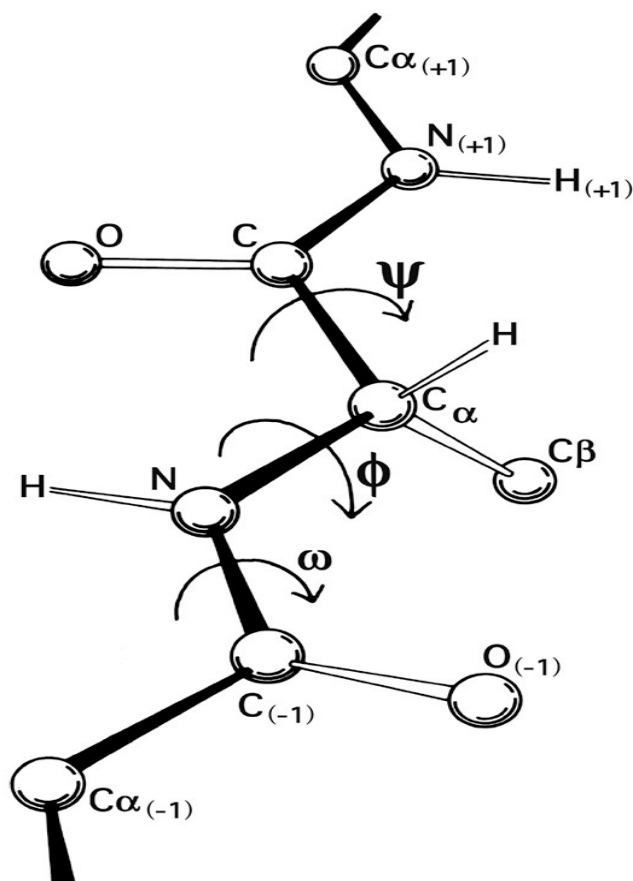
a) α hélix

b) *parallel* β -lemez

c) *antiparallel* β -lemez

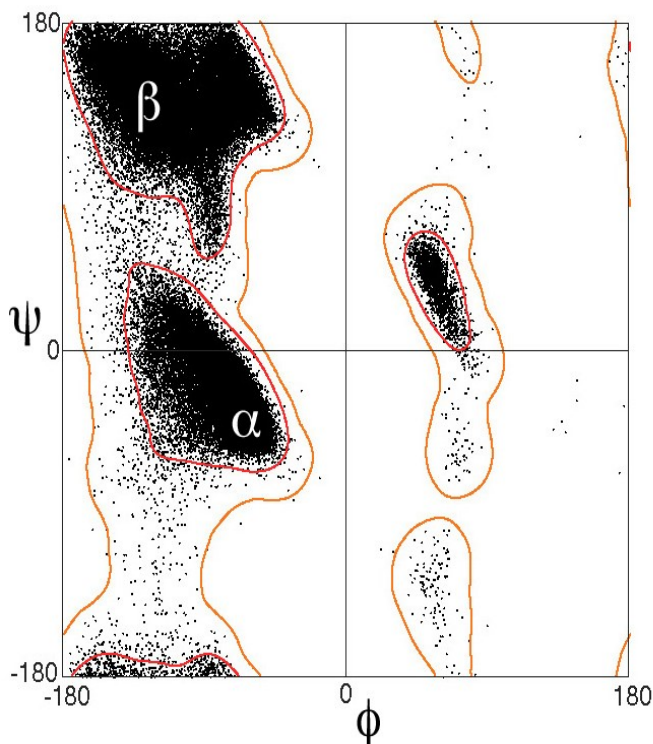
1.3 ábra másodlagos szerkezet fő alkotói. Az ábrán a szürke golyók a szénatomot, a kék golyók a nitrogénatomot, a piros golyók az oxigént, a zöld rudak az oldalláncok csatlakozási pontjait jelölik. A fekete vonalak a hidrogénkötéseket mutatják.

Ezt a szerkezeti szintet a fehérjegerinc mentén definiált torziós szögekkel jellemezhetjük. Három ilyen szöget vezettek be: Φ ($C-N-C_\alpha-C$), ψ ($N-C_\alpha-C-N$) és az ω ($C_\alpha-C-N-C_\alpha$).



1.4 ábra A peptidgerinc torziós szögei
forrás: http://en.wikipedia.org/wiki/Ramachandran_plot)

Az ω szög a gyakoribb cisz izomerek esetén 180° , a ritkább transz izomerek esetén 0° , ez a szög a peptidkötés körül gátolt rotáció miatt mindegyik szerkezeti elemre egyforma. A másik két szög, Φ és ψ szögek esetén mutatkozó eltérések alapján a másodlagos szerkezeti elemek megkülönböztethetők, Φ -t ψ függvényében ábrázolva kapjuk a Ramachadran diagramot, ahol az egyes szerkezeti elemeknek jól elkülönülő tartományok felelnek meg, amelyek az 1.5 ábrán is láthatóak.



1.5 ábra : Ramachadran diagram

(forrás: http://en.wikipedia.org/wiki/Ramachandran_plot)

1.2 A hidrogénkötés

Amint az az 1.3 ábrán látható a másodlagos szerkezetet hidrogénhidak stabilizálják.

A hidrogénkötések a másodlagos kötések közé sorolhatóak. A kötés kialakításához szükség van egy nagy elektronegativitású kisméretű atomra melyhez közvetlenül csatlakozik egy hidrogén és egy másik kisméretű nemkötő elektronpárral rendelkező atom. A fehérjékben a másodlagos szerkezetét ilyen kötések stabilizálják. A kötés az egyik aminosav C=O csoportja és egy másik aminosav N-H csoportja között jön létre. A hidrogénkötések jól leírhatóak elektrosztatikus modellel. Az elektrosztatikus kölcsönhatási energia kiszámítható [8]-ban foglaltak alapján.

$$E = f q_1 q_2 \left[\frac{1}{r(ON)} + \frac{1}{r(CH)} - \frac{1}{r(OH)} - \frac{1}{r(CN)} \right] \quad (1)$$

$$f = 1,41 \cdot 10^{-4} \frac{Jm}{C^2} \quad q_1 = 0,42e \quad q_2 = 0,2e, \text{ e az elektron töltése,}$$

$r(AB)$ az A és a B atom közötti távolság m-ben. Az energiát J/mol egységekben kapjuk meg. Ahol ez az energia kisebb, mint -2542J/mol, (-0,5kcal/mol) ott az i-ik és j-ik aminosav között hidrogénkötés található. Az (1)-es képlet alapján határoztam meg a hidrogénkötések pontos helyét (mely aminosavak között jön létre pl 4.3-ban DHFR enzim esetén). Egyes vizualizációs programok is ezzel a képlettel határozzák meg a kötések helyét.

1.3 A fehérje térszerkezetének vizsgálatára szolgáló módszerek

Három, a gyakorlatban használt módszert ismertetek, a röntgendiffrakciót, a mágneses magrezonanciát és az infravörös spektroszkópiát.

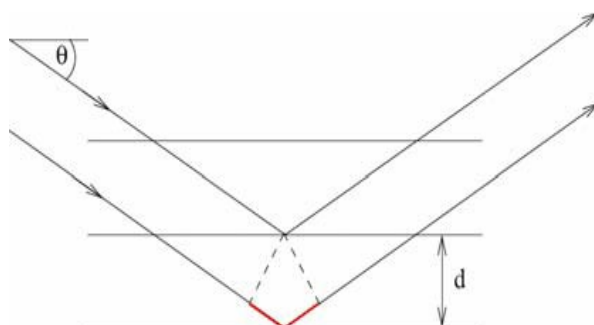
A fehérjék térszerkezetére vonatkozó adatok elérhetőek interneten keresztül adatbázisokból. Ezek közül a fehérjék térszerkezeti adatait a Protein Data Bank: (www.rcsb.org/pdb/) adatbázis tartalmazza. Ez egy nemzetközi adatbázis, amely több mint 80000 makromolekula adatait tartalmazza.

1.3.1 Röntgendiffrakció

A diffrakciós módszerek lényege, hogy a vizsgálandó anyagra annak szerkezeti periódusosságát jellemző állandónál kisebb hullámhosszúságú sugárzást bocsátunk, majd a mintán való áthaladás után detektáljuk az átjutó nyaláb Fraunhofer-féle diffrakciós képét. A rácsállandók általában 0,4nm és 0,8nm közötti tartományban esnek. A gyakorlatban alkalmazott hullámok hullámhossza általában kisebb, mint 2nm. (A röntgenhullámok elhajlása valójában a szórásról alapszik, melynek mechanizmusát például a Thomson- szórás írja le.)

Az atomi távolságok meghatározásának alapjául a Bragg-feltétel szolgál.

A Bragg-féle értelmezésben a diffrakciós képet a rácscsíkokról visszaverődő hullámok interferenciája hozza létre. $2d \sin(\theta) = n\lambda$. Így egy elemi cella méretét lehet meghatározni, egy cellában általában több atom van. A szerkezet meghatározás bonyolultabb módszerekkel történik, amelyeket nem részletezek.



1.6. ábra Bragg feltétel illusztrációja

Forrás : Cserti József Optika és

Relativitáselmélet jegyzet ELTE 2007

A röntgendiffrakciós módszerrel kristályosítható fehérjék térszerkezete határozható meg. A kristályosított fehérje szerkezete viszonylag jól közelíti a vizes oldatokban felvett szerkezetet. A fehérje kristályok jelentős mennyiségű (25-90%) vizet tartalmaznak, valamint a kristályosítás során alkalmazott módszerek is a natív szerkezetet stabilizálják, ezért napjainkban a kristályszerkezetet az oldatban felvett konformáció jó közelítésének fogadják el. A röntgendiffrakciós mérések hátránya, hogy az alacsony rendszámú elemek a kevés elektron miatt nehezen mérhetőek. A röntgendiffrakcióval a molekula elektronsűrűség térképét kapjuk meg. A fehérje elsődleges szerkezetének ismeretében ebből megalkotható a molekula térszerkezete, kiszámolhatóak az atomi koordináták.

1.3.2 Mágneses magrezonancia

A módszer alapgondolata, hogy a külső mágneses térben a magok mágneses momentumai rendeződnek. A kialakult rendezettség rádiófrekvenciás térrel változtatható. Az NMR spektroszkópia szükséges feltétele, hogy a magspin nullától különbözzön. A magspint jele: I értéke $I=0$, ha az atommagban a protonok és a neutronok száma is páros (pl. ^{12}C)

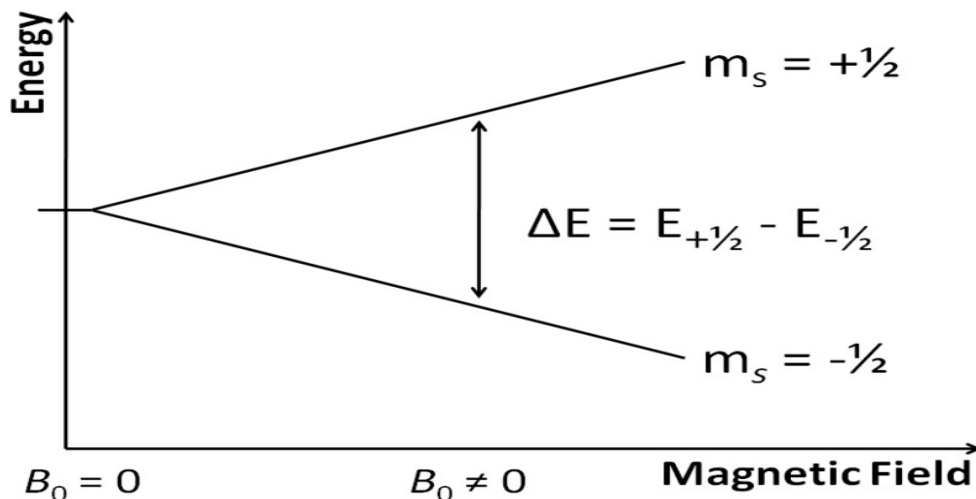
$I=1/2$ páratlan tömegszámú atommagok esetén (pl. ^1H)

$I=k$ $k=1,2,3,\dots$, amikor a protonok és neutronok száma is páratlan (pl. ^{14}N)

Z irányú mágneses térbe helyezett magok esetén az energianívók:

$\hat{H}_z = -\hat{\mu}_z B_0$ A Hamilton operátor sajátértékei $E_m = -\gamma \hbar m B_0$. Ahol μ_z a mágneses operátor z komponense, γ a giromágneses faktor.

A felhasadt energianívók távolsága $\Delta E_m = \gamma \hbar B_0$. Megengedett átmenetek csak a szomszédos energianívók között vannak. A rádiófrekvenciás gerjesztés során a minta elnyeli a sugárzást és az alacsonyabb energiájú állapotból magasabb energiájú állapotba kerül.



1.7. ábra Az atommag energianívóinak felhasadása mágneses térben 1/2-es spin esetén. (forrás: http://en.wikipedia.org/wiki/Nuclear_magnetic_resonance)

A színkép finomszerkezete (spin-spin csatolás)

Az egymással kémiai kötésben lévő atommagok között a köztük lévő elektronok miatt skaláris J csatolás van jelen. Egy NMR aktív mag „megérzi” a környezetében lévő többi magok mágneses állapotát. A hatást a kötésekön keresztül az elektronrendszer közvetíti. Ez ad lehetőséget arra, hogy az NMR-jelet molekulaszervezet meghatározására lehessen használni. A csatolást egy J csatolási állandóval jellemezhetjük.

A fehérjék szerkezetének meghatározása NMR-el

Az egymással kémiai kötésben levő magok között az elektronok közvetítésével megvalósuló J-csatolás eredményeképpen a rezonanciajel a legtöbb esetben nem szingulett, hanem finomszerkezettel rendelkezik. A multipllett vonalainak távolságából meghatározhatjuk a csatolási állandókat. A kémiai eltolódás (az abszorpciós frekvencia jellemző az atommagra, de kis mértékben függ az elektronsűrűségtől a mag környezetében, azaz jellemző a molekula szerkezetére), és a csatolási állandók ismeretében a vizsgált molekulában funkciós csoportokat, spinrendszereket azonosíthatunk. A J-csatolás fontos tulajdonsága, hogy információval szolgál a funkciós csoportok geometriájáról, például torziós szögekről. Az NMR kísérlet során rádiófrekvenciás impulzusokkal gerjesztett magok relaxációs folyamatok által térnek vissza az alapállapotba. A relaxációs mechanizmusok egyike egy a téren át ható spin-spin kölcsönhatáson, a dipoláris csatoláson

alapszik. Ez utóbbi távolságfüggésének köszönhetően mag-mag távolságinformációk kaphatók. A J-csatolással ellentétben a dipoláris kölcsönhatás nem kötések mentén, hanem téren át hat, ezáltal az intermolekuláris távolságok is megállapíthatók. A mért spektrumok segítségével spinrendszereket azonosítanak, melyek a rezonanciajelek kémiai eltolódásai és korrelációs mintázatai alapján megfeleltethetők egy-egy aminosavtípusnak. A szerkezetmeghatározás végső fázisa a szerkezetszámolás. Ennek során az aminosavsorrend ismeretében és a ^1H - ^1H távolságkényszerek birtokában molekuladinamikai módszerek alkalmazásával szerkezeteket generálnak és finomítanak a két folyamat között addig iterálnak, amíg a szerkezet kellő pontosságú nem lesz. A távolságkényszerek mellett a szerkezetszámolást segíti a kémiai eltolódásból nyert másodlagos szerkezetre vonatkozó információ. A szerkezetszámolás eredménye a röntgendiffrakcióval ellentétben nem egy, hanem a kényszerfeltételeknek eleget tevő tipikusan 10 vagy 20 legalacsonyabb energiájú szerkezetből álló sokaság. Ezért nagy felbontású NMR méréssel információt kaphatunk a fehérjék dinamikájára vonatkozóan is.

1.3.3 Infravörös spektroszkópia

Az elektromágneses sugárzásnak a 800 nm és 1 mm közé eső hullámhosszúságú tartományba eső részét infravörös fénynek nevezzük. Spektroszkópiai szempontból az infravörös tartományt három részre szokták bontani közeli 800 nm-2500 nm, közép 2500 nm-25000 nm, távoli 25 μm -1000 μm tartományra. Infravörös sugárzással gerjeszthetők a molekulák rezgési és rotációs nívói közti átmenetek. A molekulák szerkezetére és konformációjára jellemző rezgések a közép-infravörös tartományon detektálhatóak. A rezgési spektroszkópiában a hullámhossz helyett gyakran annak reciprokát a hullámszámot használják. A közép-infravörös tartomány hullámszámmal kifejezve 400 cm^{-1} -4000 cm^{-1} közé esik.

Infravörös gerjesztés hatására az elektronszerkezet nem változik. Azok a molekularezgések amelyek során a molekula dipólusmomentuma megváltozik, gerjeszthetők az infravörös sugárzással. Gerjesztéskor a gerjesztő sugárzás energiája megegyezik a végső és a kiindulási állapotok energiáinak a különbségével. $h\nu = |E_v - E_k|$ Az átmenet során a rezgési kvantumszám $n \pm 1$ -et változik. A rezgés nagyon érzékeny a rezgésben részt vevő atomokat összekötő kémiai kötések erősségére, illetve a molekula szerkezetére. A fehérjék infravörös spektroszkópiai vizsgálatainál gyakran felmerülő probléma a másodlagos szerkezet arányainak meghatározása az infravörös spektrum alapján, a fehérjét alkotó aminosavak hány százaléka alkot α -helikális, β -lemezes és egyéb, másodlagos szerkezetet.

1.3.4 Molekuladinamikai szimulációk

A molekuladinamikai szimulációkkal ismert szerkezetű fehérjék dinamikai tulajdonságait vizsgálhatjuk. Megvizsgálhatjuk például milyen mozgást végez, hogyan változik a dinamika ligandum kötés hatására. A fehérje méretű rendszerek szimulációja esetén klasszikus algoritmusokat

használnak. Minden atomra a Newton-törvényt alkalmazzuk: $m_i \frac{d^2 r_i}{dt^2} = -\frac{\partial V_K}{\partial r_i}$

Az erőt a kölcsönhatási potenciál negatív gradienseként kapjuk meg. A kölcsönhatási potenciál általában a következő tagokat tartalmazza:

$$V_K = \sum_{\text{kötések}} k_b (b - b_0)^2 + \sum_{\text{kötési szögek}} k_\Theta (\Theta - \Theta_0)^2 + \sum_{\text{dihedrális}} \sum_{n=1}^N K_\Phi^{(n)} [1 + \cos(n\Phi - \delta)] + \sum_{\text{mp}} K_\omega (\omega - \omega_0)^2 + \sum_{i,j} \frac{q_i q_j}{D r_{i,j}} + \sum_{i,j} \epsilon_{i,j} \left[\left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{12} - \left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^6 \right]$$

Az első tag a kötéstávolságokból, a második a kötésszögekből, a harmadik és negyedik tag a torziós szögekből adódik, az ötödik az elektrosztatikus Coulomb potenciál és az utolsó tag a Van der Waals potenciál. A differenciálegyenletet numerikusan megoldva megkapjuk az atomok sebességét és helyzetét adott időintervallumonként. Így egy trajektóriához jutunk. A numerikus megoldás során a Δt lépésköz fs nagyságrendbe esik, a trajektória pedig 10ns- 100 ns hosszúságú.

1.4 Molekularezgések

Képzeljünk el egy rugóra függesztett tömegpontot, amely egy irányban rezeg. Erre a rendszerre felírhatunk egy differenciálegyenletet: $m \frac{d^2 x(t)}{dt^2} = \frac{-dV(x)}{dx} = -kx(t)$ (5), amely az 1 dimenziós

harmonikus oszcillátor differenciálegyenlete. m a tömegpont tömege, $V(x)$ a rendszerhez tartozó potenciál, k a rugó direkciós ereje. A differenciálegyenlet megoldása $x(t) = A \sin(2\pi \nu t)$ (6)

Ahol ν a rezgés frekvenciája, A , a maximális kitérés. (5)-ik egyenletbe (6)-ost behelyettesítve azt kapjuk, hogy $-4\pi^2 \nu^2 m x(t) = -kx(t)$ (7)

A molekulák atomjainak mozgását leírhatjuk normál-koordináták és -rezgések bevezetésével. N atomból álló molekula esetén a mozgások előállíthatók $3N-6$, az egyensúlyi állapot körüli harmonikus rezgésekből. Egy ilyen rezgési módusban az atomok azonos frekvenciával rezegnek. A legegyszerűbb vibrációra képes molekula a kétatomos molekula, amelynek rezgései egyszerűen leírhatók. A többatomos molekulák rezgései is számolhatók, amennyiben a kötésekre jellemző erőállandók, valamint az egyensúlyi koordináták ismertek.

Többatomos molekula esetén legyenek az i -ik atom koordinátái X_i, Y_i, Z_i az egyensúlyi koordinátái X_{ie}, Y_{ie}, Z_{ie} . Képezzük az i -ik atomra az atom koordinátáinak és a hozzá tartozó egyensúlyi koordinátáinak a különbségeit $x_i = X_i - X_{ie}$. Az erőállandók a potenciális energia második

deriváltjaiként képezhetők. $k_{xx}^{11} = \frac{\partial^2 V(\underline{r})}{\partial x_1^2}$, $k_{xy}^{11} = \frac{\partial^2 V(\underline{r})}{\partial x_1 \partial y_1}$, $k_{xy}^{ij} = \frac{\partial^2 V(\underline{r})}{\partial x_i \partial y_j}$ stb

A Newton törvényt (5) az i -ik atomra felírva

$$-4\pi^2 \nu^2 m_i x_i = -k_{xx}^{i1} x_1 - k_{xy}^{i1} y_1 - k_{xz}^{i1} z_1 - k_{xx}^{i2} x_2 - \dots - k_{xz}^{iN} z_N \quad (8)$$

Így $3N$ db egyenletet kapunk, amit felírhatunk mátrixos alakban is. Ez egy sajátértékprobléma, amit numerikusan megoldhatunk, $3N \times 3N$ -es mátrix sajátértékei adják a normál módusok rezgési frekvenciáinak négyzetét ν^2 -et.

Az összetett molekulák rezgései között vannak olyanok, amelyekben csak néhány atom vesz részt. Ezeket a néhány atomra korlátozódó rezgéseket csoportrezgéseknek nevezzük. Csoportrezgések frekvenciája kis mértékben függ attól, hogy a rezgő csoport milyen molekulába van beágyazva. Ez a kismértékű eltolódás alkalmas arra, hogy az adott rezgésből információt nyerjünk a makromolekula konformációjáról. Csoportrezgésekre jó példa a lipidek zsírsavláncainak CH_2 rezgései, illetve a fehérjék konformáció-változásainak detektálására használt amid rezgések.

1.4.1 N-metilacetamid rezgései

A csoportrezgések kicsit függnak a molekula többi részétől, ez a függés első közelítésben elhanyagolható. Miyazawa és munkatársai 1958-ban kiszámolták az N-metilacetamid molekula normálrezgéseit, és rámutattak az N-metilacetamid rezgései és a polipeptid lánc gerincének rezgései közti kapcsolatra. Az N-metilacetamid molekula számunkra legfontosabb normálrezgései a 6. ábrán láthatóak. Az amid I és amid II rezgések a molekula végein található szénatomokat nem vagy csak kis mértékben érintik. Ez biztosítja, hogy az amid csoportrezgések a fehérjére is átvihetők legyenek (1.41-es táblázat és 1.7 ábra), valamint azt, hogy az egyes aminosavak aminocsoportjait, az átmeneti dipól csatolástól eltekintve független oszcillátoroknak tekinthessük. Ennek a molekulának a rezgéseit több cikkben is újrászámolták.

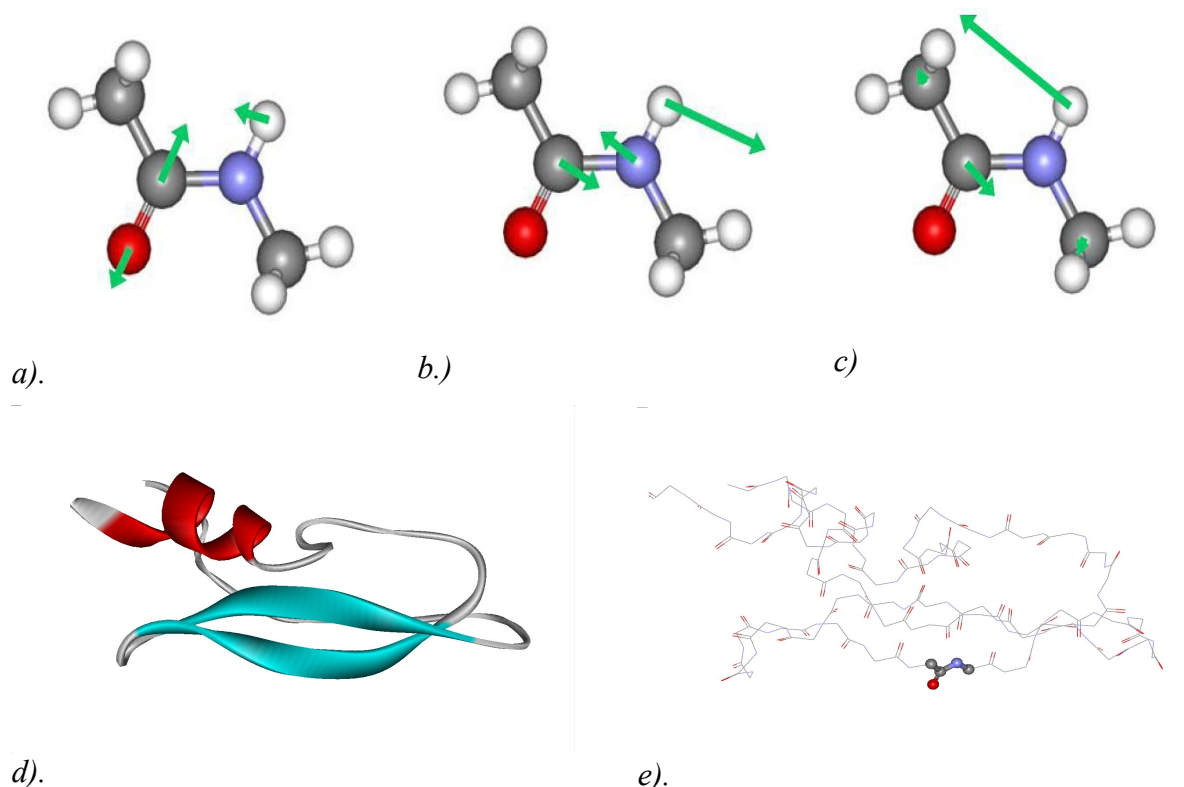
1.4.2 A fehérjék amid I rezgései

Ezen a téren Krimm és Bandekar (1986) végezték a legalaposabb számolásokat és bebizonyították, hogy az N-metilacetamidra vonatkozó erőállandók átvihetők rövid polipeptidekre és fehérjékre is. A fehérjében minden aminosavhoz egy normálrezgés rendelhető, a polipeptid lánc egy csatolt rezgő

rendszernek tekinthető, ahol az átmeneti dipólok egymással kölcsönhatásban vannak. Ez az átmeneti dipól csatolás (Torii és Tasumi, 1992) természetesen függ az egymás közelében levő dipólok orientációjától, azaz végső soron a fehérje másodlagos szerkezetétől. Az amid I rezgés a leginkább konformáció-érzékeny. A konformáció-érzékenységet a már említett csatoláson kívül az biztosítja, hogy a rezgés energiájának nagy része (83%) a C=O nyújtási módusból származik. Ezen oxigén és a fehérje egy másik aminosavjának NH hidrogénje között a konformációtól függő hidrogénhíd kötés jön létre.

Az amid I sáv komponenseinek hozzárendelése a másodlagos szerkezeti elemekhez (Byler és Susi (1986), Haris és Chapman, 1988, Ismail és munkatársai, 1992)	
hullámszám [cm^{-1}]	másodlagos szerkezet
1616	intermolekuláris béta szerkezet
1624-1637	kinyújtott láncok béta szerkezet
1645	rendezetlen
1654	alfa hélix
1662	310 helix
1663-1670	hajlatok, hurkok
1675	kinyújtott láncok (béta szerkezet)
1683-1694	hajlatok, hurkok
1685	intermolekuláris béta szerkezet

1.4.1 táblázat: Az amid I sáv komponenseinek hozzárendelése másodlagos szerkezet elemeihez



1.7. ábra Az N-metilacetamid molekula szerkezete, annak kapcsolata fehérjékkel és a molekula egyes normálrezgései: a) amid I, b): amid II c) amid III (Bandeekar (1992) alapján). A zöld nyilak az atomok elmozdulási irányait és azok relatív amplitúdóit szemléltetik, a valódi kitérések az ábrán jelzettnél jóval kisebbek d) BPTI fehérje e) BPTI peptidgerince

2. Alkalmazott módszerek

2.1 Fourier Transzformáció

A Fourier transzformáció egy elterjedt integráltranszformáció, a műszaki és természettudományok számos területén alkalmazzák.

Egy $f(t)$ függvény $F(\omega)$ Fourier transzformáltját az alábbi képlettel számolhatjuk ki:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{i\omega t} dt \quad (9)$$

Az inverz transzformációval visszakaphatjuk az eredeti függvényt.

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{-i\omega t} d\omega \quad (10)$$

A gyakorlatban sokszor a folytonos $f(t)$ függvény helyett egy diszkrét pontokat tartalmazó adatsorunk van, illetve a számítógép is diszkrét adatpontokkal számol. Ezért a Fourier transzformáció diszkrét változatát szoktuk gyakran alkalmazni.

Legyen N darab diszkrét adatunk, ekkor $f_k = f(t_k)$ $t_k = k\Delta$,
 k egy egész szám, $k=0, 1, 2, \dots, N-1$ Δ a mintavételezési idő.

A (9)-os integrált diszkrét szummával közelítjük:

$$F(f_n) = \int_{-\infty}^{\infty} f(t) e^{i\omega t} dt \approx \sum_{k=0}^{N-1} f_k e^{(2\pi f_n t_k)} \Delta = \Delta \sum_{k=0}^{N-1} f_k e^{(2\pi \frac{ikn}{N})} \quad (11)$$

A munkám során a Fourier transzformációk elvégzéséhez egy C nyelven írt programot használtam, az abban használt FFT algoritmus alapját a [7]-ben található `four1` nevű függvény képezi.

2.2 Infravörös spektrumok illesztése

A fehérjék infravörös spektrumát az egyéb módszerekkel meghatározott egyensúlyi atomkoordinátákkal próbáltam előállítani. A spektrumot egy konstans alapvonal meghatározása után

Lorentz- görbék összegével illesztettem. $f(x) = \frac{2}{\pi} \frac{w^2}{4(x-x_c)^2 + w^2}$, w a görbe félértékszélességét

adja, x_c az x tengely mentén pozicionálja a görbe középpontját, A az amplitúdója.

Az illesztést excel-ben végeztem a beépített nem lineáris megoldó segítségével. Az alábbi

függvényt minimumát kerestem: $\chi^2 = \sum_j (A \sum_i f_i(x_j) + t + (y_i - y_0)m - y_i)^2$ (12)

Minden i -ik aminosavhoz kiszámoltam egy Lorentz-görbét majd ezeket összegeztem (12).

Megszoroztam egy közös amplitúdóval és hozzáadtam az alapvonalat. A képletben $f_i(x)$ az i -k aminosav Lorentz-függvénye, y_j a j -ik hullámszámnál mért abszorbanancia, t az alapvonal értéke az $y_0 = 1600 \text{ cm}^{-1}$ -nél, m az alapvonal meredeksége.

3. Célkitűzés

Ismert szerkezetű fehérjék térszerkezetének és spektrumának összehasonlító elemzése. Az elemzés központi kérdése, hogy a fehérje adatbázisokban található szerkezeti adatokból, illetve molekuláris dinamikai számolások eredményeként kapott trajektóriákból milyen módon kaphatók meg a fehérje szerkezetre jellemző rezgések.

4. Eredmények

Munkám során kapcsolatot kerestem a fehérjék atomkoordinátáiból nyerhető szerkezeti adatai és az infravörös spektrumuk között. A spektrumuk 1600 cm^{-1} - 1700 cm^{-1} közé eső tartományán a másodlagos szerkezetre érzékeny rezgések adnak jelet. Ezek a rezgések a másodlagos szerkezetet stabilizáló hidrogénkötésekhez kötődnek, konformációérzékenyek, a rezgések energiájának döntő többsége a C=O nyújtási módusából adódik. A fehérjében minden aminosavhoz hozzárendelhető egy normálrezgés. Ezeket a rezgéseket független oszcillátoroknak tekinthetjük. Feltevésünk szerint kapcsolat van a kötéshosszak és a rezgések frekvenciái között. Klasszikusan a C=O-H-N kötés golyós rugós rendszerrel modellezhető. A rezgő rendszerhez egy $V(r)$ potenciált feltételezek. Ebből a potenciálból származtattam a rezgés frekvenciáját.

$$V(r) = V_0 + \frac{D}{2}(r-r_0)^2 + \frac{c_1}{6}(r-r_0)^3 + \hat{c}_2 V_{O-N}(r, r_{O-N}) + \hat{c}_3 V_\varphi(r, \varphi) + \dots \quad (13)$$

ahol, r a C-O távolság, r_0 az egyensúlyi távolság.

Az első tag egy konstans, a második a rugó direkción erejét tartalmazza. Az első két tag harmonikus potenciálnak felel meg. A többi tag tükrözi azt a feltételezésünket, hogy a H-kötés távolsága, szöge befolyásolja a rezgési frekvenciát. Ezek a tagok módosítják, illetve anharmonikussá teszik a potenciált. V_{O-N} függvénye az r -nek és az O-N távolságnak, V_φ függ r -től és a φ kötésszögtől. (C=O és O-N vektorok által bezárt szög, amikor 180° akkor a kötés lineáris.)

$$V(r) = V_0 + \frac{D}{2}(r-r_0)^2 + \frac{c_1}{6}(r-r_0)^3 + \frac{\hat{c}_2}{2} r_{O-N}(r-r_0)^2 + \frac{\hat{c}_3}{(r-r_0)^2} r_{C-N} \cos(\varphi) + \dots \quad (14)$$

A rendszer rezgési frekvenciája arányos a potenciál második deriváltjával.

$$f \sim \frac{\partial^2 V(r)}{\partial r^2} \sim D + c_1(r-r_0) + \hat{c}_2 \frac{\partial^2 V_{O-N}(r, r_{O-N})}{\partial r^2} + \hat{c}_3 \frac{\partial^2 V_\varphi(r, r_{C-N}, \varphi)}{\partial r^2} \quad (15)$$

(14)-ben a 3., 4. tagot kezdetben r_{O-N} és r_{C-N} szerint lineárisnak feltételeztem így (14)-ben felírt képlettel próbálkoztam. $f \sim D + c_1(r-r_0) + c_2 r_{O-N} + c_3 r_{C-N} \cos(\varphi)$ (16)

A koordinátaadatokból kiszámoltam a C=O távolságokat minden aminosavra. A másodlagos szerkezetet alkotó aminosavakhoz (amelyeknél az oxigénatom intermolekuláris hidrogénhidrokötésben vesz részt) kiszámoltam az O-N távolságot és az C=O és O-N vektorok által bezárt szöget.

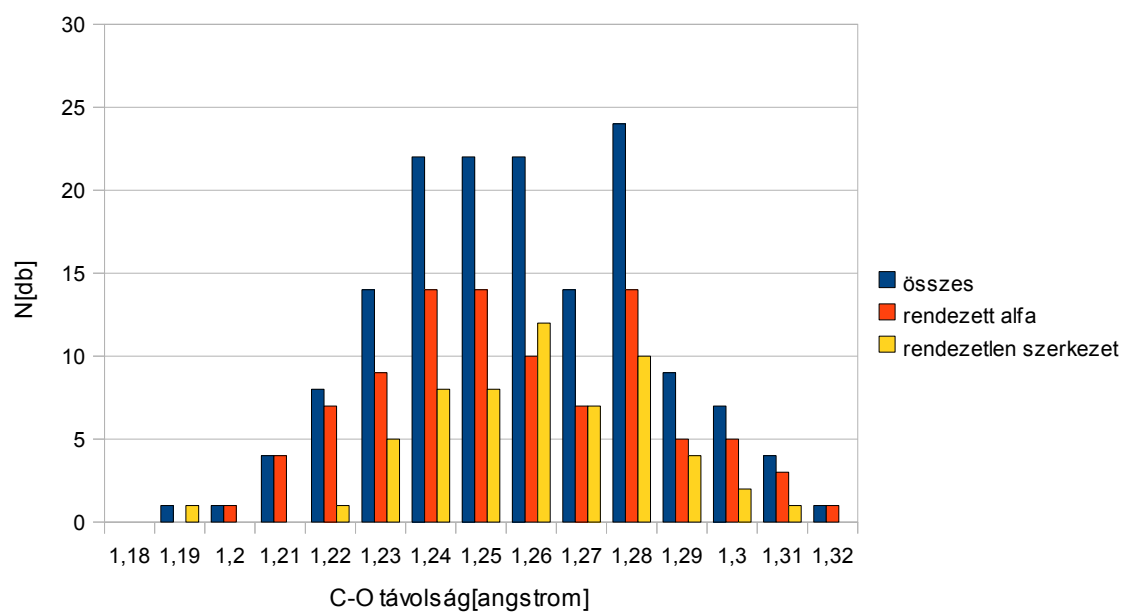
4.1 C-O és O-N távolságok különböző fehérjék esetén.

Először a lószív mioglobint és a marha tripszin inhibitor röntgendiffrakcióval meghatározott 1YMB és 4PTI szerkezeteit vizsgáltam.

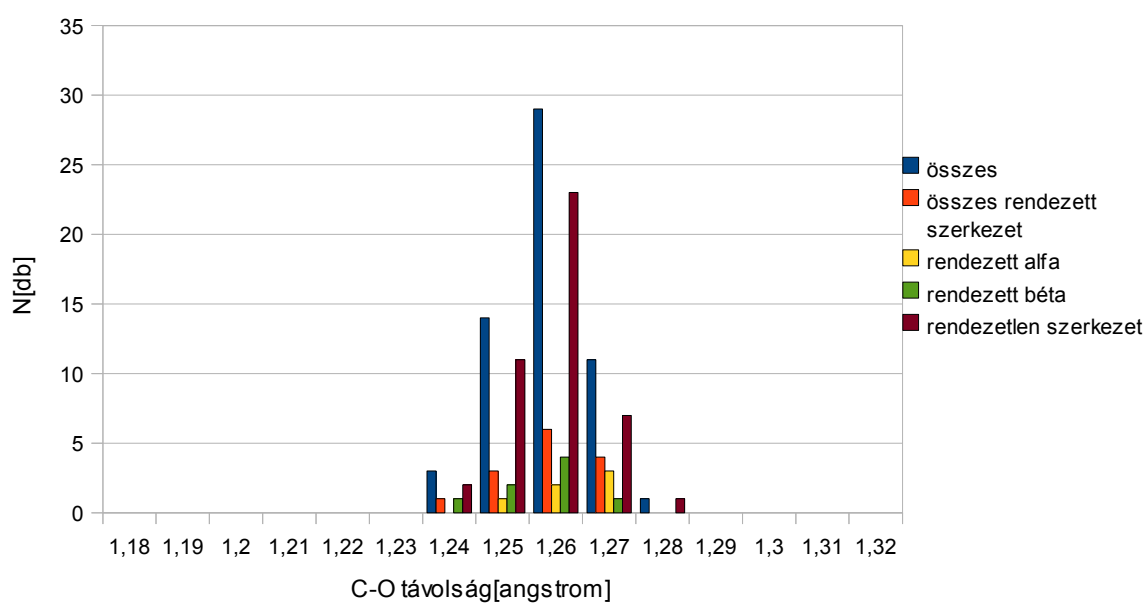
A lószív mioglobint egy 153 aminosavat, 8db alfa-hélixet tartalmazó fehérje. A marha tripszin inhibitor fehérje 58 aminosavat tartalmaz, tartalmaz egy α -hélixet és két β -redőt.

A 4.1-es diagramok alapján látható, hogy a C-O és O-N távolságok nem különülnek el a különböző szerkezeti elemekre, vannak átfedések a rendezett és rendezetlen elemekhez tartozó távolságok között. Az alfa és béta szerkezethez tartozó távolságadatok sem különülnek diszjunkt tartományokba. Ez az infravörös spektrumok illesztése szempontjából nem a legjobb, azonban még így is sikerült az illesztésnél az egyes másodlagos szerkezeti, α -hélixekhez és β -redőkhöz tartozó görbéket elkülöníteni. (ld. 4.2)

1YMB és 4PTI C-O távolságeloszlás diagramjai

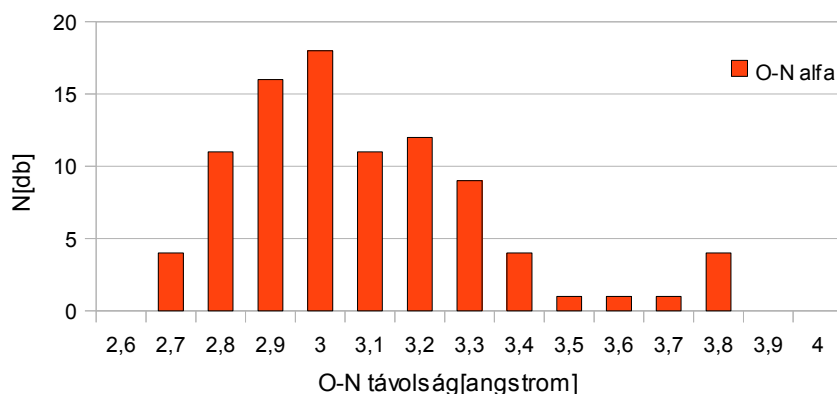


4.1.1 diagram: 1YMB szerkezet C-O távolságeloszlás diagramja

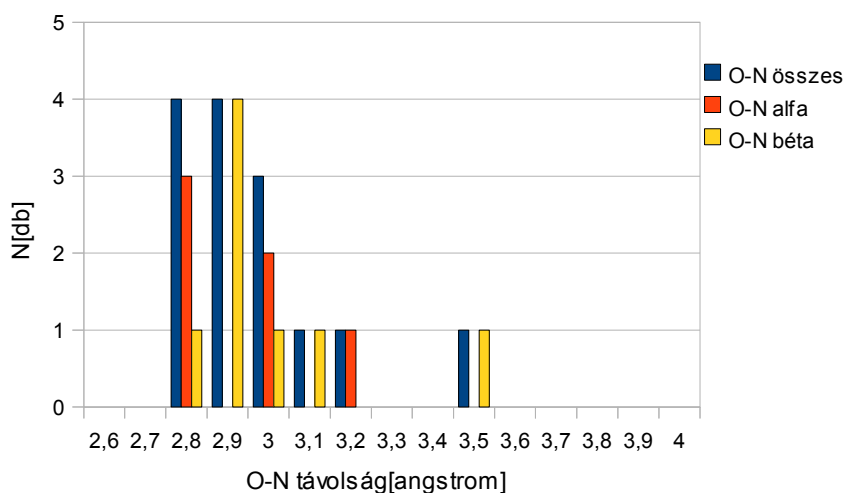


4.1.2 diagram: 4PTi C-O távolságeloszlás diagramja

1YMB és 4PTI O-N távolságeloszlás diagramjai



4.1.3. diagram: 1YMB O-N távolságeloszlása



4.1.4. diagram: 4PTI O-N távolságeloszlása

			átlag[angstrom]	szórás[angstrom]	maximum[angstrom]	minimum[angstrom]	terjedelem[angstrom]
4PTI	C-O	összes	1,254	0,008	1,276	1,235	0,041
		rendezetlen	1,254	0,008	1,276	1,235	0,041
		alfa+béta	1,255	0,008	1,269	1,235	0,034
		alfa	1,259	0,007	1,269	1,250	0,019
		béta	1,251	0,008	1,260	1,235	0,025
	O-N	összes	2,934	0,190	3,469	2,755	0,715
		alfa	2,890	0,140	3,110	2,763	0,346
		béta	2,966	0,224	3,469	2,755	0,715
1YMB	C-O	összes	1,253	0,025	1,319	1,185	0,134
		rendezetlen	1,255	0,022	1,309	1,185	0,124
		alfa	1,251	0,027	1,319	1,195	0,123
	O-N	alfa	3,030	0,254	3,765	2,636	1,128

4.1.1 táblázat: 4PTI és 1YMB szerkezet távolságadatok néhány statisztikai mutatója

Megvizsgáltam röntgendiffrakcióval meghatározott mioglobinnak szerkezetének C-O és O-N távolságait.

A 4.1.2-es, 4.1.3-as és 4.1.3-as táblázatok alapján látható, hogy a szerkezetek átlagos távolságai között nagyobb az eltérés, mint az egyes szerkezeteken belül a távolságok szórása, vagy a β és α -hoz tartozó távolságok közti különbségek. Ez arra mutat, hogy a szerkezetmeghatározás utolsó lépésében alkalmazott „finomító” programoktól nagy mértékben függenek a PDB fájlban meghatározott koordináták. Ezért megnéztem, NMR-el meghatározott szerkezeteket.

C-O távolságok összes aminosavra [angstrom]					
	max	min	terjedelem	átlag	szórás
1YMB	1,319	1,185	0,134	1,253	0,025
3LR7	1,248	1,216	0,032	1,231	0,006
1DWR	1,264	1,216	0,048	1,233	0,008
1NPF	1,237	1,225	0,012	1,231	0,002
1WLA	1,299	1,194	0,105	1,247	0,017
2NSR	1,268	1,208	0,059	1,236	0,013
2O5O	1,264	1,201	0,063	1,236	0,009
3LR9	1,245	1,218	0,026	1,231	0,005

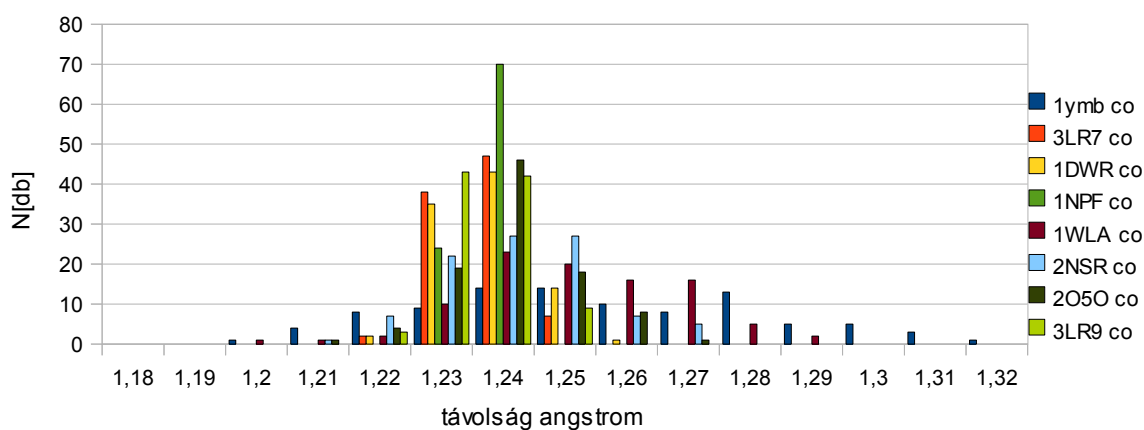
4.1.2. táblázat: röntgendiffrakcióval meghatározott szerkezetek C-O távolságadatainak néhány statisztikai mutatója

C-O távolságok alfa hélix [angstrom]					
	max	min	max-min	átlag	szórás
1YMB	1,319	1,195	0,123	1,250	0,027
3LR7	1,248	1,216	0,032	1,231	0,006
1DWR	1,256	1,218	0,039	1,232	0,007
1NPF	1,236	1,225	0,011	1,232	0,002
1WLA	1,286	1,194	0,092	1,246	0,017
2NSR	1,268	1,208	0,059	1,237	0,013
2O5O	1,264	1,201	0,063	1,236	0,010
3LR9	1,245	1,218	0,026	1,231	0,006

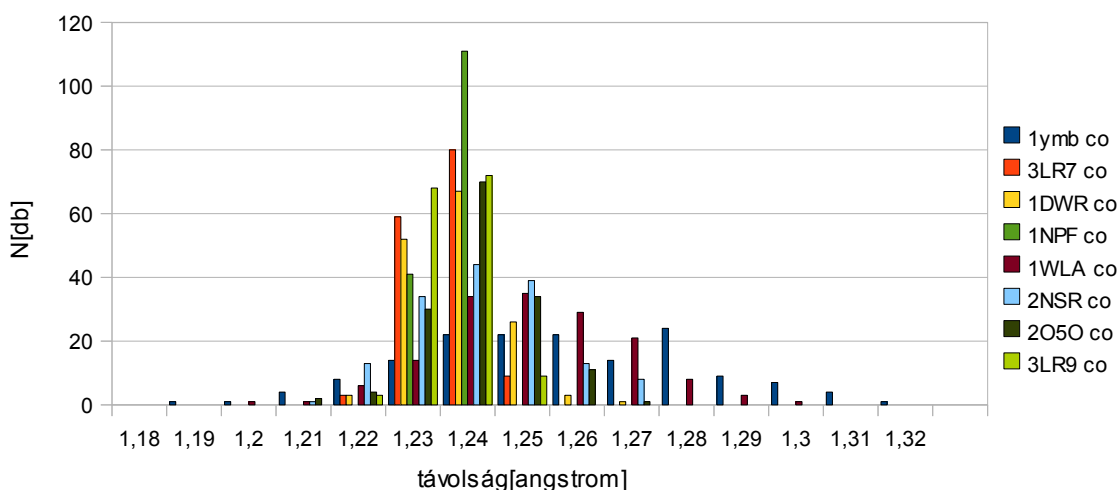
4.1.3. táblázat: röntgendiffrakcióval meghatározott szerkezetek O-N távolságadatainak néhány statisztikai mutatója

O-N távolságok [angstrom]					
	max	min	terjedelem	átlag	szórás
1ymb	3,994	2,636	1,358	3,059	0,297
3LR7	3,958	2,705	1,253	3,012	0,205
1DWR	3,997	2,748	1,249	3,036	0,231
1NPF	3,950	2,665	1,285	3,023	0,224
1WLA	3,872	2,681	1,191	3,033	0,238
2NSR	3,957	2,749	1,208	3,016	0,218
2O5O	3,975	2,694	1,281	2,991	0,232
3LR9	3,966	2,744	1,221	3,022	0,225

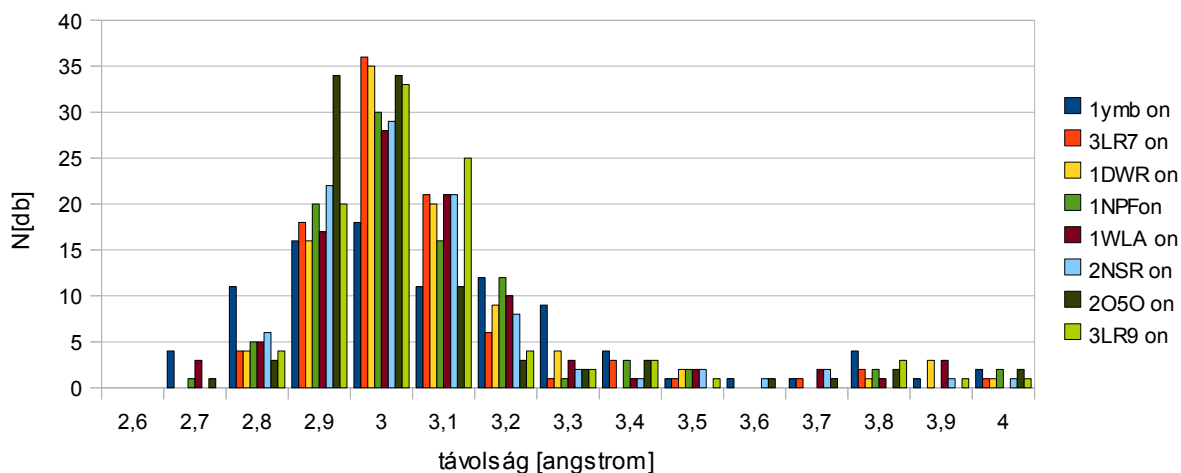
4.1.4 táblázat különböző röntgendiffrakcióval meghatározott alfa szerkezetű mioglobinnak O-N távolságadatainak néhány statisztikai mutatója



4.1.7. diagram különböző röntgendiffrakcióval meghatározott alfa szerkezetű mioglobinok C-O távolságeloszlásai rendezett alfa szerkezetben lévő aminosavakra



4.1.5. diagram különböző röntgendiffrakcióval meghatározott alfa szerkezetű mioglobinok C-O távolságeloszlásai összes aminosavakra



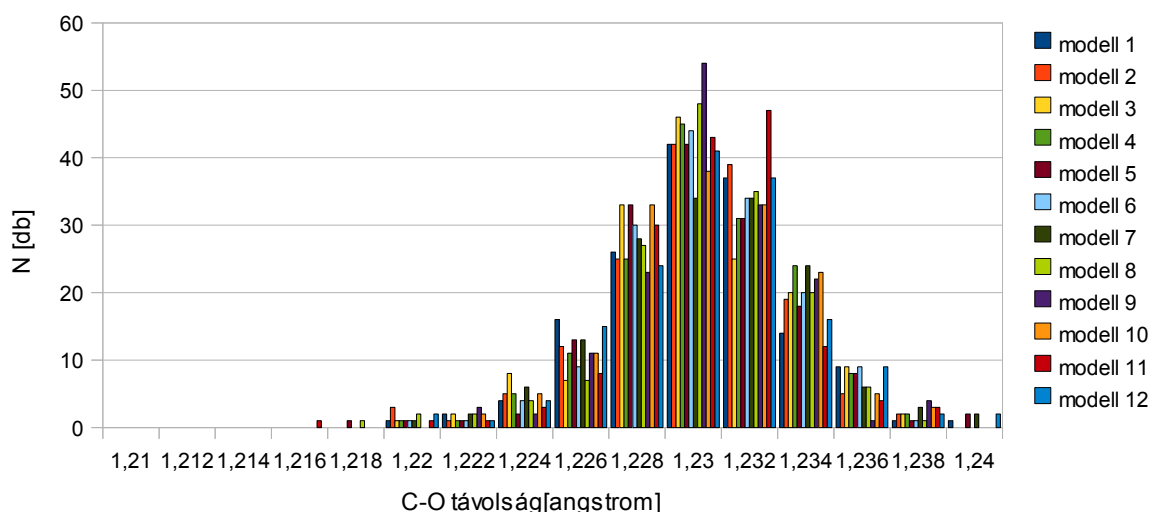
4.1.7 diagram különböző röntgendiffrakcióval meghatározott alfa szerkezetű mioglobinok O-N távolságeloszlásai

Megvizsgáltam mágneses magrezonanciával meghatározott alfa szerkezettel rendelkező fehérjék C-O és O-N távolságait is.

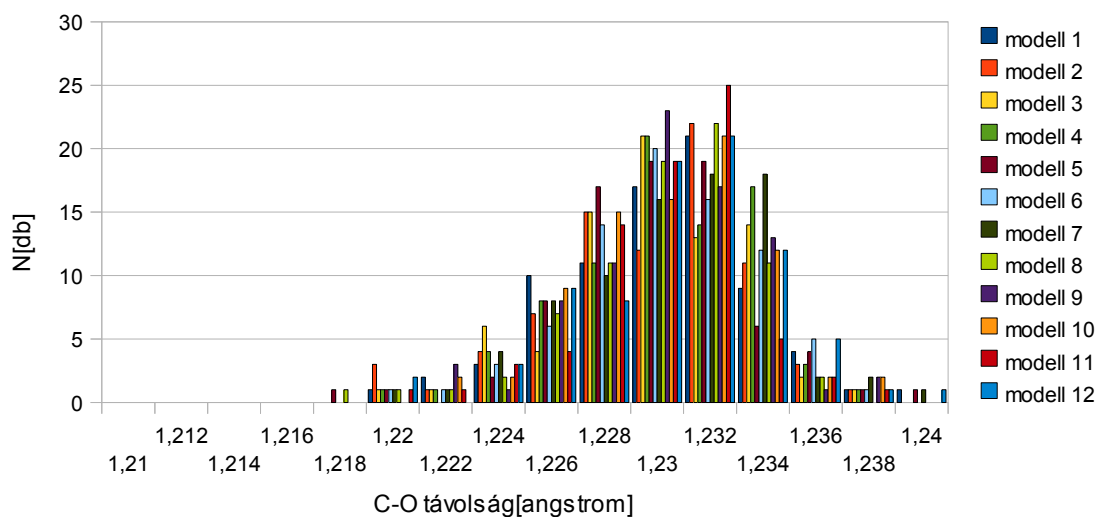
A mioglobin 1YMF PDB kódjelű NMR-el meghatározott 12 szerkezeti modelljét hasonlítottam össze. A táblázat alapján látható, hogy az egyes modellek között sokkal kisebb a különbség, mint a röntgendiffrakciós szerkezeteknél, jelen esetben ugyanazon szerkezet különböző modelljeit láthatjuk, amelyek kiszámítása azonos módszerrel történt.

	1MYF(NMR)											
	modell 1	modell 2	modell 3	modell 4	modell 5	modell 6	modell 7	modell 8	modell 9	modell 10	modell 11	modell 12
C-Oátlag	1,229	1,229	1,229	1,230	1,229	1,230	1,229	1,229	1,229	1,229	1,229	1,229
C-Oszórás	0,002	0,002	0,003	0,002	0,002	0,002	0,003	0,002	0,002	0,002	0,002	0,003
O-N átlag	2,885	2,866	2,863	2,913	2,878	2,928	2,889	2,870	2,883	2,913	2,905	2,889
O-Nszórás	0,116	0,099	0,098	0,144	0,115	0,170	0,110	0,106	0,113	0,152	0,139	0,110
C-O min	1,219	1,219	1,218	1,220	1,218	1,219	1,219	1,218	1,220	1,220	1,214	1,219
C-O max	1,240	1,237	1,238	1,238	1,239	1,238	1,239	1,237	1,238	1,237	1,238	1,240

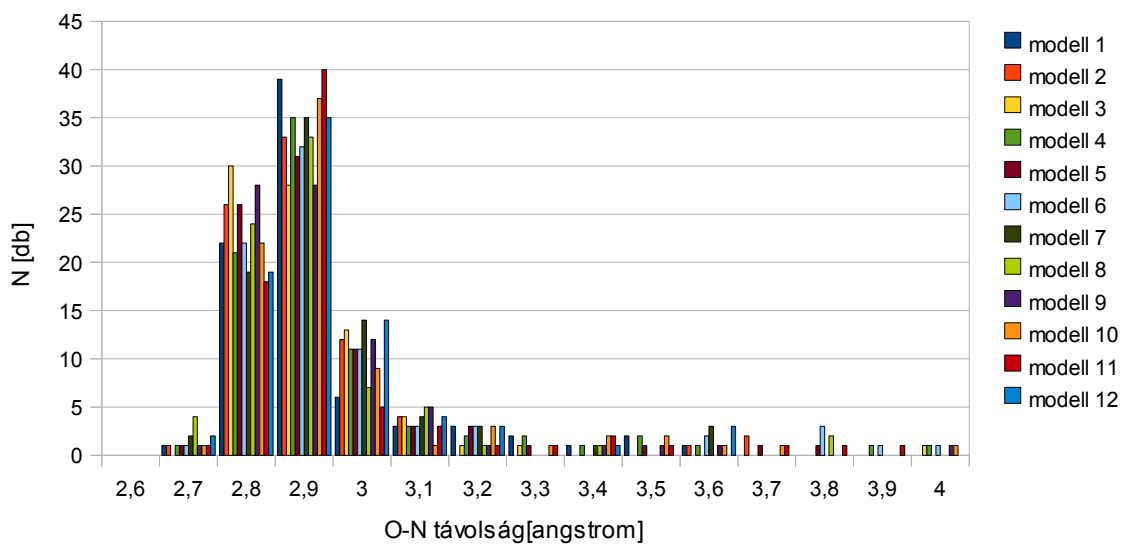
4.1.5 táblázat 1YMF NMR-el meghatározott alfa szerkezetű mioglobin O-N távolságadatainak néhány statisztikai mutatója 12 modell esetén



4.1.8 ábra 1YMF NMR-el meghatározott alfa szerkezetű mioglobin C-O távolságeloszlásai összes aminosavakra 12 modell esetén



4.1.9 ábra 1YMF NMR-el meghatározott alfa szerkezetű mioglobin C-O távolságeloszlásai rendezett aminosavakra 12 modell esetén



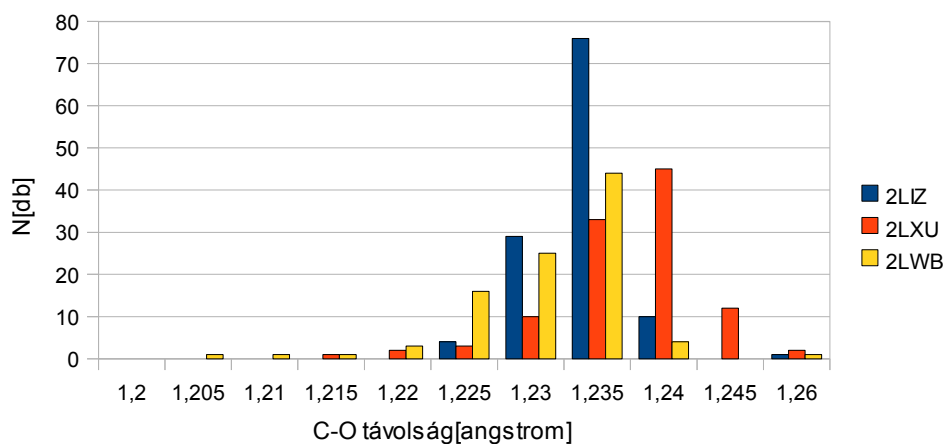
4.1.10 ábra 1YMF NMR-el meghatározott alfa szerkezetű mioglobin O-N távolságeloszlásai 12 modell esetén

Összehasonlítottam 3 db 2012-ben NMR-el készült fehérjeszerkezeti modellt is. (1-es számú modelleket vizsgáltam, alfa hélixeket tartalmaznak). Sars koronavírus, humán helikáz enzimet DDX58-at és egy heterogén ribonuklein proteint. Mindhárom fehérje helikális szerkezetű.

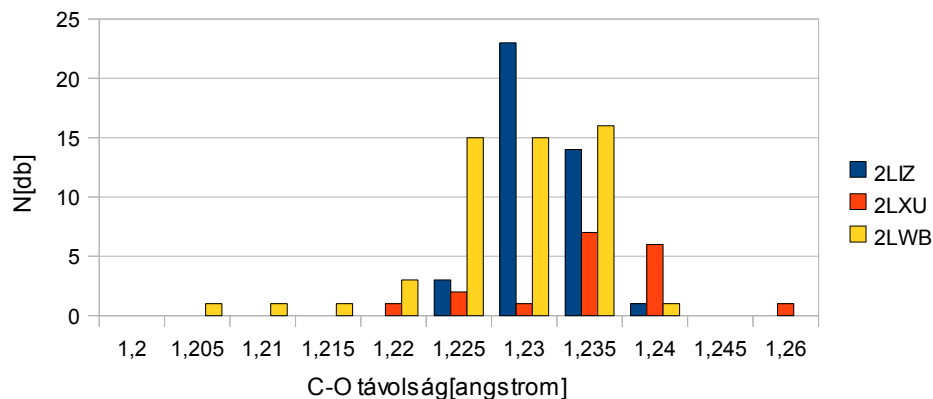
A 4.1.6-os táblázat alapján látható, hogy a különböző NMR szerkezetek távolságadatai között kisebb az eltérés, mint a röntgendiffrakciós módszerrel készült szerkezetek között. Ez is indokolja, hogy a kezdetben felhasznált röntgendiffrakciós adatokról az NMR-es adatokra való áttérést az illesztéseknél.

C-O és N-O távolságok						
		min	max	átlag	szórás	terjedelem
2LIZ	C-O	1,222	1,254	1,231	0,004	0,031
	C-Oalfa	1,222	1,235	1,229	0,003	0,013
	O-N	2,769	3,689	2,722	0,172	0,921
2LXU	C-O	1,215	1,249	1,235	0,006	0,034
	C-Oalfa	1,220	1,247	1,232	0,006	0,027
	O-N	2,741	3,562	3,030	0,178	0,821
2LWB	C-O	1,203	1,247	1,229	0,006	0,044
	C-Oalfa	1,203	1,236	1,226	0,006	0,032
	O-N	2,606	3,323	2,903	0,202	0,717

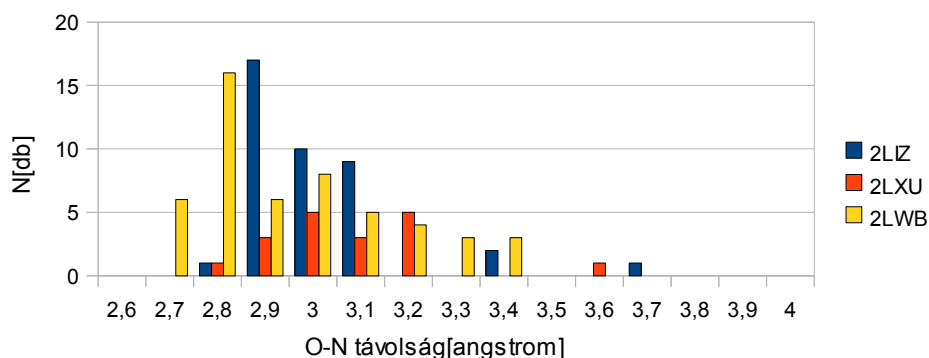
4.1.6 táblázat: 2012-es NMR szerkezeti modellek távolságadatainak néhány statisztikai mutatója



4.1.11 ábra: NMR szerkezetek C-O távolságeloszlás diagramja az össze aminosavra



4.1.12 ábra: NMR szerkezetek C-O távolságeloszlás diagramja a rendezett szerkezetben lévő aminosavra



4.1.13 ábra: NMR szerkezetek O-N távolságeloszlás diagramja

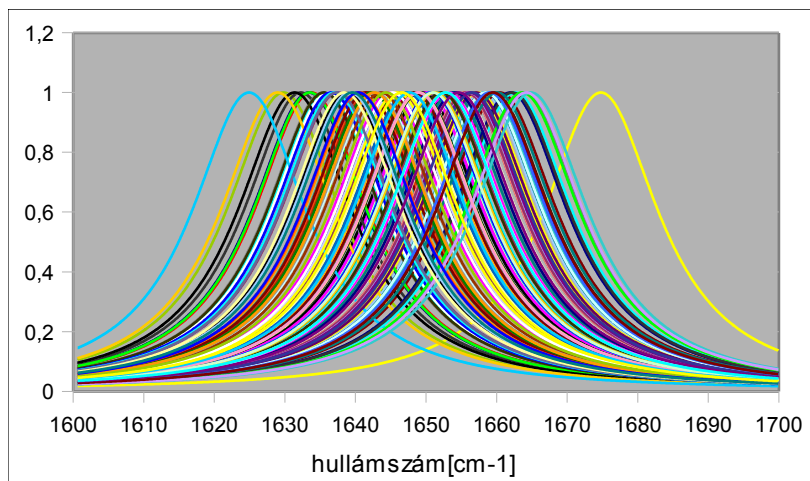
4.2 Infravörös spektrumok illesztése

A szerkezetanalízis során meghatározott C-O és O-N távolságok, valamint a C,O és O,N atomkoordinátákból előállított vektorok által bezárt C-O-N szögek segítségével megpróbáltam megilleszteni az IR spektrumokat. (Az IR spektrumok a Semmelweis egyetem Biofizika és Sugárbiológia Intézetében készültek. A mért spektrumok hibája sokkal kisebb, mint a szerkezeti adatok pontossága.) Ekkor a (13)-ban felírt $V(r)$ potenciálban található együtthatók értékeit variáltam. Elsőként röntgendiffrakciós szerkezeti adatokat felhasználva végeztem az illesztéseket. Minden aminosavhoz kiszámítottam egy w szélességű Lorentz-görbét. Ezeknek a görbéknek az összegéből próbáltam előállítani a spektrumot. A görbék pozícióját megadó X_c paraméterbe építettem bele a C-O és O-N távolságokat. Kezdetben $X_c = X_{c0} + c_1 r_{C-O} + c_2 r_{O-N}$ alakkal próbálkoztam. Ebben a formában az illesztésekre a $\chi^2 = 0,2$ közeli értékek adódtak. Módosítottam X_c -t, úgy hogy az O-N távolságot $\frac{1}{r}$ -es módon vettem figyelembe, minél kisebb az O-N távolság annál közelebb vannak egymáshoz az atomok, akkor az (9)-es képlettel számolt χ^2 is csökken.

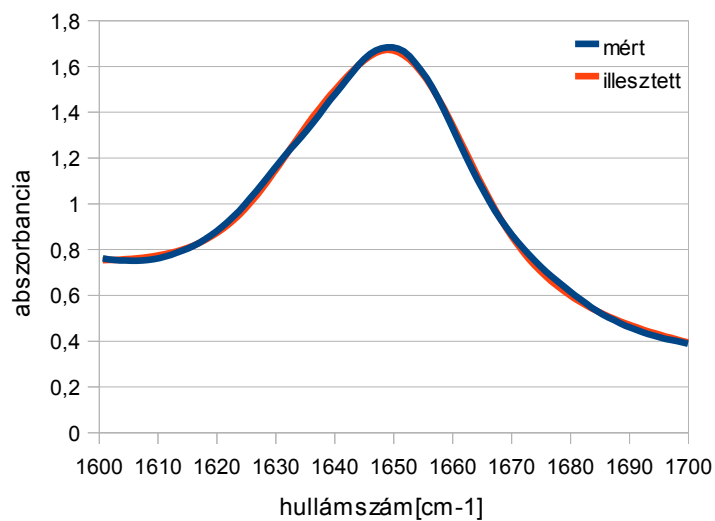
$X_c = X_{c0} + c_1 r_{C-O} + c_2 \frac{1}{r_{O-N}}$, így χ^2 értéke 0,018 illetve 0,017 lett, javult az illesztés.

1YMB							
c_1	c_2	Xc_0	w	A	tengelymetszet	meredekség	χ^2
-296,516	-41,333	2028,052	10,000	0,012	0,671	-0,003	0,018
4PTI							
c_1	c_2	Xc_0	w	A	tengelymetszet	meredekség	χ^2
-1601,046	-30,502	3658,784	10,000	0,024	0,519	-0,002	0,017

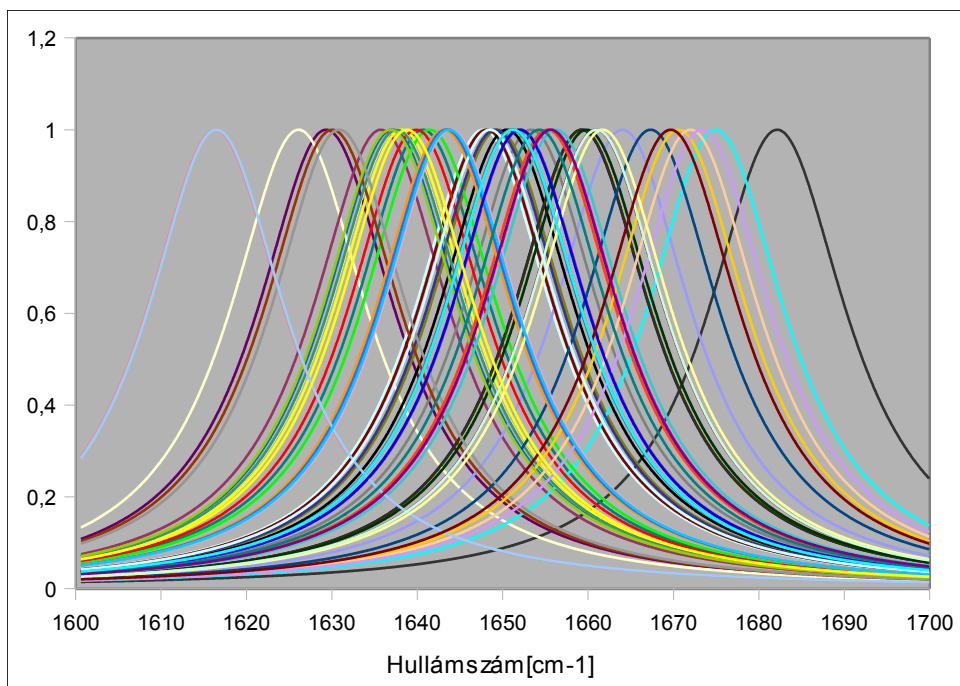
4.2.1. táblázat 1YMB és 4PTI szerkezet illesztésére kapott paraméterek



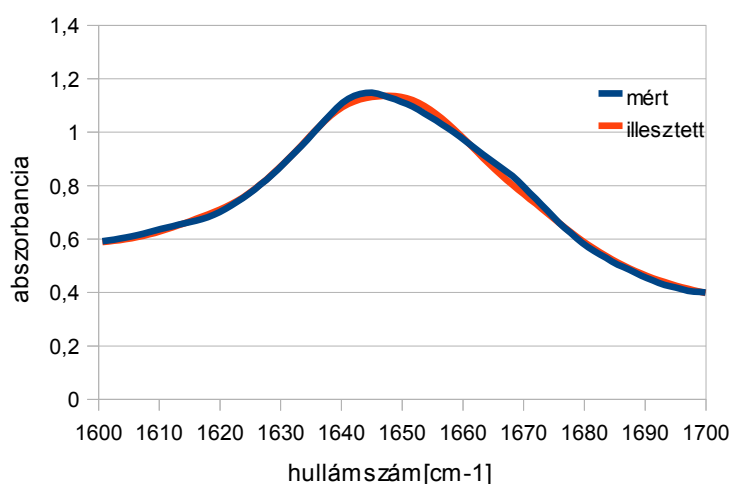
4.2.1 ábra 1YMB szerkezet egyes aminosavakhoz rendelt Lorentz-görbéi



4.2.2 ábra 1YMB szerkezet illesztése az IR spektrumra



4.2.3 ábra 4PTI szerkezet egyes aminosavakhoz rendelt Lorentz-görbéi



4.2.4 ábra 4PTI szerkezet illesztése az IR spektrumra

Illesztések közös paraméterekkel

Megpróbáltam egy közös illesztéssel, ugyanazon c_1 , c_2 , c_3 paraméterekkel megkapni a mioglobint és a BPTI(marha tripszin inhibitor) az 1YMB és 4PTI spektrumát. Fix szélességű Lorentz görbéket alkalmaztam. Egy új c_3 paramétert vettem hozzá X_c -hez, amely a C-O-N szöget veszi figyelembe. A Lorentz-görbék pozícióját a következőképpen számoltam:

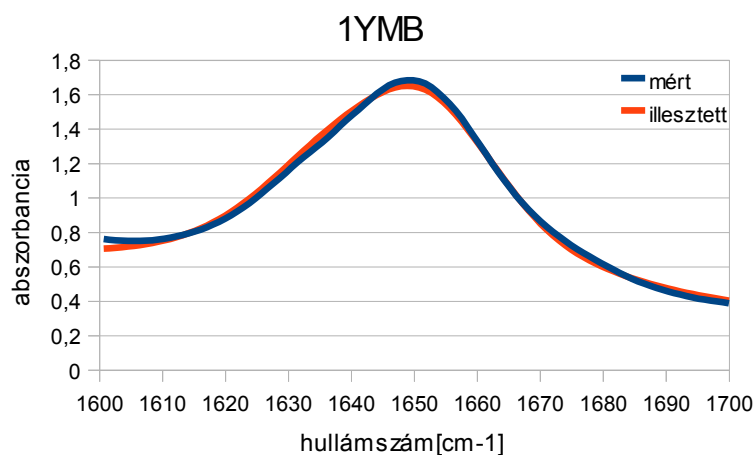
$$X_c = X_{c0} + c_1 r_{C-O} + c_2 \frac{1}{r_{O-N}} + c_3 \frac{\cos(\varphi_{C-O-N})}{r_{C-N}} . \text{ A harmadik tag akkor adja a legnagyobb}$$

járulékot, amikor C, O és N atomok egy egyenesbe esnek. Az illesztés során a program c_1 , c_2 , c_3 -at, az amplitúdókat és X_{c0} -t változtatta χ^2 minimalizálása során. Először a röntgen-kristallográfiával

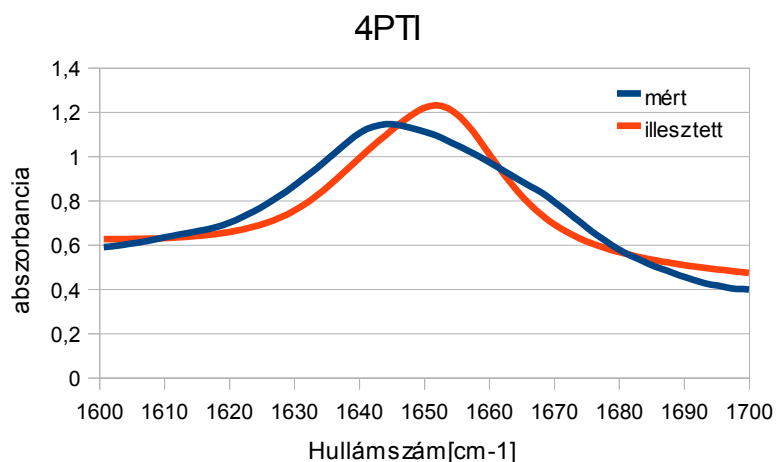
meghatározott szerkezeteket vettem alapul. 4.2.2-es táblázatban látható, hogy az illesztés 1YMB-re jobban sikerült. Az illesztések során sikerült a fordított esetet is elérnem, amikor 4PTI-re illeszkedett jobban, az illesztések eredménye attól is függött, hogy milyen kezdeti értékekről indítottam a nemlineáris megoldót, arra a spektrumra illeszkedett jobban a végeredmény amelyik a kiindulási állapotban is jobban illeszkedett. Arra a következtetésre jutottam, hogy nincsenek jó közös paraméterek amivel illeszteni lehetne, vagy az egyik, vagy a másik illeszkedik jobban.

közös paraméterek				
c_1	c_2	c_3	Xc_0	
-396,051	-44,012	-19,235	2149,559	
1YMB				
w	A	tengelymetszet	meredekség	χ^2
10,000	0,012	0,607	-0,003	0,059
4PTI				
w	A	tengelymetszet	meredekség	χ^2
10,000	0,017	0,592	-0,002	0,597

4.2.2. táblázat 1YMB és 4PTI szerkezet illesztésére kapott paraméterek



4.2.5 ábra 1YMB szerkezet illesztése az IR spektrumra



4.2.6 ábra 4PTI szerkezet illesztése az IR spektrumra

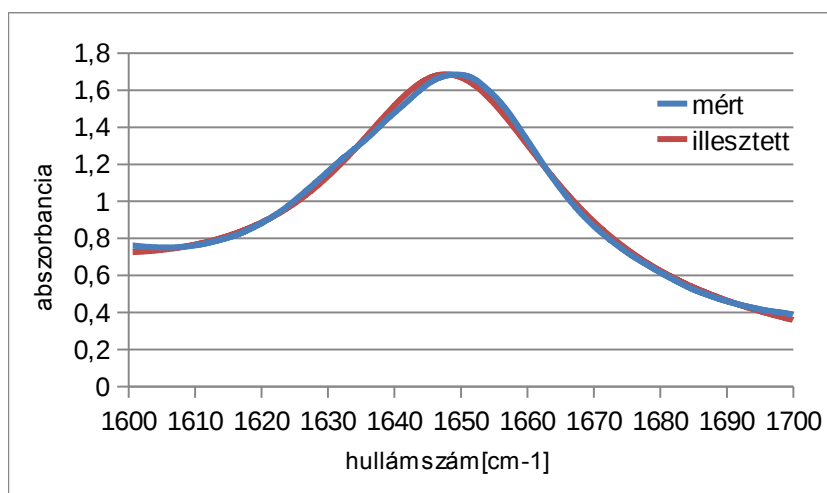
Ezután az NMR-el meghatározott szerkezetek adataival próbáltam illeszteni a mioglobin és BPTI spektrumait.

A mioglobin NMR-el meghatározott 1YMF szerkezetének 5 modelljére is végeztem egy közös illesztést. 4.2.3-as táblázatban látható, hogy vannak modellek amelyek jobban eltérnek a többitől, ilyen pl az 5-ös számú.

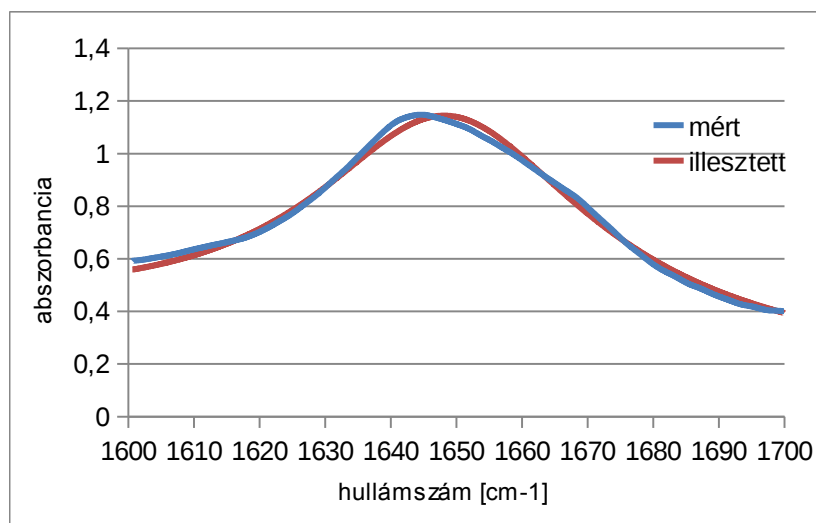
1YMF	modell 1	modell 2	modell 5	modell 9	modell 12
χ^2	0,055	0,055	0,109	0,0756	0,055

4.2.3 táblázat 1YMF különböző szerkezeti modelljeire kapott illesztések χ^2 értékei

A különböző szerkezetek közös illesztésénél kétszer illesztettem, először c_1 , c_2 , c_3 -at, az amplitúdókat és X_{c_0} -t változtattam, majd utána a görbék szélességét. Így kaptam végül az eredményeket. Így sem sikerültek az illesztések tökéletesre, a görbék kiszélesedtek χ^2 értéke viszont csökkent.



4.2.7 ábra 1YMF szerkezet illesztése az IR spektrumra

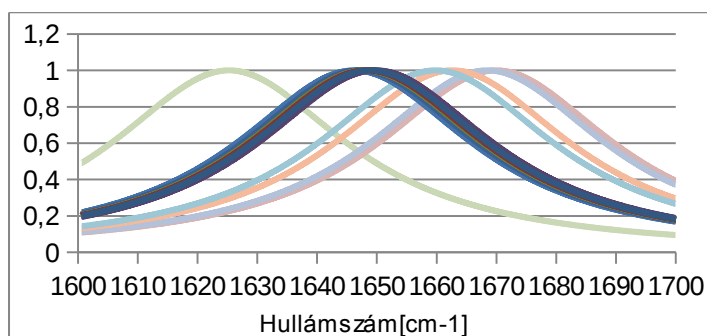


4.2.8 ábra 1PIT szerkezet illesztése az IR spektrumra

közös illesztési paraméterek				
c1	c2	c3	$X_{c,0}$	
-199,058	-342,578	-606,415	1893,025	
1YMF				
A	tengelymetszet	meredekség	w	χ^2
0,018	0,404	-0,002	20,126	0,046
1PIT				
A	tengelymetszet	meredekség	w	χ^2
0,015	0,384	-0,002	24,191	0,042

4.2.4. táblázat 1YMF és 1PIT NMR szerkezet közös illesztésére kapott paraméterek

A 4.2.9-es ábrán látható, hogy kevésbé váltak szét az egyes Lorentz-görbék. Néhány görbe középpontja 1675cm^{-1} -es hely köré esik, ami az irodalom szerint szintén a β -szerkezethez köthető. A rendezetlen szerkezeti elemeknek 1645cm^{-1} -nél van a csúcsa, a görbék többsége oda csoportosul. A szélesebb görbék miatt egy-egy aminosav több helyre is ad járulékot. Előzetesen azt vártam volna, hogy a görbék keskenyebbek és jobban szétválnak, hogy mindegyik elem a neki megfelelő helyen adjon járulékot. Amint az előzőleg a 4.2.4-es ábrán is látható 4PIT esetében. χ^2 értékére így sikerült mégis a legjobb eredményt elérnem.



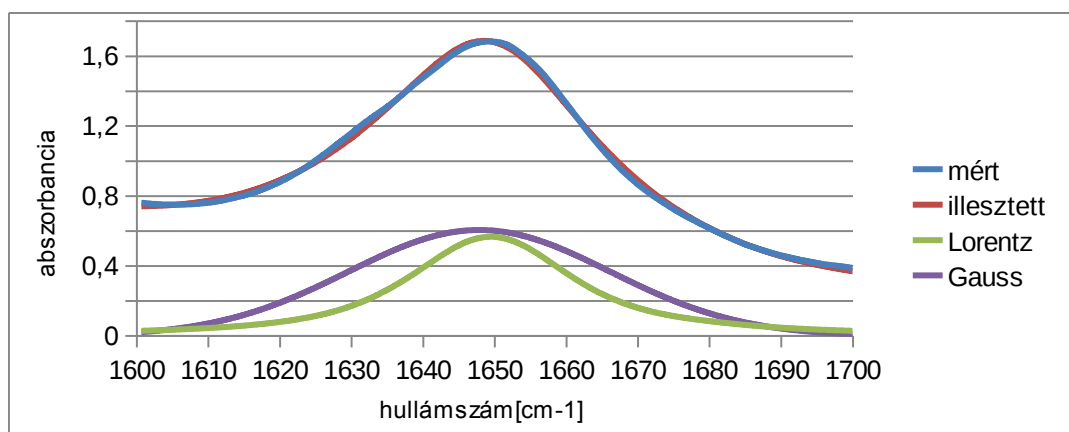
4.2.9 ábra 1PIT szerkezet egyes aminosavakhoz rendelt Lorentz-görbéi

Az 1YMF és 1PIT szerkezetre végeztem egy olyan illesztést is, ahol a rendezetlen szerkezeti elemeket lecseréltem egy a rendezetlen elemek Lorentz-görbéinek területeinek összegével

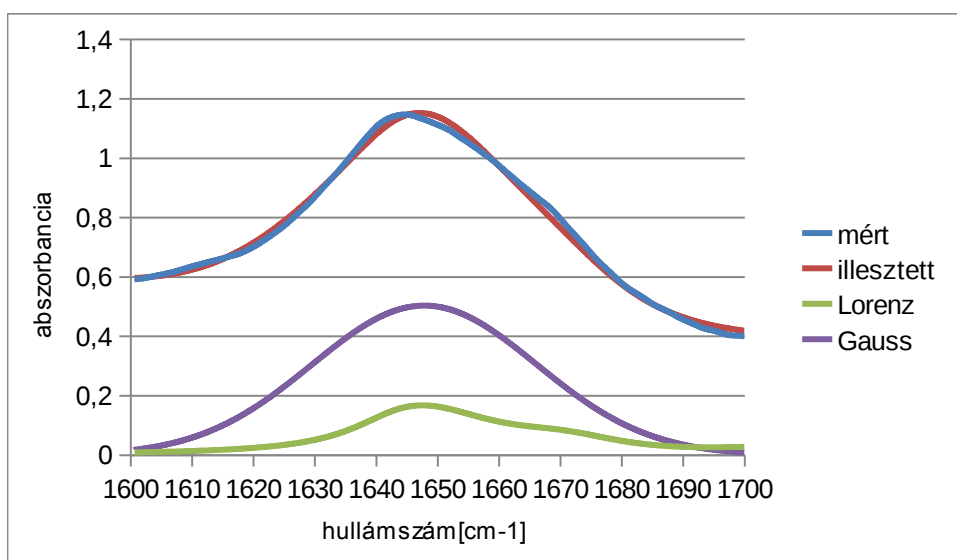
megegyező területű Gauss-görbével $\frac{A_g}{\sigma\sqrt{2\pi}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}$, amit a rendezetlen elemeknek megfelelő

1645cm^{-1} megfelelő helyre pozicionáltam. Azért helyettesítettem Gauss-görbével a rendezetlen elemeket, mert egyrészt nem tudtam hozzájuk O-N távolságokat és kötésszögeket rendelni (Nem tartozik hozzájuk hidrogén kötés), másrészt 1.4.1-es diagramon is láthatjuk, hogy a rendezetlen C=O távolságadatok. szimmetrikusan helyezkednek el az átlagos rendezetlen C=O távolság körül. Az illesztés során c_1 , c_2 , c_3 -at, az amplitúdókat és $X_{c,0}$ -t változtattam. A Gauss-görbéket közösen pozicionáltam és a félértékszélességüket is, csak az amplitúdóikban térnek el. A 4.2.10- és 4.2.11-es ábrákon láthatjuk, hogy a zöld színű görbe(Lorentz-görbék összege) az 1YMF szerkezet esetén szimmetrikus és a csúcsa 1650 körül található, 1PIT esetén asszimmetrikus, két elkülönülő csúcs látszik az egyik 1648 környékén a másik 1670 környékén ami megfelel az α - és β - szerkezeteknek. Így a Lorentz-görbék kiszélesedése nélkül is sikerült jó illesztést elérnem. Ezzel a módszerrel sem sikerült a kellő pontosságot elérnem.

1YMF						1PIT					
A	σ	x_0	χ^2				χ^2	A	σ	x_0	
2669,000	18,284	1647,822	0,021				0,024	1350,200	18,284	1647,822	
c1	c2	c3	X_0	w	A	tengelymetszet	meredekség	A	tengelymetszet	meredekség	w
-894,181	439,375	454,142	2697,443	10,000	0,010	0,693	-0,004	0,017	0,571	-0,002	10,000



4.2.10. ábra 1YMF szerkezet illesztése IR spektrumra Lorentz - és Gauss-görbék összegeként Az alsó két görbe közül a lila a Gauss-görbe, a zöld a Lorentz-görbék összegéből kirajzolódó görbe.



4.2.11. ábra 1PIT szerkezet illesztése IR spektrumra Lorentz - és Gauss-görbék összegeként Az alsó két görbe közül a lila a Gauss-görbe, a zöld a Lorentz-görbék összegéből kirajzolódó görbe.

Összefoglalás

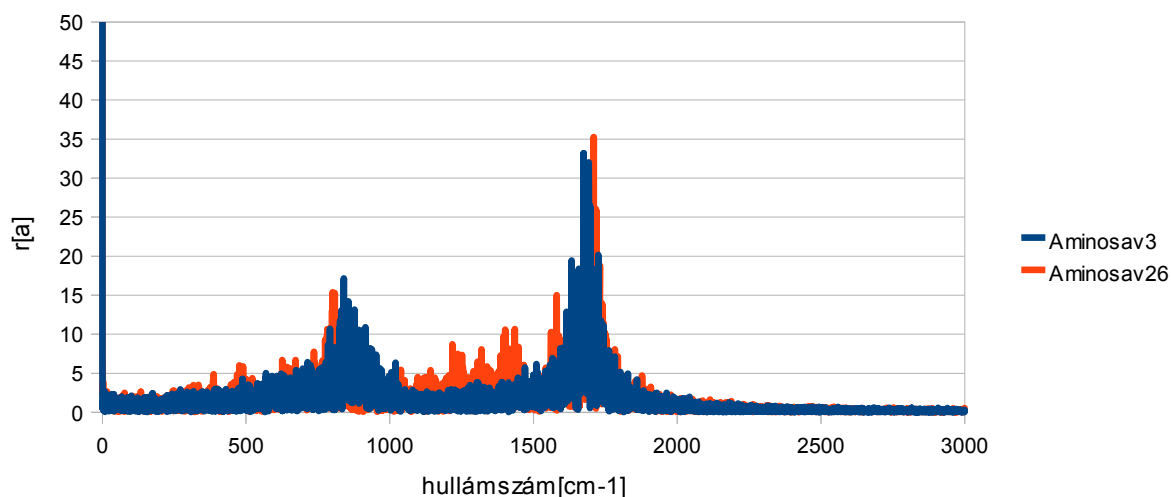
Látható, hogy az infravörös spektrum nem illeszthető meg kellő pontossággal a megfogalmazott hipotézis alapján, azaz a rezgési frekvencia nem egyértelmű függvénye a PDB adatbázisban található C-O, O-N távolságoknak és a H-híd kötősszögeknek.

4.3 Különböző fehérjeszerkezetre jellemző amid I rezgések meghatározása

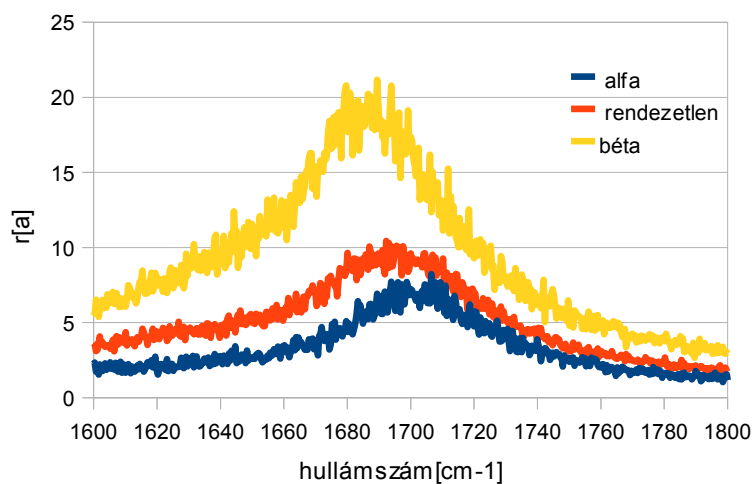
Molekuláris dinamikai szerkezeti adatokból számoltam ki az amid I rezgés értékét α - és β -szerkezeti elemekre. A dihidrofolát reduktáz (DHFR) enzim trajektóriájából meghatároztam az amid I rezgési frekvenciákat Fourier transzformáció segítségével. A DHFR enzim kulcsszerepet játszik herpesz vírusok, például a HHV-8 nukleotid anyagszere mechanizmusában. A másodlagos szerkezetében α - és β -szerkezetű részek is találhatóak (4.3.1 ábra). Egy 16384 egymást $5 \cdot 10^{-15}$ s-ként követő szerkezeti adatsokaságból, kiszámítottam minden aminosavban a C=O kötésben résztvevő atomok távolságát az adott időpillanatban, majd az így kapott $R_{C=O}(t)$ adatsorokat Fourier transzformáltam. A Fourier transzformáltakból leolvasható az egyes aminosavakhoz tartozó rezgések frekvenciái. Minden aminosavhoz kaptam egy spektrumot. Az 1600cm^{-1} - 1700cm^{-1} közötti tartományon lévő adatokat szétválogattam a másodlagos szerkezeti elemeknek megfelelően. Az α , β , valamint rendezetlen szerkezetekhez tartozó aminosavak spektrumait külön külön átlagoltam és így az egyes szerkezetekre jellemző spektrumokat kaptam (4.3.3 ábra). Átlagolt spektrumra Lorentz görbét illesztettem, melynek pozíciója megadja a keresett amid I komponensek helyét. Kiszámítottam a lépésközt a Fourier transzformált adatokhoz, vettem $R_{C=O}(t)$ adatsor Δt lépésközének reciprokát, amit leosztottam az adatok számával (16384) és így Hz-ben kaptam meg az új lépésközt, amit átszámítottam inverz cm-re ($1\text{cm}^{-1} \approx 30\text{GHz}$), így $\Delta k = 0,4069\text{ cm}^{-1}$ adódott. A kapott eredményeket 4.3.2-, 4.3.3 ábrán ábrázoltam és a 4.3.1-es táblázatban összegeztem.



4.3.1. ábra DHFR enzim Viewerlite programmal megjelenített képe



4.3.2. ábra DHFR enzim 3.(béta szerkezetben lévő) és 26(alfa-hélixben lévő). aminosavhoz tartozó C-O koordináták Fourier transzformált spektruma

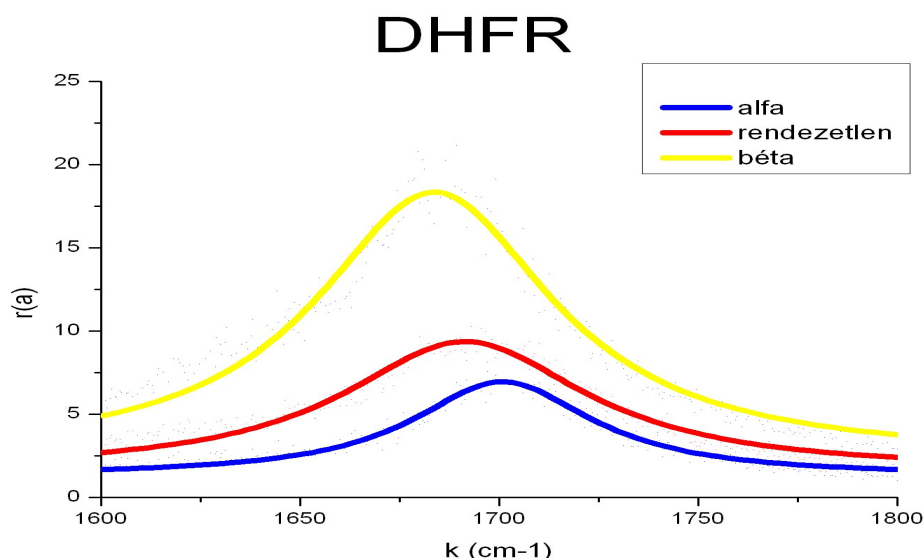


4.3.3 ábra DHFR enzim másodlagos szerkezeti elemeinek számolt átlagos Fourier-spektruma [1600cm⁻¹-1800cm⁻¹]

A 4.3.1-es táblázatban láthatóak a kapott eredmények. A Fourier transzformált majd átlagolt adatpontokra jól illeszkednek az illesztett görbék. A rezgésekre kapott értékek 2,8%-al térnek el az irodalomban fellelhető értékektől. Ez mindhárom számolt szerkezeti egységnél jelentkező azonos nagyságú szisztematikus hiba.

Illesztés után kapott értékek			
Illesztett görbe egyenlete: $y = y_0 + (2 \cdot A / \pi) \cdot (w / (4 \cdot (x - x_c)^2 + w^2))$			
	α -hélix	β -redő	rendezetlen
R^2	0,942	0,947	0,932
X_c	$1700,867\text{cm}^{-1} \pm 0,276\text{cm}^{-1}$	$1683,706\text{cm}^{-1} \pm 0,302\text{cm}^{-1}$	$1691,467\text{cm}^{-1} \pm 0,345\text{cm}^{-1}$
$X_{c \text{ korrigált}}$	$1653,867\text{cm}^{-1} \pm 0,276\text{cm}^{-1}$	$1636,706\text{cm}^{-1} \pm 0,302\text{cm}^{-1}$	$1644,467\text{cm}^{-1} \pm 0,345\text{cm}^{-1}$
irodalmi	1654cm^{-1}	1624cm^{-1} - 1637cm^{-1}	1645cm^{-1}

4.3.1 táblázat



4.3.4 ábra DHFR enzim másodlagos szerkezeti elemeinek számolt átlagos Fourier-spektrumára illesztett Lorentz-görbék

5. Összefoglalás

A különböző módszerekkel meghatározott fehérjeszerkezetekben jól látható különbség van a C-O és O-N távolságok között, a módszertől függően. Mindegyik esetben a távolságok a megfelelő nagyságrendű tartományba esnek (O-N távolságok 2,4 Å-4 Å közé, C-O 1,1 Å-1,3 Å közé), az alfa és beta szerkezetekhez kapcsolódó kötéstávolságok között is felfedezhető különbség, azonban a szerkezeti adatok nem elég pontosak ahhoz, hogy az infravörös spektrum kellő pontossággal illeszthető legyen a megfogalmazott hipotézis alapján, azaz a rezgési frekvencia nem egyértelmű függvénye a PDB adatbázisban található C-O, O-N távolságoknak és a H-híd kötőszögeknek. A molekuláris dinamikai számolásból sikerült megkapni a kísérleti spektrumokra jellemző rezgéseket az α - és β -szerkezetekre, egy 2,8%-os korrekciós faktortól eltekintve.

5. Köszönetnyilvánítás

Elsősorban szeretnék köszönetet mondani témavezetőmnek, Dr Smeller Lászlónak, hogy munkám során betekintést nyerhettem a fehérjék szerkezetekkel kapcsolatos kutatásokba, továbbá hogy elméleti tanácsaival segítette munkám elkészülését. Szeretnék köszönetet mondani Dr Balog Erikának, aki a DHFR szerkezeti adatokat a rendelkezésemre bocsájtotta. Ezenkívül köszönöm a barátaimnak és családomnak a türelmet.

7.Hivatkozások

- [1] Gábor Horváth, Judit Illényi, Levente Pusztay and Kálmán Simon RELATIONSHIP BETWEEN THE IR FREQUENCIES AND BOND LENGTHS IN CARBONYL GROUPS *Acta Chim. Hung.* 124, 1987
- [2] Jagdeesh Bandekar Amide modes and protein conformation *Biochemica et Biophysica Acta*, 1120(1992) 123-143
- [3] Havancsák Károly Fizikai mérési módszerek
- [4] Havancsák Károly Kísérleti módszerek I-II.
- [5] Haris, P. I., Chapman, D. (1988) Fourier-transform infrared-spectra of the polypeptide alamethicin and a possible structural similarity with bacteriorhodopsin, *Biochim. Biophys. Acta* 943, 375-380.
- [6] Ismail, A. A., Mantsch, H. H., Wong, P. T. T (1992) Aggregation of Chymotrypsinogen: Portrait by FT-IR Spectroscopy *Biochim. Biophys. Acta* 1121, 83-188.
- [7] Krimm, S., Bandekar J. (1986) Vibrational spectroscopy and conformation of peptides, polypeptides and proteins *Adv. Protein Chem.* 3, 181-364.
- [8] Miyazawa, T., Shimanouchi, T., Mizushima, S. I (1958) Normal Vibration of N-Methylacetamide *J. Chem Phys* 29, 611-615.
- [9] Smeller László MTA Doktori értekezés: A fehérjék konformációs és dinamikai tulajdonságai. Új eredmények nagy nyomással kombinált infravörös és fluoreszcencia spektroszkópiai módszerekkel
- [10] Susi, H., Byler, D. M., (1986) Resolution-enhanced Fourier-transform infrared spectroscopy of enzymes *Methods Enzymol.* 130, 290-311
- [11] Tőke Orsolya Mágneses magrezonancia (NMR) Spektroszkópia a fehérjekutatásban 2007
- [12] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery Numerical Recipes in C The Art of Scientific Computing Second Edition CAMBRIDGE UNIVERSITY PRESS
- [13] WOLFGANG KABSCH and CHRISTIAN SANDER Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features Biopolymers, Vol. 22, 2577-2637 (1983)