

**EGYETEMI DOKTORI (PhD) ÉRTEKEZÉS**

**A papain enzimmechanizmusának elméleti vizsgálata és a  
humán transzglutaminázok aktivációjának korai eseményei**

**Fekete Attila**

**Témavezető: Dr. Komáromi István<sup>†</sup>**



**DEBRECENI EGYETEM  
LAKI KÁLMÁN DOKTORI ISKOLA  
Debrecen, 2019**

A világhírű tudós élete Isten kezében van.

- Rejtő Jenő

## Tartalomjegyzék

1. Bevezetés.....	5
2. Irodalmi áttekintés.....	6
2.1 A papain elméleti vizsgálatának előzményei.....	6
2.2 A transzglutaminázok vizsgálatának előzményei.....	8
3. Az alkalmazott módszerek alapjai.....	11
3.1 Az alkalmazott kvantumkémiai módszerekről.....	11
3.2 Az ONIOM (QM/MM) módszer.....	15
3.3 Molekuladinamika.....	18
4. Célkitűzések.....	20
4.1 A papain teljes enzimmechanizmusának vizsgálata.....	20
4.2 Az FXIII-A <sub>2</sub> és a TG2 kalcium dependens aktivációjának vizsgálata.....	21
5. Módszerek.....	21
5.1 A papain enzimmechanizmusának vizsgálata.....	21
5.2 Az FXIII-A <sub>2</sub> és a TG2 aktivációjának vizsgálata.....	22
5.2.1 Modellrendszerek.....	22
5.2.2 Molekuladinamikai szimulációk.....	24
5.2.3 A molekuladinamikai szimulációk kiértékelése.....	26
6. Eredmények és diszkusszió.....	26
6.1 A papain enzimmechanizmusa.....	26
6.1.1 A papain nyugvó állapotának vizsgálata.....	26
6.1.2 Az acil-enzim intermedier képződése.....	31
6.1.3 Az acil-enzim intermedier hidrolízise.....	39
6.2 Az FXIII-A <sub>2</sub> és TG2 aktivációja.....	43
6.2.1 A molekuladinamikai szimulációk eredmények.....	43
6.2.2 A kalciumkötődés vizsgálata.....	48
6.2.3 Korrelált mozgások analízise.....	56
6.2.4 Főkomponens analízis és a központi domének közötti sóhíd hálózat szerepe.....	58
7. Konklúzió.....	60
7.1 A papain enzimmechanizmusa.....	60
7.2 Az FXIII-A <sub>2</sub> és a TG2 aktivációjának korai eseményei.....	61
8. A jelölt saját eredményei, új megállapításai.....	63
9. Összefoglalás.....	64
10. Summary.....	65

11. Irodalomjegyzék.....	66
12. Tárgyszavak.....	78
13. Köszönetnyilvánítás.....	79
14. Függelék.....	80
14.1 Kiegészítő ábrák.....	80
14.2 Publikációs lista.....	107
14.3. Közlemények különnyomatai.....	107

## 1. Bevezetés

A Debreceni Egyetem, Általános Orvostudományi Kar, Laboratóriumi Medicina Intézet, Klinikai Laboratóriumi Tanszékén végzett kutatómunkáim közül, a doktori disszertációm alapjául két téma szolgált. Az első a cisztein proteázok enzimechanizmusának elméleti kémiai módszerekkel történő vizsgálata a papain enzim példáján keresztül, míg a második téma a papain-szerű katalitikus doménnel rendelkező, humán transzglutaminázok aktivációjának vizsgálata a véralvadási kaszkád XIII-as faktorának A alegységén és a szöveti transzglutaminázon keresztül klasszikus molekuladinamikai módszerek segítségével.

Jóllehet, az utóbbi évtizedekben óriási fejlődésen esett át a sűrűségfüggő kémia, melynek oka részben az, hogy a meglehetősen intenzív számítási időt, amit a poszt Hartree-Fock módszerek által szolgáltatott szerkezeti, termodinamikai, spektroszkópiai és más tulajdonságokra kellene szánni, jó közelítéssel redukált számítási idő mellett tudjuk előállítani. Munkámhoz ideális alapul szolgált a papain enzim, mivel mind kísérletesen, mind pedig elméleti módszerekkel széles körben tanulmányozott, azonban több szempontból is a teljesség igénye nélkül. Modern sűrűségfüggő módszerek alkalmazásával, a báziskészletek méretének szisztematikus növelésével, geometria optimalizálással, frekvenciaanalízissel igazolt stacionárius pontokat állítottunk elő az enzimreakció potenciális energiafelületeit. Az ONIOM-EE típusú kombinált kvantummechanikai/molekulamechanikai (QM/MM) számításainkban, az enzimechanizmust az enzim nyugvó állapotából kiindulva az acil-enzim intermedier képződésén át annak hidrolíziséig vizsgáltuk.

Nem lehetünk érdektelenek a természetben lejátszódó folyamatok dinamikus viselkedésével szemben. A klasszikus mechanika jó segítséget nyújthat az olyan időtől függő jelenségek vizsgálatában, melyeket egyelőre máshogyan egyáltalán nem, vagy csak körülményesen tudnánk vizsgálni. Célszerűen, a makromolekulák esetében pikoszekundum és miliszekundum közötti időskálán lejátszódó események tanulmányozásánál alkalmazható módszer a molekuladinamika (MD). Jelenleg ezen módszer határát az elérhető számítási kapacitás, a klasszikus erőterek jósága, a számítási problémák skálázhatósága, valamint a tanulmányozni kívánt rendszer mérete szabja meg. Ezek együttesen definiálják azt az eseményablakot, mely számunkra információval szolgálhat a tanulmányozni kívánt atomi szintű változásokról. Figyelembe véve a számítási erőforrások növekedését és az elérhető szoftverek intenzív fejlesztését, feltehetően ezen módszerek jelentősége a jövőben még inkább nőni fog.

A XIII-as véralvadási faktor A alegysége (FXIII-A) és a szöveti transzglutamináz (TG2) kapcsán azon dinamikus jelenségek kerülnek tárgyalásra, melyek az aktivációjuk során bekövetkező nagy szerkezeti változásokhoz vezetnek, és amelyeket feltehetően a kalciumionok kötődése indukál

mikroszekundumos időskálán. Molekuladinamikai szimulációkat ezen rendszerekre nem, vagy csak nagyságrendekkel rövidebb időskálán végeztek. Kiemelendő, hogy magát, az ionok kötődését és azok jelenlétének tulajdonítható dinamikus változásokat még egyáltalán nem tárgyalták elméleti módszerek segítségével.

## 2. Irodalmi áttekintés

### 2.1 A papain elméleti vizsgálatának előzményei

A papain enzim (EC 3.4.22.2) a cisztein endopeptidázok C1 szuperfamilájának tagja. 1970-ben, részben az akkor újonnan nyert krisztallográfiai adatok alapján [1], Lowe javasolt egy mechanizmust a peptidkötés hasítására a szerin proteázok analógiáján [2]. Ugyanakkor a közölt eredmények alapján nyilvánvaló, hogy egyetlen, általánosan érvényes reakciómechanizmus nem írható fel ezen enzimekre [3–5].

Mind a szerin, mind a cisztein proteázok acil-enzim (AE) köztitermék képződésén keresztül hasítják a peptidkötést (ami rendre észter vagy tioészter) [4]. Lényeges különbség a kimotripszin és a papain között, hogy míg az első esetben a katalitikus triádot az Asp102, His57 és a Ser195 aminosavak alkotják, úgy a második esetben az Asn175, His159, Cys25 aminosavak találhatóak az katalitikus üregben. A nukleofil centrumot alkotó aminosavon túl lényeges különbség, hogy az Asp102-nek megfelelő aminosav a papainban a megfelelő amid, azaz aszparagin (Asn) [4], amely hozzájárul az enzim nyugvó állapotában jelen lévő tiolát-imidazólium ionpár forma stabilizálásához. Ugyanakkor az aszparagin hatása kevésbé szignifikáns, mint a szerin proteázok esetében az aszpartáté. Másik lényeges különbség, hogy a cisztein aminosav 4-8 pH tartományban deprotonált formában van jelen [6–9], mely a mechanizmus szempontjából fontos, hiszen a deprotonált cisztein jelenléte esszenciális a nukleofil támadás szempontjából.

Ha feltételezzük, hogy az aktív hely Cys25 és His159 aminosav-oldalláncai ionpárt formálnak a papain-szerű cisztein peptidázok nyugvó állapotában, az oxianionos köztitermék (OI) szerepe már kevésbé fontos (ha egyáltalán létezik). A cisztein proteázok esetében az acil-enzim intermedierben a formális töltés centrumok elmosódhatnak szemben a szerin proteázokkal, ahol az oxianionos köztitermék valóban kialakul. A Gln19 aminosav mely fontos szerepet tölt be a tetraédes köztitermék stabilizálásában is, cseréje viszont csekély mértékben befolyásolja az első (acilezési) lépést a reakciókban [10,11].

Az elmúlt négy évtized alatt számos elméleti tanulmány született az említett mechanizmus részleteinek tisztázására [12,12–26]. Az eredetitől [2] eltérő javaslatok is születettek [14,17,18], melyekben a protonátmenet a hisztidin imidazol gyűrűjéről a peptidkötés nitrogénjére koncertikusan történik a cisztein kenének támadásával, vagy megelőzi azt. Több ízben javaslatot tettek arra is,

hogy az aktív centrum ciszteinje protonált formában is jelen lehet [21,24,26]. Suhai és munkatársai, HF/B3LYP:AMBER hibrid potenciál segítségével végzett hibrid számításaikban úgy találták, hogy a tetraédes intermedier csak abban az esetben stabil, ha az amid nitrogén már protonálódott az imidazólium oldalláncról [18].

Különböző protein környezetekben vizsgálták a cisztein-hisztidin diád protonáltsági állapotát is [27–30]. Bizonyos esetekben észszerű volt megvizsgálni, az aktív helyet környező vizek szerepét a protontranszferben [30]. Ez alapján tanulmányozták a szubsztrát hatását a protonátmenet aktiválási gátjára és egyensúlyára vonatkozóan [28]. Összehasonlító QM/MM számítások segítségével vizsgálták az oldószer hatását implicit modell segítségével is. Egyszerű modell rendszerek kondenzált fázisban történő vizsgálatánál úgy találták, hogy a hemi-tioacetál anion mely megfelel a tetraédes intermediernek, a szerin proteázok esetében nem stabil közttermék [19,20]. Ma és munkatársai kimutatták, hogy a katepszin K esetében az acilezési lépés potenciális energiafelületén a tetraédes közttermék lokális minimumot képvisel, míg a dezacilezési lépésre nem találtak ilyet [22]. Wei és munkatársai az első (acilezési) és a második (dezacilezési) folyamatra is két lépéses reakciókat írtak fel [23].

A cisztein peptidázok esetében a legmagasabb szintű számításokat B3LYP, B-LYP, sűrűségfüggő (DFT) módszerekkel végezték vegyérték dupla vagy tripla zéta típusú báziskészletekkel, polarizációs és diffúz függvényekkel kiegészítve [18,23–25] az energiák számítására. Geometriai optimalásra általánosságban alacsonyabb szintű elméletet alkalmaztak. A QM/MM típusú számításokban a kvantummechanikai rétegre alkalmazott elmélet szintén alacsonyabb volt.

A DFT módszer ideális választás a hibrid számítások kvantumkémiai komponenseként, mivel a relatív és aktiválási energiákat, valamint a szerkezeteket a nagy teljesítőképességű és költséges poszt-Hartree-Fock módszerekkel összemérhetően képesek jósolni. Ugyanakkor a klasszikus sűrűségfüggő hiányossága a diszperziós kölcsönhatások megfelelő leírása, a hosszú távú kicserélődési tag hiánya vagy tapasztalati úton történő paraméterezése [31–33].

Jelenleg számos olyan DFT módszer elérhető, melyekben a London-féle diszperziós korrekciót figyelembe vették [31,32,34,35]. Ezen energiatag egyszerűen és kielégítően számítható a  $C_6/r^6$  pár tagokon keresztül, melyeken megfelelő csillapító függvényt alkalmazva a van der Waals (vdW) párkölcsönhatások helyes aszimptotikus viselkedése megkapható. Ilyen tagot alkalmaznak például az  $\omega$ B97X-D [36] és a B97-D [37] funkcionálokban. Az  $\omega$ B97X-D funkcionál tartalmaz továbbá egy hosszú távú elektron kicserélődési tagot is, amely magyarázhatja jó teljesítőképességét vdW komplexekben [33]. Az egyszerű empirikus diszperziós párkorrekció is jelentősen javítja a reakciók aktiválási gátjának számítását biokémiai reakciók DFT alapú QM/MM számításaiban [43,44].

Szintén megoldás lehet olyan funkcionálok használata is, melyek paramétereit nagy pontosságú energiaadatok felhasználásával optimalizálták. Az M06-2X [38] funkcionál harmincöt finomhangolt paramétert tartalmaz, így tipikus példa ezen funkcionálokra. Ilyen értelemben az M06-2X nem tartalmaz direkt tagot a diszperziós korrekcióra, ugyanakkor paraméterein keresztül felfogható egyfajta diszperzió-korrigált DFT módszernek. Maga a funkcionál jól teljesít az aktiválási energiák és átmeneti állapotok geometriájának számításánál is [39–42].

Nemrégiben megjelent munkákban kimutatták, hogy a megfelelő aktiválási gátak, átmeneti állapot geometriák és az asszinkronitás nagy mértékben függ a sűrűségfunkcionál módszerek és a báziskészletek megválasztásától [41,45].

## 2.2 A transzglutaminázok vizsgálatának előzményei

Mind a XIII-as véralvadási faktor A alegysége, mind a szöveti transzglutamináz a protein-glutamin gamma-glutamintranszferáz enzimek családjába tartozik (EC 2.3.2.13). A plazmában az FXIII (pFXIII)  $A_2B_2$  heterotetramerként van jelen, melyben a B alegységek hordozó szerepet töltenek be, míg a celluláris forma (cFXIII) homodimerként van jelen (FXIII- $A_2$ ) a monociták és vérlemezkék citoplazmájában. A pFXIII- $A_2$  proteolitikus aktivációja során [46] a trombin elhasítja az Arg37 és Gly38 közötti peptidkötést, így szolgáltatva az aktivációs peptideket (AP) [47]. A szarvasmarha  $\alpha$ -trombin által aktivált formában a krisztallográfiai tanulmányok alapján az elhasított N-terminális peptidek továbbra is asszociáltak maradnak a homodimer felszínhez [48], elzárva azon üreg bejáratát, ahol a katalitikus aminosavak találhatóak (Cys314, His373, Asp396) [49]. A placentából vagy plazmából származó pFXIII- $A_2$  aktiválható proteolitikus hasítás nélkül, magas, fiziológiailag nem releváns kalcium koncentráció ( $\geq 50$  mM) segítségével [50]. Ugyanakkor ezen zimogén forma aktiválható magas nátrium-klorid vagy kálium-klorid valamint 2 mM  $Ca^{2+}$  jelenlétében is, habár kalciumionok jelenléte nélkül lassú, progresszív aktiváció figyelhető meg [51]. Fiziológiailag releváns körülmények között, a proteolitikus aktivációt követően az aktivációs peptideknek disszociálniuk kell a homodimer felszínéről [52,53], valamint nagy konformációs változásnak kell bekövetkeznie annak érdekében, hogy a protein enzimatikusan aktív konformációt (FXIII- $A_2^*$ ) vehessen fel [54]. Ilyen nagy mértékű konformációs változás látható azon rekombináns cFXIII krisztallográfias szerkezetében, melyet 300 mM NaCl és 3 mM  $Ca^{2+}$  segítségével aktiváltak proteolitikus hasítás nélkül [55]. Ezen munka alapján az 1-es, valamint a 2-es  $\beta$ -hordó domén térbeli helyzete jelentősen megváltozik, az aktív hely pedig megközelíthetővé válik szubsztrátok számára, miközben maguk a hordó domének változatlanok maradnak. Mindazonáltal ezen doménmozgások alapvetően különböznek attól, melyet a szöveti (TG2) és az epidermális (TG3)

transzglutamináz enzimek esetében láthatunk, illetve szerkezeti homológia alapján feltételezték az FXIII-A<sub>2</sub> esetében is [54]. Összefoglalva elmondható, hogy a kalciumionok egyszerre felelősek a B alegységek és az aktivációs peptidek egymást követő disszociációjáért, valamint az aktív konformáció felvételéért. Bebizonyították, hogy valóban az FXIII-A alegységei felelősek a kalciumionok kötődéséért. A fluoreszcencia mérések alapján a kalciumkötődésének egyensúlyi állandója ( $K_m$ ) 0,12 mM-nak adódott. Továbbá alsó becslést adtak a dimerenként kötött kalciumionok számára vonatkozóan (1,2-1,5 egyensúlyi dialízis kísérletek alapján) és feltételezték dimerenként egy-egy erős kötőhely létezését [56]. Szerkezeti információk fellelhetőek olyan FXIII-A<sub>2</sub> szerkezetekről is, ahol Ca<sup>2+</sup>, Sr<sup>2+</sup> és Yb<sup>3+</sup> ionok kötődnek a dimerhez. Mindhárom ion vdW sugara nagyon közel esik egymáshoz (~ 200 pm), ami magyarázhatja, hogy a fő kötőhelyek (Ala457, Asn436, Asp438, Glu485 és Glu490) mindhárom kation esetében betöltöttek. Az itterbium(III) (Yb<sup>3+</sup>) esetében [57] láthatóak a legszembetűnőbb különbségek, ahol fő kötőhelyeken kívül az alegységek Asp270, Asp271, Glu272 aminosavai együttesen négy Yb<sup>3+</sup> ion számára formálnak kötőhelyet. Ez korábban nem volt megjelölve kalciumkötőhelyként, viszont a Stieler és munkatársai által közölt és aktívnek vélt monomer FXIII-A<sup>0</sup> szerkezet már tartalmaz kalciumiont ezen a helyen [55]. Helyspecifikus mutagenézis kísérletekkel igazolták, hogy a Glu485 és Glu490 valóban esszenciális a katalitikus működés szempontjából, továbbá azt is kimutatták, hogy az Asp476 és Asp479 aminosavak szerepet játszanak az FXIII-A<sub>2</sub> szerkezet konformációs változásaiban azáltal, hogy „kapcsolóznak” tekinthetőek a proteáz-érzékeny és proteáz-ellenálló szerkezeti formák között [58]. Felületi polaritás analízis és <sup>43</sup>Ca mágneses magrezonancia (NMR) kísérletek alapján feltételezhető, hogy gyengébb kötőhelyek is találhatóak a dimeren, illetve az NMR kísérletek alapján a kalciumkötődés disszociációs állandójára 0,51 mM adódott [59]. Később Turner és munkatársai kifogásolták az NMR módszerek alkalmazhatóságát azok limitált teljesítőképessége miatt, ezért H/D kicserélődési kísérleteket végeztek, majd proteolitikus emésztésnek vetették alá a fehérjét, és MALDI-TOF/MS technikával detektáltak peptideket. Az eredmények alapján következtetni lehetett a különböző kísérleti körülmények (zimogén, trombinhasított enzimek, trombinhasított + 1 mM Ca<sup>2+</sup> és zimogén + 50 mM Ca<sup>2+</sup>) által indukált, szerkezeti változásokra, amelyek az adott peptid régiók megnövekedett deuterálódásában nyilvánultak meg [60]. Az *in vitro* aktivált rekombináns zimogén FXIII-A szerkezetét Stieler és munkatársai közölték és ahogyan azt korábban említettük, nagy doménmozgások figyelhetőek meg a β-hordók helyzetét illetően [55]. Nemrégiben egy igen kiterjedt szerkezet-funkció tanulmány született, melyben számítási módszerekkel generáltak útvonalat az inaktív (zárt) és az aktívnek vélt (nyitott) FXIII-A esetében, valamint tanulmányozták a modellezett B alegységek kölcsönhatását az A dimerekkel [61]. A FXIII szerkezetéről és funkciójáról részletes összefoglalók találhatóak az irodalomban, melyek azonban nem képezik szorosan tárgyát jelen értekezésnek [54,54,62].

A transzglutaminázok családjának egy másik tagja, melyet jelen munkában vizsgálni kívántam, a szöveti transzglutamináz, ami monomer formában megtalálható különböző kompartmentekben, beleértve a plazma membránt, a citoszolt és a sejtmagot is. Mind az FXIII-A alegység, mind a TG2 esetében közös jellemző a kalciumionok által indukált nagy mértékű konformációs változás, azaz a transzglutamináz aktivitás kizárólag kalciumionok jelenlétében jelentkezik. Jóllehet számos esetben sikerült kísérletes úton eredményeket előállítani a TG2 szerkezetéről, GDP [63], GTP [64] és ATP [65] jelenlétében, sőt még az aktív (nyílt) konformációról is [66], azonban mind ez idáig nem érhető el olyan szórásképből feloldott szerkezet, ahol kalciumok lennének láthatóak. A szöveti transzglutamináz rendelkezik GTPáz aktivitással, és a guanozin-foszfátok kötődése negatívan hat a transzamidáz aktivitásra. Érdekeség, hogy egyensúlyi dialízis kísérletek alapján a kötött kalciumok száma hatra tehető, a GTP pedig csökkenti a kötött kalciumok számát [67]. A mérsékelt szekvencia homológia ellenére [55], szembevetően jól konzerválódott a másod- és harmadlagos szerkezet a nyolc humán transzglutamináz esetében. Ezen szerkezeti homológia alapján, a Yee és munkatársai által közölt FXIII-A<sub>2</sub> szerkezetet [48] felhasználva homológiamodellezéssel előállított TG2 szerkezeten 50 ps hosszúságú, szabad molekuladinamikai szimulációt végeztek, implicit oldószermodellel, és három kalciumion jelenlétében. Feltehetően az alacsony sztöchiometria és a rövid szimulációs idő miatt nem volt tapasztalható jelentős szerkezeti változás, azonban a kalciumok jelenlétében a forgási sugár növekedéséről számoltak be. Az enzim emésztése során a hurok régió a katalitikus központi és az 1-es  $\beta$ -hordó domén között hasítható [68]. A <sup>43</sup>Ca NMR kísérletek alapján a GTP teljesen megszüntette a kalciumionok kötődését [59], ami meglehetősen nehezen értelmezhető szerkezeti alapon. Király és munkatársai körültekintő *in vitro* kísérleteket végeztek, ami során szisztematikusan mutálták azon aminosavakat, melyeket számításmódellenes és a szerkezeti homológia alapján potenciális kötőhelyként azonosítottak. Ez utóbbi esetben az FXIII-A<sub>2</sub> és a TG3 enzimek esetében volt információ a kalcium kötő helyekről és öt olyan helyet azonosítottak melyek mindegyike hozzájárul a kalciumionok kötődéséhez, valamint befolyásolják az enzim GTPáz és a transzglutamináz aktivitását [68]. Fontos kiemelni, hogy a szerkezeti tanulmányok során közölt 1KV3 krisztallográfiai szerkezet [63] glicin aminosavat tartalmaz a 224-es pozícióban, ugyanakkor a natív enzim szekvenciájában ez az aminosav valin (<https://www.uniprot.org/uniprot/P21980>). Ez idáig egyetlen olyan szerkezet érhető el, melyben ugyan megtalálható a natív Val224, azonban a szerkezet három egyéb pontmutációt is tartalmaz, de azok jelenlétének oka az eredeti közleményben nem került tárgyalásra (PDB ID: 4PYG [64]). Ahogyan azt Kanchan és munkatársai kimutatták, feltehetően klónozási hiba eredménye volt a Gly224. Jelentős eltérések mutatkoztak a kalcium affinitásban és a GTPáz aktivitásban, ha a natív Val224 aminosavval expresszálták a fehérjét [69]. Az aktív (nyílt) TG2 szerkezete Pinkas és munkatársai által lett meghatározva, ahol az enzimet 150

mM NaCl és 10 mM Ca<sup>2+</sup>, valamint egy irreverzibilisen kötődő módosított pentapeptid inhibitor jelenlétében inkubálták [66]. A szöveti transzglutamináz protein-protein kölcsönhatásairól összefoglaló található az irodalomban [70].

Összefoglalva elmondható, hogy mind a XIII-as véralvadási faktor A alegysége, mind a szöveti transzglutamináz kísérletileg egyaránt széles körben vizsgált fehérjék, ugyanakkor ezidáig a kalcium-függő aktiváció számos tekintetben felderítetlen maradt. Munkámban egy sor mikroszekundum hosszúságú “*all-atom*” molekuladinamikai szimulációt végeztünk a zimogén, a proteolitikusan aktivált és az aktivációs peptidet nem tartalmazó FXIII-A<sub>2</sub> esetében. A szöveti transzglutamináz vizsgálatánál különös figyelmet szenteltünk a kalciumionok és a GDP/GTP jelenlétének egyedi, illetve azok együttes hatásának tanulmányozására. A Gly224Val pontmutáció dinamikára gyakorolt lehetséges hatásait csupán érintőlegesen kívántam tárgyalni az említett pontmutációk jelenléte miatt.

### 3. Az alkalmazott módszerek alapjai

#### 3.1 Az alkalmazott kvantumkémiail módszerekről

A kvantumkémiail számítások során a Born-Oppenheimer közelítéssel dolgozunk, tehát eltekintünk attól, hogy az elektronok és a magok között csatolás van (a mag-elektron Schrödinger egyenletében megjelenik az elektron hullámfüggvénynek a magkoordináták szerinti deriváltja) és közelítő módszerekkel oldjuk meg az elektron Schrödinger-egyenletét. A számításaim során használt kvantumkémiail módszereket két csoportba sorolhatjuk aszerint, hogy az adott elmélet a hullámfüggvényt vagy az elektronsűrűséget használja fel a számítható tulajdonságok meghatározására. Az előbbire egyszerű példa a Hartree-Fock (HF) módszer, míg az utóbbiba a sűrűségfunkcionál elméleteket (Density Functional Theory, DFT) sorolhatjuk. Mivel értekezésem alapjául kizárólag sűrűségfunkcionál módszerekkel kapott eredményeket ismertetek, így a hullámfüggvény alapú módszereket csak érintőlegesen ismertetem a HF elmélet példáján. Korábbi munkámban az alábbiakon túl a perturbációs számításról is szót ejtettem [132].

A Hartree-Fock módszernél a független részecske közelítés értelmében az  $i$  elektron az összes többi kiátlagolt potenciálterében mozog. Ez a közelítés megjelenik a  $\hat{J}_b(1) = \int dx_2 |\chi_b(2)|^2 r_{12}^{-1}$  lokális Coulomb-operátorban. Ha a  $\hat{J}_b(1)$  operátort hattatjuk a  $\chi_a$  spinpályán lévő ( $1$ ) elektronon,  $b$  szerint szummázva a teljes kiátlagolt potenciált megkapjuk, amely az összes többi spinpályán lévő  $N-1$  elektron jelenlétéből fakad. A rendszer hullámfüggvénye spinpályák antiszimetrizált szorzata  $|\Psi_0\rangle = |\chi_1 \chi_2 \dots \chi_a \chi_b \dots \chi_N\rangle$ . Mivel az  $E_0$  energia funkcionálja a spinpályáknak, így az energia pályák szerinti variálásával minimáljuk az energiát, és megkapjuk a

Hartree-Fock egyenleteket  $f|\chi_a\rangle = \varepsilon_a|\chi_a\rangle$ , ahol  $f$  a Fock operátor, amely egy egyrészescke Hamilton operátor és az  $f(1) = h(1) + \sum_b \hat{J}(1) - K_b(1)$  egyenlet szerint adódik, ahol a  $\hat{J}$  Coulomb- és a  $K_b$  kicserélődési operátor különbsége a  $v^{HF}$  potenciál [71].

Az  $n$ -elektronos molekula hullámfüggvénye  $3n$  térbeli és  $n$  spinkoordináta függvénye. A hullámfüggvény ugyanakkor sok olyan információt is hordoz, aminek kevés a közvetlen fizikai jelentősége. 1964-ben Hohenberg és Kohn bizonyította, hogy az alapállapotú molekula energiája, hullámfüggvénye és minden más elektronos tulajdonsága egyértelműen meghatározott az elektron valószínűségi sűrűsége  $\rho(x, y, z)$  által, mely csupán a három Descartes koordináta függvénye [72].

A Hohenberg-Kohn elmélet ugyanakkor nem mondja meg, hogyan számoljuk ki a  $E_0$ -t  $\rho$ -ból, azaz hogyan találjuk meg a  $\rho$ -t a nélkül, hogy előbb megkeresnénk  $\psi$ -t. A probléma megoldásában az egyik első lépést Kohn és Sham tette, akik megmutatták, hogy az  $n$ -elektronos molekula egzakt alapállapotú elektronenergiája az alapállapotú elektronsűrűséggel ( $\rho$ ) megadható,

$$E_0 = -\frac{1}{2} \sum_{i=1}^n \langle \Psi_i(1) | \nabla_i^2 | \Psi_i(1) \rangle - \sum_{\alpha} \int \frac{Z_{\alpha} \rho(1)}{r_{1\alpha}} dv_1 + \frac{1}{2} \iint \frac{\rho(1)\rho(2)}{r_{12}} dv(1) dv(2) + E_{XC}[\rho]$$

(3.1.8)

ahol  $\Psi_i(1), i=1,2,\dots,n$  a Kohn-Sham pályák és a kicserélődési-korrelációs energia ( $E_{XC}[\rho]$ ), mely  $\rho$  funkcionálja. A fenti képletben  $\Psi_i(1)$  és  $\rho_i(1)$  azt jelentik, hogy  $\Psi_i$  és a  $\rho$  az (1) elektron térbeli koordinátáinak függvénye. Kohn és Sham megmutatta ugyanakkor, hogy az egzakt alapállapotú  $\rho$ -t megtalálhatjuk a  $\Psi_i$ -k segítségével, mivel

$$\rho = \sum_{i=1}^n |\Psi_i|^2 .$$

(3.1.9.)

A Kohn-Sham pályákat megkaphatjuk az egyelektron egyenletekből

$$\hat{F}_{KS}(1) \Psi_i(1) = \varepsilon_{i,KS} \Psi_i(1)$$

(3.1.10.)

ahol az  $\hat{F}_{KS}$  az

$$\hat{F}_{KS} = -\frac{1}{2} \nabla_i^2 - \sum_{\alpha} \frac{Z_{\alpha}}{r_{1\alpha}} + \sum_{j=1}^n \hat{J}_j(1) + V_{XC}(1) .$$

(3.1.11.)

A fenti egyenletben szereplő Coulomb operátor megegyezik a hullámfüggvény alapú módszereknél használt, lokális operátorral, míg a  $V_{xc}$  a kicserélődési-korrelációs potenciál, amely  $E_{xc}$  funkcionál deriváltjával állítható elő:

$$V_{xc} = \delta E_{xc}[\rho] / \delta \rho . \quad (3.1.12.)$$

A Kohn-Sham pályák szerepe pusztán az egzakt  $\rho$  származtatása. A Kohn-Sham egyenletekkel egyedüli probléma az, hogy nem ismerjük a helyes  $E_{xc}[\rho]$  funkcionált, ezért közelítő módszereket kell alkalmaznunk [72].

A kicserélődési-korrelációs funkcionálra számos közelítő eljárást dolgoztak ki, azonban munkánk során kizárólag az általánosított gradiens közelítést (Generalized Gradient Approximation, GGA) vagy gradiens sorfejtést alkalmazó funkcionálok valamely típusát használtam. Ugyanakkor az irodalom meglehetősen bő leírást nyújt az egyes közelítő módszerek leírásáról [73,74].

A Sousa és munkatársai szerint, 2007-ben és az azt megelőző években a legtöbbet idézett sűrűségfunkcionál a B3LYP funkcionál volt [131]. A sűrűségfunkcionálok valóban előkelő helyet foglalnak el a Hartree-Fock és a poszt-HF módszerek között teljesítmény/költség (hatékonyság) tekintetében. Az említett Becke 3-paraméteres Lee-Yang-Parr kicserélődési-korrelációs funkcionál (B3LYP) egy hibrid GGA funkcionál és a következő tagokból áll

$$AE_X^{Slater} + (1-A) E_X^{HF} + B\Delta E_X^{Becke} + E_C^{VWN} + C\Delta E_C^{Non-local} \quad (3.1.13.)$$

Ahol  $A, B, C$  konstansok, melyeket G1 molekula csoport segítségével illesztett Becke. A kifejezés a LYP funkcionál nem lokális korrelációját használja, lokális korrelációra a VWN funkcionál III-at [75,76].

A diszperziós korrekciót tartalmazó funkcionálokról szóló [36] közleményben, adott tesztkészleten a legkisebb átlagos abszolút szórások (Mean Absolute Deviation, MAD) alapján az  $\omega$ B97X-D funkcionál (0,2 kcal/mol) jóval megelőzte az M06-2X (0,4 kcal/mol) funkcionált. A B3LYP itt 3,8 kcal/mol átlagos abszolút szórást produkált. Ezek után ésszerűnek tűnt a számításainkat elvégezni az  $\omega$ B97X-D funkcionállal, mely diszperziós korrekciót is tartalmaz. A DFT-D általános formalizmus szerint

$$E_{DFT-D} = E_{KS-DFT} + E_{Disp}. \quad (3.1.14.)$$

egyenlet szerint az  $E_{KS-DFT}$  energiát az  $\omega B97X$  funkcionállal számítják, melyhez korrekcióként hozzáadják az  $E_{Disp.}$  diszperziós energiát.

$$E_{Disp.} = - \sum_{i=1}^{N_{at}-1} \sum_{j=1+1}^{N_{at}} \frac{C_6^{ij}}{R_{ij}^6} f_{damp}(R_{ij}) \quad (3.1.15.)$$

Ahol  $N_{at}$  az atomok száma a rendszerben,  $C_6^{ij}$  pedig a diszperziós koefficiens az  $i, j$  atompárra,  $R_{ij}$  pedig az atomok közötti távolság. Az  $f_{damp}(R_{ij})$  tag a vdW potenciál helyes aszimptotikus viselkedése végett került bevezetésre:

$$f_{damp}(R_{ij}) = \frac{1}{1 + a(R_{ij}/R_r)^{-12}} \quad (3.1.16.)$$

Az M06-2X funkcionált bemutató cikkben [38] nem közöltek összehasonlítást az  $\omega B97X-D$ -vel, viszont 12 más funkcionállal több adatkészleten végeztek széles körű összehasonlítást.

Az értekezés alapjául szolgáló munkában alkalmazott M06-2X funkcionált kifejezetten nemfémekre paraméterezték. A funkcionál három tagtól függ és összességében 35 finomhangolt paramétert tartalmaz. Az első tag a spinsűrűség ( $\rho_\sigma$ ), a második a redukált spinsűrűség gradiens ( $x_\sigma$ ), míg a harmadik a spin kinetikus energia sűrűség ( $\tau_\sigma$ ). A funkcionálkifejezés formailag megegyezik az M06-L és M06-HF módszerek esetében alkalmazottakkal, valamint az M05 és VSXC korrelációs funkcionálok alapul, melyekben az ellentétes és a párhuzamos spinpárokat külön korrelációval írják le.

Az ellentétes spinek M06 korrelációs energia kifejezése

$$E_C^{\alpha\beta} = \int e_{\alpha\beta}^{UEG} [g_{\alpha\beta}(x_\alpha, x_\beta) + h_{\alpha\beta}(x_{\alpha\beta}, z_{\alpha\beta})] dr \quad (3.1.17.)$$

ahol  $g_{\alpha\beta}(x_\alpha, x_\beta)$  megadható az

$$g_{\alpha\beta}(x_\alpha, x_\beta) = \sum_{i=0}^n c_{c\alpha\beta, i} \left( \frac{\gamma_{C\alpha\beta}(x_\alpha^2 + x_\beta^2)}{1 + \gamma_{C\alpha\beta}(x_\alpha^2 + x_\beta^2)} \right)^i \quad (3.1.18.)$$

Párhuzamos spinekre pedig

$$E_C^{oo} = \int e_{oo}^{UEG} [g_{oo}(x_o) + h_{oo}(x_o, z_o)] D_o dr \quad (3.1.19.)$$

ahol

$$g_{oo}(x_o) = \sum_{i=0}^n c_{C_{oo},i} \left( \frac{\gamma_{C_{oo}} x_o^2}{1 + \gamma_{C_{oo}} x_o^2} \right)^i \quad (3.1.20.)$$

és  $\gamma_{C_{\alpha\beta}} = 0,0031$ ,  $\gamma_{C_{\sigma\sigma}} = 0,06$  konstansok.

Összességében tehát a korrelációs funkcionál a következő egyszerű kifejezéssel adható meg:

$$E_C = E_C^{\alpha\beta} + E_C^{\alpha\alpha} + E_C^{\beta\beta}. \quad (3.1.21.)$$

A hibrid kicserélődési-korrelációs energia pedig a következőképpen írható:

$$E_{XC}^{hyb} = \frac{X}{100} E_X^{HF} + \left( 1 - \frac{X}{100} \right) E_X^{DFT} + E_C^{DFT}, \quad (3.1.22.)$$

ahol  $E_X^{HF}$  a nem lokális Hartree-Fock kicserélődési energia,  $X$  a HF kicserélődés százalékban,  $E_X^{DFT}$  a lokális DFT kicserélődési energia és  $E_C^{DFT}$  a lokális DFT korrelációs energia.

Az M06-2X funkcionál a kötődisszociációs energiák, termokémiai jellemzők, illetve a gyenge kölcsönhatások leírásában mind közül a legjobban teljesített [38].

### 3.2 Az ONIOM (QM/MM) módszer

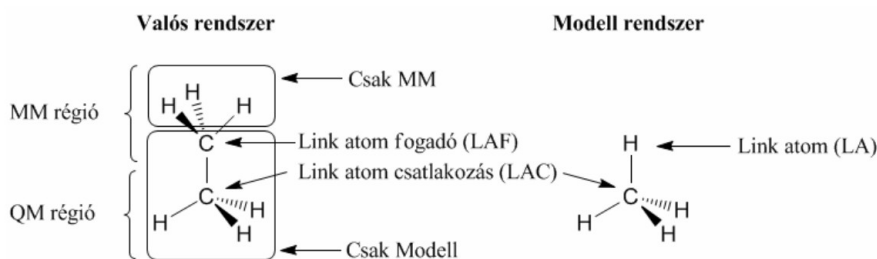
Élettani jelentőséggel bíró makromolekulák (pl: enzimek, DNS) reakcióinak modellezésére a kombinált kvantummechanikai/molekulamechanikai (QM/MM) módszerek ideálisak lehetnek. Az ilyen hibrid számítások során a rendszert több partícióra vagy rétegre osztjuk, és az egyes rétegeket eltérő elméleti szinteken vizsgáljuk. Ilyen módon lehetőség nyílik ezen molekulák reakcióinak vizsgálatára, ahol a rendszer egy részét kvantumkémiai módszerekkel számoljuk, míg a fennmaradó részt molekulamechanikai erőter alkalmazásával írjuk le.

Az ONIOM (our Own N-layered Integrated molecular Orbital molecular Mechanics) sokkal általánosabb a hagyományos QM/MM módszereknél, mivel bármilyen elméleti módszer tetszőlegesen kombinálható, így támogatja kis vagy közepes méretű rendszerek QM/QM vagy akár QM/QM/QM vizsgálatát is.

A különböző QM/MM módszerek alapján véve azonos sémát követnek, azonban számos részletben eltérnek egymástól. Lényeges eltérés lehet például az egymással szoros kölcsönhatásban álló rétegek közötti elektronos kölcsönhatások leírása, másrészt a particionálás során elvágott kovalens kötések lezárására alkalmazott elv.

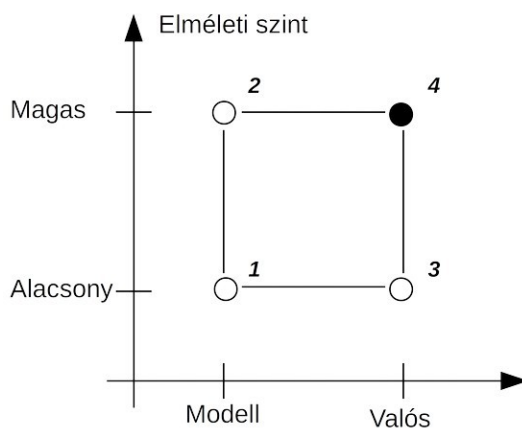
A 1. ábra segítségével könnyen értelmezhetjük az ONIOM-ban alkalmazott terminológiát az etán példáján. Meg kell különböztetnünk a valós és a modell rendszert. A valós rendszer az, amely felöleli a teljes rendszert (összekötő, „link” atom nélkül), míg a modell rendszer csupán a valós rendszer modell részét jelenti („link” atommal). Látható, hogy a QM és az MM régiót úgy hozzuk létre, hogy a szén-szén kötést formálisan elvágjuk, ennek eredményeképpen a QM rendszerben szabad vegyérték jönne létre. Ezt elkerülendő bevezetjük az *LA* link atomot, mely lezárja a szabadon maradt kötést és a *LAC* atommal létesít formálisan kapcsolatot, orientációját tekintve a *LAF-LAC* kötéssel megegyezik. A *LAC-LA* távolság pedig nem más, mint a *LAF-LAC* távolság és egy *g* skálázófaktor szorzata, amely így elfogadható *LAC-LA* távolságot szolgáltat:

$$q_{LA} = q_{LAC} + g(q_{LAF} - q_{LAC}). \quad (3.2.1.)$$



1. ábra. A QM/MM terminológia, az etán példáján.

A link atommal lezárt partícióra a továbbiakban mint „modell rendszer” hivatkozunk. Az ONIOM megközelítés alapötlete egyszerűen szemléltethető az extrapolációs séma segítségével, melyet a rendszer mérete (abszcissza) és az alkalmazott elmélet szintje (ordináta) által kifeszített két dimenziós felületen ábrázolunk (2. ábra).



2. ábra. Az extrapolációs séma a 2 rétegű ONIOM energia kifejezésre. Az ábra egyes pontjai az adott partícióra, adott szinten számolt energiát jelképezik.

Az ONIOM séma (2. ábra) segítségével a két rétegű rendszer extrapolációs energiája az alábbi kifejezés szerint adódik.

$$E_{ONIOM}^2 = E_3 - E_1 + E_2 \quad (3.2.2.)$$

A kétrétegű rendszer extrapolációs energiájának ( $E_{ONIOM}^2$ ) számítása a következő részlépésekből áll. Kiszámítjuk a „valós rendszer” ( $E_3$ ) és a „modell rendszer” ( $E_1$ ) energiáját alacsony és a különbségükhöz hozzáadjuk a „modell rendszer” energiáját, melyet magas szinten ( $E_2$ ) számoltunk.

A magkoordináták szerinti első és második deriváltak, valamint az elektromos tér szerinti magasabb deriváltak igen lényegesek, de itt csak a gradiens meghatározására szorítkozok. A gradiens kifejezés az

$$\frac{E_{ONIOM}^2}{\partial q} = \frac{E_3}{\partial q} + \frac{E_2}{\partial q_{QM}^M} J_{QM} - \frac{E_1}{\partial q_{MM}^M} J_{MM} \quad (3.2.3.)$$

szerint adódik, ahol  $q$  egy vektor, amely tartalmazza a valós rendszer Descartes koordinátáit,  $q_{QM}^M$  és  $q_{MM}^M$  pedig a QM és az MM modell rendszerek Descartes koordinátái. A  $J$  Jacobi mátrix vetíti a link atomok gradiensét a *LAF* és *LAC* koordinátákra. Ezen eljárásból fakadóan nem kerül a rendszerbe több vagy kevesebb szabadsági fok, így az ONIOM módszer jól működik a standard geometria optimaló algoritmusokkal, és közel minden más technikával, ami a potenciális energia hiperfelület felderítésére használható. A magkoordináták szerinti második és az elektromos tér szerinti magasabb deriváltak szintén meghatározhatók.

Az ONIOM módszerben a rétegek közötti elektrosztatikus kölcsönhatást kétféleképpen vehetjük figyelembe: az első a mechanikus (Mechanical Embedding, ME) a második az elektronos beágyazás (Electronic Embedding, EE). A klasszikus vagy mechanikus módszer a rétegek közti elektrosztatikus kölcsönhatást molekulamechanikai szinten számítja. A munkám során használt EE módszerben megengedett a hullámfüggvény molekulamechanikai ponttöltések általi polarizációja.

$$\hat{H}_{QM,Modell} = \hat{H}_{QM,Modell}^{(0)} + \sum_{R'} \tau_{R'} q_{R'} \sum_{Modell} \left( \frac{Z_M}{r_{R',Modell}} \right) - \sum_{R'} \tau_{R'} q_{R'} \left( \sum_{\mu} \frac{1}{r_{R'\mu}} \right), \quad (3.2.4)$$

$$E_{MM,Modell} = E_{MM,Modell}^{(0)} + \sum_{R'} \tau_{R'} q_{R'} \left( \frac{q_M}{r_{R',Modell}} \right), \quad (3.2.5)$$

$$E_{MM,Val\acute{o}s} = E_{MM,Val\acute{o}s}^{(0)} \quad (3.2.6)$$

$$E_{ONIOM}^2 = E_{MM,Val\acute{o}s} + E_{QM,Modell} - E_{MM,Modell} \quad (3.2.7)$$

A 3.2.4 egyenlet szerint a tisztán kvantummechanikailag leírt modellrendszer Hamilton operátorát hagyjuk polarizálódni a molekulamechanikai (alacsony) szinten figyelembe vett atomok  $mag^{MM}$  és  $mag^{QM}$  kölcsönhatásain (jobb oldal, első tag), illetve  $mag^{MM}$  és  $elektron^{QM}$  (jobb oldal, második tag) kölcsönhatás segítségével [77,78].

A munkám során használt Gaussian 03 RevC.01 [79] és Gaussian 09 RevB.01 [80] szoftverekben implementált ONIOM módszer segítségével tetszőleges elméletek kombinálhatóak, melyek lehetnek kvantumkémiai (HF, poszt-HF, DFT, szemiempírikus) és molekulamechanikai módszerek (UFF, AMBER), valamint ezek tetszőleges kombinációja maximum három réteg, illetve módszer erejéig.

### 3.3 Molekuladinamika

A molekuladinamikai (MD) szimulációk a klasszikus mechanikából ismert newtoni mozgásegyenleteket oldják meg az  $N$  darab kölcsönható atomból álló rendszerben

$$m_i = \frac{\partial^2 r_i}{\partial t^2} = F_i, i = 1, \dots, N. \quad (3.3.1)$$

Tekintettel arra, hogy a 3.3.1 kifejezésben megjelenik az idő ( $t$ ), az egyenleteket numerikusan kiintegrálva a rendszer időben való fejlődése megkapható. Az erők ( $F_i$ ) a  $V$  potenciálfüggvény negatív deriváltjai

$$F_i = -\frac{\partial V}{\partial r_i}. \quad (3.3.2)$$

A szimuláció során ezen egyenleteket szimultán oldjuk meg adott kis időközönként. A rendszer koordinátákat adott időközönként eltároljuk, az NpT szimulációk során a hőmérsékletet és a nyomást termoszáttal és barosztáttal kontrolláljuk. A koordináták az idő függvényében reprezentálják a rendszer trajektóriáját. Kezdeti változások után a rendszer optimális esetben bejárja az adott paramétereknek (nyomás, hőmérséklet, részecskeszám) megfelelő egyensúlyi állapotokat. Ezt követően az „egyensúlyi trajektória” analízisével sok makroszkopikus tulajdonság várható

értéke számítható és értelmezhetőek olyan jelenségek, melyek a kísérleti módszerek számára a legtöbb esetben rejtve maradnának.

A molekuladinamikai szimulációk folyamata a következő lépésekből tevődik össze. Az első lépés a kezdeti bemeneti körülmények megadása: a  $V$  potenciálfüggvény egy adott molekulamechanikai erőter alkalmazása esetén az atomok helyzetének függvénye, szükséges a rendszert alkotó atomok  $r$  koordinátáinak ismerete és az atomok kezdeti sebessége. Második

lépésben az erőket kell kiszámítani. Az  $i$ -edik atomon ható erő az  $F_i = \frac{\partial V}{\partial r_i}$  szerint adható meg,

számítása a nem kötött kölcsönhatásban lévő atompárok közti erők összegzésével  $F_i = \sum_j F_{ij}$

történik, amihez hozzáadódik még a kötött kölcsönhatásokból (1-4 atomtól függő tagok) eredő, valamint a külső és egyéb megszorításokból származó erők. Ezek után a kinetikus és potenciális energiák, valamint a nyomás tenzor kerül kiszámításra.

Az eljárás harmadik lépésében frissítjük a rendszer konfigurációját (koordinátáit), az atomok elmozdulásának szimulációja a newtoni mozgásegyenletek numerikus megoldásán alapul

$\frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i}$  vagy  $\frac{dr_i}{dt} = v_i; \frac{dv_i}{dt} = \frac{F_i}{m_i}$  szerint. A negyedik lépés a kimeneti pont, amikor mentésre

kerülhetnek a koordináták, sebességek, energiák, hőmérséklet és a nyomás. Innen az algoritmus pedig visszatérhet a második lépéshez, ahonnan kezdődik a ciklus előlről. A szimulációk időtartamától függően akár 10-10000 millió számítási ciklusra is szükségünk lehet. A szimuláció elindításának elengedhetetlen feltétele a rendszer topológiájának megléte és a molekulamechanikai erőter, (Force Field, FF), amely az atomok közötti kölcsönhatásokat definiálja. Az erőter alatt egyrészt a potenciálfüggvényt (mely alapján a rendszer az energiákat, annak deriváltjait és az erőket számolja), valamint a hozzátartozó paraméterkészletet értjük. A teljes potenciálfüggvény kötő és nemkötő kölcsönhatásokat leíró tagokat tartalmaz:

$$\sum E_{teljes} = \sum E_{kötő} + \sum E_{nemkötő} , \quad (3.3.3.)$$

ahol az egyes tagokat külön írva kapjuk:

$$\sum E_{kötő} = \sum E_{kötésnyújtás} + \sum E_{szöghajlítás} + \sum E_{torziós} \quad (3.3.4.)$$

és

$$\sum E_{nemkötő} = \sum E_{vdW} + \sum E_{Coulomb} . \quad (3.3.5)$$

A kötött kölcsönhatások leírására sok esetben az egyszerű

$$E_{nyújtás} = \sum_i k_r (r_i - r_0)^2 \quad (3.3.6.)$$

$$E_{hajlítás} = \sum_i k_\theta (\Theta_i - \Theta_0)^2 \quad (3.3.7.)$$

$$E_{torziós} = \sum_i \left\{ \frac{V_1^i}{2} [1 + \cos(\varphi_i)] + \frac{V_2^i}{2} [1 + \cos(2\varphi_i)] + \frac{V_3^i}{2} [1 + \cos(3\varphi_i)] \right\} \quad (3.3.8.)$$

kifejezéseket használják. A 3.3.6 és a 3.3.7 egyenletben szereplő  $k$  együttható tulajdonképpen a klasszikus mechanikából ismert rugóállandónak felel meg.

A nem kötött kölcsönhatások gyakran használt alakjai a következők:

$$E_{vdW(Lennard-Jones)} = \sum_{ij} \left( A_{ij} \frac{r_{ij}^{12}}{r_{ij}^{12}} - B_{ij} \frac{r_{ij}^6}{r_{ij}^6} \right) \quad (3.3.9.)$$

és

$$E_{Coulomb} = \sum_{ij} q_i \frac{q_j}{4\pi\epsilon_0 r_{ij}}. \quad (3.3.10.)$$

A molekulamechanikai erőter megalkotásánál, az atomokat atomtípusokra osztják és minden egyes atomtípusra, a paraméterek információkat tartalmaznak az atomtömegről, van der Waals paramétereiről, referencia szögekről és kötéshosszokról, a diédes szögekhez tartozó torziós potenciál paramétereiről és ezek erőállandóiról [81].

## 4. Célkitűzések

### 4.1 A papain teljes enzimmechanizmusának vizsgálata

A papain enzimmechanizmusa kapcsán fő célunk volt, hogy megvizsgáljuk hogyan teljesítenek az egyes diszperziós korrekciót tartalmazó funkcionálok az ONIOM [77] típusú hibrid számítások QM tagjaként, továbbá összehasonlítottuk őket az ilyen korrekciót egyáltalán nem tartalmazó standard B3LYP módszerrel kapott eredményekkel. Megvizsgáltuk továbbá, hogy ezen funkcionálokkal esetleg jósolható-e diverz reakcióút, vagy csupán igazolásra kerülnek korábbi elméletek a papain enzimmechanizmusára vonatkozóan. A legjelentősebb ilyen eltérés lehet az

oxianionos és ikerionos tetraédes intermedier létezése az acilezési lépés potenciális energiafelületén.

Annak érdekében, hogy igazoljuk azt, hogy a QM/MM (ONIOM) potenciális energia felületen valóban lokális minimumokat vagy elsőrendű nyeregpontokat határoztam meg, minden esetben kiszámoltuk a geometria szerinti második deriváltakat [77]. A számolásokat a Gaussian szoftverben implementált „Electronic Embedding” (EE) ONIOM módszerrel végeztem [78].

#### 4.2 Az FXIII-A<sub>2</sub> és a TG2 kalcium dependens aktivációjának vizsgálata

Figyelembe véve azon faktorok sokaságát, melyek szerepet játszanak az FXIII-A<sub>2</sub> és a TG2 fehérjék kalciumionok kötődése által indukált konformációs aktivációjában, célul tűztük ki ezen kezdeti események részleteinek vizsgálatát. Így az FXIII-A<sub>2</sub> esetében vizsgálni kívántuk a zimogén, a protetikusan aktivált és az AP feltételezett disszociációja után visszamaradt – még mindig zárt, azaz inaktív – homodimer dinamikai tulajdonságait. Vizsgálni kívántuk a kialakuló kalciumkötőhelyeket, különös tekintettel a fő kötőzsebre.

Hasonló részletességgel vizsgáltuk a TG2 esetében a GDP hatását a szöveti transzglutamináz dinamikájára, valamint a kalciumkötődésre, továbbá a dinamikai trajektóriákból nyert adatokat összehasonlítottuk a releváns kísérleti előzményekkel. A kovariancia mátrixok és főkomponens analízisek segítségével, azonosítani kívántuk azon szerkezeti elemeket melyeknek kiemelt szerep tulajdonítható az aktiváció során. Érintőlegesen vizsgáltuk az FXIII-A<sup>0</sup> és TG2 aktív formáit, azok kalcium kötő képességét. Kiegészítésként a munka során alkalmazott gömb-szerű kationokkal történő szimulációkhoz, két esetben vizsgálni kívántuk a kialakult kötőhelyek relevanciáját olyan ionmodell alkalmazásával, amelyben a kalciumion koordinációját töltéssel bíró színlelt (dummy) atomokon keresztül vesszük figyelembe („*multi-site*” modell).

## 5. Módszerek

### 5.1 A papain enzimmechanizmusának vizsgálata

Az enzimreakciók vizsgálatánál az 1PPN [82] krisztallográfiai úton meghatározott szerkezetéből indultam ki, és 250 ns hosszúságú NpT molekuladinamikai szimulációkat futtattunk, mind a neutrális (cisztein-hisztidin), mind az ionpár formában jelen levő aminosavakat tartalmazó fehérjére. Dodekaédes szimulációs cellában, TIP3P explicit víz modellel szolvatálva a neutralizált dobozban 150 mM NaCl koncentráció mellett, Amber FF99SB erőterrel [83] Berendsen barosztát [84] és velocity-rescaling termosztát [85] alkalmazásával kerültek futtatásra a szimulációk, melyekhez a GROMACS [81] szoftvert használtunk. A rövid távú és van der Waals energiatagodokra

10 Å cut-off távolságot használtam, továbbá Ewald-féle összegzés (Particle Mesh Ewald, PME) módszert [86] alkalmaztam a hosszútávú elektrosztatikus energiák számítására.

Mivel a krisztallográfias szerkezet nem tért el jelentősen a dinamikai trajektóriából származtatottól – leszámítva dinamika hőmérséklet függő fluktuációját –, így az előbbi szerkezetet az AMBER 12 sander [87] szoftverrel optimáltuk, majd használtuk fel a QM/MM (ONIOM-EE) számításokhoz.

A mechanizmus részlépéseinek értelmezéséhez használt ábrákon pálcika vagy golyó-pálcika modellekben az atomok jelölése a következő: fehérrel a hidrogént, kékkel a nitrogént, pirossal az oxigént, feketével vagy semleges színnel a szenet illetve sárgával a kenet jelöltük.

A katalitikus centrum  $\sim 15$  Å távolságon belüli környezetére (Val16-Ala30, Ser60-Trp69, Val130-Ala136, Lys156-Ala162, Ile173-Gly178, Thr204-Phe207) csonkítottam a teljes rendszert, a létrejövő N- és C-terminálisokat acetyl (ACE) és N-metil (NME) csoportokkal zártam le. Ezen rendszer Gln19-Gly-Ser-Cys-Gly-Ser-Cys25, Gly62-Cys-Asn-Gly-Gly66, Asp158-His159, Asn175-Ser-Trp177 aminosavait optimáltam a potenciális energiefelületek pásztázásánál is. A terminális ACE és NME csoportok atomjai úgy lettek elhelyezve, hogy azok rendre az  $(n-1)$  és  $(n+1)$ -ik aminosav N, C $\alpha$  illetve C, O, C $\alpha$  atomjaival azonos pozícióban legyenek. A létrejött csonkított „fehérje” tartalmazott minden aminosavat, melynek hatása lehet a mechanizmusra. A szubsztrát N-metil-acetamiddal (NMA) együtt összesen 780 atomot tartalmazott ez a kiindulási modell. Miután a reakció minden elemi lépése felderítésre került, és rendelkezésre álltak a stacionárius pontok geometriái, a trunkált rendszer vissza lett illesztve az eredeti fehérjébe a VMD [88] szoftver segítségével. A továbbiakban kizárólag a teljes rendszer hozzájárulásával nyert adatok kerülnek tárgyalásra.

A QM (magas szintű) rétegre a B97-D [37], M06-2X [38],  $\omega$ B97X-D [36] és a B3LYP [75,89] módszereket, valamint a 6-31G(d,p), 6-31+G(d,p), 6-311+G(d,p) és 6-311++G(d,p) Pople típusú báziskészleteket használtam. Az MM (alacsony szintű) rétegre az Amber erőteret használtam, mely implementálva van a Gaussian szoftverben.

## 5.2 Az FXIII-A<sub>2</sub> és a TG2 aktivációjának vizsgálata

### 5.2.1 Modellrendszerek

Annak érdekében, hogy az FXIII-A<sub>2</sub> kalciumkötődését megfelelő alapossággal tudjuk vizsgálni három szimulációs csoportot hoztunk létre. Az első csoportban (i) a zimogén enzim két formáját tanulmányoztuk. Az *i/a* modell alapjául a kötött kalciumokat tartalmazó XRD szerkezet szolgált (PDB ID: 1GGU [57]), a rendszer teljes kalcium koncentrációja 14 mM-ra lett beállítva. Az *i/b* modellben a kalciumot nem tartalmazó szerkezetet vizsgáltuk (PDB ID: 1F13 [90]), és a

rendszer kalcium koncentrációja ebben az esetben 50 mM volt. A második szimulációs csoportban (ii) az FXIII-A<sub>2</sub> a thrombin által aktivált fehérjéket tanulmányoztuk, oly módon, hogy felhasítjuk az Arg37 és Gly38 közötti kötést. Az ii/a és ii/b modellek az i/a és i/b modellek megfelelői voltak, míg a ii/c esetben i/b-vel azonos fehérjeszerkezetből indultunk ki, de csupán 14 mM Ca<sup>2+</sup> koncentráció volt jelen. A csoport utolsó modell rendszere a ii/d volt, mely kalciumionokat nem tartalmazott és referenciaként szolgált. A harmadik szimulációs csoportban (iii) az aktivációs peptidet nem tartalmazó homodimerek dinamikai vizsgálata volt a fő cél az 1F13 szerkezetből kiindulva. A csoport három modellt tartalmazott, melyek rendre a kalciumot nem tartalmazó (iii/a), (iii/b) 14 mM Ca<sup>2+</sup>-ot tartalmazó és végül (iii/c) 1000 mM Ca<sup>2+</sup>-ot tartalmazó rendszerek voltak (1. táblázat).

1. táblázat. Az FXIII-A<sub>2</sub> és TG2 szimulációs modellek részletei.

Szimulációs csoport	Szimuláció azonosítója	Protein	PDB ID	AP jelen volt?	GDP vagy GTP jelen volt?	Arg37-Gly38 hasítva volt?	[Ca <sup>2+</sup> ], mmol/dm <sup>3</sup>	Szimulációs idő, μs	Szimuláció „multi-site” Ca <sup>2+</sup> modellel
i.	a	FXIII-A <sub>2</sub>	1GGU	igen	-	nem	14	1	-
	b	FXIII-A <sub>2</sub>	1F13	igen	-	nem	50	1	-
ii.	a	FXIII-A <sub>2</sub> <sup>1</sup>	1GGU	igen	-	igen	14	1	-
	b	FXIII-A <sub>2</sub> <sup>1</sup>	1F13	igen	-	igen	50	1	-
	c	FXIII-A <sub>2</sub> <sup>1</sup>	1F13	igen	-	igen	14	1	50 ns
	d	FXIII-A <sub>2</sub> <sup>1</sup>	1F13	igen	-	igen	-	1	-
iii.	a	FXIII-A <sub>2</sub> <sup>1</sup>	1F13	nem	-	igen	-	2	-
	b	FXIII-A <sub>2</sub> <sup>1</sup>	1F13	nem	-	igen	14	2	-
	c	FXIII-A <sub>2</sub> <sup>1</sup>	1F13	nem	-	igen	1000	2	-
iv.	a	TG2	1KV3	-	-	-	-	2	-
	b	TG2	1KV3	-	-	-	8	2	-
	c	TG2	1KV3	-	GDP	-	-	1	-
	d	TG2	1KV3	-	GDP	-	8	1	50 ns
	e	TG2	4PYG	-	GTP	-	6	1	-
v.	a	FXIII-A <sup>o</sup>	4KTY	igen	-	nem	14	1	-
	b	TG2	2Q3Z	-	-	-	8	1	-

A követhetőség érdekében külön szimulációs csoportban vizsgáltuk a szöveti transzglutaminázt (iv). Az iv/a és iv/b modellek esetében a krisztallográfiai szerkezetben található, kötött GDP eltávolításra került, míg iv/b esetében 8 mM kalcium került hozzáadásra. Az iv/c és iv/d rendszerek megfeleltek az iv/a és iv/b modelleknek, azonban itt a GDP-t megtartottuk, annak érdekében, hogy vizsgálható legyen annak feltételezett hatása. Ezen szimulációkban az 1KV3 szerkezetet [63] használtuk fel, viszont ezen szerkezetben a Val224 aminosav feltehetően klónozási hiba eredménye, így a 4PYG [64] szerkezetet is vizsgáltuk (iv/e). Ez már a natív Gly224 aminosavat tartalmazza, de ezen kívül még három pontmutációt is tartalmaz, melyeket nem diszkutáltak az eredeti közleményben.

Az utolsó (v) szimulációs csoport a nyílt (aktív) konformációt tartalmazta az FXIII-A<sup>o</sup> (v/a, PDB ID: 4KTY [55]) és a TG2 (v/b, PDB ID: 2Q3Z) enzimeknek. Ezen szimulációk érdekes eseteket reprezentálnak, de nem tartoznak szorosan a disszertációban témájához.

A munka során az irodalomban elterjedt számozást alkalmaztuk a fehérjéket alkotó aminosavakra, miszerint az első (iniciátor) metionin nem került számozásra az FXIII-A<sub>2</sub> esetében, ezzel szemben a szöveti transzglutamináz vizsgálatánál a szekvencia Met1-el kezdődött.

### 5.2.2 Molekuladinamikai szimulációk

A modell rendszerek oktaédes szimulációs dobozban lettek szolvatálva TIP3P vízmodellel, a cellák méretének megválasztásánál 12 Å minimális távolságot biztosítottam a protein és a doboz fala között. A rendszerek ösztöztése minden esetben zérus volt és a teljes nátrium-klorid koncentrációt az izotóniás sókoncentrációnak megfelelően (150 mM) választottam meg. Rövid energiaminimálás után, 2 ns hosszúságú szimulált hőkezelés (simulated annealing, SA) került alkalmazásra, 310 K-en és 1 bar nyomáson a rendszer ekvilibrálása végett. Az ekvilibrálási folyamat során a nehézatomokon 1000 kJ mol<sup>-1</sup>nm<sup>-2</sup> kényszer erők kerültek alkalmazásra. A szimulációs modellektől függően 1 vagy 2 mikroszekundum hosszúságú NpT molekuladinamikai szimulációkat futtattam (1. táblázat). Annak érdekében, hogy a newtoni mozgásegyenletek megoldására 4 fs-os lépésközt lehessen alkalmazni, virtuális kölcsönható pontok kerültek elhelyezésre, melyek magas rezgési hullámhossznál adják vegyértékrezgésük első harmonikusát. A szimulációk során minden kötés meg volt kötve a LINCS algoritmus segítségével. Azon modellekben, ahol GDP vagy GTP volt jelen 2 fs-os lépésközt alkalmaztam és a szimulációk minden esetben 1 µs-ig tartottak. A Lennard-Jones és a rövid távú elektrosztatikus kölcsönhatások számítására 10,0 Å-ös cutoff érték került alkalmazásra. Simítási módszert alkalmaztam annak érdekében hogy 7,0 és 10,0 Å között az erők levágása folytonosan történjen. A hosszútávú kölcsönhatások a PME összegzés segítségével kerültek számításra. A hőmérséklet közel állandó értéken tartására a velocity-rescaling módszert [85] alkalmaztam, míg a nyomást izotróp Parinello-Rahman barosztáttal [91] szabályoztam. Amber-99SB-ILDN-NMR erőteret [83,92,93] alkalmaztam a szimulációk során, míg a GDP és GTP szubsztrátokat Amber GAFF paraméterek [94] segítségével vettük figyelembe. A használt Amber-99SB-ILDN-NMR erőteret korábban sikeresen alkalmaztuk NMR adatokból kiindulva a kalciumkötődés vizsgálatára [95]. A munka során összesen több mint 21 µs-os atomi felbontású (all-atom) molekuladinamikai szimuláció került kivitelezésre.

Két esetben (ii/c és iv/d), további 50 ns NpT MD szimulációt végeztem Amber ff14SB erőterrel [96] és “multi-site” Ca<sup>2+</sup> modell [97] alkalmazásával annak érdekében, hogy igazoljuk az egyszerű ionmodellekkel kapott eredmények relevanciáját. Ezen két esetben a mikroszekundumos

szimulációk utolsó pillanatképe került felhasználásra, valamint a kötött kalciumionokat megtartottuk, és a “*multi-site*” modell XZ atomját illesztettük mindegyik ionra. A rendszereket TIP3P vízmodellel töltöttük fel, neutralizáltuk, és 150 mM NaCl koncentráció került beállításra. A szimulációk során 10,0 Å cut-off került alkalmazásra, és energia minimálás után három fázisú, összesen 1,7 ns hosszúságú ekvilibrálás következett, melynek során a rendszer 310 K-re lett felfűtve, majd a sűrűsége került beállításra, ezt követően pedig az ekvilibrálás következett.

A molekuladinamikai szimulációkat GROMACS 5.1.4 [98,99] és az Amber csomag [100–102] pmemd szoftverével kerültek futtatásra. A dinamikai trajektóriákat a Gromacs csomag szoftvereivel valamint a cpptraj programmal [103] analizáltuk. A dinamikus kereszt-korrelációk számításához (DCCM) 12500 pillanatképet használtam és a Bio3D v2.3 [104] csomaggal kerültek számításra, ami az R statisztikai programban [105] van implementálva. A mátrixműveletek szintén az R program segítségével lettek elvégezve. A szerkezetek az UCSF Chimera 1.11.2 [106] és a Visual Molecular Dynamics (VMD) 1.9.1 [88] programokkal kerültek megjelenítésre. A fehérjék azon szegmenseit, melyek nem megfelelően vagy részben nem megfelelően voltak meghatározva a kristallográfiai szórásképből, a MODELLER 9.10 szoftver segítségével pótoltam és azon szerkezet került felhasználásra mely a legjobb DOPE score-ral [107] rendelkezett. Mindezt a UCSF Chimera grafikus felületén [108] keresztül végeztem el. Figyelembe véve, hogy az FXIII-A<sub>2</sub> 1GGU kristallográfiás szerkezetében a Thr508-Ser516 közötti hurok régió hiányzik, valamint ezen régió az 1F13 szerkezetben is két különböző állapotban található meg. Ahelyett, hogy *ab initio* modelleztem volna a kérdéses régiót, egyszerűen az 1F13-ban található hurkokat alkalmaztam az 1GGU-ban oly módon, hogy az 1F13 A főláncához tartozó hurkot az 1GGU B főláncába illeszttem. A munkában látható adatok a matplotlib [109] és a numpy [110] python könyvtárak segítségével kerültek megjelenítésre, melyekhez általam írt szkripteket alkalmaztam. A titrálható aminosavak protonáltságát a H<sup>++</sup> webszerver [111–113] segítségével határoztuk meg 7,4-es pH-n. A szöveti transzglutamináz esetében minden negatív töltésű oldallánc deprotonálnak adódott.

### 5.2.3 A molekuladinamikai szimulációk kiértékelése

A trajektóriák 80 ps-onként kerültek mentésre, és az atomi koordináták négyzetes közepe (RMSD), a forgási sugár ( $r_{\text{gyr}}$ ) és a fluktuációk (RMSF) számolására a protein nehézatommjait vettük figyelembe. Annak érdekében hogy az FXIII-A<sub>2</sub> esetében az aktivációs peptid (Ser2-Arg37) ne befolyásolja az eredmények összevethetőségét, ezen régiókat mellőztük az RMSD és  $r_{\text{gyr}}$  számításánál. Az aminosavakra átlagolt RMSF értékek mindkét alegységre külön-külön kerültek számításra, majd átlagolásra kerültek. A DCCM mátrixokhoz a C-alfa atomok trajektóriáját használtuk fel, így szimmetrikus  $N \times N$  -es mátrixokat nyertünk. Mivel az atomi mozgások

korrelációjának kiértékelése nehézkes ilyen méretű fehérjék esetében, a különbségek kiemelésére törekedtünk oly módon, hogy a különbséget vettük az adott mátrix  $C_{ij}(2)$  és egy referencia  $C_{ij}(1)$  mátrix értékeinek. A  $\text{Ca}^{2+}$  kötési tulajdonságok vizsgálatánál a savas oldalláncok karboxilát szenének és bármely kalciumionnak a távolságát vettem figyelembe. A nagy amplitudójú, alacsony frekvenciájú mozgások számításánál az első 10 sajátvektor került számításra, melyeket a ProDy szoftver [114,115] segítségével végeztem el. A domén mozgások változásának leírására egy pszeudo torziós szöveget definiáltam és a dinamikai mozgásokból számolt szórásértékek is fel lettek tüntetve minden esetben, az FXIII-A<sub>2</sub> szimulációjánál.

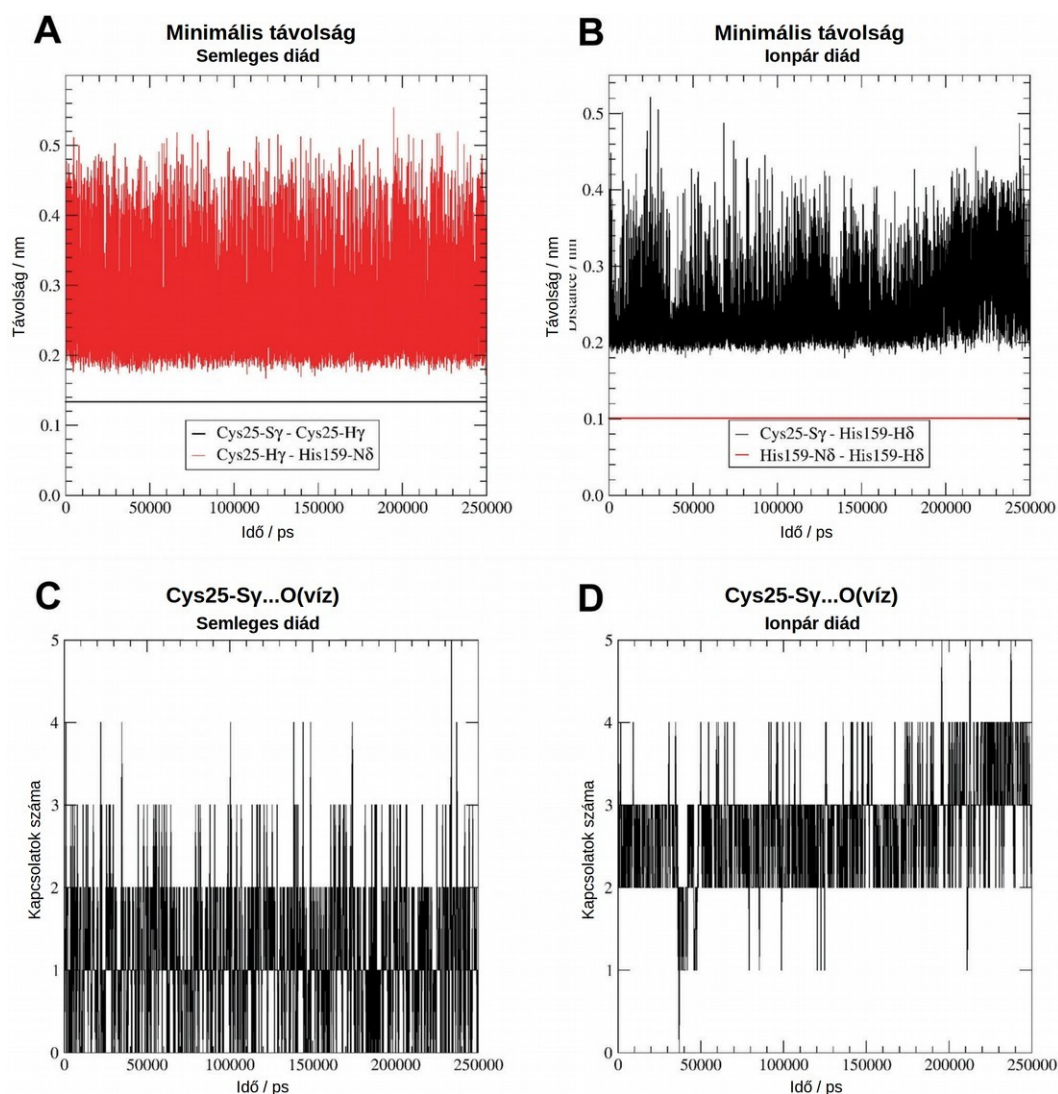
## 6. Eredmények és diszkusszió

### 6.1 A papain enzimmechanizmusa

#### 6.1.1 A papain nyugvó állapotának vizsgálata

A 250 ns hosszúságú molekuladinamikai szimulációk, a szabad (N-metil-acetamidot nem tartalmazó) papain esetében mind a neutrális, mind az ionpár formában lévő cisztein-hisztidin diád jelenlétében azt mutatták, hogy a rendszerek jól ekvilibráltak az első 100-150 ns szimulációs idő alatt (K1. ábra). Az is látható (3. ábra, A panel), hogy a neutrális forma szimulációjában a (Cys25)Hy és a (His159)Nδ közötti jelentős távolságbeli fluktuáció ellenére a legrövidebb távolságok (1,8-2,0 Å) megfelelnek az átlagos H-híd távolságnak, és mint ilyen a szimuláció jelentős részében jelen van. Ez azt sugalja, hogy a protonátmenet a (Cys25)Sy és a (His159)Nδ között előfordulhat víz jelenléte nélkül is, ezért mediátor vízmolekulát nem vettünk figyelembe a rendszer QM/MM (ONIOM-EE) vizsgálatánál. Másrészt, a protonátmenet után a (His159Nδ)-Hδ... Sy-(Cys25) távolság továbbra is megfelel a hidrogénkötésnek megfelelő távolságnak (3. ábra, B panel).

A szimulációk elemzése során jelentős eltérések tapasztalhatók a katalitikus diád környezetében előforduló vízmolekulák számában. Míg a neutrális diád szimulációjánál a (Cys25)Sy atomnál nulla és három közötti számosságú vízmolekula fordul elő (3. ábra, C panel), addig az ionpár formában jelen lévő diád esetében ez a szám kettő és négy közöttire tehető (3. ábra, D panel). Ez alapján feltételezhető, hogy az oldószer hozzájárul az egyes formák stabilizálásához. Eltekintve a hőmérsékleti fluktuációtól mindkét esetben a Cys-His diád és annak közeli környezete jelentős hasonlóságot mutatott a kiindulási röntgenkristallográfiás szerkezettel.



3. ábra. A papain neutrális és ionpár formában jelen lévő formáinak dinamikai trajektóriájából nyert távolságai a (Cys25)H $\gamma$  és (His159)H $\delta$  valamint a (Cys25)S $\gamma$  és (His159)N $\delta$  atomok között (A és B panel), valamint a kölcsönhatási távolságok belül lévő vízmolekulák számának változása (C és D panel).

A (His159)N $\delta$ -H $\gamma$ -S $\gamma$ (Cys25) H-híd kapcsolat létezése a teljes szimulációk során ahhoz az elfogadható feltételezéshez vezetett, hogy a statikus ONIOM számítások megfelelően kivitelezhetőek az optimált H-kötéses szerkezetből. A szubsztrát nélküli papain esetében az 2. táblázat foglalja össze a QM/MM számítások eredményeit. Ezen számításoknál minden esetben azt találtuk, hogy a zérus ponti vibrációs energia (ZPVE) nélkül nagyjából azonos energia értékek adódtak a semleges (N) és az ionos (Z) diád formákra is. Ezen rendszerek energiái között nem több mint 0,8 kcal mol<sup>-1</sup> volt a különbség. Az  $\omega$ B97X-D módszerrel nyert eredmények az ion pár formát jósolják a legstabilabbnak, míg a többi módszerrel meglehetősen változó eredményeket kaptunk. Az M06-2X módszerrel kapott eredmények közelebb állnak az  $\omega$ B97X-D értékeihez, ugyanakkor a B97-D és a B3LYP módszerek esetében nagyon kis különbség fedezhető fel, továbbá jellemzően

azonos függőség mutatkozik az alkalmazott báziskészlet minőségétől. Korábbi munkákban azt találták, hogy az ionpár forma (Z) valóban stabilabb, mint a neutrális diád (N) [12,17,18].

A zérusponthi energiakorrekció figyelembe vételével a neutrális forma adódik stabilabbnak minden általunk elvégzett számítás esetében. Mindazonáltal a két jellemző forma közötti korrigált energiakülönbség kicsiny marad, kevesebb mint  $2 \text{ kcal mol}^{-1}$  minden esetben. Az ionpár  $\rightarrow$  neutrális diád átmenetre az alkalmazott módszerek alacsonyabb ZPVE korrigált aktiválási gátat jósolnak, mint a neutrális  $\rightarrow$  ionpár diád irányában történő átmenetre. Ez azt jelenti, hogy spontán átmenet lehetséges a neutrális diád irányába ezen formák között, azaz még  $0 \text{ K}$  hőmérsékleten is aktiválási gát nélküli reakcióról beszélhetünk.

Ezen zérusponthi korrekciót tartalmazó elméleti energiaértékek a korábbi elméletekkel ellentétesek [6–9,116–122], melyek az ionpár forma stabilitását hangsúlyozták, ugyanakkor van valószínűsége a neutrális forma létezésének [7,119,121], azaz energiáik feltehetően közel esnek egymáshoz.

Az oldószer molekuláknak az ionpár forma relatív energiájára gyakorolt hatását szofisztikált QM/MM számításokkal demonstrálták a katepszin példáján, mely papain homológjának tekinthető. Így bizonyított, hogy az ionpár diád csupán akkor adódik stabilabbnak, ha figyelembe vesszük az azt körülvevő vízmolekulákat [25].

Ezért úgy döntöttünk, hogy elektrosztatikus oldódási szabadenergia számításokat végzünk, mind az ionpár, mind a neutrális cisztein-hisztidin diádra, annak érdekében, hogy becsülni tudjuk a solvatáció hatását. A molekuladinamikai szimulációk utolsó  $100 \text{ ns}$ -a alapján számolt Poisson-Boltzman (PB) értékeket átlagolva az ionpár formára  $20 \text{ kT}$ -vel ( $\sim 11.9 \text{ kcal mol}^{-1}$ ) nagyobb elektrosztatikus solvatációs energiatag adódik, melyeket a későbbiekben korrekciós tagként alkalmaztunk az ONIOM számításokból nyert ZPVE korrigált energiákra ( $E^{\text{ZPVEc}}$ ). Ezen korrekció elég nagy volt ahhoz, hogy megfordítsa a relatív stabilitását a két diád formának (2. táblázat), azaz az ionpár formát adja stabilabbnak minden esetben. Jóllehet az általunk alkalmazott módszer alapvetően különbözik a Mladenovic és munkatársai által alkalmazottól, ahol egyértelműen hangsúlyozzák az oldószer fontosságát az ionpár forma stabilizálásában [25].

2. táblázat. A papain nyugvó állapotának tanulmányozása során nyert válogatott szerkezeti és termodinamikai paraméterek összefoglalása. (Az atomok közötti távolságok angströmben, a termodinamikai paraméterek kcal mol<sup>-1</sup> egységekben kerültek megadásra.)

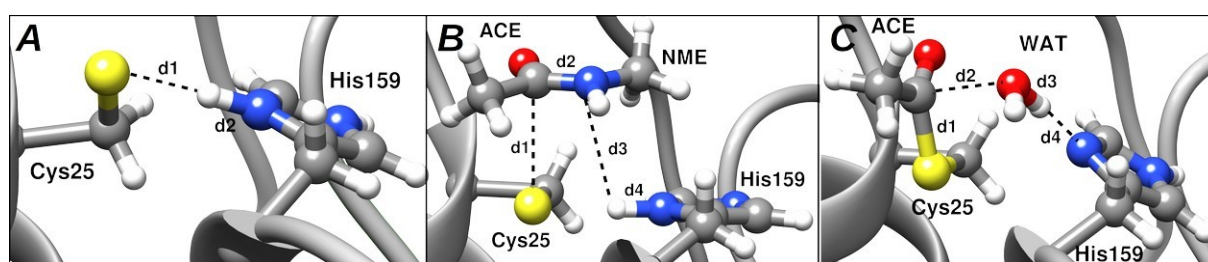
Módszer	Bázisfüggvény	Ikerionos forma (Z)		TS		Semleges forma (N)		TS – Z		TS – N		Z – N	
		d1	d2	d1	d2	d1	d2	E <sup>ONIOM</sup> , kcal mol <sup>-1</sup>	E <sup>ZPVEc</sup> , kcal mol <sup>-1</sup>	E <sup>ONIOM</sup> , kcal mol <sup>-1</sup>	E <sup>ZPVEc</sup> , kcal mol <sup>-1</sup>	E <sup>ONIOM</sup> , kcal mol <sup>-1</sup>	E <sup>ZPVEc</sup> , kcal mol <sup>-1</sup>
B3LYP	6-31G(d,p)	2,008	1,096	1,630	1,339	1,390	1,882	1,967	-0,946	2,262	0,842	0,295	1,788
	6-31+G(d,p)	2,045	1,085	1,618	1,351	1,384	1,925	2,587	-0,521	2,405	0,944	-0,182	1,465
	6-311+G(d,p)	2,052	1,083	1,626	1,346	1,382	1,940	2,581	-0,560	2,828	1,373	0,247	1,933
	6-311++G(d,p)	2,055	1,082	1,625	1,347	1,383	1,938	2,650	-0,504	2,788	1,328	0,137	1,832
B97-D	6-31G(d,p)	1,936	1,130	1,662	1,314	1,414	1,774	0,814	-1,379	1,476	0,282	0,662	1,662
	6-31+G(d,p)	1,990	1,108	1,636	1,340	1,409	1,799	1,421	-1,123	1,374	0,092	-0,047	1,215
	6-311+G(d,p)	1,999	1,105	1,641	1,337	1,405	1,822	1,461	-1,111	1,641	0,341	0,181	1,451
	6-311++G(d,p)	2,006	1,102	1,637	1,341	1,406	1,820	1,562	-1,051	1,592	0,287	0,030	1,338
ωB97X-D	6-31G(d,p)	2,013	1,085	1,596	1,358	1,381	1,888	2,625	-0,517	2,213	0,983	-0,413	1,500
	6-31+G(d,p)	2,035	1,078	1,590	1,364	1,376	1,925	3,081	-0,244	2,446	1,197	-0,635	1,441
	6-311+G(d,p)	2,041	1,077	1,599	1,357	1,374	1,947	3,033	-0,343	2,842	1,530	-0,191	1,873
	6-311++G(d,p)	2,042	1,076	1,598	1,358	1,375	1,944	3,062	-0,311	2,813	1,513	-0,249	1,824
M06-2X	6-31G(d,p)	1,985	1,096	1,578	1,388	1,377	1,899	2,151	-0,789	1,618	0,259	-0,533	1,049
	6-31+G(d,p)	2,008	1,088	1,572	1,395	1,367	1,987	2,615	-0,428	1,874	0,630	-0,740	1,058
	6-311+G(d,p)	2,014	1,086	1,581	1,384	1,363	2,033	2,863	-0,480	2,388	1,231	-0,475	1,711
	6-311++G(d,p)	2,016	1,085	1,581	1,385	1,363	2,030	2,155	-0,449	2,353	1,189	0,198	1,638

E<sup>ONIOM</sup>, ONIOM extrapolált energia

E<sup>ZPV</sup>, Zérusponti vibrációs energia

E<sup>ZPVEc</sup> = E<sup>ONIOM</sup> + E<sup>ZPV</sup>

A protonátmenetet leíró útvonal karakterisztikus geometriaparaméterei a tiol csoportról a hisztidin gyűrűre vonatkozóan az 4. ábra, A panelén, valamint a 2. táblázatban láthatóak. Az adatok azt mutatják, hogy az általunk kivitelezett ONIOM számításokban alkalmazott sűrűségfunkcionál módszerek nagyon hasonló geometriai paramétereket szolgáltatnak az S–H és a semleges imidazol oldallánc gamma protonja esetében, ugyanakkor a B97-D módszer jelentősen rövidebb N $\delta$ –H $\delta$  távolságot és kicsit hosszabb S $\gamma$ –H $\gamma$  távolságot becsül, azaz erősebb hidrogénkötést jósol az N $\delta$ –H $\gamma$ –S $\gamma$  atomok között. A báziskészletek méretének és minőségének hatása itt elhanyagolható. Szintén érdekes, hogy ugyanezen módszerrel adódtak a leghosszabbnak az S–H távolságok, míg az átmeneti állapot esetében némiképp jobban szórnak az eredmények.



4. ábra. A papain enzimmechanizmusának tanulmányozása során definiált távolságok grafikus reprezentációja a részlépések kiindulási szerkezetei alapján. (A) a papain nyugvó állapota (ikerionos diád (Z)), (B) az acilezési részlépés (Michaelis komplex (MC)), (C) az acil-enzim hidrolízise (acil-enzim intermedier-víz komplex (C)).

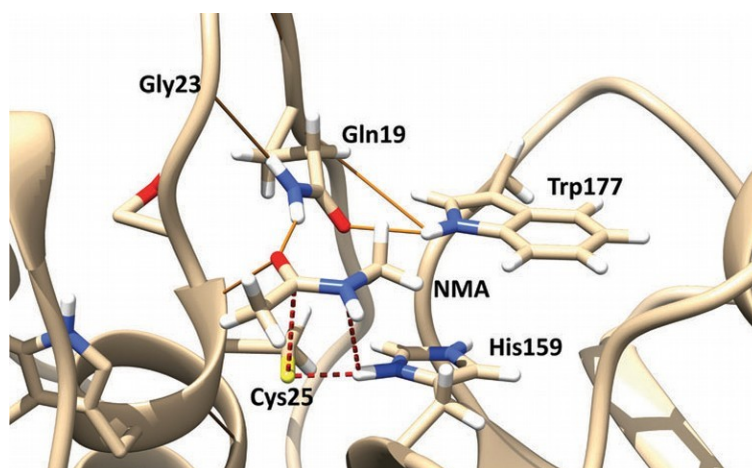
A  $d1$  és  $d2$  távolságok a semleges és az ionpár diád, valamint az átmeneti állapot esetében jó egyezést mutat a korábbi elméleti eredményekkel a papain–N-acetil-Phe-Gly-4-nitroanilid enzim–szubsztrát komplex esetében publikáltakkal [23]. Kiemelendő, hogy a  $d1$  és  $d2$  távolságok az ionpár formára szintén jól egyeznek szofisztikált QM/MM-(MD) számítások átlagolt értékeivel [25], míg a neutrális formát illetően jelentősen eltérnek a Mladenovic és munkatársai által publikáltaktól. Ez utóbbi eltérésre magyarázatként szolgálhat, hogy a neutrális diád sokkal nagyobb konformációs flexibilitással rendelkezik, melyet a mi MD szimulációink is alátámasztanak (3. ábra, A és B panel).

### 6.1.2 Az acil-enzim intermedier képződése

A cisztein proteáz reakció első lépését illetően a főbb geometriai paraméterek a 3. és 4. táblázatban találhatóak, valamint a vizsgált reakciók energiaprofiljai a 8. ábrán láthatóak. Számításaink alátámasztják az ikerionos acil-enzim intermedier létezését, így az átmeneti

állapotok energiái a Michaelis komplexek (MC) energiáival hasonlítottuk össze, a tetraédes intermedierek és az enzimreakció termékeinek energiái a 3. táblázatban láthatóak. Négy karakterisztikus távolság (4. ábra, B panel) került kigyűjtésre ezen acilezési folyamatra, melyek jól leírják a Michaelis komplexek, az átmeneti állapotok és a termékek geometriájában bekövetkező változásokat. Ellentétben eredményeinkkel Wei és munkatársai egy lépéses, elemi reakciót javasoltak a mechanizmusra, stabil (tetraédes) intermediér állapot nélkül. Számításaikat B3LYP/6-31G(d) szinten számolt reakcióút modellezésre és B3LYP/6-13++G(d,p) augmentált báziskészlettel történő energiaszámításra és pszeudo-kötés szabadenergia becslésre alapozták a QM/MM számítások keretein belül [23].

A Michaelis komplex geometriáját illetően (3. táblázat) a kötéshossz paraméterek csupán csekély függőséget mutatnak az alkalmazott elméletektől. Másrészt a nemkötő távolságok sokkal inkább mutatnak módszerfüggő eltéréseket. A leghosszabb kötéstávolságok az  $S_{\gamma}$ -C(peptid-karbonil) és a (His159N $\delta$ )H-N(peptid-amid) esetében a B3LYP módszerrel adódtak, amire magyarázatként szolgálhat az, hogy a B3LYP funkcionál nem tartalmaz megfelelő tagot a hosszútávú kölcsönhatások leírására. A legszembetűnőbb különbségek a nemkötő távolságok esetében a B3LYP és az M06-2X módszer összehasonlításánál jelennek meg. Összehasonlítva eredményeinket a Wei és munkatársai által közöltekkel az N-acetil-Phe-Gly-4-nitroanilid szubsztrát esetében kapott amid kötés hossza 1,37 Å mely kissé rövidebb mint az általunk kapott eredmények. További hasonlóságok fedezhetőek fel a korábban publikált eredmények [23] és az általunk közöltek között a nemkötő távolságok tekintetében, azzal a kivétellel, hogy az  $S_{\gamma}$ -C(peptid-karbonil) között távolság 3,7 Å-nek adódott a már említett munkában [25].



5. ábra. A papain aktív helyének grafikus reprezentációja. Az ábra az ONIOM(M06-2X/6-311++G(d,p):AMBER) módszerrel nyert optimált Michaelis komplexet (MC) mutatja. Folyamatos vonalak

mutatják a szubsztrát stabilizálásában részt vevő atomokat, szaggatott vonalak mutatják a reakció során változó kémiai kötések.

A papain-NMA Michaelis komplex hidrogénkötésekkel stabilizált szerkezet, melyben az NMA karbonil oxigénje H-híd kölcsönhatásban van a Gln19 oldalláncával, valamint a Cys25 amid hidrogénjével (5. ábra). A megfelelő orientáció kialakításában a Gln19, a Trp177 és a Gly23 aminosavak fontos szerepet töltenek be, valamint meg kell jegyeznünk, hogy a His159 imidazol gyűrű és a Cys25S<sub>y</sub> megközelítőleg egy síkban van.

A 3. táblázatban összegyűjtött távolságokból látható, hogy a tetraédes intermedier megfelel a formális ikerionos intermedier szerkezetnek. Ugyanakkor ezen szerkezet különbözik a klasszikus oxianionos intermediertől, ahol a szubsztrát amid nitrogénje nem szenved koncertikus protonálódást a nukleofil támadással egy időben, és létezése egyértelműen bizonyított a szerin proteázoknál [123]. Kijelenthető tehát, hogy esetünkben a tiolát S<sub>y</sub> nukleofil támadása a szubsztrát karbonil szén atomján egy elemi lépésben történik a hisztidin imidazólium oldalláncának gamma nitrogénjéről történő protonátmenettel az NMA amid nitrogénjére. Mindazonáltal megfigyelve a *d1*, *d3* és *d4* karakterisztikus távolságokat, azok elemi asszinkronitást adtak. Az utóbbi esetben a *d1* (S–C) távolság közel van a szerin proteázok esetében említett anionos tetraédes intermedierben (TI) feltételezhetőhöz, míg a *d3* és *d4* távolságok (protontranszfer) értékeit illetően azok inkább átmeneti állapot jellegű értékek. Ez azt jelenti, hogy az átmeneti állapot hasonlít egy olyan szerkezethez, mely összeköti a már kialakult szerin proteáz-szerű intermediert az ikerionos intermedierrel. Minden más módszernél az átmeneti állapotot főképp az S–C kötés kialakulása határozza meg és a H–N(imidazólium) távolság sokkal inkább hasonlít a Michaelis komplexben mérhetőhöz, továbbá ezen esetek mindegyikében az S–C kötés kialakulása látszólag megelőzi a protonátmenetet. A B97-D módszer esetében jelentősen hosszabb S–C kötéshosszak mérhetőek az átmeneti állapot esetében, szemben más módszerekkel, függetlenül a báziskészlettől. Hasonló, de jóval mértékletesebb tendencia figyelhető meg a tetraédes intermedier vonatkozásában.

Az átmeneti állapotokat (6. ábra) és tetraédes intermediereket jellemzően ugyanazon hidrogénkötések stabilizálják, mint amelyeket már említettünk a Michaelis komplex esetében. Jelentős különbség figyelhető meg azonban a His159 imidazol gyűrűjének orientációjában, ami elfordul kezdeti állapotához képest és az N $\delta$ –H kötés az NMA amid nitrogénjének irányába mutat az átmeneti állapotban, így az imidazol gyűrű, már nincs egy síkban a Cys25-S<sub>y</sub> atomjával.

3. táblázat. A papain acilezési mechanizmusának tanulmányozása során nyert, válogatott szerkezeti és termodinamikai paraméterek összefoglalása. (Az atomok közötti távolságok angströmben, a termodinamikai paraméterek kcal mol<sup>-1</sup> egységekben kerültek megadásra.)

Módszer	Bázisfüggvény	Michaelis komplex (MC)				TS1				Tetraédes Intermediér (TI)				TS2				Termék (AE)				TS1-MC	TS1-TI	TS2-TI	TS2-AE
		d1	d2	d3	d4	d1	d2	d3	d4	d1	d2	d3	d4	d1	d2	d3	d4	d1	d2	d3	d4				
B3LYP	6-31+G(d,p)	3,469	1,357	3,726	1,091	2,173	1,520	1,437	1,194	2,009	1,638	1,057	1,814	1,883	2,078	1,031	2,008	1,799	3,755	1,019	2,803	12,290	3,933	0,800	4,235
	6-311+G(d,p)	3,451	1,356	3,708	1,090	2,195	1,520	1,423	1,203	2,016	1,646	1,054	1,824	1,900	2,019	1,032	1,981	1,801	3,755	1,017	2,803	12,836	4,226	0,341	4,821
	6-311++G(d,p)	3,450	1,356	3,711	1,088	2,197	1,520	1,420	1,205	2,017	1,645	1,054	1,823	1,900	2,021	1,032	1,981	1,801	3,756	1,017	2,803	12,941	4,192	0,351	4,840
B97-D	6-31G(d,p)	3,375	1,367	3,337	1,137	2,852	1,435	1,812	1,059	2,072	1,704	1,067	1,741	1,969	1,926	1,047	1,819	1,821	3,503	1,024	2,360	10,212	6,793	1,538	1,851
	6-31+G(d,p)	3,364	1,364	3,387	1,111	2,786	1,441	1,763	1,065	2,070	1,662	1,071	1,731	2,053	1,683	1,068	1,731	1,813	3,555	1,023	2,440	10,261	5,440	1,726	1,724
	6-311+G(d,p)	3,355	1,363	3,380	1,109	2,746	1,447	1,739	1,069	2,077	1,670	1,067	1,744	2,057	1,696	1,064	1,745	1,815	3,552	1,021	2,428	10,385	5,243	1,626	2,503
	6-311++G(d,p)	3,354	1,363	3,385	1,105	2,711	1,453	1,714	1,075	2,079	1,669	1,067	1,744	2,059	1,695	1,064	1,744	1,815	3,553	1,021	2,429	10,201	4,900	1,621	2,484
ωB97X-D	6-31G(d,p)	3,117	1,354	3,366	1,095	2,523	1,456	1,670	1,077	1,989	1,614	1,058	1,755	1,853	2,138	1,028	1,955	1,790	3,374	1,020	2,267	11,933	7,057	2,061	1,984
	6-31+G(d,p)	3,124	1,351	3,383	1,084	2,470	1,454	1,688	1,073	1,971	1,598	1,058	1,758	1,843	2,157	1,027	1,982	1,784	3,435	1,019	2,337	12,840	7,171	3,149	2,159
	6-311+G(d,p)	3,128	1,350	3,375	1,082	2,434	1,459	1,689	1,073	1,978	1,600	1,055	1,763	1,853	2,119	1,026	1,965	1,785	3,424	1,018	2,324	13,340	7,247	2,782	2,654
	6-311++G(d,p)	3,128	1,350	3,375	1,081	2,426	1,459	1,687	1,073	1,978	1,600	1,055	1,764	1,853	2,120	1,026	1,966	1,785	3,424	1,018	2,325	13,489	7,206	2,805	2,681
M06-2X	6-31G(d,p)	3,044	1,354	3,314	1,110	2,599	1,442	1,718	1,071	1,985	1,612	1,063	1,740	1,835	2,216	1,025	2,013	1,791	3,047	1,021	2,176	11,541	6,972	2,175	0,635
	6-31+G(d,p)	3,058	1,351	3,299	1,096	2,555	1,443	1,714	1,071	1,969	1,597	1,063	1,743	1,831	2,218	1,024	2,034	1,794	2,657	1,019	2,154	12,113	6,856	3,490	1,164
	6-311+G(d,p)	3,055	1,351	3,274	1,093	2,516	1,449	1,706	1,072	1,974	1,602	1,057	1,764	1,841	2,169	1,024	2,015	1,795	2,702	1,018	2,160	12,406	6,759	2,669	1,593
	6-311++G(d,p)	3,056	1,351	3,279	1,092	2,503	1,451	1,698	1,074	1,974	1,602	1,057	1,765	1,841	2,172	1,024	2,017	1,795	2,704	1,018	2,161	12,308	6,633	2,676	1,579

$$E^{ZPVEc} = E^{ONIOM} + E^{ZPV}$$

E<sup>ONIOM</sup>, ONIOM extrapolált energia

E<sup>ZPV</sup>, Zérusponyi vibrációs energia

A B3LYP/6-31G(d,p) módszer esetében kivételes viselkedést tapasztalhatunk, mely eltér minden egyéb esettől. Ebben az egy esetben egy további átmeneti állapotot azonosítottunk, ami a Michaelis komplex és az ikerionos intermedier közötti útvonalon található (4. táblázat). A TS1a jelű átmeneti állapot igen hasonló azon átmeneti állapotokhoz, melyeket olyan módszerek esetében véltünk felfedezni, ahol valamilyen módon figyelembe vették a diszperziós korrekciót. Azonban ebben az esetben, a TS1a egy olyan intermedier képződéséhez vezet, amely megfelel az (oxi)anionos, szerin proteáz-szerű intermediernek, ezt követően pedig külön lépésben történik az ikerionos intermedier kialakulása. A lehetséges acil-enzim kialakulási útvonalak a 9. Ábrán kerültek összefoglalásra, ahol a középső sor felel meg a B3LYP/6-31G(d,p) esetében kapott eredményeknek.

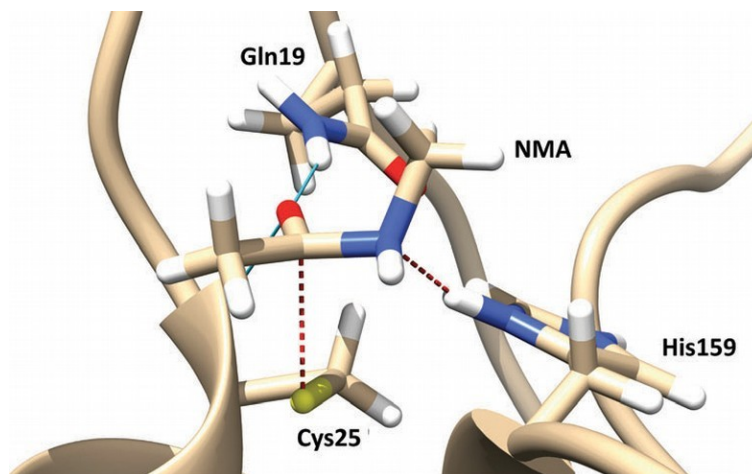
4. táblázat. A papain acilezési mechanizmusának B3LYP/6-31G(d,p):AMBER módszerrel történt tanulmányozása során nyert, válogatott szerkezeti és termodinamikai paraméterek összefoglalása. (Az atomok közötti távolságok angströmben, a termodinamikai paraméterek kcal mol<sup>-1</sup> egységekben kerültek megadásra.)

	d1	d2	d3	d4
Michaelis komplex (MC)	3,470	1,360	3,690	1,105
TS1a	2,704	1,447	1,779	1,063
Oxianionos intermedier (OI)	2,353	1,493	1,602	1,110
TS1b	2,262	1,513	1,467	1,174
Ikerionos intermedier (ZI)	2,021	1,667	1,056	1,812
TS2	1,900	2,036	1,033	1,962
Termék (AE)	1,806	3,634	1,020	2,534

	TS1a–MC	TS1a–OI	TS1b–OI	TS1b–ZI	TS2–ZI	TS2–AE
E <sup>ONIOM</sup> , kcal mol <sup>-1</sup>	12,519	0,275	0,069	6,826	0,917	2,965
E <sup>ZPVEc</sup> , kcal mol <sup>-1</sup>	13,558	0,597	-1,195	4,671	0,102	3,689

E<sup>ONIOM</sup>, ONIOM extrapolált energia  
E<sup>ZPV</sup>, Zérusponyi vibrációs energia  
E<sup>ZPVEc</sup> = E<sup>ONIOM</sup> + E<sup>ZPV</sup>



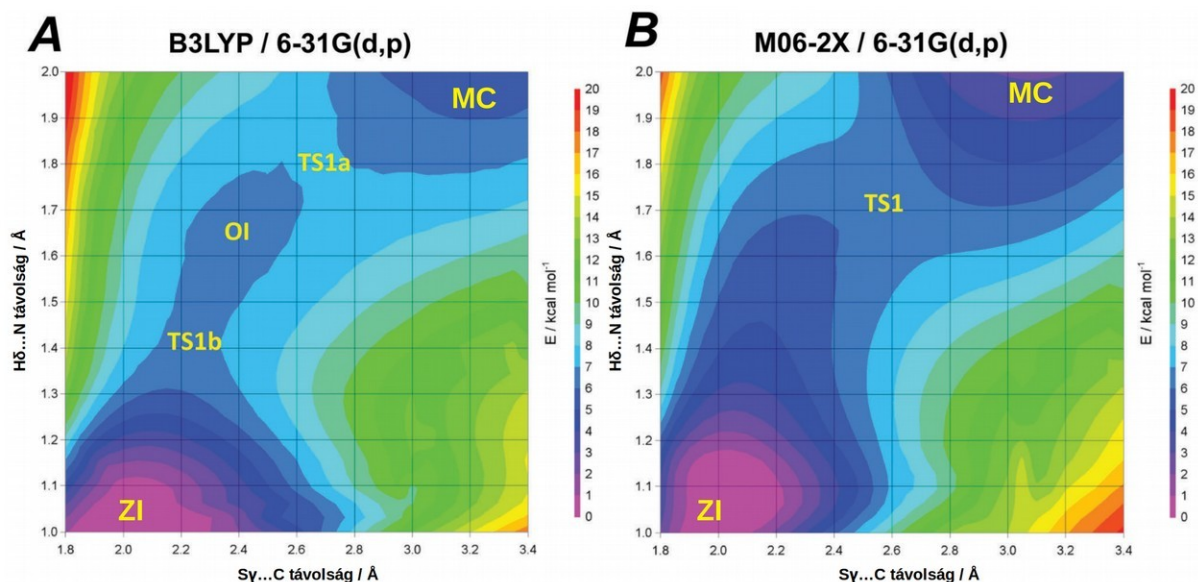
6. ábra. A papain-NMA hipotetikus reakció acilezési lépésére (ZI), ONIOM(M06-2X/6-311+ +G(d,p):AMBER) számításból kapott átmeneti állapot grafikus reprezentációja.

Az acilezési lépésre jellemző termodinamikai jellemzőket a 4. táblázatban foglaltuk össze. Látható, hogy a potenciális energiafelület a két átmeneti állapot (TS1a és TS1b) és közöttük az (oxi)anionos tetraédes intermedier (OI) közel azonos energiával foglalja el, az ikerionos intermedier (ZI) pedig jelentősen stabilabbnak adódik. Figyelembe véve a ZPVE korrekciót az energiagát, amely elkülöníti az anionos és az ikerionos tetraédes intermediereket eltűnik. Azaz ezen gátnak pusztán elméleti jelentőséget szabad tulajdonítani.

A 3. és 4. táblázatok alapján, megfigyelhető, hogy az amid kötés felszakításához szükséges energia (TS2), mely az acil-enzim köztiterméket eredményezi, jelentősen alacsonyabbnak adódik a B3LYP módszerrel. Másrészt maga a szubsztrát szerkezete, a kémiai környezet és még az alkalmazott optimálási algoritmus is könnyedén befolyásolhatja eredményeinket, feltehetően ezért találták egy elemi lépésben végbemenő reakciónak acilezési reakciót a B3LYP sűrűségfunkcionál módszer alkalmazásánál [23].

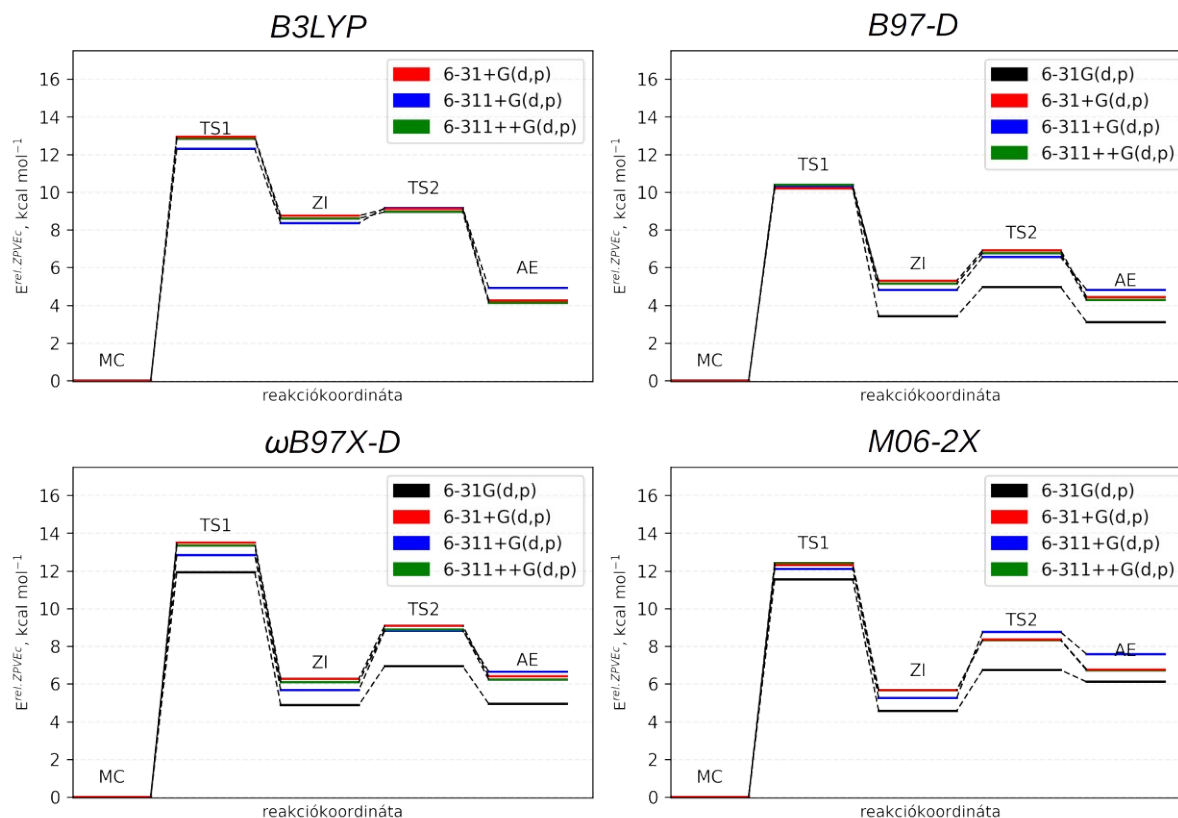
A potenciális energiafelületek, melyeket B3LYP/6-31G(d,p) (7. ábra, A panel) és M06-2X/6-31G(d,p) (7. ábra, B panel) módszerekkel végeztünk szintén megjósolják az átmeneti állapotok eltérő szerkezeteit. Ezen PES-ek alapján világosan látható, miért volt meglehetősen nehéz megtalálni az átmeneti állapotokat és intermediereket a B3LYP/6-31G(d,p) módszer esetében (és általánosságban a B3LYP számolásoknál). A B3LYP PES pásztázásánál a TS1a látható  $d1$  2,7 Å,  $d3$  1,8 Å-nél, valamint a TS1b esetén  $d1$  2,26 Å és  $d3$  1,47 Å-nél és az első köti össze a Michaelis komplexet (MC) a klasszikus oxianionos intermedierrel (OI), míg a TS1b köti össze az anionos tetraédes intermediert (TI) az ikerionos tetraédes intermedierrel (ZI). A meglehetősen lapos B3LYP PES felületen, a stacionárius pontok létezése már csekély geometriai perturbációk vagy a báziskészletek változtatása által is nagy

mértékben befolyásolható. Ezzel szemben az M06-2X PES (7. ábra, B panel) sokkal szemléletesebb képet fest a nyeregpont helyzetéről.



7. ábra. Az acil-enzim szerkezet kialakulásának modellezéséhez használt potenciális energiafelületek grafikus reprezentációja. (A), az ONIOM(B3LYP/6-31G(d,p):AMBER) és (B) az ONIOM(M06-2X/6-31G(d,p):AMBER) módszerek esetében. (A pásztázáshoz az Sy–C(NMA-karbonil) és az (imidazólium-Nδ)Hδ–N(NMA amid) belső koordináták kerültek felhasználásra.)

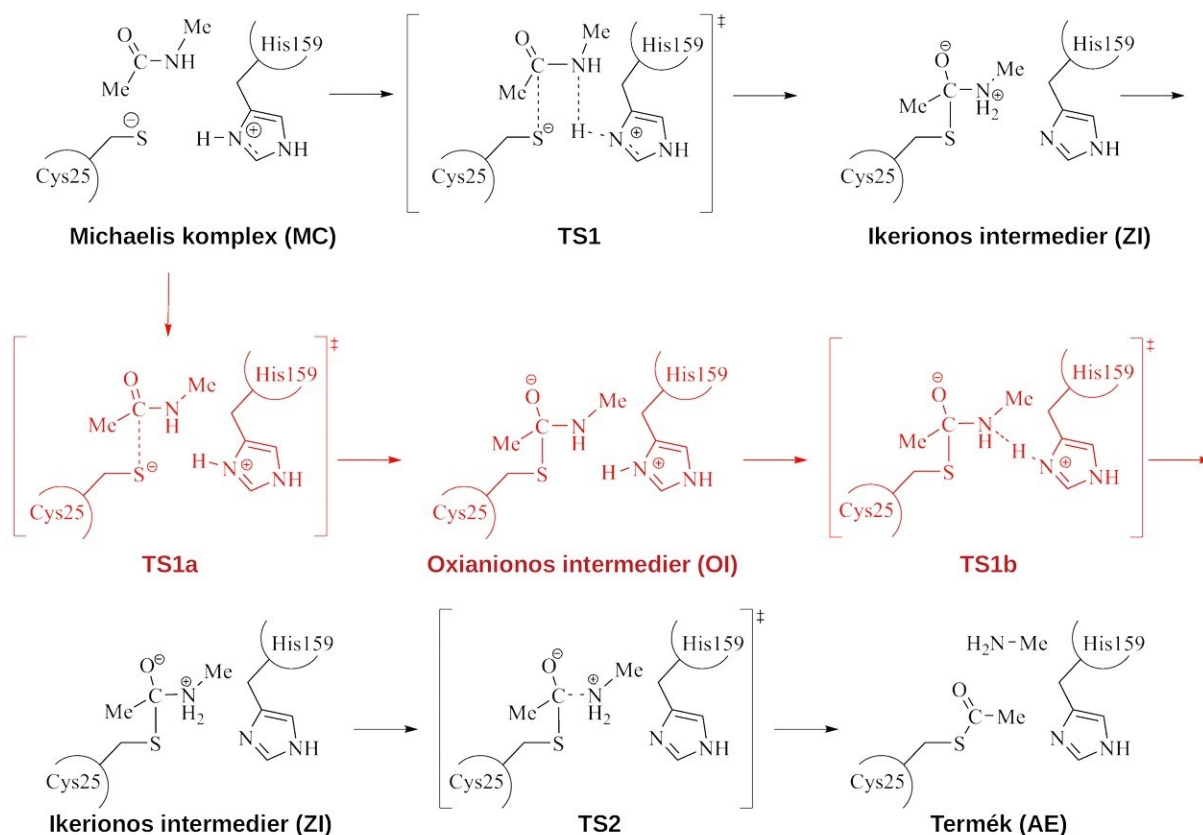
A potenciális energia felületek pásztázása során nem tapasztaltunk spontán C-N kötéshasadást az NMA szubsztrátnál, így annak érdekében hogy megvizsgáljuk ezen kötés felhasadásának jellemzőit, hasonlóan ez előbbiekhöz kétdimenziós pásztázást végeztünk a C-N(amid) és a Cys25Sy–C(amid), valamint az (amid)N–Hδ(His159) kötéstávolságok mentén. (K2. ábra). Érdekes módon, az utóbbi kombinációban a C–N kötés felhasadásával párhuzamosan lejátszódott az S–C(amid) kötés kialakulása, és az átmeneti állapot helyét ezen felület alapján becsültük meg (K2. ábra, C és D panel). Fontos megjegyeznünk, hogy a B3LYP módszer esetében a C–N kötés jelentősen megnyúlt (>2,0 Å), mely szintén magyarázhatja az egy elemi lépésben lejátszódó acilezési lépést adó korábbi eredményeket [23].



8.ábra. Az acilezési lépés különböző módszerekkel és báziskészletekkel számolt energiaprofiljai.

Harrison és munkatársai által kivitelezett hibrid QM/MM számítások alapján [17], majdnem tökéletesen szinkronban van egymással az S<sub>γ</sub> atom nukleofil támadása és a proton transzfer (B3LYP/3-21G\*:AMBER). Közel szinkronfolyamatot találtak Wei és munkatársai is ugyanezen esetben [23] az általuk publikált átmeneti állapot geometriája alapján (B3LYP/6-31++G\*\*):AMBER). Az eltérés ezen irodalmi eredmények és az általunk publikáltak között feltehetően az alkalmazott elméletek, az eltérő modellek és az eltérő QM/MM particionálási sémák különbözőségében lelhető fel. Ahogyan azt említettük, ezen részletek könnyedén befolyásolhatják megfigyeléseinket. A második átmeneti állapot, mely összeköti a tetraédres intermedier és az acil-enzim intermediert, jelentősen lazult C–N kötést tartalmaz (~2,0 Å), ami modellünkben a hasítandó peptidkötés megfelelőjének tekinthető az NMA szubsztrát jelenlétében. Meg kell említenünk hogy a B97-D módszer esetében a legkisebb alkalmazott báziskészlettől (6-31G(d,p)) eltekintve, minden más esetben figyelemreméltó eltéréseket tapasztaltunk. Szintén fontos az is, hogy a B3LYP módszer esetében jelentősen alacsonyabb aktiválási gátat kaptunk az acilezési lépés utolsó részfolyamatára. Mindazonáltal az acil-enzim képződésére az eredményeink nagyon hasonlóak. Általánosságban véve a B97-D módszer, báziskészlettől függetlenül némiképp hosszabb S–C távolságot adott, más módszerekhez képest, ugyanakkor a B3LYP módszer esetében voltak tapasztalhatóak a leghosszabb távolságok, míg más módszerek, kiváltképp az M06-2X jelentősen rövidebb

távolságokat adott. Ezen megfigyelés magyarázható azzal, hogy a B3LYP módszer nem tartalmaz hosszútávú elektron-elektron kicserélődést figyelembe vevő tagot. A 3. táblázat adatait összevetve Wei és munkatársai eredményeivel jelentősen rövidebb  $d4$  távolságot (2,09 Å) kaptak, mely erősebb hidrogénkötést sugall az NMA amid hidrogénje és a hisztidin N $\delta$  atomja között.



9. ábra. Az acilezési részlépés mechanizmusának összefoglalása.

Az aktiválási energiák tekintetében általános következtetés vonható le, miszerint az acil-enzim képződése, azaz az első elemi lépés determinálja a folyamat sebességét. A számolt ZPVE-korrigált energiák  $\sim 10-13$  kcal mol $^{-1}$  tartományban vannak, az alkalmazott módszertől függően. Minden esetben azt találtuk, hogy az ikerionos intermedier kevésbé stabilabb a Michaelis komplexnél. Az acil-enzim intermedier képződéséhez vezető, második átmeneti állapot, sokkal könnyebben végbemehet, mint az első lépés. A képződött acil-enzim + metilamin rendszer teljes energiája pedig magasabb a kiindulási komplex energiájánál (3. táblázat, (TS2-TI) – (TS2-AE) vs. (TS1-TI) – (TS1-MC)). Ez önmagában azt jelenti, hogy a reakció el van tolva a reaktánsok irányába, ugyanakkor feltételezhető, hogy élettani körülmények között specifikus kötőhelyek hiányában a képződő amin termék gyorsan disszociál az enzim felületéről, mely az acil-enzim termék irányába tolja el a reakció lefolyását.

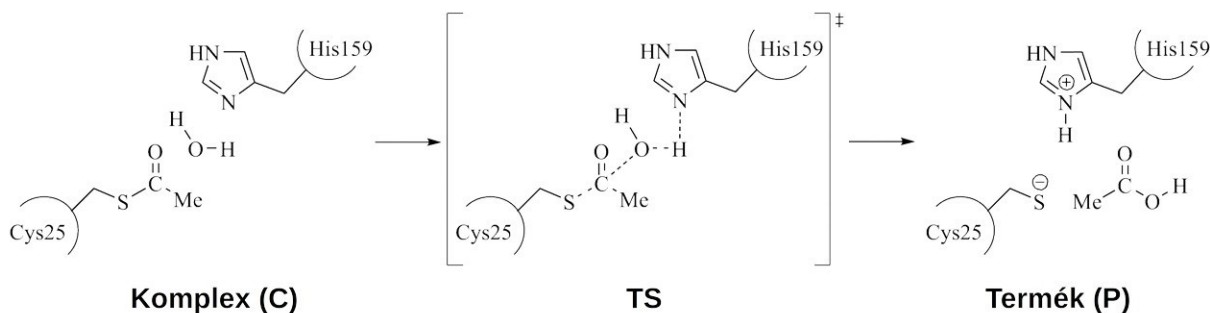
Figyelembe véve a korrelációs és kicserélődési tagokban alkalmazott paraméterek számát, melyeket a B97-D,  $\omega$ B97X-D és M06-2X módszerekben alkalmaztak, nehéz megmagyarázni az energiák és a geometriák tekintetében kapott eltérő eredmények forrását. Mindazonáltal általános megfigyelés, hogy a B3LYP módszer minden esetben jelentősen hosszabb nemkötő atomok közötti távolságot adott, mely a már korábban említett, hosszú távú kölcsönhatásokat megfelelően figyelembe vevő korrelációs és kicserélődési tagok hiányának tudható be, ezen túl a kémiai kötések tekintetében hasonlóságok fedezhetőek fel.

### 6.1.3 Az acil-enzim intermedier hidrolízise

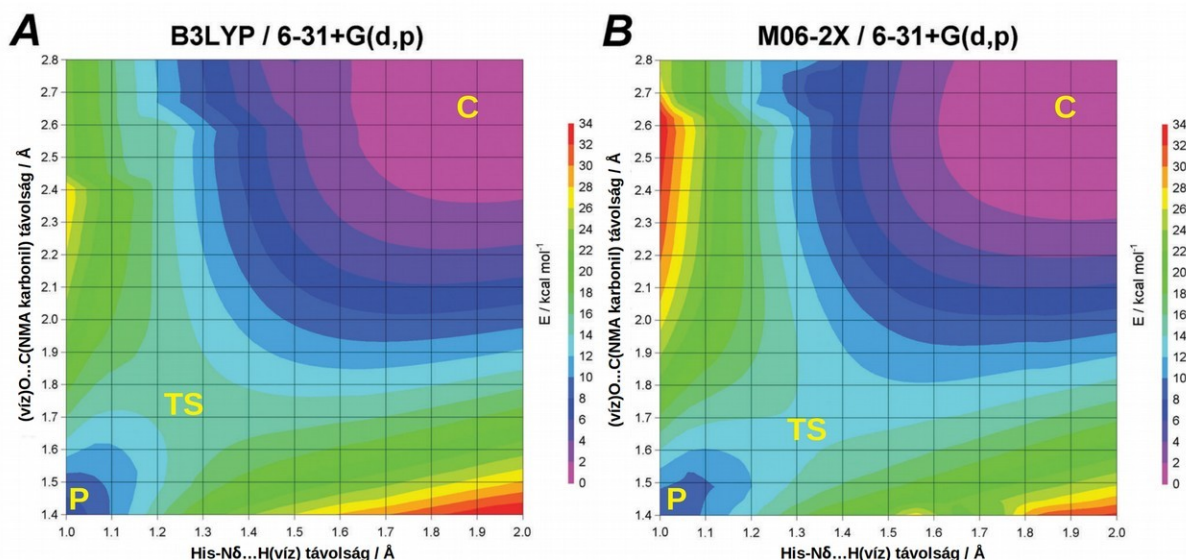
Az acil-enzim intermedier képződése után a következő lépés, annak hidrolízise (10. ábra). Ezen reakció vizsgálata során az acilezési lépés metilamin termékének metil csoportját eltávolítottuk, és a megmaradt, csonka amino csoportban a nitrogén atomot oxigénre cseréltük, azaz egy víz molekulát alkalmaztunk. A távolságok, melyeket felhasználtunk a hidrolízis követésére az S–C ( $d1$ ), O(víz)–C ( $d2$ ), H(víz)–O(víz) ( $d3$ ), és a H(víz)–N(imidazol) ( $d4$ ) távolságok voltak (4. ábra, C panel). Az acil-enzim–víz komplex esetén (C) a kapott távolságparamétereket az 5. táblázat foglalja össze. A B3LYP és B97-D módszerek esetében csupán csekély mértékben nőttek meg az S–C és C–O távolságok. Ezen paraméterek egyike sem mutatott jelentős függőséget az alkalmazott báziskészletek minőségétől.

Azon átmeneti állapotot (TS) illetően, mely az acil-enzim–víz komplexet (C) összeköti a papain–ecetsav komplexszel (P), tulajdonképpen nagyon hasonló eredményeket kaptunk a geometriákat illetően (5. táblázat). A számolt átmeneti állapotokra a  $d2$  távolságok esetében már említettem azok némiképp hosszabb voltát. A  $d3$  kötés felhasadás és a  $d4$  kötés kialakulás esetében tulajdonképpen azonos eredményeket kaptunk, csupán az  $\omega$ B97X-D módszerrel számoltunk kissé rövidebb  $d3$  távolságot a  $d4$ -hez képest.

A számolt PES-ek B3LYP/6-31+G(d,p) és M06-2X/6-31+G(d,p) elméleti szintek esetén szinkron reakciót ad a  $d2$  és  $d4$  távolságok alapján (11. ábra, A és B panel), míg a TS geometriáknál a  $d1$  távolság csekély mértékben hosszabb, mint a reaktánsokban (5. táblázat). Az S–C ( $d1$ ) kötéshasadás spontán bekövetkezik rövidebb  $d2$  és  $d4$  távolságoknál. Mivel belső reakciókoordináta (Intrinsic Reaction Coordinate, IRC) számításokra nem volt módunk, így manuálisan ellenőriztük azt, hogy a kapott átmeneti állapot valóban összeköti-e a kiindulási komplexet a termékekkel. A geometriaoptimalások a TS előtti pontból kiindulva a kiindulási acil-enzim–víz komplexhez, míg a TS utáni pontokból kiindulva papain–ecetsav komplexhez vezettek (P), ezzel az átmeneti állapot valódiságát igazoltuk.



10. ábra. Az deacilezési részlépés mechanizmusának összefoglalása



11. ábra. Az acil-enzim hidrolízis modellezéséhez használt potenciális energiafelületek grafikus reprezentációja. (A), az ONIOM(B3LYP/6-31+G(d,p):AMBER) és (B) az ONIOM(M06-2X/6-31G(d,p):AMBER) módszerek esetében. (A pártázáshoz a (víz)O–C(acil karbonil) valamint a (víz)H–Nδ(imidazol) belső koordináták kerültek felhasználásra.)

Ezen adatok alapján egy elemi lépésből álló reakció írható fel, melyben az O(víz)–C(karbonil) és a H(víz)–Nδ(hisztidin) kötések kialakulása megelőzi az S–C kötés hasadását. A termék geometriák tekintetében szintén nagyon hasonló eredményeket kaptunk, melyek közül a B3LYP módszerrel végzett számítások ismét kissé eltérő eredményeket adtak a nemkötő távolságokat illetően.

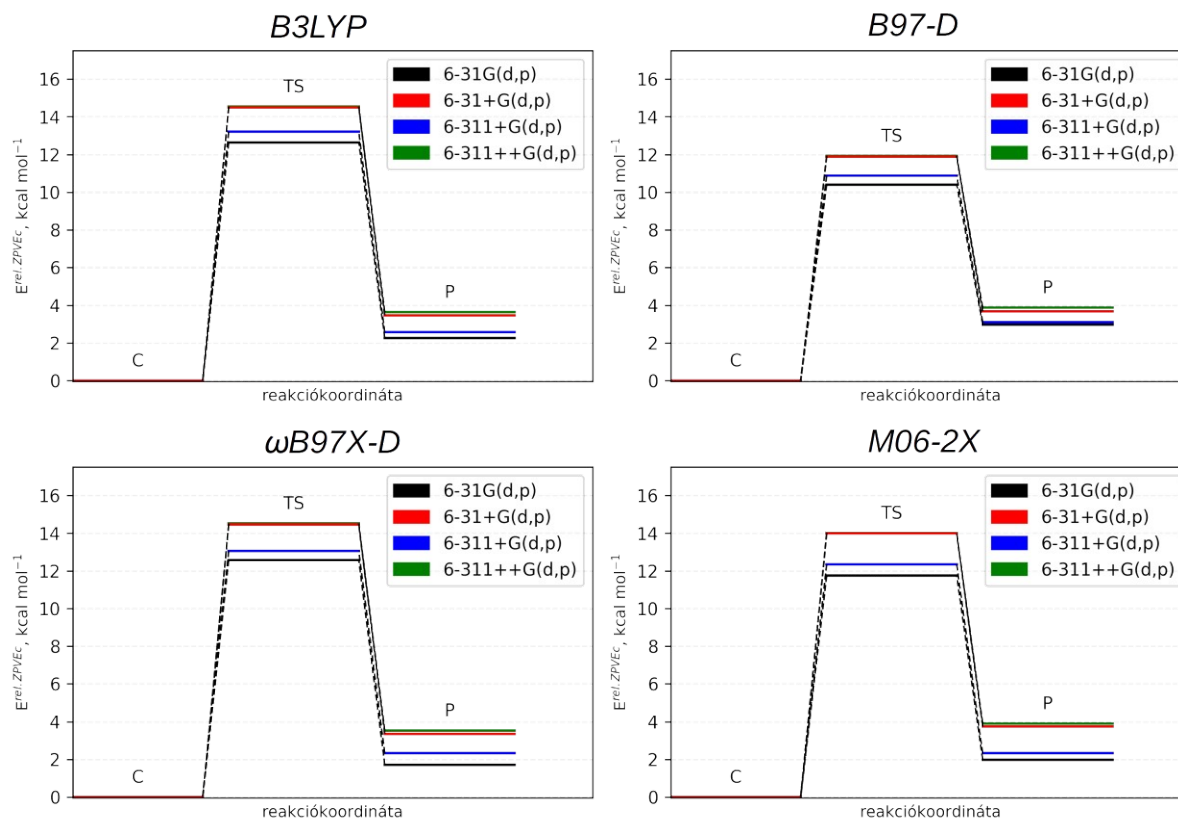
5. táblázat. A papain dezacilezési mechanizmusának tanulmányozása során nyert, válogatott szerkezeti és termodinamikai paraméterek összefoglalása. (Az atomok közötti távolságok angströmben, a termodinamikai paraméterek kcal mol<sup>-1</sup> egységekben kerültek megadásra.)

Módszer	Bázisfüggvény	Acil-enzim – víz komplex (C)				TS				Termék (P)				TS-C	TS-P
		d1	d2	d3	d4	d1	d2	d3	d4	d1	d2	d3	d4	E <sup>ZPVEc</sup> , kcal mol <sup>-1</sup>	
B3LYP	6-31G(d,p)	1,808	2,741	0,979	1,965	2,008	1,703	1,235	1,268	3,036	1,354	3,149	1,085	12,639	10,378
	6-31+G(d,p)	1,802	2,803	0,981	1,978	1,964	1,752	1,238	1,264	3,060	1,354	3,273	1,076	13,199	10,634
	6-311+G(d,p)	1,804	2,814	0,978	1,986	1,976	1,757	1,255	1,247	3,060	1,352	3,264	1,075	14,536	10,897
	6-311++G(d,p)	1,804	2,814	0,978	1,987	1,977	1,758	1,255	1,246	3,061	1,353	3,269	1,073	14,501	11,039
B97-D	6-31G(d,p)	1,819	2,801	0,984	1,938	2,079	1,712	1,263	1,249	2,969	1,369	2,851	1,096	10,397	7,432
	6-31+G(d,p)	1,813	2,791	0,986	1,936	2,024	1,756	1,271	1,240	2,979	1,369	2,984	1,085	10,878	7,773
	6-311+G(d,p)	1,815	2,808	0,982	1,946	2,038	1,756	1,282	1,229	2,981	1,367	2,976	1,082	11,928	8,045
	6-311++G(d,p)	1,815	2,808	0,982	1,947	2,038	1,757	1,281	1,230	2,985	1,367	2,981	1,080	11,894	8,203
ωB97X-D	6-31G(d,p)	1,794	2,668	0,975	1,947	1,965	1,678	1,198	1,290	2,993	1,347	2,892	1,074	12,571	10,863
	6-31+G(d,p)	1,788	2,721	0,976	1,951	1,932	1,710	1,189	1,299	3,004	1,346	2,979	1,067	13,055	10,727
	6-311+G(d,p)	1,790	2,733	0,973	1,956	1,941	1,711	1,202	1,282	3,003	1,344	2,968	1,065	14,506	10,981
	6-311++G(d,p)	1,790	2,734	0,974	1,956	1,941	1,711	1,202	1,283	3,004	1,344	2,969	1,065	14,439	11,087
M06-2X	6-31G(d,p)	1,794	2,625	0,975	1,985	1,977	1,637	1,185	1,310	2,945	1,348	2,780	1,080	11,741	9,767
	6-31+G(d,p)	1,790	2,655	0,975	1,998	1,945	1,664	1,173	1,325	2,959	1,347	2,856	1,071	12,347	10,026
	6-311+G(d,p)	1,792	2,663	0,973	1,993	1,956	1,662	1,189	1,303	2,959	1,347	2,837	1,065	14,005	10,097
	6-311++G(d,p)	1,792	2,664	0,973	1,993	1,956	1,662	1,189	1,304	2,960	1,347	2,837	1,064	13,983	10,243

E<sup>ONIOM</sup>, ONIOM extrapolált energia

E<sup>ZPV</sup>, Zérusponyi vibrációs energia

E<sup>ZPVEc</sup> = E<sup>ONIOM</sup> + E<sup>ZPV</sup>



12.ábra. Az dezacilezési lépés különböző módszerekkel és báziskészletekkel számolt energiaprofiljai.

Korábbi eredmények arra utalnak [23], hogy a reakció két különálló lépésből áll, ugyanakkor ezen eredmények alapján a második részlépés aktivációs (szabad) energiája, rendkívül alacsony, valamint az első részlépés átmeneti állapota és az általunk közölt eredmények között hasonlóság figyelhető meg. Az O(víz)–C(karbonil), O(víz)–H(víz) és a H(víz)–N $\delta$ (hisztidin) távolságok 1,68 Å, 1,16 Å és 1,32 Å-nek adódott Wei és munkatársainak munkájában, melyek jó egyezést mutatnak a  $d_2$ ,  $d_3$ ,  $d_4$  távolságokkal, amelyek a 4. táblázatban kerültek összefoglalásra. A második részlépés a S–C kötés hasadáshoz tartozik, ami a mi esetünkben nem spontán következett be.

Az átmeneti állapotok energiái az alkalmazott sűrűségfunkcionál módszerekkel a  $\sim 10$ – $14$  kcal mol $^{-1}$  tartományba estek (5. táblázat, 12. ábra). Általánosságban véve a legnagyobb báziskészletekkel számoltuk a legmagasabb aktiválási gátakat, illetve az is belátható, hogy a termékoldali komplex energiája mindig magasabb, mint az acil-enzim–víz komplex energiája. Ahogyan azt említettük az acilezési lépés esetében, nem vizsgáltuk sem az explicit szolvans hatását, sem pedig az ecetsav eltávolását a számolt reakció energetikájára vonatkozóan.

Az acilezési és a dezacilezési reakciókat összehasonlítva a sebességmeghatározó lépésekhez tartozó aktiválási gátak összemérhetőek, de a dezacilezési lépés kissé magasabb aktiválási gáttal rendelkezik. Általánosan elfogadott, hogy az amid kötés hidrolízisének sebességmeghatározó lépése az acilezési lépés szemben az észterhidrolízissel, melynél a

dezacilezési lépés a sebességhatározó [3,4]. A kevés cisztein-proteázok teljes mechanizmusát vizsgáló elméleti munka alapján, valóban összemérhető gáttal bír eme két részlépés [22,23].

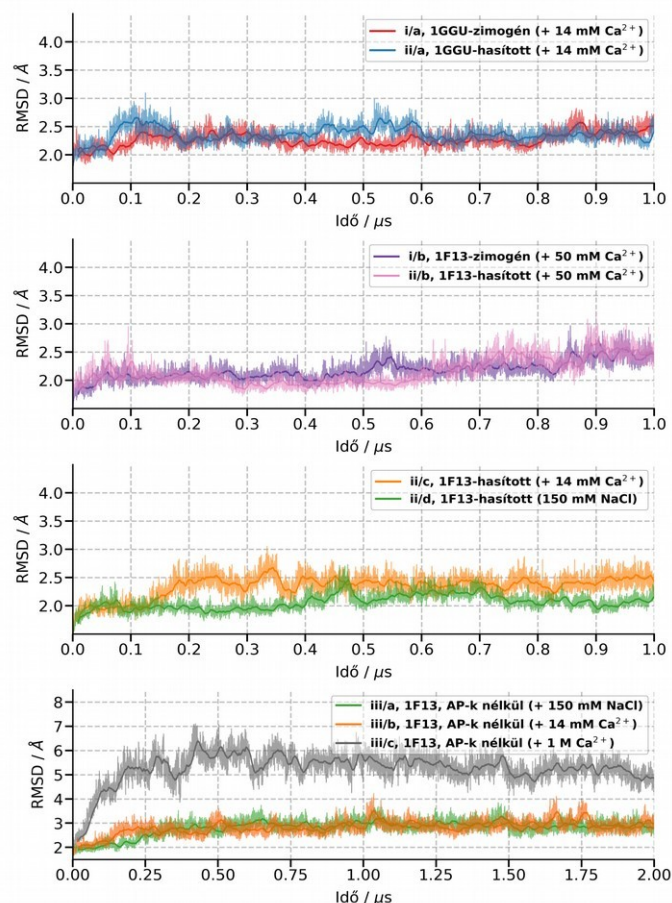
Alkalmazva az Eyring-Polányi egyenletet a  $k_{cat}$  értékekre [124–126] (melyek tipikusan az  $1-50 \text{ s}^{-1}$  tartományba esnek, szubsztráttól, hőmérséklettől és más paraméterektől függően),  $310 \text{ K}$  hőmérsékleten és a transzmissziós együtthatót egységnyinek véve az aktivációs szabadentalpia  $15,8-18,1 \text{ kcal mol}^{-1}$  közöttinek adódik. Feltételezve, hogy sem az entrópia, sem pedig az entalpia  $p\Delta V$  tagjának nincs jelentős hozzájárulása, ezen értékeknek közel kell lennie az aktiválási gáttalhoz. A reakciókinetikai paraméterekből származtatott aktivációs energiák, vagy szabadentalpiák az  $\sim 5-18 \text{ kcal mol}^{-1}$  tartományba esnek [23]. Ezen értékek származtathatóak a  $k_{cat}$  sebességi állandókból, az Arrhenius egyenlet segítségével és az általános átmeneti állapot elméletből [127]. Tehát a nyilvánvaló és szükségszerű egyszerűsítéseink ellenére, a modell rendszereink alapján számolt értékek megbízhatónak tekinthetők.

## 6.2 Az FXIII-A<sub>2</sub> és TG2 aktivációja

### 6.2.1 A molekuladinamikai szimulációk eredmények

Tekintettel a XIII-as véralvadási faktor aktivációjának bonyolult voltára, olyan megközelítést alkalmaztunk a munka során, minek segítségével elkülöníthetően tárgyalhatóak azon jelenségek, melyek hozzájárulnak az aktív enzim létrejöttéhez, kiindulva a zimogén szerkezetekből (*i/a,b*) a thrombin által aktivált formán át (*ii/a-d*) azon elméleti jelentőséggel bíró szerkezetiig, melyben az aktivációs peptidek (AP) már disszociáltak a homodimer felszínéről (*iii/a-c*). Ezen három szimulációs csoport molekuladinamikai szimulációjának során nyert RMSD értékei a 13. ábrán láthatóak.

Általánosságban elmondható, hogy a kalciumionok jelenléte növeli az RMSD értékeket minden esetben (13. ábra), a referencia modell szimulációja során kapott értékekhez képest. Ugyanakkor érdekes, hogy az  $50 \text{ mM Ca}^{2+}$  koncentráció nem befolyásolja számottevően a fehérjék dinamikai sajátságait, noha ezen rendszerek (*i/b*, *ii/b*) sokkal lassabban konvergálnak a  $2,5 \text{ \AA}$  -ös végső RMSD érték felé, mint bármely más. Nem zárható ki, hogy sokkal hosszabb időskálán komolyabb szerkezeti változások is végbe mennének, azonban erre nem találtunk közvetlen bizonyítékot.

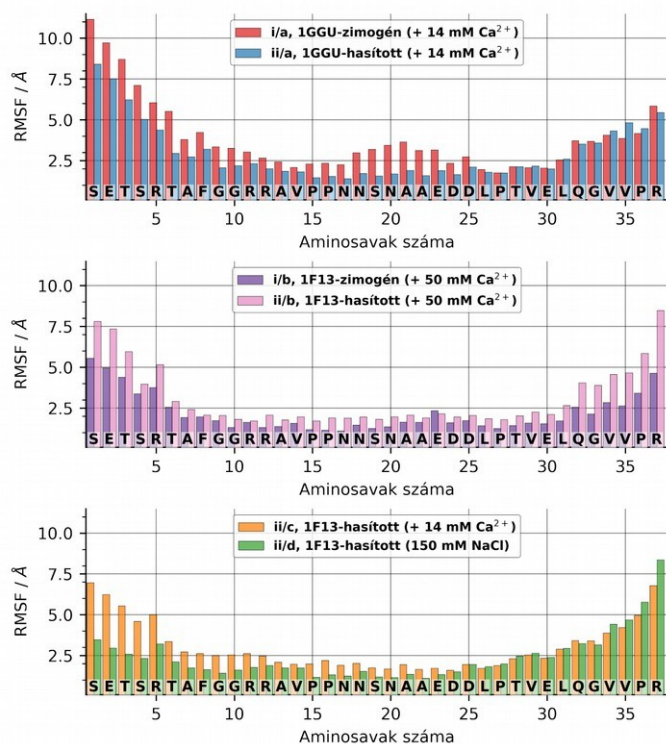


13. ábra. Nehézatom RMSD értékek az FXIII-A<sub>2</sub> szimulációkban. (Az összehasonlíthatóság érdekében az aktivációs peptideket figyelmen kívül hagytuk.)

Elmondható, hogy az *i* és *ii* szimulációs csoportok modellrendszerei alapvetően hasonló dinamikai tulajdonságokkal rendelkeznek, melynek gyökere az aktivációs peptidek jelenlétében keresendő. A kalciumionok jelenlététől függően azonban jelentősen nőtt minden esetben az N-terminálit tizenhét aminosav fluktuációja (14. ábra).

Ez a megfigyelés fontos abban a tekintetben, hogy az egyik monomer Arg11 és Arg12 aminosavai egy mély üregben találhatóak, melyek a másik alegység negatívan töltött aminosavainak oldalláncai által (Asp343, Asp345, Asp367, Glu401) meglehetősen erős sóhíd kölcsönhatásokat jelentenek [54]. Az aktivációs peptidek disszociációja a dinamikai időskálán feltehetően a ritka események körébe tartoznak, jóllehet kisebb változások megfigyelhetőek például a *ii/c* modell esetében, de az AP-k teljes disszociációja nem volt megfigyelhető. Az első harminchét aminosav esetében bekövetkezett megnövekedett fluktuációk grafikus ábrázolása a 14. ábrán látható, így feltételezhető, hogy azok képesek lehetnek disszociálni a homodimer felszínéről. A 2  $\mu$ s hosszúságú szimulációk esetében, ahol az aktivációs peptideket eltávolítottuk (*iii/a-c*), mindhárom esetben emelkedtek az RMSD értékek, ami arra enged következtetni, hogy az aktivációs peptidek fő szerepe az intakt dimer szerkezet fenntartása az

aktiváció bekövetkezéséig. A forgási (girációs) sugarak ( $r_{\text{gyr}}$ ) tekintetében szinte kivétel nélkül minden rendszer hasonló értékekkel rendelkezik, a szimulációk során átlagos értéke  $36,90 \pm 0,1 \text{ \AA}$  volt. A lánconként számított értékek megtalálhatóak a függelékben (K4 és K5. ábra).



14. ábra. Az aktivációs peptidek aminosavanként átlagolt RMSF értékei az FXIII-A<sub>2</sub> szimulációkban.

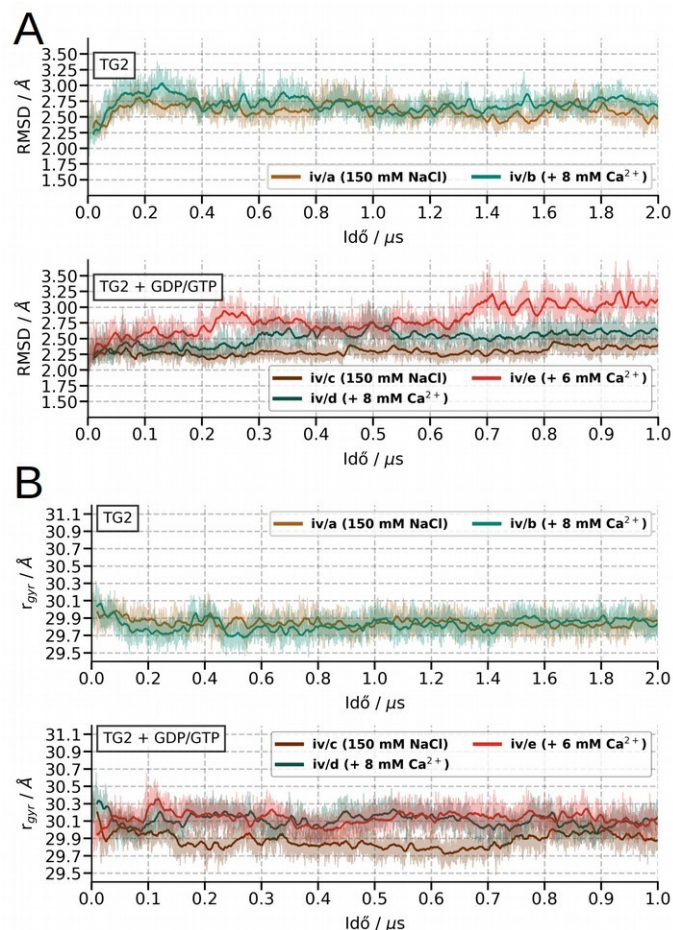
A fluktuációk tekintetében számos régió azonosítható magas lokális fluktuációval, azonban ezek jó része hurok régió, vagy az aktivációs peptid N- vagy C-terminálisához térben közeli szegmens (K6-K8. ábra). Kiemelendő azonban négy régió, melyek rendre a Val274-Asp280, Met350-Trp370, Ile440-Ile460 és Pro505-Arg515 aminosavak között található. Az utolsó három régió térben egymáshoz közel található, melyből a Met350-Trp370 és az Ile440-Ile460 három antiparallel béta-redőben, illetve egy rendezetlen hurokban van. A hurok közvetlen kapcsolatban áll a másik monomer aktivációs peptidjének N-terminálisával (többek között biztosítva a már említett sóhidat képző, negatívan töltött aminosavakat), míg az Ile440-Ile460 térben szomszédos a Pro505-Arg515 hurokrégióval, amelynek N-terminálisán található hélix szerkezet szolgáltatja a fő kalciumkötőhely aminosavait, így egyfajta allosztérikus útvonal feltételezhető ez alapján. Függetlenül attól, hogy a korábban javasolt [128] és a krisztallográfiai eredmények által sugallt [55] aktiváció során bekövetkező nagy konformációs változások közül melyiket tartjuk elfogadottnak, a Met350-Trp370 és az Ile440-Ile460 aminosavak által közrefogott régiók mindkét esetben alapvető fontossággal bírnak, mivel ezekkel szomszédos az 1-es  $\beta$ -hordó domén, melynek elmozdulása nem csak az FXIII-A<sub>2</sub> esetében kísérletileg alátámasztott, de a szöveti transzglutamináz esetében is dokumentált.

Az aktivációs peptid nélkül végzett szimulációk során ezen fluktuációk még inkább megnőnek, mely bizonyítani látszik azon hipotézist, hogy kiemelt fontosságot kell tulajdonítsunk ezen régióknak.

Követve azon elképzelésünket, hogy lépésenként tárgyaljuk az egyes faktorok jelentőségét, a szöveti transzglutamináz példáján megvizsgáltam a kalciumionok és a GDP egyedi és együttes hatását is. A GDP nélkül kivitelezett szimulációk esetében nem találtunk jelentős különbséget az RMSD és  $r_{\text{gyr}}$  értékek időbeli változása között, azonban figyelembe véve a GDP jelenlétét az alábbi tendencia fedezhető fel:  $iv/a \sim iv/b > iv/d > iv/c$  (15. ábra). Az, hogy a GDP jelenlétében jelentősen csökkent az RMSD, összhangban van azzal a megfigyeléssel, hogy a GDP (zárt) forma inaktív, azaz transzglutamináz aktivitással nem rendelkezik. A kalciumionok hatására azonban az RMSD megnő, tehát az *in vivo* aktivációt kiváltó  $\text{Ca}^{2+}$  ionok jelenléte befolyásolja a fehérje dinamikáját. A girációs sugarak vizsgálata során jelentősebb eltérés látható ez utóbbi esetben. Hasonló következtetések vonhatók el a *iv/e* (4PYG) modell szimulációja alapján, azonban ki kell emelni, hogy bár ezen szerkezet tartalmazza a natív Val224 aminosavat, de emellett három pontmutáció is jelen volt, melyek a Gln186Glu, Thr533Asn, Val655Leu aminosavakat érintik.

Az RMSD 650 ns után bekövetkezett növekedéséhez (15. ábra, A panel), ezen mutációk hozzájárulhatnak a kondenzált fázisban történő szimuláció során, azonban a girációs sugár alapján elmondható, hogy a fehérje nem szenvedett komolyabb változást, megtartotta feltekeredett szerkezetét. Az FXIII-A<sub>2</sub> esetében kiemelten nagy fluktuációkkal bíró régiók a TG2 esetében is megfigyelhetők (Ile313-Asn333, Val401-Val422, Asn460-Thr471), mely egyrészt az igen hasonló másodlagos- és harmadlagos szerkezetnek tulajdonítható, másrészt hangsúlyozza ezen szegmensek fontosságát az aktivációban betöltött szerepükre vonatkozóan. A fluktuációk változásának további tárgyalásánál az egyes szimulációk közötti különbségekre fókuszáltunk (K24. ábra).

Ahogy a fluktuációk közötti különbségekből látható, a legtöbb nagy eltérés a  $\beta$ -szendvics domént érinti, ahol számos magas fluktuációval rendelkező régió található, melyeknek feltehetően nincs jelentősége az aktiváció során (Leu12-Lys30, Gly64-Gly72, Leu79-Asp87). Az Arg240-Tyr245 régió fluktuációja kalcium jelenlétében megnő, ami olyan szempontból érdekes megfigyelés, hogy az aktív hely Cys277 aminosava ezen rendezetlen hurok régió takarásában található, valamint egy kalciumkötőhely található az Asp232 és Asp233 aminosavaknál, így ez a kötőhely indukálhat változásokat a kötőhely közelében amely hatással lehet a  $\beta$ -hordó doménekre is. A Thr343-Glu352 közötti hurokrégió szintén nagy fluktuációt mutat, de csupán a kalcium és GDP jelenlétét nélkülöző esetben (K24. ábra, A és D panel).



15. ábra. (A) Nehéz atom RMSD és (B) nehéz atom forgási sugarak a TG2 szimulációkban.

Abban az esetben ha kalcium is jelen volt a szimulációs rendszerben, ez a hurok rögzítve lett, így fluktuációja csökkent. Az Asn460-Thr471 régió tekintetében az alábbi magyarázat adható: ellentmondásos, hogy az FXIII-A<sub>2</sub>-ben a fő kötőhelyeket az Asp438, Gly457, Glu485, Glu490 aminosavak biztosítják és ezen aminosavak majdnem teljesen konzerváltak maradtak a TG2-ben is (Asp400, Ser419, Glu447, Glu452), azonban mégsem tudunk megfigyelni kalciumkötődést ezen a helyen. Az nyilvánvaló, hogy az Thr449-Ile460 aminosavak közötti  $\beta$ -szalag rendezetlen a TG2-ben (Ser411-Val422), tovább az is látható, hogy a kötőseb által biztosított zseb sokkal tágabb, mint az FXIII-A<sub>2</sub> esetében. Ezen feltételezett hélix mögötti kötőhely helyett egy másik hely került azonosításra a hélix alatt, mely tulajdonképpen lefelé húzza az Asn460-Thr471 hurok régiót. A fluktuációk változásának jelentősége lehet a hordó domének helyzetének, az aktiváció során bekövetkező változásánál, hiszen az említett hurok régió nagy valószínűséggel vesz fel egyenes  $\alpha$ -hélix konformációt, ami indukálhatja az 1-es  $\beta$ -hordó kötőhely irányába történő elmozdulását, továbbá, mindkét hordó domén fejjel lefelé történő elmozdulását, azaz az átmenetet a nyílt

konformáció irányába. Ezen hipotézis illeszkedik a svájci bicska-szerű konformációs aktiváció elméletébe [128].

### 6.2.2 A kalciumkötődés vizsgálata

Ahogy az azt korábbi tanulmányok során már igazolták, az FXIII-A<sub>2</sub> alegységként egy erős kötőhellyel rendelkezik. Számos szerkezet elérhető a protein adatbankban (<http://rcsb.org>) ahol két vagy három vegyértékű kationok vannak jelen a homodimerhez kötődve (1GGU, 1GGY, 1QRK, 1EVU [57]). Ezen tanulmányok alapján a dimer egy Ca<sup>2+</sup> vagy Sr<sup>2+</sup> iont köt alegységként ugyanazon kötőhelyen (Asp438, Gly457, Glu485, Glu490), mindazonáltal az itterbium(III) esetében végzett kísérletek során nyolc kötött Yb<sup>3+</sup> került azonosításra három különböző, részben aszimmetrikusan elhelyezkedő kötőhelyen (4 Yb<sup>3+</sup> az A/B láncok Asp270, Asp271, Glu272 aminosavai által; 2 Yb<sup>3+</sup> a B lánc Asp438, Glu485, Glu490 aminosavainál; 1 Yb<sup>3+</sup> az A lánc Asp438, Glu485, Glu490 aminosavainál és 1 Yb<sup>3+</sup> - az A lánc Asp574, Glu585 aminosavai által). Meg kell említenünk, hogy a 4KTY röntgenszerkezetben, mely a nyitottnak vélt (aktív) konformációnak felel meg, három kötött Ca<sup>2+</sup> található. Ezen kalciumionok közül egy, az itterbiumkötődés esetében már korábban azonosított helyen az Asp270, Asp271 aminosavaknál található. Ezen aminosavak egy hélix C-terminálisánál találhatóak és maga a kötőhely ismert a TG3 esetében is. Miután a krisztallográfiai tanulmányok során használt kationok vdW sugara közel megegyezik (~200 pm) így feltételezhető, hogy a dimer Asp270, Asp271 aminosavai, valóban képesek kötni kalciumiont vagy ionokat, jóllehet ezen kötőhely jelentőségét és funkcióját részleteiben eddig nem tárgyalták.

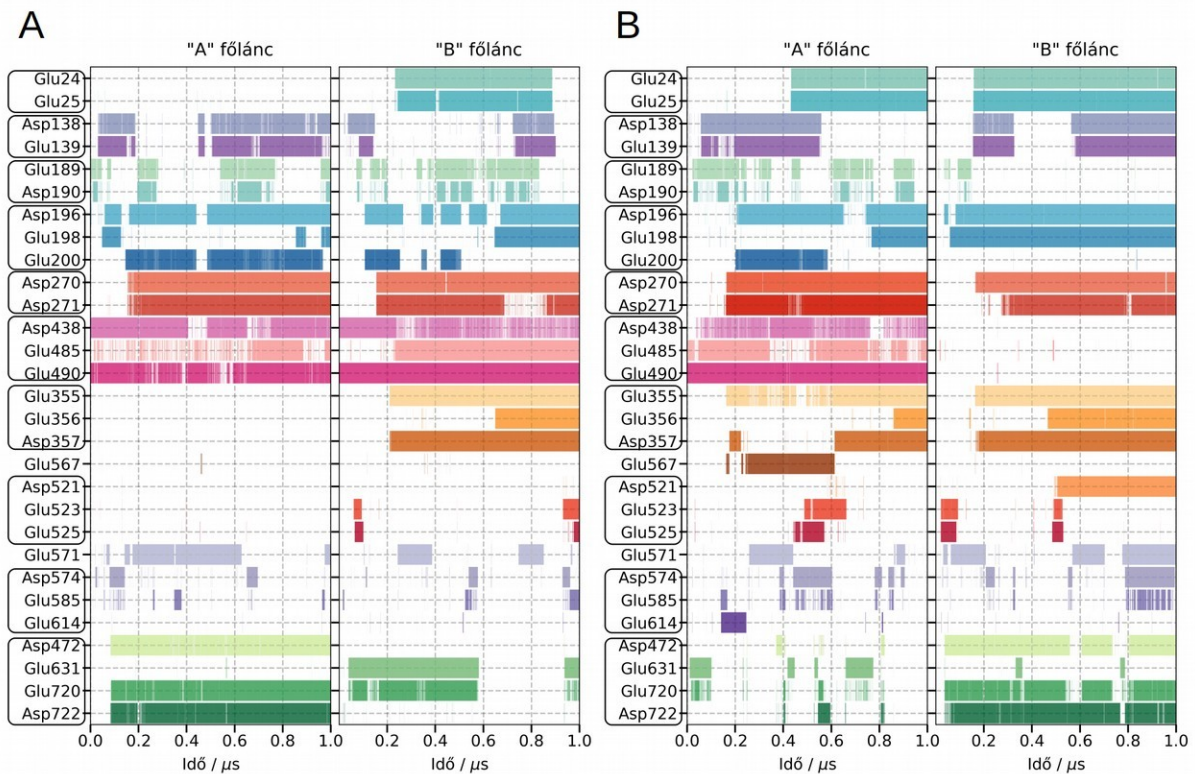
Molekuladinamikai szimulációink alapján számos lehetséges kötőhely került azonosításra (16. ábra), többségük térben közeli vagy szekvenciában egymást követő savas aminosavak által kialakított felszíni kötőhelyként került azonosításra. Annak érdekében, hogy kiszűrjük a gyenge és átmeneti kötőhelyeket, a következő kiválasztási kritériumot állítottuk fel. Az kalciumionokat akkor tekintettük kötöttnek, ha azok a megfelelő negatív töltésű oldallánc karboxilát szén atomjához képest 4,5 Å távolságon belül voltak találhatóak, a teljes szimulációs idő legalább 5 %-ában. A kalciumkötődések számszerűen teljes részletességükben megtalálhatóak az értekezés alapjául szolgáló közlemény kiegészítésben, míg a kiválasztási szabálynak megfelelő kötőhelyek grafikus ábrázolásai a függelékében is (K9-K13. ábra). Értekezésemben csupán azon kötőhelyek részletes vizsgálatára szorítkozok, melyek elsősorban konzervált kötőhelyeket jelenthetnek az FXIII-A, TG2 és TG3 enzimek vonatkozásában.

A kalciumkötődés vizsgálatánál mind az 1GGU (2 db kötött  $\text{Ca}^{2+}$ ), 1F13 (nincs kötött kalciumion jelen) szerkezetek vizsgálatát egyformán fontosnak gondoltuk. Ki kell emelnem, hogy bár az 1F13 szerkezet nem tartalmazott kötött kalciumot az *ii/c* szimuláció esetében az egyik alegység fő kötőhelye néhány 10 ns szimulációs idő elteltével betöltődött és az ion kötve maradt a szimuláció során, ugyanakkor a másik alegységnél nem detektáltam kötődési eseményt. Egy korábbi 2  $\mu\text{s}$  hosszúságú szimulációban, melynél a titrálható aminosavak protonáltságát nem vettük figyelembe 7,4 -es pH-n – ugyanakkor egyéb paramétereiben megegyezett az *ii/c* -vel – azt találtuk, hogy az A lánchoz tartozó fő kötőhely 500 ns szimulációs idő után, míg a B lánchoz tartozó fő site 900 ns után betöltődött (K12. ábra). Fontos megemlíteni, hogy ebben a szimulációban az alegység szomszédos Asp270, Asp271 aminosavai két kalciumot is képesek voltak kötni a Glu272 hozzájárulásával, ami később a protonáltsági fok számítása során protonálnak adódott. A lánconként fő kötőhelyek eltérő mértékben történő betöltődésére kézen fekvő magyarázat lenne az, hogy a Pro505-Arg515 hurokrégió főlánconként eltérő konformációban van jelen. Ezen régiók hiányoztak az 1GGU szerkezetben, így jelentőségük vizsgálatának céljából azok úgy kerültek modellezésre, ahogyan az 1F13-ban találhatóak. Az *ii/a* szimuláció során azt találtuk, hogy ezen hurokrégiók konformációjának nincs jelentősége a kalciumkötődés szempontjából, legalábbis a krisztallográfiai szerkezetben. Azaz a kalciumok helyzete változatlan maradt az 1  $\mu\text{s}$ -os dinamika során. Még részletesebben megvizsgálva ezen hurokrégiók másodlagos szerkezetét (K19. ábra), azt találtuk, hogy nincs összefüggés a kötődési események és a Pro505-Arg515 régió konformációja között (K9-K13. ábra).

A *ii/a* és *ii/c* szimulációk klaszterezése során nyert reprezentatívok grafikus megjelenítése a 17. ábrán látható. Az 1GGU-ban modellezett loop régiók, szinte tökéletesen megőrizték konformációjukat, azaz egyikük teljesen rendezetlen, másikuk gyakorta vesz fel egyeneses hélix elrendeződést, ami sokkal szembetűnőbb a DSSP analízisben, továbbá a kötőhelyek geometriája között nincs számottevő különbség. Ezen információk alapján nehéz megjósolni, hogy *in vivo* lehet-e kiemelkedő jelentősége ezen hurokrégióknak az FXIII-A<sub>2</sub> aktivációjában, miután a Pro505-Arg515 régió a köti össze a katalitikus központi és az 1-es  $\beta$ -hordó domént.

Fontosnak találtuk megvizsgálni azt, hogy az *ii/c* esetében miért nem detektáltunk kötődést a B lánc főhelye esetében, ezért az *ii/c* szimuláció utolsó pillanatképéből indulva, a kötött kalcium megtartásával újraszolvatáltuk a rendszert, 100 mM  $\text{Ca}^{2+}$  jelenlétében, majd az előkészítési lépéseket követően 17×25 ns szimulációt indítottunk, eltérő kezdeti sebességeket alkalmazva. Azt találtuk, hogy egyetlen esetben 11 ns után a B lánc fő kötőhelye is betöltődött

(K18. ábra). Így feltehetően a látens preferencia a fő kötőhelyeket illetően csupán a dinamika sztochasztikus természetének és az ionok véletlenszerű elhelyezésének tulajdonítható.

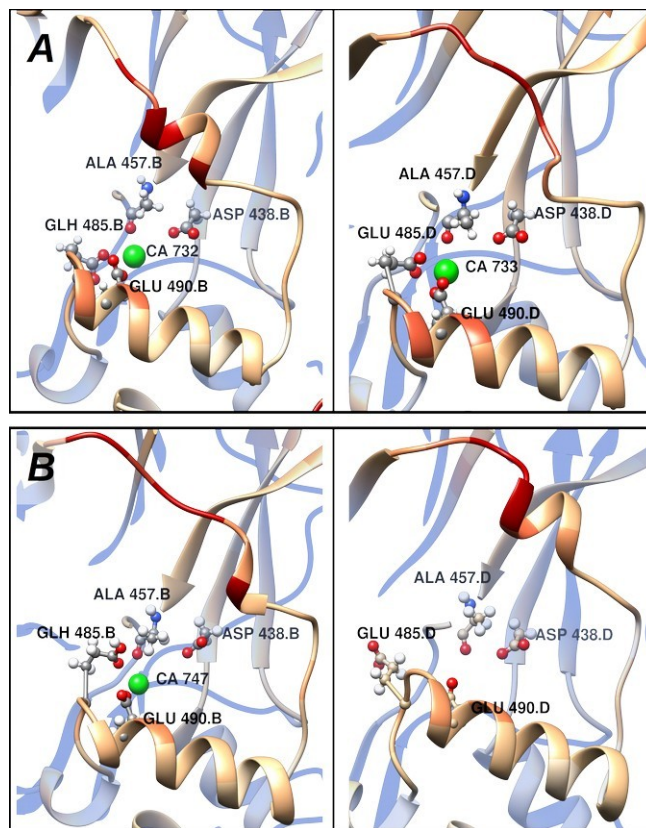


16. ábra. A dinamikai trajektóriából számolt kalciumkötődési események grafikus reprezentációja az (A) ii/a (1GGU, hasított AP-k, +14 mM  $Ca^{2+}$ ) és a (B) ii/c (1F13, hasított AP-k, +14 mM  $Ca^{2+}$ ) szimulációk esetében.

A második kötőhely, mely szinte minden esetben betöltöttnek volt tekinthető szimulációinkban az alegységek Asp270 és Asp271 aminosavainál található. Korábbi tanulmányok alapján ismert volt, hogy ezen a kötőhelyen találhatóak kötött  $Yb^{3+}$  ionok, valamint az aktív (4KTY) monomer szerkeztben egy kalciumion is fellelhető az Asp270 és Asp271 aminosavaknál, továbbá a TG3 szerkeztben (PDB ID: 1NUD [53], szintén monomer) is találunk itt kalciumot. Eredményeink azt sugallják, hogy ezen a helyen mindig található a kötéshez hozzájáruló aminosavak protonáltsági fokától (Glu272) és a kalcium koncentrációtól függően legalább egy kalcium.

Erőfeszítéseink ellenére ezen kötőhely szerepe az aktivációban továbbra is homályos maradt, ugyanakkor egyrészt meglehetősen közel található az alegységek között található sóhíd hálózathoz, másrészt ezen site-ok közvetlen kapcsolatban vannak a Ser278-Gly294 hurok régióval, amely kijelöli a katalitikus üreg határát és amely részben kapcsolatban van a 1-es  $\beta$ -hordó doménnel. Az utóbbi megállapítás érvényes a TG3 és a TG2 esetében is.

A 4KTY és az 1NUD szerkezetek alapján rendelkezésre álló szerkezeti információkból látható, hogy egy olyan kötőhely is található (Asp343, Asp345, Asp351, Asp367), amely az FXIII-A<sub>2</sub> (zárt) szerkezetben az aktivációs peptid által oldószer molekuláktól és ionoktól el van zárva. Az ionkötésért felelős oldalláncok közül az Asp343 és az Asp367 sóhíd kölcsönhatásban áll a másik alegység aktivációs peptidén található Arg11 és Arg12 aminosavakkal, míg az Asp345 és az Asp351 meglehetősen távol van a másik két aszpartáttól, így konformációs változások szükségesek ahhoz, hogy ezek valóban köthessenek kalciumot.



17. ábra. A FXIII-A<sub>2</sub> fő kötőhelye(i) a trajektória klaszterezése után kapott reprezentatívokban. (A) ii/a és (B) ii/d esetekben, bal oldalon az “A” lánc kötőhelye, a jobb oldalon a “B” főlánc kötőhelye látható.

Az aktivációs peptid jelenlétében végzett szimulációk így alkalmatlanok arra, hogy ezen feltételezett kötőhely szerepét vizsgáljuk. Az *iii/b* szimulációban, ugyan az AP-ek már nem voltak jelen, de mégsem tudtuk ezen kötőhelyet teljes egészében megfigyelni, csupán az A lánc esetében detektáltunk kalciumkötődést az Asp343 és Asp345 aminosavaknál. Ahogy azt már említettük, a rendkívül magas 1000 mM Ca<sup>2+</sup> koncentráció jelenlétében (*iii/c*), nagymértékű alegységmozgás figyelhető meg, ebben az esetben mindkét láncban detektálhatunk 250 ns után kalciumokat az említett négy aszpartát hozzájárulásával, valamint szinte minden gyengébb kötőhely betöltődött (K11. ábra, B panel). Feltételezhető tehát, hogy

ezen kötőhely csupán akkor válik fontossá, ha már az aktivációs peptidek disszociáltak a dimer felszínéről. Az alegységek elfordulása feltételezhetően összefüggésben áll azzal, hogy a nagy ionerősség letöri az alegységek közötti elektrosztatikus kölcsönhatásokat, míg a  $\beta$ -hordó domének elmozdulása nem volt megfigyelhető, továbbá az sem világos, hogy mi indukálja a figyelemre méltóan nagy elmozdulásukat. Hangsúlyoznunk kell az Asp472, Glu631, Glu720, Asp722 aminosavak által definiált kötőhely jelentőségét, mely tulajdonképpen a  $\beta$ -hordók és a katalitikus központi domén között teremt kapcsolatot.

Fontos kérdés lehet, hogy az egyszerűsített molekulamechanikai paraméterek vajon mennyire képesek pontosan prediktálni a kalciumionok kötődését. Ezért úgy döntöttünk, hogy az *ii/c* szimuláció utolsó szerkezetéből kiindulva, a kötött kalciumokat megtartva, “*multi-site*” ionmodell segítségével megvizsgáljuk a kialakult kötőhelyeket és 50 ns hosszúságú NpT dinamikát futtattunk. Eredményeink alapján egyértelműen látszik, hogy az 1  $\mu$ s hosszúságú dinamika során kialakult kötőhelyek, megtartották jelentőségüket a “*multi-site*” ionmodell alkalmazásával is (K14. ábra), ezen szimuláció utolsó képe, a koordinálódó vizekkel együtt, megtalálható a disszertáció alapját képező publikáció kiegészítésében. Megemlítenénk, hogy a 4KTY szerkezet segítségével végzett FXIII-A<sup>o</sup> szimulációban öt kötőhelyet találtunk (K13. ábra), melyek szinte folyamatosan kötöttek kalciumot a dinamika során, ez alapján jelentőség tulajdonítható nekik, az aktiváció valamely fázisában, ugyanakkor a zimogén és hasítással aktivált szimulációs modellek esetében ezen kötőhelyek közül kettő (Asp138, Glu139 or Glu355, Glu356, Asp357) gyengének adódott, illetve a rendelkezésre álló információk alapján nem tudtunk különösebb szerepet tulajdonítani nekik.

A szöveti transzglutamináz esetében végzett irányított mutagenézis kísérletek alapján, sikeresen meghatározásra került öt kötőhely a hatból [68]. A kísérletek részben szerkezeti homológia, részben pedig a felületi elektrosztatikus töltéseloszlás által definiált negatív töltésű centrumok alapján lett megtervezve. A körültekintő munka ellenére, ezen kísérletek alapjául statikus krisztallográfiai szerkezet szolgált, mely csupán egy olyan szerkezet szórásképéből adódott, mely feltehetően egy alacsony energiával bíró, így nagy számban betöltött részét képezi a valódi sokaságnak. Annak ellenére, hogy a molekuladinamikai szimulációk által leírható időtartományt behatárolja a szimulálni kívánt rendszer mérete, valamint az elérhető számítási erőforrás, a módszer minden kétséget kizáróan alkalmas módszer az olyan atomi jelenségek vizsgálatára, melyeket kísérletes úton, valós időben és megfelelő felbontásban igazolni másképp nehezen lehet. Mivel hasonlóan részletes mutagenézis kísérletek nem érhetőek el az FXIII A alegysége esetében, így ideális választásnak tűnt kihasználni a szöveti transzglutamináz esetében elérhető kísérletes adatokat és közvetlenül összevetni azokat a dinamikai szimulációink eredményeivel. Ahelyett, hogy a

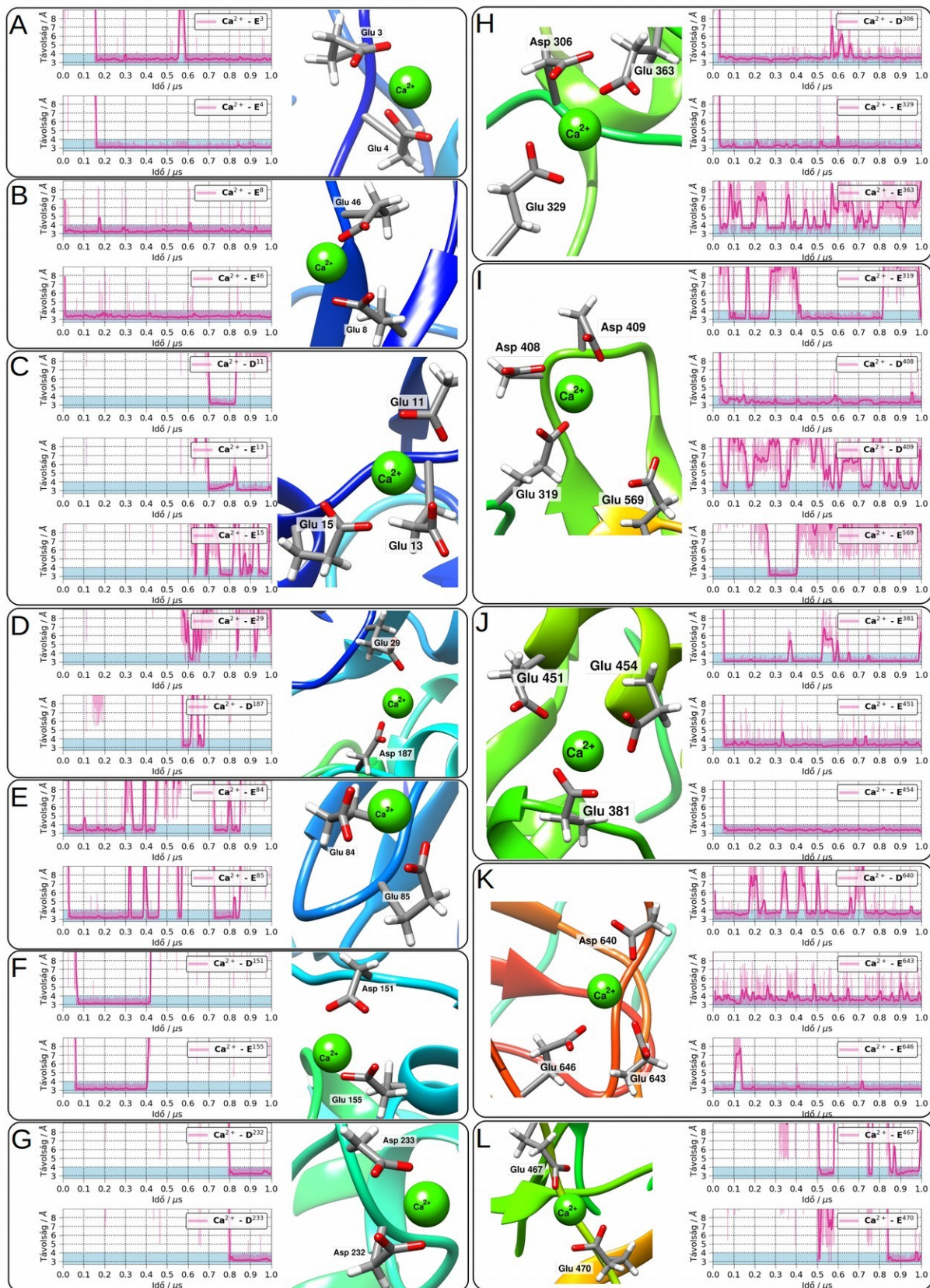
kalciumkötődés eseményeit ábrázoltuk volna, itt a megfelelő oldallánc karboxilát szene és a legközelebbi kalcium távolsága került ábrázolásra.

A kötött GDP-t tartalmazó szöveti transzglutamináz (*iv/d*) esetében kapott eredmények a 18. Ábrán láthatóak, a kötőhelyek proteinen való elhelyezkedése pedig a 19. Ábrán lettek feltüntetve. A statikus modellezés alapján végzett *in vitro* kísérletek eredményeit összevetve a molekuladinamikai szimulációk eredményeivel szembetűnő különbség, hogy legalább két olyan kötőhely található, melyeknek az aktivációban betöltött relevanciája kérdéses. Ezen kötőhelyek jelentősége mérséklődik a GTP kötött forma (*iv/e*) vizsgálatánál (K26. ábra, A-E panel), ugyanakkor teljesen elvetni nem lehet jelentőségüket, mivel a nyílt (aktív) TG2 esetében végzett szimulációban (*v/b*) is részben megfigyelhetőek (K27. ábra, A-C panel).

A kísérleti eredményekkel összhangban az S1 kötőhely a dinamikai szimulációk esetében is igazolva lett, kivételt képez ez alól a *iv/b* (GDP nélküli) 2  $\mu$ s -os szimuláció, melynek során nem sikerül kötődést detektálni. Ezen kötőhely az Asp232, Asp233 aminosavak hozzájárulásával alakul ki (18. ábra (G), K26. ábra (E), K27. ábra (F)) és megfelelnek az Asp270, Asp271 aminosavaknak az FXIII-A szekvenciájában, továbbá szintén ismert kötőhely a TG3 esetében.

Az S2A és S2B kötőhelyet illetően látható, hogy a Glu396 aminosav oldallánca a katalitikus centrum domén belseje felé orientálódik, így teljesen temetett. Ezen aminosav mellett az S2A kötőhely egy másik negatívan töltött aminosava az Asp400. Az *iv/a-d* szimulációk esetében az Asp400 C $\gamma$  és a Lys464 N $\zeta$  atomok közötti távolság  $5,55 \pm 1,25$  Å átlagosan, mely nem változik számottevően a dinamikai szimulációk folyamán. Összehasonlítva ezt az értéket a *ii/a-d* (FXIII-A<sub>2</sub>') esetekben a megfelelő aminosavak közötti távolság  $6,25 \pm 1,30$  Å, így a jelentősen rövidebb távolság felfogható úgy, mintha az Asp400-at a Lys464 blokkolva tartaná, így az már nem lesz alkalmas arra, hogy részt vegyen a kalcium kötésében.

Az S2A és S2B kötőhely helyett egy új zseb került azonosításra, mely magas betöltöttséggel rendelkezik. Ezen felületi zseb a Glu381, Glu451 és Glu454 aminosavak által alakul ki (18. ábra (J), K25. ábra (I), K27. ábra (E)) és betöltöttsége közel folyamatos. Egy másik – szintén korábban ismeretlen – kötőhely alakulhat ki a Glu381, Asp408 és Asp409 aminosavak hozzájárulásával (18. ábra (I), K25. ábra (H), K27. ábra (H)). Meg kell említeni, hogy az aktív forma szimulációjában az Asp581 is részt vesz ezen kötőhely kialakításában, valamint a  $\beta$ -hordó 1 és  $\beta$ -hordó 2 domének elmozdulása szintén szükséges, hiszen a zárt formában az Asp581  $\sim 30$  Å távolságra helyezkedik el a másik három aminosavtól.

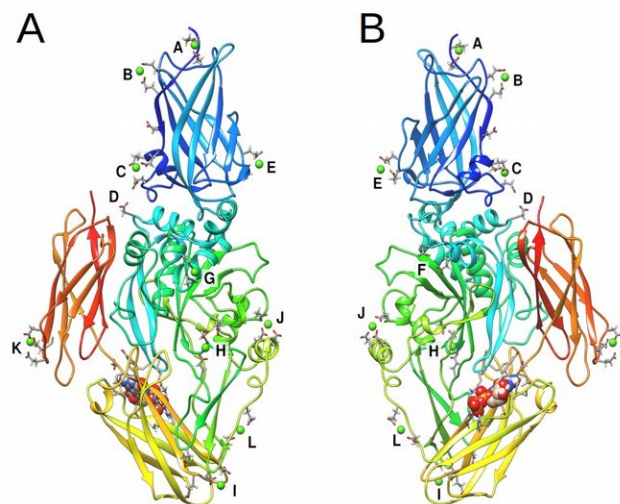


18. ábra. A TG2 iv/d szimulációjának kiértékelése után kapott kalciumkötőhelyek grafikus reprezentációja. (A karboxilát funkciós csoport szene és bármely kalcium közötti minimális távolságok kerültek ábrázolásra.)

A 460-as loop régió Glu467, Glu469, Glu470 aminosavai gyenge affinitást mutatnak, de a “multi-site” ionmodell alkalmazása mellett is közel található a kötött kalcium a Glu467 és Glu470-hez (18. ábra (L), K25. ábra (K), K26. ábra (G), K29. ábra).

Az S3A és S3B kötőzsebeket egy-egy negatív töltésű oldallánccal bíró aminosavakból által determinált és valóban mindkét aminosav részt vesz ugyanazon kalciumion kötésében, valamint néhány esetben a Glu363 is részt vesz az ion kötésében (18. ábra (H), K26. ábra (F), K27. ábra (G)). A negyedik kötőzseb (S4) az Asp151 és Glu158 közötti régió öt negatív töltésű oldallánccal rendelkező aminosavat tartalmaz és ez a kötőhely közel található az S5-höz. Szimulációink alapján az Asp151 és Glu155 hozzájárulásával valóban kialakulhat megfelelő kötőhely (18. ábra (F), K26. ábra (C), K25. ábra (F), K27. ábra (D)).

A már korábban feltételezett öt kötőhelyen túl egy olyan új kötőhely került azonosításra, mely az Asp640, Glu643 és Glu464 aminosavak segítségével alakul ki és minden szimuláció esetében megfigyelhető (18. ábra (K), K24. ábra (J), K25. ábra (J), K27. ábra (K)). Hasonlóan, ahogy azt az FXIII-A<sub>2</sub> esetben tettük, a *iv/d* szimulációt is folytattuk 50 ns szimuláció erejéig a “multi-site” kalcium modell alkalmazásával és a kapott eredmények (K29. ábra) azt mutatták, hogy a kialakult kötőhelyek túlnyomó többsége változatlan maradt.



19. ábra. A TG2 esetében kapott kalciumkötőhelyek elhelyezkedése a protein felületén, (A) előlnézetből, (B) hátnézetből. A kötőhelyek betűkódja a 16. ábrának megfelelően lett feltüntetve.

Eredményeink alapján, az alábbi általános következtetések vonhatóak le a TG2 kalcium kötő tulajdonságaival kapcsolatban, Király és munkatársai eredményeihez hasonlítva [68]. Az S1, S3A/S3B valamint az S4 kötőhelyeket illetően a dinamikai szimulációk alapján is megfelelőnek adódtak a kalciumok kötésére. Mindazonáltal az S2A/S2B kötőhelyet egyáltalán nem adódott olyan kötőhelynek, ahol megfigyelhattünk volna kalciumkötődést.

Összevetve ezen megfigyelést az FXIII-A<sub>2</sub> fő kötőhelyével kapcsolatban kapott információkkal, ahol a dimer esetében legalább az egyik lánc esetében megfigyelhető kalciumkötődés az 1 vagy 2  $\mu$ s-os szimulációs idő alatt látható, hogy a TG2 esetében is elérhető a kötőhely, mely némiképp tágabb is mint az FXIII-A<sub>2</sub> esetében. Mindazonáltal nem zárható ki, hogy a kötőhely jelentősége változik az aktiváció későbbi fázisában, de az elérhető krisztallográfiás szerkezetekből kiindulva egyetlen kötődési eseményt sem sikerült megfigyelünk. Végül, az S5 kötőhellyel kapcsolatban azt találtuk, hogy az aktiváció ezen korai fázisában, nem tulajdonítható jelentőség ezen kötőhelynek, valamint az aktív (nyílt) TG2 (*v/b*) esetében sem adódott potenciális kötőhelynek, így jelentősége megkérdőjelezhető.

### 6.2.3 Korrelált mozgások analízise

Általános konklúziót tudunk levonni az FXIII-A<sub>2</sub> esetében végzett szimulációk C $\alpha$  trajektóriái alapján végzett kereszt-korrelációk számításának eredményeivel kapcsolatban. A zimogén (*i*) és a hasított (*ii*) szimulációs csoportok osztoznak azon a közös sajátságon, hogy erős (pozitív) korreláció van az alegységek  $\beta$ -szendvics doménjei, valamint az 1-es  $\beta$ -hordók között is. Ezen korrelációk szemléletesebb értelmet nyernek, ha a pozitív korrelációt az egyes atomok azonos irányba történő, míg a negatív (anti-) korrelációt az atomok ellentétes irányba történő elmozdulásaként fogjuk fel. Az *i/a* és *i/b* valamint a *ii/a* és *ii/b* szimulációk során kapott korrelációk különbségét véve azt kapjuk, hogy ahogy azt már korábban említettük, a Met350-Trp370 és az Ile440-Ile460 régióknak kiemelt jelentősége lehet (K20. ábra). Függetlenül attól, hogy milyen úton történik az aktiváció, ezen régióknak mindkét mechanizmusnál szerepük kell hogy legyen. Ezen szegmensek korrelációja megnő az 1-es  $\beta$ -hordóval valamint a másik alegység ugyanezen régióival, míg a zimogén szerkezetek esetében ezen korrelációk kevésbé dominánsak (K20. ábra). Másik érdekes pont, a Gly210-Trp225 régiók közötti korreláció. Ezen régió az 1-es  $\beta$ -hordó C-terminálásával, az 1-es és 2-es  $\beta$ -hordót összekötő hurokkal és a 2-es  $\beta$ -hordó N-terminálásával szomszédos és a zimogén szerkezetekben szintén csökken jelentősége. A kalciumionok jelenléte a *ii/c* és *ii/d* szimulációk esetében kapott DCCM mátrixok összehasonlításából látható a legszemléletesebben, miszerint a kalciumok jelenléte egyértelműen növeli a  $\beta$ -szendvics domének közötti korrelációt (K21. ábra, A panel). Végül, minden említett korreláció megtalálható a harmadik szimulációs csoport esetében, azaz az aktivációs peptid jelenléte nélkül kivitelezett dinamikákban. Ugyanakkor a rendkívül nagy ionerősség, extrém módon megnövelte az alegységek  $\beta$ -szendvics doménjei közötti korrelációt és erős antikorrrelációt láthatunk a  $\beta$ -szendvics és a  $\beta$ -hordó domének között (K22. ábra). Ezek a felismerések

összhangban vannak az alegységek megfigyelt rotációival valamint a főkomponensanalízis eredményével.

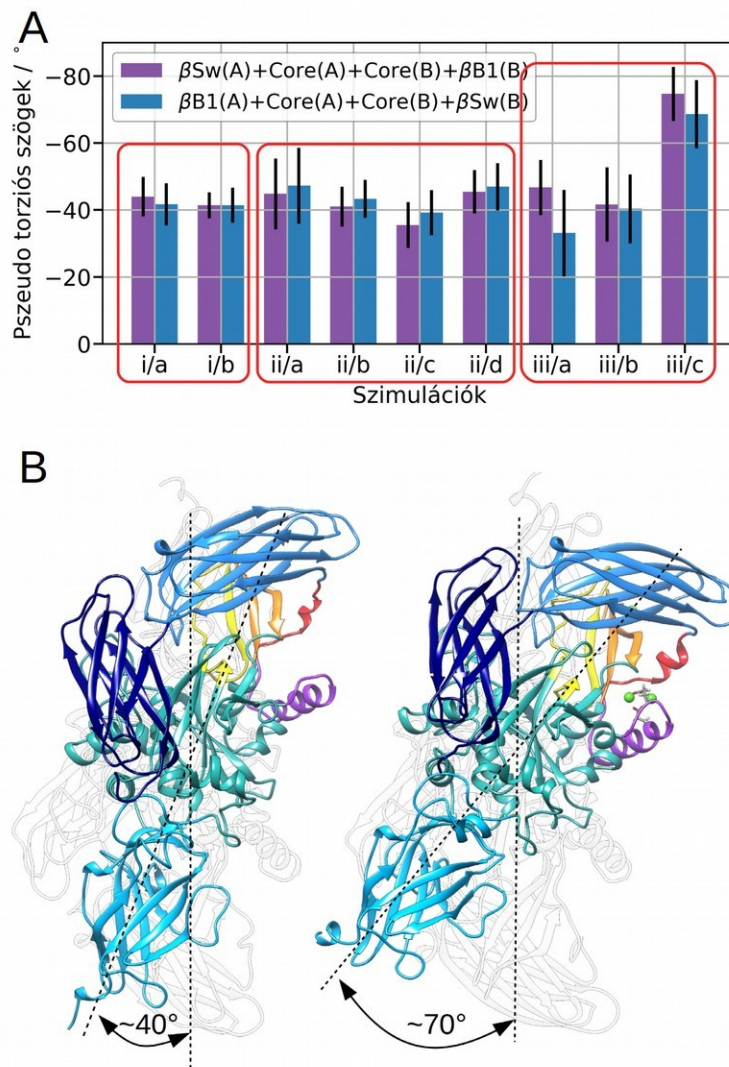
Annak ellenére, hogy az FXIII-A és a TG2 (és általánosságban a transzglutaminázok között) meglehetősen jól konzerválódtak a másodlagos és harmadlagos szerkezeti elemek, olyan különbségek figyelhetők meg a korrelációkban, melyek kizárólagos jellemzői egyiknek, s másiknak. Ilyen jellemző különbség, hogy a TG2 esetében az 1-es  $\beta$ -hordó korrelál a  $\beta$ -szendviccsel és a 2-es  $\beta$ -hordó korrelál a katalitikus doménnel és ez kizárólag a szöveti transzglutamináz esetében jellemző. Fontos kiemelni, hogy az Ile313-Asn333 és a Val401-Val422 között található három anitparallel  $\beta$ -redő (mely rendre megfelel a Met350-Trp370 és Ile440-Ile460 szegmensnek az FXIII-A<sub>2</sub> szerkezetében) korrelálnak az 1-es  $\beta$ -hordóval és a központi doménnel is, míg az FXIII-A<sub>2</sub> esetében nem figyelhető meg ilyen korreláció.

A nyílt (aktívnek vélt) szerkezetek szimulációiból (*v/a,b*) nyert korrelációs mátrixok tisztán látható különbségeket sugallnak. Nevezetesen, az *v/b* (TG2, 2Q3Z) esetében a 2-es hordó domén korrelál a  $\beta$ -szendvics doménnel, melyek térben meglehetősen távol helyezkednek el egymástól és erős antikorreláció fedezhető fel a katalitikus központi és az 1-es  $\beta$ -hordó domén között (K33. ábra, B panel). Az *v/a* (FXIII-A, 4KTY) esetében pontosan az ellenkezője érvényes az előbbi megállapításnak (K33. ábra, A panel). Jelenlegi tudásunk birtokában azt lehet feltételezni, hogy ezen enzimek vagy két, részleteiben teljesen eltérő aktivációs útvonalat követnek vagy az aktívnek vélt szerkezet valamelyike kinetikusan csapdázva lett egy olyan átmeneti termékként, mely az aktivációs útvonal valamely mozzanatában megrekedt szerkezetet reprezentálja.

#### 6.2.4 Főkomponens analízis és a központi domének közötti sóhíd hálózat szerepe

A nagy amplitúdójú (esszenciális) mozgások valós szimulációs időben történő megfigyelése, általában még hosszabb dinamikai szimulációk során is rejtve marad, így úgy gondoltuk hasznos lehet megvizsgálni ezen mozgásokat leíró sajátvektorokat. A dinamikai trajektóriákban megjelenő doménmozgásokat lehetséges leírni egy kvázi-torziós szöggel, melyet a megfelelő domének tömegközéppontjainak segítségével tudunk definiálni (20. ábra). Az általunk kivitelezett szimulációkban a 20. ábrán látható értékek nagyjából  $-40^\circ$  -ot mutatnak, míg a *iii/c* esetben ez az érték  $-70^\circ$ , amely azt jelenti hogy az alegységek  $30^\circ$ -ot mozdulnak el egymáshoz képest. Érdekes módon, ez az elmozdulás pontosan megegyezik azzal, melyet a főkomponens analízis során kaptunk és amelyet az első sajátvektor reprezentál a *iii/a-c* szimulációk esetében (21. ábra, A panel). Fontos kihangsúlyozni, hogy ezen főkomponensvektor az aktivációs peptid jelenlétében nehezebben vagy egyáltalán nem

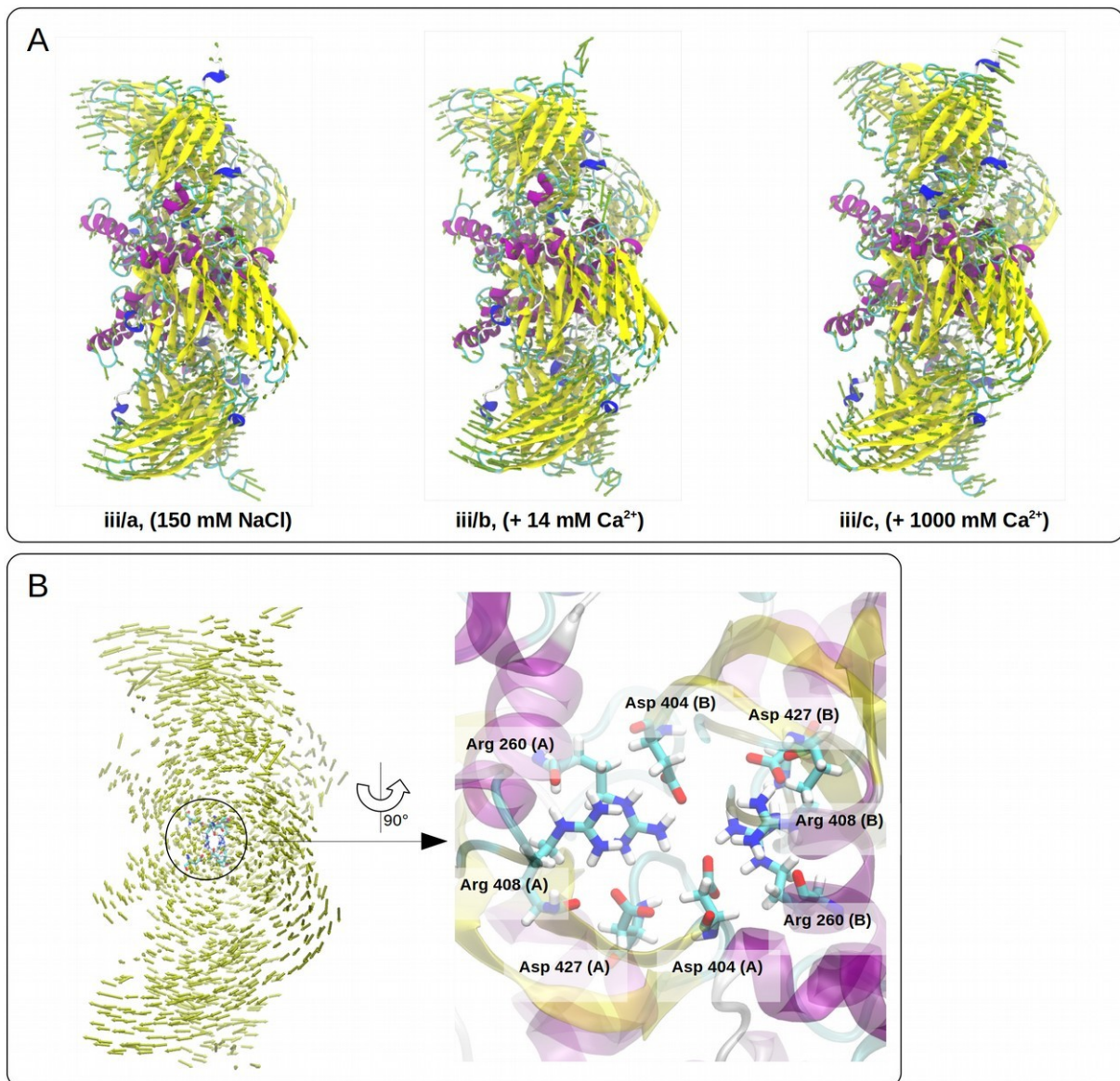
figyelhető meg. Összhangban azzal, hogy a *iii/a* esetben ( $\text{Ca}^{2+}$  nélkül) is megfigyelhető ezen sajátvektor, jól egyezik a kísérleti megfigyelés, miszerint nagy ionerősség jelenlétében kalcium nélkül is lassú, progresszív növekedés tapasztalható a zimogén FXIII-A<sub>2</sub> magas ionkoncentrációval indukált aktivációja során. Mivel ezen mozgások az aktivációs peptid jelenlétében rejtve maradnak, feltételezhető, hogy az N-terminális peptid fő szerepe az, hogy az aktivációt késleltessék. Az alegységek elfordulásával kapcsolatban elmondható, hogy az első sajátvektorok által leírt mozgás középpontjába rajzolható tengely pontosan áthalad egy olyan sóhíd hálózaton, mely az alegységek ugyanazon aminosavainak hozzájárulásával található meg a dimer szerkezetekben (21. ábra, B panel).



20. ábra. (A), A dinamikai trajektóriából számolt pszeudo-torziós szögek az FXIII-A<sub>2</sub> szimulációkban. (Az oszlopok tetején lévő fekete függőleges vonalak jelülik a szórást.) és (B), a *iii/c* szimuláció első (bal oldal) és utolsó (jobb oldal) képkockája.

Ezen aminosavak az Asp404, Asp427 és az Arg260, Arg408. Ez a központi sóhíd hálózat részleteiben még nem volt vizsgálva, ugyanakkor még a iii/c esetében is csak csekély mértékű változásokat figyelhetünk meg (K16. ábra).

Érdekességképpen megemlítenénk, hogy a sóhíd hálózat Arg260 tagjának mutációja, komoly deficienciával jár és számításos módszerekkel tanulmányozták a mutációk során kapott kölcsönhatási energiák változását [129,130]. A számításos módszerekkel készített Arg260Cys mutáns esetében általunk kivitelezett 1  $\mu$ s-os szimulációk alatt nem volt megfigyelhető változás a homodimer szerkezetében, annak ellenére, hogy a Cys260 nyilvánvalóan képtelen natív kapcsolatokat kialakítani (K17. ábra).



21. ábra. (A), a C-alfa trajektória főkomponens analízisének során kapott első sajátvektor grafikus reprezentációja, valamint (B) az alegységek közötti, központi sóhíd hálózat és annak elhelyezkedése.

Lehetséges magyarázat lehet ezen mutáció által okozott súlyos működésbeli zavarokra, hogy a fehérje feltekeredésében következnek be változások, melyek után az intakt dimer szerkezet már nem tud megfelelően összeszerelődni. Ezt a hipotézist támogatja az a gondolat, hogy az Arg408 aminosav a TG2 szerkezetében a Cys370 aminosavnak felel meg, melynek jelentősége a Cys371 aminosavval alkotott diszulfid-kapcsolóként ismert. Feltételezhető tehát, hogy amennyiben ez a mutáció az oka annak, hogy a szöveti transzglutamináz monomerként működik, úgy vélhetően ugyanez az oka annak, hogy az Arg260Cys mutáció a felelős azért, hogy a mutáns FXIII-A képtelen a dimer szerkezet kialakítására.

## 7. Konklúzió

### 7.1 A papain enzimmechanizmusa

A papain proteolitikus reakcióját vizsgáltuk ONIOM típusú QM/MM módszerekkel, melyben az egyszerű N-metilacetamid szubsztátot alkalmaztuk. Az ONIOM QM/MM számításainkban, QM komponensként a hibrid GGA (B3LYP) módszer mellett, olyan újabb DFT módszereket is alkalmaztunk, melyekben a hosszútávú korrekciós tag vagy azt megfelelően leíró paraméterek már jelen voltak a kölcsönhatás leírására.

Számításaink megmutatták, hogy a papain nyugvó állapotában a semleges és az ionpár formában jelen lévő cisztein-hisztidin diád energiája igen közel esik egymáshoz és csupán csekély aktiválási gát választja el őket. A zérus ponti energia korrekciót figyelembe véve, az egyensúly enyhén eltolódik a semleges forma irányába, míg a Poisson-Boltzman (PB) számítások alapján becsült, implicit oldószer modell korrekciót figyelembe véve az ionpárként jelen levő diád tűnik a leginkább betöltött populációnak, mely jó egyezést mutat a kísérleti adatokkal.

Az enzimmechanizmussal kapcsolatban elmondható, hogy a diszperziós korrekciót tartalmazó DFT módszerek, valamint a B3LYP módszerek a nagyobb (6-31+G(d,p) - 6-311++G(d,p)) báziskészletek alkalmazása mellett két elemi lépés jósolnak az acilezési lépésre, míg egy lépésben történik az acil-enzim hidrolízise. Számításaink alapján az acilezési lépésben az ikerionos tetraédes intermedier létezik. A származtatott aktivációs energiák jól egyeznek az irodalomban található kísérleti úton meghatározott kinetikai konstansokból származtatottakkal. Habár az alkalmazott sűrűségfüggő módszerek asszinkron sebességmeghatározó első lépést jósoltak az acilezési folyamatra, ki kell emelnünk, hogy a diszperziós korrekciót tartalmazó DFT módszerek az ONIOM QM/MM számításainkban jelentősen eltérő átmeneti állapotot adtak a B3LYP (6-31+G(d,p) - 6-311++G(d,p)) módszerhez képest. Ugyanakkor a

protonátmenet lemarad (de legalábbis nem előzi meg) az S-C kötés kialakulását. A B3LYP/6-31G(d,p) módszer esetében három elemi lépést találtunk, mely a klasszikus oxianionos, majd az ikerionos intermedieren keresztül vezet az acil-enzim intermedier kialakulásához. Ezen két intermedier közötti aktiválási gát igen alacsony volt. Mivel a B3LYP módszer által az amid kötés felhasadására jóvolt gát nagyon alacsony volt, feltételezhetően a kémia környezettől és az alkalmazott módszer részleteitől függően az acilezési lépésre 2-3 elemi lépésben lejátszódó mechanizmus felírása lehetséges.

Eredményeink alapján a cisztein proteináz reakciók tanulmányozásában erősen ajánlott a diszperzió-korrigált DFT módszerek használata a QM/MM számítások, tiszta QM rétegében, mivel ezek nagyban befolyásolhatják a tanulmányozott reakciók kimenetelét, ellentétben a szerin-proteázok mechanizmusával.

## 7.2 Az FXIII-A<sub>2</sub> és a TG2 aktivációjának korai eseményei

Mikroszekundumos hosszúságú „all-atom” molekuladinamikai szimulációk kerültek kivitelezésre annak érdekében, hogy fényt derítsünk azon finom részletekre, melyek a tanulmányozott transzglutaminázokban nagy konformációs változásokat indukálnak és első sorban a kalciumkötődéshez köthetőek.

A FXIII-A<sub>2</sub> esetében kimutattuk, hogy a kalciumionok jelenléte egyértelműen befolyásolja az aktivációs peptidek mozgékonyágát. Javaslatot tettünk ezidáig keveset vagy egyáltalán nem vizsgált kalcium kötő helyekre, azok vélhető szerepére. Aktivációs peptidek hiányában az egyensúlyi szimulációk eredményeként a nagy kalcium koncentráció az A alegységek egymástól való elfordulását okozta, mely látható a C-alfa atomok dinamikus keresztkorrelációjából és a főkomponens analízis alapján prediktált első sajátvektor által jóvolt mozgásokból is. Az elfordulás középpontjában egy alegységek között sóhíd hálózat található, melyet részleteiben eddig nem tárgyaltak, így ezidáig jelentősége nem került hangsúlyozásra.

A szerkezetileg az FXIII-A-hoz nagyon hasonló, ám eltérő biológiai szereppel bíró szöveti transzglutamináz enzim esetében megvizsgáltuk a Király és munkatársai által végzett kalciumkötőhelyek relevanciáját, számításos eredményeink az esetek többségében jó egyezést mutattak a kísérleti előzményekkel, illetve egy eddig nem ismert, azonban nagy affinitással rendelkező kötőhely került azonosításra.

## 8. A jelölt saját eredményei, új megállapításai

- Az értekezésben szereplő és alapjául szolgáló publikációkban közölt számításos eredmények teljes mértékben a saját munkámat tükrözik.
- A papain enzimmechanizmusának vizsgálata során igazoltuk, hogy annak nyugvó állapotában a semleges és az ionpár formában jelen lévő cisztein-hisztidin diád között a protonátmenetnek nem szükséges mediátor vízmolekula.
- A protonátmenet a zérusponthoz közeli vibrációs energia korrekció alkalmazásával spontán átmenetnek adódott.
- A szolvatációs szabadentalpia korrekcióval az ionpár formát minden esetben stabilabbnak számoltuk.
- A papain acilezési lépése két lépéses folyamat, melyben az ikerionos intermedier létezik. Az intermedier képződéséhez vezető asszinkron átmeneti állapotban koncertikusan történik a protonátmenet és a nukleofil támadás.
- Az acil-enzim hidrolízise egy koncertikus átmeneti állapoton keresztül történik egy vízmolekula jelenlétében, mely termékként ecetsavat ad és visszaalakul az enzim nyugvó állapotának ionpár formájára.
- A humán transzglutaminázok dinamikáját befolyásolja a kalcium jelenléte.
- A FXIII-A<sub>2</sub> esetében a fő kötőhelyek betöltődése között nincs preferencia, de allosztérikus útvonal írható fel az aktivációs peptid Arg11 és Arg12 aminosavait rögzítő üreg irányába.
- A FXIII A alegységeinek Asp270 és Asp271 aminosavai közösen alakítanak ki egy kötőhelyet, mely minden esetben megfigyelhető és eddig a zárt forma esetében nem volt igazolva. Ezen túl számos gyengébb kötőhelyet azonosítottunk.
- Magas (1 M) kalcium koncentráció jelenlétében, az A alegységek 30°-kal elfordulnak, melyet igazolnak a dinamikus kereszt-korrelációk és a főkomponensanalízis is.
- A szöveti transzglutamináz (TG2) esetében a korábban azonosítottaként vélt kötőhelyek közül az S1, S3, S4 kötőhelyek valóban minden esetben jelen vannak.
- Az S2A/S2B (FXIII-A analóg) kötőhelyek és az S5 kötőhelynek esetében nem tulajdonítható szerep a tanulmányozott mikorszekundumos időskálán.
- Három új kötőhely került azonosításra a TG2 esetében, melyek közül egy a mind ez idáig ismeretlen S6 kötőhely.

## 9. Összefoglalás

A cisztein proteázok enzimmechanizmusát több ízben vizsgálták, mind kísérleti, mind pedig elméleti módszerekkel a papain példáján keresztül, azonban mind ez idáig nem került publikálásra olyan közlemény, melyben egyrészt a teljes enzimreakciót összefoglalóan tárgyalták volna, másrészt pedig a mechanizmust optimált, frekvenciaanalízissel igazolt stacionárius pontokkal jellemezték volna. Jelen munkában ONIOM-EE típusú hibrid számításokkal kívántuk vizsgálni a mechanizmus részleteit. A QM/MM számítások kvantumkémiai rétegiére a sűrűségfunkcionál módszereket alkalmaztuk és teljesítőkéességüket is összehasonlítottuk. Számításaink alapján a semleges és az ionpár (tiolát-imidazólium) formában jelen lévő katalitikus diád közötti átmenet spontán gáz fázisú szimuláció szerint, melyet a solvatáció az ionpár forma irányába tol el. Az acilezési lépést illetően koncertikus reakciót találtunk, melyben a cisztein tiolát kenének nukleofil támadása és az imidazólium formában jelen lévő hisztidin delta nitrogénjéről, a szubsztrát amid nitrogénjére történő protonátmenet erősen csatolt, de nem teljesen szinkronban történik. A peptidkötés hasadása csekély energiagáttal rendelkezik, külön részlépésnek tekintendő és csupán elméleti jelentőséggel bír. A kialakult acil-enzim intermedier hidrolízisének aktiválási gátja összemérhető az acilezési részlépés esetében szükséges energiamennyiséggel, ezen – szintén koncertikus – reakció eredményeként, megfelelő karbonsav termék és a regenerálódott aktív centrum marad vissza.

A papain-szerű katalitikus, központi doménnel rendelkező humán transzglutaminázok, kalciumionok által indukált konformációs aktivációját két fehérje példáján keresztül tárgyaltuk. Ugyan meglehetősen jól konzerválódott másod- és harmadlagos szerkezet fedezhető fel a véralvadási kaszkád XIII-as faktorának A alegysége (FXIII-A) és a szöveti transzglutamináz (TG2) között, azonban a szekvenciahomológiáról ugyanez már nem mondható el, mely talán már előre vetíti azon diverz biológiai funkciókat, melyek ezen enzimekhez társíthatóak.

Munkánk során első ízben tárgyaltuk mikroszekundumos molekuladinamika segítségével, ezen fehérjék aktivációjának azon korai fázisát, melyben a kalciumionok kötődése történik. Szimulációink azt mutatták, hogy az FXIII-A<sub>2</sub> esetében korábban feltételezett gyengébb kötőhelyek valóban léteznek, melyek közül kiemelendő az alegységek ugyanazon aminosavai (Asp270, Asp271) által formált kötőhely. A TG2 szöveti transzglutamináz vizsgálatánál kiemelt figyelmet fordítottunk a korábban mutagenézis kísérletekkel felderített kalciumkötőhelyek vizsgálatára. A már azonosítottak vélt zsebek vizsgálata mellett, egy új – a dinamikai szimulációk alapján - erősnek vélt kötőhelyet is azonosítottunk, mely feltehetően a mind ez idáig ismeretlen hatodik kötőhely.

## 10. Summary

The enzyme mechanism of cysteine proteases has been already investigated with both theoretical and experimental methods using papain as model system. However, by this time the full catalytic cycle was not studied in one place with the means of verified stationary points. In present work we used ONIOM-EE type hybrid calculations to explore the elementary step of the reaction mechanism and the effect of the choice of density functional (including B3LYP functional) and basis set for the QM layer had been investigated.

Based on our results, the transition between the neutral and the thiolate-imidazolium ion pair form of the catalytic diad is with in gas phase simulation and the solvation (treated by Poisson-Boltzmann solvation correction) pushes the equilibrium toward the ion pair form. We have found that the first acylation step of the studied mechanism is a concerted reaction in which the attack of thiolate nucleophile and the proton transfer from the imidazolium ring to the substrates amide nitrogen is strongly coupled, but not fully synchronous process (Proton coupled electron transfer, PCET). The cleavage of peptide bond possesses only a small energy barrier and its importance is only theoretical. The activation barrier regarding the hydrolysis of acyl-enzyme intermediate is comparable to the energy of the acylation step. This final step was found to be also concerted process which yield the corresponding carboxylic acid product and the regenerated active centre.

The activation of human transglutaminases with papain-like catalytic core domain are clearly in connection with the presence of calcium ions, we presented the study of two human transglutaminases (FXIII-A<sub>2</sub> and TG2) and the effects of the presence of calcium ions were deeply investigated. These two enzymes were the coagulation FXIII-A<sub>2</sub> and the tissue transglutaminase (TG2). Both enzymes share highly conserved secondary and tertiary structures despite the relatively low sequence homology, the latter fact may predict the diverse biological functions of these enzymes.

During our extensive, microsecond timescale, *in silico* molecular dynamics simulations, the calcium binding properties of these transglutaminases were carefully examined. Based on our results at least two new different binding sites had been identified beside the known main site of A subunits in the case of FXIII-A<sub>2</sub>. The importance of the binding site formed by the Asp270 and Asp271 residues of both main chains is highlighted. At the study of TG2 tissue transglutaminase our computational results had been compared to the existing *in vitro* experiments which based on site-directed mutagenesis. The most important finding was a new, strong, sixth calcium binding site of TG2.

## 11. Irodalomjegyzék

- [1] Drenth J, Jansonius JN, Koekoek R, Swen HM, Wolthers BG. Structure of Papain. *Nature* 1968;218:929–32. doi:10.1038/218929a0.
- [2] Lowe G. The Structure and Mechanism of Action of Papain. *Philosophical Transactions of the Royal Society B: Biological Sciences* 1970;257:237–48. doi:10.1098/rstb.1970.0023.
- [3] Rawlings ND, Salvesen G, ProQuest (Firm). *Handbook of proteolytic enzymes Volume 1*. Oxford: Academic; 2012.
- [4] Polgár L, Halász P. Current problems in mechanistic studies of serine and cysteine proteinases. *Biochemical Journal* 1982;207:1–10. doi:10.1042/bj2070001.
- [5] Polgár L. The mechanism of action of thiolenzymes. *International Journal of Biochemistry* 1977;8:171–6. doi:10.1016/0020-711X(77)90143-4.
- [6] Polgar L. On the Mode of Activation of the Catalytically Essential Sulfhydryl Group of Papain. *European Journal of Biochemistry* 1973;33:104–9. doi:10.1111/j.1432-1033.1973.tb02660.x.
- [7] Polgár L. Mercaptide-imidazolium ion-pair: The reactive nucleophile in papain catalysis. *FEBS Letters* 1974;47:15–8. doi:10.1016/0014-5793(74)80415-1.
- [8] Sluyterman LA æ., Wijdenes J. An agarose mercurial column for the separation of mercaptopapain and nonmercaptopapain. *Biochimica et Biophysica Acta (BBA) - Protein Structure* 1970;200:593–5. doi:10.1016/0005-2795(70)90122-4.
- [9] Johnson FA, Lewis SD, Shafer JA. Determination of a low pK for histidine-159 in the S-methylthio derivative of papain by proton nuclear magnetic resonance spectroscopy. *Biochemistry* 1981;20:44–8. doi:10.1021/bi00504a008.
- [10] Menard R, Carriere J, Laflamme P, Plouffe C, Khouri HE, Vernet T, et al. Contribution of the glutamine 19 side chain to transition-state stabilization in the oxyanion hole of papain. *Biochemistry* 1991;30:8924–8. doi:10.1021/bi00101a002.
- [11] Menard R, Plouffe C, Laflamme P, Vernet T, Tessier DC, Thomas DY, et al. Modification of the Electrostatic Environment is Tolerated in the Oxyanion Hole of the Cysteine Protease Papain. *Biochemistry* 1995;34:464–71. doi:10.1021/bi00002a010.
- [12] van Duijnen PT, Thole BT, Broer R, Nieuwpoort WC. Active-site  $\beta$ -helix in papain and the stability of the ion pair  $RS^- - ImH^+$ . *Ab initio* molecular orbital study. *International Journal of Quantum Chemistry* 1980;17:651–71. doi:10.1002/qua.560170407.
- [13] Rullmann JAC, Bellido MN, van Duijnen PT. The active site of papain. *Journal of Molecular Biology* 1989;206:101–18. doi:10.1016/0022-2836(89)90527-5.

- [14] Arad D, Langridge R, Kollman PA. A simulation of the sulfur attack in catalytic pathway of papain using molecular mechanics and semiempirical quantum mechanics. *Journal of the American Chemical Society* 1990;112:491–502. doi:10.1021/ja00158a004.
- [15] Lin Y, Welsh WJ. Molecular modeling of substrate-enzyme reactions for the cysteine protease papain. *Journal of Molecular Graphics* 1996;14:62–72. doi:10.1016/0263-7855(96)00028-8.
- [16] Welsh WJ, Lin Y. Discussion of the catalytic pathway of cysteine proteases based on AM1 calculations. *Journal of Molecular Structure: THEOCHEM* 1997;401:315–26. doi:10.1016/S0166-1280(97)00025-0.
- [17] Harrison MJ, Burton NA, Hillier IH. Catalytic Mechanism of the Enzyme Papain: Predictions with a Hybrid Quantum Mechanical/Molecular Mechanical Potential. *Journal of the American Chemical Society* 1997;119:12285–91. doi:10.1021/ja9711472.
- [18] Han W-G, Tajkhorshid E, Suhai S. QM/MM Study of the Active Site of Free Papain and of the NMA-Papain Complex. *Journal of Biomolecular Structure and Dynamics* 1999;16:1019–32. doi:10.1080/07391102.1999.10508311.
- [19] Štrajbl M, Florián J, Warshel A. Ab Initio Evaluation of the Free Energy Surfaces for the General Base/Acid Catalyzed Thiolytic of Formamide and the Hydrolysis of Methyl Thiolfornate: A Reference Solution Reaction for Studies of Cysteine Proteases. *The Journal of Physical Chemistry B* 2001;105:4471–84. doi:10.1021/jp010279l.
- [20] Byun K, Gao J. A combined QM/MM study of the nucleophilic addition reaction of methanethiolate and N-methylacetamide. *Journal of Molecular Graphics and Modelling* 2000;18:50–5. doi:10.1016/S1093-3263(00)00035-8.
- [21] Miscione GP, Calvaresi M, Bottoni A. Computational Evidence for the Catalytic Mechanism of Caspase-7. A DFT Investigation. *The Journal of Physical Chemistry B* 2010;114:4637–45. doi:10.1021/jp908991z.
- [22] Ma S, Devi-Kesavan LS, Gao J. Molecular dynamics simulations of the catalytic pathway of a cysteine protease: a combined QM/MM study of human cathepsin K. *J Am Chem Soc* 2007;129:13633–45.
- [23] Wei D, Huang X, Liu J, Tang M, Zhan C-G. Reaction Pathway and Free Energy Profile for Papain-Catalyzed Hydrolysis of *N*-Acetyl-Phe-Gly 4-Nitroanilide. *Biochemistry* 2013;52:5145–54. doi:10.1021/bi400629r.

- [24] Shokhen M, Khazanov N, Albeck A. Challenging a paradigm: Theoretical calculations of the protonation state of the Cys25-His159 catalytic diad in free papain. *Proteins: Structure, Function, and Bioinformatics* 2009;77:916–26. doi:10.1002/prot.22516.
- [25] Mladenovic M, Fink RF, Thiel W, Schirmeister T, Engels B. On the Origin of the Stabilization of the Zwitterionic Resting State of Cysteine Proteases: A Theoretical Study. *Journal of the American Chemical Society* 2008;130:8696–705. doi:10.1021/ja711043x.
- [26] Ke Z, Zhou Y, Hu P, Wang S, Xie D, Zhang Y. Active Site Cysteine Is Protonated in the PAD4 Michaelis Complex: Evidence from Born–Oppenheimer *Ab Initio* QM/MM Molecular Dynamics Simulations. *The Journal of Physical Chemistry B* 2009;113:12750–8. doi:10.1021/jp903173c.
- [27] Lee W, Engels B. The Protonation State of Catalytic Residues in the Resting State of KasA Revisited: Detailed Mechanism for the Activation of KasA by Its Own Substrate. *Biochemistry* 2014;53:919–31. doi:10.1021/bi401308j.
- [28] Paasche A, Zipper A, Schäfer S, Ziebuhr J, Schirmeister T, Engels B. Evidence for Substrate Binding-Induced Zwitterion Formation in the Catalytic Cys-His Dyad of the SARS-CoV Main Protease. *Biochemistry* 2014;53:5930–46. doi:10.1021/bi400604t.
- [29] Lee W, Engels B. Clarification on the Decarboxylation Mechanism in KasA Based on the Protonation State of Key Residues in the Acyl-Enzyme State. *The Journal of Physical Chemistry B* 2013;117:8095–104. doi:10.1021/jp403067m.
- [30] Lee W, Luckner SR, Kisker C, Tonge PJ, Engels B. Elucidation of the Protonation States of the Catalytic Residues in *mt* KasA: Implications for Inhibitor Design. *Biochemistry* 2011;50:5743–56. doi:10.1021/bi200006t.
- [31] Grimme S. Density functional theory with London dispersion corrections: Density functional theory with London dispersion corrections. *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2011;1:211–28. doi:10.1002/wcms.30.
- [32] Marom N, Tkatchenko A, Rossi M, Gobre VV, Hod O, Scheffler M, et al. Dispersion Interactions with Density-Functional Theory: Benchmarking Semiempirical and Interatomic Pairwise Corrected Density Functionals. *Journal of Chemical Theory and Computation* 2011;7:3944–51. doi:10.1021/ct2005616.
- [33] Tsuneda T, Hirao K. Long-range correction for density functional theory: Long-range correction for DFT. *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2014;4:375–90. doi:10.1002/wcms.1178.
- [34] Roy D, Marianski M, Maitra NT, Dannenberg JJ. Comparison of some dispersion-corrected and traditional functionals with CCSD(T) and MP2 *ab initio* methods:

- Dispersion, induction, and basis set superposition error. *The Journal of Chemical Physics* 2012;137:134109. doi:10.1063/1.4755990.
- [35] Marianski M, Asensio A, Dannenberg JJ. Comparison of some dispersion-corrected and traditional functionals as applied to peptides and conformations of cyclohexane derivatives. *The Journal of Chemical Physics* 2012;137:044109. doi:10.1063/1.4737517.
- [36] Chai J-D, Head-Gordon M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Physical Chemistry Chemical Physics* 2008;10:6615. doi:10.1039/b810189b.
- [37] Grimme S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry* 2006;27:1787–99. doi:10.1002/jcc.20495.
- [38] Zhao Y, Truhlar DG. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theoretical Chemistry Accounts* 2008;120:215–41. doi:10.1007/s00214-007-0310-x.
- [39] Zhao Y, Truhlar DG. Density Functional Theory for Reaction Energies: Test of Meta and Hybrid Meta Functionals, Range-Separated Functionals, and Other High-Performance Functionals. *Journal of Chemical Theory and Computation* 2011;7:669–76. doi:10.1021/ct1006604.
- [40] Zhao Y, Truhlar DG. Applications and validations of the Minnesota density functionals. *Chemical Physics Letters* 2011;502:1–13. doi:10.1016/j.cplett.2010.11.060.
- [41] Rokob TA, Hamza A, Pápai I. Computing Reliable Energetics for Conjugate Addition Reactions. *Organic Letters* 2007;9:4279–82. doi:10.1021/ol701872z.
- [42] Walker M, Harvey AJA, Sen A, Dessent CEH. Performance of M06, M06-2X, and M06-HF Density Functionals for Conformationally Flexible Anionic Clusters: M06 Functionals Perform Better than B3LYP for a Model System with Dispersion and Ionic Hydrogen-Bonding Interactions. *The Journal of Physical Chemistry A* 2013;117:12590–600. doi:10.1021/jp408166m.
- [43] Lonsdale R, Harvey JN, Mulholland AJ. Effects of Dispersion in Density Functional Based Quantum Mechanical/Molecular Mechanical Calculations on Cytochrome P450 Catalyzed Reactions. *Journal of Chemical Theory and Computation* 2012;8:4637–45. doi:10.1021/ct300329h.

- [44] Tóth L, Muszbek L, Komáromi I. Mechanism of the irreversible inhibition of human cyclooxygenase-1 by aspirin as predicted by QM/MM calculations. *Journal of Molecular Graphics and Modelling* 2013;40:99–109. doi:10.1016/j.jmgs.2012.12.013.
- [45] Linder M, Brinck T. On the method-dependence of transition state asynchronicity in Diels–Alder reactions. *Physical Chemistry Chemical Physics* 2013;15:5108. doi:10.1039/c3cp44319a.
- [46] Lorand L, Konishi K. ACTIVATION OF THE FIBRIN STABILIZING FACTOR OF PLASMA BY THROMBIN. *Arch Biochem Biophys* 1964;105:58–67.
- [47] Takagi T, Doolittle RF. Amino acid sequence studies on factor XIII and the peptide released during its activation by thrombin. *Biochemistry* 1974;13:750–6. doi:10.1021/bi00701a018.
- [48] Yee VC, Pedersen LC, Bishop PD, Stenkamp RE, Teller DC. Structural evidence that the activation peptide is not released upon thrombin cleavage of factor XIII. *Thrombosis Research* 1995;78:389–97. doi:10.1016/0049-3848(95)00072-Y.
- [49] Pedersen LC, Yee VC, Bishop PD, Trong IL, Teller DC, Stenkamp RE. Transglutaminase factor XIII uses proteinase-like catalytic triad to crosslink macromolecules. *Protein Science* 1994;3:1131–5. doi:10.1002/pro.5560030720.
- [50] Credo RB, Curtis CG, Lorand L. Ca<sup>2+</sup>-related regulatory function of fibrinogen. *Proceedings of the National Academy of Sciences* 1978;75:4234–7. doi:10.1073/pnas.75.9.4234.
- [51] Polgár J, Hidasi V, Muszbek L. Non-proteolytic activation of cellular protransglutaminase (placenta macrophage factor XIII). *Biochem J* 1990;267:557–60.
- [52] Ahvazi B. Three-dimensional structure of the human transglutaminase 3 enzyme: binding of calcium ions changes structure for activation. *The EMBO Journal* 2002;21:2055–67. doi:10.1093/emboj/21.9.2055.
- [53] Ahvazi B, Boeshans KM, Idler W, Baxa U, Steinert PM. Roles of Calcium Ions in the Activation and Activity of the Transglutaminase 3 Enzyme. *Journal of Biological Chemistry* 2003;278:23834–41. doi:10.1074/jbc.M301162200.
- [54] Muszbek L, Bereczky Z, Bagoly Z, Komáromi I, Katona É. Factor XIII: A Coagulation Factor With Multiple Plasmatic and Cellular Functions. *Physiological Reviews* 2011;91:931–72. doi:10.1152/physrev.00016.2010.
- [55] Stieler M, Weber J, Hils M, Kolb P, Heine A, Büchold C, et al. Structure of Active Coagulation Factor XIII Triggered by Calcium Binding: Basis for the Design of Next-Generation Anticoagulants. *Angewandte Chemie International Edition* 2013;52:11930–4. doi:10.1002/anie.201305133.

- [56] Lewis BA, Freyssinet JM, Holbrook JJ. An equilibrium study of metal ion binding to human plasma coagulation factor XIII. *Biochem J* 1978;169:397–402.
- [57] Fox BA, Yee VC, Pedersen LC, Le Trong I, Bishop PD, Stenkamp RE, et al. Identification of the calcium binding site and a novel ytterbium site in blood coagulation factor XIII by x-ray crystallography. *J Biol Chem* 1999;274:4917–23.
- [58] Lai TS, Slaughter TF, Peoples KA, Greenberg CS. Site-directed mutagenesis of the calcium-binding site of blood coagulation factor XIIIa. *J Biol Chem* 1999;274:24953–8.
- [59] Ambrus A, Bányai I, Weiss MS, Hilgenfeld R, Keresztessy Z, Muszbek L, et al. Calcium Binding of Transglutaminases: A <sup>43</sup>Ca NMR Study Combined with Surface Polarity Analysis. *Journal of Biomolecular Structure and Dynamics* 2001;19:59–74. doi:10.1080/07391102.2001.10506720.
- [60] Turner BT, Maurer MC. Evaluating the Roles of Thrombin and Calcium in the Activation of Coagulation Factor XIII Using H/D Exchange and MALDI-TOF MS †. *Biochemistry* 2002;41:7947–54. doi:10.1021/bi025630n.
- [61] Gupta S, Biswas A, Akhter MS, Krettlner C, Reinhart C, Dodt J, et al. Revisiting the mechanism of coagulation factor XIII activation and regulation from a structure/functional perspective. *Scientific Reports* 2016;6. doi:10.1038/srep30105.
- [62] Lorand L. Factor XIII: structure, activation, and interactions with fibrinogen and fibrin. *Ann N Y Acad Sci* 2001;936:291–311.
- [63] Liu S, Cerione RA, Clardy J. Structural basis for the guanine nucleotide-binding activity of tissue transglutaminase and its regulation of transamidation activity. *Proceedings of the National Academy of Sciences* 2002;99:2743–7. doi:10.1073/pnas.042454899.
- [64] Jang T-H, Lee D-S, Choi K, Jeong EM, Kim I-G, Kim YW, et al. Crystal Structure of Transglutaminase 2 with GTP Complex and Amino Acid Sequence Evidence of Evolution of GTP Binding Site. *PLoS ONE* 2014;9:e107005. doi:10.1371/journal.pone.0107005.
- [65] Han B-G, Cho J-W, Cho YD, Jeong K-C, Kim S-Y, Lee BI. Crystal structure of human transglutaminase 2 in complex with adenosine triphosphate. *International Journal of Biological Macromolecules* 2010;47:190–5. doi:10.1016/j.ijbiomac.2010.04.023.
- [66] Pinkas DM, Strop P, Brunger AT, Khosla C. Transglutaminase 2 Undergoes a Large Conformational Change upon Activation. *PLoS Biology* 2007;5:e327. doi:10.1371/journal.pbio.0050327.

- [67] Bergamini CM. GTP modulates calcium binding and cation-induced conformational changes in erythrocyte transglutaminase. *FEBS Letters* 1988;239:255–8. doi:10.1016/0014-5793(88)80928-1.
- [68] Király R, Csósz É, Kurtán T, Antus S, Szigeti K, Simon-Vecsei Z, et al. Functional significance of five noncanonical Ca<sup>2+</sup>-binding sites of human transglutaminase 2 characterized by site-directed mutagenesis: Ca<sup>2+</sup>-binding sites of TG2. *FEBS Journal* 2009;276:7083–96. doi:10.1111/j.1742-4658.2009.07420.x.
- [69] Kanchan K, Ergülen E, Király R, Simon-Vecsei Z, Fuxreiter M, Fésüs L. Identification of a specific one amino acid change in recombinant human transglutaminase 2 that regulates its activity and calcium sensitivity. *Biochemical Journal* 2013;455:261–72. doi:10.1042/BJ20130696.
- [70] Kanchan K, Fuxreiter M, Fésüs L. Physiological, pathological, and structural implications of non-enzymatic protein–protein interactions of the multifunctional human transglutaminase 2. *Cellular and Molecular Life Sciences* 2015;72:3009–35. doi:10.1007/s00018-015-1909-z.
- [71] Szabo A, Ostlund NS. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Mineola, N.Y: Dover Publications; 1996.
- [72] Hohenberg P, Kohn W. Inhomogeneous Electron Gas. *Physical Review* 1964;136:B864–71. doi:10.1103/PhysRev.136.B864.
- [73] Koch W, Holthausen MC. *A chemist’s guide to density functional theory*. New York: Wiley-VCH; 2001.
- [74] Fiolhais C, Nogueira F. *A primer in density functional theory*. Berlin: Springer; 2003.
- [75] Becke AD. Density functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* 1993;98:5648–52. doi:10.1063/1.464913.
- [76] Becke AD. A new mixing of Hartree–Fock and local density functional theories. *The Journal of Chemical Physics* 1993;98:1372–7. doi:10.1063/1.464304.
- [77] Dapprich S, Komáromi I, Byun KS, Morokuma K, Frisch MJ. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *Journal of Molecular Structure: THEOCHEM* 1999;461–462:1–21. doi:10.1016/S0166-1280(98)00475-8.
- [78] Vreven T, Byun KS, Komáromi I, Dapprich S, Montgomery JA, Morokuma K, et al. Combining Quantum Mechanics Methods with Molecular Mechanics Methods in ONIOM. *Journal of Chemical Theory and Computation* 2006;2:815–26. doi:10.1021/ct050289g.

- [79] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. Gaussian 03 Revision C.02. n.d.
- [80] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. Gaussian 09 Revision B.01. n.d.
- [81] Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 2013;29:845–54. doi:10.1093/bioinformatics/btt055.
- [82] Pickersgill RW, Harris GW, Garman E. Structure of monoclinic papain at 1.60 Å resolution. *Acta Crystallographica Section B Structural Science* 1992;48:59–67. doi:10.1107/S0108768191006572.
- [83] Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* 2006;65:712–25. doi:10.1002/prot.21123.
- [84] Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* 1984;81:3684–90. doi:10.1063/1.448118.
- [85] Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* 2007;126:014101. doi:10.1063/1.2408420.
- [86] Darden T, York D, Pedersen L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *The Journal of Chemical Physics* 1993;98:10089–92. doi:10.1063/1.464397.
- [87] Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, Duke RE, et al. AMBER 12. University of California, San Francisco; 2012.
- [88] Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* 1996;14:33–8. doi:10.1016/0263-7855(96)00018-5.
- [89] Lee C, Yang W, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* 1988;37:785–9. doi:10.1103/PhysRevB.37.785.
- [90] Weiss MS, Metzner HJ, Hilgenfeld R. Two non-proline cis peptide bonds may be important for factor XIII function. *FEBS Lett* 1998;423:291–6.
- [91] Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* 1981;52:7182–90. doi:10.1063/1.328693.

- [92] Li D-W, Brüschweiler R. NMR-Based Protein Potentials. *Angewandte Chemie International Edition* 2010;49:6778–80. doi:10.1002/anie.201001898.
- [93] Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics* 2010:NA-NA. doi:10.1002/prot.22711.
- [94] Meagher KL, Redman LT, Carlson HA. Development of polyphosphate parameters for use with the AMBER force field. *Journal of Computational Chemistry* 2003;24:1016–25. doi:10.1002/jcc.10262.
- [95] Fizil Á, Sonderegger C, Czajlik A, Fekete A, Komáromi I, Hajdu D, et al. Calcium binding of the antifungal protein PAF: Structure, dynamics and function aspects by NMR and MD simulations. *PLOS ONE* 2018;13:e0204825. doi:10.1371/journal.pone.0204825.
- [96] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* 2015;11:3696–713. doi:10.1021/acs.jctc.5b00255.
- [97] Saxena A, Sept D. Multisite Ion Models That Improve Coordination and Free Energy Calculations in Molecular Dynamics Simulations. *J Chem Theory Comput* 2013;9:3538–42. doi:10.1021/ct400177g.
- [98] Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 2015;1–2:19–25. doi:10.1016/j.softx.2015.06.001.
- [99] Páll S, Abraham MJ, Kutzner C, Hess B, Lindahl E. Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. In: Markidis S, Laure E, editors. *Solving Software Challenges for Exascale*, vol. 8759, Cham: Springer International Publishing; 2015, p. 3–27. doi:10.1007/978-3-319-15976-8\_1.
- [100] Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, et al. The Amber biomolecular simulation programs. *J Comput Chem* 2005;26:1668–88. doi:10.1002/jcc.20290.
- [101] D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, et al. *AMBER 2016*. University of California, San Francisco: 2016.
- [102] Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. *WIREs Comput Mol Sci* 2012;3:198–210. doi:10.1002/wcms.1121.

- [103] Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* 2013;9:3084–95. doi:10.1021/ct400341p.
- [104] Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* 2006;22:2695–6. doi:10.1093/bioinformatics/btl461.
- [105] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
- [106] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera?A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 2004;25:1605–12. doi:10.1002/jcc.20084.
- [107] Šali A, Blundell TL. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology* 1993;234:779–815. doi:10.1006/jmbi.1993.1626.
- [108] Yang Z, Lasker K, Schneidman-Duhovny D, Webb B, Huang CC, Pettersen EF, et al. UCSF Chimera, MODELLER, and IMP: An integrated modeling system. *Journal of Structural Biology* 2012;179:269–78. doi:10.1016/j.jsb.2011.09.006.
- [109] Droettboom M, Caswell TA, Hunter J, Firing E, Nielsen JH, Lee A, et al. Matplotlib/Matplotlib V2.2.2 2018. doi:10.5281/zenodo.1202077.
- [110] van der Walt S, Colbert SC, Varoquaux G. The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science & Engineering* 2011;13:22–30. doi:10.1109/MCSE.2011.37.
- [111] Anandakrishnan R, Aguilar B, Onufriev AV. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Research* 2012;40:W537–41. doi:10.1093/nar/gks375.
- [112] Gordon JC, Myers JB, Folta T, Shoja V, Heath LS, Onufriev A. H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Research* 2005;33:W368–71. doi:10.1093/nar/gki464.
- [113] Myers J, Grothaus G, Narayanan S, Onufriev A. A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins: Structure, Function, and Bioinformatics* 2006;63:928–38. doi:10.1002/prot.20922.
- [114] Bakan A, Meireles LM, Bahar I. ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* 2011;27:1575–7. doi:10.1093/bioinformatics/btr168.

- [115] Bakan A, Dutta A, Mao W, Liu Y, Chennubhotla C, Lezon TR, et al. Evol and ProDy for bridging protein sequence evolution and structural dynamics. *Bioinformatics* 2014;30:2681–3. doi:10.1093/bioinformatics/btu336.
- [116] Lowe G, Whitworth AS. A kinetic and fluorimetric investigation of papain modified at tryptophan -69 and -177 by *N*-bromosuccinimide. *Biochemical Journal* 1974;141:503–15. doi:10.1042/bj1410503.
- [117] Sluyterman LAA, Wijdenes J. Proton Equilibria in the Binding of Zn<sup>2+</sup> and of Methylmercuric Iodide to Papain. *European Journal of Biochemistry* 1976;71:383–91. doi:10.1111/j.1432-1033.1976.tb11125.x.
- [118] Lewis SD, Johnson FA, Shafer JA. Effect of cysteine-25 on the ionization of histidine-159 in papain as determined by proton nuclear magnetic resonance spectroscopy. Evidence for a histidine-159-cysteine-25 ion pair and its possible role in catalysis. *Biochemistry* 1981;20:48–51. doi:10.1021/bi00504a009.
- [119] Lewis SD, Johnson FA, Shafer JA. Potentiometric determination of ionizations at the active site of papain. *Biochemistry* 1976;15:5009–17. doi:10.1021/bi00668a010.
- [120] Halasz P, Polgar L. Negatively Charged Reactants as Probes in the Study of the Essential Mercaptide-Imidazolium Ion-Pair of Thiolenzymes. *European Journal of Biochemistry* 1977;79:491–4. doi:10.1111/j.1432-1033.1977.tb11832.x.
- [121] Creighton DJ, Gessouroun MS, Heapes JM. Is the thiolate-imidazolium ion pair the catalytically important form of papain? *FEBS Letters* 1980;110:319–22. doi:10.1016/0014-5793(80)80101-3.
- [122] Creighton DJ, Schamp DJ. Solvent isotope effects on tautomerization equilibria of papain and model thiolamines. *FEBS Letters* 1980;110:313–8. doi:10.1016/0014-5793(80)80100-1.
- [123] Topf M, Várnai P, Richards WG. Ab Initio QM/MM Dynamics Simulation of the Tetrahedral Intermediate of Serine Proteases: Insights into the Active Site Hydrogen-Bonding Network. *Journal of the American Chemical Society* 2002;124:14780–8. doi:10.1021/ja026219q.
- [124] Lowe G, Yuthavong Y. Kinetic specificity in papain-catalysed hydrolyses. *Biochemical Journal* 1971;124:107–15. doi:10.1042/bj1240107.
- [125] Gauthier F, Moreau T, Lalmanach G, Brillardbourdet M, Ferrerdimartino M, Juliano L. A New, Sensitive Fluorogenic Substrate for Papain Based on the Sequence of the Cystatin Inhibitory Site. *Archives of Biochemistry and Biophysics* 1993;306:304–8. doi:10.1006/abbi.1993.1516.

- [126] Papamichael EM, Roustas MK, Bieth JG. Detection of S1-P1 and S3-P3 interactions between papain and four synthetic substrates. *Brazilian Archives of Biology and Technology* 1999;42:277–80. doi:10.1590/S1516-89131999000300002.
- [127] Alvarez-Idaboy JR, Galano A, Bravo-Pérez G, Ruiz ME. Rate Constant Dependence on the Size of Aldehydes in the  $\text{NO}_3 + \text{Aldehydes}$  Reaction. An Explanation via Quantum Chemical Calculations and CTST. *Journal of the American Chemical Society* 2001;123:8387–95. doi:10.1021/ja010693z.
- [128] Komáromi I, Bagoly Z, Muszbek L. Factor XIII: novel structural and functional aspects: FXIII: structure and function. *Journal of Thrombosis and Haemostasis* 2011;9:9–20. doi:10.1111/j.1538-7836.2010.04070.x.
- [129] Ichinose A, Tsukamoto H, Izumi T, Yamazaki T, Togashi M, Takamatsu J, et al. Arg260-Cys mutation in severe factor XIII deficiency: conformational change of the A subunit is predicted by molecular modelling and mechanics. *British Journal of Haematology* 1998;101:264–72. doi:10.1046/j.1365-2141.1998.00698.x.
- [130] Peyvandi F, Tagliabue L, Menegatti M, Karimi M, Komáromi I, Katona É, et al. Phenotype-genotype characterization of 10 families with severe a subunit factor XIII deficiency: MUTATIONS IN BRIEF. *Human Mutation* 2004;23:98–98. doi:10.1002/humu.9206.
- [131] Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. General Performance of Density Functionals. *The Journal of Physical Chemistry A* 2007, 111 (42), 10439–10452. <https://doi.org/10.1021/jp0734474>.
- [132] Fekete A., Vegyész MSc Diplomamunka, Debreceni Egyetem, Természettudományi Kar, 2012.

## 12. Tárgyszavak

### Tárgyszavak

cisztein proteázok  
hibrid (QM/MM) módszerek  
kalciumkötődés  
kvantumkémia  
molekuladinamika  
ONIOM módszer  
papain  
sűrűségfüggő elmélet  
számításos kémia  
szöveti transzglutamináz  
transzglutaminázok  
XIII-as véralvadási faktor

### Subject words

cysteine proteases  
hybrid (QM/MM) methods  
calcium binding  
quantum chemistry  
molecular dynamics  
ONIOM method  
papain  
density functional theory  
computational chemistry  
tissue transglutaminase  
transglutaminases  
blood coagulation FXIII

### **13. Köszönetnyilvánítás**

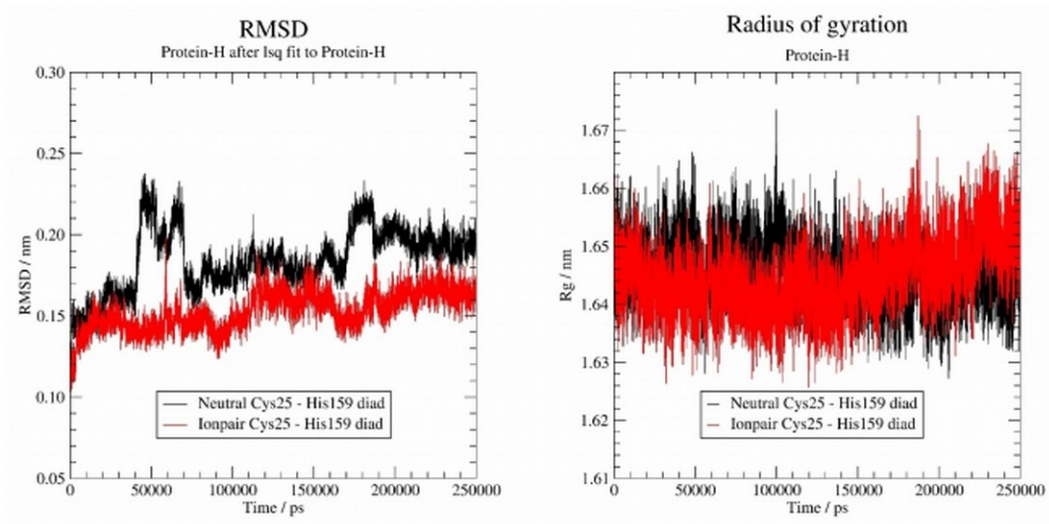
Köszönetet szeretnék mondani témavezetőmnek Dr. Komáromi Istvánnak a Debreceni Egyetem Általános Orvostudományi Kar, Laboratóriumi Medicina Intézet, Klinikai Laboratóriumi Kutató Tanszék tudományos főmunkatársának.

Köszönet illeti a Debreceni Egyetem Általános Orvostudományi Kar, Laboratóriumi Medicina Intézet, Klinikai Laboratóriumi Kutató Tanszék vezetőjét Dr. Bereczky Zsuzsanna egyetemi adjunktust a PhD tanulmányaimhoz szükséges feltételek biztosításáért.

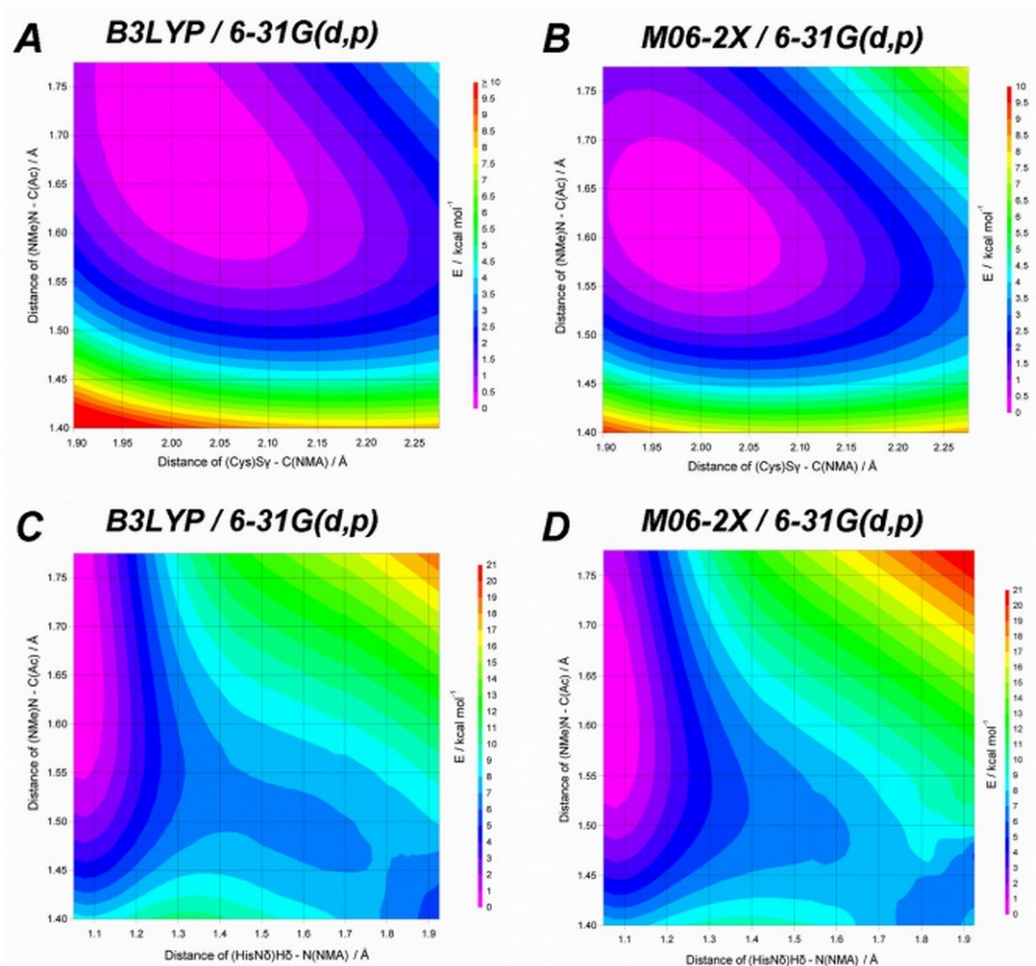
Köszönöm a Richter Gedeon Talentum Alapítványnak a PhD tanulmányaim támogatására folyósított ösztöndíjat.

## 14. Függelék

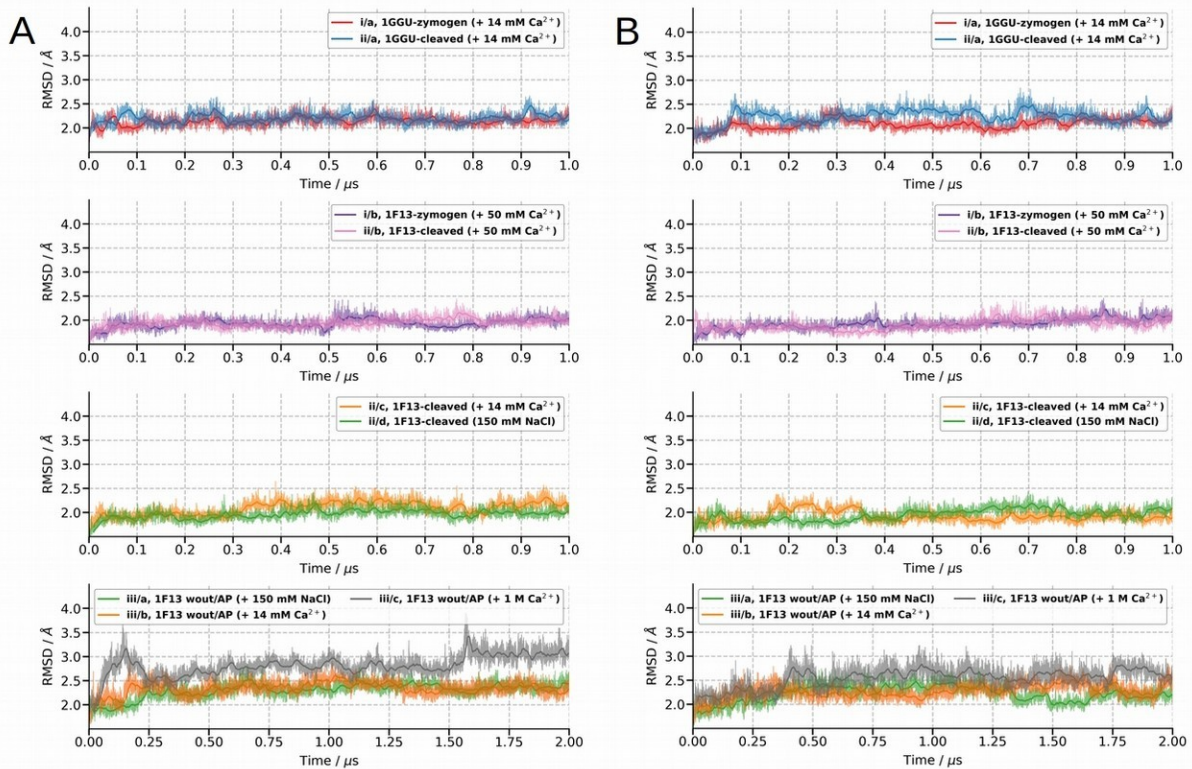
### 14.1 Kiegészítő ábrák



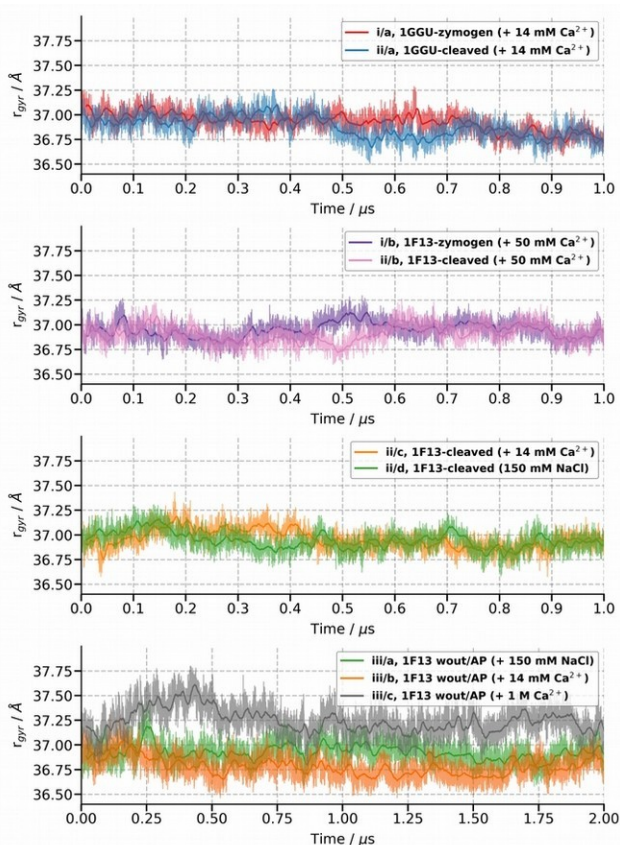
K1. ábra. A papain molekuladinamikai trajektóriájából számolt nehéz atom RMSD és girációs sugár értékek.



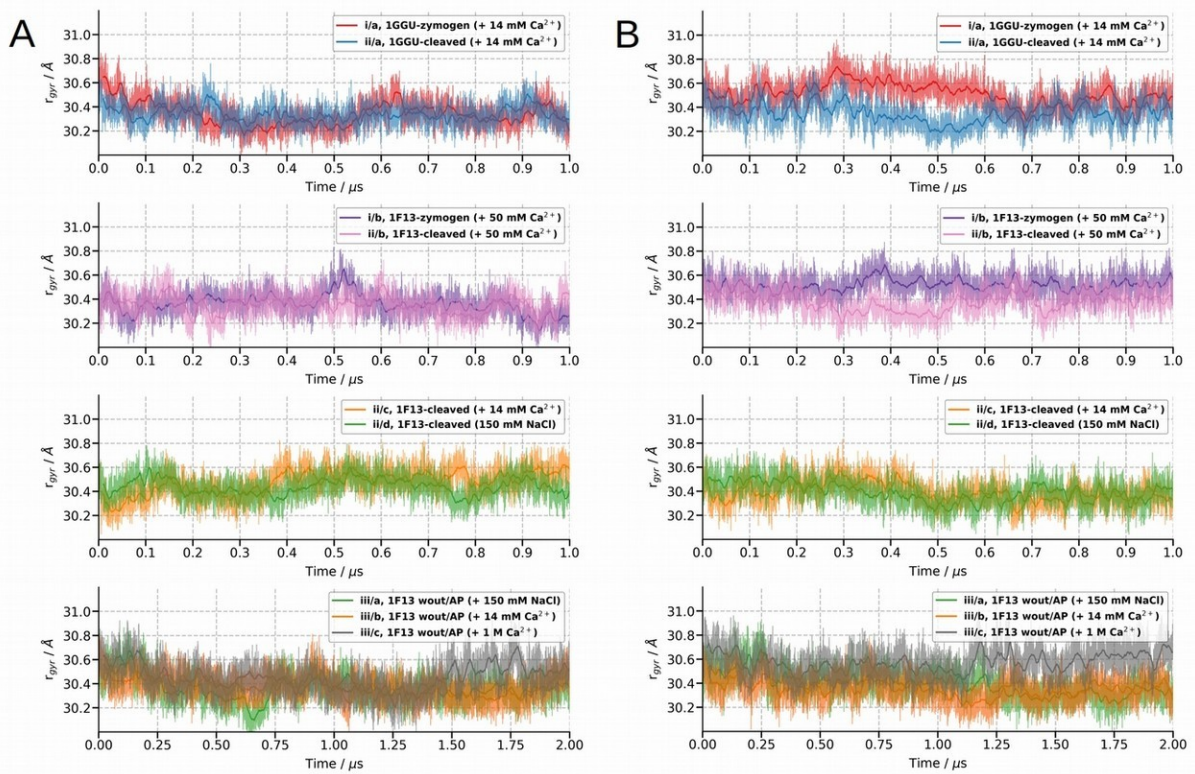
K2. ábra. ONIOM B3LYP/6-31G(d,p):AMBER (A és C panel) és M06-2X/6-31G(d,p):AMBER (B és D panel) módszerekkel számolt potenciális energiafelületek a papain acilezési lépésében a peptidkötés hasadására vonatkozóan ((NMA)N–C(NMA)).



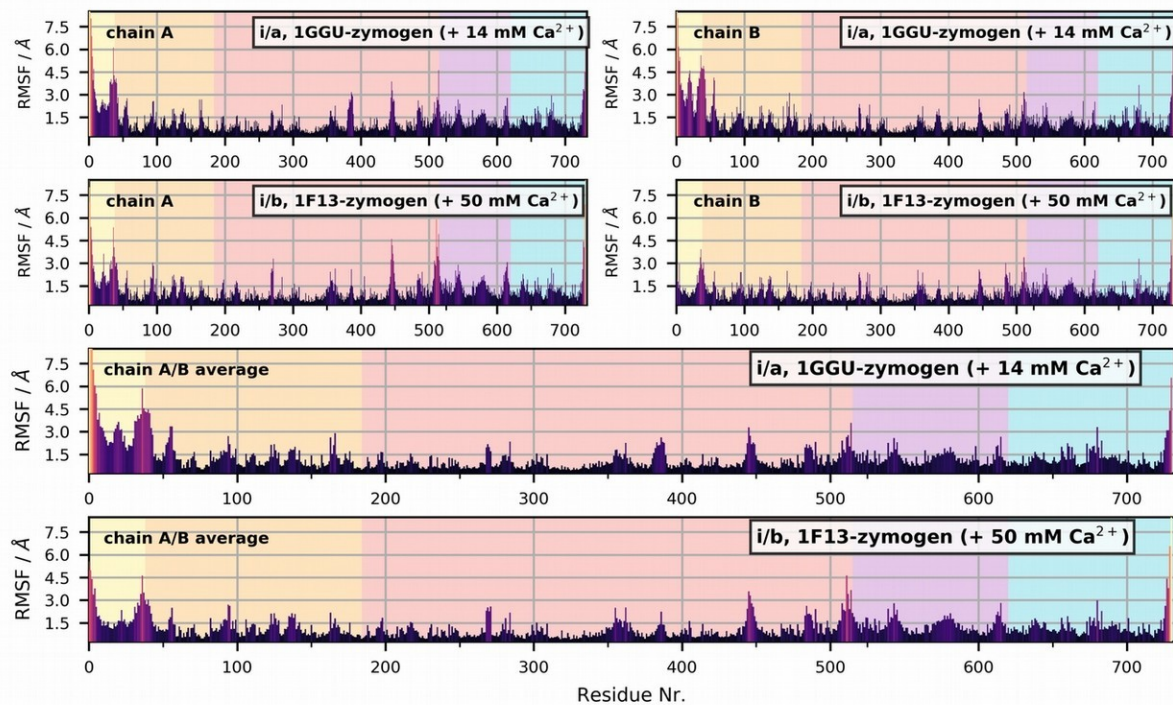
K3. ábra. A nehéz atom RMSD-k változása az FXIII-A<sub>2</sub> szimulációiban, lánconként. (A) „A” főlánc, (B) „B” főlánc, az aktivációs peptidet figyelmen kívül hagytuk az összehasonlíthatóság megtartásának érdekében.



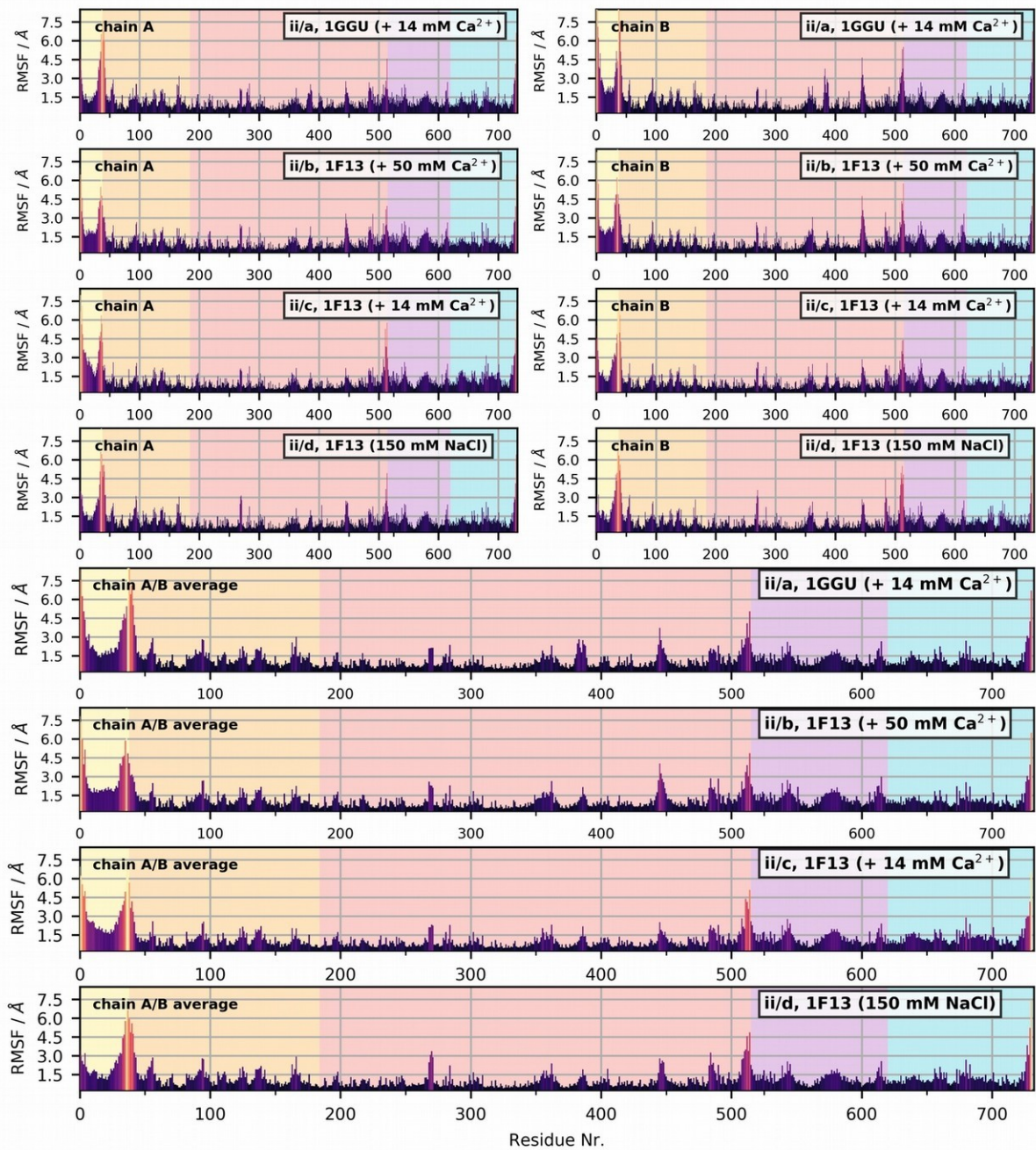
K4. ábra. Teljes (átlagolt) nehéz atom forgási sugarak az FXIII-A<sub>2</sub> szimulációinak esetében. Az aktivációs peptidet figyelmen kívül hagytuk az összehasonlíthatóság megtartásának érdekében.



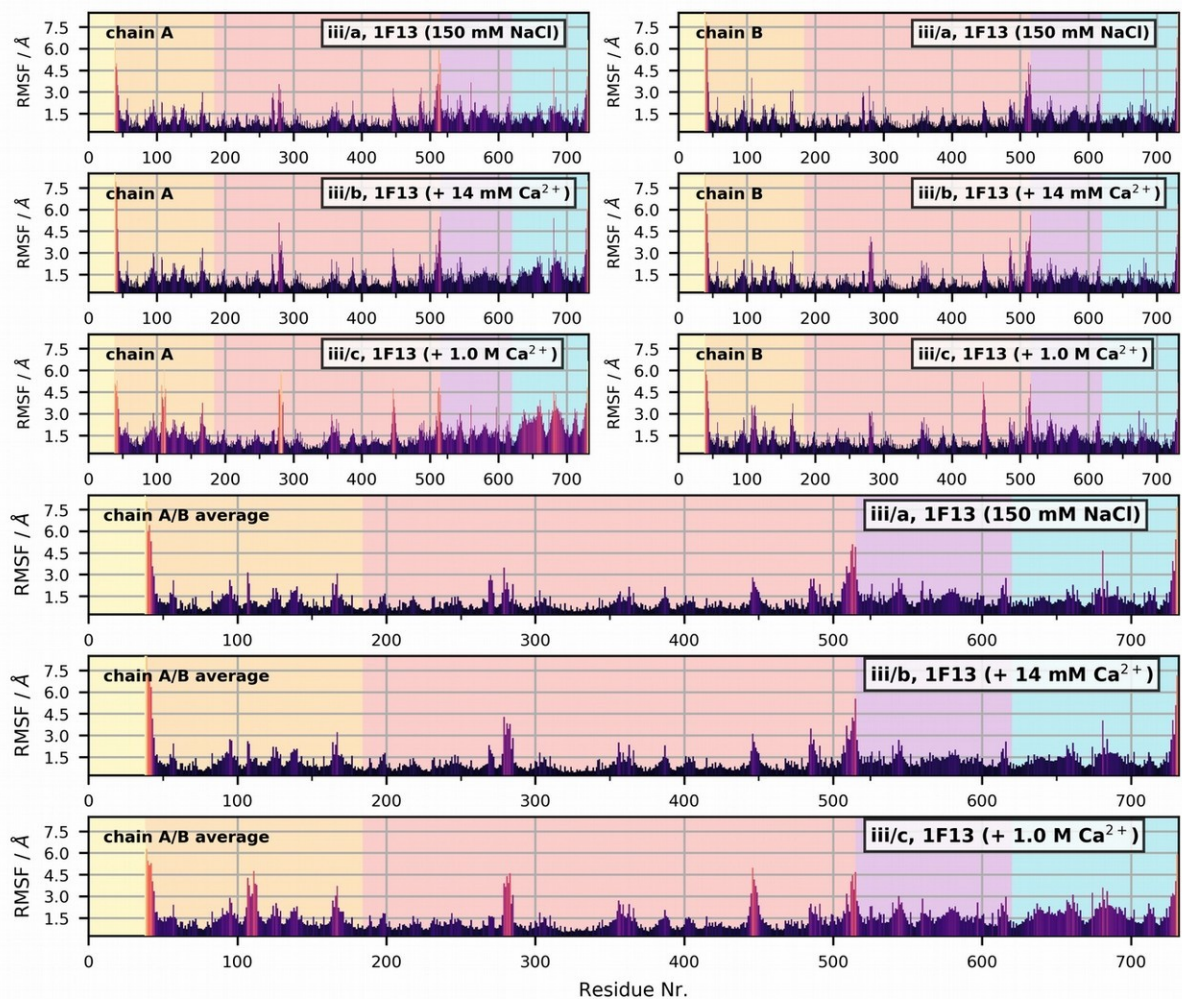
K5. ábra. A nehéz atom forgási sugarak változása az FXIII-A<sub>2</sub> szimulációiban, lánconként. (A) „A” főlánc, (B) „B” főlánc, az aktivációs peptidet figyelmen kívül hagytuk az összehasonlíthatóság megtartásának érdekében.



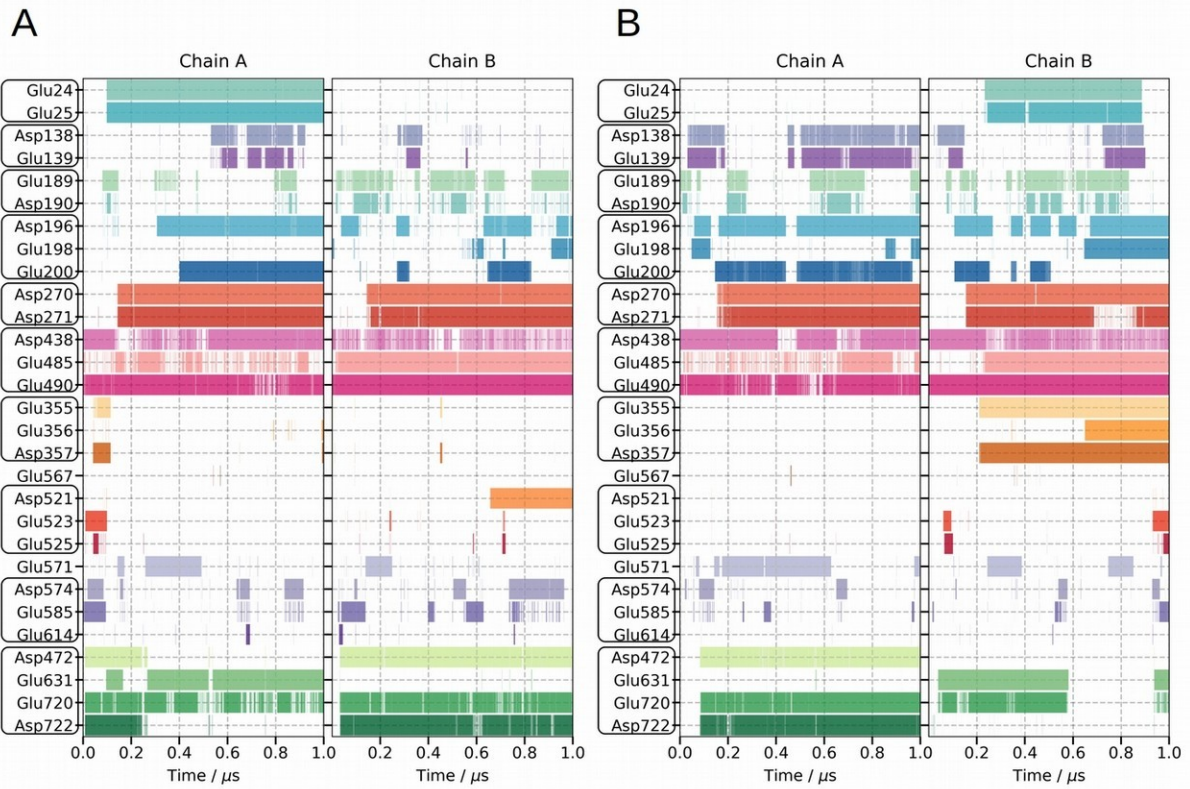
K6. ábra. Aminosavankénti fluktuációk a zimogén FXIII-A<sub>2</sub> (i/a,b) szimulációnak esetében.



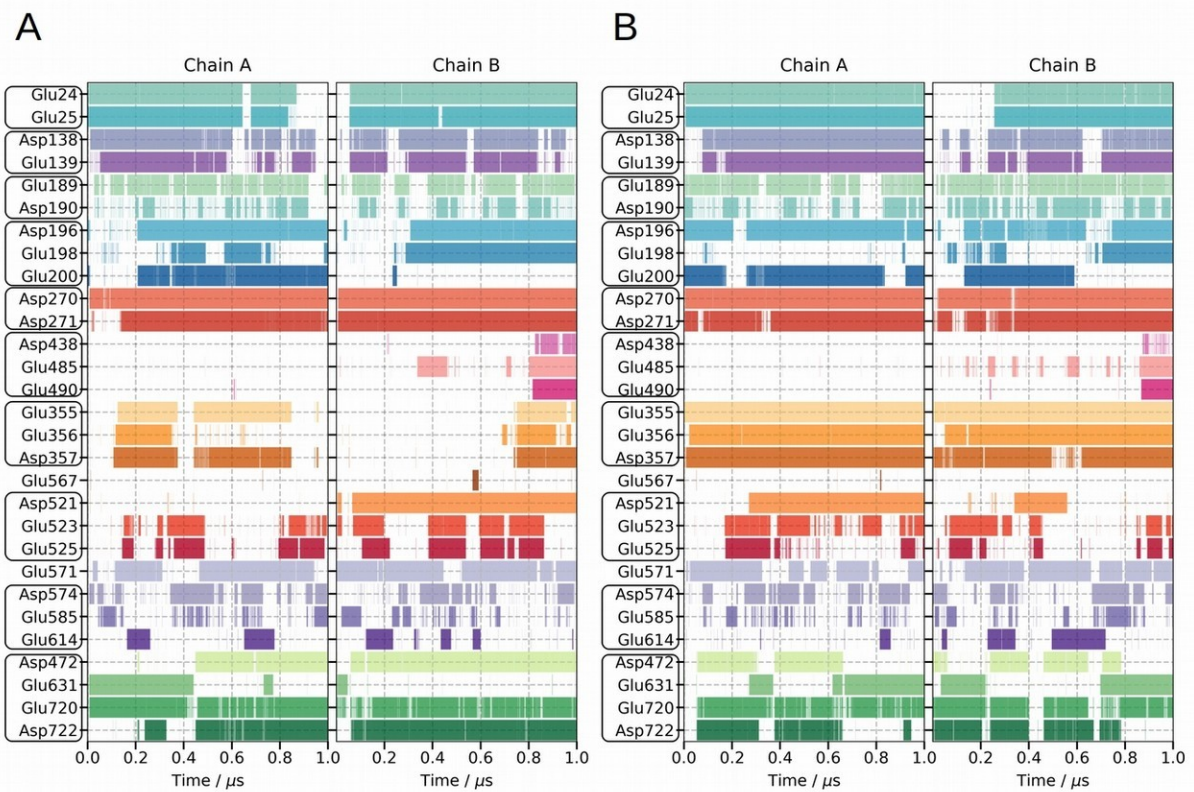
K7. ábra. Aminosavankénti fluktuációk a hasított aktivációs peptid jelenlétében történt FXIII-A<sub>2</sub> (ii/a-d) szimulációnak esetében.



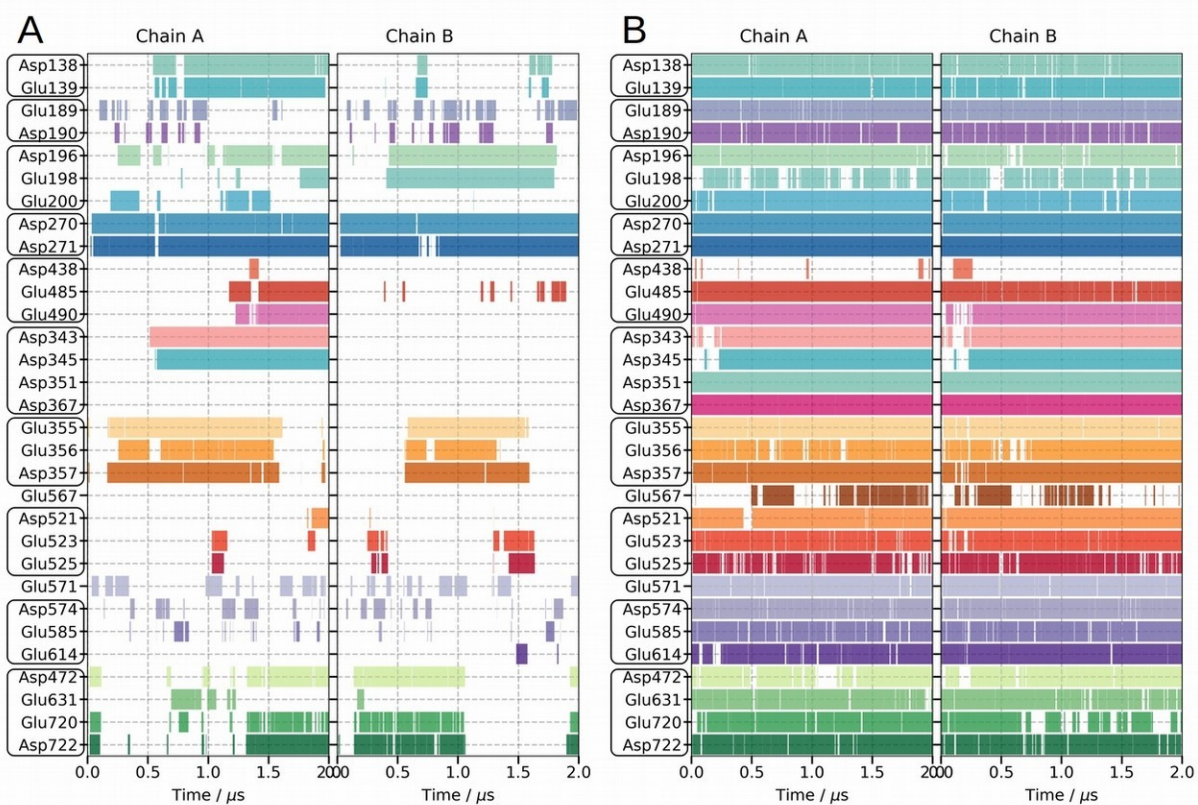
K8. ábra. Aminosavankénti fluktuációk az aktivációs peptidek nélkül végzett FXIII-A<sub>2</sub> (iii/a-c) szimulációnak esetében.



K9. ábra. A dinamikai trajektóriából számolt kalciumkötődési események grafikus reprezentációja az (A) i/a (1GGU, zimogén, +14 mM  $Ca^{2+}$ ) és a (B) ii/a (1GGU, hasított AP-k, +14 mM  $Ca^{2+}$ ) szimulációk esetében.



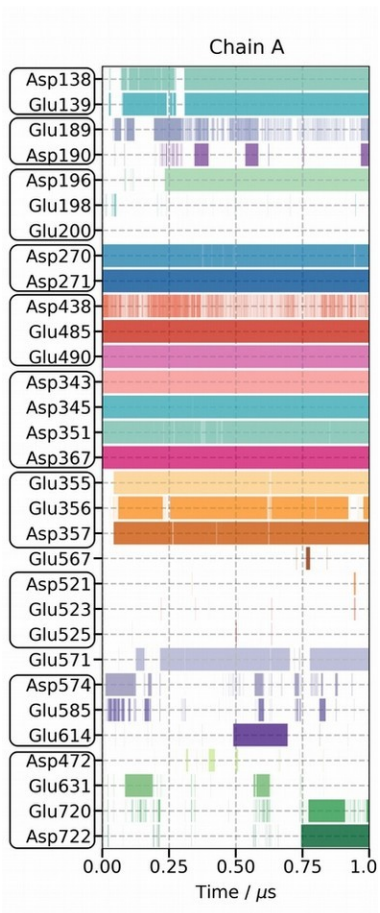
K10. ábra. A dinamikai trajektóriából számolt kalciumkötődési események grafikus reprezentációja az (A) i/b (1F13, zimogén, +50 mM  $Ca^{2+}$ ) és a (B) ii/b (1F13, hasított AP-k, +50 mM  $Ca^{2+}$ ) szimulációk esetében.



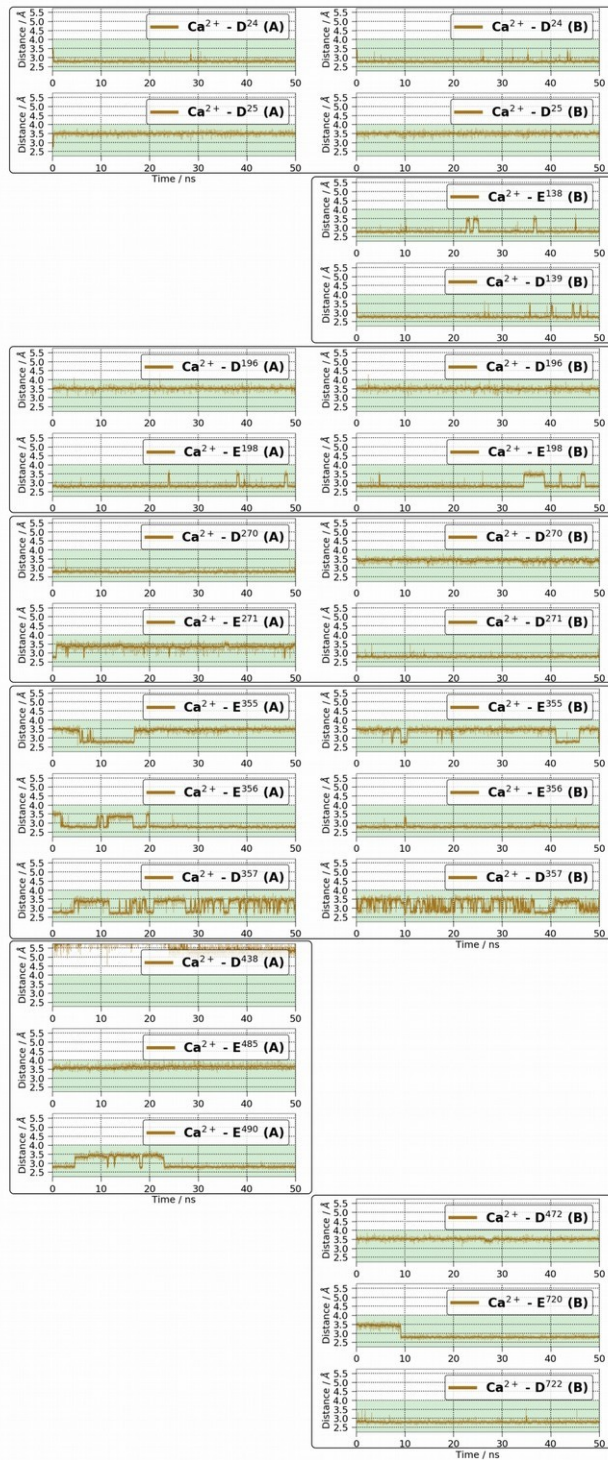
K11. ábra. A dinamikai trajektóriából számolt kalciumkötődési események grafikus reprezentációja az (A) iii/b (1F13, AP-k nélkül, +14 mM  $\text{Ca}^{2+}$ ) és a (B) iii/c (1F13, AP-k nélkül, +1000 mM  $\text{Ca}^{2+}$ ) szimulációk esetében.



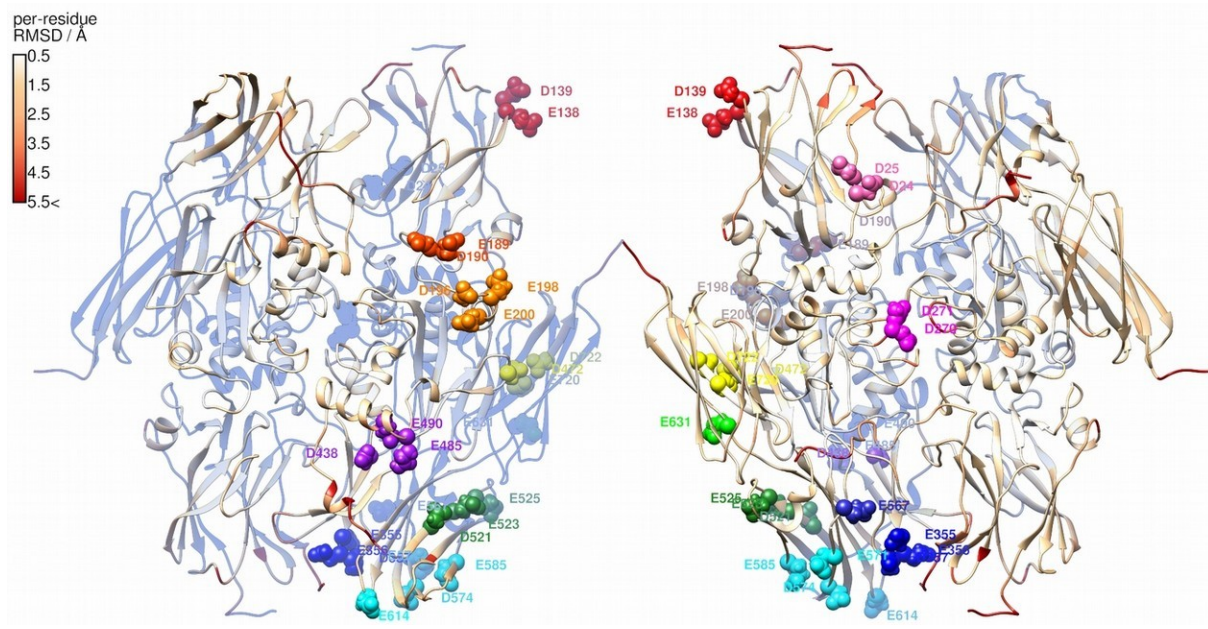
K12. ábra. Korábban kivitelezett dinamikai szimuláció trajektóriájából számolt kalciumkötődési események grafikus reprezentációja, a ii/c (1F13, hasított AP-k, +14 mM  $\text{Ca}^{2+}$ ) szimulációval azonos paraméterek mellett, azonban a titrálható aminosavak protonáltsági foka nem lett számítósos módszerekkel jósolva.



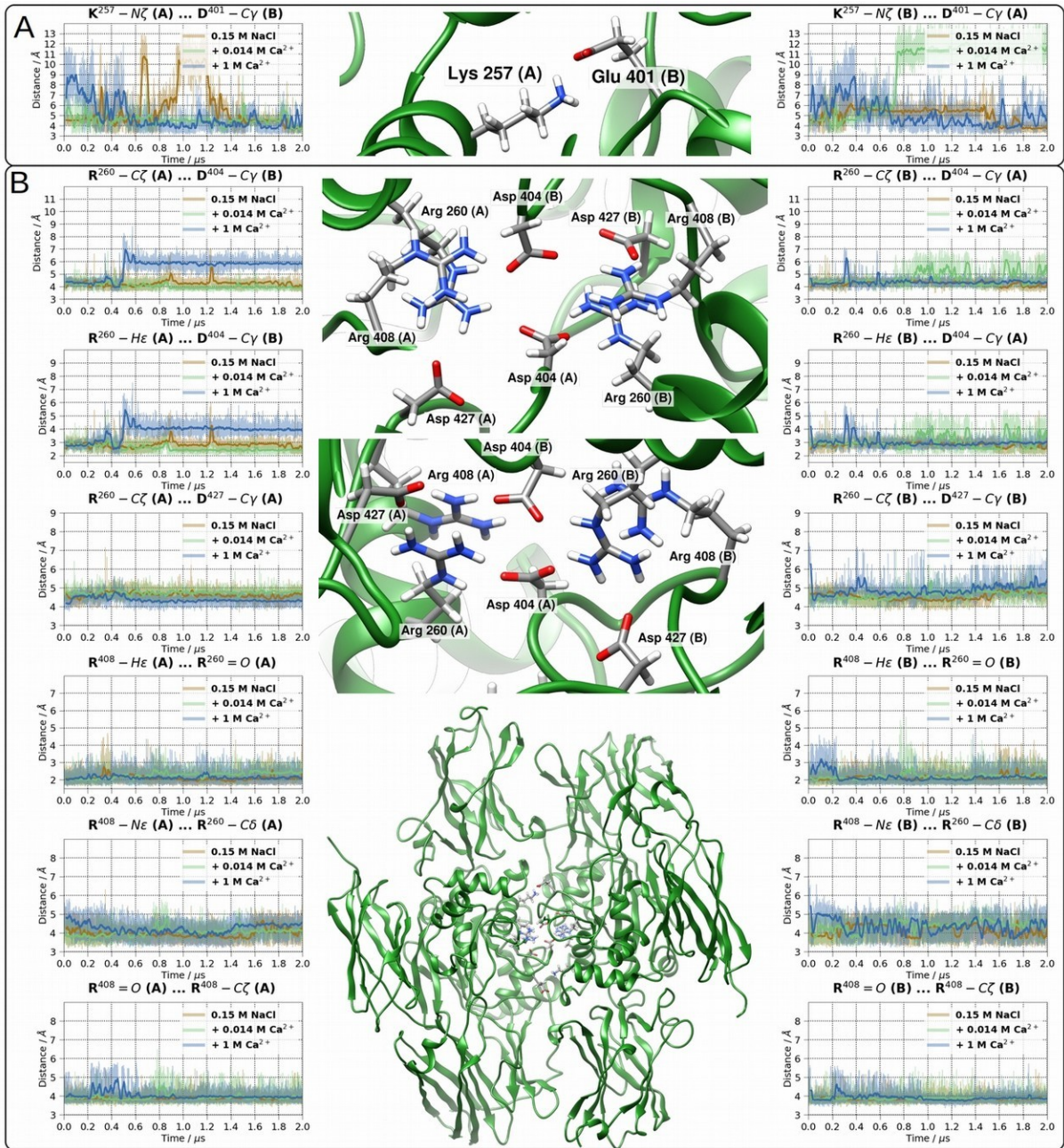
K13. ábra. A dinamikai trajektóriából számolt kalciumkötődési események grafikus reprezentációja az *v/a* (monomer, nyílt konformáció, 4KTY, +14 mM  $\text{Ca}^{2+}$ ).



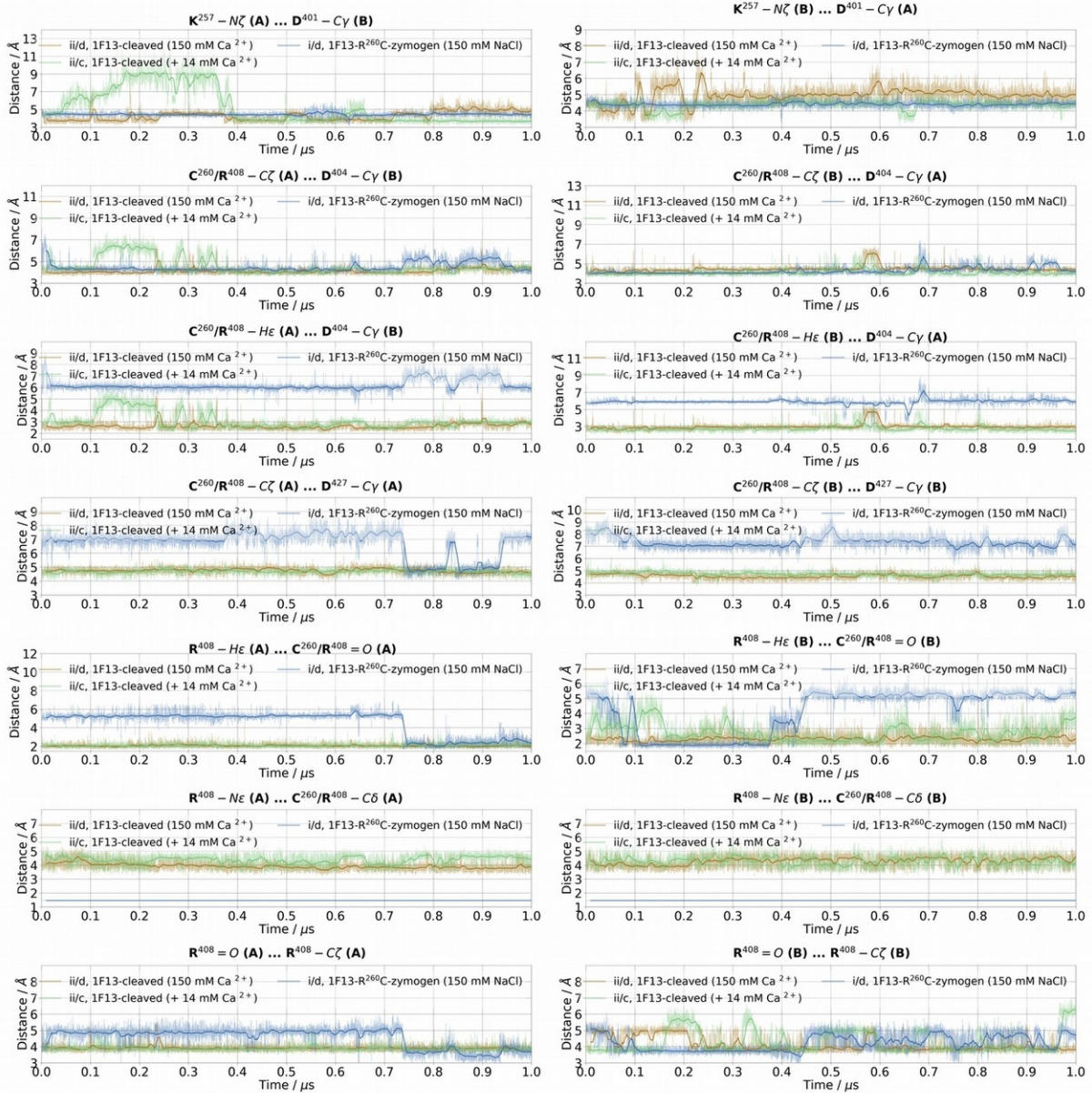
K14. ábra. A FXIII-A<sub>2</sub> i/c szimulációjának utolsó frame-jéből indított “multi-site” kalcium modellel kivitelezett, 50 ns hosszúságú szimuláció eredménye.



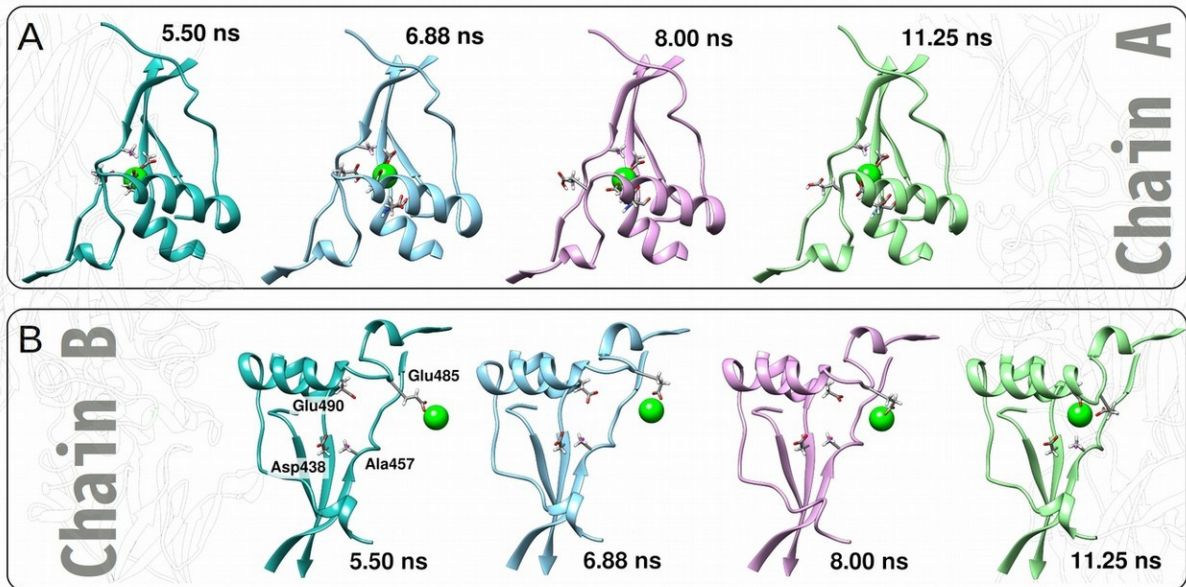
K15. ábra. A kalciumkötőhelyek helyzete az FXIII-A dimer felületén.



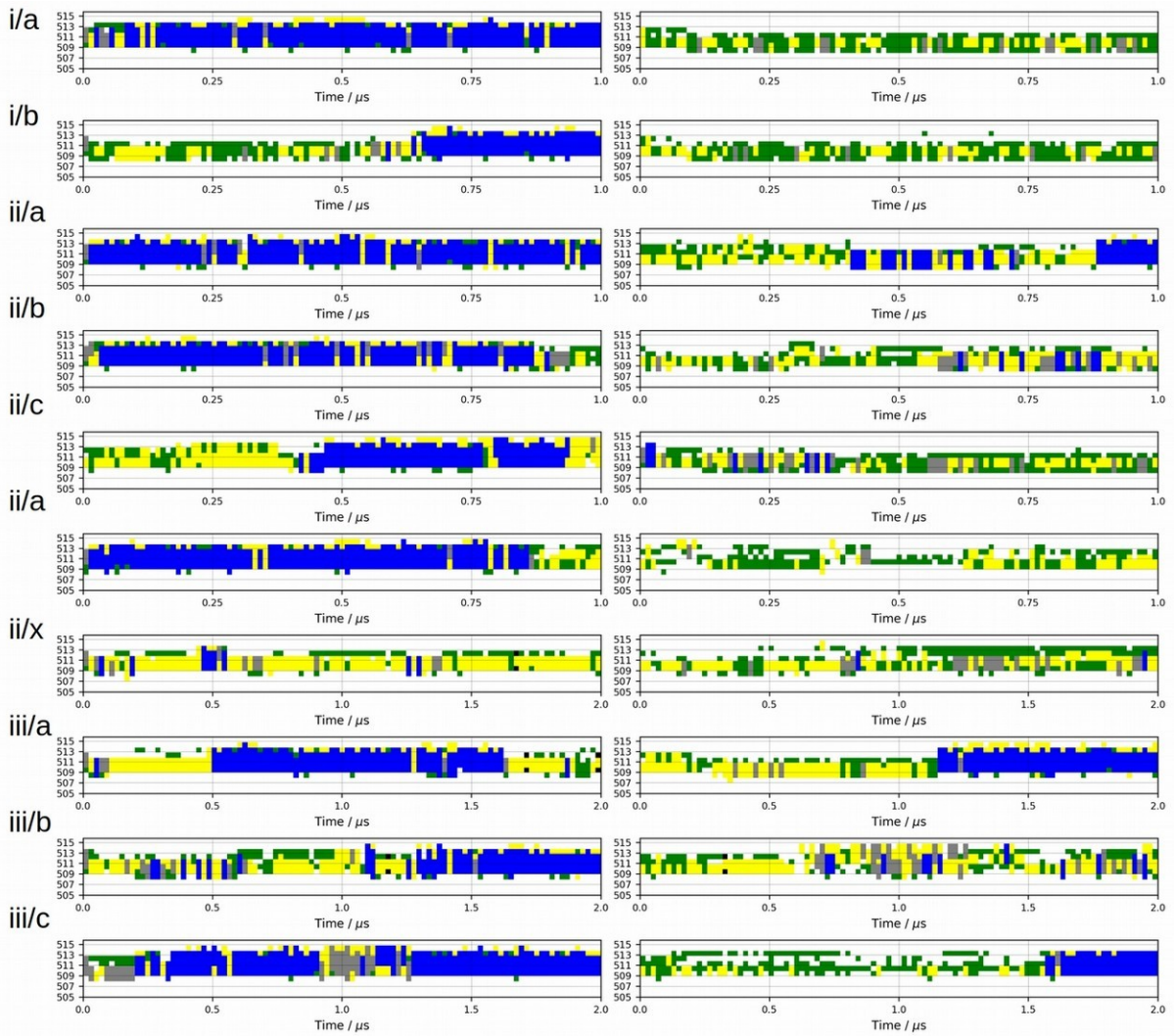
K16. ábra. A központi, A alegységek közötti sóhíd hálózat karakterisztikus távolságai a iii/a-c szimulációk esetén és azok elhelyezkedése az A homodimeren.



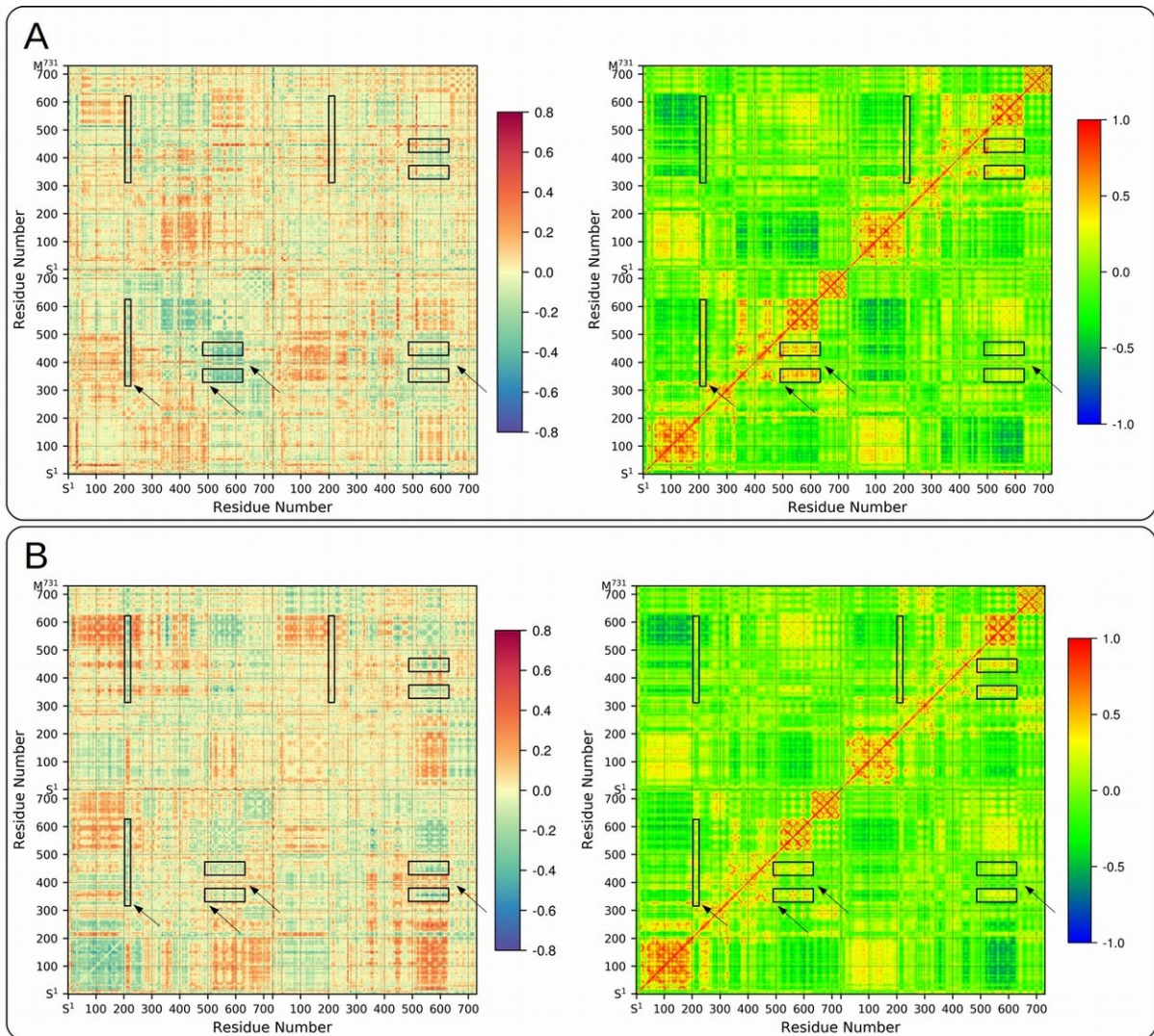
K17. ábra. A központi, A alegységek közötti sóhíd hálózat karakterisztikus távolságai a ii/c, ii/d és az i/d ( $R^{260}C$  mutációt tartalmazó dimer szerkezet, kalcium nélkül) szimulációk esetében.



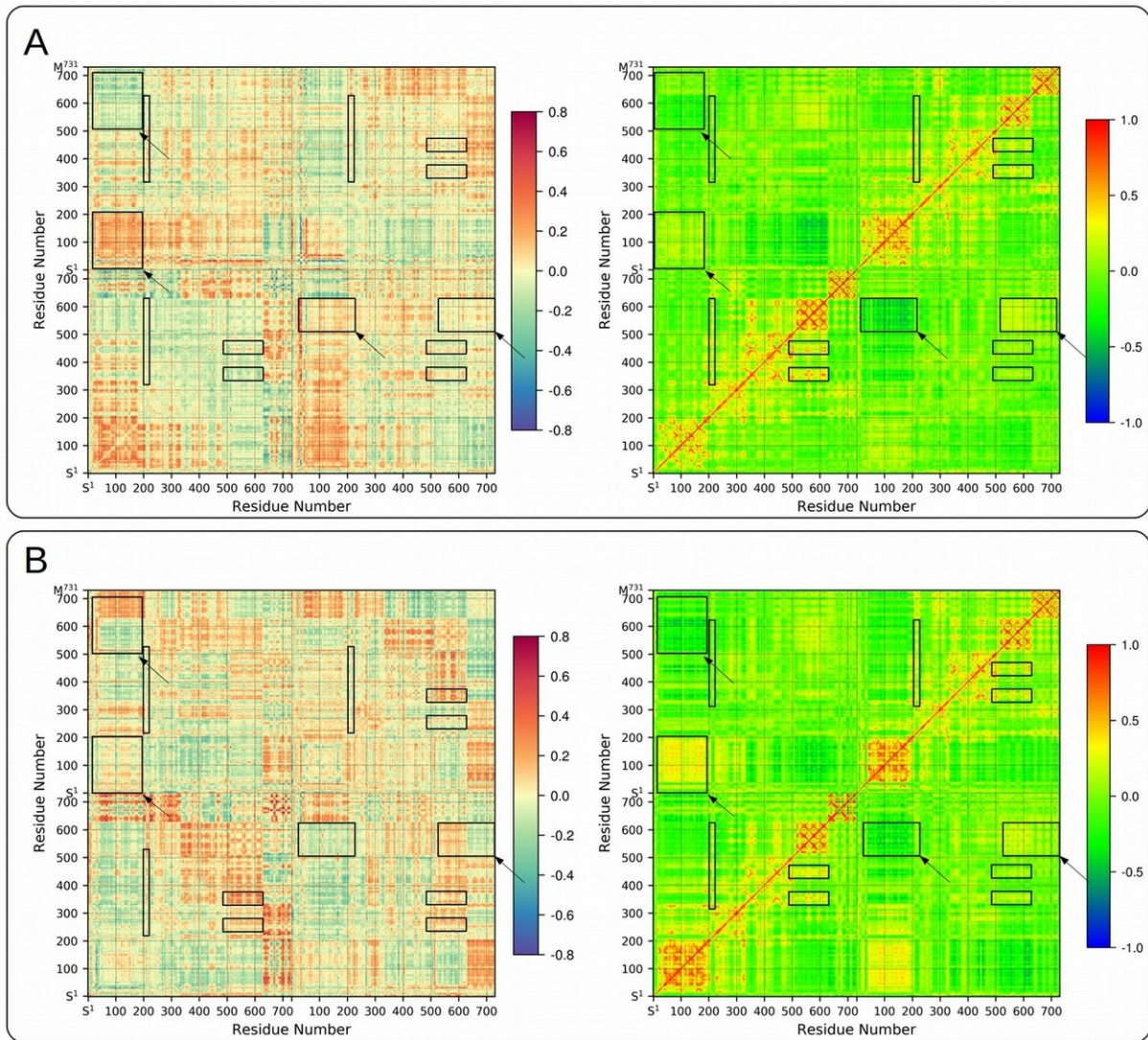
K18. ábra. A ii/c szimuláció utolsó frame-jéből, a kötött kalciumok megtartásával, 100 mM  $\text{Ca}^{2+}$  jelenlétében eltérő kezdősebességgel indított 17 x 25 ns MD szimulációk azon eredménye, melynél a második fő kötőhely betöltődése (B) is megfigyelhető volt, míg a már betöltött kötőhely változatlan maradt (A).



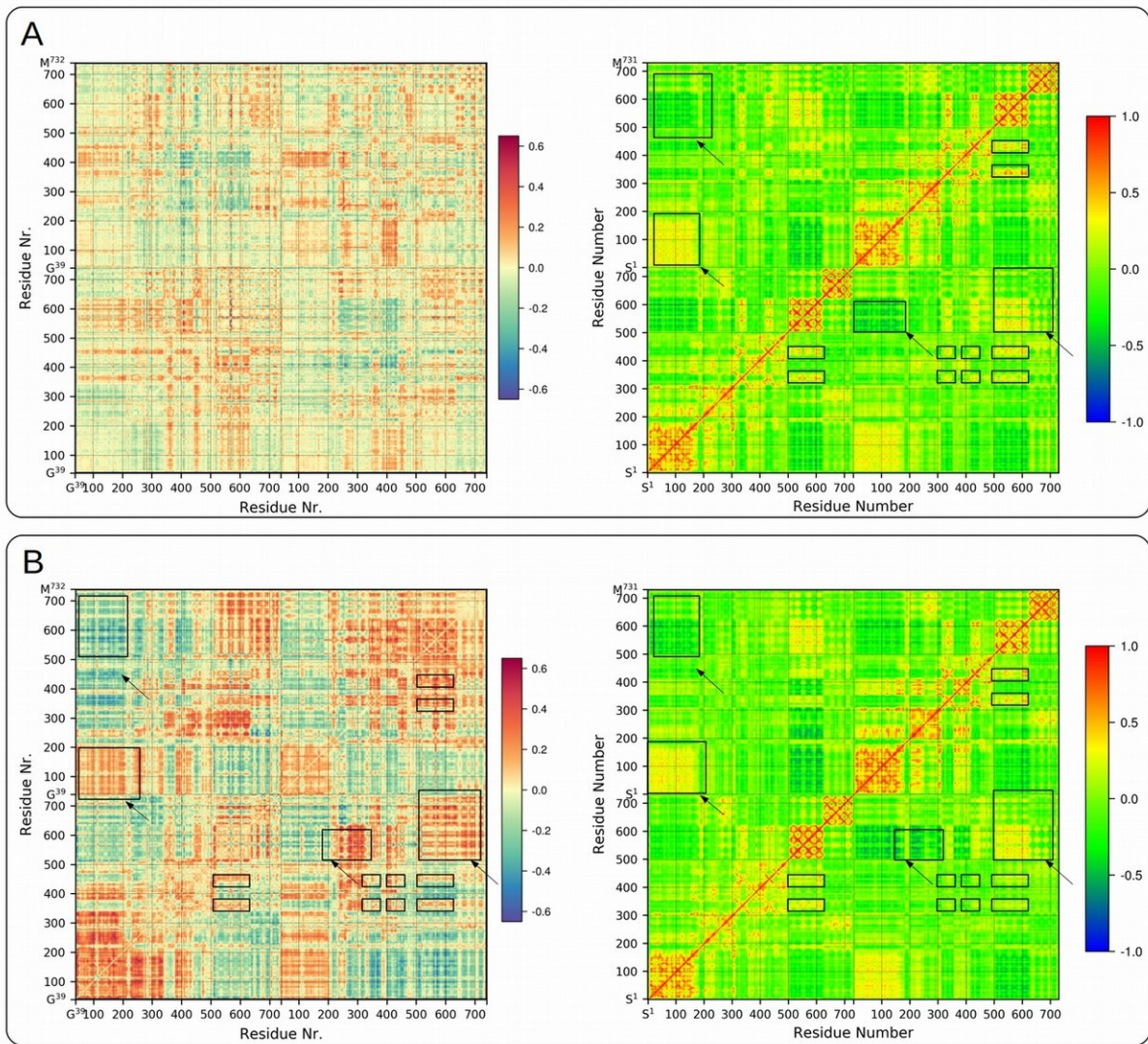
K19. ábra. A FXIII-A<sub>2</sub> Pro505-Arg515 közötti rendezetlen loop régiójának másodlagos szerkezet analízise. Bal oldalon az “A” főlánc, jobb oldalon a “B” főlánc. (Színkódok: fehér = coil, zöld = bend, sárga = turn, kék = alfa hélix, szürke = 3-hélix.)



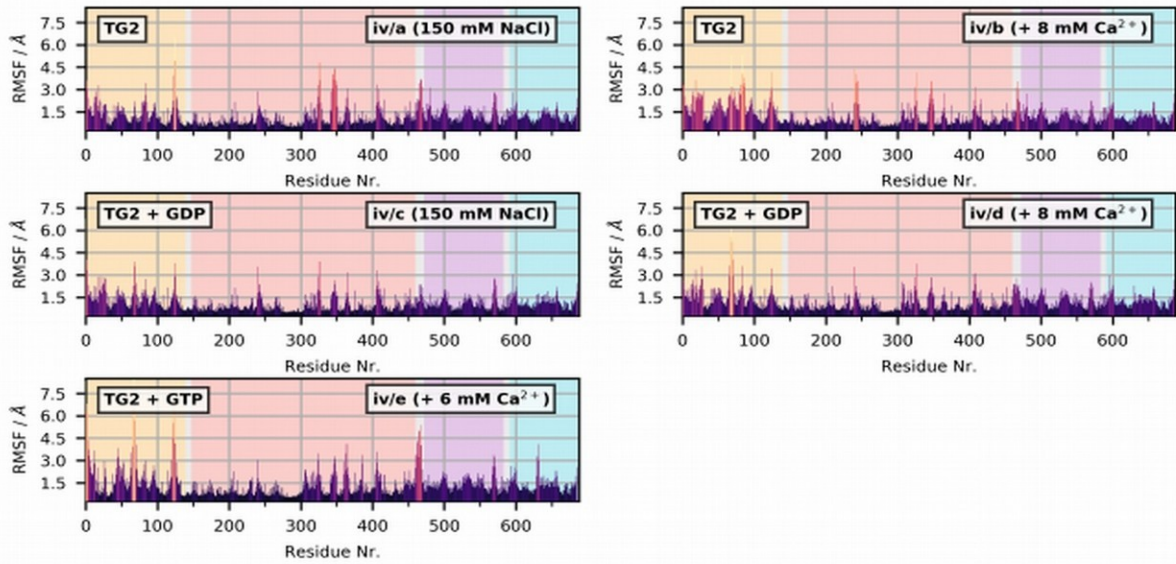
K20. ábra. Dinamikus kereszt-korreláció mátrixok (referencia a jobb oldalon) és a megfelelő különbség mátrixok a bal oldalon. (A), i/a (zimogén) és ii/a (hasított AP-k), valamint (B) i/b (zimogén) és ii/b (hasított AP-k). A hasított AP-k esetében végzett szimulációk eredményei szolgáltak referenciaként (jobb oldal).



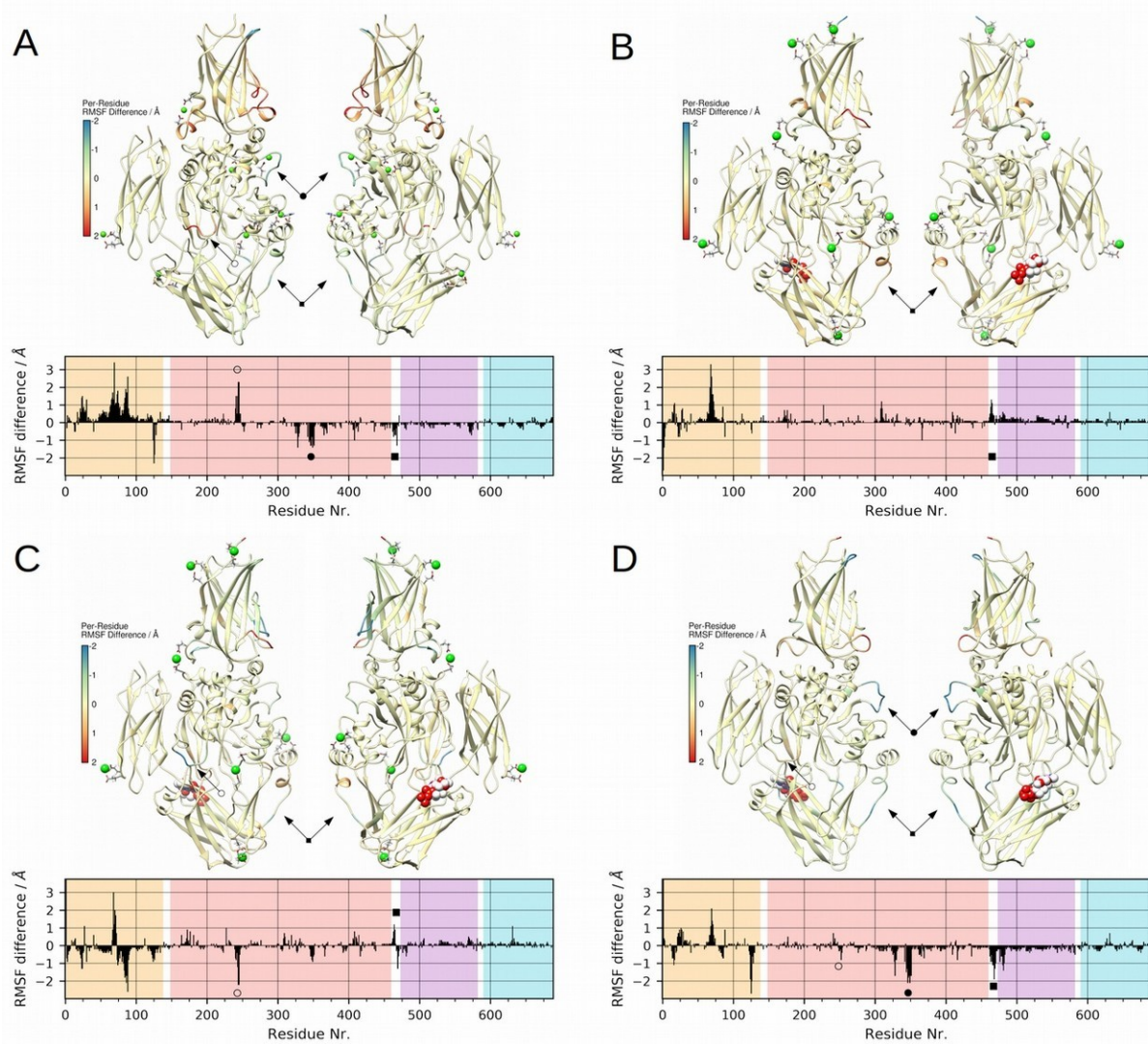
K21. ábra. Dinamikus kereszt-korreláció mátrixok (referencia a jobb oldalon) és a megfelelő különbség mátrixok a bal oldalon. (A) ii/d (hasított AP-k, 1F13, 150 mM NaCl) és ii/c (hasított AP-k, 1F13, +14 mM  $\text{Ca}^{2+}$ ) valamint (B) ii/a (hasított AP-k, 1GGU, +14 mM  $\text{Ca}^{2+}$ ) és ii/c (hasított AP-k, 1F13, +14 mM  $\text{Ca}^{2+}$ ). A ii/d és ii/c szimulációk eredményeit tekintettük referenciának (jobb oldal).



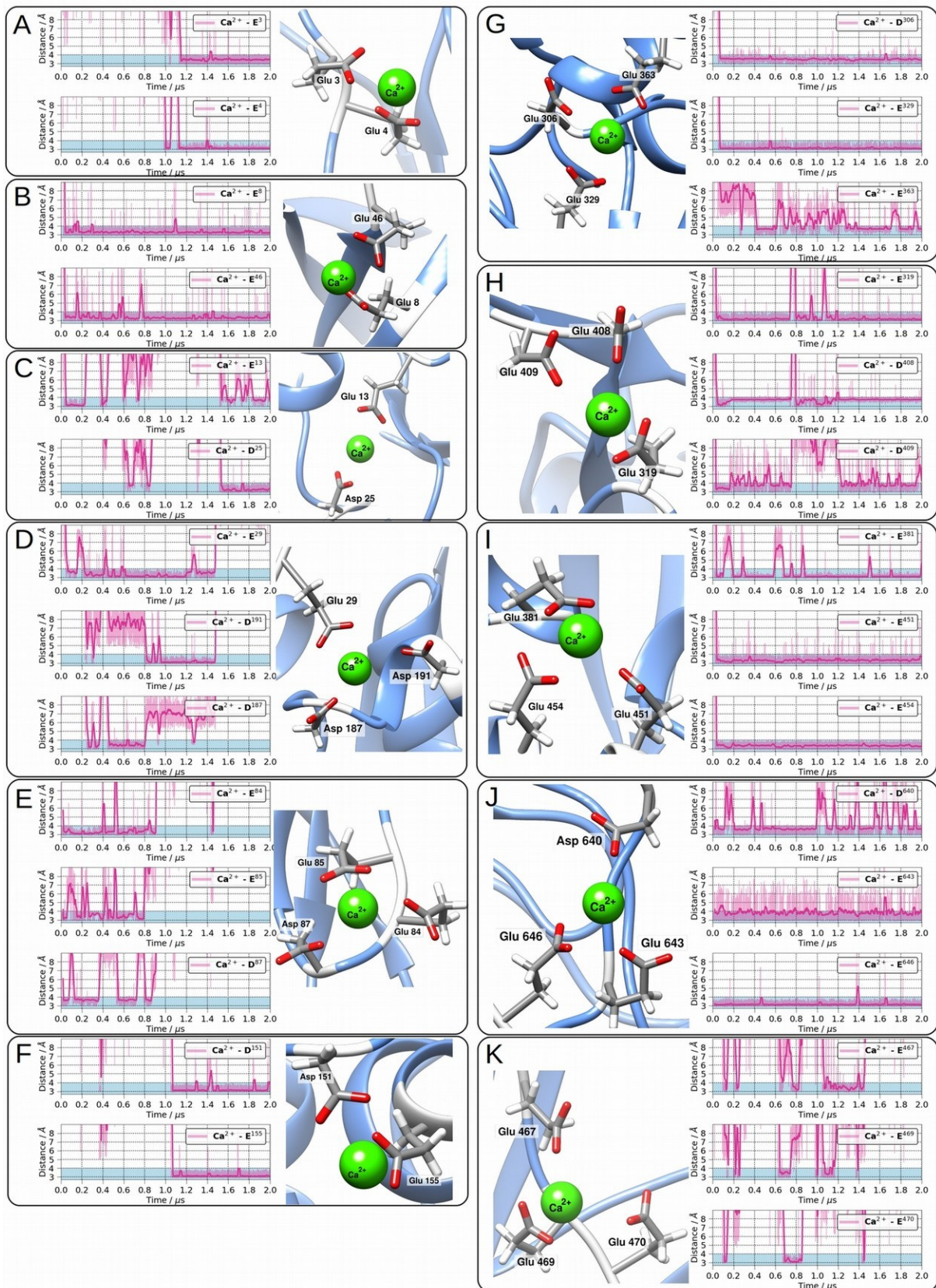
K22. ábra. Dinamikus kereszt-korreláció mátrixok (referencia a jobb oldalon) és a megfelelő különbség mátrixok a bal oldalon. (A), iii/a (AP-k nélkül, 1F13, 150 mM NaCl) és iii/b (AP-k nélkül, 1F13, +14 mM Ca<sup>2+</sup>), valamint (B) iii/b (AP-k nélkül, 1F13, +14 mM Ca<sup>2+</sup>) és iii/c (AP-k nélkül, 1F13, +1000 mM Ca<sup>2+</sup>). A iii/a és iii/b szimulációk eredményeit tekintettük referenciának (jobb oldal).]



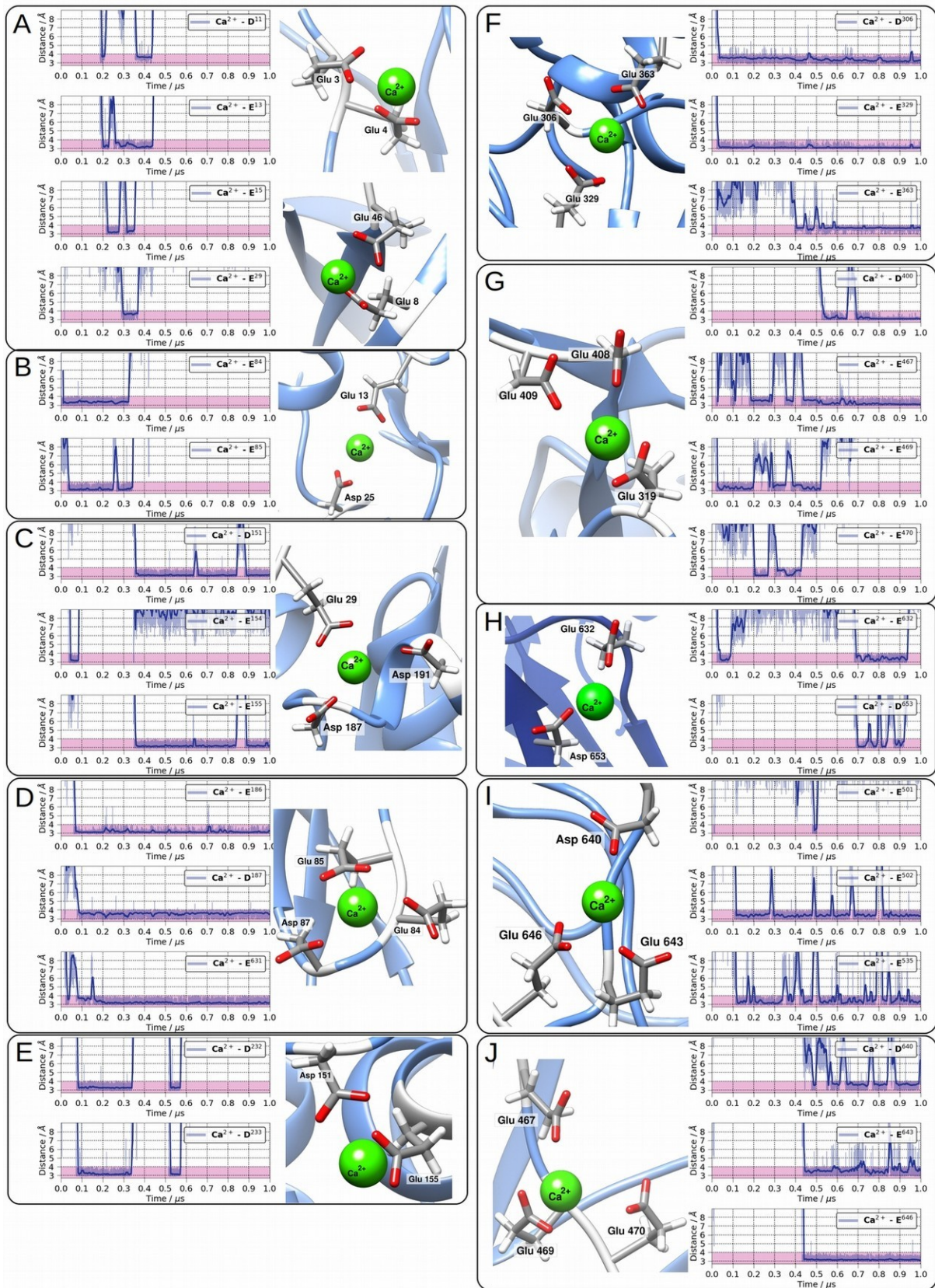
K23. ábra. Aminosavankénti fluktuációk a TG2 (iv/a-e) szimulációnak esetében.



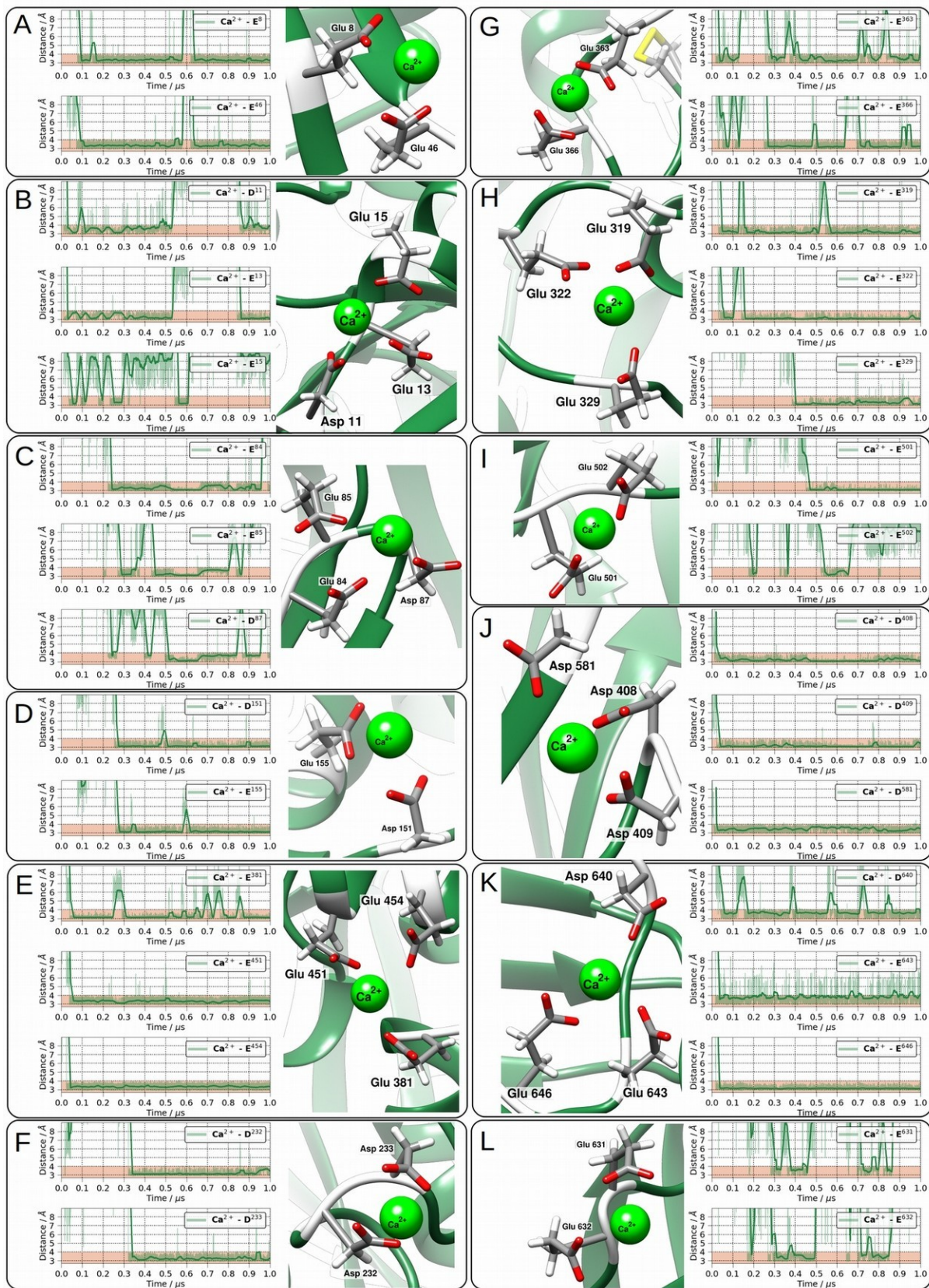
K24. ábra. A TG2 szimulációk esetében kapott aminosavankénti RMSF értékek különbségei. (A) a iv/b és iv/a, (B) iv/d és iv/c, (C) iv/d és iv/b, (D) iv/c és iv/a különbségei.



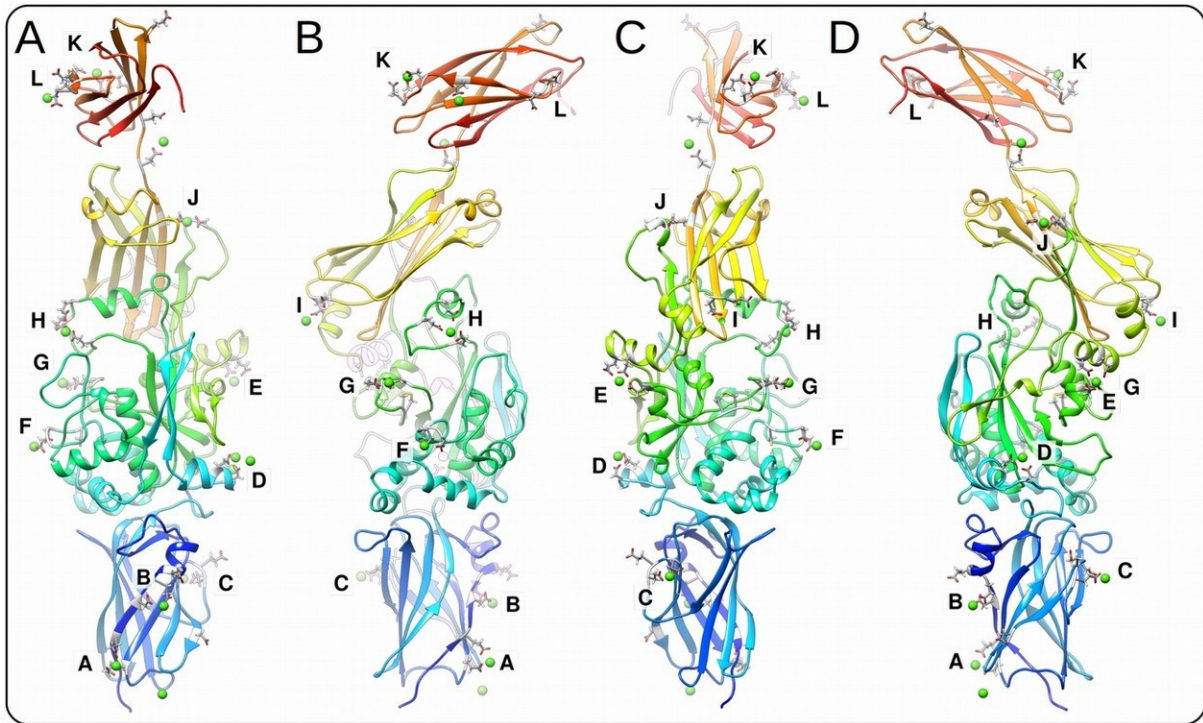
K25. ábra. kalciumkötődések a TG2 (GDP nélkül) iv/b szimulációjának esetében.



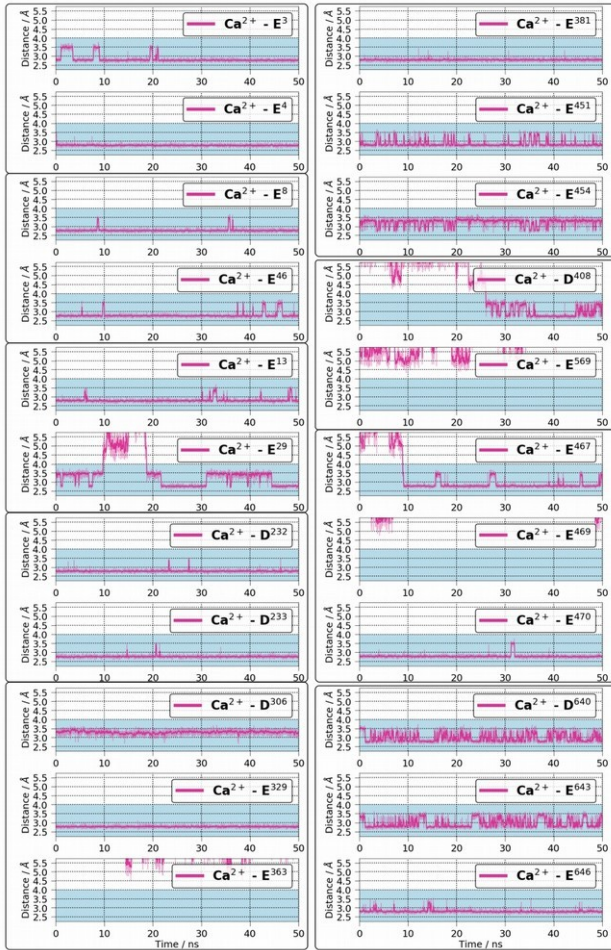
K26. ábra. kalciumkötődések a TG2 (GTP jelenlétében) iv/e szimulációjának esetében.



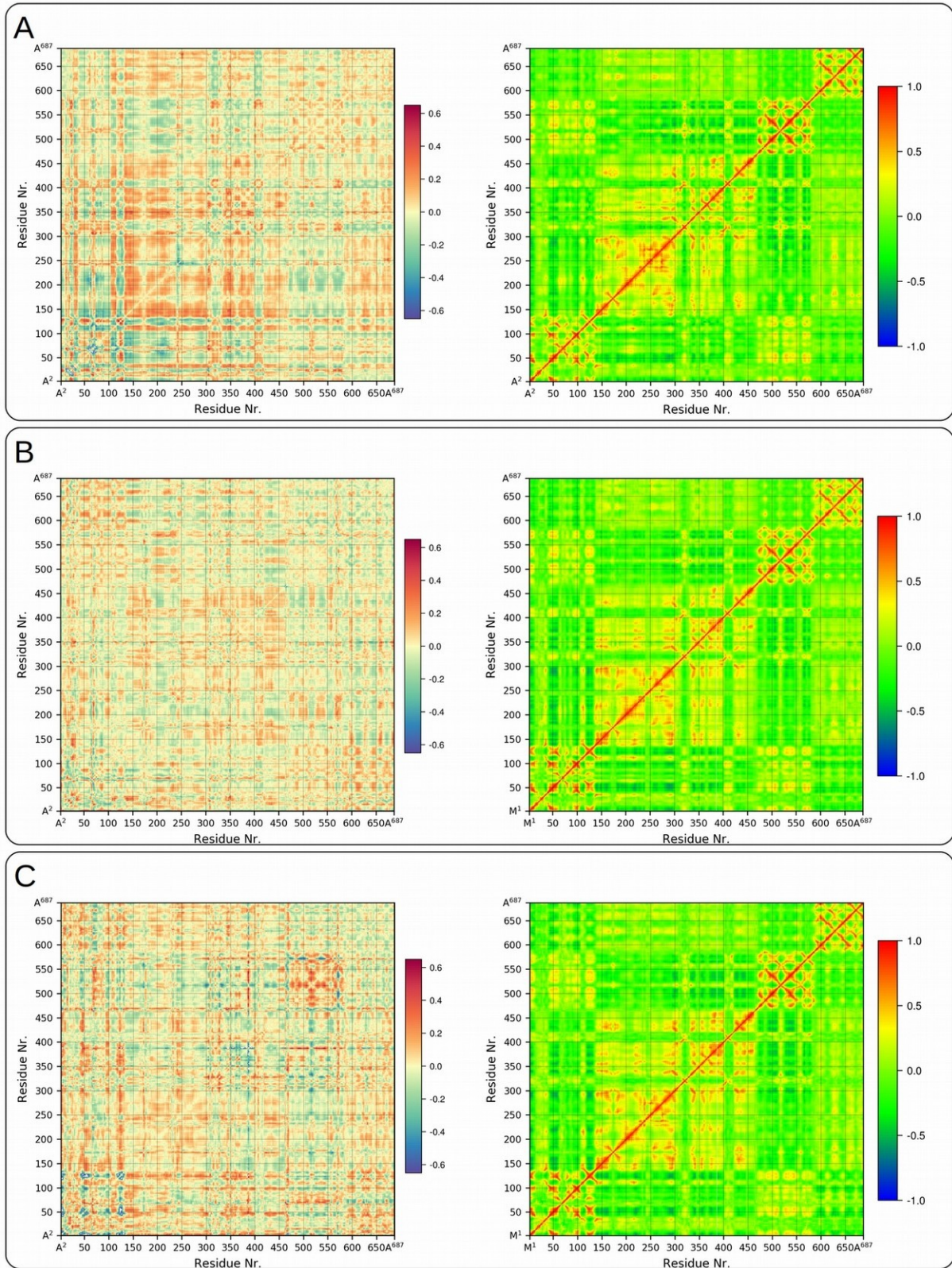
K27. ábra. kalciumkötődések a nyílt TG2 (v/b) szimulációjának esetében.



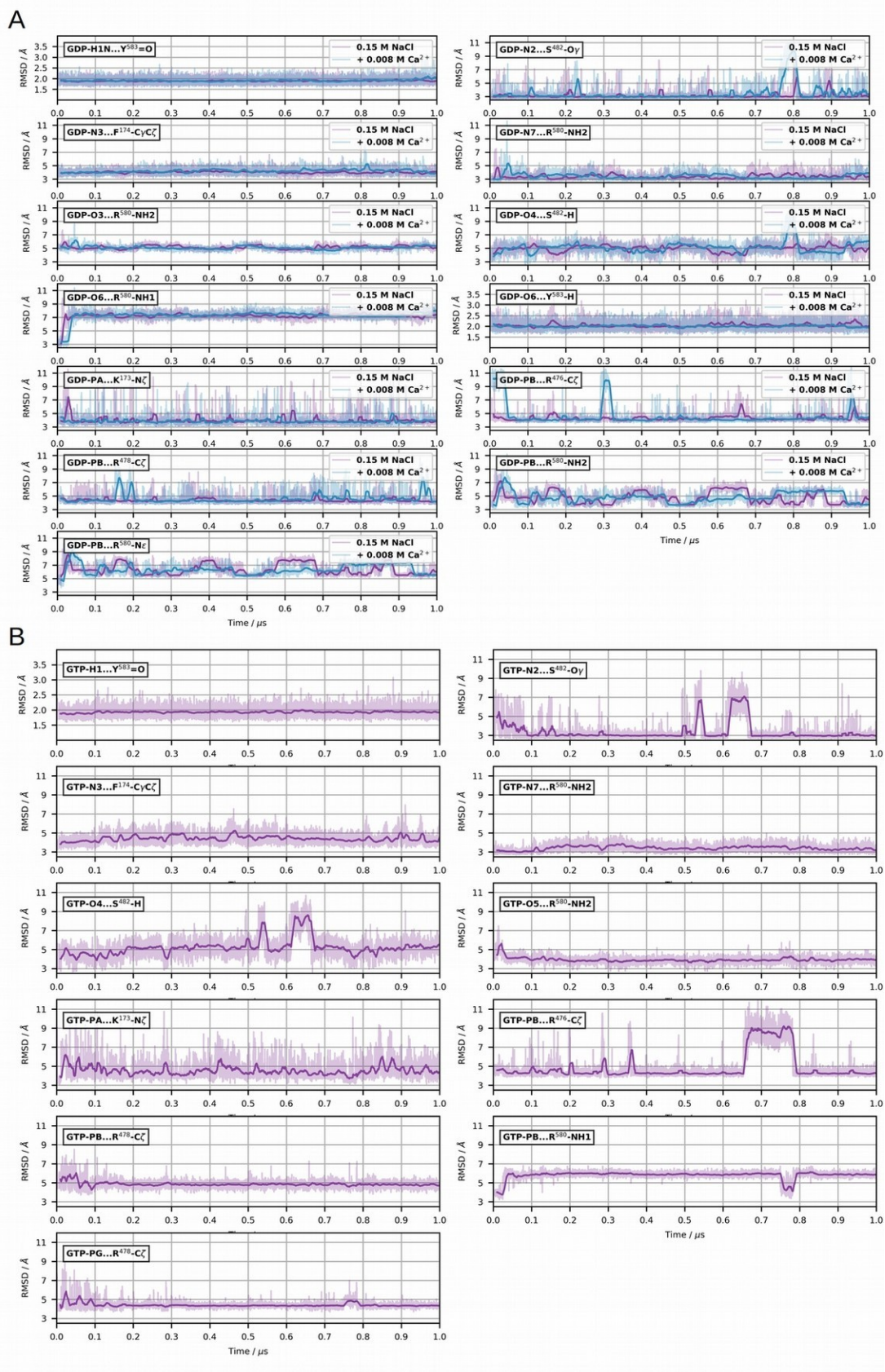
K28. ábra. A nyílt TG2 kalciumkötőhelyeinek elhelyezkedése a protein felületén. A kötőhelyek betűjelei megegyeznek az K27. Ábrán feltüntetetteknek.



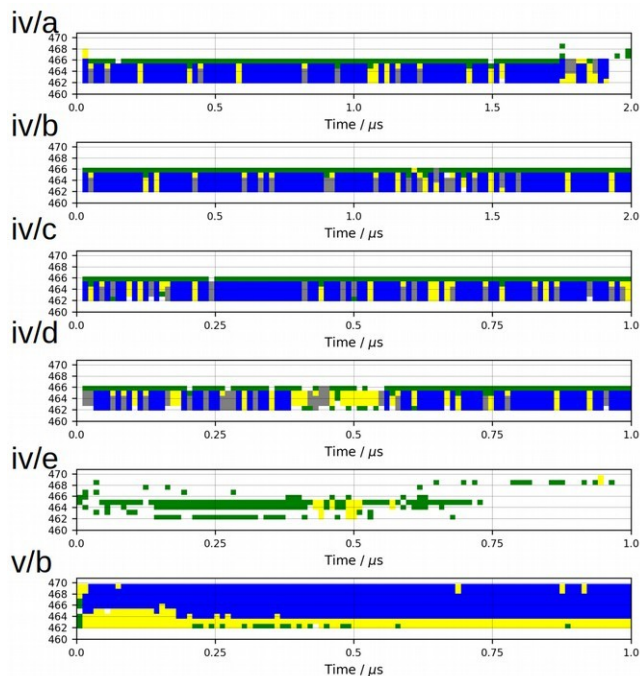
K29. ábra. A TG2 iv/d szimulációjának utolsó frame-jéből indított “multi-site” kalcium modellel kivitelezett, 50 ns hosszúságú szimuláció eredménye.



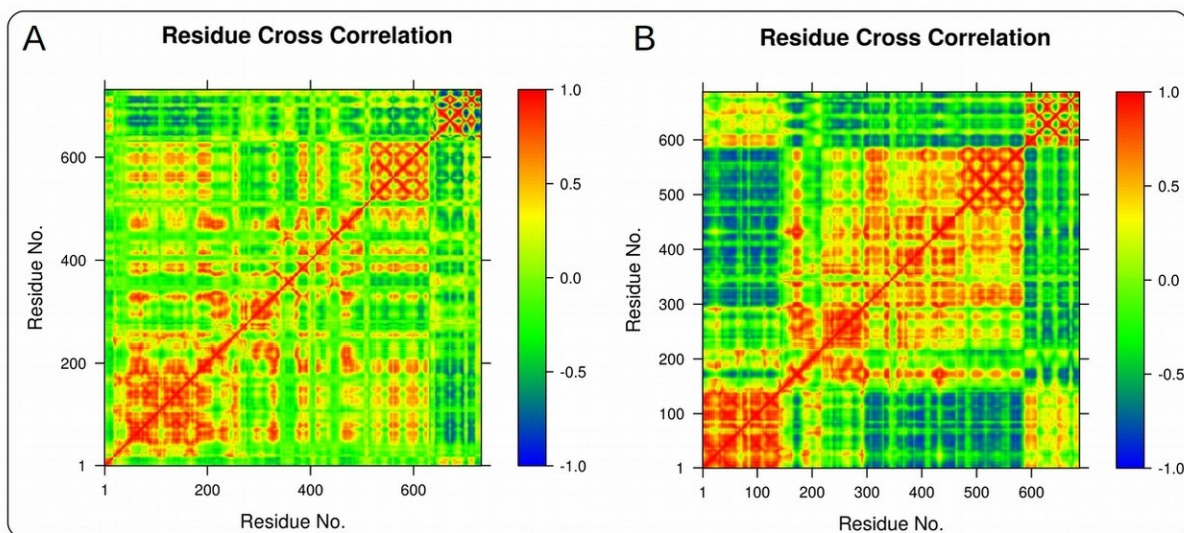
K30. ábra. Dinamikus kereszt-korreláció mátrixok (referencia a jobb oldalon) és a megfelelő különbség mátrixok a bal oldalon. (A) iv/b és iv/a, (B) iv/d és iv/c, (C) iv/e és iv/d különbség mátrixok (bal oldalon) valamint a jobb oldalon a referenciák rendre: iv/a, iv/c és iv/d szimulációkból nyert mátrixok.



K31. ábra. A TG2 GDP/GTP kötődései során kapott karakterisztikus távolságok. (A) iv/c és iv/d, valamint (B) a iv/ szimuláció(k) eredményei.



K32. ábra. A TG2 Asn460-Thr471 közötti rendezetlen loop régiójának másodlagos szerkezet analízise. Színek: fehér = coil, zöld = bend, sárga = turn, kék = alfa hélix, szürke = 3-hélix.)



K33. ábra. Dinamikus kereszt-korreláció mátrixok az (A) v/a (FXIII-A<sup>0</sup>, 4KTY, +14 mM Ca<sup>2+</sup>) és (B) (nyílt TG2, 2Q3Z, +14 mM Ca<sup>2+</sup>) esetében.



Nyilvántartási szám: DEENK/51/2019.PL  
Tárgy: PhD Publikációs Lista

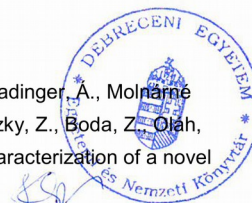
Jelölt: Fekete Attila  
Neptun kód: L0TV9S  
Doktori Iskola: Laki Kálmán Doktori Iskola  
MTMT azonosító: 10062749

### A PhD értekezés alapjául szolgáló közlemények

1. **Fekete, A.**, Komáromi, I., Mucs, D.: On the early events of the calcium induced activation of coagulation factor XIII-A2 and tissue transglutaminase: an in silico study.  
*J. Biomol. Struct. Dyn. [Epub ahead of print]*, 1-16, 2019.  
IF: 3.107 (2017)
2. **Fekete, A.**, Komáromi, I.: Modeling the archetype cysteine protease reaction using dispersion corrected density functional methods in ONIOM-type hybrid QM/MM calculations: the proteolytic reaction of papain.  
*Phys. Chem. Chem. Phys.* 18 (48), 32847-32861, 2016.  
DOI: <http://dx.doi.org/10.1039/C6CP06869C>  
IF: 4.123

### További közlemények

3. Fizil, Á., Sonderegger, C., Czajlik, A., **Fekete, A.**, Komáromi, I., Hajdu, D., Marx, F., Batta, G.: Calcium binding of the antifungal protein PAF: Structure, dynamics and function aspects by NMR and MD simulations.  
*PLoS One.* 13 (10), 1-19, 2018.  
DOI: <http://dx.doi.org/10.1371/journal.pone.0204825>  
IF: 2.766 (2017)
4. Selmeczi, A., Gindele, R., Ilonczai, P., **Fekete, A.**, Komáromi, I., Schlammadinger, A., Molnár, Rázsó, K., Kovács, K. B., Bárdos, H., Ádány, R., Muszbek, L., Bereczky, Z., Boda, Z., Olah, Z.: Antithrombin Debrecen (p.Leu205Pro) - Clinical and molecular characterization of a novel mutation associated with severe thrombotic tendency.  
*Thromb. Res.* 158, 1-7, 2017.  
DOI: <http://dx.doi.org/10.1016/j.thromres.2017.07.023>  
IF: 2.779





5. Tóth, L., **Fekete, A.**, Balogh, G., Bereczky, Z., Komáromi, I.: Dynamic properties of the native free antithrombin from molecular dynamics simulations: computational evidence for solvent-exposed Arg393 side chain.  
*J. Biomol. Struct. Dyn.* 33 (9), 2023-2036, 2015.  
DOI: <http://dx.doi.org/10.1080/07391102.2014.986525>  
IF: 2.3
6. Kovács, K. B., Pataki, I., Bárdos, H., **Fekete, A.**, Pfliegler, G., Haramura, G., Gindele, R., Komáromi, I., Balla, G., Ádány, R., Muszbek, L., Bereczky, Z.: Molecular characterization of p.Asp77Gly and the novel p.Ala163Val and p.Ala163Glu mutations causing protein C deficiency.  
*Thromb. Res.* 135 (4), 718-726, 2015.  
DOI: <http://dx.doi.org/10.1016/j.thromres.2015.01.011>  
IF: 2.32
7. Bokor, É., **Fekete, A.**, Varga, G., Szócs, B., Czifrák, K., Komáromi, I., Somsák, L.: C-( $\beta$ -d-Glucopyranosyl)formamidrazones, formic acid hydrazides and their transformations into 3-( $\beta$ -d-glucopyranosyl)-5-substituted-1,2,4-triazoles: a synthetic and computational study.  
*Tetrahedron.* 69 (48), 10391-10404, 2013.  
DOI: <http://dx.doi.org/10.1016/j.tet.2013.09.099>  
IF: 2.817

**A közlő folyóiratok összesített impakt faktora: 20,212**

**A közlő folyóiratok összesített impakt faktora (az értekezés alapjául szolgáló közleményekre):  
7,23**

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, 2019.03.05.



### 14.3. Közlemények különnyomatai