

DOKTORI (PhD) ÉRTEKEZÉS

**A venezuelai ló-láz encephalitis vírus
cisztein proteázának szerkezet-alapú vizsgálata**

Hoffka Gyula

Témavezető: Dr. Mótyán János András



DEBRECENI EGYETEM

MOLEKULÁRIS SEJT- ÉS IMMUNBIOLÓGIA DOKTORI ISKOLA

Debrecen, 2024.

TARTALOMJEGYZÉK

1. RÖVIDÍTÉSEK JEGYZÉKE.....	4
2. BEVEZETÉS.....	5
3. IRODALMI ÁTTEKINTÉS	7
3.1. A venezuelai ló-láz encephalitis vírus (VEEV).....	7
3.1.1. A VEEV, mint kórokozó vírus.....	7
3.1.2. Az nsP2pro fehérje.....	9
3.1.3. Az alfavírus nsP2pro fehérjék kristályszerkezete	13
3.2. Kristályszerkezet-meghatározás	15
3.3. Molekulamechanika	17
3.3.1. Történeti összefoglalás.....	18
3.3.2. Erőterek.....	18
3.3.3. Előkészítés	20
3.3.4. Szimulációk lefuttatása	24
3.3.5. Szimulációk analízise.....	27
3.4. A szimulációs szoftverek alkalmazása	29
3.4.1. Molekuladinamikai szoftverek.....	29
3.4.2. Szuperszámítógépes környezet	30
4. CÉLKITŰZÉSEK	31
5. ANYAGOK ÉS MÓDSZEREK	32
5.1. A VEEV nsP2pro szubsztrát-specifitásának vizsgálata	32
5.1.1. Az enzim-szubsztrát komplex előkészítése.....	32
5.1.2. Molekuladinamikai szimulációk	33
5.1.3. Az enzim-szubsztrát kölcsönhatások vizsgálata	34
5.2. A VEEV nsP2pro kristályszerkezetének meghatározása	35
5.3. VEEV nsP2pro kristályszerkezetek molekuladinamikai vizsgálata	35
5.3.1. A szerkezetek előkészítése.....	35
5.3.2. Molekuladinamikai szimulációk	36
5.3.3. A szerkezetek analízise	37
6. EREDMÉNYEK.....	39
6.1. A VEEV nsP2pro szubsztrát-specifitásának vizsgálata	39
6.1.1. Az enzim-szubsztrát komplex felépítése.....	39
6.1.2. Az enzim-szubsztrát komplex számításhoz vizsgálata.....	40

6.1.3. Az enzim-szubsztrát interakciók vizsgálata	40
6.2. A VEEV nsP2pro kristályszerkezetének meghatározása és elemzése	45
6.3. A VEEV nsP2pro kristályszerkezetek molekuladinamikai vizsgálata	48
6.3.1. A molekuladinamikai szimulációk ellenőrzése.....	48
6.3.2. A VEEV nsP2pro N-terminálisának interakciói	50
6.3.3. A β -hajtű és Asn545 aminosav	53
6.3.4. Az Asn545 és Val476 közötti kölcsönhatás.....	54
6.3.5. Az aktív hely kölcsönhatásai	56
6.3.6. Az aktív hely hozzáférhetősége	59
6.3.7. A katalitikus diád stabilitása	60
6.3.8. A K741A és K767A felszíni mutációk hatása	62
7. MEGBESZÉLÉS	65
8. ÖSSZEFOGLALÁS	73
9. SUMMARY.....	74
10. IRODALOMJEGYZÉK	75
11. TÁRGYSZAVAK.....	85
12. KÖSZÖNETNYILVÁNÍTÁS	86
13. PUBLIKÁCIÓS LISTA.....	88
14. FÜGGELÉK	90

1. RÖVIDÍTÉSEK JEGYZÉKE

CHIKV	Chikungunya vírus
EEEV	Keleti ló-láz encephalitis vírus (<i>Eastern equine encephalitis virus</i>)
GPU	grafikus processzoregység (<i>Graphical Processing Unit</i>)
HPC	„szuperszámítógép” (<i>High Performance Computer</i>)
krio-EM	krio-elektronmikroszkópia
MD	molekuladinamika
NPT	izotermális-izobárikus sokaság, állandó anyagmennyiség, nyomás és hőmérséklet (<i>isothermal-isobaric ensemble, Number of particles, Pressure, Temperature</i>)
nsP2	nem-szerkezeti fehérje 2
nsP2pro	nem-szerkezeti fehérje 2 proteáz
NVT	kanonikus sokaság, állandó anyagmennyiség, térfogat és hőmérséklet (<i>canonical ensemble, Number of particles, Volume, Temperature</i>)
P1, P2, P3, P4, P5	szubsztrát aminosavrésze (Schechter és Berger [1967] alapján)
PDB/pdb	fehérje adatbank (<i>Protein Data Bank</i>), ill. fájlformátum
PDB ^(MD-) aktív/öninaktivált	egy adott PDB azonosítójú szerkezet, amely VEEV nsP2pro aktív vagy öninaktivált konformerét reprezentálja (az MD előtag a molekuladinamikai szimulációval vizsgált szerkezetekre utal)
pK _a	savi disszociációs állandó
RMSD/r.m.s.d.	négyzetes közép eltérés (<i>Root mean square deviation</i>)
ORF	nyitott leolvasási keret (<i>Open Reading Frame</i>)
S1, S2, S3, S4, S5	az enzim szubsztrátkötő alhelyei (Schechter és Berger [1967] alapján)
SAM MTáz	S-adenozil-L-metionin (SAM)-függő RNS metiltranszferáz
SFV	Semliki erdei vírus (<i>Semliki forest virus</i>)
SINV	Sindbis vírus
TIP3P	<i>Transferable Intermolecular Potential with 3 Points</i>
VEEV	venezuelai ló-láz encephalitis vírus (<i>Venezuelan equine encephalitis virus</i>)
WEEV	Nyugati ló-láz encephalitis vírus (<i>Western equine encephalitis virus</i>)

2. BEVEZETÉS

A vírusos eredetű betegségek korunk egyik legnagyobb kihívását jelentik, amiktől még a modern orvostudomány eszköztára sem tud teljes mértékben megvédeni minket. Míg a bakteriális eredetű megbetegedésekkel szemben számos antibiotikum áll rendelkezésre, a vírusokkal szembeni védekezés jelentősebb kihívásokat tartogat. Ahogy az elmúlt évek során valamennyien megtapasztalhattuk, egy új, veszélyes vírus, mint például a súlyos akut légzőszervi szindróma koronavírus 2 (SARS-CoV-2) koronavírus megjelenése akár globális katasztrófát okozhat és komoly társadalmi, gazdasági és egészségügyi következményekkel járhat. Azonban számos más, már ismert vírus is súlyos veszélyforrást hordozhat magában.

Ilyen többek között az alfavírusok közé tartozó venezuelai ló-láz encephalitis vírus (VEEV) is, amely elsősorban állati kórokozóként ismert, de emberi megbetegedést is képes okozni. A fertőzés következtében maradandó károsodással járó súlyos agyvelőgyulladás alakulhat ki, és a fertőzés súlyos esetekben halálos következményekkel is járhat. A VEEV vírus az általa jelentett egészségügyi kockázat miatt nem csak az antivirális szerek, hanem a vakcinák fejlesztésének is az egyik célpontja. Ezen túl, kutatások irányultak a virális proteáz lehetséges biotechnológiai és molekuláris biológiai területen való alkalmazhatóságának vizsgálatára is. Ahhoz, hogy a vírusokkal szemben sikeresen védekezni tudjunk, mindenképp szükséges, hogy molekuláris szinten megértsük működésüket. Ehhez nem elég ismernünk a virális genomot, szükséges a virális fehérjék szerkezetének minél behatóbb ismerete is. Azonban a fehérjék kísérletes módszerekkel meghatározott szerkezete és az ezekre épülő homológia modellek, valamint a mesterséges intelligencia-alapú módszerekkel végzett predikciók sem nyújtanak minden esetben elegendő információt a fiziológias körülmények között valóban megjelenő szerkezetekről, vagy a ligandumokkal kialakított kölcsönhatásairól. A valóság minél pontosabb leírásához a számítási fehérjemodellezés eszköztárához kell nyúlnunk.

A számítástechnika robbanásszerű fejlődése lehetővé tette, hogy ma már olyan számítási vizsgálatokat folytassunk le, amelyekre korábban még csak lehetőség sem volt. Mindez számos tudományterület, köztük a biológiai rendszerek kutatásának szempontjából is különösen lényeges. A számítási kémia és a bioinformatika módszereinek alkalmazásával olyan kutatásokat végezhetünk, amelyek vagy az elméleti megközelítések, vagy a hatalmas adatmennyiség miatt megkövetelik a számítástechnika alkalmazását. A fehérjemolekulák vizsgálatának szempontjából kiemelkedő jelentőségű a molekuladinamikai módszerek alkalmazása. A klasszikus fizikai megközelítésen alapuló módszerrel atomi sokaságok időbeli

viselkedését modellezhetjük. Olyan, kísérleti módszerekkel csak nehezen vagy egyáltalán nem megfigyelhető tulajdonságokról szerezhethetünk vele információkat, mint az egyes oldalláncok különböző konformációinak viselkedése, hidrogénhidás kölcsönhatások megjelenése, vagy éppen az egyes konformációs állapotokat stabilizáló kölcsönhatási hálózatok.

Kutatócsoportunk, a Dr. Tózsér József professzor úr által alapított és vezetett Retrovirális Biokémiai Kutató Laboratórium több évtizedes múltja tekint vissza a virális eredetű fehérjék vizsgálatában, melyben kiemelt szerepet kaptak a proteolitikus enzimek, köztük a VEEV nem-szerkezeti fehérje 2 proteáza (nsP2pro) is. Kutatócsoportunk tagjai - az Amerikai Egyesült Államokbeli National Cancer Institute (Frederick) krisztallográfiai laboratóriumának munkatársaival együttműködésben - korábban már vizsgálták a VEEV nsP2pro proteázát. A kutatások elsősorban a VEEV nsP2pro enzimátikus sajátságainak vizsgálatára irányultak és céljuk volt annak meghatározása is, hogy az nsP2pro alkalmas eszköz lehet-e fúziós címkék proteolitikus eltávolítására. A korábbi kutatások folytatásaként célunk volt tanulmányozni a VEEV élelciklusa szempontjából kiemelt fontosságú nsP2pro proteáz működésének molekuláris hátterét, a szerkezeti sajátságok kísérletes és számítógépes módszerekkel történő vizsgálatával.

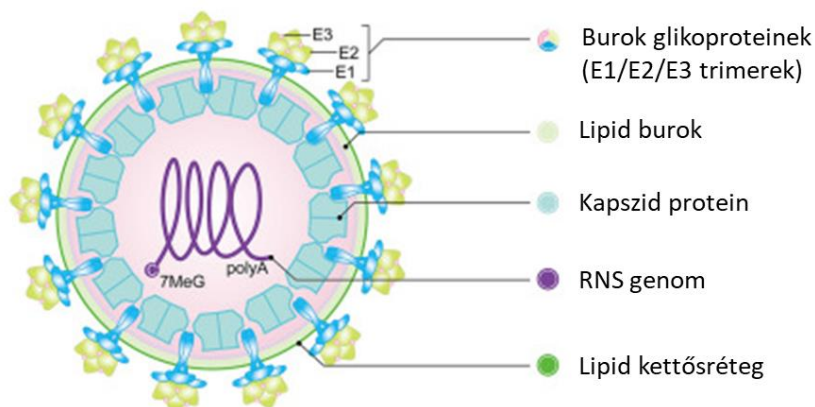
Munkánk során - a kísérletes vizsgálatok eredményének értelmezése érdekében - a molekuladinamikai modellezés eszköztárát felhasználva tártuk fel az nsP2pro specificitását meghatározó enzim-szubsztrát kölcsönhatásokat. Továbbá, a szakmai együttműködés keretén belül meghatároztuk az nsP2pro kristályszerkezetét, melyet molekuladinamikai szimulációk segítségével elemeztünk. Az *in silico* elemzések eredményeit felhasználva feltérképeztük az enzim aktív és inaktív konformációs állapotai közötti hasonlóságokat és különbségeket, valamint az öninaktivált konformer kialakulásáért felelős kölcsönhatások és a dinamikai tulajdonságok alapján meghatároztuk az enzim stabilitását meghatározó szerkezeti sajátságokat.

3. IRODALMI ÁTTEKINTÉS

3.1. A venezuelai ló-láz encephalitis vírus (VEEV)

3.1.1. A VEEV, mint kórokozó vírus

A venezuelai ló-láz encephalitis vírus (*Venezuelan equine encephalitis virus*, VEEV) a vírusok IV(+) csoportjába tartozik, rendszertanilag a *Togaviridae* családba tartozó alfavírusok nemzetségébe, számos más vírussal együtt. Az alfavírusok származásukat tekintve két csoportra oszthatók, amelyek néhány ezer évvel ezelőtt váltak el egymástól az evolúció során. A VEEV vírus az ún. újvilági (*New World*) alfavírusok közé tartozik, a Keleti ló-láz encephalitis vírussal (EEEV) és a Nyugati ló-láz encephalitis vírussal (WEEV) együtt. A másik csoportba az ún. óvilági vírusok (*Old World* alfavírusok) tartoznak, mint például a Sinbdis vírus (SINV), a Semliki erdei vírus (SFV) és a Chikungunya vírus (CHIKV) (**Powers és mtsai, 2001; Zacks és mtsai, 2010**). Az alfavírusok általános jellemzője, hogy egyszálú, pozitív irányítottágú RNS genommal (ss(+))RNS rendelkeznek (**Rupp és mtsai, 2015**), gömb alakúak, és membránjukban glikoproteineket tartalmaznak, amelyek a receptor-mediált felismeréshez nélkülözhetetlenek (**1. ábra**).



1. ábra. A VEEV vírus felépítése. Megfigyelhető az alfavírusokra jellemző gömb alakú forma és az egyszálú RNS genetikai állomány. Az ábrán jelölve szerepelnek továbbá a burok glikoproteinjei (E1-E3), a lipid burok, a kapszid protein és a lipid kettősréteg. Az ábra Han és mtsai. közleménye (**2023**) alapján készült.

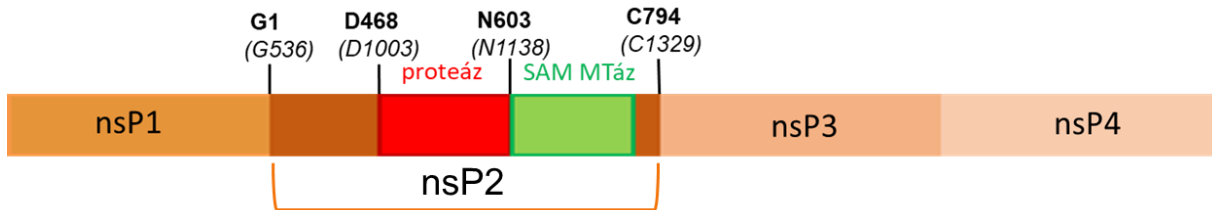
Általánosságban elmondható, hogy az újvilági alfavírusok általi fertőzés embereken agyvelőgyulladást (encephalitis) okozhat, amelynek tünetei az enyhétől a komoly panaszokig terjedhetnek, a fertőzés mortalitása pedig vírusonként változó lehet. A VEEV legjellemzőbb elterjedési területei Közép- és Dél-Amerikában találhatóak meg, terjedéséért nagymértékben a szúnyogok tehetők felelőssé, amelyek madarakat és emlősöket is képesek megfertőzni

csípésükkel. A VEEV fertőzés különösen veszélyes lovak esetében, a fertőzött egyedek 20-80%-a elpusztul (**Deardor és mtsai, 2009**). A VEEV vírus általi fertőzés embereknél általában enyhe, influenzaszerű tüneteket okoz, és az immunrendszer képes leküzdeni, azonban az esetek körülbelül 1%-a halálos kimenetelű, amelyért vírusfertőzés által okozott súlyos agyvelőgyulladás tehető felelőssé (**Garmashova és mtsai, 2006; Johnson és mtsai, 1974**). Az esetek további kb. 14%-ában a fertőzés változó súlyosságú neurológiai tüneteket okoz. További érdekesség, hogy a vírus aeroszol formában belélegezve képes bejutni az agyba (a szaglóiidegekben található neuronokat megfertőzve). Mivel virális részecskéi nagyon ellenállóak a szárítással szemben, így könnyen előállítható aeroszolizált vírus a fagyasztva szárítás módszerével. Mindezek miatt a VEEV-re potenciális biológiai fegyverként tekintettek sőt, jelentések szerint az Amerikai Egyesült Államok és a volt Szovjetunió is kísérletezett vele (**Bronze és mtsai, 2002**), és a biológiai terrorizmus szempontjából is potenciális veszélyforrásnak is tekinthető (**Meyer és mtsai, 2008**).

A VEEV fertőzés kezelésére jelenleg nincs az amerikai Élelmiszer- és Gyógyszer-Engedélyeztetési Hivatal (*Food and Drug Administration*, FDA) által elfogadott antivirális gyógyszer (**Carey és mtsai, 2019**), ezért a fertőzést megelőzni képes vakcina fejlesztésére irányuló kutatások jelenleg is folyamatban vannak (**Stromberg és mtsai, 2020; Han és mtsai, 2023**). Az elmúlt években egy vírus-szerű részecskéken alapuló vakcinával már klinikai vizsgálatokat is végeztek (**Coates és mtsai, 2022**), továbbá, egy DNS-alapú vakcina hatását makákók esetében igazolták (**Suschak és mtsai, 2022**). Az alfavírusok elleni védekezés lehetőségeinek vizsgálata általánosságban is számos kutatás tárgyát képezi (**Abdelnabi és Delang, 2020**), amit jól mutat, hogy napjainkban is zajlanak hatékony vakcina kifejlesztésére irányuló kutatások, így például egy, az újvilági alfavírusok (VEEV, EEEV és WEEV) ellen tervezett vakcina fejlesztése jelenleg a klinikai vizsgálatok I. fázisában van (**Fierro és mtsai, 2024**). Emellett számos, az újvilági alfavírusokkal szemben alkalmazható kismolekulás inhibitor is vizsgálják (**Ogorek és Golden, 2022**). Többek között a hatékony megelőzési és kezelési lehetőségek hiánya miatt tekintettek a VEEV-re potenciális biológiai fegyverként, de ezek a lehetőségek sajnos még napjainkban sem állnak rendelkezésre. Érdeemes megemlíteni a vírus potenciális terjedésében a klímaváltozás szerepét, ami miatt a vírus olyan helyeken is felbukkanhat, ahol korábban még nem fordult elő. A globális utazási hálózatok segítségével a különféle vírusok - mint például a SARS-CoV-2 és humán immundeficiencia-vírusok (HIV) is - a világ minden pontjára eljuthatnak és okozhatnak egészségügyi problémákat. Mindezeket figyelembe véve egy VEEV ellen kifejlesztendő antivirális gyógyszer vagy vakcina nagy fontossággal bírna (**Sharma és mtsai, 2019; Guzmán-Terán és mtsai, 2020**).

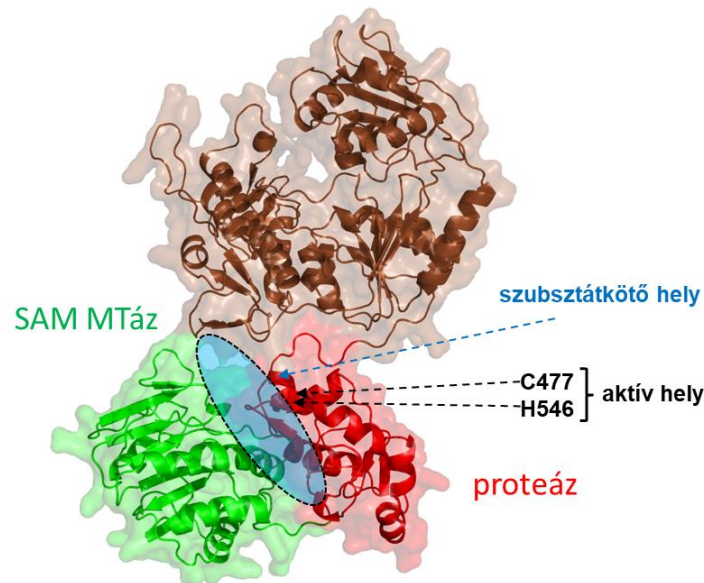
3.1.2. Az nsP2pro fehérje

A VEEV virális genom két nyitott leolvasási keretet (ORF-et) tartalmaz. Az első egy, a nem-szerkezeti fehérjéket (*non-structural protein*, nsP) tartalmazó poliproteint kódol, a második pedig a szerkezeti fehérjéket (Strauss és Strauss, 1994). A transzláció után a poliprotein (nsP123 vagy nsP1234, 2. ábra) proteolitikus hasítással darabolódik önálló nsP1, nsP2, nsP3 és nsP4 nem-szerkezeti fehérjékre.



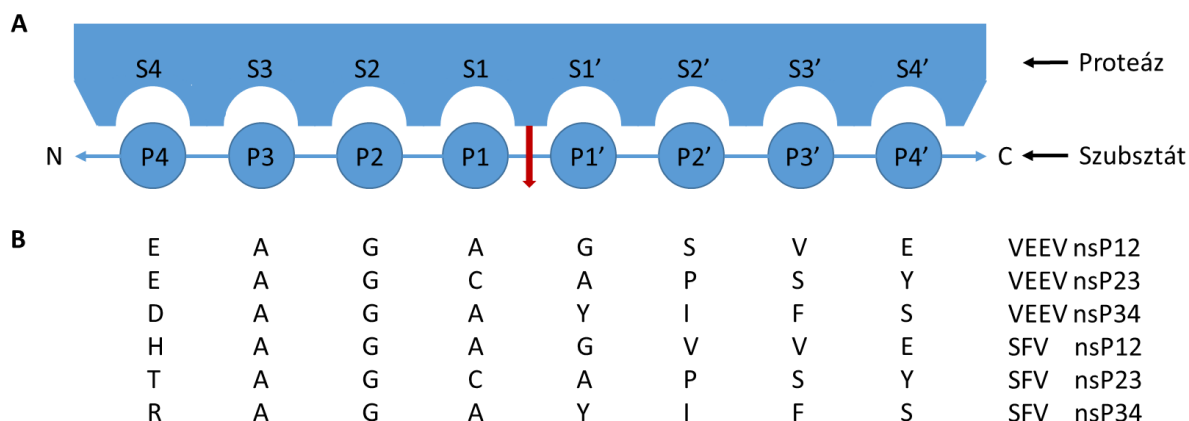
2. ábra. A VEEV nsP1234 poliprotein sematikus ábrája. Az nsP2 fehérje esetében feltüntettük a proteáz és az S-adenozil-L-metionin (SAM)-függő RNS metiltranszferáz (SAM MTáz) doméneket. A vastagon szedett számok az aminosavak az nsP2 fehérje-, míg a zárójelben lévő számok az nsP1234 poliprotein-alapú számozását jelölik.

A nem-szerkezeti fehérje 2 (nsP2) kiemelt fontosságú a vírus életsiklusának szempontjából. Az alfavírusok nsP2 fehérjéjét vizsgálva több fontos régiót határoztak meg, többek között az N-terminális régióról (Gly1-Ile456) megállapították, hogy ATPáz és GTPáz aktivitással (Rikonnen és mtsai, 1994), továbbá RNS helikáz aktivitással is rendelkezik (de Cedrón és mtsai, 1999). A C-terminális régiót (Met457-Cys794) alkotja többek között egy papain-szerű cisztein proteáz domén, amit pedig egy S-adenozil-L-metionin (SAM)-függő RNS metiltranszferáz (SAM MTáz) domén követ (Vasiljeva és mtsai, 2001; Vasiljeva és mtsai, 2003). A VEEV nsP2 teljes szerkezetét kísérleti módszerekkel még nem határozták meg, a teljes szerkezet AlphaFold alkalmazásával modellezhető (3. ábra).



3. ábra. A teljes hosszúságú VEEV nsP2 AlphaFold-dal készített modellje. A proteáz domén vörös, a SAM MTáz domén zöld, az nsP2 fehérje N-terminális doménje pedig barna színnel megjelenítve szerepel. A színezés megfelel a poliprotein esetében használt színezésnek (2. ábra). A proteáz szubsztrátkötő helye kékkel van jelölve, a katalitikus aminosavakat nyilak mutatják.

A cisztein proteázok családjába tartozó VEEV nem-szerkezeti fehérje 2 proteáznak (nsP2pro) (EC 3.4.22.B79) különösen fontos szerepe van a vírus életciklusában, ez felelős a virális poliproteinek funkcionális alegységekre hasításáért, limitált proteolízis által. A proteáz össze van kötve a SAM MTáz doménnel, és annak proteolízisben mutatott szerepét is sikerült igazolni: a doménben lévő Arg662, Lys705 és Lys706 aminosavak elősegítik a szubsztrát felismerését, és a szubsztrát kötő hely térszerkezetének kialakításában vesznek részt. A katalitikus helyen egy cisztein és egy hisztidin katalitikus aminosav található, a Cys477 és a His546, a szubsztrátkötő hely pedig a proteáz és a SAM MTáz domének között helyezkedik el (Russo és mtsai, 2006). Az nsP2pro legfontosabb szerepe az nsP1234 fehérje hasítása az nsP1/nsP2, nsP2/nsP3 és nsP3/nsP4 helyeken. A szubsztrátkötő alhelyek és az ezekbe kötődő szubsztrát egyes aminosavainak elnevezéséhez a Schechter és Berger által 1967-ben bevezetett nevezéktant vesszük alapul (Schechter és Berger, 1967). Ennek megfelelően a szubsztrát esetében a hasítóhelyet, mint központi elemet tekintve az N-terminális irányába haladva az aminosavak P1, P2, stb., míg a C-terminális irányába haladva P1', P2', stb. elnevezéssel illetjük. Az enzim szubsztrátkötő helyét alhelyekre osztjuk, valamennyi egy-egy szubsztrát aminosavrész kötéséért felel, és az előzőeknek megfelelően S2, S1, S1', S2', stb. néven hivatkozunk ezekre (4A. ábra).



4. ábra. A szubsztrátkötő alhelyek és a szubsztrát aminosavak nevezéktana és alfavírus nsP2 proteázok természetes hasítóhely szekvenciá. (A) A szubsztrátkötő alhelyek és a szubsztrát aminosavak nevezéktana. A hasítási pozíciót vörös nyíl jelöli. Az egyes szubsztrát aminosavak a hasítóhelytől az N-terminális irányába P1, P2, stb. jelöléssel, a C-terminális irányába P1', P2', stb. elnevezéssel szerepelnek. Az enzimnek a megfelelő szubsztrát aminosavat kötő katalitikus alhelyei S1, S2, stb., és S1', S2', stb., elnevezéssel szerepelnek, a Schechter és Berger által bevezetett nevezéktannak megfelelően (**Schechter és Berger, 1967**). **(B)** A VEEV és SFV nsP2pro természetes hasítóhely szekvenciái a virális poliproteinekben, a P4-P4' pozíciókban szereplő aminosavak feltüntetésével.

Az alfavírusok hasítóhely szekvenciái egyes pozíciókban jelentős hasonlóságot mutatnak (**4B. ábra**). Egy korábbi tanulmány megállapításai szerint a P1 helyen egy kisebb oldallánccal rendelkező aminosav található, amit a P2 pozícióban egy glicin követ mindhárom esetben (**Strauss és Strauss, 1994; Morazzani és mtsai, 2019**). A P3 helyzetben lévő aminosav konzervált, míg a P4 jellemzően eltérő lehet. A P1'-P4' pozíciókban lévő aminosavak azonban nagyobb mértékű változatosságot mutatnak (**Abu Bakar és Ng, 2018**).

A VEEV nsP2pro esetében az S1 alhely kialakításában az Asn475, Val476 és Ala509 aminosavaknak van fontos szerepe, amelyhez hasonló aminosavak találhatóak meg más alfavírusok esetében is (**Russo és mtsai, 2006**). A Val476 és Cys477 aminosavak főlánci amidjai oxianionos üreget képezhetnek. Az S2 alhely az ún. „glicin-specifikus motívumra” emlékeztet, amely más cisztein proteázokban is előfordul (**Golubtsov és mtsai, 2006**), egy magasan konzervált Trp547 és az előtte lévő katalitikus His546 alkotja, és a P2 helyen lévő glicin kötéséért felel. Az S3 alhely kialakításában fontos szerepe van a SAM MTáz doménnek, ezt az alhelyet az Ile698 és Met702 alkotja, amelyek mellett az Ala509 és His510 aminosavak szerepelnek.

Fontos megjegyezni, hogy több vizsgálat is célozta a VEEV nsP2pro biotechnológiai és molekuláris biológiai területen történő felhasználásának lehetőségét is. Kutatócsoportunk korábban - a SIN és SFV nsP2pro proteázokkal együtt - vizsgálta a VEEV nsP2pro alkalmazhatóságát fúziós címkék enzimátikus eltávolítására (**Zhang és mtsai, 2009**). A VEEV nsP2pro enzim kísérletileg meghatározott k_{cat}/K_m értéke kb. 600-szor alacsonyabbnak bizonyult a papainhoz képest (**Hu és mtsai, 2016**). A mesterséges oligopeptid szubsztrátokkal végzett *in vitro* hasítási reakciók eredményei alapján a VEEV, SIN és SFV proteázok katalitikus hatékonysága meglehetősen alacsonynak bizonyult, ami miatt ezek az alfavírus proteázok sokkal kevésbé bizonyultak alkalmas enzimátikus eszköznek a fúziós címkék proteolitikus eltávolítására a legelterjedtebben használt potivírusok, így a dohány karcolatós vírus (*tobacco etch virus*, TEV) (**Kapust és mtsai, 2001; Kapust és mtsai, 2002**) valamint a dohány érfoltosság vírus (*tobacco vein mottling virus*, TVMV) (**Nallamsetty és mtsai, 2004**) proteázaihoz képest.

A VEEV nsP2pro fontos gyógyszercélpont, egyrészt a vírus életciklusában betöltött kritikus szerepe miatt, másrészt mivel központi szerepet tölthet be az interferon védekezési mechanizmus hátráltatásában. Ezt részben a megfertőződött gazdasejtekben termelődő, a veleszületett immunválasz kialakításáért felelős fehérjék elhasításával éri el, aminek a patogenitás szempontjából is fontos jelentősége van. Több virális proteázról is kimutatták, hogy képesek felismerni a megfertőzött sejtek fehérjeiben a virális poliproteinben lévő hasítóhelyekhez hasonló szekvenciákat, és azokon belül elhasítani ezeket a fehérjéket. Ezeket a hasonlóságot mutató szekvenciákat a gazdaszervezet és a patogén fehérje szekvenciák rövid homológ szakaszainak nevezzük (*short stretches of homologous host-pathogen protein sequences*, SSHPS) (**Morazzani és mtsai, 2019**). A VEEV esetében ilyen hasítást igazoltak a TRIM14 (*Tripartite motif-containing protein 14*) fehérje esetében (**Morazzani és mtsai, 2019**), melynek szerepe van egy interferonok termelődéséért felelős kaszkádban (**Zhou és mtsai, 2014**), ezért hasítása valószínűleg megzavarja ezt a folyamatot. Érdekes, hogy az újvilági alfavírusok (VEEV, EEEV, WEEV) nsP12 hasítóhely szekvenciája azonos (**Morazzani és mtsai, 2019**) (a VEEV hasítóhely szekvenciáját lásd a **4. ábrán**), ezért mindhárom vírus proteáza képes a TRIM14 fehérje hasítására. Jelenleg nem áll rendelkezésre gyógyszeres védekezési lehetőség a VEEV fertőzéssel szemben, ezért az nsP2pro jelenleg is a gyógyszertervezés egyik fontos célpontjának tekinthető, a hatékony gátlószer azonosítására tett erőfeszítések folyamatosak (**Campos-Gomez és mtsai, 2016; Hu és mtsai, 2016; Kumar és mtsai, 2019; Zhang és mtsai, 2020**), a legutóbbi átfogó gátlószer-vizsgálat eredményeit 2023. júliusában tették közzé (**Hu és mtsai, 2023**).

Az elmúlt években kifejlesztett virális proteáz-gátló gyógyszerek közül kiemelhetjük a SARS-CoV-2 vírusfertőzés kezeléséhez elsőként elfogadott antivirális gyógyszert, a nirmatrelvirt, amely szintén egy cisztein proteázt, a vírus fő proteázát (M^{pro} vagy 3CL^{pro}) gátolja, kovalensen kötődve az aktív hely ciszteinjéhez (Mótyán és mtsai, 2022).

3.1.3. Az alfavírus nsP2pro fehérjék kristályszerkezete

Általánosságban elmondható, hogy jelenleg kevés alfavírus proteáz kristályszerkezet áll rendelkezésünkre. A 2024 év elején az RCSB Protein Data Bank adatbázisban elérhető szerkezeteket az **1. táblázat** foglalja össze. Ezidáig csak a VEEV, a CHIKV és a SINV vírus nsP2 proteázok szerkezetét határozták meg, a legtöbb szerkezeti információ a VEEV nsP2pro esetében érhető el. A VEEV nsP2pro esetében két, inhibitorot kötő szerkezetet is meghatároztak, azonban szubsztrátot kötő (mutáns) nsP2pro szerkezet nem áll rendelkezésre, az enzimszubsztrát kölcsönhatások vizsgálata napjainkban kizárólag modell szerkezet alapján lehetséges (Hu és mtsai, 2016).

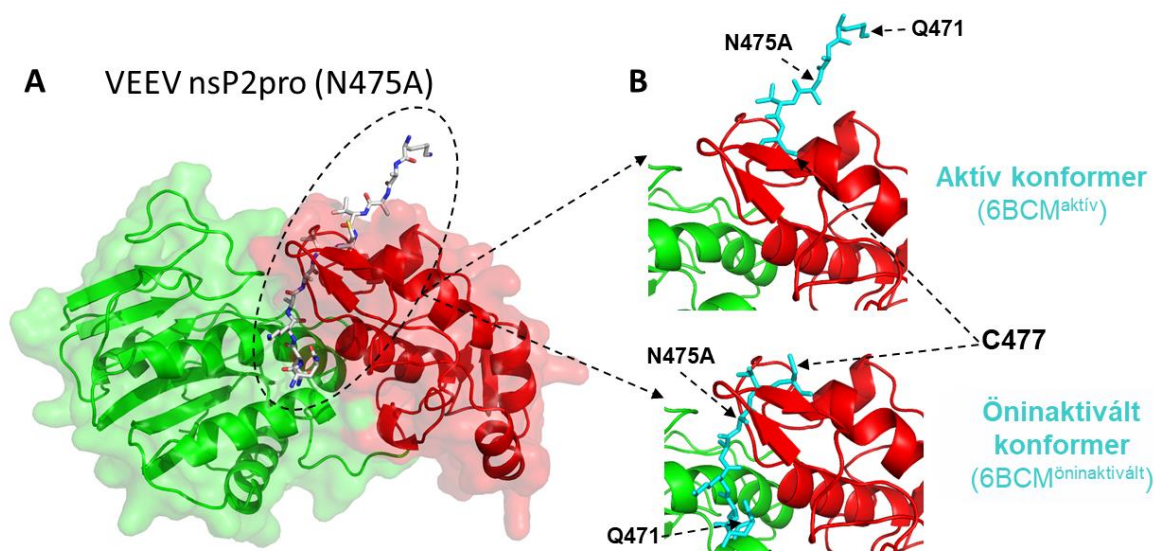
1. táblázat: A Protein Data Bank adatbázisban elérhető VEEV, CHIKV és Sindbis vírus proteáz szerkezetek adatai. Az adatbázis legutolsó megtekintésének ideje: 2024.02.04. A mutációk számozása az nsP2 protein alapján történt. Valamennyi szerkezetet röntgenkristallográfiával határozták meg. Az E64d és CA074 ligandumok peptid-szerű kovalens cisztein proteáz inhibitorok.

PDB ID	Vírus	Felbontás	Konformer	Mutáció	Ligandum	Hivatkozás
8DUF	VEEV	1,46 Å	aktív + öninaktivált	K741A, K767A	nincs	Hoffka és mtsai, 2023
6BCM	VEEV	2,10 Å	aktív + öninaktivált	N475A	nincs	Compton és mtsai, 2017
2HWK	VEEV	2,45 Å	aktív	nincs	nincs	Russo és mtsai, 2006
5EZS	VEEV	2,16 Å	aktív	nincs	E64d	Hu és mtsai, 2016
5EZQ	VEEV	1,66 Å	aktív	nincs	nincs	Hu és mtsai, 2016
8T8N	VEEV	2,34 Å	aktív	nincs	CA074	Hu és mtsai, 2023
3TRK	CHIKV	2,40 Å	aktív	nincs	nincs	nem elérhető
4ZTB	CHIKV	2,59 Å	aktív	nincs	nincs	Narwal és mtsai, 2018
4GUA	SINV	2,85 Å	aktív	nincs	nincs	Shin és mtsai, 2012

Időrendi sorrendben tekintve a VEEV nsP2pro fehérje szerkezetét először 2006-ban sikerült leírni (**Russo és mtsai, 2006**). A kutatás során az nsP2 457-794 régióját tartalmazó fehérjét expresszálták és kristályosították, majd vizsgálták szerkezetét röntgenkristallográfiás módszerrel. A szerkezetet 2,45 Å felbontással sikerült meghatározni, a kristályszerkezetben az Asp468-Ser787 régió szerkezetét sikerült azonosítani, amely magába foglalja a proteáz és a SAM MTáz doméneket is.

Egy másik tanulmányban a VEEV nsP2pro enzimet kinetikai és szerkezeti tulajdonságai tekintetében vizsgálták (**Hu és mtsai, 2016**). Ennek során a 457-792 régiót tartalmazó fehérjét kristályosították, a kristályszerkezetet az apoenzim formában és egy inhibitorral gátolt formában is meghatározták. Az apoenzim esetében (PDB ID: 5EZQ) a felbontás 1,66 Å volt, és az nsP2 Asp468-Ser787 régió szerkezetét sikerült meghatározni. Az inhibitorral alkotott komplex szerkezet (PDB ID: 5EZX) esetében ugyanezen fehérjeszerkezetét írták le, 2,16 Å felbontással. Egy későbbi kutatás során a VEEV nsP2pro egy mutáns változatát is vizsgálták (**Compton és mtsai, 2017**), amely esetében alacsonyabb enzimaktivitást határoztak meg. Az enzim az N475A mutációt tartalmazta, meghatározott kristályszerkezetében (PDB ID: 6BCM) az N-terminális régió egy alternatív konformerét figyelték meg. Ebben a szerkezetben az enzim N-terminális szakasza (Gln471-Cys477) az aktív hely irányába fordulva szerepelt, elfoglalva a szubsztrátkötő helyet. A tanulmány szerzői, mint „öninaktivált” formára utaltak rá (6BCM^{öninaktivált}), szemben a fő konformerrel, amely az aktív konformernek tekinthető (6BCM^{aktív}). Az elektronsűrűségi térkép alapján az aktív és inaktív konformerek 60 és 40%-os eloszlásban voltak jelen a szerkezetben. A kristályszerkezet felbontása 2,10 Å, abban a Gln471-Glu791 régió szerkezetét sikerült meghatározni. Az aktív konformáció esetében az enzim rövid N-terminális szakasza (Gln471-Cys477) az oldószer felé, míg az öninaktivált konformáció esetében az aktív hely felé fordult állapotban figyelhető meg (**5. ábra**).

A legújabb, 2023-ban publikált vizsgálat keretében egy *in silico* módszerek felhasználásával fejlesztett inhibitorral együtt kristályosított szerkezetet is sikerült meghatározni (**Hu és mtsai, 2023**). Ezeket a szerkezeti koordinátákat csak kutatásaink befejezése után publikálták, így azokat nem volt lehetőségünk bevonni a vizsgálatainkba. A nagyfokú szerkezeti hasonlóság miatt a VEEV nsP2pro gátlására képes molekulák potenciálisan gátolhatják más, a VEEV-vel közeli rokonságban álló alfavírusok proteázát is (pl. CHIKV, WEEV), ami szintén kiemelt jelentőséget biztosít a VEEV szerkezetének és gátolhatóságának megismerését célzó vizsgálatoknak.



5. ábra. A VEEV nsP2pro N475A mutánsának kristályszerkezete. (A) A VEEV nsP2pro N475A kristályszerkezete (PDB ID: 6BCM), bekarikázva az N-terminális régió látható. A szerkezetben az N-terminális aktív és öninaktivált konformere figyelhető meg. A proteáz domént vörös, a SAM MTáz domént zöld színnel jelöltük. (B) Az N-terminális régió közelített ábrázolása, az N-terminális régió kék színnel van megjelenítve az aktív és öninaktivált konformerek esetében egyaránt.

3.2. Kristályszerkezet-meghatározás

A doktori disszertációban bemutatott munka alapját egy, a kollaborációs partnereink által meghatározott kristályszerkezet molekuladinamikai tanulmányozása képezi, ezért fontosnak tartom röviden megemlíteni a kristályszerkezeti vizsgálatok alapjait, amelynek biológiai jelentőségét nem lehet elégszer hangsúlyozni (Shi, 2014). A módszer ma már a fehérjetudomány egyik alapkövének tekinthető (McPherson és Gavira, 2014; Gavira, 2016), azonban számos kihívást tartogat magában (Timofeev és Samygina, 2023). A módszer leírásának kiváló összefoglalása a magyar nyelvű szakirodalomban is megtalálható (Harmat és Vértessy, 2018).

A modern enzimológia tudománya nagymértékben épül a fehérjék atomi szintű szerkezetének ismeretére. A fehérje szerkezet kísérletes vizsgálatának számos módja ismert, melyek közül a legfontosabb módszerek közé tartozik a röntgenkristallográfia, krio-elektronmikroszkópia (krio-EM), mágneses magrezonancia (NMR) spektroszkópia, valamint a neutrondiffrakciós módszerek. Ezek közül kulcsfontosságú a röntgenkristallográfia, amely mint tudomány, már több mint egy évszázados múltra tekint vissza. Kifejezetten fehérjék szerkezetének meghatározására is már több mint fél évszázada alkalmazzák. Elsőként az

ámbráscet mioglobín szerkezetét sikerült meghatározni (**Kendrew és mtsai, 1958**), amelyért Sir J.C. Kendrew-t Nobel díjjal jutalmazták. A technológia fejlődésével az évek során egyre nagyobb felbontást sikerült elérni.

A kísérleti módszer lépéseit röviden a következőkben foglalhatjuk össze. Első lépésben szükség van a természetesen előforduló fehérje kinyerésére, vagy amennyiben az nem áll rendelkezésre, mesterséges (pl. kémiai szintézissel vagy rekombináns eljárással történő) előállítására. A bakteriális expressziós rendszerek (pl. *E. coli*) alkalmazása a legelterjedtebb a fehérjék szerkezet-vizsgálathoz történő előállítására. A megtisztított fehérjét ezután kristályosítani kell, a kristályosítás menete, az alkalmazott körülmények, és annak sikere fehérjénként eltérő lehet, és számos nehézséget tartogathat, pl. sikertelen kristályosodás rendezetlen szerkezet miatt, aggregáció, stb. Egyes esetekben további anyagok, pl. inhibitor hozzáadása lehet szükséges. A röntgendiffrakció alkalmazásához ún. egykristályra van szükség, amelynek belső szerkezete rendezett, lehetőség szerint nem tartalmaz szennyezőket, kristályhibákat. Fontos a megfelelő hőmérséklet biztosítása, melynek érdekében a kristályokat általában hűteni kell. A kristályt monokromatikus röntgensugárással sugározzák be, amely a kristályt alkotó molekulák elektronfelhőjével kölcsönhatásba lép, és reflektálódva szóródik. Az ún. diffrakciós képeket detektorok érzékelik, az adatfelvétel során a kristályt több irányban is forgatják, hogy minden irányból történjen adatfelvétel. A detektált diffrakciós képekből számítják ki az elektronsűrűségi térképet. A konkrét atomok pozíciójának meghatározása az elektronsűrűségi térkép alapján történik, ebbe a térképbe helyezik be az atomokat és építik fel a fehérje háromdimenziós szerkezetét. A modellépítés nehézsége alapvetően függ az elektronsűrűségi térkép részletgazdagságától, azaz felbontásától, minél kisebb a felbontás, annál „nehezebb” pontos modell felépítése. Gyenge felbontás esetén (3-3,5 Å) még az oldalláncok helyzete is bizonytalan, közepes felbontás esetében (1,5-2,5 Å) már azonosíthatók eltérő konformerek is, míg magas felbontás esetében (kb. 1 Å) az egyedi atomokat is azonosítani lehet. A modellépítést több módszer is segítheti, pl. molekuláris helyettesítés (**Scapin, 2013**) egy ismert, hasonló szekvenciával rendelkező fehérje szerkezete alapján, izomorf helyettesítés (**McCoy és Read, 2010**), izostrukturális kristályok alapján, ill. az anomális szórás módszere. A szerkezet felépítése során a geometriára vonatkozó ismereteket kényszerfeltételek formájában kell megadni (pl. kötőszögek, kötéstávolságok). A modell felépítése és javítása iteratívan történik, ami automatizálható is (**Badger, 2003**). Ezt követően a szerkezet finomítása történik, a modellt leíró paramétereket illesztve a kísérletileg mért adatokhoz. A szerkezetfinomítás geometriai kényszerfeltételek (ún. *restraint*-ek) alapján történik, a folyamat során minimalizálási lépést, esetleg molekuladinamikai szimulációt is alkalmazhatnak. A szerkezetet

ezután validálni kell, amely során fizikai, kémiai és biológiai ismereteink alapján várt tulajdonságait vizsgáljuk (Wlodawer és mtsai, 2008; Kleywegt, 2009). Elterjedt gyakorlat az R-faktor és a szabad R-faktor (Kleywegt és Brünger, 1996) használata a mért adatokkal való egyezés leírására (Vaguine és mtsai, 1999). Továbbá vizsgálni kell az ideális geometriától való eltéréseket és másodlagos kölcsönhatásokat is (Chen és mtsai, 2010).

Általánosan elmondható, hogy a módszer alkalmazásának előfeltétele a fehérje kristályosítása, amely sok esetben elvi (pl. rendezetlen fehérjék) vagy gyakorlati nehézségekbe ütközhet (pl. transzmembrán fehérjék). Bizonyos esetekben szükséges lehet a szekvencia kismértékű megváltoztatása (pl. mutációk beépítése), amely elősegítheti a kristályképződést. Az így meghatározott szerkezeteket többek között az RCSB Protein Data Bank (PDB) adatbázisba (Berman és mtsai, 2000) feltöltve őrizzük meg. Napjainkban ez az adatbázis már kizárólag a kísérletesen meghatározott szerkezeti koordinátákat gyűjti össze, a homológ modellezéssel felépített szerkezeteket nem tartalmazza. A szerkezeti koordináták esetében a pdb kiterjesztésű a legáltalánosabban használt szabvány fájlformátum, mely tartalmazza a meghatározott fehérje legfontosabb biológiai jellemzőit, a kristályszerkezet fizikai jellemzőit és az atomok Descartes-féle koordináta rendszerben megadott háromdimenziós koordinátáit (University of California, 2002).

3.3. Molekulamechanika

Számos molekuláris tulajdonság kísérleti vizsgálata ütközik gyakorlati nehézségekbe, amelynek oka részben a kísérleti módszerek elvi hiányosságaira, részben elméleti okokra vezethető vissza. Az egyik legnagyobb kihívás az egyedi molekulák vizsgálata, többek között a konformációs változások időbeli megfigyelése. A molekulamechanikai módszerek az egyik lehetőséget biztosítják arra, hogy megvizsgáljuk egyedi molekulák, atomszerveződések dinamikai tulajdonságainak változásait. A módszer alapelvei a klasszikus, newtoni mechanikát követik, amelyben a molekulákat rugókkal összekötött „atomokként” írjuk le. A molekulamechanikai „atom” az elektronokkal körülvevett atommagot egyben, mint töltéssel, tömeggel és mérettel rendelkező gömböt modellezi. A molekulamechanika molekuladinamikai módszerei különösen hasznosak szerkezetvizsgálati tanulmányok esetében. A molekuladinamikai módszerek elméleti alapjainak összefoglalását Andrew R. Leach mára már alapműnek számító könyvében találhatjuk meg (Leach, 1996).

3.3.1. Történeti összefoglalás

A klasszikus, newtoni fizikai világkép alapján a világ leírása determinisztikus módon képzelhető el, az egyes atomok időbeli mozgása a többi jelen lévő atommal való kölcsönhatása alapján meghatározott. Bármely pillanatban megmérve az atomok pozícióját és sebességét, abból a rendszer múltbeli és jövőbeli állapotait is kiszámíthatjuk. A kvantummechanika 1900-as évek elején történt felfedezése, majd axiomatikus leírása (**Neumann, 1980**) óta azonban tudjuk, hogy ez a leírás nem felel meg a valóságnak. Mégis, megfelelő közelítések alkalmazásával az atomi sokaságok egyes tulajdonságait vizsgálhatjuk a klasszikus leírás alapján is, amennyiben a vizsgált rendszerben nem tételezzük fel elektronátmenettel járó kémiai reakció lezajlását. Az ilyen rendszerek vizsgálatára fejlesztették ki a molekuladinamika módszerét, az 1950-es években. Eleinte a szerkezeti vizsgálatok során a leírt rendszer véletlenszerű megváltozásait vizsgáló Monte Carlo módszert alkalmazták (**Metropolis és mtsai, 1953; Harrison, 2010**), azonban a számítástechnika rohamos léptékű fejlődése lehetővé tette az egyes rendszerek, pl. fehérjék, oldószer sokaságok teljes, atomi szintű leírását, és az egyes atomok hosszú időtartamot felölelő viselkedésének determinisztikus vizsgálatát is. Az első igazi, determinisztikus molekuladinamikai szimulációt gömbök modellezésére használták (**Alders és Wainright, 1957**). A későbbi évtizedekben a molekuladinamikát fehérjék modellezésére is elkezdték használni (**McCammon és mtsai, 1977**), ma pedig már a mindennapos gyakorlat része.

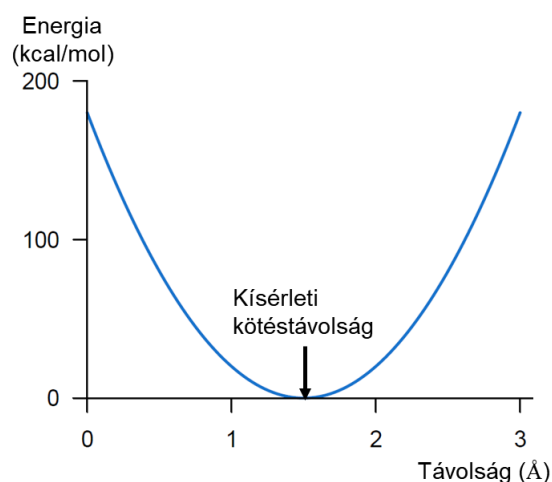
3.3.2. Erőterek

A szerkezetmeghatározó módszerek, a szerkezetek értelmezése, kiegészítése valamint számítógépes modellezése szempontjából kiemelt jelentőségű olyan erőterek igénybe vétele, amelyek a vizsgálni kívánt rendszert a valóságot minél jobban megközelítően írják le. Fehérjék vizsgálatához már számos erőtér áll rendelkezésre, ezek közül kiemelhetjük az Amber szoftverben használt (**Case és mtsai, 2016**) ff14SB (**Maier és mtsai, 2015**), vagy a GROMACS (**Bekker és mtsai, 1993**) által használt GROMOS erőtereket (**Gunsteren és mtsai, 2002**). Mivel a klasszikus modellben az atomokat rugókkal összekötött, gömb alakú, töltésekkel rendelkező modell-atomokként írjuk le, az erőtér ezen rugók és modell atomok kölcsönhatásainak számszerűsítéséért felel. Az erőtér általánosan egy több tagból álló függvény, amely az atomi minőségek és pozíciók összességéhez, azaz konformációjához hozzárendel egy energia-értéket.

Az Amber esetében (**Cornell és mtsai, 1995**) az alkalmazott erőter az alábbi képletben foglalható össze:

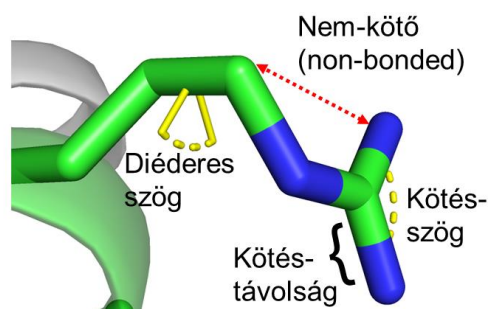
$$U = \sum_{\text{kötések}} k_r (r - r_0)^2 + \sum_{\text{szögek}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{diéderes}} k_\phi [1 + \cos(n\phi + \phi_0)] + \sum_{i.\text{atom}} \sum_{j \neq i} 4\varepsilon_{i,j} \left[\left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{12} - \left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^5 \right] + \sum_{i.\text{atom}} \sum_{j \neq i} \frac{q_i q_j}{\varepsilon_0 r_{i,j}} \quad (1)$$

Az ebben szereplő tagok közül az első a kötésben lévő atomok távolságától függ, ezt a kötetést jellemző rugóállandóval (k_r), a kötésre jellemző idealizált (vagy nyugalmi) kötéstávolsággal r_0 , és az atomok közötti aktuális távolsággal r , írjuk le. Az atomok távolságának ideálistól való eltérése nagy energiaváltozással jár, ez kényszeríti a rendszert az ideálshoz minél közelebb eső kötéstávolságok felvételére (**6. ábra**).



6. ábra. Molekuladinamikai erőter kötéstávolsági tagját jellemző harmonikus potenciál függvény. Az ideális, (kísérletileg meghatározott, esetleg magasabb szintű elméleti módszerekkel számított) kötéstávolságtól eltérve az energia négyzetes függvényt követve emelkedik.

A képlet második tagja hasonló összefüggést fogalmaz meg a kötésszögekre, míg a harmadik szög a diéderes szögekre, amely a periodicitását is figyelembe veszi. A következő két tag a kötésben nem lévő atomok közötti kölcsönhatások leírását foglalja össze. A negyedik tag a Lennard-Jones potenciál (**Lennard-Jones, 1931**) alapján írja le az atomok kölcsönhatását nagyon alacsony távolságok esetén (megakadályozza, hogy az ellentétes töltésű atomok túlságosan közel kerüljenek). Az ötödik tag a ponttöltések közötti Coulomb kölcsönhatást írja le (**7. ábra**).



7. ábra. Molekuladinamikai szimuláció során alkalmazott erőterek tagjai. A kötésben lévő atomok esetében kötéstávolság, kötésszög és diéderes szög alapján számítunk energiát, a kötésben nem lévő atompárok között ún. nem-kötő (*non-bonded*) interakciós energiát számítunk, amely Coulomb és van der Waals energiából tevődik össze.

Az erőterre a képlet formája és a képletben szereplő állandók is jellemzőek, annál pontosabban adja vissza a valóságot, minél pontosabban vannak meghatározva a paraméterei. Ez kísérleti módszerekkel (pl. kötéstávolság mérése) és magas szintű számítási módszerekkel (pl. kvantummechanikai módszerek) történhet (Weiner és mtsai, 1983; Brooks és mtsai, 1983). A paraméterek az atomtípustól is függenek, minél részletesebben különítünk el atomtípusokat, annál pontosabb erőtereket, modelleket készíthetünk. Pl. fontos különbséget tenni a gyűrűben lévő, vagy a lineáris molekulában lévő szénatomok között, stb.

3.3.3. Előkészítés

3.3.3.1. Kiindulási szerkezetek

A molekuladinamika módszereinek fehérijék esetében való alkalmazásához szükséges egy megfelelő szerkezeti modell, amellyel a szimulációkat elvégezzük. Ehhez több lehetőség áll előttünk, amelyek közül a leggyakrabban alkalmazott a kísérletileg meghatározott szerkezet használata. Ma már a fejlett szerkezet-meghatározási módszereknek köszönhetően (többek között röntgenkristallográfia és krio-EM) számos fehérje szerkezete ismert. Az elérhető szerkezetek könnyen hozzáférhetőek az RCSB Protein Data Bank (Berman és mtsai, 2000) adatbázisában, és számos vizualizációs szoftver segíti ezek megjelenítését, pl. Pymol (Schrödinger LLC., 2010), YASARA (Krieger és Vriend, 2018). Amennyiben a vizsgálni kívánt fehérje szerkezet elérhető, elfogadhatjuk, mint kiindulási szerkezetet.

Azonban számos esetben a vizsgálandó fehérje szerkezetét nem határozták meg kísérletesen, vagy bizonyos esetekben olyan mutáns változatok vizsgálata lehet szükséges,

melyek szerkezete nem elérhető. Pontmutációk beépítésére lehetőségünk van több szoftverrel is, ezek különböző elérhető rotamer adatbázisok (**Dunbrack, 2002**) alapján építik be a kívánt mutációt. Amennyiben a vizsgálni kívánt fehérje szerkezete egyáltalán nem elérhető, homológ modellezés, vagy újabban gépi tanuláson alapuló módszerek alkalmazásával van lehetőségünk annak előállítására. Homológia modellezés esetében a szoftver - pl. SWISS-MODEL (**Waterhouse és mtsai, 2018**) vagy MODELLER (**Webb és Sali, 2016**) - korábban már meghatározott szerkezetek alapján keresi ki a leginkább megfelelőt, amely a modellünk váza lesz. A módszer különösen hasznos lehet például nem-fehérje makromolekulákkal interakcióban lévő fehérjék szerkezetének meghatározására. A mesterséges intelligencia-alapú módszerek szintén korábban meghatározott kristályszerkezetek alapján állapítanak meg szerkezeteket, azonban azt a gépi tanulás módszereivel teszik. Ezek közül a legfontosabb az AlphaFold (**Jumper és mtsai, 2021; Mirdita és mtsai, 2022**), amely nagyon magas pontossággal képes szerkezetek becslésére (predikciójára), de sajnos az oldalláncok konformációja esetében súlyos hibák fordulhatnak elő, amely szükségessé teszi a szerkezetek további vizsgálatát.

A kristályszerkezetek, az ezek alapján készített homológiamodellek, és a tanított mesterséges intelligencia-alapú módszerek által becsült szerkezetek egyaránt magukban hordoznak egy hibalehetőséget: a kristályfázisban reprezentálják a fehérje szerkezetét, amely nem biztos, hogy megegyezik az élő sejtben, vagy pedig az *in vitro* kísérletek esetében valóban jelenlévő szerkezettel. A homológia modellezés esetében alkalmazott szekvencia-illesztés módszereinek és kihívásainak tárgyalása meghaladja a doktori dolgozat keretét.

3.3.3.2. A szerkezet előkészítése

A fehérje szerkezetet ezután szükséges előkészítenünk a szimulációk lefuttatására. Ellenőriznünk kell, hogy a kristályszerkezet adatait tartalmazó pdb fájl monomert vagy oligomert tartalmaz-e, és ez összeegyeztethető-e a vizsgálni kívánt fehérje biológiailag aktív formájával. Eltérés esetén szükséges lehet a felesleg lánc törlése vagy a megfelelő módon összerendezett oligomer készítése. Az aminosavak alternatív konformereinek megléte esetén szükséges lehet azokat külön-külön is figyelembe venni. Meg kell vizsgálni, hogy vannak-e hiányzó szakaszok, vagy beépítetlen oldalláncok, amelyek nem jelentek meg az elektronsűrűségi térképen, ilyenek lehetnek például relatíve nagyobb flexibilitású terminális szakaszok vagy intramolekuláris hurkok, felszíni (általában töltéssel rendelkező) aminosavak. Ezek pótlása homológ modellezés alapján történhet, vagy egyszerűen egy modellező szoftverrel

is beépíthetőek, ekkor az oldalláncot rotamer adatbázisból adhatjuk hozzá, ha ezek kevésbé lényegesek, csak a kristályszerkezetben lévő atomokat vesszük figyelembe.

Az oldószerként vagy inhibitorként jelen lévő molekulákat, kristályosítószereket, stabilizáló adalékokat és egyéb molekulákat sok esetben el kell távolítani a szerkezetből. Azonban mindenképp szükség lehet egyes, strukturális szempontból lényeges vízmolekulák, valamint a fehérjéhez kapcsolódó molekulák (pl. kofaktor, szubsztrát, inhibitor) szerkezetben tartására (amihez megfelelő paraméterezés is szükséges).

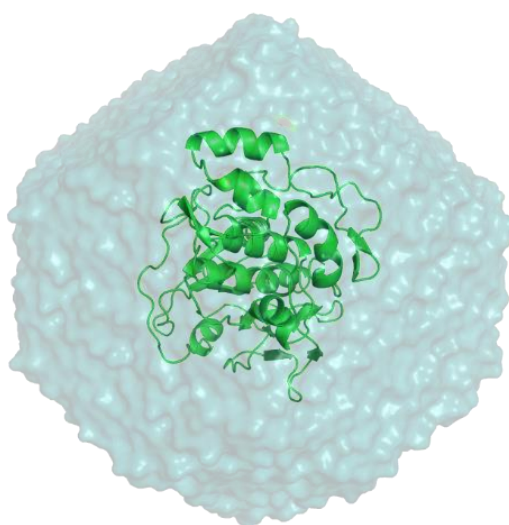
Meg kell határozni a fehérje aminosav oldalláncainak protonáltsági fokát: változó protonáltsági fokú további aminosavak (leginkább aszpartát, glutamát, lizin, arginin és hisztidin, esetleg cisztein) esetében savi disszociációs állandót (pK_a) prediktáló szoftvereket előnyös alkalmaznunk, pl. PROPKA 3.1 (Søndergaard és mtsai, 2011; Olsson és mtsai, 2011) vagy H++ webszerver (Gordon és mtsai, 2005; Anandakrishnan és mtsai, 2012) neutrális hisztidinek esetében a Chimera szoftvert (Pettersen és mtsai, 2004) alkalmazhatjuk a proton pozíciójának meghatározására. Fontos lehet megvizsgálnunk a szerkezetben lévő diszulfid hidakat, valamint a fehérjénk terminálisait is. Ha ezek a valóságnak megfelelő aminosavak, a modellező programok általában kiegészítik a láncvégi csoportokra jellemző amin- és karboxilcsoportokkal. Ha a terminális szakasz rövidebb, mint a valós fehérje, érdemes a terminális aminosavak töltés nélküli, ún. *end-capping* csoportokra módosítani. Általánosan ezek az N-terminálison acetyl-, míg a C-terminálison N-metilamid csoportok szoktak lenni. Ha a láncvégi oldallánc jelenléte kiemelt fontosságú, ez a lépés mellőzhető.

A hiányzó hidrogén atomokat a legtöbb molekuladinamikai szoftver előkészítő programja automatikusan pótolja. Az így összeállított fehérjével implicit oldószer alkalmazva futtathatunk számításokat, ám a valóság jobb közelítése explicit oldószer jelenlétét, ill. lehetőség szerint a fiziológiásnak megfelelő só-koncentrációt követeli meg, így azzal ki kell egészíteni rendszerünket.

3.3.3.3. A szimulációs fázis előkészítése

A fehérje szerkezetének előkészítése után szükséges az explicit oldószer is tartalmazó szimulációs „doboz” elkészítése. Ehhez különböző előkészítő programokat vehetünk igénybe, mint pl. az AmberTools *tleap* modulját, vagy webszervereket, pl. CHARMM-GUI (Jo és mtsai, 2008). A molekuladinamikai szimulációk általános esetben periodikus környezetben zajlanak (*Periodic Boundary Condition*), ami azt jelenti, hogy az adott, sík oldalakkal határolt szimulációs doboz minden oldalról egy további másolatával határos. A módszer több évtizedes múltra tekint vissza (Rahman, 1964) és az erőteret számításokat többnyire ilyen környezetben

végzik (Case és mtsai, 2005; Sieradzan, 2015). Előnye, hogy garantálja az anyagmegmaradást, ha egy atom a szimulációs dobozon kívül kerülne, az a doboz másik oldalán visszakerül a rendszerbe. Ezen felül az elektrosztatikus kölcsönhatás lassan lecsengő, a szimulációs dobozon kívül eső komponenseinek összegzésére is lehetőséget ad. A szimulációs doboz geometriája lehet pl. téglatest alakú (leginkább transzmembrán fehérjék szimulációja során), vagy csonka oktaédes (8. ábra). A doboz általában egy adott távolságra terjed ki a fehérje valamennyi atomjától, amelynek elég nagyoknak kell lenni ahhoz, hogy véletlenül se történjen meg például az, hogy a fehérje a saját atomjaival ütközik a túl kicsi szimulációs doboz miatt. Ezt a távolságot az előkészítő programokban adhatjuk meg.



8. ábra. Molekuladinamikai szimulációkra előkészített fehérjeszerkezet. A fehérje csonkolt oktaédes geometriájú, explicit vizekkel feltöltött szimulációs dobozban van.

Ezután szükséges az oldószer molekulák hozzáadása, amely általános esetben víz. A víz leírására több, kísérleti értékek és elméleti megfontolások alapján készített modellt alkalmaznak, ezek közül kiemelhetjük a régóta használt TIP3P modellt (Jorgensen és mtsai, 1983). Ezután fontos megfontolnunk a modell neutralizálását: ha a fehérjénk töltéssel rendelkezik, az erősen befolyásolhatja a szimuláció lefutását, akár a valóságban nem létező konformerek felvételére, vagy letekeredésre is kényszerítheti a fehérjét. Rövid szimulációk esetében ez a hatás nem jelentős, azonban hosszabb szimulációk esetén mindenképp szükséges legalább a rendszer semleges töltésének beállítása. Ezt ionok hozzáadásával érhetjük el, kationként leggyakrabban Na^+ vagy K^+ ionokat, anionként Cl^- ionokat alkalmazunk. Fontos megfontolni, hogy fiziológiás sókoncentrációt alkalmazzunk-e. Ennek beállítása további pontatlanságot hordoz magában, ha mind az anionok, mind a kationok koncentrációját 0,15 M-

ra állítjuk, a rendszer ismét csak nem lesz semleges. Erre kínál megoldást az ún. *Split charge* módszer (**Machado és mtsai, 2020**), amelynek keretében az egyes ionok számát a fehérje töltésének függvényében módosítjuk, az alábbi képletek alapján:

$$N^+ = N^{avg} - \frac{Q}{2} \quad (2)$$

$$N^- = N^{avg} + \frac{Q}{2} \quad (3)$$

A képletekben Q a fehérje töltését jelenti, N^{avg} az ionok átlagos számát, N^+ és N^- a kationok és az anionok számát jelenti. A módszer helyes alkalmazására több lehetőség is van, pl. szolvatálhatjuk a fehérjét, majd a hozzáadott vízmolekulák számából és a fehérje töltéséből kiszámíthatjuk a valóban hozzáadandó ionok számát, és ugyanazt a szolvatációs algoritmust módosíthatjuk, úgy, hogy egyes vízmolekulákat ionokra cserélünk. Így végül egy olyan neutrális rendszert kapunk, amely jól közelíti a kívánt ionkoncentrációt.

3.3.4. Szimulációk lefuttatása

3.3.4.1. Topológiai leírás

Az előkészített szerkezetnek az alkalmazott erőternek megfelelő leírásához szükséges adatokat a topológia fájl írja le, amihez feltételezzük, hogy nincs elektronátmenet a rendszerben, a kötések hálózata a teljes molekuláris rendszert leírja (**Bader és mtsai, 1980**). A topológia fájlban az szerepelnek mind a molekula teljes szerkezetét leíró adatok (parciális töltések, atomok közötti kötések), és az erőternek megfelelő paraméterek (kötésekre jellemző állandók értékei, atomtípusok, Lennard-Jones paraméterek. A topológia fájlban a teljes rendszer, azaz a lehetséges inhibitor, szubsztrát és vízmolekulák, ionok is szerepelnek, ezt a legtöbb szoftver automatikusan generálja. Ezen kívül szükség van az egyes atomok térbeli pozícióját leíró koordinátákra is. A klasszikus molekuladinamikai szimulációk során csak ezek változhatnak a szimuláció során, a topológiában megadott paraméterek nem.

3.3.4.2. A szimuláció körülményei

A molekuladinamikai szimulációk lefuttatásának első lépése általában a szerkezet konformációs energiájának minimalizálása, amelynek során a túl közeli atomok, vagy geometriai deformációk miatt jelentkező feszültségeket szüntetjük meg (**Levitt és Warshel, 1975**). A minimalizáláshoz több algoritmus is rendelkezésünkre áll, pl. a legmeredekebb ereszkedés (*steepest descent*) algoritmus, amely a minden lépésben a legnagyobb mértékű

energiacsökkenés irányába mozgatja a szerkezetet és a konjugált gradiens (*conjugate gradient*) algoritmus, amely összetettebb, figyelembe veszi a megelőző lépések irányát is. Az algoritmus megválasztása a vizsgált szerkezettől függ.

A minimalizált szerkezeteket ezután különböző szimulációs lépéseknek vetjük alá. Első lépésben valamilyen hőmérsékletnek megfelelően, véletlenszerű sebességeket adunk meg az atomok számára. Ezt követően, a szimulációk során az algoritmus lépésenként kiszámítja az egyes atomokra ható, az erőtértől és a pillanatnyi elrendeződéstől függő erőket.

$$F_i = -\frac{\partial V}{\partial r_i} \quad (4)$$

A képletben F_i az i -edik atomra ható erő, V a potenciális energia (értékét az erőtér adja meg), r_i az i -edik atom pozícióját jellemző vektor. A parciális deriválás után az i -edik atomra ható erőt kapjuk meg. Az erő nagyságának értékéből megkapjuk az egyes atomok gyorsulását, ill. sebességét. A konkrét számítás szoftverenként és algoritmusonként eltérhet, azonban általánosan elmondható, hogy a kiszámolt értékek alapján, és az általunk megadott időköz függvényében (az adott időköz alatt a számított gyorsulást állandónak tekintve) számítjuk ki az atomok új helyzetét. Ehhez gyakran a Verlet integrálás módszerét használjuk (**Verlet, 1967**). A lépésköz általában 1-2 fs, ugyanis ez a legtöbb vizsgálni kívánt mozgásnál gyorsabb. A lépésköz alatt elmozdult, módosított pozíciók fogják megadni az atomok új pozícióit. A következő lépésben az algoritmus újra kiszámítja a megváltozott erőket, és ismét lépteti az atomokat.

Nagymértékben kontrollálhatjuk a rendszer egyes tulajdonságait. Szükséges megadnunk, hogy milyen körülmények között futtatjuk a szimulációt, milyen tulajdonságok jellemzik majd a sokaságot. A leggyakrabban alkalmazottak az ún. NVT (*Number of particles, Volume, Temperature*), azaz kanonikus (**Nosé, 1983; Hoover, 1985**), valamint az ún. NPT (*Number of particles, Pressure, Temperature*), vagyis izotermális-izobárikus körülmények (**Parrinello és Rahman, 1980; Berendsen és mtsai, 1984**). A képletekben N az atomok számát, V a térfogatot, P a nyomást, T pedig a hőmérsékletet jelenti. A N jelentése az, hogy a rendszerben a kezdetben megadott atomokon, molekulákon kívül nem szerepelhet más atom. Létezik olyan szimulációs körülmény is, ahol az atomok száma változik, ez az ún. nagykanonikus szimuláció, ahol N helyett μ , azaz a kémiai potenciál várható értékét rögzítjük. A térfogat NPT körülmények között a nyomás függvényében változhat. A hőmérséklet és a nyomás várható értékét dinamikusan követjük. Az atomok átlagos sebességeloszlásától függ a hőmérséklet, ennek megfelelő skálázásra van szükség. Számos lehetőség áll rendelkezésre, itt

is részben a szoftvertől, és részben saját választásunktól függ, hogy melyiket alkalmazzuk, az egyik leggyakrabban alkalmazott a Langevin dinamika (**Langevin, 1908**) módszere. A nyomás beállítása hasonlóan történik, pl. a Berendsen (**Berendsen és mtsai, 1984**) vagy Monte Carlo (**Faller és de Pablo, 2002**) barosztát alkalmazásával, amelyek lehetővé teszik a nyomás értékének egy várható érték közelében tartását.

A molekuladinamikai szimulációk során a rendszert az algoritmus a megadott időközönként lépteti. Lehetőségünk van megadni a mintavételezési időtartamot: azt az időszakot, amennyi lépés után a rendszer elmenti a kapott koordinátákat. Érdemes ezt úgy beállítani, hogy megfelelően alapos mintavételezés történjen. A rendszer adott időközönként más adatokat is ment, pl. a szimuláció becsült hátralévő hosszát, ill. az atomokhoz tartozó sebesség értékeket.

A szimulációk során többféle lehetőségünk van ún. pozicionális megkötések alkalmazására, amelyek általunk meghatározott erősségű négyzetes függvényekkel rögzítik az atomok pozícióját. Az első minimalizációs lépések során nagy feszültségek lehetnek a rendszerben, pl. az újonnan beépített szegmensek környékén vagy a fehérje-oldószer határfelületen, ezek minimalizációja során célszerű lehet a többi régió rögzítése. Ez az ekvilibráció első lépéseiben különösen fontos jelentőségű, megakadályozza a valóságnak nem megfelelő konformációk gyors megváltozását.

3.3.4.3. *Simulated annealing*

A molekuladinamikai szimulációk ugyan képesek reprezentálni a vízben oldott fehérjét, azonban a kristályszerkezet vagy modell szerkezet jelentős mértékben eltérhet a fiziológiás oldatban valójában megjelenő szerkezettől. Ezek az eltérések azonban olyan nagy energiájú konformációs változásokkal küszöbölhetőek csak kik, amelyeket szobahőmérsékletű szimuláció esetén nem mintavételez a rendszer. A probléma megoldására nyújt lehetőséget az ún. *simulated annealing* módszer alkalmazása (**Kirkpatrick és mtsai, 1983; Henderson és mtsai, 2003**). Ennek során a cél, hogy egy lokális energiaminimumból eljuttassuk a rendszert a globális minimumba, amely leginkább jellemzi a szerkezetet. A rendszert fiziológiás hőmérsékletnél magasabb hőmérsékletre felfűtjük, azon a hőmérsékleten ekvilibráljuk, majd alacsony hőmérsékletre hűtjük. A lépéseket több ciklusban megismételve nagyobb eséllyel jutunk el a globális minimumba. A hőmérsékletek megválasztásánál ügyelni kell arra, hogy viszonylag alacsonyabb hőmérsékleten (pl. 360 K) nem feltétlenül kap a rendszer elég energiát az egyes energiagátak átugrására. Túl magas hőmérsékleten pedig esély van a rendszer destabilizálódására és olyan kölcsönhatások megszűnésére, amelyek a valóságban szobahőmérsékleten stabilak lennének, így pedig mesterségesen indukált átrendeződések vagy

konformáció-változások következhetnek be. Továbbá, figyelembe kell venni, hogy az esetleg jelen lévő ligandum magasabb hőmérsékleten egyszerűen „kirepülhet” a rendszerből.

3.3.4.4. Megkötések alkalmazása

Egyes esetekben, leginkább fehérje-ligandum komplex vizsgálata során szükség lehet további megkötések alkalmazására. Ezek bevezetését az indokolhatja például, hogy a valóságban a bekötődés után a szubsztrát minden valószínűség szerint rövid időn belül reakcióba lép az enzimmal, míg egy molekuladinamikai szimuláció során (lévén elektrontranszfert nem képes modellezni) nincs lehetőség reakcióba lépésre, ezért szükséges lehet a szubsztrátot mesterségesen az aktív helyben tartani. Erre nyújt lehetőséget a megkötések alkalmazása, amellyel meghatározhatunk egy adott távolságtartományt, amin belül a szubsztrát szabadon mozoghat, és olyan távolsági határokat, amelyeket átlépve a távolsággal négyzetesen arányosan az energia növekedni fog. A megkötés az erőterben ismertetett energiatagokhoz hasonló, azonban tágabb határok között mozoghat. Az Amber szoftverben (**Case és mtsai, 2016**) további távolsági limitek is megadhatóak, amelyeken túl négyzetes helyett lineáris a függvény változása.

3.3.5. Szimulációk analízise

3.3.5.1. Az r.m.s.d. érték

A szimulációk minőségi ellenőrzésének egyik lehetősége a négyzetes közép eltérés (*root mean square deviation*, r.m.s.d.) értékek kiszámítása, amelynek jelentőségét már a korai fehérjemodellezési tanulmányokban is hangsúlyozták (**Cohen és mtsai, 1979; Havel és mtsai, 1979**). A mozgékony, dinamikus fehérjék vagy fehérjeszakaszok (pl. rendezetlen fehérjék) esetében konvergenciáról akkor beszélhetünk, ha az r.m.s.d. értéke egy meghatározott tartományon belül fluktuál ugyan, de átlagértéke már nem változik. A kristályszerkezetekből nyert, határozott szerkezettel rendelkező fehérjék oldószerben történő viselkedésének megfelelő reprodukálását azonban kiválóan vizsgálhatjuk az r.m.s.d. értékek ellenőrzésével.

A számítás alapját az alábbi képlet képezi.

$$r. m. s. d. = \sqrt{\frac{\sum_{i=0}^N [m_i * (X_i - Y_i)^2]}{M}} \quad (5)$$

A képletben N az atomok száma, m_i az i . atom tömege, X_i az adott atom koordinátája a vizsgált szimulációs szerkezetben, Y_i az adott atom referencia szerkezetbeli koordinátája, M a teljes tömeg. Amennyiben tömeggel nem súlyozott r.m.s.d. értéket számolunk, $m_i=1$ és $M=N$.

A képletben szereplő referencia egyes esetekben a kristályszerkezet is lehet, de érdekesebb a szimuláció során kapott összes mintavételezett szerkezet átlagát alkalmazni, hiszen ez jobban reprezentálja az oldószerben jelen lévő fehérjét. A r.m.s.d. értékeket a szimulációs idő függvényében ábrázolva stabil szerkezet esetén egy nem túl nagy fluktuációkat mutató, közel egyenest kell kapnunk, amely igazolja, hogy nem történtek jelentős szerkezeti változások a szimuláció során.

3.3.5.2. Dinamikai tulajdonságok vizsgálata

Számos esetben szükséges lehet különböző, dinamikusan változó tulajdonságok vizsgálatára is (**Karplus és Petsko, 1990**). A fehérjék térszerkezetének leírásában nagy szerepet kapnak az olyan paraméterek, mint például az egyes aminosavak közötti távolságok, szögek és torziók, amelyek kölcsönhatások megjelenését igazolhatják, vagy pl. láncszakaszok stabilitását jellemezhetik. Ezek az ún. geometriai paraméterek mérése a trajektória analízisével történhet, a szimulációk lefutását követően. Erre specifikus programokat alkalmazhatunk, pl. a korábban említett *cpptraj* scriptet. Távolságmérésre több lehetőség is van, az egyik legerjedtebb az aminosavak tömegközéppontjai közötti távolság mérése. Ezt mintavételezett szerkezetenként számítjuk ki, és a szimulációs idő függvényében ábrázoljuk. A diagramok alapján kölcsönhatások megváltozására, esetleg megszűnésére következtethetünk.

3.3.5.3. Hidrogénhid kölcshatások vizsgálata

A hidrogénhidak rendkívül fontos jelentőséggel bírnak a fehérjék szerkezetének stabilizálásában (**Pace és mtsai, 2014**), szerepüket már korai fehérjeszerkezeti tanulmányok megmutatták (**Pauling és mtsai, 1951**). Kialakulásukért erősen polarizált kötésben, pl. hidroxil csoportban lévő hidrogének, mint donorok, és negatív parciális töltésű atomok, leginkább oxigén és nitrogén, mint akceptorok felelősek. A másodlagos kölcsönhatások közül az egyik legerősebb kölcsönhatást biztosítják, amelynél fehérjékben csak nagyon erős ionos (sóhid) kölcsönhatások lehetnek stabilabbak. Míg egy kristályszerkezet csak a fehérje egy adott állapotát (a kristályban lévő molekulák átlagát) reprezentálja, és nem is oldószer fázisban, a szimulációk lehetőséget biztosítanak a hidrogénhidak által alkotott kölcsönhatási rendszerek vizsgálatára. A hidrogénhidak elemzése szintén pl. *cpptraj* scripttel történhet, ami az atomokra jellemző távolságértékek alapján határozza a hidrogénhid jelenlétét. A hidrogénhidak megjelenésének százalékos arányát is kiszámíthatjuk, ebből következtethetünk azok relatív stabilitására. Az eltérő konformerek közötti százalékos eltérések alapján pedig következtethetünk a konformerek közötti stabilitásbeli különbségekre is.

3.3.5.4. Klaszterezés

A szimulációs trajektóriák rendszerint számos, többé-kevésbé eltérő szerkezetet tartalmaznak. Sok esetben szükség van egy-egy, a fehérjére leginkább jellemző szerkezet kiválasztására, pl. további vizsgálatok kezdeti szerkezeteként vagy reprezentatív vizualizáció céljából. Ehhez célszerű a kapott szerkezetek egymással való összehasonlítása, csoportosítása, és a csoport(ok)ra leginkább jellemző szerkezet(ek) kiválasztása, az ún. klaszterezés módszerével, amelynek során különböző megközelítésű algoritmusok alkalmazásával csoportosíthatjuk a szerkezeteket. Szerepe a molekuladinamikai szimulációk időtartamának növekedésével lett egyre fontosabb (**Karpen és mtsai, 1993; Shenkin és McDonald, 1994**) A klaszterek előállítására számos mód van, és nekünk is lehetőségünk van a paraméterek szabályozására, vagy a klaszterezés alapjául szolgáló aminosavak, domének, stb. kiválasztására (**Shao és mtsai, 2007**). A klaszterezés alapjául szolgáló paraméterek, mint csoport, összes tagjának lehetőség szerinti „centrális” elemét, azaz a „klaszter közepén” található centroid szerkezetet használhatjuk a klaszter reprezentálására.

3.4. A szimulációs szoftverek alkalmazása

3.4.1. Molekuladinamikai szoftverek

Ma már számos különböző molekuladinamikai szoftver áll rendelkezésre, ezek közül több ingyenesen elérhető. A leggyakrabban alkalmazottak közül kiemelhetjük az Amber (**Case és mtsai, 2016**), GROMACS (**Bekker és mtsai, 1993**) és NAMD (**Phillips és mtsai, 2008**) szoftvereket. Ezek közül jelen kutatásaink során teljes mértékben az Amber16 szoftverre hagyatkoztunk, a szám az adott verzió megjelenési évére utal. Helyesebb kifejezés lenne szoftvercsomagot használni, ugyanis az Amber számos, különböző funkciójú programot, scriptet tartalmaz. A programcsomag két részre osztható, az ingyenesen elérhető AmberTools és a korábban licenszhez kötött, ma már akadémiai felhasználók számára ingyenesen elérhető Amber csomagokra. Az AmberTools gyakorlatilag minden scriptet tartalmaz, amelyekkel egy fehérjét előkészíthetünk, lefuttathatjuk a szimulációkat, és az eredményeket is értékelhetjük. A csomaghoz tartoznak az egyes, Amber által használt erőterek paraméterei (pl. ff14SB erőtér (**Maier és mtsai, 2015**) a fehérjék leírásához, GAFF erőtér az egyéb molekulák leírására (**Wang és mtsai, 2004**), TIP3P modell (**Jorgensen és mtsai, 1983**) a víz leírására, stb.). A csomag tartalmazza a szerkezetek előkészítéséhez, topológia fájlok előállításához szükségeltet *tleap* scriptet, továbbá az egyéb molekulák, pl. inhibitorok paraméterezését segítő *antechamber* programot (**Wang és mtsai, 2006**). A szimulációk lefuttatására a *sander* programmal van

lehetőség. A lefutott szimulációk ellenőrzését, eredményeinek analízisét a *cpptraj* scripttel tudjuk elvégezni. Az AmberTools-on felül az Amber csomagra a szimulációk gyorsabb, pl. grafikus processzoregység (GPU)-gyorsítást használó alkalmazásának lefuttatásához van szükség, amelyre a *pmemd* script segítségével van lehetőség (Le Grand és mtsai, 2013; Salomon-Ferrer és mtsai, 2013). Valamennyi, a disszertációban bemutatott számítást az Amber16 szoftverrel futtattunk le, igénybe véve a *pmemd* algoritmus GPU-gyorsított változatát. A szerkezetek előkészítését, az eredmények analízisét pedig az AmberTools-ban elérhető scriptek segítségével végeztük el.

3.4.2. Szuperszámítógépes környezet

A molekuladinamikai szimulációk általában nagy számítási kapacitást igényelnek, és a szimulációk során nagy mennyiségű adat generálódik. Ezeket figyelembe véve a számítások lefuttatására leginkább szuperszámítógépes (HPC) környezetben van lehetőség. Az ilyen klaszterekben számos nagy teljesítményű CPU mag és GPU áll rendelkezésre, amelyek nagyban megkönnyítik munkánkat, hiszen egyszerre, párhuzamosan futtatható több szimuláció, nem kell megvárni az egyes számítások lefutását. A szuperszámítógépes környezetben előre telepített szoftverek segítik munkánkat, amelyeket az adott HPC erőforrásainak megfelelően paramétereztek, így nekünk már ezek beállításával nem kell foglalkoznunk. Továbbá, akadémiai licenccel olyan szoftverek is elérhetőek lehetnek, amelyek külön megvásárlása jelentős összegeket emésztene fel. Fontos kiemelnünk a GPU egységek molekuladinamikában betöltött fontos szerepét. A modern MD programok erősen optimalizáltak a GPU-kon való futtatásra, sokszorosán meggyorsítva munkánkat ezzel a hagyományos CPU futásokkal szemben. Ebben kiemelhetjük a CUDA programozási interfész szerepét, amelyre a legtöbb GPU gyorsítást igénybe vevő molekuladinamikai szoftver támaszkodik.

4. CÉLKITŰZÉSEK

A doktori disszertáció alapjául szolgáló kutatásaink célja a VEEV nem-szerkezeti fehérje 2 proteáz (nsP2pro) működését meghatározó szerkezeti sajátságok számításon alapuló kémiai módszerekkel történő vizsgálata volt, különös tekintettel az aktív helyen kialakuló inter- és intramolekuláris kölcsönhatásokra. A szerkezeti vizsgálatok részét képezték azon kísérletes vizsgálatoknak, melyek célja az enzim specificitásának *in vitro* aktivitásmérésekkel történő meghatározása, valamint az enzim szerkezetének röntgenkristallográfiával való feloldása volt.

I. A VEEV nsP2pro szubsztrát-specificitásának vizsgálata

A kísérletes specificitás vizsgálatok kiegészítéseként célul tűztük ki a VEEV nsP2pro szubsztráttal alkotott komplex szerkezetének elkészítését, a komplex molekuladinamikai szimulációkkal történő elemzését, valamint különböző szubsztrát variánsok szerkezeti és energetikai vizsgálatát, az aminosav preferenciákat meghatározó szerkezeti sajátságok azonosítása érdekében.

II. A VEEV nsP2pro kristályszerkezetének vizsgálata

A VEEV nsP2pro kristályszerkezetében megfigyelt aktív és öninaktivált konformációs állapotok részletesebb vizsgálata érdekében célunk volt az enzim stabilitását meghatározó szerkezeti sajátságok feltérképezése az elérhető kristályszerkezetek összehasonlítása és molekuladinamikai számítások segítségével, különös tekintettel az enzimet érintő mutációk hatásainak valamint az aktív hely kölcsönhatási hálózatainak feltérképezésére.

5. ANYAGOK ÉS MÓDSZEREK

A jelen munkánkban bemutatott valamennyi számítás a Kormányzati Informatikai Fejlesztési Ügynökség (KIFÜ, korábban NIIF) által rendelkezésünkre bocsájtott HPC erőforrások igénybe vételével készült. Munkánk egyes részeit, pl. adatelemzést részben a Debrecen1 szuperszámítógépen, a szimulációkat pedig a Debrecen2 szuperszámítógépen végeztük. A Debrecen2 szuperszámítógép alkalmazásakor igénybe vettük az elérhető GPU magokat, a szimulációk lefuttatásához pedig a *pmemd* GPU gyorsítást használó verzióját alkalmaztuk.

5.1. A VEEV nsP2pro szubsztrát-specifitásának vizsgálata

5.1.1. Az enzim-szubsztrát komplex előkészítése

A fehérjemodellezés kiindulási szerkezete a specificitás vizsgálata során a VEEV nsP2pro és a természetes VEEV nsP1/nsP2 hasítóhelyet reprezentáló oligopeptid szubsztrát (LQEAGA↓GSVETP) által alkotott komplex modell szerkezete volt. A fehérje szerkezete a 469-767 pozícióban lévő aminosavak tartalmazta, a korábban felépített komplex szerkezeti koordinátáit Dr. Patricia M. Legler bocsájtotta rendelkezésünkre (Center for Bio/molecular Science and Engineering, U.S. Naval Research Laboratory) (**Hu és mtsai, 2016**).

Az eredeti modell és a kísérletileg vizsgált fehérje szekvenciáit a Clustal Omega (**Sievers és Higgins, 2014**) szoftver alkalmazásával illesztettük egymásra. Annak érdekében, hogy az enzim-szubsztrát komplexben lévő oligopeptid az *in vitro* kísérletek során is alkalmazott SFV nsP1/nsP2 hasítóhelyet (EYHAGA↓GVVETP) reprezentálja, szükség volt a rendelkezésünkre bocsájtott komplexben lévő peptid (LQEAGA↓GSVETP, **Hu és mtsai, 2016**) módosítására. Továbbá, szükség volt az nsP2pro C-terminális szakaszán két aminosav módosítására is, hogy a szekvencia megegyezzen a kísérletesen vizsgált fehérjéjével. Mindegyik aminosav esetében a FoldX szoftverrel (**Schymkowitz és mtsai, 2005**) végeztük el a mutálást. A N-terminálist mind a proteáz fehérje, mind a szubsztrát esetében acetyl csoporttal egészítettük ki, a C-terminálosokon metilamid csoportot alkalmaztunk. A titrálható aminosavak (aszpartát, glutamát, cisztein, lizin, arginin és hisztidin) protonáltsági állapotait a PROPKA 3.1 szoftverrel (**Søndergaard és mtsai, 2011; Olsson és mtsai, 2011**) állapítottuk meg. A neutrális hisztidinek esetében ezen felül Chimera szoftverrel (**Pettersen és mtsai, 2004**) meghatároztuk, melyik tautomer állapot jellemzi az adott oldalláncot. Az aktív hely katalitikus cisztein

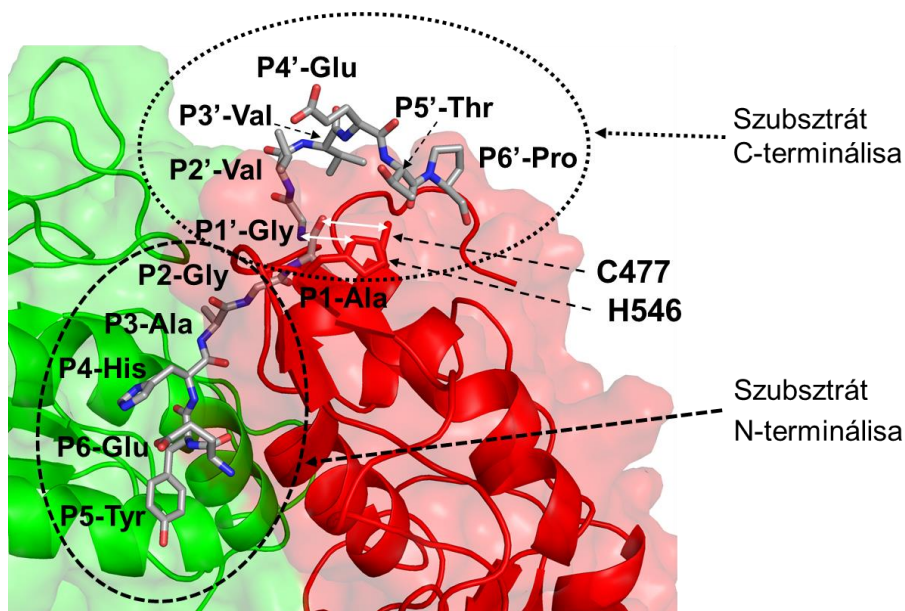
aminosavát deprotonált, a katalitikus hisztidint protonált oldallánccal modelleztük. Az előkészítést az Amber16 (Case és mtsai, 2016) *tleap* moduljával végeztük el. A szerkezeteket kiegészítettük hidrogénekkal, majd a fehérje és a szubsztrát komplexéhez vízmolekulákat adtunk. Mind a proteáz, mind a szubsztrát leírásához a ff14SB (Maier és mtsai, 2015) erőteret alkalmaztuk. A vízmolekulák leírása a TIP3P (Jorgensen és mtsai, 1983) modell alapján történt.

5.1.2. Molekuladinamikai szimulációk

A számítások lefuttatásához az Amber16 szoftvert alkalmaztuk (Case és mtsai, 2016), GPU gyorsítást igénybe véve (Le Grand és mtsai, 2013; Salomon-Ferrer és mtsai, 2013). Az előkészített rendszert elsőként minimalizáltuk, először a legmeredekebb leereszkedés (*steepest descent*) módszert alkalmazva, 5000 lépésen keresztül, majd a konjugált gradiens módszer alkalmazásával további 5000 lépésben. A szimulációk során 1 fs-os lépésközt alkalmaztunk, és a sokaságot jellemző körülmények közül a nyomást határoztuk meg állandóként (NPT sokaság). A rendszert elsőként 300 K-re fűtöttük egy 2 ns hosszú szimuláció során, majd 300 K hőmérsékleten ekvilibráltuk 2 ns időtartamon keresztül. A hőmérséklet szabályozásához a Langevin dinamika módszerét alkalmaztuk. Az egyik atomként hidrogént tartalmazó kötésekre a SHAKE algoritmust (Miyamoto és mtsai, 1992) alkalmaztuk. A rendszert ezután tovább fűtöttük 400 K-re egy 0,5 ns hosszú szimuláció során, majd 400 K-en ekvilibráltuk további 0,5 ns ideig. Ezután 5 K hőmérsékletre hűtöttük egy 1 ns hosszú szimuláció során, a *simulated annealing* módszerének megfelelően. Ezután ismét 400 K-re melegítettük egy 0,5 ns hosszú szimuláció során, ezen a hőmérsékleten egy további 0,5 ns hosszú ekvilibrációs szimulációt futtattunk. Ezt a szimulációs ciklust még kétszer megismételtük, azonban a második esetben a rendszert csak 300 K hőmérsékletre fűtöttük fel. Ezen a hőmérsékleten további 0,5 ns ideig egy ekvilibrációs szimulációt futtattunk.

A szimulációk során egy $50 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ értékű megkötést alkalmaztunk a katalitikus diád cisztein aminosavának (C477) kén atomja és a szubsztrát hasítási helyen lévő alanin aminosavának karbonil oxigénje között (9. ábra). Egy ugyanilyen értékű megkötést alkalmaztunk a katalitikus hisztidin (H546) donor nitrogén atomja és a szubsztrát hasítási helyén lévő glicin aminosav főlánci nitrogén atomja között, amelyet a hisztidin a ciszten proteázok általános mechanizmusának megfelelően protonálhat. A megkötések alkalmazását az indokolta, hogy azok nélkül futtatva a szimulációt a katalitikus diád ciszteinjének (Cys477) kén atomja és a szubsztrát P1-Ala karbonil szene közötti távolság a szimuláció utolsó lépésére 7,73 Å-re növekedett a kiindulási 2,80 Å értékről, emellett vízmolekulák kerültek a két aminosavréz

közé. A szubsztrát aktív helyről kimozdulásának oka feltételezhetően az volt, hogy a kiindulási komplexben lévő szubsztrát módosítása következtében olyan mértékben megváltoztak az enzim-szubsztrát kölcsönhatások, amik a molekuladinamikai szimuláció során nagy erők fellépését okozták.



9. ábra. VEEV nsP2pro modell és SFV nsP1/nsP2 hasítóhelyet reprezentáló oligopeptid szubsztrát komplex kiindulási szerkezete. A proteáz színezése a korábbi ábráknak megfelelő. A szubsztrát szén, oxigén és nitrogén atomjait szürke, piros és kék színek jelölik. Fehér nyilak jelölik a bevezetett távolság-megszorításokat a szubsztrát és a katalitikus aminosavak atomjai között.

5.1.3. Az enzim-szubsztrát kölcsönhatások vizsgálata

A szimulációk konvergenciáját az utolsó, ekvibrációs szakaszban meghatározott r.m.s.d. értékek ellenőrzésével igazoltuk. Ennek során a referencia szerkezet az ekvibráció során kapott szerkezetek atomi koordinátáinak átlaga volt. A számítások elvégzéséhez a *cpptraj* programot használtuk (Roe és mtsai, 2013).

A különböző pozíciókban módosított szubsztrátot tartalmazó komplexek elkészítésére a DynaMut webszert alkalmaztuk (Rodrigues és mtsai, 2018), amelynek segítségével a szubsztrát pontmutációnak hatására bekövetkező feltekeredési energia-változásokat is becsültük. A DynaMut a mutációk hatására bekövetkező energetikai változásokat a pdb formátumban feltöltött szerkezet alapján, normál mód analízissel becsüli, paraméterezését a gépi tanulás módszerével végezték. Összesen 1 db P5, 4 db P4, 2 db P2, 2db P1, 19 db P1' és

2 db P2' mutáns szubsztrátot vizsgáltunk. A feltekeredési energia (szabadentalpia) definíció szerint ≥ 0 kcal/mol változásait ($\Delta\Delta G$) stabilizáló, míg a $\Delta\Delta G < 0$ kcal/mol mértékű változásait destabilizáló hatásként adtuk meg, a DynaMut szoftver számítási algoritmusának megfelelően. A jelentős szerkezeti változást okozó mutációk esetében a DynaMut kevésbé pontosan képes prediktálni az energiaváltozást (**Rodrigues és mtsai, 2018**). A szerver validálásához használt vakteszt esetében a kísérletesen igazolt nagymértékű feltekeredési energia-változások esetében, ahol $|\Delta\Delta G| \geq 2$ kcal/mol volt, sok kiugró értéket kaptak. Az interakciók 2D vizuálizációjához a LigPlot+_v2.1 szoftvert (**Laskowski és mtsai, 2011**) használtuk, amely a szerkezeti koordináták alapján határozza meg a hidrogénhidakat és apoláros kölcsönhatásokat, míg a szerkezetek 3D megjelenítéséhez a PyMOL szoftvert (**Schrödinger LLC., 2010**) használtuk.

5.2. A VEEV nsP2pro kristályszerkezetének meghatározása

A VEEV nsP2pro kristályszerkezetét kollaborációs partnereink határozták meg. A rekombináns fehérje előállítás, kristályosítása és a szerkezet feloldása George T. Lontos, Danielle Needle, Alexander Wlodawer és David S. Waugh munkája (**Hoffka és mtsai, 2023**). A szintetizált fehérje a VEEV nsP2 nem-szerkezeti fehérje Arg463-Thr785 régióját tartalmazta. A K741A és K767A felszíni entrópiát csökkentő mutációkat is tartalmazta a szerkezet, amelyek a SAM MTáz szekvenciában voltak jelen, a cisztein proteáz doménben nem szerepelt mutáns aminosav. A rekombináns fehérjét expresszálták, majd a kapott fehérjét kristályosították. A kristályt röntgendiffrakciós vizsgálatnak vetették alá, amelynek kísérleti leírása nem tartozik jelen dolgozat leírásába. A kapott kristályszerkezet a Protein Data Bank adatbázisban érhető el, azonosítója: 8DUF (<https://www.rcsb.org/structure/8DUF>).

5.3. VEEV nsP2pro kristályszerkezetek molekuladinamikai vizsgálata

5.3.1. A szerkezetek előkészítése

A szerkezeti vizsgálatok kiindulási szerkezeteit az RCSB Protein Data Bank (**Berman és mtsai, 2000**) adatbázisból töltöttük le. Három nsP2pro szerkezetet vizsgáltunk: az újonnan meghatározott kristályszerkezetünket, amely a K741A és K767A mutációkat tartalmazta (PDB ID: 8DUF), az elsőként meghatározott, vad típusú nsP2pro szerkezetet (PDB ID: 2HWK, **Russo és mtsai, 2006**) és az N475A mutáns szerkezetét (PDB ID: 6BCM, **Compton és mtsai, 2017**). A szerkezeteket Chimera (**Pettersen és mtsai, 2004**) segítségével a 6BCM szerkezethez igazítottuk. A fehérjeláncokon és a kristályszerkezetekben lévő vízmolekulákon kívül minden egyéb jelenlévő atomot eltávolítottunk. A 8DUF esetében a szerkezetet az Ala768 aminosavig

vizsgáltuk. Az aktív és öninaktivált szerkezeteket külön vizsgáltuk, ehhez mindkét esetben az egyes aminosavak adott szerkezetnek megfelelő konformációját vettük figyelembe. A titrálható aminosav oldalláncok protonáltsági állapotainak meghatározása a PROPKA 3.1 szoftvert (Søndergaard és mtsai, 2011; Olsson és mtsai, 2011) alkalmaztuk. A szoftver által a térszerkezetek alapján az egyes aminosavakhoz kiszámított pK_a értékeket vettük figyelembe, amelyek alapján beállítottuk az aminosavak protonáltsági állapotait. A terminálisokhoz GaussView 6 (Dennington és mtsai, 2016) segítségével acetyl-, és N-metilamid csoportokatadtunk. A hidrogéneket az AmberTools16 program *tleap* moduljának segítségéveladtuk hozzá a szerkezetekhez. A katalitikus aminosavak ebben az esetben is aktív, töltött formában szerepeltek (deprotonált cisztein, ill. protonált hisztidin).

A fehérjéket minden esetben explicit oldószer vízmolekulákkal egészítettük ki, továbbá megfelelő számú vízmolekulát a *Split charge* módszer (Machado és mtsai, 2020) alkalmazásával a fiziológiás ionkoncentrációt és az egyes szerkezetek töltését figyelembe véve nátrium és klorid ionokra cseréltük. Első lépésben a vízmolekulákatadtunk a rendszerekhez, majd azok száma és az ionok koncentrációja (0,15 M) alapján kiszámoltuk az ionok átlagos számát, amely mind a nátrium, mind a klorid ionokra jellemző. Ezután az ionok számát a módszernek megfelelően módosítottuk, a kapott ionkoncentrációk nem tértek el jelentős mértékben a fiziológiás 0,15 M értéktől, azonban így a rendszert sikerült neutrális állapotba hoznunk. Mindezekhez a lépésekhez a *tleap* programot alkalmaztuk. A *tleap* alkalmazásávaladtuk hozzá a kristályszerkezetekből hiányzó hidrogéneket is. A protein paraméterezésére az ff14SB (Maier és mtsai, 2015) erőteret használtuk, míg a vízmolekulákat a TIP3P modell segítségével írtuk le (Jorgensen és mtsai, 1983).

5.3.2. Molekuladinamikai szimulációk

A szimulációkhoz az Amber16 szoftvert (Case és mtsai, 2016) vettük igénybe, az ebben elérhető *pmemd* program GPU gyorsítást használó változatával futtattuk a számításokat (Le Grand és mtsai, 2013; Salomon-Ferrer és mtsai, 2013).

A molekuladinamikai szimulációk előtt a rendszereket minimalizáltuk, a legmeredekebb leereszkedés (*steepest descent*) algoritmussal, 100000 lépésen keresztül, amit a konjugált gradiens módszerrel folytattunk, további 25000 lépésen keresztül. Ezután molekuladinamikai szimulációk futtattunk, amelyek során mindvégig a Langevin dinamika módszerét alkalmaztuk a hőmérséklet szabályozására, és alkalmaztuk a SHAKE algoritmust (Miyamoto és mtsai, 1992) a hidrogéneket tartalmazó kötések hosszának megkötésére. A szimuláció lépésköze 1 fs volt. A kötésben nem lévő atomok közötti (*non-bonded*) interakciók

leírására egy 10 \AA mértékű limitet szabtuk meg. A nagy távolságú elektrosztatikus kölcsönhatások leírására a Particle Mesh Ewald (PME) módszert alkalmaztuk (**Darden és mtsai, 1993**).

Első lépésként a vizsgált rendszereket 300 K -re fűtöttük egy 1 ns hosszú szimuláció során, NVT sokaságot alkalmazva, amit egy 1 ns hosszú NVT ekvilibrációs szimuláció követett 300 K hőmérsékleten. Ennek során $10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ mértékű helyzeti megkötést alkalmaztuk a fehérje atomjain. Ezután egy további ekvilibrációt futtattunk, NPT körülmények között, $0,5 \text{ ns}$ időtartamon át. Az ekvilibrációt követően a rendszereket 400 K hőmérsékletre fűtöttük, egy $0,5 \text{ ns}$ hosszú szimuláció során. Ezen a hőmérsékleten további 1 ns időtartamig ekvilibráltuk a rendszereket. Ezután egy $2,5 \text{ ns}$ hosszú szimuláció során lehűtöttük azokat 5 K hőmérsékletre. Ezután a rendszert ismét 400 K -re fűtöttük, ekvilibráltuk, és 5 K -re hűtöttük, a *simulated annealing* protokollnak megfelelően, az előbbieken megadott időtartamokat alkalmazva. Azonban ezen ciklus során már csak egy kisebb, $5 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ mértékű helyzeti megkötést használtunk a fehérje atomjainak megkötésére. Ezután további két, ugyanilyen ciklust futtattunk le, melyek során rendre $3 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ és $1 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ mértékű pozicionális megkötést alkalmaztunk, de már csak a fehérje főláncban lévő atomjain. Végül megkötések alkalmazása nélkül a rendszereket 5 K -en $1,5 \text{ ns}$ időtartamig ekvilibráltuk, majd 300 K hőmérsékletre fűtöttük egy $1,5 \text{ ns}$ hosszú szimuláció során, és 300 K -en ekvilibráltuk további 5 ns időtartamig. Ezután a kiértékelésre tervezett, adatgyűjtési szakaszok következtek, amelyek során megkötések nélkül, 300 K -en futtattuk a szimulációkat 50 ns időtartamig. A szerkezeteket jellemző koordináták mintavételezése 5 ps időtartammal történt. Minden általunk vizsgált szerkezet esetében 3-3 különálló szimulációt futtattunk le.

5.3.3. A szerkezetek analízise

Az utolsó lépésben futtatott adatgyűjtési szakasz (*production run*) szimulációk során kapott szerkezetekre kiszámítottuk az r.m.s.d. értékét, referenciaként a minden tizedik szerkezet összegzésével kapott koordináták átlagát adtuk meg. Az átlag számítása során a három különálló szimulációt egy közös adattömbként kezeltük, azonban az r.m.s.d. értéket futásonként külön tüntettük fel. A számítást a *cpptraj* program alkalmazásával futtattuk le (**Roe és mtsai, 2013**), amely molekuladinamikai trajektóriák elemzésére alkalmas, beleértve dinamikai tulajdonságok, kölcsönhatások kvantitatív, és statisztikai vizsgálatát is. A *cpptraj* alkalmazásával vizsgáltuk az N-terminális szakaszok által képzett hidrogénhidak hálózatát is, amelynek meghatározásához ugyancsak a *cpptraj* programot alkalmaztuk, az egyes szerkezetekkel számított három trajektóriát egyben értékeltük ki. A *hbond* parancs

használatával írtuk ki a hidrogénhid kölcsönhatásokat. A továbbiak során a trajektória legalább 10%-ában jelenlévő hidrogénhidakat vettük figyelembe, mint jelentős, valóban megjelenő kölcsönhatások. A fehérjék dinamikai tulajdonságainak, és az egyes szerkezetek közötti különbségek meghatározásának céljával különböző távolságméréseket is végeztünk. Továbbra is a *cpptraj* programot alkalmazva, a *distance* parancs segítségével mértük meg az egyes, fontosnak vélt aminosavak tömegközéppontjainak távolságát. A számítások során az egyes, független szimulációkat külön-külön kezeltük, az így kapott, szimulációs idő függvényében mért távolságvértékeket is külön-külön tüntettük fel. A távolságmérésekkel különböző szerkezeti tulajdonságokat szeretnénk volna vizsgálni. Vizsgáltuk az aktív hely kölcsönhatásait is, ehhez az öninaktivált konformerekkel számított szimulációkat klaszterezésnek vetettük alá. A klaszterezés után a legnagyobb mértékben jelen lévő klaszter centroidját vettük, mint a szimulációt reprezentáló szerkezetet. A kölcsönhatásokat ebben a szerkezetben vizsgáltuk a LigPlot+ szoftver (Laskowski és mtsai, 2011) segítségével, amely a fehérje-ligandum, vagy fehérjén belüli kölcsönhatások 2D diagramon történő ábrázolására alkalmas, beleértve a hidrogénhid és hidrofób kölcsönhatásokat is. Az összehasonlításhoz a korábban leírt, szubsztrát-kötött modellt vettük alapul.

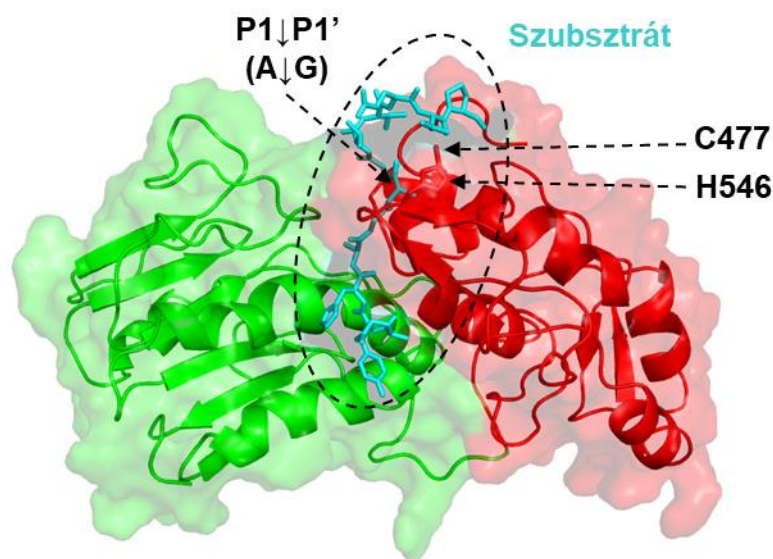
A VEEV nsP2pro szubsztrát-specifitásának vizsgálata során alkalmazott protokolhoz képest az aktív és öninaktivált konformációs állapotok molekuladinamikai vizsgálata hosszabb számítási idő, valamint *simulated annealing* protokoll alkalmazásával történt azért, hogy fel tudjuk térképezni a N-terminális kölcsönhatásait és azok dinamikus változásait. A szubsztrát-specifitás vizsgálat során célunk egy olyan enzim-szubsztrát komplex ekvibrálása volt, amelyből kiindulva különböző variánsokat tudunk vizsgálni, DynaMut szoftver alkalmazásával. A molekuladinamikai szimuláció ott lényegében az energetikai vizsgálatok készítette elő, míg az N-terminális konformációs állapotainak vizsgálata során a tanulmány leglényegesebb eleme a molekuladinamikai szimuláció eredményeinek értékelése volt.

6. EREDMÉNYEK

6.1. A VEEV nsP2pro szubsztrát-specifitásának vizsgálata

6.1.1. Az enzim-szubsztrát komplex felépítése

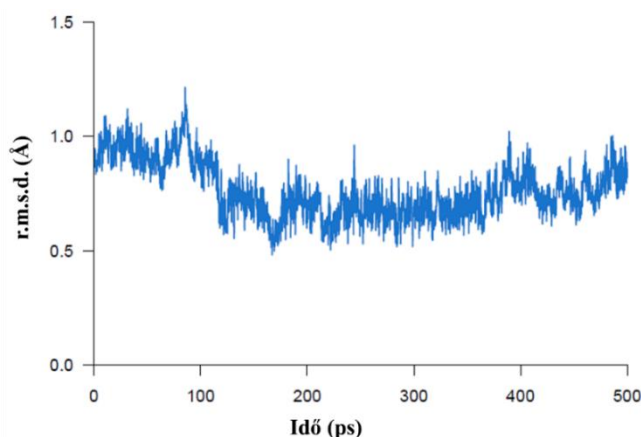
Munkánk során kutatócsoportunk célja a VEEV nsP2pro specifitásának kísérletes és számítógépes vizsgálata volt. Az enzim S5, S4, S2, S1, S1' és S2' szubsztrátkötő alhelyeinek aminosav preferenciáit olyan rekombináns fehérje szubsztrátokkal vizsgáltuk *in vitro*, melyek az SFV vírus nsP1/nsP2 természetes hasítási helyének vad típusú (EYHAGA↓G \downarrow VVETP), valamint P5, P4, P2, P1, P1' és P2' pozíciókban módosított variánsait reprezentálták. A vizsgálatokhoz összesen 31 különböző szubsztrátot hoztunk létre és használtunk fel (**Bozóki és mtsai, 2020**). Annak érdekében, hogy magyarázatot találjunk a kísérletesen meghatározott hasítási hatékonyságok és kinetikai paraméterek közötti különbségekre, elvégeztük a szerkezetek *in silico* elemzését. Ehhez felépítettem a kísérletes vizsgálatok során is alkalmazott enzim és a vad típusú SFV nsP1/nsP2 hasítóhelyet reprezentáló oligopeptid szubsztrát komplexének szerkezetét (**10. ábra**).



10. ábra. A VEEV nsP2pro enzim EYHAGA↓G \downarrow VVETP szubsztráttal alkotott komplexének molekuladinamikai szimulációval ekvibrált szerkezete. A proteáz domént vörös, a SAM MTáz domént zöld színnel, a szubsztrátot késsel jelöltük. Az ábrán jelöltük a szubsztrát hasítási helyét, a P1 és P1' pozícióban lévő alanin és glicin aminosavak között és a katalitikus diád aminosavait.

6.1.2. Az enzim-szubsztrát komplex számításos vizsgálata

Az enzim-szubsztrát komplex felépítése után elvégzett molekuladinamikai szimulációkat követően a kísérletes vizsgálatokkal megegyező pozíciókban módosítottuk a szubsztrátot. Ezt követően energetikai és szerkezeti szempontok alapján egyaránt vizsgáltuk, hogy a szubsztrát módosítása milyen hatással van az enzim-szubsztrát kölcsönhatásokra. A molekuladinamikai szimuláció utolsó szimulációs lépése, az adatgyűjtési szakasz lefuttatása során számított r.m.s.d. értéket a szimulációs idő függvényében ábrázoltuk (**11. ábra**).

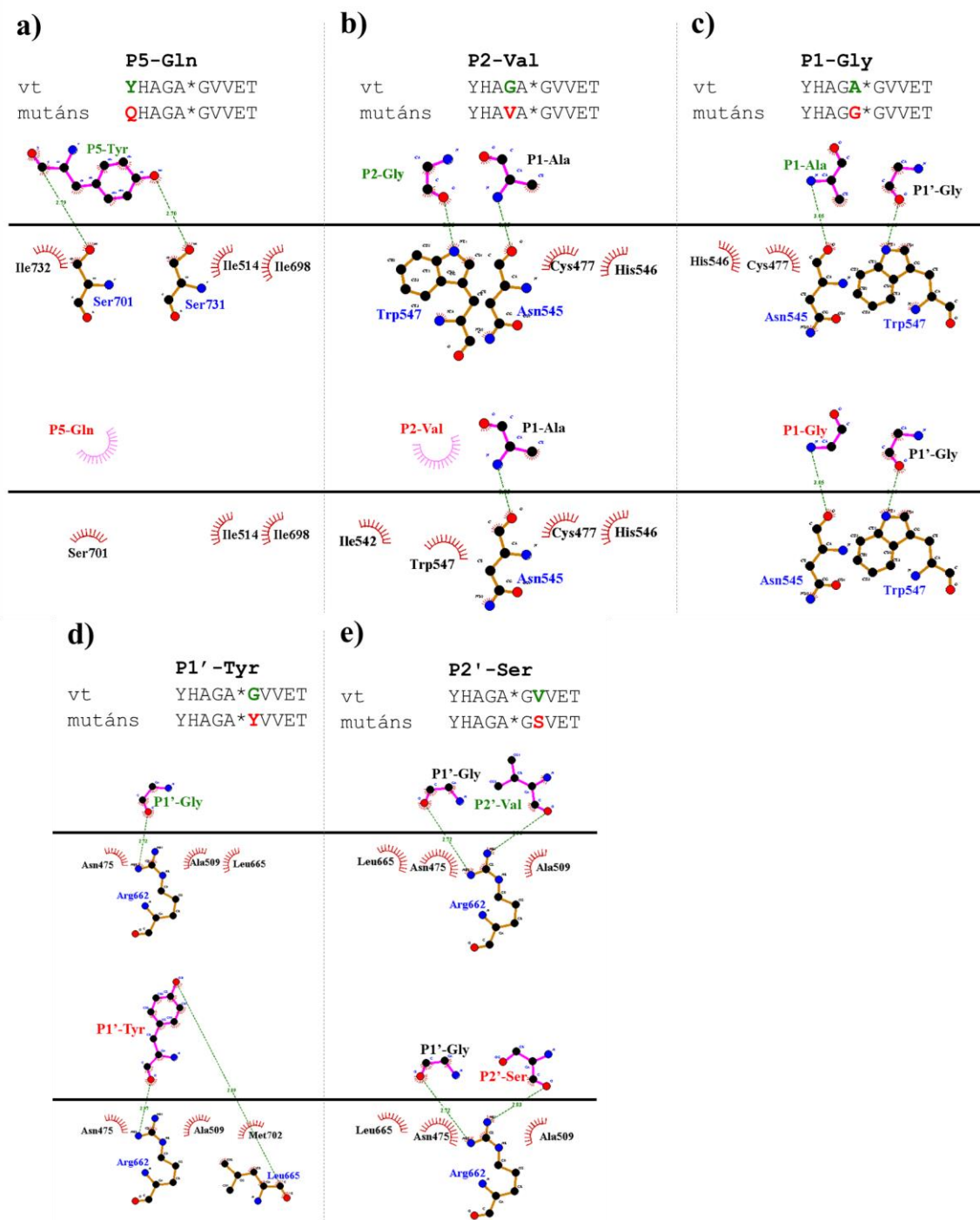


11. ábra. A VEEV nsP2pro enzim EYHAGA↓GVVETP szubsztráttal alkotott komplexének molekuladinamikai szimulációja során kapott r.m.s.d. értékek (Å) a szimulációs idő (ps) függvényében. A számítás során a szubsztrát főlánci és oldallánci atomjait egyaránt figyelembe vettük.

Mivel az r.m.s.d. nem változott jelentős mértékben a szimuláció során, a szimuláció utolsó lépésében kapott szerkezetet fogadtuk el, mint a VEEV nsP2pro enzim és az oligopeptid szubsztrát komplexének szerkezetét. Munkatársaim az egyes mutáns szubsztrát változatok esetében számos kísérleti értéket meghatároztak (Dr. Bozóki Beáta munkája), és az általam készített ekvilibrált szerkezet alapján kapott feltekeredési energia különbségét is kiszámították a DynaMut szoftver alkalmazásával (Dr. Mótyán János András munkája).

6.1.3. Az enzim-szubsztrát interakciók vizsgálata

Az általam létrehozott enzim-szubsztrát komplex szerkezetek DynaMut szoftverrel történő vizsgálata az aktivitásbeli változások atomi szintű megértéséhez járult hozzá. A szubsztrát egyes aminosavainak különböző mutáns variánsait vizsgáltuk, néhány reprezentatív szerkezet esetében kapott eredmény az alábbi ábrán látható (**12. ábra**).

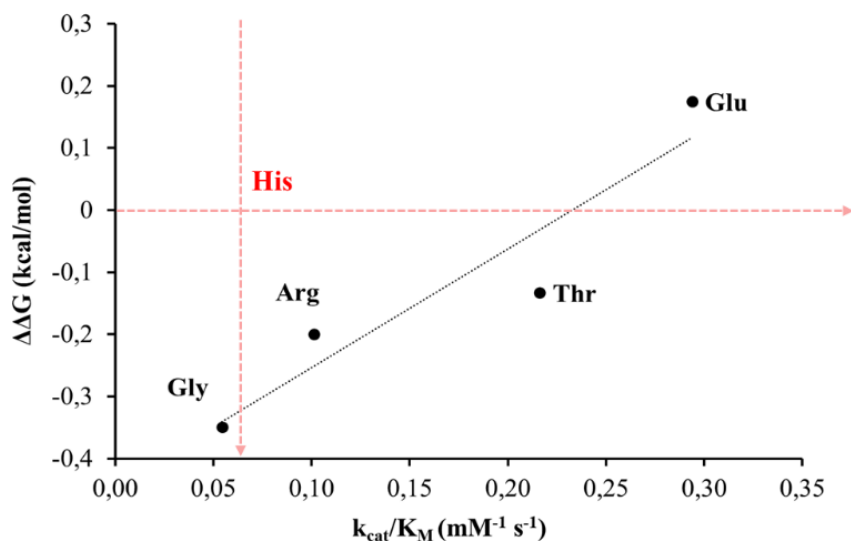


12. ábra. A vad típusú és a módosított szubsztrátok kölcsönhatásai az enzim aktív helyén.

A szubsztrátok módosításához a DynaMut webszert, míg az enzim-szubsztrát kölcsönhatások vizualizálásához a LigPlot+ v2.1 szoftver DIMPLOT modulját használtuk (Dr. Mótyán János András munkája). A panelek felső része minden esetben a vad típusú (vt), míg alsó része a mutáns szubsztrátokkal kialakított kölcsönhatásokat mutatja be. A hidrogénhíd kölcsönhatásokat zöld szaggatott vonal jelöli. A szubsztrát adott pozícióban lévő aminosavával apoláros kölcsönhatásokat kialakító aminosavakat vörös félkör jelöli. (A) P5-Gln mutáns, (B) P2-Val mutáns, (C) P1-Gly mutáns, (D) P1'-Tyr mutáns, (E) P2'-Ser mutáns.

A P5 helyen lévő tirozint glutaminra módosítottuk. A DynaMut számítások alapján a mutáció hatására az enzim-szubsztrát kölcsönhatás gyengül ($\Delta\Delta G = -0,222$ kcal/mol), a kísérletes vizsgálatok eredménye ezzel összhangban volt, a katalitikus aktivitás a mutáns szubsztrát esetében nagyon alacsonynak bizonyult a vad típusú szubsztráttal végzett reakcióhoz képest. A P5-Gln-t tartalmazó szubsztrát esetében a kevésbé hatékony *in vitro* hasítás hátterében az állhat, hogy a vad típusú tirozin két hidrogénhidat is kialakít az S5 alhelyen (oldallánc- és főlánc-mediált kölcsönhatások a Ser371 és Ser701 aminosavakkal), mely kölcsönhatásokat nem mutattuk ki a P5-Gln-mutáns szubsztrát esetében (**12A. ábra**).

A P4 helyen lévő hisztidin esetében több mutánst is megvizsgáltunk. A DynaMut-tal számított $\Delta\Delta G$ értékeket a kísérletileg meghatározott k_{cat}/K_M függvényében ábrázoltuk (**13. ábra**). A számított értékek jó korrelációt mutattak a kísérletileg meghatározott értékekkel. Valamennyi vizsgált variáns esetében a vad típusnál (P4-His) is magasabb proteolitikus aktivitás volt mérhető. Kísérletes vizsgálatainkban a P4-Glu hasítása hatékonyabbnak bizonyult a vad típusú szubsztráthoz képest, amit magyarázhat, hogy a VEEV nsP1/nsP2 és nsP2/nsP3 természetes hasítóhelyek is glutamint tartalmaznak ebben a pozícióban. A legalacsonyabb aktivitást a P4-Gly mutáns esetében tapasztaltuk, ami arra utal, hogy az S4 alhelyen az oldallánc-mediált poláris kölcsönhatások fontos szerepet töltenek be a szubsztrát-felismerésben.

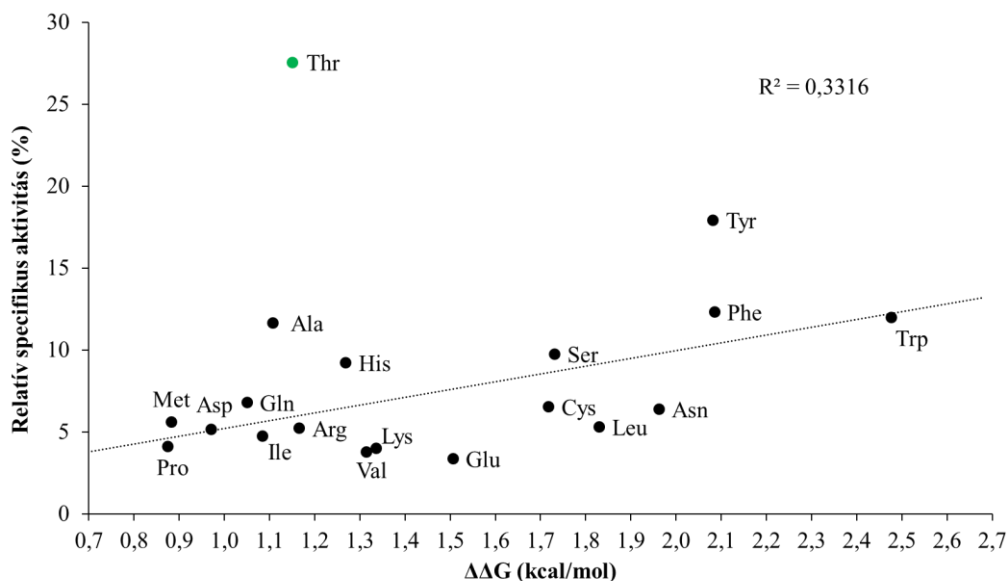


13. ábra. A P4 helyen módosított szubsztrátok esetében DynaMut-tal számított feltekeredési energia-különbség ábrázolása a katalitikus hatékonyság függvényében. Referenciaként vörössel a vad típusú szubsztrátra (P4-His) meghatározott k_{cat}/K_M értéket is feltüntettük, a $\Delta\Delta G$ értéke ebben az esetben értelemszerűen nulla.

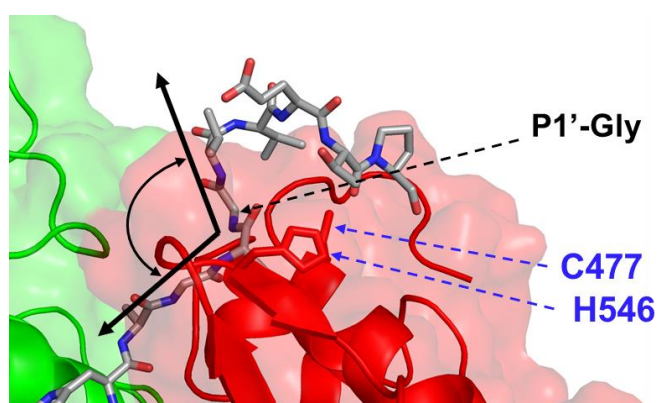
A vad típusú hasítóhely P2 pozícióban glicint tartalmaz, aminek alaninra vagy valinra módosítása esetében egy főlánci hidrogénhidat megszüntetése jár destabilizáló hatásokkal (**12B. ábra**). A DynaMut-tal számított feltekeredési energia-változás P2-Ala esetében -0,536, míg P2-Val esetében 0,495 kcal/mol volt. A kísérletes vizsgálatokban mindkét mutáns szubsztrátra a vad típusúhoz képest egyaránt jelentősen alacsonyabb aktivitás értéket határoztunk meg, ami alapján arra következettünk, hogy a glicinnél nagyobb méretű valin és alanin oldalláncok kötődése az S2 alhelyre kedvezőtlen.

A vad típusú szubsztrát esetében a P1 helyzetben előforduló alanin hidrofób kölcsönhatásokat alakít ki az S1 alhely aminosavaival. A glicinre való módosítás esetében (**12C. ábra**) mérsékelt változást jósoltunk a szerkezet vizsgálata alapján (0,188 kcal/mol), ami összhangban volt a kísérletes vizsgálatok eredményével, a vad típusú (P1-Ala) és mutáns (P1-Gly) szubsztrátok esetében meghatározott aktivitásértékek nagymértékben hasonlóak voltak. Ezzel szemben az enzim-szubsztrát kölcsönhatások jelentősebb változását jósoltuk a P1-Val mutáció esetén (1,207 kcal/mol), a kísérletes vizsgálatok során pedig ezen mutáns szubsztrátot az enzim csupán elhanyagolható mértékben hasította.

A vizsgált vad típusú hasítóhely szekvencia P1' pozícióban egy glicin aminosavat tartalmaz, amelynek főlánci hidrogénhidat tud képezni az Arg662 aminosav oldalláncával (**12D. ábra**). A kísérletes vizsgálatok során a pozíció valamennyi mutáns változatát sikerült megvizsgálni, és a meghatározott aktivitások leginkább a teljes hasítóhely hidrofobicitással mutattak korrelációt, amely összhangban van korábbi megfigyelésekkel (**Russo és mtsai, 2010**), miszerint az enzim a szubsztrát főláncával alakít ki hidrogénhidat ebben a pozícióban. Ennek ellenére úgy tűnik, hogy az enzim P1' aminosav preferenciája mégsem teljes mértékben független az oldalláncától. A vad típusú glicinen kívül a legnagyobb aktivitás értékeket poláros (Thr és Ser) vagy aromás (Tyr, Trp, Phe) oldalláncok, továbbá az apoláros Ala esetében tapasztaltuk. Ezen variánsok közül a legnagyobb aktivitást a P1'-Tyr esetében határoztuk meg, amely esetében az aromás oldallánc miatt további hidrofób kölcsönhatások is megjelenhetnek, továbbá a láncvégi hidroxil csoport egy további hidrogénhidat alakíthat ki a Leu665 főláncával. Az említett hidrofób kölcsönhatások Phe vagy Trp oldallánc esetében is jelen lehetnek, ami megmagyarázza a tapasztalt jelentősebb aktivitást. A számított feltekeredési energia-különbségek ($\Delta\Delta G$) és relatív specifikus aktivitás értékek (%) között nem volt megfigyelhető jelentő korreláció (**14. ábra**).



14. ábra. A P1' helyen módosított szubsztrátok esetében kísérletileg meghatározott relatív specifikus aktivitás értékek ábrázolása a DynaMut-tal számított feltekeredési energia-különbségek függvényében. Zölddel kiemelve a P1'-Thr aminosav. A vad típusú szubsztrátra (P1-Gly) meghatározott aktivitás értéket tekintettük 100%-nak.



15. ábra. A VEEV nsP2pro aktív helyére kötődött oligopeptid szubsztrát konformációja. Balra és fölfelé mutató nyilakkal jelöltük a P1'-Gly helyzethez képest N- és C-terminális irányba elhelyezkedő szubsztrát aminosavakat. A P1' pozícióban egy „törés” figyelhető meg a szerkezetben. A katalitikus aminosavakat kék nyilak mutatják.

Az alacsony korreláció oka feltételezhetően a szubsztrát szerkezetének relatíve nagyobb mértékű megváltozása. A P1' pozícióban a vad típusú szubsztrátban egy „törés” látható (**15. ábra**), ezt a konformációs állapotot jelentősen megváltoztathatja egy nagyobb méretű oldallánc bevezetése, amely nem kedvez egy ilyen szerkezet kialakulásának. A nagymértékű szerkezeti

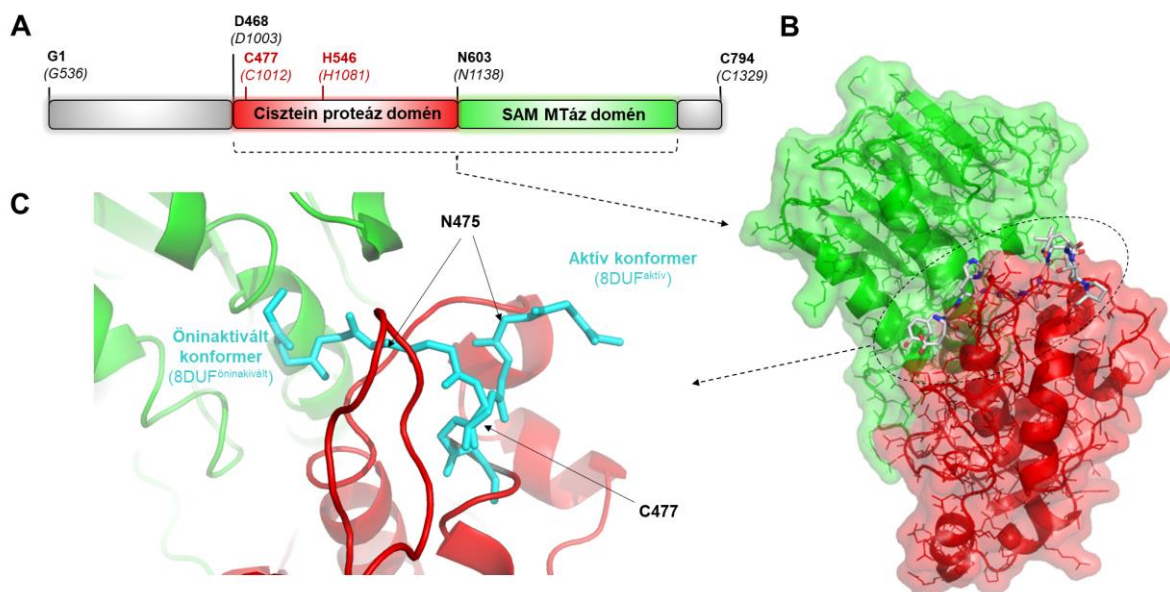
változásokat követő energiaváltozásokat a DynaMut relatíve kisebb pontossággal képes prediktálni (Rodrigues és mtsai, 2018).

A kísérletes vizsgálatok során kollégáim az SFV nsP1/nsP2 vad típusú hasítóhelyének valin aminosavát szerinre vagy prolinra mutálták. A vad típusú SFV nsP1/nsP2 hasítási helyre jellemző valin szerinre történő módosítása nem mutatott jelentős különbségeket a kinetikai paraméterekben, a k_{cat}/K_M értéke mindkét esetben $0,06 \text{ mM}^{-1}\text{s}^{-1}$ értéknek adódott. Az eredmények összeegyeztethetők azzal a ténnyel, hogy a P2' oldallánc az oldószer irányába fordul. Ebben a pozícióban leginkább a szubsztrát főlánca alakít ki interakciót az enzimmal (12E. ábra). A számított feltekeredési energia-különbség $-0,068 \text{ kcal/mol}$ volt, ami nagyon kis különbséget mutat. A P2'-Pro esetében számított energia $0,427 \text{ kcal/mol}$ volt. A vad típusú, valamint a P2'-Pro és P2'-Ser variáns szubsztrát esetében a kísérletesen meghatározott specifikus aktivitás értéke nagymértékben hasonlóan bizonyult. Az S2' alhely esetében számított és mért értékek azzal magyarázhatók, hogy az S4-S1' alhelyekhez képest az S2' alhely kevésbé jól definiált, az oldószer felé nagyobb mértékben nyitott.

6.2. A VEEV nsP2pro kristályszerkezetének meghatározása és elemzése

A VEEV nsP2pro szerkezetét kollaborációs partnereink röntgendiffrakcióval határozták meg (Hoffka és mtsai, 2023). A szerkezet meghatározásához előállított rekombináns fehérje a VEEV nsP2 Arg463-Thr785 régióját tartalmazta. Az enzim SAM MTáz doménje a megfelelő diffrakció és nagy felbontás - a felületi entrópia csökkentése révén történő - elérése két felületi mutációt is tartalmazott (K741A és K767A). Ezek a módosított lizin aminosavak nem vesznek részt az aktív hely felépítésében és a ligandum-kötésben sem. A cisztein proteáz domén nem tartalmazott módosításokat.

A vizsgálatok eredménye alapján az Arg463-Thr785 rekombináns fehérje szerkezetét csaknem teljes hosszában sikerült feloldani (Asn472-Thr785), azonban a Tyr764-Tyr774 hurok részét képező Arg769-Tyr774 aminosavak nem voltak láthatóak az elektronsűrűségi térképen, feltehetően a régió relatíve nagyobb mértékű flexibilitása miatt. A meghatározott szerkezet felbontása relatíve nagy, $1,46 \text{ \AA}$ volt. Meglepő módon, az elektronsűrűségi térkép alapján meghatározott kristályszerkezetben (PDB ID: 8DUF) az enzim N-terminális régiójának két eltérő konformációja volt megfigyelhető ($8DUF^{\text{aktív}}$ és $8DUF^{\text{öninaktív}}$) (16. ábra).

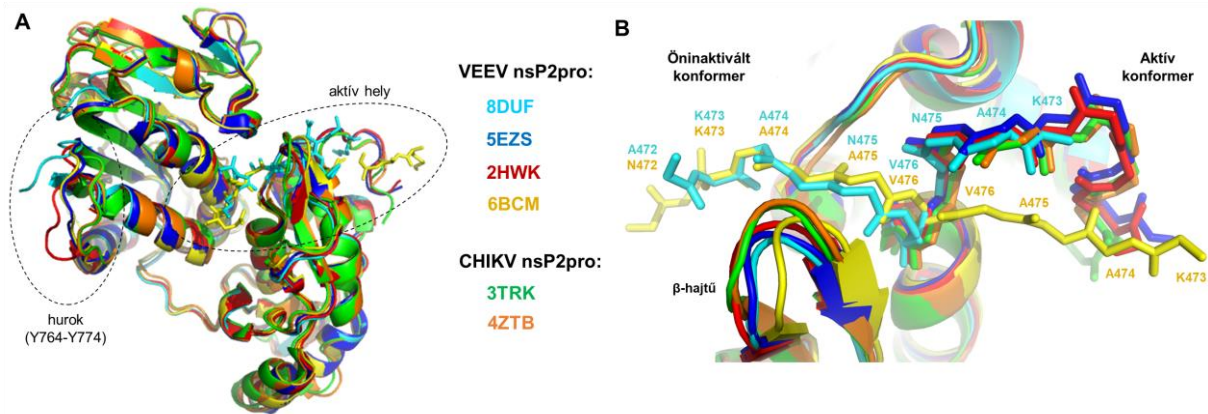


16. ábra: A VEEV nsP2pro szerkezete. (A) A kristályszerkezet meghatározásához használt VEEV nsP2pro szekvencia a teljes nsP2 fehérje szekvenciában. A számozás során az nsP2 számozás mellett zárójelben a teljes nsP1234 szekvenciát jellemző számozást is feltüntettük. A proteáz szekvenciát vörös, a SAM MTáz domént zöld színnel jelöltük. (B) A VEEV nsP2pro kristályszerkezete, bekarikázva az N-terminális régió látható. (C) Az N-terminális régió kék színnel vannak megjelenítve, az aktív és öninaktivált konformer esetében egyaránt.

A $8DUF^{aktív}$ konformer esetében az elektronsűrűségi térkép alapján a Lys473 aminosavtól kezdődően volt látható szerkezet, míg a $8DUF^{öninaktivált}$ konformer esetében az Asn472-től kiindulva. Az N-terminálison minden aminosav rendelkezett alternatív konformerrel, a Cys477-tel bezáróan. Az N-terminális szakasz a $8DUF^{aktív}$ konformer esetében az oldószer irányába fordulva jelent meg, míg a $8DUF^{öninaktivált}$ konformer esetén az aktív hely irányába fordulva volt látható, a domének közötti részbe kötődve, közvetlenül a β -hajtú alatt (**16C. ábra**). Az öninaktivált konformerben az N-terminális szakasz az S5-S1' szubsztrátkötő helyeket foglalja el, ezért katalitikusan inaktívnak tekinthető, lévén az aktív helyet elfoglalva megakadályozza a szubsztrát bekötődését. Ez megfelelt a korábban az N475A aktív hely mutáns enzim esetében megfigyelt konformációs állapotnak ($6BCM^{öninaktivált}$) (**Compton és mtsai, 2017**), bár érdekes módon korábban ezt az öninaktivált állapotot nem figyelték meg a nem-módosított aktív helyet tartalmazó enzim esetében ($2HWK^{aktív}$) (**Russo és mtsai, 2006**).

Az ismert VEEV nsP2pro és a CHIKV nsP2pro szerkezetek összehasonlítása alapján megállapítottuk, hogy az aktív konformáció esetén az N-terminális szakasz az oldószer irányába fordulva látható. Ez a szakasz jól illeszkedik az egyes szerkezetekben. Érdekes módon az N-

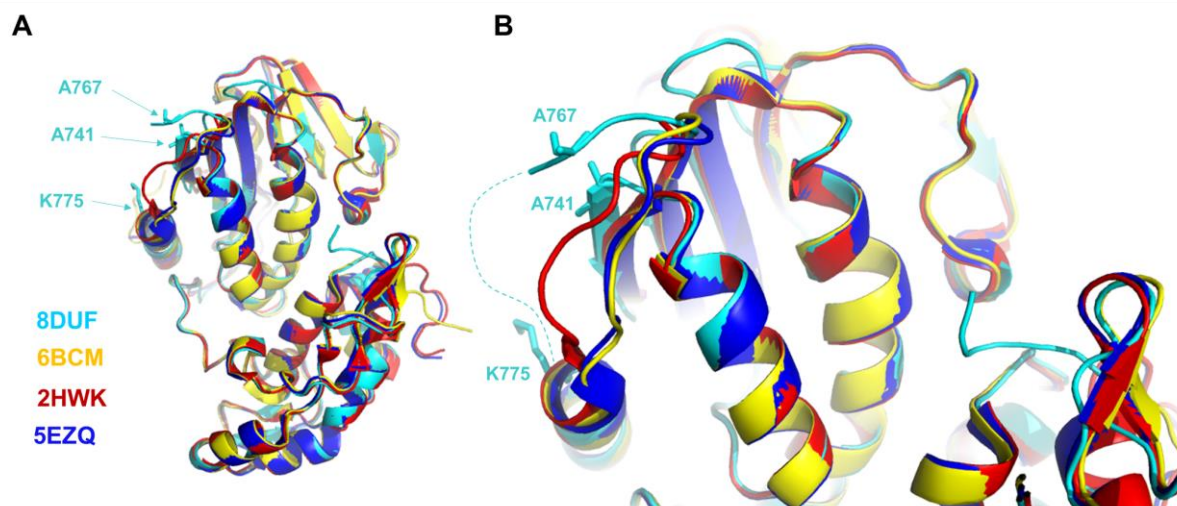
terminális konformációja eltérést mutatott abban az esetben, ha az enzimben jelen volt az N475A mutáció is (6BCM^{aktív}), feltehetően a mutáció hatásának következtében változik meg az N-terminális konformációja az oldószer felé nyitott állapotban (**17. ábra**).



17. ábra: Alfavírus nsP2pro szerkezetek összehasonlítása. (A) A VEEV és CHIKV nsP2pro szerkezeteinek illesztése. Az egyes szerkezetek színezése az ábrán szereplő színek szerint történt. Az aktív hely és a Tyr764-Tyr774 felszíni hurok bekarikázva szerepel az ábrán. (B) Az N-terminális régiók konformációjának összehasonlítása az aktív és öninaktivált konformerek esetében. A színezés az A panelével egyezik, az N-terminális régió aminosavai (472-477) pálcika formátumban láthatóak. Az ábrát Dr. Mótyán János András készítette.

A 8DUF kristályszerkezetben nem jelentek meg az Arg769-Tyr774 aminosavak a C-terminális régióban, illetve különbségek figyelhetők meg ezen régió esetében a SAM MTáz domén mutációit nem tartalmazó szerkezetekhez képest (PDB ID: 5EVS, 2HWK, 6BCM), ami arra utal, hogy a Tyr764-Tyr774 hurok flexibilisebbé vált a mutációk következtében (**18. ábra**).

Korábbi tanulmányok megállapították, hogy a CHIKV nsP2pro esetében a katalitikus His548 aminosavat tartalmazó β-hajtú a proteáz és SAM MTáz domének közötti távolságot csökkentve meg tudja akadályozni a szubsztrát bekötődését (Narwal és mtsai, 2018), és ehhez hasonló szerkezetet láttak az N475A VEEV nsP2pro mutáns esetében is (Compton és mtsai, 2017). A kutatásaink keretében meghatározott kristályszerkezetben (PDB ID: 8DUF) azonban nem találtunk kisebb távolságot, ami arra utal, hogy a N475A mutációnak feltehetően nincs szerepe az öninaktivált forma kialakulásában.



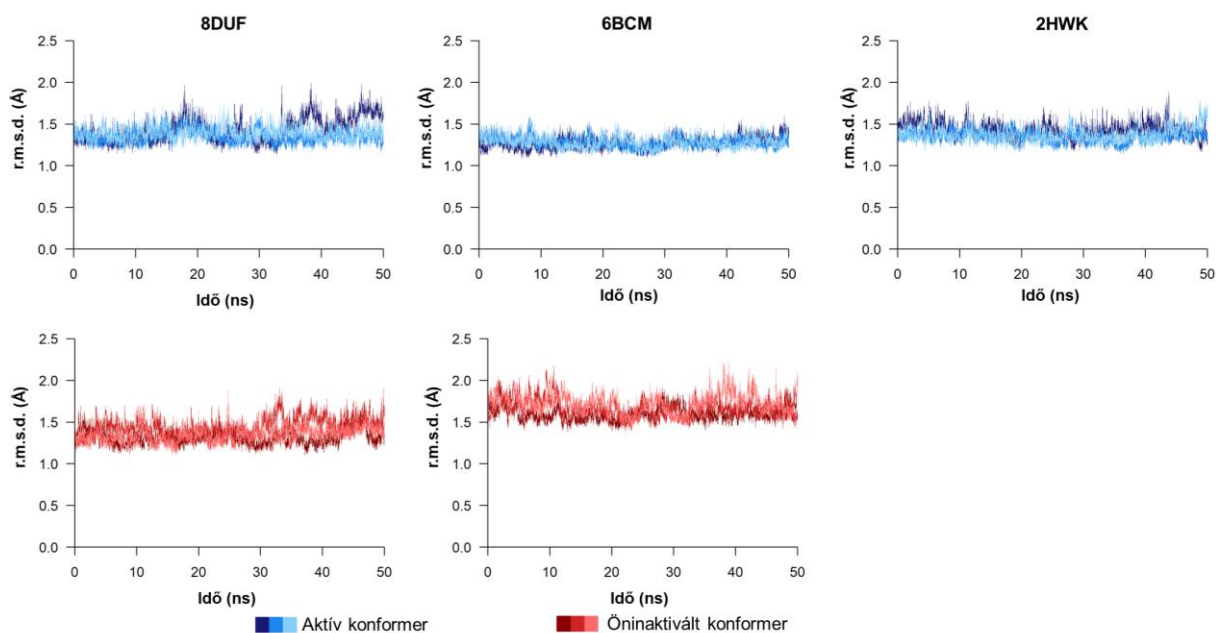
18. ábra: A VEEV nsP2pro SAM MTáz doménjében található Tyr764-Tyr774 hurok konformációjának összehasonlítása különböző kristályszerkezetekben. (A) Az egymásra illesztett nsP2pro szerkezetek esetében az egyes láncok a PDB azonosítók jelmagyarázat szerinti színezését alapul véve vannak megjelenítve. A Lys741, Lys767 és Lys775 aminosavakat nyilak jelölik **(B)** A Tyr764-Tyr774 hurok szerkezete. A kék szaggatott vonal az elektronsűrűségi térképben nem látható aminosavakat (Arg769-Tyr774) jelöli a 8DUF szerkezet esetében. Az ábrát Dr. Mótyán János András készítette.

6.3. A VEEV nsP2pro kristályszerkezetek molekuladinamikai vizsgálata

6.3.1. A molekuladinamikai szimulációk ellenőrzése

Molekuladinamikai szimulációkkal vizsgáltuk a vad típusú ((PDB ID: 2HWK), valamint a K741A/K767A (PDB ID: 8DUF) és az N475A mutációt tartalmazó (PDB ID: 6BCM) VEEV nsP2pro fehérjék szerkezeteit. Célunk volt, hogy megismerjük az öninaktivált konformer megjelenéséért felelős kölcsönhatásokat, továbbá összehasonlítsuk a vad típusú Asn475 aminosavat valamint az N475A mutációt tartalmazó szerkezeteket. Első lépésben a szimulációk minőségét ellenőriztük.

A protokoll „*simulated annealing*” szakasza során az egyes lépésekben a szerkezetek geometriáját vizualizációval ellenőriztük. Az egyensúlyi szakasz utáni adatgyűjtési szakasz számítások lefutását az r.m.s.d. értékek kiszámításával ellenőriztük, amelynek referenciáját minden fehérje esetében a teljes trajektóriából minden tizedik mintavételezett szerkezet atomi pozícióinak átlagából számított szerkezetként adtuk meg **(19. ábra)**.

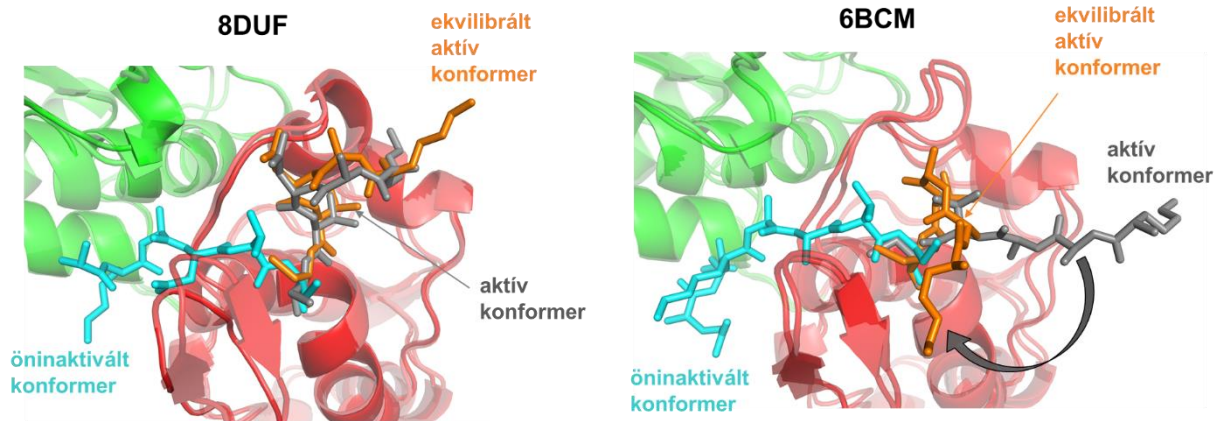


19. ábra: A molekuladinamikai számítások során kapott trajektóriák alapján számított r.m.s.d. értékek. A kiszámított r.m.s.d. értékek (Å) a 8DUF, 2HWK és 6BCM szerkezetek esetében vannak feltüntetve. A számítás során mind a főlánci, mind az oldallánci atomokat figyelembe vettük. Az eltérő színezés az eltérő konformációs állapotokat jelöli (kék – aktív konformer, vörös – öninaktivált konformer), az eltérő árnyalatok a három független számítás során kapott eredményeket jelölik.

Az r.m.s.d. értékek mind 1,5 Å körül fluktuáltak az aktív és az öninaktivált szerkezetek esetében egyaránt. Fontos azonban megemlíteni, hogy az értékek az N475A mutáns öninaktivált szerkezetében ($6BCM^{\text{öninaktivált}}$) rendre kissé magasabbak voltak az aktív konformációhoz képest ($6BCM^{\text{aktív}}$), és a görbék szemmel láthatóan is nagyobb mértékű fluktuációt mutattak. Ez arra utalt, hogy az aktív hely mutáns öninaktivált szerkezete flexibilisebb, dinamikusabb az aktív konformernél.

A három párhuzamos adatgyűjtési szakasz utolsó lépésében kapott szerkezeteket összehasonlítottuk egymással és a kiindulási szerkezetekkel is. Az aktív konformerek esetében a kiindulási szerkezethez képest legnagyobb eltérést mutató szerkezeteket a **20. ábrán** mutatjuk be. Az összehasonlítás alapján megfigyelhető, hogy a $8DUF^{\text{MD-aktív}}$ szerkezet esetében nem történt jelentő változás a kiindulási állapothoz képest, az N-terminális konformációja az adatgyűjtési szakasz végén a legnagyobb különbséget mutató szerkezetben is nagymértékben hasonló volt a kiindulási szerkezetben megfigyelthez. Ezzel ellentétben a $6BCM^{\text{MD-aktív}}$ szerkezet esetében olyan végállapotot is megfigyeltünk, melyben az N-terminális szakasz az aktív és az öninaktivált konformerekre jellemző állapotok közötti, „átmeneti” szerkezetet vette

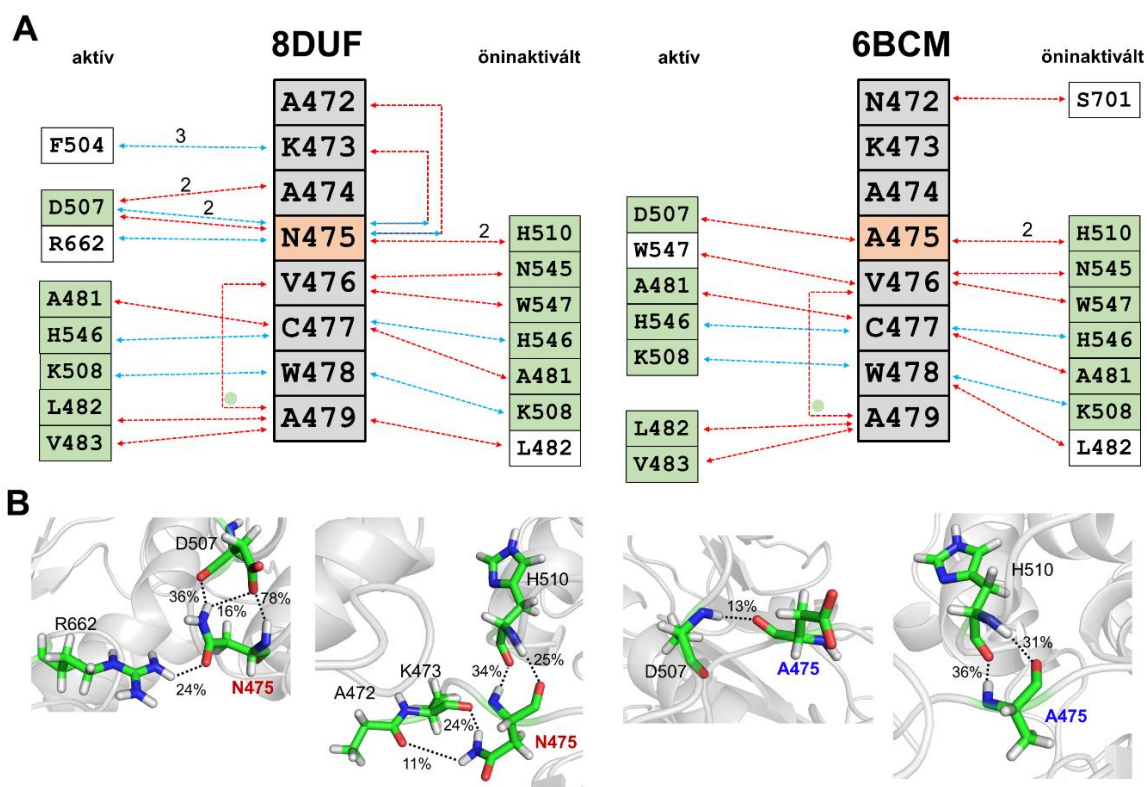
fel. Eredményeink alapján ugyan nem történt teljes átalakulás a konformerek között egyik vizsgált szerkezet egyik konformere esetében sem a szimulációk alatt, a megfigyeléseink mégis arra utalnak, hogy az N475A mutáció hatására bekövetkezhet az aktív állapot öninaktiválttá alakulása.



20. ábra: Az N-terminális konformációs állapotainak összehasonlítása a kiindulási és ekvilibrált szerkezetekben. Az N-terminális kiindulási konformációs állapotait az aktív és öninaktivált konformerek esetében szürke és kék színek jelölik, az ekvilibrált szerkezetben az N-terminálisnak a kiindulási állapothoz képest legnagyobb változást mutató végállapotait narancssárga színűek, az N-terminális elmozdulását szürke nyíl jelöli.

6.3.2. A VEEV nsP2pro N-terminálisának interakciói

Összehasonlító vizsgálataink egyik legfontosabb pontja az N-terminális szakasz intramolekuláris kapcsolati hálózatának feltérképezése volt, az aktív helyen módosított (N745A mutáns, $6BCM^{MD-aktív}$ és $6BCM^{MD-öinaktivált}$) és a vad típusú aminosavat tartalmazó (N475, $8DUF^{aktív}$ és $8DUF^{öinaktivált}$) nsP2pro esetében. Az N-terminális régió aktív helyen kialakított kölcsönhatásainak elemzésével lényeges információkat tudhattunk meg mind a konformerek közötti eltérésekről, mind pedig az N475A mutáció hatásáról (**21. ábra**).



21. ábra. Az N-terminális aminosavak intermolekuláris kölcsönhatásai. (A) A trajektóriák több, mint 10%-ában megjelenő kölcsönhatások nyíllal ábrázolva. Az N-terminális aminosavakat szürke háttér előtt láthatjuk, a 475. pozícióban lévő aminosavakat narancssárga színnel emeltük ki. Az N-terminális által kialakított interakciókat nyilak jelölik, a vörös és a kék nyilak a főlánc ill. az oldallánc által alkotott hidrogénhidakat. A nyíl felett számmal jelöltük, hogy adott aminosav mennyi azonos típusú hidrogénhidat alakít ki. Zöld színezéssel, ill. az N-terminális belső interakciói esetében zöld körrel jelöltük ki a mindkét szerkezetben (8DUF és 6BCM) előforduló interakciókat, páronként összehasonlítva mind az aktív, mind az öninaktivált konformereket. **(B)** A 475. pozícióban lévő aminosav által kialakított hidrogénhid kölcsönhatások szerkezeti ábrája, valamennyi vizsgált szerkezet esetében. Szaggatott vonallal ábrázoltuk a trajektóriák legalább 10%-ában megjelenő kölcsönhatásokat és az előfordulási gyakoriságot (%) is feltüntettük.

A trajektóriában megjelenő hidrogénhidak nem mutattak jelentős különbségeket az N475A mutáns (PDB ID: 6BCM) és a vad típusú Asn475 aminosavval rendelkező K741A/K767A mutáns (PDB ID: 8DUF) aktív és öninaktivált konformerei között. Az ellentétes töltéssel rendelkező katalitikus aminosavak (Cys477 és His546) hidrogénhid kölcsönhatást alakítottak ki valamennyi trajektóriában. Fontos kiemelni a 475. aminosavat jellemző különbségeket. Az általunk meghatározott szerkezetben, a vad típusú Asn475

jelenlétében az aktív konformer négy hidrogénhid kölcsönhatást alakított ki (8DUF^{MD-aktív}). Ezek közül két oldallánc és egy főlánc által kialakított hidrogénhid az 507. pozícióban lévő Asp aminosavval jelent meg, míg egy oldallánc által kialakított hidrogénhid kapcsolódik a 662. helyzetben lévő Arg aminosavhoz. Az oldallánc által kialakított hidrogénhidak lehetősége az aszparagin aminosav szerkezetéből adódik: az amid csoportban lévő NH₂ hidrogénjei jelentős elektronhiánnyal, míg ugyanezen csoport oxigén atomja jelentős elektrontöbblettel rendelkezik, amelyek elősegítik a hidrogénhid kölcsönhatások kialakulását. Azonban az oldallánc flexibilitását figyelembe véve nem meglepő, hogy ezek a kölcsönhatások minden esetben csak a trajektória kevesebb, mint felében jelentek meg.

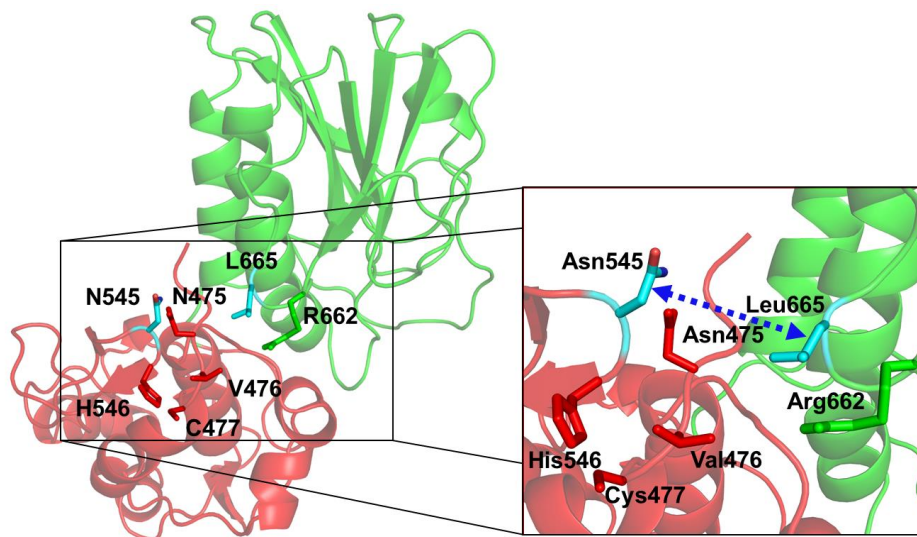
Ezzel szemben az említett főlánc által kialakított hidrogénhid a trajektóriák 78%-ában volt jelen, ami erősebb kölcsönhatásra utal. Az általunk meghatározott szerkezet esetében az öninaktivált forma (8DUF^{öninaktivált}) egészen más interakciós hálózatot mutat: Asn475 oldallánca hidrogénhidakat alakít ki az N-terminálison található 472-es és 473-as pozícióban lévő aminosavakkal. Bár ezek megjelenése nem jelentős (rendre 11% és 24%), mindez az N-terminális erősen fixált, belsőleg rendezett szerkezetére utal. Érdeemes megjegyezni, hogy a Lys473 és Ala474 további hidrogénhid kölcsönhatásokat alakít ki a fehérje más aminosavaival is. Ezen felül Asn475 ezen kívül két hidrogénhidat alakít a His510 aminosavval. Ezek jelentősebb mértékben vannak jelen a szimulációk során (25% és 34%), utalva a szerkezet kialakulásában betöltött jelentősebb szerepükre.

Az N475A mutációval rendelkező szerkezet aktív konformerét vizsgálva (6BCM^{aktív}) egyértelműen nem kell számítani az oldallánc által kialakított hidrogénhid kölcsönhatásokra, lévén a metil csoport nem képes kialakítani ilyeneket. Azonban ennél jelentősebb különbségeket is megfigyelhetünk: az N475A mutáns aminosav és Asp507 közötti oldallánc hidrogénhid jóval kisebb mértékben jelenik meg, mint a vad típusú Asn475 esetében (13%, 78% helyett) ami egy jóval gyengébb interakcióra utal. Továbbá sem Lys473, sem Ala474 nem alakít ki hidrogénhidakat egy másik aminosavval sem. Mindez a szerkezet kisebb stabilitására utal. Az N475A mutáns öninaktivált konformerében (6BCM^{öninaktivált}) viszont nem látszik, hogy az oldallánc hidrogénhidak megszűnése N475A főlánci hidrogénhidjait jelentősen befolyásolta volna, megjelenésük százalékos aránya közel esik az Asn475 esetében kapott értékekhez. Összefoglalva elmondható, a 475-ös helyzetben lévő aszparagin oldallánca több hidrogénhid kialakítására is képes, amelyek szerepe jelentős lehet mind az aktív, mind az öninaktivált konformerek stabilizálásában is. Az N475A mutációt követően ezek megszűnése jelentős különbségeket okoz az N-terminális régió hidrogénhid kölcsönhatási hálózatában, és ez az aktív konformerben még inkább jelentősebb.

Feltételeztük, hogy az öninaktivált konformáció kialakulása hatással lehet a domének egymáshoz viszonyított helyzetére és ezáltal az aktív hely szerkezetére és hozzáférhetőségére, ezért a továbbiakban megvizsgáltunk több, a doménon belül vagy a domének között kialakuló kölcsönhatást is és meghatároztuk az egyes aminosavak közti távolságokat.

6.3.3. A β -hajtú és Asn545 aminosav

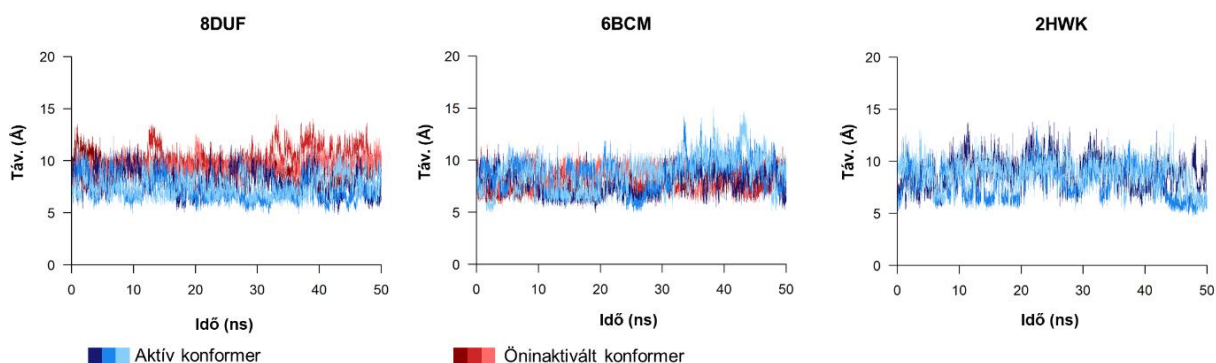
Az Asn545 aminosav a cisztein proteáz doménben lévő egyik hurokban, az ún. β -hajtúben helyezkedik el, mely hajtú alacsony relatív B faktor (másnéven *atomic displacement parameter*, ADP) értéket mutatott mindkét öninaktivált konformer esetében a kristályszerkezetekben. Az aktív és öninaktivált konformereket azonban a molekuladinamikai szimulációk trajektóriáiban is megvizsgáltuk, célunk az interdomén távolság változásának követése volt, amelyhez az Asn545 és Leu665 aminosavak tömegközéppontjának távolságát mértük meg (22. ábra).



22. ábra: Az Asn545 és Leu665 aminosavak helyzete a VEEV nsP2pro szerkezetében. A dinamikai tulajdonságok vizsgálata szempontjából lényeges aminosavak pálcika formátummal vannak megjelenítve, az Asn545 és Leu665 kézzel színezve láthatók. A kék nyíl a vizsgált távolságot jelöli, a meghatározott értékek változása a trajektória során a 23. ábrán kerül bemutatásra. Az ábra a 8DUF^{aktív} szerkezet alapján készült.

Az Asn545 és Leu665 aminosavak a VEEV nsP2pro proteáz és SAM MTáz doménjei kölcsönhatási felszínén helyezkednek el, így feltételeztük, hogy a köztük lévő távolság esetleges megváltozása ezek egymáshoz viszonyított távolságának megváltozására utalhat. A távolság hasonló mértékű fluktuációit figyeltük meg mind az aktív, mind az öninaktivált konformerekben (23. ábra).

A trajektóriákban számított távolság valamennyi esetben 5-10 Å között volt. A β -hajtú flexibilitása az aktív konformerek esetében mind a vad típusú, mind a N475A mutációt tartalmazó szerkezetben is közel azonos értéket mutatott, ami arra utalt, hogy az N475A mutáció nem változtatja jelentősen ennek a huroknak a mobilitását.

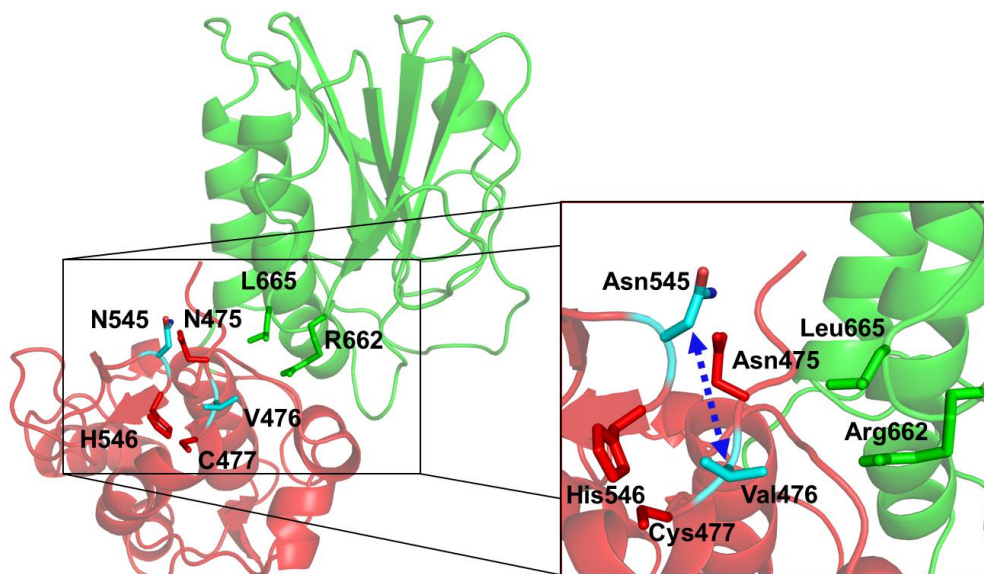


23. ábra: Az interdomén távolság megváltozása a trajektóriákban. Az Asn545 és Leu665 aminosavak tömegközéppontjainak távolsága (Å) a trajektóriákban, a 8DUF, 2HWK és 6BCM szerkezetek esetében. Az aktív konformerekre jellemző értékeket kézzel, az öninaktiváltak esetében kapottakat vörössel jelöltük. A különböző árnyalatok az eltérő szimulációkban meghatározott értékeket jelölik.

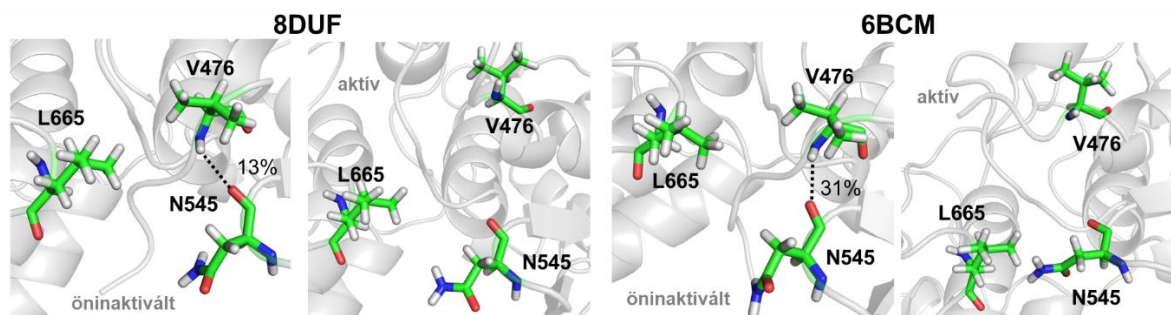
6.3.4. Az Asn545 és Val476 közötti kölcsönhatás

A β -hajtú hurok mozgását korábban megfigyelték az N475A mutációt tartalmazó szerkezetben (Compton és mtsai, 2017), megállapítva, hogy az N545-O és V476-NH atomok közötti hidrogénhid kölcsönhatás felelhet a β -hajtú relatív alacsonyabb flexibilitásért az öninaktivált konformerben (24. ábra). Ezt az interakciót számításaink során is megfigyeltük, a 8DUF^{öninaktivált} kristályszerkezetből indított szimulációk során Asn545 főlánci oxigénje és Val476 főláncja között hidrogénhid volt megfigyelhető. Ezt a megfigyelést megerősíti a N-terminális intramolekuláris kapcsolatainak vizsgálata is, ugyanis az öninaktivált konformerben a terminális interakciókat alakít ki β -hajtúval (Asn545-tel), azonban az aktív konformerben nem figyelhető meg ilyen kölcsönhatás.

A molekuladinamikai szimulációk alapján továbbá az Asn545 és Val476 közötti hidrogénhid kölcsönhatás (24. ábra) még nagyobb mértékben volt jelen az N475A mutáns öninaktivált szerkezetben. A 475. pozícióban vad típusú aminosavat tartalmazó szerkezetben (Asn 475) a hidrogénhid a trajektória 13%-ában jelent meg, míg az N475A mutáns esetében 31%-ban volt jelen, ami relatívan erősebb kölcsönhatásra utal (25. ábra).

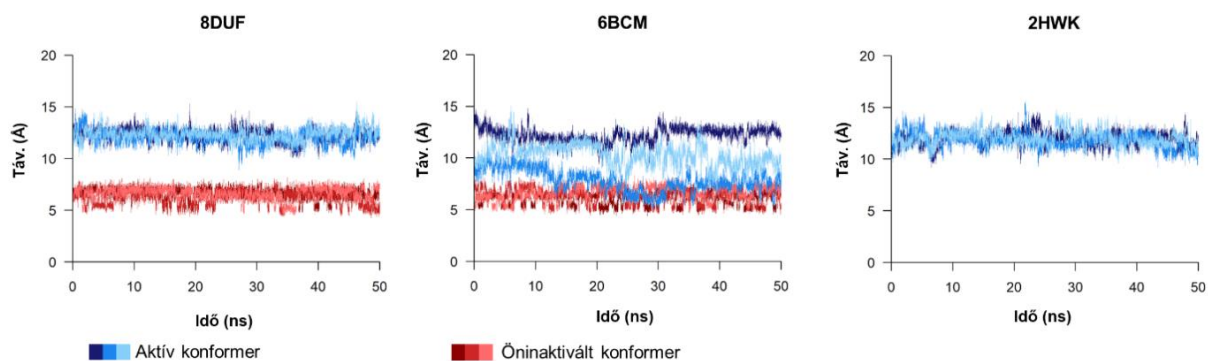


24. ábra: Az Asn545 és Val476 aminosavak helyzete a VEEV nsP2pro szerkezetében. A dinamikai tulajdonságok vizsgálata szempontjából lényeges aminosavak pálcica formátummal vannak megjelenítve, az Asn545 és Val476 kézzel színezve láthatók, a kék nyíl a két aminosav közt vizsgált távolságot jelöli. Az ábra a 8DUF^{aktív} szerkezet alapján készült.



25. ábra: Az Asn545 és Val476 aminosavak interakciója. A két aminosav főlánci atomjai között megfigyelhető hidrogénhid kölcsönhatás százalékos megjelenési aránya van feltüntetve minden vizsgált szerkezet esetében. A vizsgált távolságvértékek változása a trajektória során a **26. ábrán** kerül bemutatásra.

A hidrogénhid megjelenésének vizsgálata mellett a konformációs változásokat alaposabban is tanulmányoztuk és kiszámítottuk a trajektóriákban az Asn545 és Val476 tömegközéppontjainak távolságát is (**26. ábra**).



26. ábra: Az Asn545-Val476 interakció vizsgálata. A Val476 és Asn545 aminosavak tömegközéppontjainak távolsága (Å) a trajektóriákban. Az aktív konformerekre jellemző értékeket kézzel, az öninaktiváltak esetében kapottakat vörössel jelöltük. A különböző árnyalatok az eltérő szimulációkban meghatározott értékeket jelölik.

A 8DUF szerkezetben az aktív és az öninaktivált konformer esetében kapott távolságértékek jól elkülönültek, az aktív konformer esetében a távolság stabilan 10-15 Å közé esett. Az öninaktivált konformerre kapott távolságérték jelentősen alacsonyabb, kb. 5-8 Å határok között fluktuált. A 2HWK^{MD-aktív} számítás esetében kapott értékek jó egyezést mutattak a 8DUF^{MD-aktív} esetében számított értékkel. Ehhez képest jelentős különbséget tapasztaltunk az N475A mutáns vizsgálatokor. A 6BCM^{MD-öninaktivált} és a 8DUF^{MD-öninaktivált} szerkezetének vizsgálatokor kapott értékek hasonlóak voltak, ami az öninaktivált konformerek közötti strukturális hasonlóságra utal. Azonban az N475A mutáns enzim esetében (6BCM^{MD-aktív}) számított értékek több trajektóriában is nagy fluktuációt mutattak az ebben a pozícióban nem módosított enzimekhez képest (2HWK^{MD-aktív} és 8DUF^{MD-aktív}), és a mért távolság elérte az öninaktivált konformerre jellemző értékeket is. Ezek a nagymértékű fluktuációk az N475A mutációnak az aktív szerkezetet esetlegesen destabilizáló hatását mutatták, azaz egy jóval flexibilisebb, kevésbé jól definiált aktív konformerre utaltak. A távolságérték csökkenésének következménye lehet továbbá az is, hogy a mutánsban az aktív konformer könnyebben alakulhat át az öninaktivált konformerré.

6.3.5. Az aktív hely kölcsönhatásai

Az általunk meghatározott nsP2pro kristályszerkezetben - mely a K741A és K767A felszíni mutációkat is tartalmazta - megjelent öninaktivált konformer (8DUF^{öninaktivált}) a korábban meghatározott N475A mutáns öninaktivált konformerének feleltethető meg (6BCM^{öninaktivált}), amelyben az N-terminális az aktív hely irányába fordul és a ⁴⁷²AKANVC⁴⁷⁷ aminosavak kötődnek az aktív helyre. A bekötési mód hasonlóságát a korábban az N475A

mutációt tartalmazó 6BCM^{öninaktivált} konformerrel már vizuális megfigyeléssel is láthattuk (**16. ábra**). Mindkét esetben a Cys477 aminosav foglalja el a S1' alhelyet, még az N-terminális további aminosavai (472-476) az S5-S1' alhelyekre kötődnek. Az öninaktivált konformer esetében az enzim N-terminális régiója az aktív helyre kötődik be, a természetes szubsztráthoz hasonló módon, azért ezt a szakaszt tulajdonképpen pszeudo-szubsztrátnak tekinthetjük ezekben a konformerekben. Azzal a céllal, hogy összehasonlítsuk a vad típusú Asn475 és az N475A mutánsok esetében látott öninaktivált szerkezetek kölcsönhatási hálózatát, feltérképeztük a pszeudo-szubsztrát által az aktív helyen kialakított legfontosabb kölcsönhatásokat (intramolekuláris kölcsönhatások), amit összehasonlítottunk egy oligopeptid szubsztrát által kialakított interakciókkal (intermolekuláris kölcsönhatások) (**2. táblázat**). Ennek alapján fontos következtetéseket vonhattunk le a szubsztrát bekötődésért valamint az öninaktivált konformer kialakulásáért felelős kölcsönhatások különbségeivel és esetleges hasonlóságaival kapcsolatban.

Az oligopeptid szubsztrát és az enzim közt az S2-S7 helyeken kialakuló kölcsönhatások vizsgálatához a VEEV nsP2pro és a VEEV vad típusú nsP1/nsP2 hasítóhelyet reprezentáció oligopeptid szubsztrát modell szerkezetét használtuk fel, melyet Dr. Patricia M. Legler bocsájtott rendelkezésünkre (**Hu és mtsai, 2016**). Azért döntöttünk ezen szubsztrát vizsgálata mellett, mert ez a VEEV poliproteinben lévő természetes hasítóhely szekvenciát reprezentálja, amit a proteáz *in vivo* is hasít a poliprotein processzálása során, ezért valós enzim-szubsztrát kölcsönhatásokról kaphattunk információt. Ezzel szemben a specificitás szerkezet-alapú vizsgálata során az enzimnek egy olyan modell szubsztráttal alkotott komplexét vizsgáltuk, mely egy másik alfavírus, az SFV természetes hasítóhelyét reprezentálta (nsP1/nsP2). Ezt az indokolta, hogy ezt a modell szubsztrátot alkalmaztuk az aminosav-preferenciák *in vitro* vizsgálatában is, így lehetőségünk volt az *in vitro* vizsgált kölcsönhatásokat *in silico* is tanulmányozni. Annak érdekében, hogy a pszeudo- és az oligopeptid szubsztrátok kötődését össze tudjuk hasonlítani, a peptid szubsztrát P6-P1' aminosavainak (LQEAGA↓G) kölcsönhatásait térképeztük fel. A szerkezetet vizsgálva azt láttuk, hogy összességében relatíve alacsony az enzim és a szubsztrát között kialakuló hidrogénhidak száma, ami egyezik a korábban leírt megállapításokkal (**Russo és mtsai, 2010**), miszerint a szubsztrát aktív helyhez kötődéséért leginkább a P4-P1' aminosavak által kialakított van der Waals interakciók felelnek.

Az inter- és intramolekuláris kölcsönhatások összehasonlítása alapján megállapíthatjuk, hogy az öninaktivált konformerek esetében a kölcsönhatások között nincsenek jelentős különbségek a vad típusú és az N475A mutáns enzim esetében. Fontos megjegyezni, hogy az N475A mutáns esetében valamivel több apoláros kölcsönhatás megjelenését figyeltük meg.

Mindez a konformerek nagymértékű hasonlóságára utal. A pszeudo- és az oligopeptid szubsztrátok kötődése alapvető hasonlóságokat mutat, mindkét „szubsztrát” aminosavai az S5-S1’ alhelyeket foglalják el, és jelentős hasonlóság látható kölcsönhatásaik között is: az S4-S1 alhelyek mindegyike esetében legalább egy aminosav részt vesz a peptid szubsztrát vagy az N-terminális P4-P1 aminosavainak kötésében (**2. táblázat**). A hasonlóságok mellett azonban számottevő különbségek is láthatóak az inter- és intramolekuláris kölcsönhatások hálózatában, amelynek esetlegesen szerepe lehet egy esetleges inhibitor tervezése esetén.

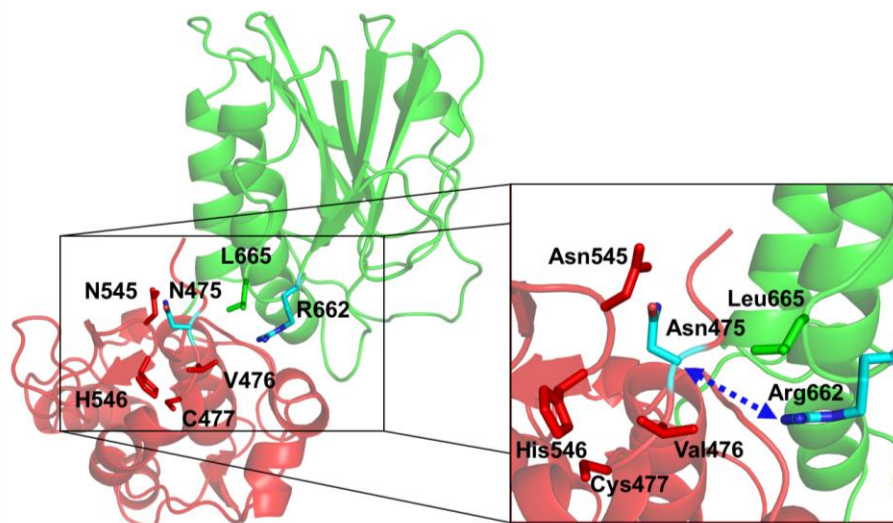
2. táblázat. A VEEV nsP2pro aktív helyén lévő aminosavak interakciói peptid és pszeudo-szubsztráttal. A hidrogénhidakat kialakító aminosavakat aláhúzással jelöltük, zárójelben a hidrogénhidak számát is feltüntettük. A hidrofób kölcsönhatásokat az enzim-szubsztrát komplex modellje, ill. az öninaktivált szerkezetek esetében a szimulációk eredményeként kapott, klaszterezett szerkezetek alapján határoztuk meg. A hidrofób kölcsönhatások megadásához a LigPlot+ szoftvert (**Laskowski és mtsai, 2011**) használtuk. *A szubsztrátkötő alhelyeket alkotó aminosavak a VEEV nsP2pro és a EAGA↓G oligopeptid által alkotott komplex szerkezete alapján lettek meghatározva (**Russo és mtsai, 2010**). **Az enzim-szubsztrát interakciókat a VEEV nsP2pro és az LQEAGA↓GSVETP oligopeptid alkotott komplex modell szerkezete alapján határoztuk meg (**Hu és mtsai, 2016**).

Kötőhely aminosavai *	Szubsztrát aminosav**	Enzim-szubsztrát kölcsönhatások **	N-terminális aminosav (8DUF / 6BCM)	8DUF ^{MD-öninaktivált}	6BCM ^{MD-öninaktivált}
-	P6-Leu1	<u>S511</u> , <u>E513</u> , I542	- / Q471	-	-
-	P5-Gln2	I514, I698, <u>S701</u> , A728, I732	A472 / N472	M702	I514, I698, <u>S701</u> , A728, S731, I732
S511, E513, W547, M702, <u>K706</u> *	P4-Glu3	H510, S511, W547, I698, S701, M702, <u>K706</u>	K473 / K473	I542, W547	S511, I542, W547
H510, I542, W547, I698, M702	P3-Ala4	A509, H510, I542, D664, I698, M702	A474 / A474	A509, H510, D664, I698, M702	A509, H510, D664, M701
W478, C477, A509, <u>H510</u> , W547	P2-Gly5	W478, A509, <u>H510</u> , N545, W547	N475 / A475	<u>A472</u> , <u>K473</u> , A509, <u>H510 (2)</u> , I542, N545, W547	A509, <u>H510 (2)</u> , I542, N545, H546, W547
N475, C477, <u>W478</u> , A509, <u>N545</u> , H546, L665	P1-Ala6	C477, W478, <u>N545</u> , H546, W547	V476 / V476	<u>N545</u> , H546, <u>W547</u>	N545, H546, <u>W547</u>
A474, N475, C477, K480, <u>H546</u>	P1'-Gly7	N475, C477, <u>R662 (2)</u>	C477 / C477	A479, K480, <u>A481</u> , <u>H546</u>	A479, K480, <u>A481</u> , <u>H546</u> , W547
			W478 / W478	K480, L482, F504, <u>K508</u> , H510, W547, V615	A479, K480, A481, <u>L482</u> , V501, D502, Y503, F504, <u>K508</u>

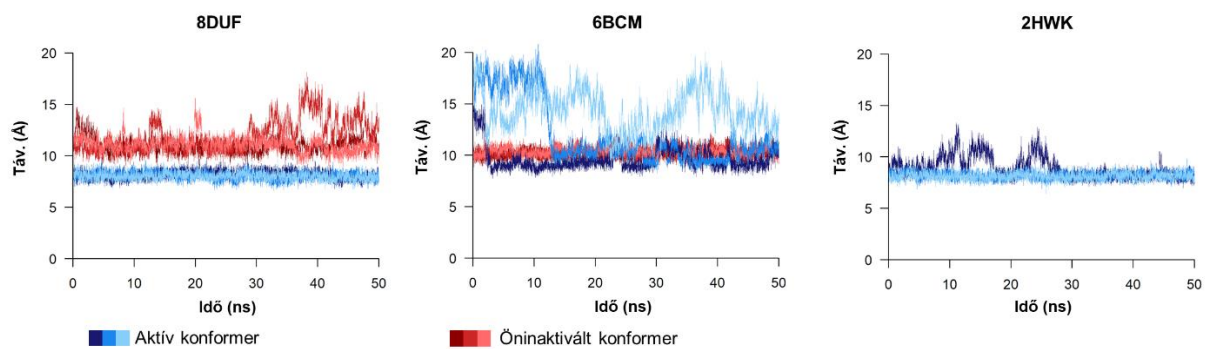
6.3.6. Az aktív hely hozzáférhetősége

Korábbi molekuladinamikai szimulációk során megállapították, hogy a VEEV nsP3/nsP4 hasítóhelyet reprezentáló szubsztrát (RFDAGA↓YIFSSD, amely a P6-P6' aminosavakat tartalmazza) a számítások során jelentős konformációs változásokon mehet keresztül, és a P1'-P6' helyen lévő aminosavak nem csak a proteáz domén, hanem az SAM MTáz domén felé is fordulhatnak (Hu és mtsai, 2016). Feltételezhető, hogy az enzim konformációs flexibilitásának jelentős szerepe van a szubsztrát bekötődésében, a flexibilitás ugyanis összefüggésben lehet az aktív hely hozzáférhetőségével, amely kulcsfontosságú a szubsztrát, vagy az N-terminális bekötődéséhez. Egy korábbi tanulmány arra utalt (Hu és mtsai, 2016), hogy az Asn475 és Arg662 aminosavak közötti interakciónak jelentős szerepe lehet a SAM MTáz és a proteáz domének komplexének stabilizálásában (27. ábra).

Mindezeket figyelembe véve a két aminosav tömegközéppontjának távolságát megmértük szimulációink során is azzal a céllal, hogy információkat kapjunk a két domén közötti interakciók erősségéről és a rendszerek flexibilitásáról.



27. ábra: Az Asn475 és Arg662 aminosavak helyzete a VEEV nsP2pro szerkezetében. A dinamikai tulajdonságok vizsgálata szempontjából lényeges aminosavak pálcika formátummal vannak megjelenítve, az Asn475 és Arg662 kézzel színezve láthatók. A kék nyíl a vizsgált távolságot jelöli, a meghatározott értékek változása a trajektória során a 28. ábrán kerül bemutatásra. Az ábra a 8DUF^{aktív} szerkezet alapján készült.

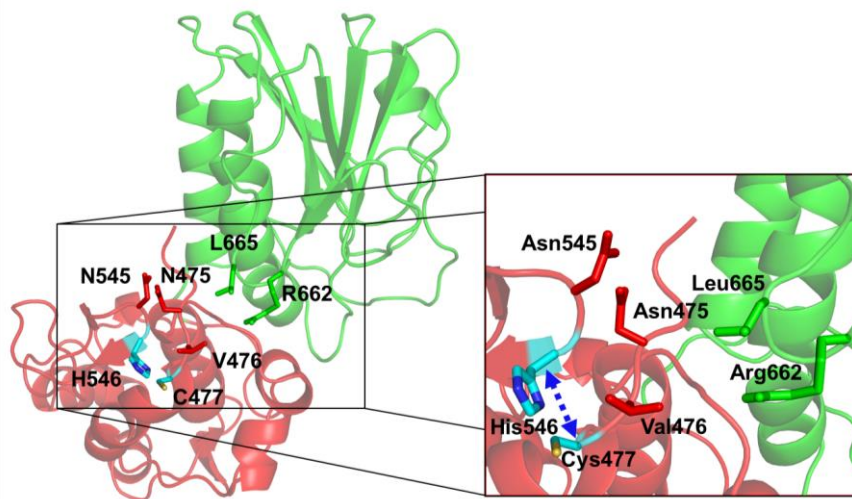


28. Ábra: Az Asn475-Arg662 aminosavak távolsága. Az Asn475/Ala475 és Arg66 aminosavak tömegközéppontjai között meghatározott távolság (Å) a trajektóriákban. Az aktív konformerekre jellemző értékeket késsel, az öninaktiváltak esetében kapottakat vörössel jelöltük. A különböző árnyalatok az eltérő szimulációkban meghatározott értékeket jelölik.

Nem jelentős mértékben, de szisztematikusan a vad típusú Asn475 aminosavat tartalmazó szerkezetek esetében (8DUF és 2HWK) a két konformer között eltérés volt tapasztalható, az öninaktivált szerkezetre valamivel magasabb távolságérték volt jellemző minden szimuláció során (**28. ábra**). Az N475A mutáció esetében azonban az aktív konformerre (6BCM^{MD-aktív}) jellemző érték volt nagyobb, elérte a 8DUF^{MD-öninaktivált} esetében tapasztalt értékeket és kiugró fluktuációkat mutatott. Ez a kristályszerkezetből nem volt egyértelműen megállapítható (PDB ID: 6BCM, **Compton és mtsai, 2017**), az N475A mutáció nem módosította a mutáció az Arg662 konformációját, de szimulációink megmutatták, hogy hatására csökkent a két domén közötti interakció erőssége. Eredményeink arra utalnak, hogy az Asn475 és Arg662 aminosavak közötti interakcióknak szerepe lehet a domének stabilitásában.

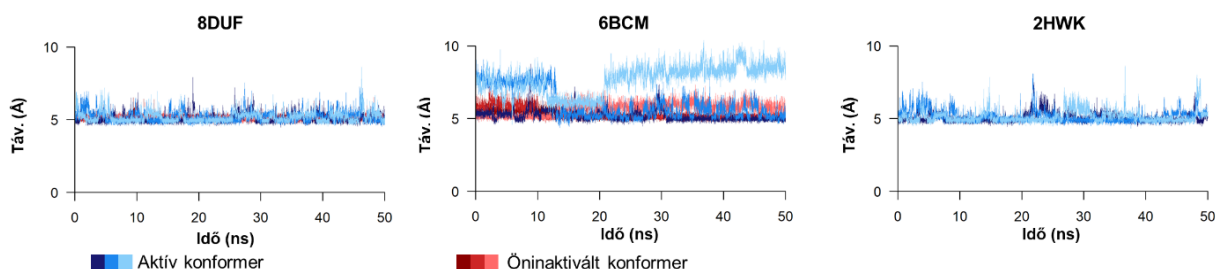
6.3.7. A katalitikus diád stabilitása

Egy enzim működéséhez létfontosságú a reakció lefutását elősegítő, megfelelően rendezett aktív hely, ezért feltételezhető, hogy a katalitikus aminosavak közötti távolság fluktuálása negatívan befolyásolja az enzim működését. Ezért megvizsgáltuk a katalitikus diád aminosavai közötti távolságot is annak meghatározására, hogy az aktív helyet érintő mutáció (N475A) hatással van-e a hidrolízisért felelős aminosavak egymáshoz viszonyított helyzetére.



29. ábra: A Cys477 és His546 aminosavak helyzete a VEEV nsP2pro szerkezetében. A dinamikai tulajdonságok vizsgálata szempontjából lényeges aminosavak pálcika formátummal vannak megjelenítve, Cys477 és His546 késsel színezve látható. A kék nyíl a vizsgált távolságot jelöli, a meghatározott értékek változása a trajektória során a **30. ábrán** kerül bemutatásra. Az ábra 8DUF^{aktív} szerkezet alapján készült.

A Cys477 és His546 katalitikus aminosavak közti távolság (**29. ábra**) az aktív helyen a vad típusú Asn475-öt tartalmazó szerkezet esetében (8DUF és 2HWK) stabilan 5 Å-höz közeli értéket mutatott, csupán kisebb mértékű fluktuációkkal (**30. ábra**). A N475A mutáns esetében azonban a távolságérték fluktuációja jóval nagyobb volt az aktív konformer esetében (6BCM^{MD-aktív}), egyes szimulációkban elérte akár a 10 Å értéket is, ami egy jóval kevésbé stabil aktív hely konformációra utalt. Ennek rendkívül jelentős szerepe lehet az enzim aktivitása szempontjából: a destabilizálódott aktív hely nem feltétlenül képes ellátni feladatát, azaz a peptidkötés hidrolízisét.



30. ábra: Az aktív hely stabilitása. A katalitikus diád Cys477 és His546 aminosavainak tömegközéppontjai között meghatározott távolság (Å) a trajektóriákban. Az aktív konformerekre jellemző értékeket késsel, az öninaktiváltak esetében kapottakat vörössel jelöltük. A különböző árnyalatok az eltérő szimulációkban meghatározott értékeket jelölik.

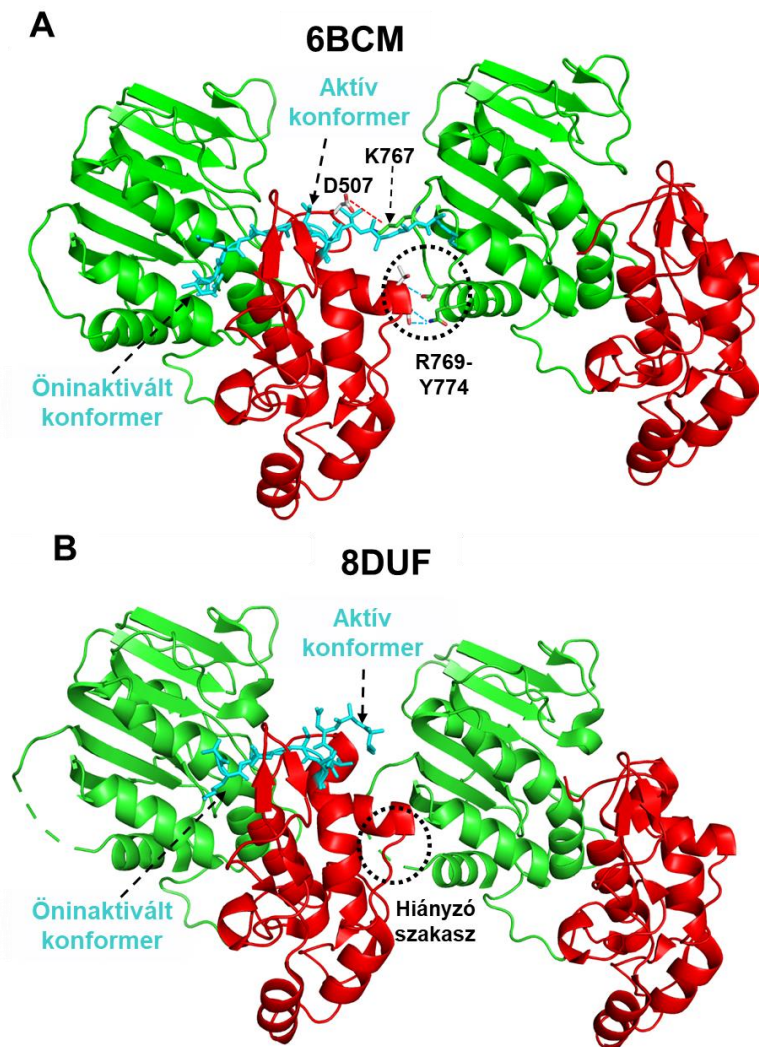
Az Asn475 aminosav által kialakított kölcsönhatások N475A mutáció következtében történő megszűnése a SAM MTáz és a proteáz domének között olyan konformációs változásokat okozhat, amely hatással lehetnek a szubsztrát bekötődésére is. Az Asn545 és Val476 közötti interakciók korábbiak során bemutatott megváltozása is erre utal (**25. és 26. ábra**). A SAM MTáz és proteáz domének között megnövekedett távolság elviekben segíthetné a szubsztrát és az N-terminális aktív helyre való bekötődését is. Azonban az N475A mutánsban az aktív hely konformációja sem megfelelő, amint az láttuk (**30. ábra**), a katalitikus diád távolsága kimagasló fluktuációkat mutat. Így az N-terminális bekötődésének valószínűsége nagyobb lehet (lévén azzal nincs reakcióba lépésre lehetőség). Elmondható tehát, hogy az N475A mutánsban nagyobb lehet az öninaktivált konformer kialakulásának valószínűsége, és az aktív hely kedvezőtlen megváltozása negatív hatással lehet az nsP2pro aktivitására. Ezzel szemben az aktív helyen a vad típusú Asn475 aminosavat - de a felszínen K741A és K767A mutációkat - tartalmazó enzim esetében a molekuladinamikai vizsgálatok alapján az öninaktivált konformer az, amely kevésbé stabil.

6.3.8. A K741A és K767A felszíni mutációk hatása

Az általunk vizsgált VEEV nsP2pro esetében a kristályszerkezet meghatározásához - a felszíni entrópia csökkentésének céljából - kerültek bevezetésre a K741A és K767A felszíni mutációk. Szerkezeti vizsgálataink eredményei alapján nem jelenthető ki egyértelműen, hogy a K741A és K767A mutációknak van-e közvetlen szerepe az öninaktivált forma kialakulásában. Ezek a mutációk egyik korábban vizsgált VEEV nsP2pro szerkezetében sem szerepeltek, és vizsgálataink esetében is csak a kristályszerkezet meghatározása érdekében kerültek beépítésre. Az elektronsűrűségi térkép alapján nem sikerült meghatározni annak a felszíni Tyr764-Tyr774 hurok egy részének a szerkezetét (Arg769-Tyr774, lásd korábban, **18. ábra**). Ennek oka feltételezhetően az, hogy a mutációk hatással vannak a hurok dinamikai tulajdonságaira, megnövelik flexibilitását. Bár a Lys741 és a Lys767 aminosavak az aktív helytől távol helyezkednek el és nem vesznek részt a ligandum kötésében, feltételezzük, hogy módosításuk - bizonyos allosztérikus kölcsönhatások megváltoztatása - miatt mégis hozzájárulhat az öninaktivált forma megjelenéséhez vagy annak stabilizálásához. Ezt a feltételezést támasztja alá az a megfigyelés is, hogy az öninaktivált konformer megjelenését korábban nem tapasztalták a vad típusú enzim esetében, mely sem az aktív helyen, sem pedig az enzim felszínén nem tartalmazott mutációkat (PDB ID: 2HWK) (**Russo és mtsai, 2006**).

A kapott eredmények értelmezéséhez és a további kutatások tervezéséhez egyaránt fontos lehet figyelembe venni, hogy milyen hatással lehetnek a Tyr764-Tyr774 hurokrész

stabilitására a kristálykontaktusok, azaz a kristályban lévő fehérjék között kialakuló kölcsönhatások. Ennek érdekében összehasonlítottuk az elemi cellákat a 6BCM és a 8DUF szerkezetek esetében (31. ábra).



31. ábra: A Tyr764-Tyr774 hurok stabilitásáért felelős kristálykontaktusok. Kristálykontaktusok bemutatása (A) az N475A aktív hely mutáns és (B) a K741A és K767A felszíni mutációkat tartalmazó enzimek esetében. A színezés a korábbi ábrákkal egyező, a monomerek elrendezése a pdb fájlokban elérhető elemi cella paraméterek alapján történt. A szaggatott vonallal bekarikázott rész az Arg769-Tyr774 hurokszakaszt, ill. ezeket a hiányzó aminosavakat jelöli. Az ábrán piros szaggatott vonal jelöli a Lys767-Asp507 sóhidat, kék szaggatott vonalak az Arg769-Tyr774 aminosavak által kialakított intermolekuláris hidrogénhidakat.

Megfigyelhető, hogy az aktív hely mutáns esetében (PDB ID: 6BCM) a felszíni hurkot több kölcsönhatás is stabilizálja a kristályszerkezetben, így hidrogénhidak, valamint egy sóhíd a Lys767 és a szomszédos monomer Asp507 oldallánca között. Ezek az intermolekuláris kölcsönhatások nem figyelhetők meg a felszíni mutációkat tartalmazó szerkezet esetében (PDB ID: 8DUF), a K767A mutáció következtében nem alakulhat ki sóhíd az Asp507-tel, az Arg769-Tyr774 szakasz pedig nem látható a szerkezetben, így kölcsönhatásai sem vizsgálhatók. Továbbá, a két szerkezetben a két monomer relatív orientációja is eltérő. Lehetséges, hogy a Tyr764-Tyr774 hurok flexibilitásának megnövekedéséért elsősorban a sóhíd kialakításáért felelős aminosav (Lys767) mutációja okozza, és nem a K741A és a K767A mutációk együttes hatása. Az elemi cellát megvizsgáltuk a 2HWK esetében is, a monomerek relatív orientációja ebben az esetben inkább a 8DUF szerkezetre emlékeztet, a Lys767 a szomszédos monomer Glu505 oldalláncával alakít ki sóhidat. A monomerek orientációja azokban a szerkezetekben hasonló, ahol a 475. pozícióban a vad típusú aszparagin található, sóhidak viszont csak azokban a szerkezetekben megfigyelhetőek, ahol az ionos kölcsönhatás kialakítására képes bázikus oldallánc található meg a 767. pozícióban.

7. MEGBESZÉLÉS

Munkám során azokba a kutatásokba kapcsolódtam be, melyeknek célja a VEEV nsP2pro szubsztrát-specifitásának kísérletes eljárásokkal történő vizsgálata volt, ehhez kapcsolódóan végeztem el az enzimnek a kísérletes aktivitásmérésekben is alkalmazott szubsztráttal alkotott komplexének *in silico* módszerekkel történő vizsgálatát. A modell komplex elkészítéséhez egy másik kutatócsoport által készített (és Patricia M. Legler kutatócsoportja által a rendelkezésünkre bocsájtott) komplexben lévő oligopeptidet úgy módosítottam, hogy az megfeleljen az *in vitro* aktivitásvizsgálatokban is alkalmazott és az SFV nsP1/nsP2 hasítóhelyét reprezentáló (EYHAGA↓GVVETP) szekvenciának. Ezután molekuladinamikai módszert alkalmazva a megfelelően előkészített szerkezetet explicit oldószeres szimuláció keretében vizsgáltam (**10. ábra**), a szimuláció megfelelő konvergenciáját az r.m.s.d. értékek kiszámításával ellenőriztem (**11. ábra**). Az ekvilibráció utolsó lépésben kapott szerkezetet ezután DynaMut szoftver segítségével módosítottuk az aktív helyen lévő szubsztrát megfelelő pozíciókban történő mutációjával, majd vizsgáltuk a számított feltekeredési energia-különbségeket és az enzim-szubsztrát kölcsönhatások megváltozását is nyomon követtük.

A kapott szerkezetek hozzájárultak, hogy molekuláris szinten is értelmezni tudjuk a kísérletesen meghatározott szubsztrát-specifitást (**12. ábra**). Eredményeink alapján a szubsztrát P5 helyen lévő Tyr aminosavának Gln-ra való módosítása kedvezőtlen hatású, a vad típusú szubsztrát esetében azonosított hidrogénhid kölcsönhatás megszűnése miatt (**12A. ábra**). A P4 helyzetben előforduló hisztidin módosítása esetében a kísérletesen mért k_{cat}/K_M és a számított feltekeredési energia-eltérések korrelációját állapítottuk meg (**13. ábra**). A P4-Glu hasítása hatékonyabb volt, mint a vad típusú P4-His szubsztráté, ami összhangban van azzal, hogy a VEEV nsP1/nsP2 és nsP2/nsP3 hasítóhelyei is glutamint tartalmaznak P4 pozícióban, továbbá a korábbi vizsgálatok alapján az S4 alhelyen elsősorban a poláris enzim-szubsztrát kölcsönhatások a dominánsak és az S4-Lys706 oldallánc sóhidat képes kialakítani a P4-Glu oldallánccal (**Russo és mtsai, 2010; Hu és mtsai, 2016**). A vizsgált mutánsok közül a P4-Gly esetében tapasztaltuk a legalacsonyabb aktivitást, ami az oldallánc-mediált poláris kölcsönhatások fontosságára utal. A P2 helyzetben lévő glicin alaninra vagy valinra cseréje pedig egy főlánci hidrogénhid megszűnésével járt (**12B. ábra**), ez destabilizáló hatású lehet, a kísérletes mérések mindkét mutáns esetében a vad típusnál alacsonyabb aktivitást határoztak meg, ami a nagyobb méretű oldalláncok kedvezőtlen kötődésére utal. A vad típusú szubsztrát P1 pozíciójában található alanin hidrofób kölcsönhatásokat tud kialakítani az S1 alhellyel,

glicinre mutálása (**12C. ábra**) a számított energiaértékek alapján kismértékű változást okoz. Ez összhangban volt a kísérletes méréseink és korábbi kutatások eredményeivel is, melyek alapján az elágazó apoláros oldalláncok kötődése az S1 alhelyre kevésbé kedvező, az elágazó oldalláncok alacsonyabb mértékű flexibilitása miatt (**Russo és mtsai, 2010**). Az SFV nsP1/nsP2 hasítóhely minden lehetséges P1'-mutáns változatát megvizsgáltuk, a mért aktivitások leginkább a teljes hasítóhely hidrofobicitásával mutattak kismértékű korrelációt. Saját megfigyeléseink alapján - és az irodalmi adatokkal is összhangban (**Russo és mtsai, 2010**) - az enzim elsősorban a szubsztrát P1' aminosavak főláncával alakít ki hidrogénhidakat. Ennek ellenére a hasítási hatékonyság az oldalláncoktól nem teljes mértékben független, a vad típusú glicin mellett magas aktivitást tapasztaltunk aromás (Tyr, Trp, Phe) és poláros (Thr, Ser) oldalláncok esetében. Az alfavírusok valamennyi természetes nsP1/nsP2, nsP2/nsP3 és nsP3/nsP4 hasítóhelye rendre Gly, Ala és Tyr aminosavat tartalmaz a P1' pozícióban (**Lulla és mtsai, 2006; Hu és mtsai, 2016; Bozóki és mtsai, 2020**), ez igaz a VEEV-re is (**4B. ábra**). Ezek közül a legnagyobb aktivitást a P1'-Tyr esetében kaptuk (**14. ábra**). A P2' pozícióban módosított szubsztrátok vizsgálata során nem tapasztaltuk a kísérletes hasítási hatékonyságok számottevő megváltozását. Ennek az lehet a szerkezeti oka, hogy az S2' alhely kevésbé jól definiált, és a P2' pozícióban a szubsztrát oldallánca az oldószer felé fordulva helyezkedik el (**15. ábra**), ami miatt a szubsztrátnak elsősorban a főlánc alakít ki interakciókat az enzimmel (**Hu és mtsai, 2016**). A legtöbb alfavírus esetében a poliproteinek nsP1/nsP2, nsP2/nsP3 és nsP3/nsP4 hasítóhelyei P2' helyzetben hidrofób/apoláros aminosavakat tartalmaznak (Pro, Leu, Ile, Val) (**Lulla és mtsai, 2006; Hu és mtsai, 2016**), míg egyes alfavírusok esetében az nsP1/nsP2 hasítóhelyen poláros Ser található ebben a pozícióban, ez jellemző többek között a VEEV-re is (**4B. ábra; Hu és mtsai, 2016**). A szerkezetileg kevésbé jól definiált alhelyekre kötődő és az oldószer irányába orientáltan elhelyezkedő P2'-P6' pozíciókban lévő aminosavak esetében inkább a főlánc-mediált kölcsönhatások szerepe jelentős az enzim aktív helyére való kötődés esetén (**Hu és mtsai, 2016**), ezért a saját és más kutatások eredménye alapján is az S4-S1' alhelyek hozzájárulása tekinthető jelentősebbnek a szubsztrát felismeréséhez (**Russo és mtsai, 2010; Lulla és mtsai, 2006**). Összességében elmondható, hogy eredményeink hozzájárulnak a VEEV nsP2pro aminosav preferenciáinak és az azokat meghatározó szerkezeti követelmények jobb megértéséhez.

A kutatások másik részéhez kapcsolódóan munkatársaim a kísérletes vizsgálatok során az nsP2pro egy, az N-terminális régióban hosszabb láncal rendelkező változatával dolgoztak, az *in vitro* vizsgálatok során az enzim aktivitásának relatíve gyors csökkenését tapasztalták. Ennek szerkezeti okait felderítendő, kollaboráció keretében meghatároztuk az enzim

kristályszerkezetét. A meghatározott szerkezetben (PDB ID: 8DUF) az Asn472-Cys477 aminosavakat magába foglaló N-terminális régió két, jelentősen eltérő konformerét láttuk (**16. ábra**). Az egyik konformer a korábban a VEEV és más alfavirus proteázok esetében is megfigyelt aktív konformernek volt megfeleltethető, melyben az N-terminális régió az oldószer irányába volt fordulva. Ezzel szemben a másik konformer esetében - melyet korábban az N475A mutáns VEEV nsP2pro esetében (6BCM) figyeltek meg és annak öninaktivált formájaként azonosítottak - az enzim N-terminális szakasza az aktív hely irányába fordul, elfoglalva a szubsztrátkötő helyet (**5. ábra**). Az általunk meghatározott szerkezet esetében az öninaktivált konformer 57%-ban, míg az aktív 43%-ban volt jelen, azaz megközelítőleg azonos arányban voltak jelen. Az N475A mutációt tartalmazó szerkezetet (PDB ID: 6BCM) leíró közlemény (**Compton és mtsai, 2017**) szerzői azt feltételezték, hogy a szokatlan konformációs állapot az aktív hely mutációja (N475A) következtében alakul ki. Erről az aminosavról korábbi kísérletes vizsgálatok már megállapították, hogy fontos az enzim katalitikus aktivitásának szempontjából, amelynek egyik erős bizonyítéka volt az, hogy az N475A mutáció nagymértékben csökkentette az enzim aktivitását (**Compton és mtsai, 2017**).

Munkánk során a VEEV nsP2pro kristályszerkezeteiben megfigyelhető konformereket molekuladinamikai szimulációkkal vizsgáltuk, az aktív hely 475. pozíciójában aszparagint (Asn485, PDB ID: 8DUF és 2HWK) és alanint (N475A mutáns, PDB ID: 6BCM) tartalmazó enzimek esetében egyaránt. Az r.m.s.d. értékek (**19. ábra**) nem utaltak olyan konformációs változásokra, melyek befolyásolták volna a szerkezetek további vizsgálatát.

A hidrogénhidak százalékos megjelenési aránya alapján vizsgáltuk az N-terminális régiók kölcsönhatási hálózatát az általunk meghatározott szerkezet és az N475A mutáns esetében (8DUF és 6BCM) (**21. ábra**). A kölcsönhatások nem mutattak jelentős különbségeket a két vizsgált változat között, azonban fontos kiemelni, hogy a 475-ös pozícióban lévő mind az aszparagin, mind az alanin képes volt jelentős hidrogénhidak kialakítására, amelyekben főlánci atomjaival vett részt. Ezen felül az aszparagin további hidrogénhidakat is kialakított az N-terminális további aminosavaival. Mindez arra utal, hogy az öninaktivált forma stabilitásáért nem feltétlen az N475A mutáció felel, ez a forma a vad típusú Asn475 aminosavval is megjelenik, sőt, az poláros aszparagin oldallánc olyan további kölcsönhatásokat is kialakíthat, amelyek elősegítik ennek a konformernek a stabilizálását.

Megvizsgáltuk a β -hajtú és az Asn545 aminosav között mért távolságot (**22. és 23. ábrák**), amellyel az interdomén távolság jellemzése volt célunk. Mind az öninaktivált, mind az aktív konformerek esetében, valamennyi vizsgált szerkezet esetében hasonló értékeket kaptunk, ami arra utalt, hogy az N475A mutáció nem változtatja meg a hurok mobilitását.

Vizsgáltuk továbbá az N-terminálison lévő Val476 aminosav és Asn545 kölcsönhatását is (**24. ábra**). Az N475A mutáció tanulmányozása során ugyanis azt állapították meg, hogy ezek interakciója felelhet a β -hajtú hurok relatíve alacsonyabb flexibilitásáért. Hidrogénhid kölcsönhatást figyeltünk meg a két aminosav között mindkét öninaktivált szerkezet (8DUF és 6BCM) trajektóriája esetében (**25. ábra**). Az N475A mutánsban ez a hidrogénhid a trajektória 31%-ában jelent meg, míg az általunk meghatározott, vad típusú Asn475 aminosavval rendelkező szerkezet esetében csak 13%-ban, ami arra utalt, hogy a mutáns esetében erősebb kölcsönhatásról lehet szó. Vizsgáltuk a két aminosav távolságát is (**26. ábra**), amelyre a kapott értékek szerkezetünk esetében jól elkülönültek, azonban az N475A mutáns esetében az aktív konformer több esetben is az öninaktivált konformerre jellemző értékeket vette fel. Ez arra utalt, hogy az aktív konformer kevésbé stabil az N475A mutáns esetében.

Megvizsgáltuk az aktív helyen kialakuló kölcsönhatásokat is mindkét öninaktivált formában, valamint szubsztrátot kötő szerkezet esetében is (**2. táblázat**). A kölcsönhatási hálózatok a két öninaktivált esetben nem tértek el jelentős mértékben. A szubsztrát kevés hidrogénhidat alakít ki az enzimmel, leginkább apoláros kölcsönhatások felelnek annak bekötésért.

Korábbi tanulmányok arra utaltak, hogy az Asn475 és Arg662 közötti kölcsönhatásnak fontos szerepe lehet a proteáz domén és a SAM MTáz domén komplexének stabilizálásában (**27. ábra**). Vizsgáltuk a két aminosav tömegközéppontjának távolságát a szimulációs trajektóriákban (**28. ábra**), azzal a céllal, hogy feltérképezzük az esetleges különbségeket. Az Asn475 aminosavval rendelkező szerkezetek esetében nem láttunk jelentős különbséget a két konformer között, az öninaktiváltra jellemző érték volt valamivel magasabb. Azonban az N475A mutáns esetében az aktív konformerre jellemző érték volt nagyobb, ami arra utal, hogy az N475A mutáció csökkentette a domének közötti interakció erősségét.

Szerettünk volna információkat kapni arról, hogyan befolyásolhatja az N475A mutáció az enzim aktív helyét, ezért megvizsgáltuk a katalitikus diád aminosavainak távolságát is (**29. ábra**). Jelentős különbség volt látható a vad típusú Asn475 aminosavat tartalmazó és az N475A mutáns enzimek szerkezete között (**30. ábra**). A mutáns aktív konformere esetében a diád aminosavainak távolsága egészen nagy értékeket vett fel, ami az aktív hely relatíve nagyobb instabilitására utal. Ez az enzim aktivitása szempontjából kedvezőtlen következményekkel járhat, csökkentheti az aktivitását. Az eredmények azt mutatják, hogy az N475A mutáció megváltoztatja az enzim dinamikai tulajdonságait az aktív hely környékén, ami csökkenti az aktív konformer stabilitását. Az instabilabb aktív hely magában is elegendő lehet az enzim aktivitásának csökkentéséhez, továbbá, az aktív konformer nagyobb flexibilitása azzal a

következménnyel járhat, hogy az könnyebben alakul át az öninaktivált konformerré. Ennek igazolására összehasonlítottuk a kiindulási és az ekvibrált szerkezeteket. Az N-terminálisnak az N475A mutáns esetében tapasztalt jelentős elmozdulása (**20. ábra**) arra utalt, hogy bekövetkezhet az aktív konformáció átalakulása az öninaktivált állapotba, azonban ennek igazolása hosszabb szimulációk vizsgálatával lenne lehetséges.

Az általunk meghatározott 8DUF szerkezet két, a felszíni entrópia csökkentésének céljával beépített mutációt tartalmazott (K741A és K767A). A kristályszerkezetek elemzése alapján feltételezhető, hogy ezek a mutációk növelik a Tyr764-Tyr774 felszíni hurok flexibilitását, ugyanis a hurok egy szakasza (Arg769-Tyr774) nem volt meghatározható a kristályszerkezetben (**18. ábra**). Mivel a kristályszerkezetben megjelent az N-terminális öninaktivált formája, feltételezhető, hogy a felszíni mutációk érzékenyebbé tehetik az enzimet az öninaktivált forma kialakulására vagy annak stabilizálódására, az allosztérikus interakciók megváltozása révén. Nem áll rendelkezésre információ arról, hogy a VEEV replikációs ciklusára hatással vannak-e a felszíni mutációk, de az alfavirus proteázok közti nagymértékű hasonlóság alapján következtethetünk a módosított aminosavak jelentőségére. A VEEV nsP2pro K741A és K767A aminosavai megfeleltethetők a SINV nsP2 SAM MTáz domén R755 és R781 aminosavainak. Az R781 aminosav a SINV SAM MTáz doménjének felszínén helyezkedik el, az nsP2 fehérje valamint az nsP3 „macro” és cink-kötő doménjét összekötő linker közti felszínen. Az nsP2 és nsP3 közti interakciókban az R755 aminosav is részt vehet (**Shin és mtsai, 2012**). Az R718A SINV mutációról korábban kimutatták, hogy nem okoz változást a vírus életképességében (**Mayuri és mtsai, 2008**). Azt, hogy a VEEV replikációs ciklusára van-e hatással ez a két felszíni mutáció, nem ismerjük, erre vonatkozóan nem állnak rendelkezésre irodalmi adatok és korábban ilyen mutációkat tartalmazó enzimformát nem vizsgáltak. A jelen disszertációban bemutatott munkák célja egy rekombináns fehérje szerkezetének *in vitro* és *in silico* vizsgálata volt, ezért kutatásaink nem terjedtek ki az életciklus vizsgálatára sejtes rendszerben. A szerkezetek azonos pozícióiban elhelyezkedő aminosavak hasonló szerepet tölthetnek be a fehérjék működésében, így - a SINV nsP2pro-hoz hasonlóan - a VEEV pozitívan töltött Lys741 és Lys767 oldalláncai is szerepet játszhatnak az nsP2 és nsP3 közti kölcsönhatások kialakításában. Vizsgálataink ezekről a kölcsönhatásokról nem szolgáltatott információkat, mert az általunk vizsgált fehérje csupán az nsP2pro enzimátikus domént reprezentálta. A vad típusú enzimmel és a K741A/K767A mutánsal végzett *in vitro* aktivitásvizsgálatok adhatnak választ arra a kérdésre, hogy ezek a mutációk hogyan befolyásolják az enzimátikus sajátságokat.

A teljes hosszúságú nsP2pro, nsP2 vagy nsP1234 poliprotein fehérjék *in vitro* és/vagy *in silico* vizsgálata segíthet jobban megérteni azt is, hogy az egyes konformációs állapotok kialakulása hogyan befolyásolja az nsP2pro működését *in vivo*, hiszen az enzim aktivációja és inaktivációja egyaránt kiemelt jelentőségű a virális poliprotein és a celluláris fehérjék processzáálásában is. Az *in silico* vizsgálatok azonban számos kihívást tartogathatnak, az egyik limitáló tényező, hogy a nagy méretű fehérjék és makromolekuláris komplexek vizsgálatához szükséges számítási igény igen jelentős lehet. A kutatásunkban vizsgált nsP2pro-nál hosszabb fehérjeformák, így az nsP2 molekuladinamikai szimulációja is jelentősen nagyobb számítási kapacitást követelne meg, míg a teljes nsP1234 poliprotein esetén már a modellt sem lehetne felépíteni az AlphaFold felhasználói felületén keresztül, ilyen léptékű modell elkészítése szuperszámítógépes környezetben lehetséges. A számítási korlátok áthidalásával vagy az arra alkalmas kísérletes rendszerek alkalmazásával érdemes lehet megvizsgálni, hogy az öninaktiváció az enzim működését szabályozó endogén mechanizmus-e vagy az csak az enzim irányított mutagenézise következtében megfigyelhető. Ismertek példák inaktiválódó enzimekre, mint például a T7 bakteriofág kináz is, mely képes foszforilálni önmagát (**Pai és mtsai, 1975**), ami aktivitásának elvesztését okozza, vagy a NADH dehidrogenáz, melynek aktivitását az általa termelt szuperoxid molekula gátolja az enzim oxidációja révén (**Chen és mtsai, 2005**). Ezek az öninaktivációs mechanizmusok intermolekuláris úton zajlanak, azaz az egyes enzim molekulák képesek a további, azonos enzimek inaktiválására. Egyes enzimek azonban saját magukat is képesek lehetnek inaktiválni, mint pl. az aszpartil proteázok közé tartozó kimozin emésztőenzim (**Andreeva és mtsai, 1992**) és egy rovar-eredetű aszpartil proteáz (**Gustchina és mtsai, 2005**), amelyek esetében megfigyeltek egy olyan mechanizmust, melynek során az enzim katalitikus helye közelében lévő aminosavainak konformációs változásai vezetnek öninaktiváláshoz. Lehetséges, hogy a VEEV nsP2pro esetében megfigyelt konformációs állapot az enzim önszabályozási mechanizmusa miatt jön létre, azonban nem ismerjük azokat a faktorokat, amik ennek az öninaktivált formának a kialakulását segíthetik vagy gátolhatják. Eredményeink alapján az Asn475 aminosavnak fontos szerepe lehet az aktív konformáció stabilizálásában és gátolhatja az öninaktivált konformáció kialakulását.

A virális poliprotein egyes régiói és proteáz aktív helye között kialakuló kölcsönhatások megismerése azért is lehet fontos, mert a virális poliprotein részeként expresszálandó funkcionális domének működőképessé válása más vírusok életciklusában is kulcsfontosságú. Ilyen funkcionális domének lehetnek szerkezeti (például mátrix és kapszid) vagy enzimaktivitással rendelkező nem-szerkezeti fehérjék is (például integráz, proteáz). A poliproteinből történő kihalással aktiválódik a HIV vírus aszpartil proteáza (**Louis és mtsai,**

2000) és a SARS-CoV-2 cisztein proteáza is (fő proteáz, Mpro) (Aniana és mtsai, 2023), de a vírusokéhoz hasonló módon hasítja ki magát a teljes hosszúságú prekursor fehérjéből az emberi szervezetben is megtalálható retrovírus-szerű aszpartil proteáz 1 (ASPRV1) fehérje proteáz doménje is, ami a proteáz aktiválódásához vezet (Golda és mtsai, 2020). Kimutatták, hogy a HIV-1 és SARS-CoV-2 proteázok aktivációja a proteolitikus aktiváció során a proteáz domén N-terminálisán bekövetkező hasítás intramolekuláris, míg a C-terminális hasítása intermolekuláris módon történik meg (Louis és mtsai, 2000; Aniana és mtsai, 2023). Jövőbeli kutatások tisztázhatják, hogy a VEEV esetében is jellemzőek-e ezek a lépések az nsP2pro kihaladására, és hogy az autoaktiváció mechanizmusa általánosan jellemző-e azokra a proteázokra, melyek a prekursor fehérjéből történő kihalás révén aktiválódnak. Ezeknek a folyamatoknak a megértése segíthet választ adni arra a kérdésre is, hogy a VEEV nsP2pro öninaktivációja is megtörténhet-e intra- és intermolekuláris kölcsönhatások által egyaránt.

Az alfavírus proteázok vizsgálatának jelentőségét adja többek között a fehérjéik, így a nem-szerkezeti proteázok szerkezete (17. ábra) és hasítóhely-szekvenciáik (4. ábra) közti nagyfokú hasonlóság. A hasítóhely szekvenciák hasonlóságának jelentőségét a VEEV, EEEV és WEEV vírusok proteázai esetében is vizsgálták már (Morazzani és mtsai, 2019; Lulla és mtsai, 2006; Russo és mtsai, 2006; Russo és mtsai, 2010, Hu és mtsai, 2023). Kimutatták többek között, hogy ezen vírusok nsP12 autoproteolitikus hasítóhely szekvenciái teljes mértékben azonosak, és ez az szekvencia megtalálható a veleszületett immunrendszer működésében szerepet játszó TRIM14 fehérjében is, amit így ezek proteázok mindegyike képes elhasítani (Morazzani és mtsai, 2019). A VEEV nsP2pro vizsgálatával tehát fontos információk nyerhetők többek között az EEEV és WEEV vírusok azonos fehérjéről is, hiszen például az azonos és/vagy hasonló szekvenciákat felismerő és hasítani képes alfavírus proteázok a fertőzött sejtekben is potenciálisan ugyanazokat a fehérjéket lehetnek képesek hasítani. A specificitás vizsgálatával szerzett új ismertetek tehát a jövőben segíthetik az nsP2pro új celluláris szubsztrátjainak az azonosítását, akár a vírus és a celluláris fehérjék közti szekvencia-azonosság vizsgálatával (SSHPS elemzés; Morazzani és mtsai, 2019). Ezen túlmenően a hasítóhelyek közti hasonlóság nem csak a rokon vírusok működését segíthet jobban megérteni, de nagyban segítheti például egyszerre több vírus replikációját is gátolni képes inhibitor kifejlesztését. Az alfavírusok más fehérjéi, így a felszíni burokfehérjék közti szerkezetek homológia már napjainkban is alapját képezi az újvilági alfavírusok elleni kombinált vakcina kifejlesztésének (Fierro és mtsai, 2024). A vakcinák fejlesztése mellett széles-spektrumú gátlószerek kifejlesztéséhez is hozzájárulhat azoknak az enzim-szubsztrát és enzim-ligand kölcsönhatásoknak az azonosítása, melyek mindegyik vírus proteáza esetében

kialakulnak. A VEEV nsP2pro öninaktiváló konformációja esetében az enzim N-terminális pseudo-szubsztrátként kötődik az enzim aktív helyére. Mivel az enzim N-terminális régiója képes az aktív helyre kötődni, az ezt a szekvenciát mimikáló molekulák akár szubsztrát-analóg gátlószer fejlesztésének is alapját képezhetik. A virális proteáz természetes hasítóhely szekvenciájához hasonló, a szubsztrátot mimikáló proteáz inhibitorokat a HIV-1 proteáz esetében is kifejlesztettek (Tózsér, 2001; Flexner, 2007; Tie és mtsai, 2012). Vizsgálataink során több olyan kölcsönhatást is azonosítottunk a VEEV nsP2pro aktív helyén, amelyek elősegíthetik az öninaktivált konformer stabilizálását az N-terminális bekötődése esetén, de szerepet játszhatnak akár egy szubsztrát-analóg inhibitor kötésében is. Jövőbeli vizsgálatok célja lehet például olyan inhibitor kifejlesztése is, amely kovalens kötést alakít ki a katalitikus ciszteinnel (Huang és mtsai, 2022), melyhez leggyakrabban nitrilcsoportot alkalmaznak (Brogi és mtsai, 2022). Ilyen csoportot találunk a SARS-CoV-2 fő proteázt gátolni képes inhibitorban, a nirmatrelvirben is (Owen és mtsai, 2021), melynek kölcsönhatásait a proteáz aktív helyével már kutatócsoportunk is vizsgálta (Mótyán és mtsai, 2022). A közelmúltban oldották fel a VEEV nsP2pro-nak egy peptid-szerű, epoxid csoportot tartalmazó kovalens cisztein proteáz inhibitorral (CA074) alkotott komplexének kristályszerkezetét (PDB ID: 8T8N, Hu és mtsai, 2023). Kimutatták, hogy az inhibitor hatékonyan képes gátló hatást kifejteni a VEEV, a WEEV és a Chikungunya vírus proteáz esetében, mely enzimek szerkezete nagyfokú hasonlóságot mutat. A CA074 inhibitor bekötődésében kifejezetten azon aminosavaknak van szerepe, amelyek jelentős konzerváltságot mutatnak ezen alfavírusok szekvenciáiban. Az enzim-inhibitor kölcsönhatást molekuladinamikai szimulációkkal történő vizsgálata segíthet felderíteni, hogy hogyan lehet tovább növelni a gátlószer hatékonyságát. A számítógépes modellezés módszereit már korábbi tanulmányok során is alkalmazták lehetséges alfavírus inhibitorok vizsgálatában (Byler és mtsai, 2016), a jövőben pedig az inhibitorok tervezésekor akár az öninaktivált N-terminális konformer esetében meghatározott interakciókat is figyelembe lehet venni.

8. ÖSSZEFOGLALÁS

A doktori értekezésemben bemutatott kutatások célja a venezuelai ló-láz encephalitis vírus (VEEV) nem-szerkezeti fehérje 2 proteázának (nsP2pro) számításhoz kémiai módszerekkel történő vizsgálata volt, amelyhez a molekuladinamika módszereit alkalmaztam. Vizsgálataink során az enzim specificitását, valamint az öninaktivált konformáció kialakulását és az enzim stabilitását meghatározó szerkezeti sajátosságok vizsgálatát tűztük ki célul.

A szubsztrát-specificitás vizsgálatához az nsP2pro oligopeptid szubsztráttal alkotott komplex modelljét hoztam létre, mely szubsztrát a kísérletes vizsgálatok során is használt SFV nsP1/nsP2 természetes hasítóhely szekvenciát reprezentálta. Az így felépített enzim-szubsztrát komplexet molekuladinamikai szimulációkkal vizsgáltam, majd az ekvilibrált komplexben lévő szubsztrátot különböző (P5-től P2'-ig) pozíciókban mutáltuk *in silico*, és meghatároztuk, hogy a szubsztrát módosítása hogyan befolyásolja az enzim-szubsztrát kölcsönhatásokat. Eredményeinket a kísérletes specificitási vizsgálatok eredményeinek értelmezéséhez és az aminosav-preferenciákat meghatározó szerkezeti sajátosságok azonosításához használtuk fel.

A VEEV nsP2pro szerkezetének vizsgálatát kollaborációs partnereink röntgenkristallográfiával is elvégezték. Az aktív helyen nem (de a felszínen két pozícióban is) módosított enzim szerkezetében az N-terminális aktív hely irányába fordult konformerét észleltük, amely jelentős hasonlóságot mutatott a korábban az N475A aktív hely mutáns enzim esetében leírt, ún. öninaktivált formával. Az öninaktivált konformációs állapot kialakulásának szerkezeti hátterét molekuladinamikai szimulációkkal vizsgáltam a munkánk során meghatározott, valamint korábban publikált szerkezeti koordináták alapján. Vizsgáltam az N-terminális hidrogénhid-interakciós hálózatát és a dinamikai szempontok alapján fontos aminosavak távolságának változását is. Továbbá, összehasonlítottam az aktív helyen kialakuló kölcsönhatásokat az enzim N-terminális régiója (öninaktivált konformer) vagy oligopeptid szubsztrát kötődése esetében, meghatározva a kölcsönhatási hálózatok legfontosabb hasonlóságait és különbségeit. A szimulációs eredmények felfedték a vad típusú Asn475 aminosav esetében is megjelenő, az öninaktivált forma kialakulásért felelős kölcsönhatásokat.

Eredményeink hozzájárulnak a VEEV nsP2Pro specificitását valamint stabilitását meghatározó szerkezeti követelmények jobb megértéséhez és potenciálisan felhasználhatóak lehetnek más homológ virális proteázok vizsgálatában, valamint a proteázt gátolni képes molekulák tervezésekor.

9. SUMMARY

The goal of the research - presented in my doctoral thesis - was to study the Venezuelan equine encephalitis virus (VEEV) non-structural protein 2 protease (nsP2pro) by using computational approaches, for which I have applied the molecular dynamics methods. During our studies, we aimed to investigate the structural features that determine the specificity and stability of the enzyme and have correlation with the formation of the self-inactivated conformation

To investigate the substrate specificity, I prepared a model of the nsP2pro in complex with an oligopeptide substrate representing the SFV nsP1/nsP2 natural cleavage site sequence, this substrate used in the *in vitro* specificity study, as well. The enzyme-substrate complex was investigated with molecular dynamics simulations, followed by the modification of the substrate in the equilibrated complex in various positions (from P5 to P2') *in silico*. We have determined how the modification of the substrate affects the enzyme-substrate interactions. The results of the computational analyses were used to interpret those of the experimental specificity assays and to identify the structural features that determine amino acid preferences.

The structure of VEEV nsP2pro was determined by our collaboration partners using X-ray crystallography. We found that the N terminus of the enzyme lacking any mutation at the active site (but containing two in its surface) flipped into the active site, which so-called self-inactivated conformation closely resembled that of the previously described N475A active site-mutant enzyme. The structural requirements of the self-inactivated conformation were investigated by using molecular dynamics simulations, based on the structural coordinates determined in this work or published previously. We examined the hydrogen bond interaction network of the N terminus and the changes in the distances between the most important residues. In addition, the similarities and differences of the interaction networks were also determined in the case of binding of the N terminus (self-inactivated conformer) or an oligopeptide substrate to the active site. The simulations revealed the interactions that are responsible for the formation of the self-inactivated conformation and also appears in the case of the wild-type Asn475 residue.

Our results help better understanding of the structural determinants of the specificity and stability of VEEV nsP2Pro, and can potentially be used in the investigation of other homologous viral proteases as well as can aid the design of protease inhibitors.

10. IRODALOMJEGYZÉK

- Abdelbani, R.**, Delang, L. Antiviral Strategies against Arthritogenic Alphaviruses. *Molecules*, 2020, 8(9), 1365.
- Abu Bakar, F.**, Ng, L.F.P. Nonstructural Proteins of Alphavirus—Potential Targets for Drug Development. *Viruses* 2018, 10(2), 71.
- Alder, B.J.**, Wainwright, T.E. Phase Transition for a Hard Sphere System. *J. Chem. Phys.* 1957, 27, 1208-1209.
- Anandkrishnan, R.**, Aguilar, B., Onufriev, A.V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acid Res* 2012, 40, W537-541.
- Andreeva, N.**, Dill, J., Gilliland, G.L. Can enzymes adopt a self-inhibited form? Results of X-ray crystallographic studies of chymosin. *Biochem. Biophys. Res. Commun.* 1992, 184(2), 1074-1081.
- Aniana, A.**, Nashed, N.T., Ghirlando, R., Coates, L., Kneller, D.W., Kovalevsky, A., Louis, J.M. Insight into the mechanism of SARS-CoV-2 main protease autocatalytic maturation from model precursors *Commun. Biol.* 2023, 6, 1159.
- Bader, R.F.W.**, Tal, Y., Anderson, S.G., Nguyen-Dang, T.T. Quantum Topology: Theory of Molecular Structure and its Change. *Isr. J. Chem.* 1980, 19(1-4), 8-29.
- Badger, J.** An evaluation of automated model-building procedures for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 2003, 59(5), 823-827.
- Bekker, H.**, Berendsen, H.J.C., Dijkstra, E.J., Achterop, S., van Drunen, R., van der Spoel, D., Sijbers, A, Keegstra H., és mtsai. “Gromacs: A parallel computer for molecular dynamics simulations”; *Physics computing 92*. 1993, Edited by R.A. de Groot and J. Nadrchal. World Scientific, Singapore, 252–256.
- Berendsen, H.J.C.**, Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak, J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 1984, 81(8), 3684-3690.
- Berman, H.M.**, Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* 2000, 28(1), 235–242.
- Byler, K.G.**, Collins, J.T., Ogungbe, I.V., Setzer, W.N. Alphavirus protease inhibitors from natural sources: A homology modeling and molecular docking investigation. *Comput. Biol. Chem.* 2016, 64, 163-184.
- Brogi, S.**, Ibba, R., Rossi, S., Butini, S., Calderone, V., Gemma, S., Campiani, G. Covalent Reversible Inhibitors of Cysteine Proteases Containing the Nitrile Warhead: Recent Advancement in the Field of Viral and Parasitic Diseases. *Molecules* 2022, 27(8), 2561.
- Bronze, M.S.**, Huycke, M.M., Machado, L.J., Voskuhl, G. W., Greenfield, R.A. Viral Agents as Biological Weapons and Agents of Bioterrorism. *Am. J. Med. Sci.* 2002, 323, 316–325.

- Bozóki, B.**, Mótyán, J.A., Hoffka, G., Waugh, D.S., Tózsér, J. Specificity Studies of the Venezuelan Equine Encephalitis Virus Non-Structural Protein 2 Protease Using Recombinant Fluorescent Substrates, *Int. J. Mol. Sci.* 2020, 21(20), 7686.
- Campos-Gomez, J.**, Ahmad, F., Rodriguez, E., Faeed, M.F. A novel cell-based assay to measure activity of Venezuelan equine encephalitis virus nsP2 protease. *Virology*, 2016, 496, 77–89.
- Case, D.A.**, Betz, R.M., Cerutti, D.S., és mtsai. AMBER 2016. University of California, San Francisco., 2016.
- Case, D.A.**, Cheatham, T.E. 3rd, Darden, T., Gohlke, H., Luo, R., Merz K.M. Jr., Onufriev, A., Simmerling, C., Wang, B., Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* 2005, 26(16), 1688-1688.
- Carey, B.D.**, Bakovic, A., Callahan, V., Narayanan, A., Kehn-Hall, K. New World alphavirus protein interactomes from a therapeutic perspective. *Antiviral Res.* 2019, 163, 125–139.
- Chen, V.B.**, Arendall, W.B. 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 2010, 66(1), 12-21.
- Chen, Y-R.**, Chen, C-L., Zhan, L., Green-Church, K.B., Zweier, J. Superoxide Generation from Mitochondrial NADH Dehydrogenase Induces Self-inactivation with Specific Protein Radical Formation. *J. Biol. Chem.* 2005, 280(45), 37339-37348.
- Coates, E.E.**, Edupuganti, S., Chen, G.L., Happe, M., Strom, L., Widge, A., Florez, M.B., Cox, J.H. és mtsai. Safety and immunogenicity of a trivalent virus-like particle vaccine against western, eastern, and Venezuelan equine encephalitis viruses: a phase 1, open-label, dose-escalation, randomised clinical trial. *Lancet Infect. Dis.* 2022, 22(8): 1210-1220.
- Cohen, F.E.**, Sternberg, M.J.E. On the prediction of protein structure: The significance of the root-mean-square deviation. *J. Mol. Biol.* 1980, 138(2), 321-333.
- Compton, J.R.**, Mickey, M.J., Hu, X., Marugan, J.J., Legler, P.M. Mutation of Asn-475 in the Venezuelan Equine Encephalitis Virus nsP2 Cysteine Protease Leads to a Self-Inhibited State. *Biochemistry* 2017, 56(47), 6221–6230.
- Cornell, W.D.**, Cieplak, P., Bayly, C.I., Gould I.R., Merz, K.M. Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., Kollman, P.A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* 1995, 117(19) 5179-5197.
- Darden, T.**, York, D., Pedersen, L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 1993, 98, 10089–10092.
- de Cedrón, M.G.**, Ehsani, N., Mikkola, M.L., García, J.A., Kääriäinen, L. RNA helicase activity of Semliki Forest virus replicase protein NSP2. *FEBS Lett.* 1999, 448, 19–22.

- Deardor, E.R.**, Forrester, N.L., da Rosa, A.P.T., Estrada-Franco, J.G., Navarro- Lopez, R., Tesh, R.B., Weaver, S.C. Experimental Infection of Potential Reservoir Hosts with Venezuelan Equine Encephalitis Virus, Mexico. *Emerg. Infect.Dis.* 2009, 15, 519-525.
- Dennington, R.**, Keith, T.A., Millam, J.M. GaussView, Version 6.1.1. Semichem Inc., Shawnee Mission, KS. 2016
- Dunbrack, R.L.** Rotamer Libraries in the 21st Century. *Curr. Opin. Struct. Biol.* 2002, 12(4), 431–440.
- Faller, R.**, de Pablo, J.J. Constant pressure hybrid Molecular Dynamics –Monte Carlo simulations. *J. Chem. Phys.* 2002, 116, 55-59.
- Fierro, C.**, Weidenthaler, H., Vidojkovic, S., Schmidt, D., Gafoor, Z., Stroukova, D., Zwiers, S., Müllet, J., Volkmann, A. Safety and immunogenicity of a novel trivalent recombinant MVA-based equine encephalitis virus vaccine: A Phase 1 clinical trial. *Vaccine* 2024, 42(10), 2695-2706.
- Flexner C.** HIV drug development: the next 25 years. *Nat. Rev. Drug. Discov.* 2007, 6(12), 959-966.
- Garmashova, N.**, Gorchakov, R., Volkova, E., Paessler, S., Frolova, E., Frolov, I. The Old World and New World alphaviruses use different virus-specific proteins for induction of transcriptional shutoff. *J. Virol.* 2007, 81(5), 2472–2484.
- Gavira, J.A.** Current trends in protein crystallization, *Arch. Biochem. Biophys.* 2016, 602, 3-11.
- Golda, M.**, Mótyán, J.A., Nagy, K., Matúz, K., Nagy, T., Tózsér, J. Biochemical Characterization of Human Retroviral-Like Aspartic Protease 1 (ASPRV1). *Biomolecules* 2020, 10(7), 1004.
- Golubtsov, A.**, Kääriäinen, L., Caldenty, J. Characterization of the cysteine protease domain of Semliki Forest virus replicase protein nsP2 by in vitro mutagenesis. *FEBS Lett.* 2006, 580, 1502–1508.
- Gordon, J.C.**, Myers, J.B., Folta, T., Shoja, V., Heath, L.S., Onufriev, A. H⁺⁺: a server for estimating pK_as and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* 2005, 33, W368-371.
- Gunsteren, W.F.V.**, Daura, X., Mark, A.E., Schleyer, P.V.R. GROMOS force field. *Encyclopedia of Computational Chemistry* 1998, 1211-1216.
- Gustchina, A.**, Li, M., Wünschmann, S., Chapman, M.D., Pomés, A., Wlodawer, A., Crystal Structure of Cockroach Allergen Bla g 2, an Unusual Zinc Binding Aspartic Protease with a Novel Mode of Self-inhibition. *J. Mol. Biol.* 2005, 348(2), 433-444.
- Guzmán-Terán, C.**, Calderón-Rangel, A., Rodríguez-Morales, A., Mattar, S. Venezuelan equine encephalitis virus: the problem is not over for tropical America. *Ann. Clin. Microbiol. Antimicrob.* 2020, 19(1), 19.

- Han, L.**, Song, S., Feng, H., Ma, J., Wei, W., Si, F. A roadmap for developing Venezuelan equine encephalitis virus (VEEV) vaccines: Lessons from the past, strategies for the future *Int. J. Biol. Macromol.* 2023, 245, 125514.
- Harmat, V.**, Vértessy, G.B. Fehérjekrisztallográfia. *Ezerarcú fehérjék (Szerkesztette: Buday, L., Nyitray, L., Perczel, A.) – Semmelweis Kiadó*, 2018, 181-201.
- Harrison, R.L.** Introduction To Monte Carlo Simulation, *AIP Conf. Proc.* 2010, 17-21.
- Havel, T.F.**, Crippen, G.M., Kuntz, I.D. Effects of distance constraints on macromolecular conformation. II. Simulation of experimental results and theoretical predictions. *Biopolymers* 1979, 18(1), 73-81.
- Henderson, D.**, Jacobson, S.H., Johnson, A.W. The theory and practice of simulated annealing. *Handbook of metaheuristics*, 2003, 57, 287-319.
- Hoffka, G.**, Lountos, G.T., Needle, D., Wlodawer, A., Waugh, D.S., Tózsér, J., Mótyán, J.A. Self-inhibited State of Venezuelan Equine Encephalitis Virus (VEEV) nsP2 Cysteine Protease: A Crystallographic and Molecular Dynamics Analysis. *J. Mol. Biol.* 2023, 435(6), 168012.
- Hoover, G.W.** Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* 1985, 31(3), 1695-1697.
- Hu, X.**, Compton, J.R., Leary, D.H., Olson, M.A., Lee, M.S., Cheung, J., Ye, W., Ferrer, M. és mtsai. Kinetic, Mutational, and Structural Studies of the Venezuelan Equine Encephalitis Virus Nonstructural Protein 2 Cysteine Protease. *Biochemistry* 2016, 55(21), 3007–3019.
- Hu, X.**, Morazzani, E., Compton, J.R., Harmon, M., Soloveva, V., Glass, P.J., Garcia, A.D., Marugan, J.J., Legler, P.M. In silico screening of inhibitors of the Venezuelan equine encephalitis virus nonstructural protein 2 cysteine protease. *Viruses* 2023, 15(7), 1503.
- Huang, F.**, Han, X., Xiao, X., Zhou, J. Covalent Warheads Targeting Cysteine Residue: The Promising Approach in Drug Development, *Molecules* 2022, 27(22), 7728.
- Jo, S.**, Kim, T., Iyer, V.G., Im, W. CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. *J. Comput. Chem.* 2008, 29, 1859-1865.
- Johnson, K.M.**, Martin, D.H. Venezuelan equine encephalitis. *Adv. Vet. Sci. Comp. Med.* 1974, 18, 79–116.
- Jorgensen, W.L.**, Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 1983, 79, 926.
- Jumper J.**, Evans, R., Pritzel, A., Green, T., Figurnov, M., és mtsai. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021, 596, 583–589.
- Kapust, R.B.**, Tózsér, J., Copeland, T.D., Waugh, D.S. The P1' specificity of tobacco etch virus protease. *Biochem. Biophys. Res. Commun.* 2002, 294(5), 949–955.

- Kapust, R.B.**, Tózsér, J., Fox, J.D., Anderson, D.E., Cherry, S., Copeland, T.D., Waugh, D.S. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency *Protein. Eng.* 2001, 14(12) 993–1000.
- Karpen, M.E.**, Tobias, D.J., Brooks, C.L. 3rd. Statistical clustering techniques for the analysis of long molecular dynamics trajectories: analysis of 2.2-ns trajectories of YPGDV. *Biochemistry* 1993, 32(2), 412-420.
- Karplus, M.**, Petsko, G.A. Molecular dynamics simulations in biology. *Nature*, 1990, 347(6294), 631-639.
- Kendrew, J.C.**, Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H.A. Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* 1958, 181, 662–666.
- Kirkpatrick, S.**, Gelatt, C.D., Vecchi, M. Optimization by Simulated Annealing. *Science* 1983, 220(4598) 671–680.
- Kleywegt, G.J.**, Brünger, A.T. Checking your imagination: applications of the free R value. *Structure*, 1996, 4(8), 897-904.
- Kleywegt, G.J.** On vital aid: the why, what and how of validation. *Acta Crystallogr. D Biol. Crystallogr.* 2009, 65(2), 134-139.
- Krieger, E.**, Vriend, G. YASARA View—molecular graphics for all devices—from smartphones to workstations. *Bioinformatics* 2014, 30(20)m 2981-2982.
- Kumar, P.**, Kumar, D., Giri, R. Targeting the nsp2 Cysteine Protease of Chikungunya Virus Using FDA Approved Library and Selected Cysteine Protease Inhibitors. *Pathogens*, 2019, 8(3), 128
- Langevin, P.** Sur la théorie du mouvement brownien. *C. R. Acad. Sci. Paris* 1908, 146, 530-533.
- Laskowski, R.A.**, Swindells, M.B. LigPlot+: Multiple ligand–protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* 2011, 51, 2778–2786.
- Leach, A.R.** Molecular modelling: Principles and applications, Longman Harlow, 1996
- Le Grand, S.**, Gotz, A.W., Walker, R.C. SPFP: Speed without compromise – a mixed precision model for GPU accelerated molecular dynamics simulations. *Comput. Phys. Commun.* 2013, 184, 374–380.
- Lennard-Jones, J.E.** Cohesion. *Proc. Phys. Soc.* 1931, 43(5), 461.
- Levitt, M.**, Warshel, A. Computer simulation of protein folding. *Nature* 1975, 253, 694-698.
- Louis, J.M.**, Weber, I.T., Tózsér, J., Clore, G.M., Gronenborn, A.M. HIV-I Protease: Maturation, Enzyme Specificity, and Drug Resistance. *Advances in Pharmacology* 2000, Volume 49, 111-146.
- Lulla, A.**, Lulla, V., Tints, K., Ahola, T., Merits, A. Molecular determinants of substrate specificity for Semliki Forest virus nonstructural protease. *J. Virol.* 2006, 80(11), 5413-5422.

- Machado, M.R.**, Pantano, S. Split the Charge Difference in Two! A Rule of Thumb for Adding Proper Amounts of Ions in MD Simulations. *J. Chem. Theory Comput.* 2020, 16(3), 1367–1372.
- Maier, J.A.**, Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 2015, 11(8), 3696–3713.
- Mayuri, G.T.W.**, Smith, J.L., Kuhn, R.J. Role for conserved residues of sindbis virus nonstructural protein 2 methyltransferase-like domain in regulation of minus-strand synthesis and development of cytopathic infection. *J. Virol.* 2015, 82(15), 7284–7297.
- McCammon, J.A.**, Gelin, B.R., Karplus, M. Dynamics of folded proteins, *Nature* 1977, 267, 585-590.
- Metropolis, N.**, Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E.D. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* 1953, 21, 1087-1092.
- Meyer, R.F.**, Morse, S.A. Viruses and Bioterrorism. *Encycloped. Virol.* 2008, 406–411.
- McCoy, A.J.**, Read, R.J. Experimental phasing: best practice and pitfalls *Acta Crystallogr. D Biol. Crystallogr.* 2010, 66(4), 458-469.
- McPherson, A.**, Gavira, J.A. Introduction to protein crystallization. *Acta Crystallogr. F. Struct. Biol. Commun.* 2014, 70(1), 2-20.
- Mirdita, M.**, Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., Steinegger, M. ColabFold: making protein folding accessible to all. *Nat. Methods*, 2022, 19, 679-682
- Miyamoto, S.**, Kollman, P.A. SETTLE – An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* 1992, 13, 952–962.
- Morazzani, E.M.**, Compton, J.R., Leary, D.H., Berry, A.V., Hu, X., Marugan, J.J., Glass, P.J., Legler, P.M. Proteolytic cleavage of host proteins by the Group IV viral proteases of Venezuelan equine encephalitis virus and Zika virus. *Antiviral Res.* 2019, 164, 106–122.
- Mótyán, J.A.**, Mahdi, M., Hoffka, G., Tózsér, J. Potential Resistance of SARS-CoV-2 Main Protease (Mpro) against Protease Inhibitors: Lessons Learned from HIV-1 Protease. *Int. J. Mol. Sci.* 2022, 23(7), 3507
- Nallamsetty, S.**, Kapust, R.B., Tózsér, J., Cherry, S., Tropea, J.E., Copeland, T.D., Waugh, D.S. Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro. *Protein. Expr. Purif.* 2004, 38(1), 108–115.
- Narwal, M.**, Singh, H., Pratap, S., Malik, A., Kuhn, R.J., Kumar, P., Tomar, S. Crystal structure of chikungunya virus nsP2 cysteine protease reveals a putative flexible loop blocking its active site. *Int. J. Biol. Macromol.* 2018, 116, 451–462.
- Neumann, J.** A kvantummechanikai matematikai alapjai, 1980, Akadémiai Kiadó (Eredeti címe: Mathematische Grundlagen der Quantenmechanik, 1932.)

- Nosé, S.** A molecular dynamics method for simulations in the canonical ensemble *Mol. Phys.* 1984, 52(2), 255-268.
- Ogorek, T.J.,** Golden, J.E. Advances in the Development of Small Molecule Antivirals against Equine Encephalitic Viruses. *Viruses* 2023, 15(2), 413.
- Olsson, M.H.M.,** Søndergaard, C.R., Rostkowski, M., Jensen, J.H. PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* 2011, 7(2), 525–537.
- Owen, D.R.,** Allerton, C.M.N., Anderson, A.S., Aschenbrenner, L., Avery, M., Berritt, S., Boras, B., Cardin, R.D., Carlo, A., Coffman, K.J., Dantonio, A., Di, L., Eng, H., Ferre, R., Gajiwala, K.S., Gibson, S., Greasley, S.E., Hurst, B.L., Kadar, E., P., Kalgutkar, A.S., Lee, C.J., Lee, J., Liu, W., Mason, S., W., Noell, S., Novak, J.J., Obach, R.S., Ogilvie, K., Patel, N.C., Pettersson, M., Rai, D.K., Reese, M.R., Sammons, M.F., Sathish, J.G., Singh, R.S.P., Stepan, C.M., Stewart, A.E, Tuttle, J.B., Updyke, L., Verhoest, P.R., Wei, L., Yang, Q., Zhu, Y. An oral SARS-CoV-2 M^{Pro} inhibitor clinical candidate for the treatment of COVID-19. *Science* 2021, 374(6575), 1586-1593.
- Pace, C.N.,** Scholtz, J.M., Grimsley, G.R. Forces stabilizing proteins. *FEBS Lett.*, 2014, 588(14), 2177-2184.
- Pai, S.H.,** Rahmsdorf, H.J., Ponta, H., Hirsch-Kauffmann, M., Herrlich, P., Schweiger, M. Protein kinase of bacteriophage T7. 2. Properties, enzyme synthesis in vitro and regulation of enzyme synthesis and activity in vivo. *Eur. J. Biochem.* 1975, 55(1), 305-314.
- Parrinello, M.,** Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* 1980, 45(14), 1196-1199.
- Pauling, L.,** Corey, R.B., Branson, H.R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *PNAS*, 1951, 37(4), 205-211.
- Pettersen, E.F.,** Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 2004, 25(13), 1605–1612.
- Phillips, J.C.,** Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, D.R., Kalé, L., Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* 2005, 26(16), 1781–1802.
- Powers, A.M.,** Brault, A.C., Shirako, Y., Strauss, E.G., Kang, W., Strauss, J.H., Weaver, S.C. Evolutionary relationships and systematics of the alphaviruses. *J. Virol.* 2001, 75(21), 10118–10131.
- Rahman, A.** Correlations in the Motion of Atoms in Liquid Argon. *Phys. Rev.* 1964, 136(2A), A405-A411.
- Rikkonen, M.,** Peränen, J., Kääriäinen, L. ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. *J. Virol.* 1994, 68, 5804–5810.

- Rodrigues, C.H.**, Pires, D.E.V., Ascher, D.B. DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res.* 2018, 46, W350–W355.
- Roe, D.R.**, Cheatham 3rd, T.E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* 2013, 9(7), 3084–3095.
- Rupp, J.C.**, Sokoloski, K.J., Gebhart, N.N., Hardy, R.W. Alphavirus RNA synthesis and non-structural protein functions. *J. Gen. Virol.* 2015, 96, 2483–2500.
- Russo, A.T.**, Malmstrom, R.D., White, M.A., Watowich, S. J. Structural basis for substrate specificity of alphavirus nsP2 proteases. *J. Mol. Graph. Model.* 2010, 29(1), 46–53.
- Russo, A.T.**, White, M.A., Watowich, S.J. The crystal structure of the Venezuelan equine encephalitis alphavirus nsP2 protease. *Structure* 2006, 14(9), 1449–1458.
- Salomon-Ferrer, R.**, Gotz, A.W., Poole, D., Le Grand, S., Walker, R.C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* 2013, 9(9), 3878–3888.
- Scapin, G.** Molecular replacement then and now. *Acta Crystallogr. D Biol. Crystallogr.* 2013, 69(11), 2266–2275.
- Schrödinger, LLC.** The PyMOL Molecular Graphics System, 2010
- Schymkowitz, J.**, Borg, J., Stricher, F., Nys, R., Rousseau, F., Serrano, L. The FoldX web server: An online force field. *Nucleic Acids Res.* 2005, 33, W382–W388.
- Schechter, I.**, Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* 1967, 27(2), 157–62.
- Sharma, A.**, Knollmann-Ritschel, B. Current Understanding of the Molecular Basis of Venezuelan Equine Encephalitis Virus Pathogenesis and Vaccine Development. 2019, *Viruses* 11 (2), 164.
- Shao, J.**, Tanner, S.W., Thompson, N., Cheatham, T.E. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. *J. Chem. Theory. Comput.* 2007, 3(6), 2312–2334.
- Shenkin, P.S.**, McDonald, D.Q. Cluster analysis of molecular conformations. *J. Comput. Chem.* 1994, 15(8), 899–916.
- Shi, Y.** A Glimpse of Structural Biology through X-Ray Crystallography. *Cell.* 2014, 159(5), 995–1014.
- Shin, G.**, Yost, S.A., Miller, M.T., Elrod, E.J., Grakoui, A., Marcotrigiano, J. Structural and functional insights into alphavirus polyprotein processing and pathogenesis. *PNAS* 2012, 109(41), 16534–16539.
- Sievers, F.**, Higgins, D.G. Clustal Omega. *Curr. Protoc. Bioinform.* 2014, 48.

- Søndergaard, C.R.**, Olsson, M.H.M., Rostkowski, M., Jensen, J.H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *J. Chem. Theory Comput.* 2011, 7(7), 2284–2295.
- Strauss, J.H.**, Strauss, E.G. The alphaviruses: gene expression, replication, evolution. *Microbiol. Rev.* 1994, 58, 491–562.
- Stromberg, Z.R.**, Fischer, W., Bradfute, S.B., Kubicek-Sutherland, J.Z., Hraber, P. Vaccine Advances against Venezuelan, Eastern, and Western Equine Encephalitis Viruses. *Vaccines (Basel)*. 2020, 8(2), 273.
- Suschak, J.J.**, Bixler, S.L., Badger, C.V., Spik, K.W., Kwilas, S.A., Rossi, F.D., Twenhafel, N., Adams, M.L. és mtsai. A DNA vaccine targeting VEE virus delivered by needle-free jet-injection protects macaques against aerosol challenge. *npj Vaccines*. 2022, 7(1), 46.
- Tie, Y.**, Wang, Y.-F., Boross, P. I., Chiu, T.-Y., Ghosh, A.K., Tózsér, J., Louis, J.M., Harrison, R.W., Weber, I.T. Critical differences in HIV-1 and HIV-2 protease specificity for clinical inhibitors. *Prot. Sci.* 2012, 21(3), 339-350.
- Tózsér J.** HIV inhibitors: problems and reality. *Ann. N. Y. Acad. Sci.* 2001, 946(1), 145-159.
- Timofeev, V.**, Samygina, V. Protein Crystallography: Achievements and Challenges, *Crystals* 2023, 13(1), 71.
- University of California.** The Regents of the University of California. Introduction to Protein Data Bank Format. 2002
- Vaguine, A.A.**, Richelle, J., Wodak, S.J. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D Biol. Crystallogr.* 1999, 55(1), 191-205.
- Vasiljeva, L.**, Merits, A., Golubtsov, A., Sizemskaja, V., Kääriäinen, L., Ahola, T. Regulation of the Sequential Processing of Semliki Forest Virus Replicase Polyprotein. *J. Biol. Chem.* 2003, 278, 41636–41645.
- Vasiljeva, L.**, Valmu, L., Kääriäinen, L., Merits, A. Site-specific Protease Activity of the Carboxyl-terminal Domain of Semliki Forest Virus Replicase Protein nsP2. *J. Biol. Chem.* 2001, 276, 30786–30793.
- Verlet, L.** Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules *Phys. Rev*, 1967, 159, 98–103.
- Wang, J.**, Wang, W., Kollman P.A., Case, D.A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* 2006, 25, 247260.
- Wang, J.**, Wolf, R.M., Caldwell, J.W., Kollman, P.A., Case, D.A. Development and testing of a general AMBER force field. *J. Comput. Chem.* 2004, 25, 1157-1174.
- Waterhouse, A.**, Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumenny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL:

homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018, 46, W296-W303

Webb, B., Sali, A. Protein Structure Modeling Using Modeller. *Curr. Protoc. Bioinform.* 2016, 54(1), 5.6.1-5.6.37.

Woldawer, A., Minor, W., Dauter, Z., Jaskolski, M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J.* 2008, 275(1), 1-21.

Zacks, M.A., Paessler, S. Encephalitic alphaviruses. *Vet. Microbiol.* 2010, 140(3–4), 281–286.

Zhang, H., Harmon, M., Radoshitzky, S.R., Soloveva, V., Kane, C.D., Duplantier, A.J., Ogungbe, I.V. Vinyl Sulfone-Based Inhibitors of Nonstructural Protein 2 Block the Replication of Venezuelan Equine Encephalitis Virus. *ACS Med. Chem. Lett.* 2020, 11(11), 2139–2145.

Zhang, D., Tózsér, J., Waugh, D.S. Molecular cloning, overproduction, purification and biochemical characterization of the p39 nsp2 protease domains encoded by three alphaviruses. *Protein Expr. Purif.* 2009, 64(1), 89–97.

Zhou, Z., Jia, X., Xue, Q., Dou, Z., Ma, Y., Jiang, Z., He, B., Jin, Q., Wang, J., TRIM14 is a mitochondrial adaptor that facilitates retinoic acid-inducible gene-I-like receptor-mediated innate immune response. *PNAS* 2014, 111(2), E245-254.

11. TÁRGYSZAVAK

dinamikai tulajdonságok

dynamic properties

fehérje szerkezet

protein structure

molekuladinamika

molecular dynamics

nem-szerkezeti fehérje

non-structural protein

öninaktivált konformer

self-inactivated conformation

proteáz

protease

számítógépes biokémia

computational biochemistry

szubsztrát-specifitás

substrate specificity

venezuelai ló-láz encephalitis vírus

Venezuelan equine encephalitis virus

12. KÖSZÖNETNYILVÁNÍTÁS

Szeretnék köszönetet mondani témavezetőmnek, **Dr. Mótyán János Andrásnak** a kutatómunkámhoz és doktori dolgozatom elkészüléséhez nyújtott rengeteg segítségért, továbbá a számos pályázati anyagom, előadásom elkészüléséhez nyújtott önzetlen segítségért, ami hatalmas mértékben járult hozzá szakmai fejlődésemhez.

Köszönetet szeretnék mondani **Dr. Tózsér József** intézetvezető professzor úrnak, mind a szakmai segítségért, mind a lehetőségért, hogy a Retrovirális Biokémiai Kutató Laboratórium kutatócsoporthoz csatlakozva kutatómunkát folytathassak, és számos kutatási projektben, konferencián részt vehessek.

Szeretnék köszönetet mondani **Dr. Bozóki Beátának** a doktori munkám elkészüléséhez nélkülözhetetlen szubsztrát-specifitást vizsgáló kísérletes munkák elvégzéséért.

Köszönettel tartozom **Dr. David S. Waugh** professzor úrnak, továbbá **Dr. George T. Lountos**-nak, **Dr. Danielle Needle**-nek és **Dr. Alexander Wlodawer**-nek, a VEEV nsP2pro kristályszerkezetének meghatározásáért, amely nélkül e doktori dolgozat nem születhetett volna meg.

Köszönetet mondanék a Retrovirális Biokémiai Kutató Laboratórium munkatársainak, **Dr. Mohamed Mahdinak**, **Dr. Szabó Andrásnak**, **Dr. Golda Máriának** és **Irene Wanjiru Kiarie**-nek a közös munkáért és a lehetőségért, hogy további érdekfeszítő projekteknél is részt vehettem. Köszönettel gondolok a munkacsoport további tagjaira, megköszöném a segítséget **Miltner Noéminek**, **Ambrus Viktornak**, továbbá **Nagyné Veres Ágotának** és **Janics-Pető Szilviának**. Kiemelt hálával tartozom **Kunkli Balázs Tibornak** a szakmai segítségért, a baráti beszélgetésekért és közös pályázatunk sikeréért.

Köszönettel tartozom **Dr. Shina Caroline Lynn Kamerlin** professzor asszonynak, mind az EBSA ösztöndíjas lehetőségért, mind a számításműködés kémia területén nyújtott rengeteg segítségért, tanácsért.

Szeretnék köszönetet nyújtani tanulmányaim során korábbi mentoraimnak, **Dr. Mándi Attilának**, aki megismertetett a számításműködés kémia alapjaival. Köszönetet mondanék **Dr. Fuxreiter Mónika** professzor asszonynak a szakmai segítségért. Hálával szeretnék megemlékezni **Dr. Komáromi Istvánról**[†], aki megismertetett a fehérjemodellezés számos módszerével.

Szeretném megköszönni **Dr. Fésüs László** professzor úrnak, **Dr. Karancsiné Dr. Menyhárd Dórának** és **Dr. Fehér Krisztinának** a komplex vizsgán nyújtott szakmai tanácsokat, ami jelentős mértékben járult hozzá a doktori dolgozat színvonalának emeléséhez.

Köszönet szeretnék mondani **Dr. Fekete Attilának** a számításhoz kémia területén nyújtott tanácsokért, és a szuperszámítógépek használatában nyújtott segítségért, türelméért. Szeretném megköszönni a számos szakmai tanácsot **Dr. Miskei Mártonnak**. Köszönettel tartozom **Dr. Hajdu Péter Bélának** a fizika tudományával kapcsolatos számos tanácsáért. Köszönettel tartozom **Dr. Kurtán Tibor** professzor úrnak, aki egyetemi tanulmányaim alatt segítette munkám. Köszönet illeti **Tornyai Ilonát**, a bioinformatikai módszerekkel kapcsolatos tanácsaiért.

Szeretnék köszönetet mondani **Dr. Tüü-Szabó Boldizsárnak**, **Duró Norbertnek** és **Kertiné Ferenczi Renátának** a korábbi kutatómunkám során nyújtott segítségért.

Köszönettel tartozom **Tölli Lajosnak** a számítástechnikai segítségért.

Köszönettel tartozom **Oláh Zsuzsannának**, **Szabó Orsolyának**, **Nagy Anettnek** és **Vida Ildikónak** az adminisztratív feladatokkal kapcsolatos segítségért.

Köszönetet mondok családom tagjainak is. Különösen nagy hálával tartozom **édesanyámnak** a tanulmányaim során nyújtott támogatásért, türelméért. Szeretném megköszönni a rengeteg önzetlen segítséget és támogatást **nagynénémnek**. Őszinte hálával gondolok **nővéremre** és **bátyámra**.

Őszinte hálával gondolok **édesapámra**, aki mindig támogatott tanulmányaimban, és bátorított a felmerülő nehézségekkel szemben. Emlékét mindig őrizni fogom.

A doktori dolgozatban tárgyalt munkák a felsorolt pályázatok támogatásával készültek: Adatintenzív és Nyitott Tudomány Program ösztöndíj (Hoffka Gyula részére), Tématerületi Kiválósági Program (TKP2021-EGA-20), MTA Post-covid jelenségek kutatására irányuló pályázati program (POST-COVID₂₀₂₁₋₁₆), „Idea to Life” Tudományos Ötletfejlesztési Program (EFOP-3.6.1-16-2016-00022) és EBSA Bursary (Hoffka Gyula részére). A Kulturális és Innovációs Minisztérium ÚNKP-23-5-DE-486 kódszámú (Mótyán János András részére) és ÚNKP-23-3-II-DE-456 kódszámú (Hoffka Gyula részére) Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült, továbbá Mótyán János András munkáját az MTA Bolyai János Kutatási Ösztöndíj támogatja (BO/00110/23/5).

Köszönetet mondok a KIFÜ által biztosított számítási időért.

13. PUBLIKÁCIÓS LISTA



**DEBRECENI
EGYETEM**

**DEBRECENI EGYETEM
EGYETEMI ÉS NEMZETI KÖNYVTÁR**

H-4002 Debrecen, Egyetem tér 1, Pf.: 400
Tel.: 52/410-443, e-mail: publikaciok@lib.unideb.hu

Nyilvántartási szám: DEENK/302/2024.PL
Tárgy: PhD Publikációs Lista

Jelölt: Hoffka Gyula
Doktori Iskola: Molekuláris Sejt- és Immunbiológia Doktori Iskola
MTMT azonosító: 10080018

A PhD értekezés alapjául szolgáló közlemények

1. **Hoffka, G.**, Lountos, G. T., Needle, D., Wlodawer, A., Waugh, D. S., Tózsér, J., Mótyán, J. A.:
Self-inhibited state of Venezuelan equine encephalitis virus (VEEV) nSP2 cysteine protease:
a crystallographic and molecular dynamics analysis.
J. Mol. Biol. 435 (6), 1-20, 2023.
DOI: <http://dx.doi.org/10.1016/j.jmb.2023.168012>
IF: 5.6 (2022)
2. Bozóki, B., Mótyán, J. A., **Hoffka, G.**, Waugh, D. S., Tózsér, J.: Specificity Studies of the
Venezuelan Equine Encephalitis Virus Non-Structural Protein 2 Protease Using Recombinant
Fluorescent Substrates.
Int. J. Mol. Sci. 21 (20), 1-26, 2020.
DOI: <http://dx.doi.org/10.3390/ijms21207686>
IF: 5.924

További közlemények

3. Golda, M., **Hoffka, G.**, Cherry, S., Tropea, J. E., Lountos, G. T., Waugh, D. S., Wlodawer, A.,
Tózsér, J., Mótyán, J. A.: P1' specificity of the S219V/R203G mutant tobacco etch virus
protease.
Proteins. [Epub ahead of print], 2024.
DOI: <http://dx.doi.org/10.1002/prot.26693>
IF: 2.9 (2022)
4. Mótyán, J. A., Mahdi, M., **Hoffka, G.**, Tózsér, J.: Potential Resistance of SARS-CoV-2 Main
Protease (Mpro) against Protease Inhibitors: lessons Learned from HIV-1 Protease.
Int. J. Mol. Sci. 23 (7), 1-20, 2022.
DOI: <http://dx.doi.org/10.3390/ijms23073507>
IF: 5.6





5. Ambrus, V. A., **Hoffka, G.**, Fuxreiter, M.: Asymmetric dynamic coupling promotes alternative evolutionary pathways in an enzyme dimer.
Sci. Rep. 10 (1), 1-9, 2020.
DOI: <http://dx.doi.org/10.1038/s41598-020-75772-5>
IF: 4.379
6. Tüü-Szabó, B., **Hoffka, G.**, Duró, N., Fuxreiter, M.: Altered dynamics may drift pathological fibrillization in membraneless organelles.
Biochim. Biophys. Acta. Proteins Proteom. 1867 (10), 988-998, 2019.
DOI: <http://dx.doi.org/10.1016/j.bbapap.2019.04.005>
IF: 2.371
7. Ferenczi, R., Illyés, T. Z., Király, S. B., **Hoffka, G.**, Szilágyi, L., Mándi, A., Antus, S., Kurtán, T.: Evaluation of Different Synthetic Routes to (2R,3R)-3-Hydroxymethyl-2-(4-hydroxy-3-methoxyphenyl)-1,4-Benzodioxane-6-Carbaldehyde.
Curr. Org. Chem. 23 (26), 2960-2968, 2019.
DOI: <http://dx.doi.org/10.2174/1385272823666191212113407>
IF: 1.933

A közlő folyóiratok összesített impakt faktora: 28,707

**A közlő folyóiratok összesített impakt faktora (az értekezés alapjául szolgáló közleményekre):
11,524**

A DEENK a Jelölt által az iDEa Tudóstérbe feltöltött adatok bibliográfiai és tudományometriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.

Debrecen, 2024.05.24.



14. FÜGGELÉK

A doktori disszertáció alapjául szolgáló közlemények különlenyomatai.