

Lokális kölcsönhatások által modulált  
belső dinamika mint biológiai szabályozó  
mechanizmus vizsgálata globuláris  
fehérjékben



Doktori értekezés

Kovács Bertalan

2019

Pázmány Péter Katolikus Egyetem

Roska Tamás Műszaki és Természettudományi Doktori Iskola

Témavezető: Dr. Gáspári Zoltán

ipjC'í·m öy/

*Utúlie'n aure!*

„Ám ekkor kiáltás szállt föl, s röpült a szél szárnyán völgyről völgyre, s a tündék meg az emberek ámulattal és örömmel emelték föl hangjukat. Mert kéretlenül és váratlanul Turgon megnyitotta Gondolin körül a zárat, s előmasírozott tízezres serege fényes páncélban, lándzsák erdejével. S amikor Fingon meghallotta a távolból fivérének, Turgonnak a kürtjét, elszállt szívéből az árnyék, s hangosan kiáltotta:

– *Utúlie'n aure! Aiya Eldalie ar Atanári, utúlie'n aure!* Eljött a nap! Nézzétek, eldák és emberek atyáinak népe, eljött a nap!”

J. R. R. Tolkien: A szilmarilok

Gálvölgyi Judit fordítása

E dolgozat létrejöttéért köszönet illeti

elsősorban témavezetőmet, Gáspári Zoltánt, aki a tudományos kutatásban egyszerre volt vezető és útítárs, aki egyszerre inspirált és támogatott, folyamatosan szem előtt tartva, hogy mi az, ami leginkább a javamat szolgálja, akinek az ajtaja mindig nyitva volt előttem, és szinte felfoghatatlan módon bármikor volt fél órája, hogy átbeszéljük éppen aktuális nehézségeimet, és akiben a szakmai tudás és az alázat olyan ötvözetét ismertem meg, ami mindig példaértékű marad számomra,

bírálóimat, Dr. Hegedűs Tamást és Dr. Beke-Somfai Tamást, akik észrevételeikkel, kérdéseikkel és építő kritikájukkal nagyban segítették e szöveg csiszolódását,

a Pázmány Péter Katolikus Egyetemet és a Roska Tamás Műszaki és Természettudományi Doktori Iskolát a belém vetett bizalomért, a lelki és anyagi támogatásért, amelyet nem fukarkodott többször is meghosszabbítani az állami ösztöndíjam lejárta után, de eközben mégsem szűnt meg szelíden, de határozottan biztatni engem a disszertáció befejezésére,

a Doktori és Habilitációs Iroda munkatársát, Vida Katinka nénit, akinek áldozatos segítségével nélkül számtalanszor elvesztem volna a leadási határidők és adminisztratív kötelességek útvesztőjében,

drága menyasszonyomat, Annamarit, aki az elmúlt hónapok során angyali türelemmel viselte, hogy közös vacsoráinkat egyre inkább a doktori munkám előrehaladása, semmint közös jövőnk tervezése töltse ki, s akinek keze munkája díszíti e disszertáció mottóját,

sorstársaimat, különösen Meszéna Domit és Csány Gergőt, akikkel vállt vállnak vetve álltunk helyt mind az egyetemen, mind a színpadon, és együtt fejeztük be a doktori iskolát,

egész családomat, és különösen édesanyámat, aki jóformán egyetlen alkalmat sem mulasztott el, hogy diszkréten a disszertációm állapota felől érdeklődjön,

leendő apósomat és anyósomat, akik végtelen odaadással biztosítottak nekem helyet a saját házukban a karantén három hónapja alatt, és ezáltal lényegében otthont adtak a disszertációmnak is,

kollégáimat a ChemAxon Kft.-nél a kedves biztatásért a munka hajrájában,

és végül a Fokoláre mozgalmat és lelki édesanyámat, az Egyházat, ahova akár győzelemittasan, akár szégyenkezve, de mindig hazatérhetek, bárhová is sodorjon az élet.

# Tartalomjegyzék

Tartalomjegyzék	I
Kivonat	III
Abstract	IV
Közlemények	V
Ábrajegyzék	VI
Ábrák	VI
Táblázatok	IX
Függelékek	IX
Rövidítések listája	XII
1 Bevezetés	1
2 Irodalmi áttekintés	4
2.1 A PSD-95 N-terminális PDZ1-2 tandemjének szerkezete és belső dinamikája	4
2.1.1 A serkentő posztszinaptikus denzitás (PSD)	4
2.1.2 A PSD-95 és a MAGUK fehérjecsalád	7
2.1.3 PDZ domén szerkezete és belső dinamikája	9
2.1.4 PDZ domének alkotta szupramoduláris szerkezetek és dinamikájuk	12
2.2 A peptidil-prolil izomerázok	16
2.2.1 A peptidil-prolil izomerázok sejtbiológiai szerepe	16
2.2.2 A peptidil-prolil izomerázok szerkezete, dinamikája és mechanizmusa	20
2.3 A fehérjemozgások jellemzése	25
2.3.1 A fehérjék belső dinamikája	25
2.3.2 NMR-rel mérhető, gyors dinamikát tükröző paraméterek	26
2.3.3 Dinamikus fehérjeszerkezeti sokaságok	29
2.3.4 A szerkezeti sokaságok előállítása mint matematikai probléma	31
2.4 A molekuladinamikai szimuláció	32
3 Célkitűzés	35
4 Alkalmazott módszerek	36
4.1 Külső kényszerfeltételekkel megkötött molekuladinamikai szimuláció	36
4.1.1 Kényszerfeltételként alkalmazott kísérletes paraméterek	36
4.1.2 A kényszerfeltételek alkalmazása	38
4.1.3 A MUMO-protokoll	38
4.1.4 Lokális illesztés	39
4.1.5 A szimulációk felparaméterezése	40



4.2	Dinamikus szerkezeti sokaságok validálása	41
4.3	Fehérjeszerkezeti adatfájlok elemzése a ProDitor saját fejlesztésű programcsomaggal	42
4.3.1	RMSD-számolás	42
4.3.2	Sokaságok illesztése	43
4.4	A leghangsúlyosabb belső mozgások azonosítása főkomponens analízissel	43
4.5	Kísérletes <sup>1</sup> H- <sup>1</sup> H NOE-ekből származó távolság-kényszerfeltételek sztereoszimmetrizálása	44
4.6	Felhasznált szoftverek	45
5	A PSD-95 PDZ1-2 tandemjének ligandumfüggő intra- és interdomén dinamikai vizsgálata	46
5.1	A szerkezeti sokaságok előállítás	46
5.1.1	Kiindulási adatok előkészítése	46
5.1.2	A molekuladinamikai szimulációk futtatása	48
5.2	A szerkezeti sokaságok elemzése	51
5.2.1	A szerkezeti sokaságok validálása	51
5.2.2	Ligandumfüggő intradomén mozgások	51
5.2.3	Ligandumfüggő interdomén mozgások	55
5.2.4	Szoros illeszkedésű interdomén orientációk klaszterezése	59
5.2.5	Atomi szintű kölcsönhatások a szoros illeszkedésű klaszterekben	61
5.2.6	A klaszterekben található interdomén orientációk összevetése kísérletesen meghatározott szerkezetekkel	63
5.2.7	Az intra- és interdomén mozgások összefüggésének vizsgálata	66
5.3	Diszkusszió – Első tézis csoport	68
6	A parvulin-típusú cisz-transz izomerázok belső mozgásainak vizsgálata	73
6.1	A felhasznált szerkezeti sokaságok	73
6.2	A szerkezeti sokaságok elemzése	75
6.2.1	A kísérleti adatoknak való megfelelés	75
6.2.2	A leghangsúlyosabb belső mozgások	76
6.2.3	Összevetés kísérletes parvulin szerkezetekkel	79
6.2.4	A konzervált hisztidinek és a hidrogénkötés-hálózat vizsgálata	81
6.3	Diszkusszió – Második tézis csoport	84
7	Konklúzió és kitekintés	88
8	Irodalomjegyzék	90
9	Függelékek	105

## Kivonat

A PSD-95 fehérje N-terminális két PDZ doménje egy független szerkezeti és dinamikai egységet alkot, amelyet PDZ tandemnek hívunk. A PDZ domén az idegsejtekben igen gyakran előforduló fehérje-fehérje kölcsönhatási alegység: szerepet játszik a membránreceptorok klaszterezésében és a szinaptikus plaszticitás szabályozásában. A PSD-95 PDZ1-2 tandemjének számtalan bizonyított kötőpartnerre van, amely igen nehezé teszi egy atomi szintű szerkezeti-dinamikai modell megalkotását. Az elérhető kísérletes szerkezetek egymásnak ellentmondanak a PDZ1-2 tandem szupramoduláris szerkezetét illetően, valamint az is bebizonyosodott, hogy a kötőpartnerrel kialakított kölcsönhatás szignifikáns interdomén reorientációt indukál a fehérjében.

A parvulinok a peptidil-prolil cisz-transz izomeráz (PPIáz) vagy rotamáz enzimek egyik családja. A PPIázok a prolint megelőző peptidkötés cisz-transz izomerizációját katalizálják, és szinte mindegyik sejtalkotóban megtalálhatók. Szerepet játszanak többek között a sejtciklus szabályozásában, az immunválaszban és a génexpresszióban is. Bár általánosan elfogadott, hogy a reakció a peptidkötés felszakadása nélkül, egy csavart amid átmeneti állapoton keresztül játszódik le, a katalízis pontos mechanizmusa nem tisztázott. Felderítetlen továbbá a parvulinokban található két konzervált hisztidin, egy 5 aminosavból álló hidrogénkötés-hálózat, valamint a bizonyos parvulinokban jelen lévő N-terminális WW domén szerepe a fehérje biológiai funkciójában.

Azért, hogy pontosabb képet kapjak e két fehérje belső dinamikája és biológiai funkciója közötti különbségről, dinamikus fehérjeszerkezeti sokaságokat állítottam elő külső kényszerfeltételekkel megkötött molekuladinamikai szimulációkkal. Az értekezésben bemutatom a PSD-95 fehérje PDZ1-2 tandemjéről készült szerkezeti sokaságok előállítását és elemzését, valamint három parvulin-típusú rotamáz, a SaPrsA, a TbPin1 és a CsPinA fehérjékről korábban előállított sokaság elemzését.

A molekuladinamikai szimulációkhoz kísérletesen meghatározott  $S^2$  rendparamétereket és  $^1\text{H}$ - $^1\text{H}$  NOE távolság-megkötéseket alkalmaztam külső kényszerfeltételként, amelyek a fehérje gyors, ps-ns időskálán történő mozgásait tükrözik. A PDZ tandem esetén a kényszerfeltételeket lokális illesztési algoritmussal csak az intradomén mozgások megkötésére alkalmaztam. A leghangsúlyosabb belső mozgásokat főkomponens-analízissel vizsgáltam, és a szerkezeti sokaságokat összevettem a nyilvánosan elérhető adatbázisban lévő kísérletes szerkezetekkel.

Az előállított sokaságok elemzése során kiderült, hogy a kötőzseb kinyíló-becsukódó mozgásának modulációja egy általános szabályozó mechanizmus lehet a vizsgált fehérjékben. A PDZ1-2 tandemben a kötőpartner jelenléte megváltoztatja a flexibilis  $\beta$ 2-  $\beta$ 3 hurok dinamikáját, amely pedig jelentős részben az interdomén orientáció szabályozásáért is felel. A parvulinokban ezzel szemben mind a hidrogénkötés protonáltsági állapota, mind pedig a WW domén és a kötőzseb mozgását lehetővé tevő könyökrégió közötti kölcsönhatás modulálja a kinyíló-becsukódó mozgást. A különböző parvulinok eltérő biológiai aktivitása részben megmagyarázható e mozgás eltérő mértékével.

## Abstract

The two N-terminal PDZ domains of the postsynaptic density protein 95 (PSD-95) constitute an independent structural and dynamical unit termed PDZ tandem. The PDZ domain is a frequently occurring protein-protein interaction subunit in nerve cells. It plays a key role in clustering membrane receptors and regulating synaptic plasticity. The PDZ1-2 tandem of PSD-95 has several proven binding partners, therefore creating an atomic-level structural-dynamical model of it remains a challenging task. The available experimental structures of PDZ1-2 tandem contradict each other in terms of its supramodular conformation. Furthermore, ligand binding was shown to induce significant interdomain reorientation.

Parvulins are a family of the peptidyl-prolyl cis-trans isomerase (PPIase) or rotamase enzymes. PPIases, being present in practically all subcellular components, catalyze the cis-trans isomerization of the peptide bond preceding proline residues. They are involved, among other functions, in cell-cycle regulation, immune response and gene expression. Even though the catalyzed reaction is generally accepted to occur through a twisted amide transition state, without breaking and reforming the peptide bond, the exact catalytic mechanism remains elusive. Also, it is unknown how the two conserved histidines, the 5-residue hydrogen-bonding network and the N-terminal WW domain, if present, are involved in the parvulins' biological function.

To uncover the connection between the internal motions and the biological function of these two systems, I generated dynamic structural ensembles with externally restrained molecular dynamics simulations. Hereby I present the generation and analysis of structural ensembles of the PDZ1-2 tandem of PSD-95 and the analysis of previously generated ensembles of the parvulin-type SaPrsA, TbPin1 and CsPinA rotamase enzymes.

In the molecular dynamics simulations, I used experimental  $S^2$  order parameters and  $^1\text{H}$ - $^1\text{H}$  NOE distances as external restraints, which reflect motions on the fast, ps-ns time scale. In case of the PDZ tandem, only the intradomain motions were restrained by applying a local fitting algorithm. The most predominant internal motions were analyzed with principal component analysis, and the generated structural ensembles were compared to publicly available experimental structures.

When analyzing the generated structural ensembles, the opening-closing motion of the ligand binding site emerged as a possible general regulation mechanism in the investigated proteins. In the PDZ1-2 tandem, the binding partner alters the dynamics of the flexible  $\beta 2$ -  $\beta 3$  loop, which in turn is partially responsible for regulating the interdomain orientation. In parvulins, on the other hand, both the protonation state of the hydrogen-bonding network and the interaction between the WW domain and the hinge region allowing for the fluctuation of the binding site are modulating the opening-closing motion. The different biological activity of the distinct parvulin subtypes might be explained with their differences in the extent of this motion.

## Közlemények

E dolgozat a következő közlemények alapján készült:

Kovács Bertalan, Zajác-Epresi Nóra, Gáspári Zoltán (2020): “Ligand-dependent Intra- and Interdomain Motions in the PDZ12 Tandem Regulate Binding Interfaces in Postsynaptic Density Protein-95.” *FEBS Letters* 594(5):887–902.  
<https://doi.org/10.1002/1873-3468.13626>

Czajlik András\*, Kovács Bertalan\*, Perttu Permi, Gáspári Zoltán (2017): “Fine-Tuning the Extent and Dynamics of Binding Cleft Opening as a Potential General Regulatory Mechanism in Parvulin-Type Peptidyl Prolyl Isomerases.” *Scientific Reports* 7. 44504.  
<https://doi.org/10.1038/srep44504>

\* megosztott első szerzők

További közlemények:

Dudola Dániel, Kovács Bertalan, Gáspári Zoltán (2017): “CoNSEnsX+ Webserver for the Analysis of Protein Structural Ensembles Reflecting Experimentally Determined Internal Dynamics.” *Journal of Chemical Information and Modeling*, 57(8), 1728–1734.  
<https://doi.org/10.1021/acs.jcim.7b00066>

Dudola Dániel, Kovács Bertalan, Gáspári Zoltán. (2020): “Evaluation and Selection of Dynamic Protein Structural Ensembles with CoNSEnsX+.” In: *Structural Bioinformatics*, (Gáspári Zoltán szerk.) New York (NY), Amerikai Egyesült Államok : Springer US, (2020) pp. 241-254.  
[https://doi.org/10.1007/978-1-0716-0270-6\\_16](https://doi.org/10.1007/978-1-0716-0270-6_16) ISBN: 978-1-0716-0270-6

Továbbá 8 hazai és 8 nemzetközi konferenciamegjelenés.

# Ábrajegyzék

## Ábrák

1. ábra: Fehérjekomplexek szerveződése a PSD-ben (Neff et al., 2009), reprodukálva: (W. Feng & Zhang, 2009).
2. ábra: A néhány legfontosabb, PDZ domént tartalmazó fehérje felépítése (H.-J. Lee & Zheng, 2010).
3. ábra: A PDZ domén szerkezete és kölcsönhatása a ligandummal. **A:** a PDZ domén a PSD-95 PDZ1 példáján bemutatva. A hidrofób zsebhez tartozó aminosavak explicit módon vannak jelölve. A GLGF motívum 4 aminosavja sötétkéssel van kiemelve. A ligandum C-terminális aminosavjai pirossal vannak színezve. **B:** a C-terminális ligandum és a PDZ domén közötti kölcsönhatás. A ligandum atomjai pirossal vannak színezve. A GLGF motívum aminosavjai csillaggal vannak jelölve. A ligandum és a domén közötti hidrogénkötések kék, a hidrofób kölcsönhatások zöld szaggatott vonallal vannak jelölve. Reprodukálva Doyle és mtsai. alapján (Doyle et al., 1996).
4. ábra: A PSD-95 PDZ1-2 tandemjéről készült elérhető kísérletes szerkezetek. Az interdomén orientáció összehasonlíthatósága érdekében a szerkezetek közös PDZ1 templátra illeszkednek. A 3ZRT (Bach et al., 2012), 3GSL (Sainlos et al., 2010), valamint a 6SPV és 6SPZ (Ahmad Rodzli et al., 2019) kristályszerkezetek szabadon elérhetők a PDB adatbázisban (rcsb.org). A FRET-szerkezetet (J. J. J. McCann et al., 2011), valamint az abból származó nyitott és zárt szerkezetet (Yanez Orozco et al., 2018) a tanulmány szerzői személyes úton rendelkezésünkre bocsátották.
5. ábra: A peptidil-prolil cisz-transz izomerázok által katalizált reakció.
6. ábra: A PPIázok három családjának egy-egy képviselője (felső sor) és kompetitív inhibitoraik (alsó sor), a parvulinok képviselőjének, a Pin1-nek a szerkezete részletesen bemutatva. **A:** Ciklofilin A (PDB kód: 1M9Y). **B:** FKBP12, (PDB kód: 1B6C). **C:** Pin1, egy parvulin (PDB kód: 1PIN). Az ábrán elöl a katalitikus domén látható, háttérben szürkével a WW domén. Külön jelölve vannak a doménszerkezet főbb részei, valamint a hidrogénkötés-hálózatot kialakító aminosavak, beleértve a két konzervált hisztidint. **D:** a cyclosporin A, a ciklofilinek inhibitora. **E:** az FK506, az FKBP-k inhibitora. **F:** a juglone, a parvulinok inhibitora.
7. ábra: A peptidkötés két határszerkezete. A  $C_i$ - $N_{i+1}$  kötés kettőskötés jellege miatt  $Ca_i$  -  $C_i$  -  $N_{i+1}$  -  $Ca_{i+1}$  atomok egy síkba kényszerülnek.
8. ábra: Enzimreakciók lehetséges mechanizmusának sematikus ábrázolása. **A:** Az Emil Fischer-féle kulcs-zár modell. **B:** Indukált illeszkedés. **C:** Konformerszelekció.
9. ábra: Az amid N-H kötésvektorok gyors dinamikája az általános rendparaméterekkel ( $S^2$ ) kifejezve:  $S^2 = 1$  esetén egyáltalán nincs belső mozgás, míg  $S^2 = 0$  esetén a kötésvektor egyforma valószínűséggel mutat a tér bármely irányába.
10. ábra: A MUMO-algoritmus (minimal under-restraining, minimal over-restraining) lényegének sematikus ábrázolása. Az algoritmus során több, kényszerfeltételekkel megkötött párhuzamos replikát futtatunk. A  $S^2$  rendparaméter-megkötéseket mindegyik replikára egyformán alkalmazzuk, még az NOE távolság-kényszerfeltételeket párosával átlagolva olyan módon, hogy minden egyes replikára a két-két szomszédjával vett átlagot határozzuk meg (Richter et al., 2007).
11. ábra: A PSD-95 fehérje PDZ1 és PDZ2 doménjének szekvencia- és szerkezetillesztése. **A:** a PDZ1 és PDZ2 domén szekvenciaillesztése. A PDZ2 N-terminálisa (linker régió) és farkrégiója dőlt betűvel van jelölve. A GLGF motívumhoz tartozó aminosavak aláhúzással vannak kiemelve. A hidrofób kötőzseb aminosavjai félkövérrel vannak szedve. Alul a 2KA9 PDB szerkezetben lévő cypin ligandum C-terminális szekvenciája látható. **B:** A PDZ1 és PDZ2 domén szerkezetillesztése.

12. ábra: Intradomén távolság-kényszerfeltételekkel (NOE) megkötött és megkötések nélkül futtatott PDZ1-2 tandem tesztszimulációk összehasonlítása. A különböző színek a párhuzamos trajektóriákat reprezentálják. A nem megkötött sokaság 4, a megkötött sokaság 8 párhuzamos replikából készült. A megkötött sokaságokban az NOE kényszerfeltételeket párosával alkalmaztam (ld. 10. ábra.) A szimulációk 20 ns-ig futottak, a kiértékeléskor az első 2 ns elhagyásával. A főkomponenseket a szűkített doménhatárookra (4-88, illetve 99-183 aminosav) határozta meg.
13. ábra: A PDZ1 és PDZ2 domén belső mozgásainak összevetése. **A:** A teljes sokaságban szereplő PDZ1 és PDZ2 doménről készült főkomponens-analízis. **B:** Az egymásra illesztett átlagos PDZ1 és PDZ2 szerkezetek (csak a Ca atomok). **C-D:** Az első két főkomponensből számolt atomi elmozdulások amplitúdója az aminosav sorszám függvényében ábrázolva a PDZ1, illetve PDZ2 doménre.
14. ábra: Intradomén mozgások vizsgálata a PDZ1-2 tandemben. A főkomponens elemzést a szűkített PDZ1 doménre (4-88 aminosav), illetve a szűkített PDZ2 doménre (99-183 aminosav) végeztem el, ez utóbbiból kihagyva a flexibilis hurokrégiókat (105-109 és 117-127 aminosav). A kötőzseb nyitottságát a Leu77 és Phe17, valamint a Leu172 és Phe112 aminosavak Ca atomjai közötti távolsággal mértem. **A,C:** A szabad és a komplex sokaság eloszlása az első főkomponens és a kötőzseb nyitottság mentén, feltüntetve a korrelációt. **B,D:** Az első főkomponens által reprezentált mozgás két extrém konformációja, a legjobban elmozduló régiók jelölésével. A ligandumot fekete szalag jelöli.
15. ábra: A szabad (**A,C**) és a komplex (**B,D**) sokaság főkomponenseinek átfedése a teljes sokaság főkomponenseivel. A főkomponens-analízist az intradomén mozgásokra, a szűkített doménhatárokat (4-88, illetve 99-183 aminosav) felhasználva határozta meg.
16. ábra: Mindhárom generált PDZ1-2 tandem sokaság, egy közös PDZ1 templátra illesztve. Az rajzolószoftver számítási kapacitásának végessége miatt az ábrán csak minden tizedik replika szerepel.
17. ábra: **A:** A PDZ1 domének közös templátra való illesztését követő, PDZ2 domének rövidített szakaszán (99-183 aminosav) elvégzett főkomponens-elemzés ábrázolása az első két főkomponens mentén. **B:** A pontokra legjobban illeszkedő kör sugarának 0,8, illetve 1,1-szeres nagyítása 16 ekvidisztáns pontban mintavételezve, a mintavételezési pontokhoz legközelebb eső PCA-pontok kiemelésével. **C, D:** A kiemelt PCA-pontoknak megfelelő PDZ szerkezetek Ca atomjainak tömegközéppontja térben ábrázolva, és az azokra illesztett háromdimenziós kör.
18. ábra: interdomén szögek ábrázolása a PDZ1-2 tandemben. **A:** az interdomén szög meghatározása szemléletesen ábrázolva (ld. YYY fejezet). **B:** az itt bemutatott ( $\alpha$ ), valamint a Wang és mtsai. által publikált ( $\Phi$ ) módszerrel számolt interdomén szögek korrelációja az általunk előállított sokaságra meghatározva.
19. ábra: PDZ1-2 tandem szerkezetek kategorizálása az interdomén kontaktusok és távolság szerint. Az ábrák az interdomén főkomponensek szerint vannak ábrázolva (ld. szöveg), és a **D** panel kivételével a teljes sokaságot tartalmazzák. **A:** a teljes sokaság, az interdomén távolság szerint színezve. **B:** Szoros és laza szerkezetek eloszlása az interdomén PCA-ábrán. **C:** A teljes sokaság, az atom-atom kontaktusok száma szerint színezve. **D:** A szoros szerkezetekből előállított klaszterek.
20. ábra: PDZ1-2 tandem klaszterek jellemzése az interdomén interfész szempontjából. **A:** Az egyes klaszterek reprezentatív szerkezetei, közös PDZ1 templátra illesztve. **B:** A PDZ domén felületének felosztása öt különböző régióra. **C:** A PDZ domén felületének öt különböző régiója az aminosavsorszám függvényében ábrázolva. Külön jelölve vannak a PDZ domének másodlagos szerkezeti elemei. **D:** Az interdomén interfész kialakulása az egyes klaszterekben, az imént definiált régiók szerint színezve. Az ábrán szerepel a szabad és komplex szerkezetek aránya az egyes klaszterekben, illetve a két ligandum egymáshoz viszonyított orientációja.
21. ábra: Atom szintű interdomén kölcsönhatások az 5. klaszter példáján bemutatva. **A:** Az interdomén hidrogénkötések (kék) és hidrofób kölcsönhatások (piros), kiemelve az azokat kialakító aminosavakat. **B:** Az interdomén kölcsönhatási hajlandóság (ld. szöveg) az aminosavsorszám függvényében ábrázolva. Külön

- jelöltük azokat a PDZ régiókat, amelyek az 5. klaszterben az interfész kialakításában részt vesznek (ld. 20. ábra).
22. ábra: Az elérhető kísérletes PDZ1-2 tandem szerkezetek és az általunk előállított klaszterek közös interdomén PCA-ábrán ábrázolva.
23. ábra: Az általunk előállított teljes sokaság és a kísérletes FRET adatok közötti korreláció. Az általunk előállított sokaság egyes szerkezeteiből visszszámoltam azt a 10 interdomén atom-atom távolságot, amelyek a FRET-mérések során adódtak (Yanez Orozco 2018), és meghatároztam az adatpontok korrelációját a finomítás során előállt három adatkészlettel ( $\langle R_{DA} \rangle_1$ ,  $\langle R_{DA} \rangle_2$  és  $\langle R_{DA} \rangle_3$ ).
24. ábra: Az intra- és interdomén mozgások közötti összefüggés vizsgálata a PDZ1-2 tandemben. Az ábrán az általunk előállított klaszterek adatpontjai vannak ábrázolva a PDZ1- és PDZ2-intradomén mozgásaira meghatározott főkomponensek szerint, valamint az egyes klaszterek eloszlása az első főkomponens mentén.
25. ábra: Az elemzéshez felhasznált SaPrsA, TbPin1 és CsPinA parvulin-típusú cisz-transz izomerázok egymáshoz illesztett szerkezete és szekvenciája. Az illesztés a MAMMOTH-MULT (Lupyan et al 2005) programmal készült. **A:** A három fehérje reprezentatív szerkezete a  $C_\alpha$  atomok pozíciója szerint egymáshoz illesztve. A szerkezetek csak a konszenzus aminosavakat tartalmazzák. **B:** A hidrogénkötés-hálózatot alkotó 5 aminosav elhelyezkedése a TbPin1 szerkezetén bemutatva. A foszfátkötő hurok külön van jelölve ezen az ábrán, ugyanis ezt a régiót a három ábrázolt fehérje közül csak a TbPin1 tartalmazza. **C:** A három fehérje szekvencia-illesztése, külön kiemelve a hidrogénkötés-hálózatot alkotó aminosavakat.
26. ábra: A három vizsgált fehérjéből készült sokaságok egymásra illesztve (felső sor), illetve a megkötött, nem megkötött és az eredeti sokaságokat tartalmazó összevont sokaságokon elvégzett főkomponens-elemzés ábrázolása. Külön ábrák mutatják a teljes fehérjéken (középső sor), illetve csak a konszenzus szekvenciákon elvégzett (alsó sor) főkomponens-elemzés eredményeit.
27. ábra: A három különböző parvulin fehérjét tartalmazó egyesített sokaságok, és az azokra elvégzett főkomponens-analízis eredménye. Az elemzéshez csak a konszenzus szekvencia aminosavjait használtuk fel (ld. 25. ábra). A főkomponens-elemzést elvégeztük csak a megkötött, csak a nem megkötött, valamint a megkötött és nem megkötött sokaságokból álló kombinált sokaságra.
28. ábra: A megkötött parvulin-sokaságokra meghatározott első főkomponens és a kötőzseb nyitottsága közötti összefüggés. **A:** A megkötött és nem megkötött sokaságok első öt főkomponense közötti átfedés, külön feltüntetve az egyik kiugró értéket. **B:** Az első főkomponens által reprezentált leghangsúlyosabb mozgás. A két szerkezet a két extrém konformációnak felel meg. Piros színnel van jelölve az a két aminosav, amelyeket a kötőzseb nyitottságának definíciójához felhasználtam. **C:** A három sokaság eloszlása a kötőzseb nyitottsága és az első főkomponens szerint, feltüntetve a korrelációt is.
29. ábra: A három általunk előállított megkötött sokaságból, valamint a 100 kísérletes rotamáz doménből álló szerkezet 53 aminosavat tartalmazó szűkített konszenzus régiójának ábrázolása és az így létrejött sokaságon végzett főkomponens-elemzése. Az ábrán külön színnel vannak jelölve azon fehérjékből származó rotamáz domének, amelyek pontosan egy rotamáz, két rotamáz, vagy pedig egy rotamáz és egy WW domént tartalmaznak. **A:** Az adatpontok ábrázolása az első két főkomponens mentén. **B:** A három általunk előállított sokaság és a 100 kísérletes rotamáz egymásra illesztett szűkített konszenzus régiója. **C:** A 100 kísérletes rotamáz egymásra illesztett szűkített konszenzus régiója.
30. ábra: Az első két főkomponens négyzetes fluktuációja a szűkített konszenzus aminosavszám függvényében. A főkomponens-elemzést a három általunk előállított megkötött sokaságból, valamint a 100 kísérletes rotamáz doménből álló szerkezet 53 aminosavat tartalmazó szűkített konszenzus régióján végeztük el. A jobb oldali ábrán a Pin1 fehérje azon régiói vannak jelölve, amelyek az első két főkomponens által reprezentált belső mozgásban a legnagyobb fluktuációt mutatják.
31. ábra: A két konzervált hisztidin orientációja és pKa értéke. Az ábrázolt adatpontok a három vizsgált fehérjére előállított megkötött sokaságokhoz tartoznak. **A:** A hisztidin aminosav oldalláncának orientációja  $\chi_1$  és  $\chi_2$  torziós szögekkel definiálva. **B:** A két hisztidinhez tartozó  $\chi_1$  és  $\chi_2$  torziós szögek közötti

különbségek abszolút értékei, egymás függvényében ábrázolva. **C:** A két hisztidin pKa értéke, egymás függvényében ábrázolva. A predikció a PropKa programmal készült.

32. ábra: A három vizsgált fehérjéhez tartozó kiindulási szerkezetekben található két konzervált hisztidin protonáltsági állapota, és a protonáltsági állapot függvényében létrejövő hidrogénkötés-hálózat a nagy lebeny 5 érintett aminosav-oldallánca között (ld. 25. ábra).
33. ábra: A hidrogénkötés-hálózatban érintett 5 aminosavra elvégzett főkomponens-elemzés. Az érintett aminosavak a konszenzus-szekvencia szerint: His9, Cys/Asp40, Ser42, Thr/Ser79 és His84 (ld. 25. ábra). Az ábrázolt adatpontok a három vizsgált fehérjére előállított megkötött sokaságokhoz tartoznak. **A:** Az adatpontok eloszlása az 5, hidrogénkötés-hálózatban érintett aminosavra meghatározott első és második főkomponens mentén. **B:** A hidrogénkötés-hálózatra meghatározott első főkomponens és a három megkötött sokaságra meghatározott első főkomponens egymás függvényében ábrázolva, a korreláció feltüntetésével.

## Táblázatok

1. táblázat: A humán membrán-asszociált guanilát-kináz (MAGUK) fehérjecsalád tagjai. A leggyakrabban használt elnevezések dőlttel vannak szedve (Kim & Sheng, 2004; Kwan et al., 2016; Sheng & Sala, 2001).
2. táblázat: Egy fehérjemolekula jellegzetes belső mozgásai és az azokhoz tartozó belső korrelációs idők ( $\tau_c$ ) időskálája. A táblázat tartalmazza a fehérjemolekulák (mérettől függő) rotációs korrelációs idejének jellemző nagyságrendjét ( $\tau_c$ ) (Gáspári & Perczel, 2010).
3. táblázat: A 2KA9 szerkezethez tartozó sztereoszimmetrizált  $^1\text{H}$ - $^1\text{H}$  távolság-kényszerfeltételek megoszlása a PDZ domének, valamint a ligandumok között.
4. táblázat: A három előállított PDZ tandem-sokaság (oszlopok), valamint az azok kombinációjával alkotott további sokaságok (sorok), és ezek elnevezése.
5. táblázat: Az előállított PDZ1-2 tandem sokaságok validációja a kísérletes paramétereknek való megfeleléssel. Az RMSD értékeket a kiindulási szerkezet (PDB: 2KA9) számozása szerinti 1-91 és 96-186 aminosavakra határoztuk meg a PDZ1 és a PDZ2 domén esetén, csak a peptidgerinc  $\text{C}\alpha$ , N és C atomjait felhasználva. A korrelációt a visszszámolt kémiai eltolódással és  $\text{S}^2$  rendparaméterekkel 4.2. fejezetben bemutatottak szerint végeztem.
6. táblázat: Az előállított parvulin-sokaságok gerinc RMSD-je és megfelelése a kísérletes paramétereknek. Az eredeti sokaságok a 2JZV, 2LJ4 és 2RQS PDB kóddal jelölt szerkezetek voltak sorrendben az SaPrsA, TbPin1 és CsPinA fehérjékre. Az RMSD értékek meghatározásakor a peptidgerinc mindegyik atomját figyelembe vettük. A merev régiókra meghatározott RMSD értékek számolásakor elhagytuk a fehérjék viszonylag nagy flexibilitást mutató farokrégióját, így a fennmaradó régiók a 6-111, 3-115 és 6-97 aminosavakat tartalmazták sorrendben az SaPrsA, TbPin1 és CsPinA fehérjékre. Az RMSD értékeket a MOLMOL programmal, a kémiaieltolódás- és  $\text{S}^2$ -korrelációkat a CoNSEnsX programmal számoltuk. \*A CsPinA fehérje esetén a Gly97 aminosavra visszszámolt  $\text{S}^2$  értéket elhagytuk a korreláció számolásakor.

## Függelék

- F1. táblázat: A PSD-t felépítő fehérjék kategóriái és legfontosabb képviselőik. A második oszlop az egyes kategóriák százalékos előfordulási arányát tartalmazza, a harmadik oszlopban pedig a nagyobb, jól mérhető számban fellelhető fehérjék neve mellett lévő szám azok abszolút gyakoriságát jelzi egy átlagos méretű PSD-ben (Sheng & Hoogenraad, 2007; Sheng & Kim, 2011).
- F2. ábra: **A:** Lokális RMSD-k az aminosav sorszám függvényében, az eredeti PDB sokaságra, valamint az általunk előállított szabad, komplex és megkötött komplex sokaságra. A lokális RMSD számolásához a peptidgerinc N, C és  $\text{C}\alpha$  atomjait vettem figyelembe (ld. 4.3.1 fejezet). **B:** Távoli NOE távolság-



kényszerfeltételek száma aminosavanként. Azon aminosavak közötti NOE-kényszerfeltételeket tekintettem távolinká, amelyek legalább 2 távolságra vannak egymástól a szekvenciában.

- F3. táblázat: A PSD-95 fehérje felépítése. A doménhatárok kijelölése a BLAST (Altschul et al., 1990) keresőalgoritmus által is használt SMART (Letunic & Bork, 2018) adatbázis alapján történt.
- F4. ábra: Az egérből, patkányból és emberből származó PSD-95 teljes szekvencia-illesztése a Clustal Omega programmal. Külön jelölve vannak a doménhatárok, valamint azok az aminosavak, amelyekben különbözik a szekvencia.
- F5. táblázat: A PSD-95 fehérje PDZ1 és PDZ2 doménjéről készült, PDB adatbázisban elérhető kísérletes szerkezetek, valamint a hozzáférhető NMR-s kísérletes adatok (kémiai eltolódás:  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ , rendparaméter:  $S^2$ , valamint  $^1\text{H}$ - $^1\text{H}$  NOE-k). A nyilvánosan nem elérhető adatok zárójelben szerepelnek.
- F6. ábra: A megkötött komplex, és a nem megkötött komplex sokaság intradomén mozgásainak összehasonlítása a PDZ1 és PDZ2 doménekben. **A:** A PDZ1 és PDZ2 doménekre elvégzett főkomponens-analízis ábrázolása. **B:** A nem megkötött és a megkötött komplex sokaság interdomén főkomponenseinek átfedése. **C:** A megkötött komplex és a nem megkötött komplex sokaság első két főkomponense által reprezentált négyzetes fluktuációk az aminosav sorszám függvényében.
- F7. ábra: A három előállított PDZ1-2 tandem sokaságra elvégzett interdomén főkomponens-analízis (ld. 5.2.3. fejezet).
- F8. táblázat: A PDZ1-2 tandemről készített klaszterek mérete (konformerek száma), a visszszámolt kémiai eltolódások korrelációja és a globális RMSD értékek a peptidgerinc atomjaira meghatározva.
- F9. táblázat: A PDZ1-2 tandemről készített klaszterek mérete (konformerek száma), a komplex forma aránya, az interfészt kialakító régiók (ld. 20.B ábra), valamint a ligandumok pozíciója: orientációja egymáshoz, illetve az interdomén tengelyhez viszonyítva, és elhelyezkedésük (ld. F11. ábra).
- F10. ábra: A lehetséges ligandum-orientációk eloszlása a PDZ1-2 tandemben az interdomén PCA-ábrán. A ligandum-orientációkat a  $-2$  aminosav  $\text{Ca}$  atomjából a  $0$  aminosav  $\text{Ca}$  atomjába mutató vektorok által bezárt szög alapján definiáltam, a következő elv szerint.  $<45^\circ$ : parallel,  $>45^\circ$  és  $<135^\circ$ : merőleges,  $>135^\circ$ : antiparallel.
- F11. ábra: A PDZ1-2 tandemről készített klaszterek ábrázolásai. Fönt: az interdomén kölcsönhatási hajlandóság minden egyes aminosavra (ld. 5.2.5. fejezet). A hidrogénkötések kék, a hidrofób kölcsönhatások piros színnel vannak jelölve. Középen: minden egyes klaszter reprezentatív szerkezete, külön jelölve az interdomén hidrogénkötéseket kézzel, valamint az interdomén hidrofób kölcsönhatásokat pirossal. Alul: minden egyes klaszter reprezentatív szerkezete, a ligandum 4 C-terminális aminosavját pirossal jelölve. Mindegyik ábrán a PDZ1 domén bal, a PDZ2 domén jobb oldalon van.
- F12. táblázat: Az interdomén hidrogénkötések és hidrofób kölcsönhatások előfordulási aránya az egyes PDZ1-2 tandem klaszterekben. A legalább 10%-os előfordulási arányok zöld háttérrel vannak kiemelve. Csak olyan kölcsönhatások vannak feltüntetve a táblázatban, amelyek valamelyik klaszterben legalább 10%-ban megjelennek.
- F13. táblázat: A PDZ domén is a ligandum között létrejövő hidrogénkötések az egyes klaszterekben.
- F14. táblázat: Egyedi konformerek és a kísérletes FRET-adatok közötti korreláció. Az egyes konformerek a 7 klaszter reprezentatív szerkezetei, valamint a kísérletes szerkezete. A FRET-adatsorok 10-10 interdomén atom-atom távolságot tartalmaznak (Yanez Orozco et al., 2018) (ld. 5.2.6. fejezet).
- F15. ábra: Az egyes parvulin-típusú fehérjékből készült megkötött sokaságok leghangsúlyosabb belső mozgásainak átfedése a mindhárom sokaságot tartalmazó, kombinált sokaság mozgásaival. A főkomponens-elemzést a konszenzus szekvencia aminosavjaira végeztük el.
- F16. táblázat: A kísérletes parvulin-típusú rotamáz-doméneket tartalmazó teljes fehérjék listája. A táblázatban külön szerepel az egyes fehérjék egyedi Uniprot (uniprot.org) azonosítója, teljese neve, valamint hogy a

teljes fehérje összesen egy (rot) vagy két (rot + rot) rotamáz domént, illetve egy rotamáz és egy WW domént (rot + WW) tartalmaz.

- F17. táblázat: Az elemzéshez felhasznált, parvulin-típusú rotamáz doméneket tartalmazó kísérletes szerkezetek PDB kódjai. A szerkezetek mellet fel van tüntetve, hogy a teljes fehérje összesen egy (rot) vagy két (rot + rot) rotamáz domént, illetve egy rotamáz és egy WW domént (rot + WW) tartalmaz.
- F18. ábra: Az elemzéshez felhasznált 100 darab, parvulin-típusú rotamáz doméneket tartalmazó kísérletes szerkezet többszörös szekvencia-illesztése. Az illesztés a MAMMOTH-Mult programcsomaggal készült. A konszenzus szekvencia 53 aminosavat tartalmaz.

## Rövidítések listája

<b>Rövidítés</b>	<b>Magyar kifejezés</b>	<b>Angol kifejezés</b>
AMPA	2-amino-3-(5-metil-3-oxo-1,2-oxazol-4-il)propánsav	$\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
CaMKII	Ca <sup>2+</sup> /kalmodulin-függő proteinkináz-II	Ca <sup>2+</sup> /calmodulin-dependent protein kinase II
CASK	Ca <sup>2+</sup> /kalmodulin-függő szerin proteinkináz	Ca <sup>2+</sup> /calmodulin-dependent serine protein kinase
CDK	ciklin-dependens kináz	cyclin-dependent kinase
CS	kémiai eltolódás	chemical shift
DHR	-	discs-large homology region
DLG	-	discs large
DMD	diszkrét molekuladinamika	discrete molecular dynamics
ePSD	serkentő posztszinaptikus denzitás	excitatory postsynaptic density
EPSP	serkentő posztszinaptikus potenciál	excitatory postsynaptic potential
FCS	fluoreszcencia-korrelációs spektroszkópia	fluorescence correlation spectroscopy
FKBP	FK506 kötő fehérje	FK506 binding protein
FRET	Förster-féle rezonancia-energiaátadás / fluoreszcens rezonáns energiaátadás	Förster resonance energy transfer / fluorescence resonance energy transfer
GK	guanilát-kináz	guanylate kinase
GKAP	guaniáltkínáz-asszociált fehérje	guanylate kinase associated protein
GPCR	G-fehérje-kapcsolt receptor	G protein-coupled receptor
hPTP1E	humán protein-tirozin-foszfátáz 1E	human protein tyrosine phosphatase 1E
IPSP	gátló posztszinaptikus potenciál	inhibitory postsynaptic potential
ITC	izotermális titrációs kalorimetria	isothermal titration calorimetry
LTD	hosszútávú depresszió	long term depression
LTP	hosszútávú potenciáció	long term potentiation
MAGUK	membrán-asszociált guanilát kináz	membrane associated guanylate kinase
MAPK	mitogén-aktivált proteinkináz	mitogen-activated protein kinase
NMDA	N-metil-D-aszparaginát	N-Methyl-D-aspartate
NMR	mágneses magrezonancia	nuclear magnetic resonance

nNOS	neurális nitrogén-monoxid-szintáz	neural nitric oxide synthase
NOE	nukleáris Overhauser-effektus	Nuclear Overhauser Effect
PCA	főkomponens-elemzés	principal component analysis
PDB	fehérje-adatbázis	Protein Data Bank
PKA/PKC	proteinkináz-A/proteinkináz-C	protein kinase A / protein kinase C
PPIáz	peptidil-prolil izomeráz	peptidyl-prolyl isomerase (PPIase)
PSD	posztoszínaptikus denzitás	postsynaptic density
PSD-95	posztoszínaptikus denzitás fehérje-95	postsynaptic density protein 95
RDC	reziduális dipoláris csatolás	residual dipolar coupling
RMSD	az átlagos négyzetes eltérés négyzetgyöke	root mean square deviation
SPAR	gerincvelő-asszociált Rap GTPáz-aktiváló fehérje	spine-associated Rap GTPase activating protein
SynGAP	színaptikus Ras GTPáz-aktiváló fehérje	synaptic Ras GTPase-activating protein

# 1 Bevezetés

Az elmúlt két évtizedben egyre inkább tért hódított, s végül általánosan elfogadottá vált a szerkezeti biológiában az a szemléletmód, amely szerint a fehérjék biológiai funkciójában meghatározó szerepet játszik azok dinamikus volta. A fehérjemolekulákat merev, háromdimenziós szerkezettel jellemző modelleket szép lassan kiegészítették, sőt felváltották az olyan leírások, amelyek számot vetnek a belső mozgásokkal. Ebben a szemléletformálásban meghatározóak voltak bizonyos elméleti megfontolások (például a belsőleg rendezetlen fehérjék biológiai funkciójának felismerése vagy épp az allosztérikus szabályozás leírása a fehérjemolekula konformációs változásaival) és a technológiai fejlődés is, mind a fehérjemozgások vizsgálatára alkalmas kísérleti eljárások, mind a modellezéshez szükséges számítási kapacitás területén.

A korábbi felfogással szemben egy fehérjemolekula nem egyetlen stabil szerkezettel rendelkezik, hanem a termális fluktuáció eredményeképpen számtalan lehetséges konformációja van egymással dinamikus egyensúlyban, amelyek a potenciálisenergia-felületen egy-egy lokális minimumot képviselnek, egymástól kisebb vagy nagyobb potenciálgáttal elválasztva. A belső mozgások így lehetőséget biztosítanak több lehetséges konformáció felvételére, amelyek akár különböző biológiai funkcióval is rendelkezhetnek (például nyílt és zárt forma). A belső mozgások, illetve azok megváltozása ezenkívül jelentős entropikus hozzájárulással bírhatnak a kötőpartnerrel való kölcsönhatás energetikájában.

A kísérleti fehérjeszerkezet-kutató eljárások közül egyik sem ad közvetlen képet a fehérje belső mozgásairól: a dinamikus kép minden esetben a nyers mérési eredmények interpretációjának és egy modellalkotási folyamatnak az eredménye. Maguk a mérési körülmények is meghatározzák, mely belső mozgások vizsgálhatók az adott kísérlettel. A röntgenkristallográfiában például a molekulákat kristályos állapotban vizsgáljuk. Ez a fehérjemolekulák esetében azt jelenti, hogy egy adott szerkezet kimerevített állapotát térképezzük fel. Az atomi mozgástényezők, azaz B-faktorok a hurokrégiók fluktuációjáról és az oldalláncok gyors reorientációjáról szolgáltatnak némi közvetett információt, de ezektől eltekintve nem tudjuk megfigyelni a molekula különböző konformációit, illetve az azok közötti átmeneteket.

Az atomi felbontású szerkezetvizsgáló módszerek közül a mágneses magrezonancia (NMR) spektroszkópia a legszélesebb körben alkalmazott eljárás a fehérjemolekula különböző időskálán történő belső mozgásainak vizsgálatára. A mérés maga oldatfázisban történik, ami lehetőséget biztosít a fehérjének a natív környezetéhez leginkább hasonló, fiziológias közegben történő megfigyelésére. Bár a háromdimenziós szerkezet NMR-rel történő meghatározása kis és közepes méretű fehérjék esetén már legalább rutinszerű, de sok esetben automatizálható, a belső mozgások vizsgálatáról ez közel sem mondható el. A különböző időskálán történő mozgásokhoz különböző NMR-kísérleteket terveznek, és a közvetlenül mérhető információkból áttételesen következtetnek a fehérjedinamikára. Külön

megnehezíti a helyzetet, hogy a mérés időskálájánál gyorsabb mozgások kiátlagolódnak, és az eredmények értelmezésekor figyelembe kell vennünk, hogy a közvetlenül mérhető mennyiségek több lehetséges szerkezet hozzájárulásával jöttek létre.

Mind a fehérjemolekula mérete és bonyolultsága, mind az NMR-rel közvetlenül mérhető dinamikai paraméterek nehéz interpretálhatósága szükségessé teszi, hogy a kísérleti eredményeket számítógépes modellekkel és szimulációkkal egészítsük ki ahhoz, hogy képet kapjunk a fehérje belső mozgásairól. Ebben nélkülözhetetlen szerepet játszik a molekuladinamikai (MD) szimuláció, melynek használata már az NMR-rel történő fehérjeszerkezet-meghatározáshoz is szükséges. A számítási kapacitás növekedése és az atomi szintű kölcsönhatásokat leíró erők fejlődése mára már az NMR-es mérések állandó társává tették az MD szimulációt mint eszközt. Lehetőségünk van az NMR-rel mért információkat mint külső kényszerfeltételeket figyelembe venni a szimuláció során: ilyen módon a szimulált rendszert mintegy rákényszerítjük, hogy visszatükrözze a kémcsőben közvetlenül megfigyelt paramétereket.

A mérések és számolások eredményeképpen a fehérjét merevnek tekintő háromdimenziós szerkezetet egy olyan modellre cserélhetjük le, amely jobban tükrözi a molekula dinamikus voltát: ez a dinamikus fehérjeszerkezeti sokaság. Ez összhangban van azzal a feltételezésünkkel, amely szerint a fehérje számtalan konformáció sokaságaként jelenik meg, és a mérés során e lehetséges szerkezetek időátlagát figyeljük meg. A szerkezeti sokaság alapvetően különbözik az NMR-rel rutinszinten meghatározott szerkezetek sorozatától, amelyekkel nyilvános szerkezeti adatbázisokban találkozunk. Ez utóbbiakban ugyanis minden egyes szerkezettől azt várjuk, hogy megfeleljen a mérési eredményeknek, míg a sokasággal kapcsolatban nem élünk ilyen elvárással: ott az összes szerkezetből visszszámolt szerkezeti és dinamikai paraméterek átlagát hozzuk összefüggésbe a mérési eredményekkel.

Doktori értekezésemben kétféle fehérje (az idegsejtekben található PSD-95, valamint a parvulin-típusú peptidil-prolil cisz-transz izomerázok családja) példáján bemutatom azt a folyamatot, amely során a nyers NMR-es mérési eredményeket egy külső kényszerfeltétellel megkötött molekuladinamikai szimulációban felhasználva olyan modellt – dinamikus fehérjeszerkezeti sokaságokat – hozunk létre, amelyek pontosabb képet adnak a fehérjének egy bizonyos időskálán (ps-ns) történő mozgásairól, mint a nyilvános adatbázisban elérhető háromdimenziós szerkezetek. Mindkét fehérje esetén egy olyan, a dinamika modulálásával megvalósított szabályozó mechanizmust valószínűsítettünk, amely lokális kölcsönhatások által valósul meg. A létrehozott sokaságok elemzése olyan összefüggésekre világít rá e két fehérje belső mozgásai és biológiai funkciója között, amelyre a kísérletes dinamikai paraméterek és a számítógépes szimuláció integrálása nélkül nem lett volna lehetőség.

Különösen is kihívást jelent a többdoménes fehérjék belső dinamikájának leírása. A bevett nyelvi eszköztárunk a szerkezet-funkció paradigma mentén történő gondolkodásmódot tükrözi: a harmadlagos szerkezet fogalma egy stabil, feltekeredett domént, míg a negyedleges szerkezet több ilyen alegység

egymáshoz történő – szintén merev – kapcsolódását jelöli. Az utóbbi években tért hódított “szuperharmadlagos” vagy “szupramoduláris” szerkezet és dinamika fogalma ezzel szemben feltételezi annak lehetőségét, hogy egy adott polipeptidláncon belül előforduló domének egymáshoz viszonyított orientációja nem rögzített, hanem egy adott konformációs sokasággal írható le (Tomba, 2012a). A többdoménes fehérjék szupramoduláris dinamikájának mind a kísérletes vizsgálata, mind a modellszintű leírása újfajta megközelítést igényel, melyet a PSD-95 fehérje PDZ12 tandemjéről előállított szerkezeti sokaságok elemzésén keresztül mutatok be.

## 2 Irodalmi áttekintés

### 2.1 A PSD-95 N-terminális PDZ1-2 tandemjének szerkezete és belső dinamikája

#### 2.1.1 A serkentő posztszinaptikus denzitás (PSD)

Mind a központi, mind a környéki idegrendszer működése az egymással szinaptikus kapcsolatban lévő idegsejtek közötti jelátvitelre épül. Gerincesekben ennek legerjedtebb formája a kémiai szinapszis. Ennek során az ingerület hatására a preszinaptikus idegsejtben felszabadulnak a szinaptikus vezikulákban tárolt neurotranszmitterek, melyek a membránhoz való dokkolást és fúziót követően exocitózissal távoznak a két idegsejt közötti szinaptikus résbe, majd a posztszinaptikus sejt membránján található receptorokhoz kötődve valamilyen választ váltanak ki.

A posztszinaptikus denzitás (PSD) a posztszinaptikus sejt membránjához kötődő sűrű fehérjehálózat, amely az elektronmikroszkópos felvételeken szembetűnő, lemez alakú sötét foltként jelenik meg (M. B. Kennedy, 2000; Sheng & Hoogenraad, 2007; Ziff, 1997). Valójában egy rendkívül összetett jelátviteli komplexről van szó, amely méreténél és jelentőségénél fogva a neuronokra specifikus sejszervecskének tekinthető (W. Feng & Zhang, 2009; Sheng & Hoogenraad, 2007; Sheng & Sala, 2001).

A PSD egyaránt megtalálható a serkentő és gátló szinapszisokban is, ám ezek mind kémiailag, mind morfológiailag különböznek egymástól. A serkentő szinapszisa jellemzők a nagyobb, gömb alakú szinaptikus vezikulák. A gátló szinapszis esetén jóval vékonyabb a PSD, és a szinaptikus vezikulák is kisebbek és laposak. A serkentő szinapszisok legnagyobb részt axo-dendritikusak, vagyis a preszinaptikus vezikulák egy axonvégződésben találhatóak, a posztszinaptikus sejt pedig egy jellegzetes alakú kitüremkedéssel, ún. szinaptikus tüskével kapcsolódik az axonhoz. Gátló szinapszisok nem jellemzőek szinaptikus tüskéken: gyakrabban alakulnak ki a dendritek nyelénél, a szómánál, illetve az axonok kezdeti szakaszán (Sheng & Kim 2011). Gyakori előfordulása és jellegzetes morfológiája miatt sokkal behatóbban tanulmányozták a serkentő PSD-t, és a szakirodalom is részletesebben tárgyalja (Z. Feng et al., 2019; Sheng & Kim, 2011; Ziff, 1997). A továbbiakban a szakirodalmi gyakorlatnak megfelelően a PSD kifejezésen – ezzel ellentétes utalás hiányában – a serkentő posztszinaptikus denzitást (excitatory postsynaptic density - ePSD) értem.

A kémiai szinapszisa jellemző a szinaptikus plaszticitás, amely során megváltozik a két idegsejt között szinaptikus erősség, azaz hogy egyetlen jelátviteli folyamat mekkora potenciálváltozást okoz a posztszinaptikus sejt membránján. Ez a PSD méretével, különösen is a receptorok számával függ össze. Serkentő szinapszisokra jellemző a hosszútávú potenciáció (long-term potentiation - LTP). Ennek során az NMDA-receptorok kitartó stimulációjának hatására  $Ca^{2+}$ -ionok áramlanak a sejtbe, amelyek –



számos, rövid- és hosszútávú biokémiai válasz kiváltása mellett – megnövelik az sejtmembránba ágyazódott AMPA-receptorok számát, és ilyen módon növelik meg a szinaptikus erősséget. Ellenkező irányú folyamat a hosszútávú depresszió (long-term depression - LTD), amely során kitartó ingerület hiányában lecsökken az AMPA-receptorok száma, és ezáltal a szinaptikus erősség.

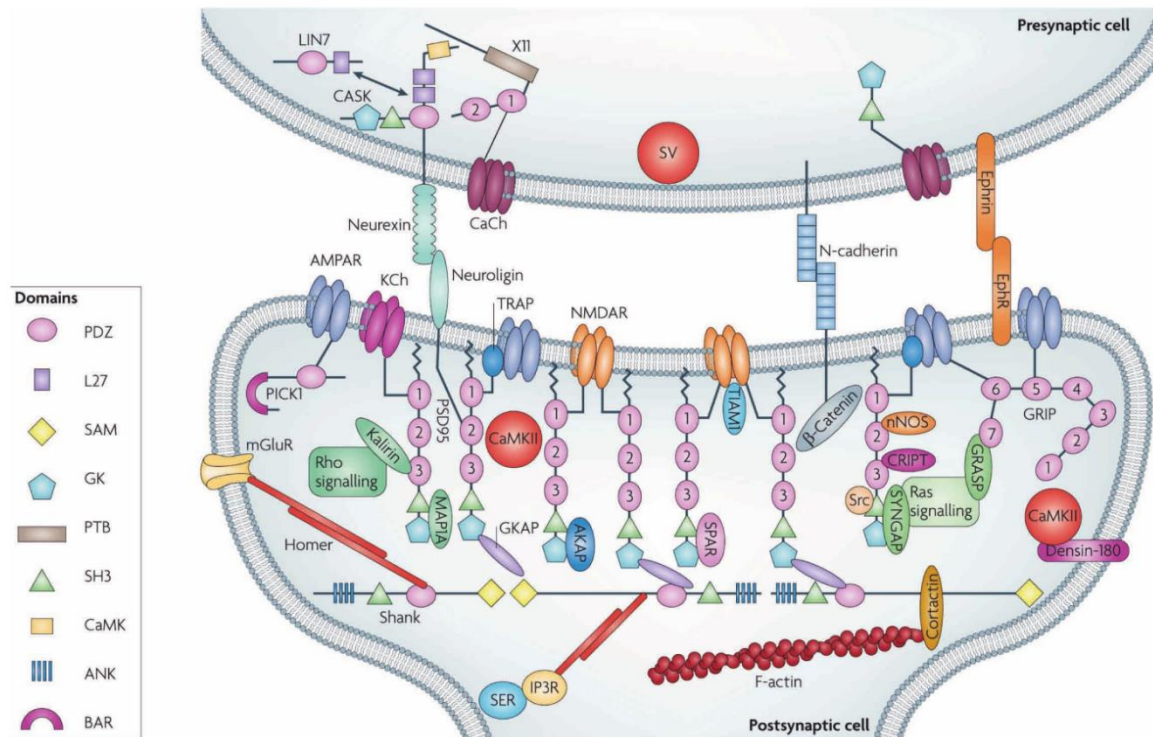
A PSD legfontosabb biológiai funkciója az idegsejtek közötti kommunikációért felelős membrán-asszociált molekuláris gépezet működtetése és finomhangolása. A PSD felel a membránfehérjék trafficking folyamatának és klaszterezésének szabályozásáért, valamint a jelátviteli fehérjekomplex felépítéséért. Számos jel arra utal, hogy a szinaptikus plaszticitás szabályozásán keresztül a PSD központi szerepet játszik a tanulásban és az emlékezésben, ugyanakkor e szabályozó folyamatnak a pontos molekuláris mechanizmusa egyelőre nagyrészt tisztázatlan (M. B. Kennedy, 2000; Kim & Sheng, 2004; Ziff, 1997). Az mindenesetre nagy bizonyossággal elmondható, hogy a PSD patológiás fejlődése vagy működése hozzájárul bizonyos neuropszichiátriai és neurodegeneratív betegségek (pl. skizofrénia, autizmus, valamint az Alzheimer-kór vagy a szélütés) kialakulásához (Coley & Gao, 2018; Gardoni et al., 2009; Lau & Zukin, 2007; Sheng & Hoogenraad, 2007; Sheng & Kim, 2011).

A PSD az egymástól 1-10  $\mu\text{m}$  távolságra lévő, 0,5-2  $\mu\text{m}$  hosszú dendritikus tüskék csúcán lokalizálódik (Sheng & Hoogenraad, 2007). Mérete tág határok között változik, mivel korrelál mind a tüskék hosszával, mind az aktív zóna nagyságával: szélessége 200-800 nm-re, vastagsága 30-50 nm-re tehető (W. Feng & Zhang, 2009; Sheng & Hoogenraad, 2007; Sheng & Sala, 2001). Egy átlagos, kb. 360 nm átmérőjű PSD molekulatömege  $1,10 \pm 0,36$  GDa-nak adódik (Sheng & Hoogenraad, 2007).

Az analitikai kémiai módszerek fejlődésével sorra jelentek meg a PSD kvalitatív és kvantitatív molekuláris összetételét taglaló kutatási eredmények. Bár az azonosított fehérjék száma mára már több mint ezerre rúg (Collins et al., 2006), pontosabb képünk mégis csak azokról a PSD-komponensekről van, amelyek a legnagyobb arányban alkotják azt, és ezért a kutatás története során már viszonylag hamar azonosították.

Az F1. táblázat a PSD-t alkotó fehérjék egy lehetséges kategorizálását, valamint az egyes kategóriákba tartozó fehérjék százalékos eloszlását tartalmazza. A leggyakrabban előforduló, így viszonylag könnyen azonosítható és mérhető fehérjék neve mellett az abszolút gyakoriságuk is szerepel egy átlagos méretű PSD-re vetítve. A PSD legfontosabb fehérjéinek molekuláris felépítését az 1. ábra mutatja be. A vázfehérjék elrendeződése és egymással kialakított kölcsönhatása biztosítja a PSD fehérjehálózatának megfelelő szerveződését. A vázfehérjék által alkotott fehérjehálózat három jelentősebb rétegre különíthető el (W. Feng & Zhang, 2009). Az első réteget alkotják a membránreceptorok, ioncsatornák és a transzmembrán sejtadhéziós molekulák. Hozzájuk kapcsolódnak a második réteget alkotó vázfehérjék, a membrán-asszociált guanilát kináz (MAGUK) fehérjecsald, melynek legfontosabb képviselője a PSD-95. Ezek a fehérjék a sejtmembránra merőlegesen helyezkednek el, és az N-terminálisukkal kapcsolódnak a membránhoz vagy a membránfehérjékhez. A harmadik rétegben további vázfehérjék találhatók, mint például a guanilát-kináz asszociált fehérjék

(GKAP), melyek a MAGUK fehérjék C-terminális SH3-GK doménjéhez kapcsolódnak, a Shank fehérjék, melyek a GKAP-hoz kötődnek, valamint a Homer, amelynek dimerjei egyfelől az mGlu-receptorokhoz, másfelől a Shank fehérjéhez kötődnek (W. Feng & Zhang, 2009; Z. Feng et al., 2019). A sejtmembrán harmadik rétegének fehérjei párhuzamosan szerveződnek, és szintén kötődnek a citoszkeleton további intracelluláris fehérjeihez, például az aktinhoz (W. Feng & Zhang, 2009; Z. Feng et al., 2019). A Homer kölcsönhat a sima felszínű endoplazmatikus retikulumban (SER) található inozitol-1,3,5-trifoszfát receptorral is, tehát adapterként viselkedik a szignáltranszdukció folyamatában (Sheng, 2001).



1. ábra: Fehérjekomplexek szerveződése a PSD-ben (Neff et al., 2009), reprodukálva: (W. Feng & Zhang, 2009).

Mind a PSD szerveződésében, mint a szinaptikus plaszticitás szabályozásában kulcsszerepet játszik a folyadék-folyadék fázisszeparáció, amely lényegében a vázfehérjék közötti rendkívül összetett és finomhangolt kölcsönhatásokkal van összefüggésben. A PSD-t nem határolja sejtmembrán, így az azt alkotó fehérjekomplex önszerveződését sokáig nem tudták mivel magyarázni. Az elmúlt évtizedben azonban egyre több kutatási eredmény igazolja, hogy a PSD leggyakoribb vázfehérjei az egymással kialakított szelektív kölcsönhatásoknál fogva asszociációra képesek és gyakorlatilag egy külön fázist alkotó kondenzátumot hoznak létre (Z. Feng et al., 2019; Zeng et al., 2016). Erre a kondenzációra már a négy leggyakrabban előforduló vázfehérje is képes: a PSD-95, a Shank, a GKAP és a Homer (Zeng et al., 2018). Az ilyen kondenzátumok klaszterezik a glutamát receptorokat, megnövelik a szinaptikus enzimek koncentrációját és szelektíven kilökik magukból a gátló posztzinaptikus denzitásra jellemző fehérjéket, mint például gephyrin vázfehérjét (Zeng et al., 2018).

A PSD molekuláris mechanizmusának motorjai a szabályozó- és jelzőfehérjék. A szinaptikus Ras GTPáz aktiváló fehérje (SynGAP), illetve a Rap GTPáz aktiváló fehérje (SPAR) az egyik legfontosabb útvonalban vesznek részt: a Ras/Rap jelátviteli útvonal szabályozásában játszanak szerepet a Ras, illetve a Rap enzimek GTPáz aktivitásának megnövelésével (M. B. Kennedy, 2000). A neurális nitrogén-monoxid-szintáz (nNOS) az NMDA receptorok által mediált  $Ca^{2+}$ -koncentráció növekedése által aktiválódik, és nitrogén-monoxidot termel, ami szintén egy fontos másodlagos jelzőmolekula a serkentő szinapszisokban. A  $Ca^{2+}$ /kalmodulin-függő proteinkináz (CaMKII) az NMDA receptorok szabályozásáért felelős, amit az NR2A és NR2B alegységeikhez való kötődéssel és foszforilációjával ér el (M. B. Kennedy, 2000; Mary B. Kennedy, 1997). Bár nem szinapszis-specifikusak, de mégis a legfontosabb szabályozó fehérjék közé tartoznak, ezért fontos megemlíteni a proteinkináz A (PKA) és proteinkináz C (PKC), a protein foszfatáz-1 (PP1) és a  $Ca^{2+}$ /kalmodulin-függő szerin proteinkináz (CASK) fehérjéket.

### 2.1.2 A PSD-95 és a MAGUK fehérjecsald

A posztzinaptikus állványfehérjék között különös jelentősége van a membrán-asszociált guanilát-kinázoknak (MAGUK), és közülük is kiemelkedik a PSD-ben egyik leggyakrabban előforduló MAGUK, a PSD-95. E fehérjecsald első tagját, a tumorszupresszor *discs large* (DLG) fehérjét a *Drosophila melanogaster*-ben azonosították, és megállapították, hogy a hámsejtek polaritásának szabályozásában játszik szerepet (Roberts et al., 2012). Humán homológjainak listáját az 1. táblázat tartalmazza. Történelmi okokból több különböző elnevezésük is elterjedt, mielőtt egyértelmű bizonyítást nyert volna rokonságuk.

A DLG humán homológjai	További elnevezések
DLG1	<i>SAP97</i> , hDlG
DLG2	<i>PSD-93</i> , chapsyn-110
DLG3	<i>SAP102</i>
DLG4	<i>PSD-95</i> , <i>SAP90</i>
<i>DLG5</i>	-

1. táblázat: A humán membrán-asszociált guanilát-kináz (MAGUK) fehérjecsald tagjai. A leggyakrabban használt elnevezések dőlttel vannak szedve (Kim & Sheng, 2004; Kwan et al., 2016; Sheng & Sala, 2001).

A MAGUK fehérjék homológiáját a nagyfokú szerkezeti azonosság bizonyítja: mindegyik három N-terminális PDZ domént, valamint egy SH3 és egy guanilát-kináz (GK) domént tartalmaz – a *SAP97* kivételével, amelyhez az imént felsoroltakon kívül egy L27 domén is tartozik (Kim & Sheng, 2004; Sheng & Kim, 2011; Sheng & Sala, 2001). Az első két PDZ domén, valamint a C-terminális SH3 és GK domén egy-egy szupramodult alkot (W. Feng & Zhang, 2009). A PDZ és SH3 domének más fehérjékben is gyakran előforduló fehérje-fehérje kölcsönhatási alegységek. Különösen érdekes, hogy a MAGUK fehérjék C-terminális doménje bár szerkezetileg homológ a GK doménnel, nem mutat kináz aktivitást, ugyanis az ATP-kötőhelyén található aminosavak nem konzerváltak (Craven & Brecht, 1998).

Ez is azt jelzi, hogy a MAGUK fehérjék elsődleges szerepe nem az enzimatis folyamatokban, hanem a molekuláris állványzat kialakításában van.

A PSD-95 egy átlagos PSD-ben mintegy 300 példányban fordul elő (Sheng & Hoogenraad, 2007), ezzel megelőzve a többi MAGUK fehérjét, de még az NMDA- és AMPA-receptoroknál is egy nagyságrenddel gyakoribb. Emiatt már három évtizede komoly érdeklődés övezi, és biológiai funkciójával számos publikáció foglalkozik. A PSD-95 minden jel szerint jóval fontosabb feladatokat is ellát a PSD molekuláris architektúrájának fenntartásánál, ám tekintve természetes kötőpartnereinek számát, pontos működése máig nem teljesen felderített.

A MAGUK fehérjék több PDZ domént is tartalmaznak, ami lehetővé teszi, hogy a PSD-95 is fontos szerepet játsszon a membránfehérjék klaszterezésében (J. F. Long et al., 2003). Ez a  $K_v1.4$  feszültségfüggő  $K^+$ -csatorna esetén bizonyítást nyert (Imamura et al., 2002), ugyanakkor a PSD-95 kiütése nem okozott különösebb eltérést az NMDA-receptorok elhelyezkedésénél - bár ez akár azzal is magyarázható, hogy a különböző MAGUK fehérjék részben átfedő biológiai funkcióval rendelkeznek, és így mások vették át a PSD-95 feladatát (Sheng & Sala, 2001).

A PSD-95 az adapterfehérje szerepet is betölt, ugyanis kapcsolatot teremt a membránreceptorok és a jelátviteli kaskád további intracelluláris fehérjéi között. Erre jó példa a NMDA-receptor-PSD-95-nNOS komplex (Kim & Sheng, 2004; Sheng & Sala, 2001).

Számos jel utal arra, hogy a hosszútávú potenciációban és depresszióban (LTP, LTD), melyek a tanulás és a memória mögött rejlő sejtszintű kulcsfolyamatok, szintén szabályozó szerepet tölt be a PSD-95 (Sheng & Sala, 2001). Az AMPA-receptorok alegységei nem tartoznak a PSD-95 kötőpartnerei közé, ám mivel az LTP-t az NMDA-receptorok kitartó stimulációja váltja ki, aminek egyik adapterfehérjéje a PSD-95, feltételezhető, hogy a PSD-95 valamilyen módon szabályozza az LTP és LTD indukciójához szükséges küszöbértéket (Kim & Sheng, 2004; Sheng & Sala, 2001). Számos kísérlet igazolta, hogy a PSD-95 kiütése drámaian lecsökkenti az LTP mértékét a hippokampuszban (Migaud et al., 1998).

Mindezekon felül a legfontosabb bizonyíték a PSD-95 központi szerepére a kötőpartnereinek száma (Kim & Sheng, 2004, 2009). A membránfehérjék közül kötődik NMDA-receptorhoz, kainátreceptorhoz, G-fehérje kapcsolt receptorhoz (GPCR), acetilkolin-receptorhoz és befelé irányító és feszültségfüggő  $K^+$ -ioncsatornákhöz. Kötődik a neuroliginhez, amely egy jellemző sejtadhéziós membránfehérje. Az intracelluláris jelző- és szabályozó fehérjék közül kölcsönhatásba lép az nNOS-sel, a SynGap-pal, SPAR-ral, Src-homológ kinázokkal vagy nem receptor tirozinkinázokkal, csak hogy a legfontosabbakat említsük. A vázfehérjék közül kötődik GKAP-hoz, AKAP79-hez, CRIPT-hez és a cypin-hez.

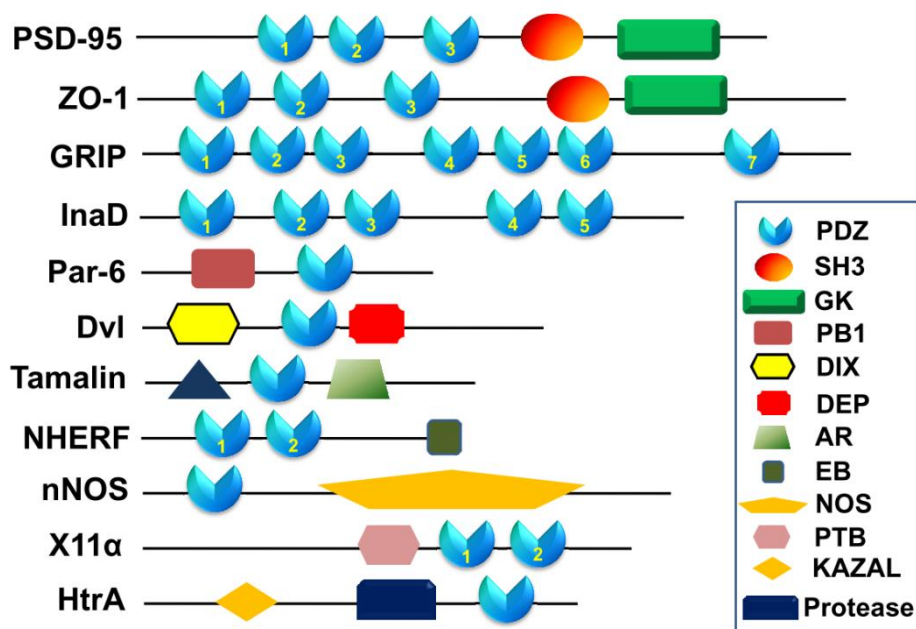
Több neurológiai rendellenesség molekuláris mechanizmusáról is bebizonyosodott, hogy vagy összefüggésben van a PSD-95 hibás működésével, vagy épp e rendellenességek kezelésében eredményesnek mutatkoztak olyan gyógyszerek, amelyek célpontja a PSD-95 volt. Ezek közé tartoznak

bizonyos neurodegeneratív betegségek (Alzheimer-kór és Parkinson-kór), a sztrók, a neuropátiás fájdalom, a skizofrénia és az autizmus (Coley & Gao, 2018; Gardoni et al., 2009).

### 2.1.3 PDZ domén szerkezete és belső dinamikája

A PDZ domén olyan gyakran fordul elő a poszt-szinaptikus fehérjékben, hogy akár a PSD kötőanyagának is tekinthető (Fanning & Anderson, 1999). Elsődleges funkcióját tekintve egy fehérje-fehérje kölcsönhatási alegység, így a PSD számos vázfehérjében több, akár 5-6 példányban is szerepel. Valójában a PSD óriási molekuláris komplexeinek megfelelő szerveződését irányító legalapvetőbb egység nem más, mint maga a PDZ domén (Craven & Brecht, 1998; Kim & Sheng, 2004).

A PDZ domén ugyanakkor nem korlátozódik a szinaptikus fehérjékre: számos olyan fehérjének is alkotóeleme, amelyek nem specifikusak a PSD-re (2. ábra). A teljes humán proteomban nagyjából 270 példányban fordul elő mintegy 150 fehérjében (Ernst et al., 2014; Luck et al., 2012). Szintén szerepet tölt be a szoros csomópontok létrehozásában, a hámsejtek polaritásának fenntartásában és jelútvonalak szabályozásában is (Harris & Lim, 2001; Hung & Sheng, 2002; Luck et al., 2012; Nourry et al., 2003).



2. ábra: A néhány legfontosabb, PDZ domént tartalmazó fehérje felépítése (H.-J. Lee & Zheng, 2010).

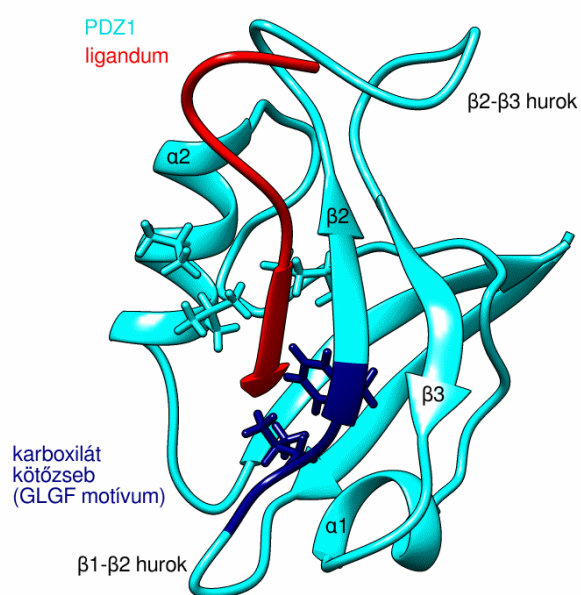
Felfedezése különböző fehérjék szekvencia-azonosságának elemzéséhez köthető. Több csoport egymástól függetlenül egy új, független szerkezeti alegységre bukkant, ám a rokonság ebben a fázisban még nem volt bizonyított. (Cho et al., 1992; Woods & Bryant, 1991). Az elnevezés egységesítésére először 1995-ben érkezett javaslat: a három olyan fehérje kezdőbetűjét összevonva nevezték el, amelyben legelőször azonosították. Ezek a PSD-95, DlgA (szoros csomópontokért felelős fehérje *Drosophila*-ban) és a ZO-1 (hámsejtek szoros illeszkedéséért felelős fehérje gerincesekben). Történelmi okokból néhány helyen DHR doménként (discs-large homology region) és GLGF ismétlődésként is

hivatkoznak rá – ez utóbbi név a domén ligandumkötő zsebében található magasan konzervált négy aminosavra utal.

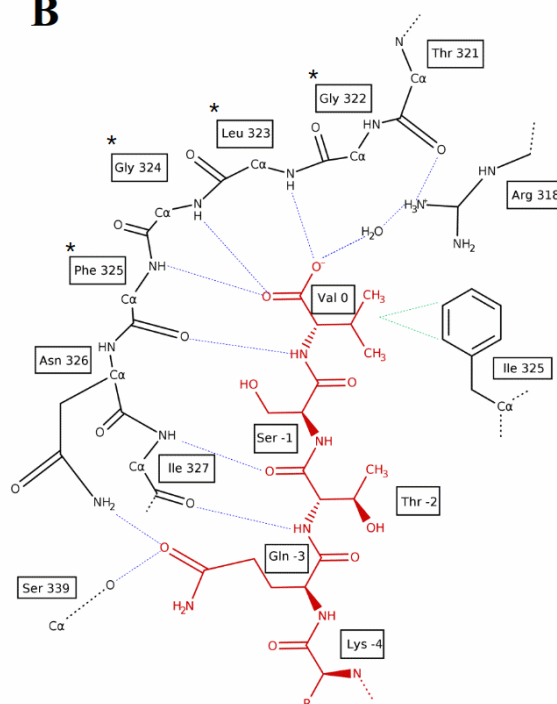
A kanonikus PDZ-fold nagyjából 90 aminosav hosszú, 6  $\beta$ -redőből áll, melyek egy félig nyitott hordóba rendeződnek, és ezt kétoldról egy-egy  $\alpha$ -hélix fogja közre (3.A ábra) (Doyle et al., 1996; Morais Cabral et al., 1996). A ligandum a  $\beta$ -hordó kiterjesztéseként kötődik a doménhez, egy extra antiparallel  $\beta$ -szálat alkotva a  $\beta 2$  szál mellett.

A PDZ domének túlnyomó többsége rövid, C-terminális szekvenciákat ismer fel, amelyek utolsó aminosavja szinte minden esetben apoláris (Doyle et al., 1996). Épp ezért van nagy jelentősége a konzervált GLGF motívumnak, amelyet karboxilát-kötőzsebek is szoktak hívni. Ezenkívül szintén fontos szerep jut a  $\beta 2$ -szálaban és  $\alpha 2$ -hélixben található apoláris aminosavaknak, amelyek a ligandum C-terminális aminosavjának apoláris oldalláncát stabilizálják (3. ábra).

**A**



**B**



3. ábra: A PDZ domén szerkezete és kölcsönhatása a ligandummal. **A:** a PDZ domén a PSD-95 PDZ1 példáján bemutatva. A hidrofób zsebez tartozó aminosavak explicit módon vannak jelölve. A GLGF motívum 4 aminosavja sötétkéssel van kiemelve. A ligandum C-terminális aminosavjai pirossal vannak színezve. **B:** a C-terminális ligandum és a PDZ domén közötti kölcsönhatás. A ligandum atomjai pirossal vannak színezve. A GLGF motívum aminosavjai csillaggal vannak jelölve. A ligandum és a domén közötti hidrogénkötések kék, a hidrofób kölcsönhatások zöld szaggatott vonallal vannak jelölve. Reprodukálva Doyle és mtsai. alapján (Doyle et al., 1996).

Kevésbé tipikus ugyan, de a PDZ domén képes a C-terminális motívumokon kívül más kötőpartnerekkel is kölcsönhatásba lépni. Ilyenek a foszfatidilinozitol-lipidek, amelyekhez a PDZ domén N-terminálisán keresztül kötődik, és így fontos szerepet játszanak a membrán-asszociált PDZ-fehérjék rögzítésében a lipid kettősréteghez (Nourry et al., 2003). Szintén ritkaságszámba menő kötési mód a belső szekvenciákkal kialakított kapcsolat. Erre példa a PSD-95 PDZ2 doménje, amely a nNOS

PDZ2-nek egy  $\beta$ -hajtú motívumát ismeri fel (Harris & Lim, 2001; N. X. Wang et al., 2008). És végül meg kell említeni a PDZ domének dimerizációját, amelyre a későbbiekben részletesen is kitérünk.

Hagyományosan a PDZ doméneket három osztályba szokás sorolni a kötőpartnereik C-terminálisán található 4 aminosav alapján (Hung & Sheng, 2002). Az I., II. és III. osztályba tartozó PDZ domének kötési motívumai rendre  $-X-[S/T]-X-\Phi$ ,  $-X-\Phi-X-\Phi$ , és  $-X-[D/E]-X-\Phi$ , ahol  $\Phi$  apoláris aminosavat, X pedig tetszőleges aminosavat jelöl. Ahogy a 3.B ábrán is látszik, a kölcsönhatás szempontjából fontos oldalláncok a ligandum C-terminális és -2 aminosavjához tartoznak, és ez a megállapítás összhangban van az imént bemutatott kötési motívumokkal is. A III. osztályba tartozó PDZ domének továbbá egy fontos szerkezeti különbséget mutatnak: 6 aminosavval rövidebb  $\beta$ 2- $\beta$ 3 hurokkal rendelkeznek (Imamura et al., 2002; Luck et al., 2012).

A ligandumkötés vizsgálata során azonban belátták, hogy nagyon valószínűtlen, hogy a PDZ domének rendkívül finomhangolt specificitását egyedül a ligandum -2 aminosavja határozza meg. Egyre részletesebb PDZ-családfák születtek, amelyek a GLGF-motívumon és a hidrofób zseben kívüli kontextust is figyelembe veszik a PDZ domének ligandumaffinitásának és -specificitásának magyarázatok (Ernst et al., 2014; Luck et al., 2012). Fontos megemlíteni azokat az egyre szaporodó kísérleti megfigyeléseket, amelyek a  $\beta$ 2- $\beta$ 3 hurok és a ligandum felső (-3-tól -7-ig) aminosavjai közötti kapcsolat jelentőségét igazolják (Mostarda et al., 2012; Tonikian et al., 2008). Később is látni fogjuk, hogy a  $\beta$ 2- $\beta$ 3 huroknak döntő szerep jut mind a ligandumspecificitás finomhangolásában, mind a PDZ domén belső dinamikájának szabályozásában.

Több mágneses magrezonancia kísérlet eredménye is arra enged következtetni, hogy a PDZ domének belső mozgásai a biológiai funkció integráns részét képezik. Az első, oldatfázisban meghatározott PDZ szerkezetből (a PSD-95 PDZ2 doménjéről), illetve a kísérleti adatokon elvégzett, az amid  $^1\text{H}$  és  $^{15}\text{N}$  atomokra mért relaxációs interpretációjára alkalmas Lipari-Szabó-féle modellfüggetlen elemzésből úgy tűnt, hogy a GLGF-motívumon kívül a fehérje gerince merev szerkezetű (Tochio et al., 2000). Ennek némileg ellentmondani látszik egy későbbi tanulmány, amelyben jelentős  $R_{\text{ex}}$  értékeket (kémiai kicserélődésből – *chemical exchange* – adódó relaxációs állandót) mértek a  $\beta$ 2- $\beta$ 3 hurokrégióra, amely annak lassú,  $\mu\text{s}$ -ms skálájú mozgására enged következtetni (Fuentes et al., 2004). Mindemellett azt találták, hogy a ligandum kötésekor a  $\beta$ 2- $\beta$ 3 hurok megmerevedése figyelhető meg, azaz mind a gyors, mind a lassú időskálán lecsökkennek a belső mozgások.

A PDZ doménben bizonyos aminosavak egy dinamikus hálózatot alkotnak. A humán protein-tirozin-foszfátáz 1E (hPTP1E) PDZ2 doménjén mért oldallánc  $S^2$ -adatokból arra következtettek, hogy a fehérjének a kötőzsebtől távoli, és azzal látszólag összeköttetésben nem lévő két régió aminosavjai a ligandumkötéskor szisztematikus oldallánc- $S^2$ -növekedést mutatnak, azaz a gyors mozgások intenzitásának csökkenése történik (Fuentes et al., 2004, 2006). A kísérletes  $S^2$ -paramétereket külső kényszerfeltételekként felhasználva egy kutatócsoport megkötött molekuladinamikai (MD) szimulációkat végzett, és így egészen konkrétan is azonosítani tudták, mely aminosavak tartoznak az



első, illetve a második hálózathoz (Dhulesia et al., 2008). Ezeknek az eredményeknek a segítségével az is kiderült, hogy a dinamikusan csatolt aminosavak többnyire statisztikailag is csatoltak, azaz a fehérje-evolúció során egyszerre konzerválódnak vagy mutálódnak (Dhulesia et al., 2008; Fuentes et al., 2004).

A belső dinamika szintén meghatározó szerepet tölt be a PSD-95 fehérje PDZ3 doménjének a ligandumaival kialakított kölcsönhatás szabályozásában. Ebben a doménben egy extra C-terminális hélix található, amely nem tartozik hozzá a kanonikus PDZ foldhoz, és nem is alakít ki közvetlen kölcsönhatást sem a ligandummal, sem a kötőzsebbel. Ennek ellenére a mesterségesen csonkított PDZ3 variáns, amely nem tartalmazza ezt a hélixet, 21-szeres csökkenést mutat a kötési affinitásban (Petit et al., 2009). Az oldallánc-S<sup>2</sup>-értékek ebben az esetben is azt bizonyították, hogy a hélix szabályozó szerepe egyfajta delokalizált entropikus mechanizmussal működik: a hélix hiányában nagyobbak az oldalláncok belső mozgásai, ami csökkenti a ligandumkötéskor létrejövő entrópianyereséget (Petit et al., 2009). Ezt az oldallánc-S<sup>2</sup>-értékekkel megkötött MD szimuláció is megerősítette (Mostarda et al., 2012). A C-terminális hélix dinamikája és a ligandumkötési affinitás közötti összefüggés arra is magyarázatot ad, hogy miért szabályozható a PSD-95 fehérje PDZ3 doménje a hélix foszforilációjával.

Mind a hPTP1E PDZ2 doménjének, mind a PSD-95 PDZ3 doménjének esete jó példát mutat arra, hogy a fehérje belső dinamikájának tanulmányozása ligandumkötés új mechanizmusait is feltárhatja, amelyben a kötőzsebtől távoli régiók gyors mozgásai lehetnek a meghatározóak. Külön figyelmet érdemel, hogy mindkét esetben pusztán az oldalláncok dinamikájának a változása detektálható, amely a fehérje szerkezetét vagy a gerinckonformáció változásait nem érinti.

A belső mozgásokkal kapcsolatos tanulmányok jelentős része bizonyítja, de legalábbis utal rá, hogy a  $\beta$ 2- $\beta$ 3 hurokrégió fontos szerepet tölt be a ligandum kötéskor: ezt vagy a ligandum felső aminosavjaival való közvetlen kölcsönhatással (sóhid, hidrogénkötések vagy apoláris kölcsönhatások), vagy pedig a belső dinamikában betöltött szabályozó szerepével éri el (Fuentes et al., 2004; Mostarda et al., 2012; Tochio et al., 2000). Ez nem is meglepő annak fényében, hogy a különböző  $\beta$ 2- $\beta$ 3 hurokkal rendelkező PDZ izoformák különböző kötési affinitásokat mutatnak ugyanazok ligandumok felé (Kozlov et al., 2002).

#### 2.1.4 PDZ domének alkotta szupramoduláris szerkezetek és dinamikájuk

A szekvenciában egymáshoz közel elhelyezkedő PDZ domének olykor egy független szerkezeti és funkcionális szupramoduláris egységet alkotnak, melyet PDZ tandemnek hívunk (W. Feng & Zhang, 2009). A domének ezekben az esetekben nem tekinthetők úgy, mint a gyöngyök egy füzéren, amihez gyakran hasonlítják a többdoménes vázfehérjéket: a PDZ tandem biológiai funkciója jellemzően több vagy más, mint a két domén funkciójának az összege (Ye & Zhang, 2013). Több esetben is a domének közötti linker mesterséges meghosszabbításakor a PDZ tandem funkciójának a csökkenése, szélsőséges esetben teljes megszűnése tapasztalható (J. F. Long et al., 2003).



A különböző PDZ tandemekben viszonylag széles skálán mozog a két domén közötti kölcsönhatás erőssége (W. Feng & Zhang, 2009; Ye & Zhang, 2013). A GRIP fehérjék PDZ1-2 és PDZ4-5 tandemje is olyan példa, amelyben az egyik PDZ domén funkciója egyfajta intramolekuláris chaperon-effektus, azaz a másik PDZ domén megfelelő szerkezetének biztosítása. A PDZ4-5 tandem például csak 1:1 sztöchiometriával képes kötődni a GluR2/3 receptor-alegységhez, ami azt sugallja, hogy a két domén között csak az egyik vesz részt a kölcsönhatásban (W. Feng et al., 2003; Q. Zhang et al., 2001). Azonban a doméneket külön-külön vizsgálva azt találták, hogy egyedül egyikük sem képes a ligandumkötésre. A PDZ4-nek torz az  $\alpha$ 2- $\beta$ 2 hurokrégiója, ami a kötőzsebet is érinti, a PDZ5 pedig oldatban teljesen rendezetlen. Érdekes módon csak akkor stabil mindkét domén, ha kovalensen kapcsolódnak egymáshoz, és a PDZ4 N-terminális kiterjesztése, amely kölcsönhatásba lép a kovalens linkerrel, szintén sértetlen. Hasonlóan viselkedik a PDZ1-2 tandem: a kristályszerkezet alapján PDZ2 kötőzsebet teljesen elfoglalja az interdomén interfész, így a ligandum csak a PDZ1-hez kötődhet (J. Long et al., 2008). Ugyanakkor a PDZ1 oldatban rendezetlen, és szintén kovalens kötésre van szüksége a megfelelő biológiai funkcióhoz.

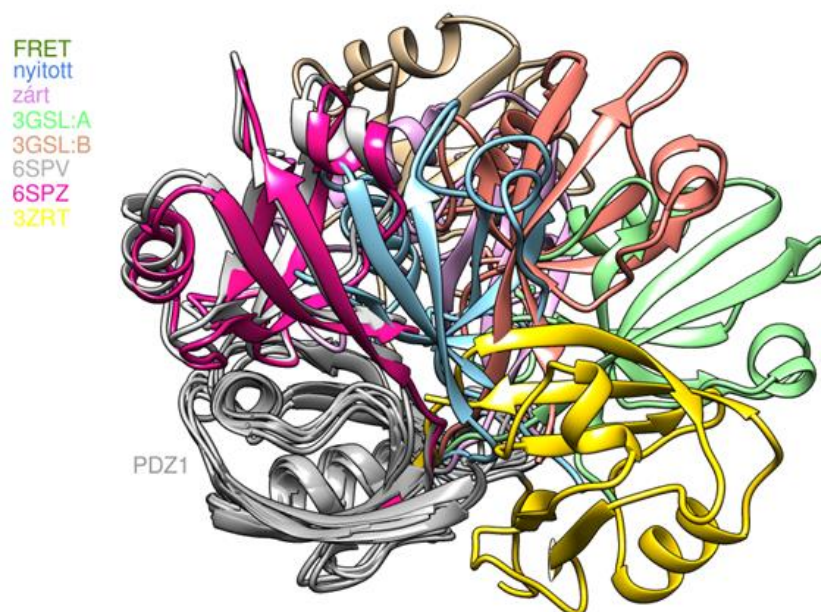
Az X11 és a humán syntenin fehérjék PDZ1 és PDZ2 doménje szintén egy-egy tandemet alkot. Ezek oldatban külön-külön is stabil domének, ám a tandem a kanonikustól eltérő kötési módot mutat. A syntenin esetében azt találták, hogy bizonyos ligandumok 2:1 sztöchiometriával kötődnek az izolált tandemhez, ám 1:1 sztöchiometriával a teljes fehérjéhez, sőt, bizonyos ligandumok egyáltalán nem kötődnek az izolált PDZ doménekhez, csak a tandemhez vagy a teljes fehérjéhez (Cierpicki et al., 2005; Kang et al., 2003). E jelenség oka, hogy a tandemhez tartozó N-terminus a kötéskor szabályozó szerepet tölt be. Nagyon hasonló a helyzet az X11 PDZ1-2 tandemjében, ahol a C-terminális farokrégió foszforilációja szabályozza a két domén ligandumkötését (J.-F. Long et al., 2005).

Az INAD a drosophila fotoreceptorokban található vázfehérje, amely a jelátvitelben tölt be szerepet. Ennek a fehérjének a PDZ4 és PDZ5 doménje is tandemet alkot. A fény erősségétől függően a PDZ5-ben található egyik diszulfid híd vált redukált és oxidált állapotai között. Azt találták, hogy ezt az átmenetet allosztérikusan szabályozza a PDZ4 domén: az izolált PDZ5 az oxidált állapotában stabil, ám a PDZ45 szupramodul a redukált állapotba kényszeríti a diszulfid hidat (Liu et al., 2011). A fény valójában, egy áttételes biokémiai kaszkád során, a két domén közötti kölcsönhatást gyengíti meg.

Talán a legérdekesebb PDZ szupramodul a PSD-95 PDZ1-2 tandemje. Mindeddig egymásnak részben ellentmondó eredmények születtek mind a szerkezetét, mind a dinamikáját illetően (4. ábra). Az első kristályszerkezetet közlő tanulmány szerzői egy olyan szupramoduláris szerkezetet javasoltak, amelyben a két PDZ domén kiterjesztett konformációban van, és a kötőzsebek egy irányba állnak (J. F. Long et al., 2003). Ezt a megfigyelést összhangba hozták a PSD-95 membránreceptor-klaszterező képességével, így ugyanis a két PDZ domén olyan orientációban lenne, amely lehetővé teszi két membránreceptor egyidejű kötését.

Ugyanebben a tanulmányban beszámolnak oldatfázisú NMR kísérletekről is (J. F. Long et al., 2003). Az eredményekből arra következtetnek, hogy a szupramoduláris szerkezet nem képes lényeges

konformációs változásokra. Ezt azzal indokolták, hogy a linker régióra mért  $^1\text{H}$ - $^{15}\text{N}$  heteronukleáris NOE-k (az NMR-aktív atommagok csatolásából adódó értékek) jóval magasabbak, mint a rendezetlen régióra várt értékek. Ugyanakkor a két domén között nincsen mérhető távolság-kényszerfeltétel, és a tandem HSQC (az amid N és H atomok csatolásait tartalmazó) spektruma csak igen kevésbé tér el a két izolált domén spektrumától, ami azt jelenti, hogy a két domén között csak nagyon csekély kölcsönhatás figyelhető meg.



4. ábra: A PSD-95 PDZ1-2 tandemjéről készült elérhető kísérletes szerkezetek. Az interdomén orientáció összehasonlíthatósága érdekében a szerkezetek közös PDZ1 templátra illeszkednek. A 3ZRT (Bach et al., 2012), 3GSL (Sainlos et al., 2010), valamint a 6SPV és 6SPZ (Ahmad Rodzli et al., 2019) kristályszerkezetek szabadon elérhetők a PDB adatbázisban (rcsb.org). A FRET-szerkezetet (J. J. J. McCann et al., 2011), valamint az abból származó nyitott és zárt szerkezetet (Yanez Orozco et al., 2018) a tanulmány szerzői személyes úton rendelkezésünkre bocsátották.

A későbbiekben három további kristályszerkezet készült el a PSD-95 PDZ1-2 tandemjére. Ezek PDB kódja 3ZRT (Bach et al., 2012), 3GSL (Sainlos et al., 2010), 6SPV a szabad, valamint 6SPZ a komplexált állapotra (Ahmad Rodzli et al., 2019). Ezekben a szerkezetekben mind különböző interdomén orientációt figyeltek meg (sőt, a 3GSL kristályszerkezetben az asszimmetrikus egységben lévő két PDZ tandem molekula egymástól is eltér), amelynek eredményeképpen nem azonos sem a ligandum orientációja, sem a két domén közötti kölcsönhatási interfész.

Egy oldatfázisú NMR-rel végzett kísérlet során azt vizsgálták, milyen szerkezetet és szupramoduláris dinamikát mutat a PSD-95 PDZ1-2 tandemje szabad, illetve komplex állapotában (W. Wang et al., 2009). Ehhez a mért  $^1\text{H}$  és  $^{15}\text{N}$  relaxációs időkből a Lipari-Szabó-féle modellfüggetlen elemzés segítségével oldallánc  $S^2$  adatokat számoltak. Azt találták, hogy a szabad állapotában a tandem merev, ám a ligandumkötés jelentős interdomén-reorientációt indukál. Abból a megfigyelésből jutottak erre a következtetésre, hogy a komplex formában a két doménre külön-külön meghatározott  $\tau_c$  rotációs-korrelációs idők egymástól valamelyest eltérnek (14,4 ms, ill. 12,5 ms). Abban az esetben, ha a tandem

teljesen merev lenne, illetve ha a két domén teljesen szabadon mozogna, mint egymástól független “gyöngyök egy füzéren”, akkor azonos rotációs korrelációs időket mértek volna.

Részben azért, hogy ezeket az ellentmondásokat feloldják, részben pedig, hogy azonosítsák a szupramoduláris orientációt szabályozó interdomén kölcsönhatásokat, megvizsgálták a PSD-95 fehérjék, és a külön annak PDZ1-2 tandemjét is a Förster-féle rezonancia-energiaátadás (FRET) módszerével is (J. J. J. McCann et al., 2011). Ezekkel a kísérletekkel egyrészt bebizonyították, hogy bármilyen orientációt, illetve interdomén dinamikát mutat az izolált PDZ1-2 tandem, ezt a teljes fehérjében is megtartja. Sőt, arra jutottak, hogy a teljes PSD-95 fehérje valójában két szupramodulból, a PDZ1-2 tandemből és a PDZ3-SH3-GK-szupramodulból áll (J. J. McCann et al., 2012). Fluoreszcencia-korrelációs spektroszkópiával (FCS) azt is megmutatták, hogy ténylegesen történik szupramoduláris reorientáció a PDZ1-2 tandemben, mégpedig a  $10^{-5}$  s-os (lassú) időskálán. Megállapították, hogy a linker mutáns formái nem befolyásolják a reorientációt. Ugyanakkor az izotermális titrációs kalorimetria (ITC) nem mutatott ki mérhető kölcsönhatási energiát a két izolált PDZ domén között, tehát ha van is ilyen kölcsönhatás, akkor az rendkívül csekély, “ultra-gyenge”, a szerzők kifejezését használva.

A FRET-mérések eredményeképpen egy újabb PDZ1-2 tandem szerkezet született, amely azonban továbbra sem volt összhangban a kristály és NMR-szerkezetekkel (J. J. J. McCann et al., 2011). Ezért később ezt a szerkezetet a kísérleti FRET-adatokból kiindulva diszkrét molekuladinamikai szimulációval (DMD) modellezték (Yanez Orozco et al., 2018). Ennek során különböző interdomén konformációk egész sorozatát kapták, amelyből két energiaminimum bontakozott ki. A két szerkezetet, amely ehhez a két minimumhoz tartozik, “nyitott” és “zárt” szerkezetnek hívták, az interdomén interfész nagysága alapján megkülönböztetve őket ( $440 \text{ \AA}^2$ , valamint  $701 \text{ \AA}^2$ ). Ez a két forma, bár nem fedik le a teljes interdomén konformációs teret, jól példázza a szupramoduláris átrendeződés működését, ugyanis azonosíthatók bennük azok a tranziens, ultra-gyenge kölcsönhatások, amely stabilizálják a PDZ tandemet az egyes energiaminimumokban. Ezek nyitott konformáció esetében főleg sóhidak, a zárt konformáció esetében hidrofób kölcsönhatások.

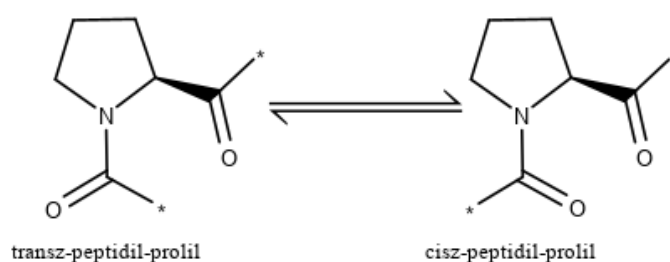
Összefoglalásként a PSD-95 PDZ12 tandemjéről álljon itt egy idézet Mark Bowen kutatócsoportjának 2011-es publikációjából, amely jól illusztrálja az eddigi ismereteink töredezettségét:

*“At mature excitatory synapses, PSD-95 controls localization of three different glutamate receptor subtypes; all have been shown to be dependent on an interaction with the PDZ tandem. How PSD-95 selects between ligands and recruits the appropriate signaling machinery remains a mystery.”* (J. J. J. McCann et al., 2011)

## 2.2 A peptidil-prolil izomerázok

### 2.2.1 A peptidil-prolil izomerázok sejtbiológiai szerepe

A peptidil-prolil izomerázok (PPIázok) vagy rotamázok (EC 5.2.1.8) olyan enzimek, amelyek egy fehérjében vagy polipeptid láncban egy adott prolin N-terminális peptidkötésének cisz-transz izomerizációját katalizálják (5. ábra) (Hanes, 2015). A spontán izomerizáció időállandója a 10-100 s-os időskálára tehető, így ez a fehérjefeltekeredés sebességmeghatározó lépése (Alderson et al., 2018; Matena et al., 2018; Nguyen et al., 2010; Schmid et al., 1986). A PPIáz enzimek kb. 8 kcal/mol-lal csökkentik az izomerizáció aktiválási energiáját, és így annak sebességét akár 3-6 nagyságrenddel is növelhetik (Eberhardt et al., 1992; Hanes, 2015; Kofron et al., 1991).



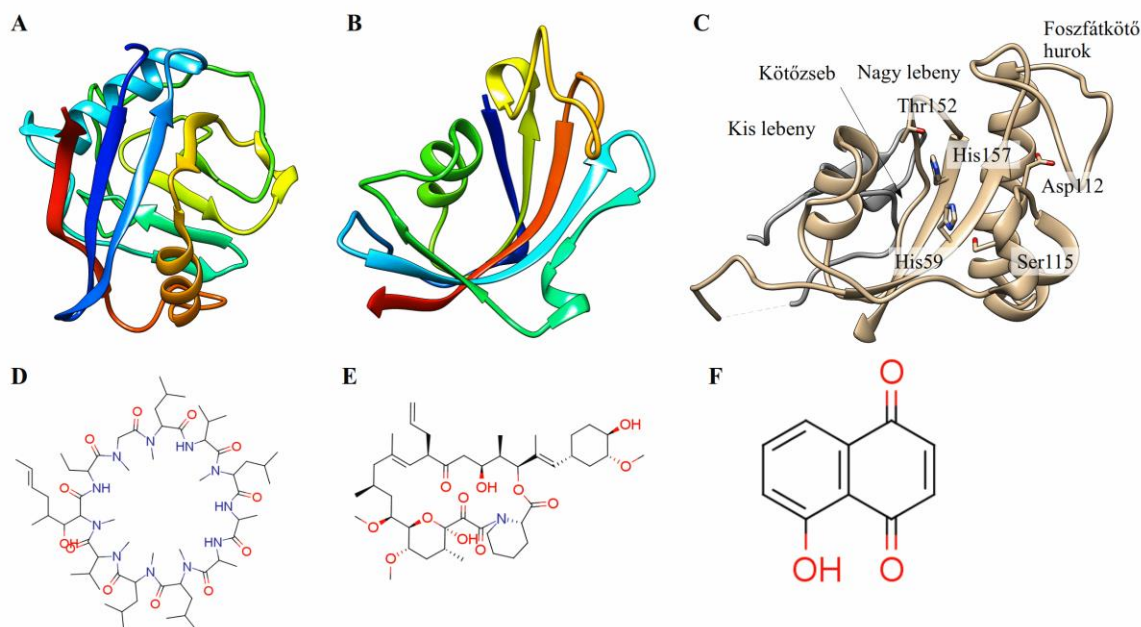
5. ábra: A peptidil-prolil cisz-transz izomerázok által katalizált reakció.

A PPIázok csoportjába három nem homológ fehérjecsald tartozik: a ciklofilenek, az FK506-kötő fehérjék vagy FKBP-k és a parvulinok (6. ábra). Az archeáktól kezdve a főemlősökig szinte minden élőlényben megtalálhatók, és az eukarióta sejten belül is mindegyik sejtalkotóban jelen vannak, ideértve a citoplazmát, az endoplazmatikus retikulumot, a mitokondriumot, a nukleuszt és a nukleoluszt. A humán genomban valószínűsíthetően 20 ciklofilint, 15 FKBP-t és 3 parvulint kódoló gén található (Nath & Isakov, 2015).

A peptidil-prolil cisz-transz izomerázok az 1984-es felfedezésüket követően hamar az érdeklődés középpontjába kerültek, ugyanis bebizonyosodott róluk, hogy komplexet képeznek olyan immunosuppresszáns kismolekulákkal, amelyeket a gyógyászatban transzplantáció után alkalmaznak a kilökődés megelőzése céljából (G Fischer et al., 1984; Gunter Fischer et al., 1998; Lang et al., 1987; Rahfeld, Rücknagel, et al., 1994; Rahfeld, Schierhorn, et al., 1994). A PPIázok két családját a kompetitív inhibitoraikról nevezték el: a ciklofilenek a ciklosporin A-val, az FKBP-k pedig az FK506-tal és a rampamicynnel alakítanak ki komplexet, és az immunválaszban betöltött szerepük alapján összefoglaló néven immunofilekként is hivatkoznak rájuk.

A parvulinokat csak később azonosították olyan enzimként, amelyeknek szintén van peptidil-prolil izomeráz aktivitásuk, ám az előbbi két csoporttól eltérően nem kötnek immunosuppresszáns molekulákat (Rahfeld, Rücknagel, et al., 1994; Rahfeld, Schierhorn, et al., 1994). Az elnevezés az első, *E. coli* baktériumból izolált parvulinmolekulának a másik két család tagjaihoz viszonyított kicsi méretére utal (*parvulus* - kicsi): a parvulinok mindössze kb. 10 kDa molekulatömegűek, szemben a kb.

18 kDa molekulatömegű FKBP-kkel és a kb. 12 kDa molekulatömegű ciklofilinokkal. A parvulinok inhibitora az 5-hidroxi-1,4-naftaléndion, hétköznapi nevén *juglone* (Göthel & Marahiel, 1999). Az elsőként azonosított humán parvulin-típusú enzim a Pin1, amelyről bebizonyosodott, hogy az aktivitása foszforiláció-függő: olyan peptidkötések izomerizációját katalizálja, amelyekben a prolint megelőző aminosav foszfo-szerin vagy foszfo-treonin.



6. ábra: A PPIázok három családjának egy-egy képviselője (felső sor) és kompetitív inhibitoraik (alsó sor), a parvulinok képviselőjének, a Pin1-nek a szerkezete részletesen bemutatva. **A:** Ciklofilin A (PDB kód: 1M9Y). **B:** FKBP12, (PDB kód: 1B6C). **C:** Pin1, egy parvulin (PDB kód: 1PIN). Az ábrán elől a katalitikus domén látható, háttérben szürkével a WW domén. Külön jelölve vannak a doménszerkezet főbb részei, valamint a hidrogénkötés-hálózatot kialakító aminosavak, beleértve a két konzervált hisztidint. **D:** a cyclosporin A, a ciklofilinok inhibitora. **E:** az FK506, az FKBP-k inhibitora. **F:** a juglone, a parvulinok inhibitora.

A PPIázok közé tartozó három család nem mutat olyan szintű szekvenciaazonosságot, amelyből evolúciós rokonságra lehetne következtetni. A harmadlagos szerkezeteik meghatározását követően azonban találtak hasonló elemeket a feltekeredési mintázatokban, ami felveti mind a konvergens evolúciós fejlődés, mind egy közös katalitikus mechanizmus lehetőségét (Nath & Isakov, 2015). A ciklofilinok egy 8 antiparallel  $\beta$ -szálból álló  $\beta$ -hordót alkotnak, amelyet három  $\alpha$ -hélix és hurkok kötnék össze (Davis et al., 2010; Göthel & Marahiel, 1999; Ke, 1992). Az FKBP-k 5 antiparallel  $\beta$ -szálból és egy  $\alpha$ -hélixből, míg a parvulinok 4 antiparallel  $\beta$ -szálból és 4  $\alpha$ -hélixből állnak (Göthel & Marahiel, 1999; Matena et al., 2018; Somarelli et al., 2008). Hasonlóságot találtak a ciklofilinok és az FKBP-k aktív helyében: mindkét régió tartalmaz például két konzervált triptofánt, amelyek mindkét család katalitikus aktivitásához szükségesek (Denesyuk et al., 1993; Nath & Isakov, 2015). Ezenkívül a parvulinok és FKBP-k topológiája nagyon sok hasonló elemet tartalmaz, amire több helyen *FKBP superfold*ként is hivatkoznak (G. Fischer & Aumüller, 2003; Sekerina et al., 2000).

A szuperharmadlagos szerkezet kapcsán mindhárom PPIáz családról elmondható, hogy egyaránt előfordulnak egyszárműves és többszárműves formában (Hanes, 2015). A katalitikus domén mellett

megjelenő többi domén általában a kötőpartnerrel való kölcsönhatás kialakulását segíti. Külön figyelemre méltó, hogy számos parvulinban, köztük a Pin1-ben is a C-terminális katalitikus domén mellett megjelenik egy N-terminális WW-domén, amiről sokáig azt feltételezték, hogy a foszforilált pSer/Thr-Pro motívumok felismerésében és kötésében játszik szerepet (Hanes, 2015; Ranganathan et al., 1997). Ez alapján megkülönböztetünk foszforilációfüggő Pin1-típusú, és foszforilációfüggetlen nem Pin1-típusú parvulinokat. Érdekes módon előfordulnak olyan Pin1-típusú parvulinok, amelyek nem tartalmaznak WW-domént (Sun et al., 2012; Yao et al., 2001).

Mivel a prolin N-amidkötésének cisz-transz izomerizációja a fehérjelánc konformációs mozgásaihoz képest lassú folyamat, ezért a PPIáz molekulákat legelőször a fehérjefeltekeredésben betöltött segítő funkcióval hozták kapcsolatba (G. Fischer & Aumüller, 2003; Wedemeyer et al., 2002). A fehérjék megfelelő feltekeredéséhez szükséges, hogy a peptidláncban lévő összes prolin a natív konformációba kerüljön. Egy tanulmány a ribonukleáz A fehérje példáján mutatta be a különbséget az esszenciális és nem esszenciális prolinok között: az előbbiek nem natív konformációja drasztikus torzulást tud okozni a fehérje harmadlagos szerkezetében (Wedemeyer et al., 2002). A PPIáz katalitikus hatása szükségessé válik a megfelelő feltekeredési útvonal bejárásához. Ugyanakkor több PPIáz enzimről bebizonyosodott, hogy nem önmagában katalizálja a feltekeredést, hanem együttműködve más dajkafehérjékkel (ún. *chaperonokkal*), vagy éppen maga a PPIáz enzim tartalmaz egyúttal chaperon domént is. Az előbbi kategóriába tartozik az FKBP51, FKBP52 és ciklofilin 40, amelyek közvetlenül a hsp90 (heat shock protein) chaperon fehérjéhez kötődnek (Schiene-Fischer & Yu, 2001), az utóbbiba pedig a trigger factor fehérje, SlyD (sensitive to lysis D) fehérje vagy az FkpA (Jakob et al., 2009). Bizonyos PPIázok a saját feltekeredésüket katalizálják (autokatalízis) (Göthel & Marahiel, 1999).

Egy-egy enzim aktív centrumában található pre-prolin peptidkötésének izomerizációja az adott enzim részleges le- és feltekeredését, ezáltal pedig ki- és bekapcsolását eredményezheti. A pre-prolin amidkötés cisz-transz oszcillációja kielégíti azokat a feltételeket, amelyek alapján molekuláris kapcsolóként tud működni: (i) két állapota lehetséges, (ii) ellenáll a véletlenszerű átkapcsolásnak az izomerizáció magas aktiválási energiája miatt, (iii) a kapcsoló további jelek felerősítésére képes (emelőrúd-effektus) és (iv) egyéb enzimek (jelen esetben a PPIázok) *in vivo* szabályozzák a kapcsoló állapotát. Mindezek alapján a PPIáz enzimek szerepet játszanak a sejtciklus szabályozásában, a jelutakban, a génexpresszióban, az immunválaszban, neuronális folyamatokban és több esetben a fertőzésben is (Göthel & Marahiel, 1999; K. P. Lu et al., 2007).

A PPIázok mint molekuláris időzítők szerepet játszhatnak a sejt folyamatok hosszúságának és amplitúdójának szabályozásában. A sejtciklus szabályozása egyes enzimek foszforilációjával, defoszforilációjával vagy ubiquitinációjával (a degradációra való előkészítésével), illetve az enzim szerkezetének ezáltal történő módosításával történik. A PPIáz enzimek többek között ezekben a folyamatokban töltenek be irányító szerepet (Göthel & Marahiel, 1999; K. P. Lu et al., 2007). A Pin1-nek például fontos szerep jut a sejtciklus során a G2/M mitotikus átmenet szabályozásában (Göthel & Marahiel, 1999; K. P. Lu et al., 2007). A mitotikus fázisok időzítését a ciklin-dependens kináz (CDK)

enzimek aktiválódása és inaktiválódása határozza meg. A Pin1 kölcsönhat számos mitózis-specifikus foszfofehérjével, például a Cdc25C foszfatázzal, ami a CDK1 defoszforilációjával a mitózis metafázisának beindítását szabályozza. Néhány újabb eredmény szerint a Pin1 a ciklin D1 aktiválásával szerepet játszik a G0/G1 és S fázis közti átmenet irányításában is (K. P. Lu et al., 2007). A Pin1 továbbá a Raf-Ras-MAPK kaszkád több szintjén is beavatkozik a jelút szabályozásába a foszforiláció-függő peptidil-prolil kötés izomerációjával (K. P. Lu et al., 2007).

A PPIáz enzimek sokrétű módon befolyásolják a génexpresszió folyamatát (Göthel & Marahiel, 1999; Hanes, 2015; K. P. Lu et al., 2007; Shaw, 2002). Mindhárom PPIáz család valamilyen módon összefüggésbe hozható a génexpresszió valamely fázisával: (i) szabályozzák a transzkripciós faktorok lokalizációját, stabilitását, valamint kötődését a DNS-hez, (ii) szintén szabályozzák az RNS-polimeráz működését, és (iii) módosítják a kromatin szerkezetét (K. P. Lu et al., 2007). Mivel a Pin1 által kötött motívumok könnyen azonosíthatók (pThr/Ser-Pro), és ezeknek az aminosavaknak a foszforilációját a mindenütt jelenlévő ciklin-dependens (CDK) és mitogén-aktivált kinázok (MAPK) végzik, számos, a transzkripció folyamatában résztvevő fehérjéről bebizonyították, hogy kölcsönhatásba lép a Pin1-gyel. A Pin1 többek között modulálja a  $\beta$ -katenin, a NF- $\kappa$ B és a c-Myc transzkripciós faktorok stabilitását, promotálja a ciklin-D nukleáris lokalizációját, valamint kötődik az RNS-polimeráz legnagyobb alegységének C-terminális doménjéhez (CTD) (K. P. Lu et al., 2007). A foszforilációfüggetlen peptidil-prolil izomerizáció is szabályozza a génexpressziót. A ciklofileinek és az FKBP-k is kölcsönhatásba lépnek transzkripciós faktorokkal, de a legmeghatározóbb szerepük a hiszton módosításában van. Az Fpr4 például, egy FKBP, a H3 és H4 hisztonokhoz kötődik, és közvetlenül izomerizál két adott prolin előtti peptidkötést (K. P. Lu et al., 2007). Emellett szerepet játszanak hiszton-chaperonként is, és irányítják a hiszton módosító és kromatin újramodellező fehérjék működését (Hanes, 2015).

A ciklofileinek és az FKBP-k szerepét az immunválasz kialakításában nagyon hamar felfedezték, mivel ezek a PPIázok komplexet képeznek bizonyos immunosuppresszor molekulákkal. Meglepő módon később bebizonyosodott, hogy a gyógyszerek immunosuppresszív hatása nem a PPIázok inhibíciójához köthető. Éppen ellenkezőleg: bizonyos esetekben a PPIáz kölcsönhatása a ligandummal az enzimaktivitás gátlása mellett akár a fehérje egyéb biológiai funkciójának növekedését is eredményezheti. A ciklofilin A-ciklosporin A, valamint az FKBK12-FK506 komplexek például gátló hatást fejtenek ki a calcineurin foszfatáz enzimre, amely alapesetben aktiválná az NF-AT transzkripciós faktort és ezzel promotálná a T-sejtek aktiválódását (Hanes, 2015; K. P. Lu et al., 2007).

A PPIáz enzimeket – különösen a Pin1-et – a sejt folyamatokban játszott sokrétű szerepüknél fogva számos rendellenességgel is összefüggésbe hozták, mint például az Alzheimer-kór, a Parkinson-kór, a rák számos fajtája és az öregedés (Ayala et al., 2003; Balastik et al., 2007; Bao et al., 2004; Butterfield et al., 2006; T. H. Lee et al., 2011; K. P. Lu et al., 2007; M. Lu et al., 2006; Rudrabhatla et al., 2008; Ryo et al., 2006; Wulf et al., 2001). Egéren végzett vizsgálatok megmutatták, hogy a Pin1 kiütése megnöveli a korfüggő amiloid neurodegeneratív betegség kialakulását, ami a humán Alzheimer-kór analógja (T. H. Lee et al., 2011; K. P. Lu et al., 2007). Néhány tanulmány arra mutatott rá, hogy a

Pin1 túlszabályozódik a kóros idegsejtekben és nagy arányban lokalizálódik az abnormálisan feltekeredett fehérjeaggregátumokból létrejövő Lewy-testekben, amelyek a Parkinson-kór során keletkeznek (Ghosh et al., 2013; T. H. Lee et al., 2011; Ryo et al., 2006).

Mivel a Pin1 inhibíciója mitotikus leállást és apoptózist eredményez, a Pin1 szabályozását a rákterápiában is célravezetőnek tartják (Göthel & Marahiel, 1999). Valóban: számos ráktípusról bebizonyosodott, hogy a Pin1 túlexpresszáldott a kóros sejtekben, egészen annyira, hogy a megnövekedett koncentrációja akár a tumor markerének is használható (Ayala et al., 2003; K. P. Lu et al., 2006; Z. Lu & Hunter, 2014; Theuerkorn et al., 2011).

Amint látjuk, a PPIáz enzimek számos sejtfollyammal és rendellenességgel hozhatók összefüggésbe, és sok esetben ígéretes terápiás célpontnak is mutatkoznak. Annak érdekében azonban, hogy mélységében megértsük a PPIázok szerepén a molekuláris biológiában, szükség van azok szerkezetének és katalitikus mechanizmusának a beható tanulmányozására is.

## 2.2.2 A peptidil-prolil izomerázok szerkezete, dinamikája és mechanizmusa

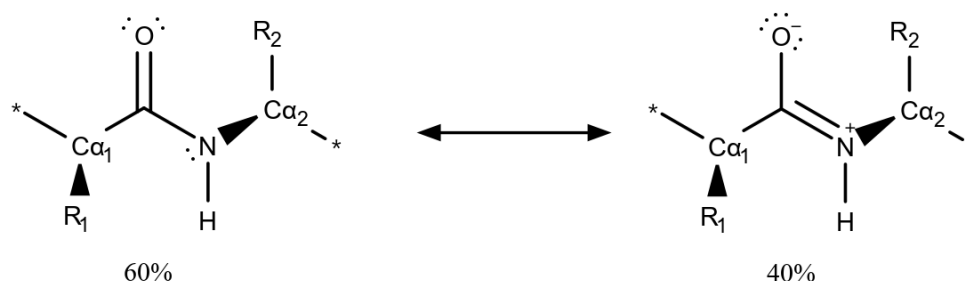
Annak ellenére, hogy a PPIáz enzimek egy viszonylag egyszerű reakció, a peptidil-prolil amidkötések cisz-transz izomerizációjának katalizátorai, a mai napig nem sikerült egyértelműen megállapítani a katalitikus mechanizmusukat, és az sem világos, hogy a különböző családba tartozó PPIázok azonos mechanizmus szerint működnek-e. Bár az elmúlt két évtizedben jónéhány alternatív javaslat született, mára már nagyjából elfogadottá vált, hogy a reakcióban nem szakad fel a peptidkötés, hanem egy csavart amid átmeneti állapoton keresztül valósul meg az izomerizáció (Tork Ladani et al., 2015). Szintén egyre inkább előtérbe kerülnek azok az eredmények, amelyek arra utalnak, hogy a katalitikus mechanizmus közvetlen összefüggésben van a PPIáz enzimek belső mozgásaival (Eisenmesser et al., 2005).

A rezonanciaelmélet alapján a peptidkötésnek két határszerkezete lehetséges, így egy polipeptid láncban az  $C_i$  és  $N_{i+1}$  atomok között a kovalens kötésnek mintegy 40%-ban kettőskötés jellege van (7. ábra) (Matena et al., 2018). Az  $sp^2$  hibridállapotban lévő atomok közötti kovalens kötés mentén gátolt a rotáció, és a hozzájuk kapcsolódó atomok szükségszerűen egy síkban vannak. Ennek következménye, hogy a peptidkötés az olefinekre jellemző cisz-transz izomériát mutat, és a  $Ca_i - C_i - N_{i+1} - Ca_{i+1}$  atomok által definiált  $\omega$  torziós szög egyensúlyi állapotban  $0^\circ$  (cisz) és  $180^\circ$  (transz) értékeket vehet fel.

A prolin kivételével mindegyik aminosav által kialakított peptidkötés esetében a transz konformáció energetikailag jóval kedvezőbb. Ennek oka az  $O_i$  és  $C_{i+1}$  atomok között fellépő elektrosztatikus kölcsönhatás, valamint a  $Ca_i$  és  $Ca_{i+1}$  atomok közötti szterikus tasztítás a cisz konformáció esetén (Wedemeyer et al., 2002). A két konformer közötti energiakülönbség N-metilacetamidra 2,5 kcal/mol-nak, az állapotok közötti energiagát pedig 20 kcal/mol-nak adódott, amely alapján a cisz konformer 1,5%-os előfordulási aránya jósolható (Pal & Chakrabarti, 1999). Ez a szám jóval nagyobb, mint a nemredundáns adathalmazon végzett statisztikai elemzés, amely azt találta, hogy



a nem prolin aminosavak közötti peptidkötések mindössze 0,03%-a volt cisz konformációban (Craveur et al., 2013; Weiss et al., 1998). Ugyanakkor az N-metilacetamid – egy egyszerű dipeptid – cisz-transz preferenciáját árnyalják azok az eredmények, amelyek a szomszédos aminosavak hatását, az oldószerhatást, valamint a fehérje feltekeredésekor létrejövő extra sztérikus feltételeket is figyelembe veszik (Alderson et al., 2018; Eberhardt et al., 1992; Reimer et al., 1998).



7. ábra: A peptidkötés két határszerkezete. A  $C_i-N_{i+1}$  kötés kettőskötés jellege miatt  $C_{i-1} - C_i - N_{i+1} - C_{i+1}$  atomok egy síkba kényszerülnek.

A prolin előtti peptidkötések jóval csekélyebb preferenciát mutatnak a transz konformáció irányába, ugyanis a  $C_{\alpha_i}$  és a  $C_{\alpha_{i+1}}$  atomok sztérikus hatása között kisebb a különbség, valamint a  $O_i$  és  $C_{i+1}$  atomok közötti elektrosztatikus kölcsönhatás mértéke is lecsökken (Wedemeyer et al., 2002). Itt a két konformer közötti energiakülönbség mindössze 0,5 kcal/mol, és az energiagát is valamivel alacsonyabb: 13 kcal/mol. Ez alapján a peptidil-prolil kötések cisz konformációjának az aránya 30% (Pal & Chakrabarti, 1999). Ez kevésbé tér el a nemredundáns adatkészleten megfigyelt aránytól (5,2%), mint a nem prolin aminosavak esetén, ugyanakkor további megfigyelések szerint a random coil polipeptidekben a cisz konformáció előfordulása 5-30% között mozog (Craveur et al., 2013; Weiss et al., 1998).

A peptidkötés cisz-transz izomerizációjának sebességére hatással vannak nem-enzimikus effektusok is, amelyek tanulmányozása fontos hozzájárulást ad az enzimekatalízis mechanizmusának megértéséhez. Mivel a peptidkötés töltésszeparációja nagyobb az egyensúlyi sík állapotokban ( $\omega = 0^\circ$  vagy  $\omega = 180^\circ$ ), mint az ortogonális átmeneti állapotokban, ami a csavart amid szerkezetet jellemzi ( $\omega = \pm 90^\circ$ ), ezért a kevésbé poláris oldószerek megnövelik az izomerizáció sebességét (Fanghänel & Fischer, 2004). Hasonló megfontolással belátható, hogy hidrogénkötés donálása a karbonil oxigénre vagy elektron donálása az amid nitrogénre csökkenti, míg hidrogénkötés donálása a nitrogénre vagy elektron donálása az oxigénre megnöveli az energiagátat, ez pedig az izomerizáció sebességét is befolyásolja (Fanghänel & Fischer, 2004).

Az enzimikus katalízis mechanizmusára kezdetben több alternatív javaslat született. Ezek közül a legfontosabbak (i) a szubsztrát deszolvatáció (a karbonil oxigén által kialakított hidrogénkötés stabilizálja az amidsíkot), (ii) a szubsztrát autokatalízis (amelyben a prolint követő aminosav amid hidrogénje kölcsönhat a prolin amid nitrogénjével), (iii) a nukleofil támadás a prolin karbonil szenére és (iv) a csavart amid átmenet stabilizálása (S. Fischer et al., 1993; K. P. Lu et al., 2007). Az elmúlt

három évtizedben gyűjtött kísérleti eredmények a csavart amid elméletet erősítették meg és egyértelműen cáfolták a többi lehetséges mechanizmust.

Mind a szubsztrát deszolvatáció, mind a szubsztrát autokatalízis mechanizmusát cáfolni látszanak a másodlagos kinetikus izotópeffektus és oldószer kinetikus izotópeffektus kísérletek, amelyek alapján a reakcióban nem vesz részt kicserélődésre képes proton (Mercedes-Camacho et al., 2013).

Több PPIáz esetében is azt feltételezték, hogy a katalitikus mechanizmust egy, az aktív centrumhoz közeli konzervált cisztein (Pin1-ben Cys133) tiol oldalláncának nukleofil támadása indítja be (Bailey et al., 2008; Fanghänel & Fischer, 2004; K. P. Lu et al., 2007; Ranganathan et al., 1997). Ezt követően a karbonil szén  $sp^2$  hibridállapota  $sp^3$ -má alakul, amely megszünteti a szén-nitrogén kötés mentén történő rotáció gátlását. Ezt a feltételezést is több kísérleti eredmény megcáfolta. Bizonyos parvulinoknak már a vad típusában is aszparaginsav helyettesíti ezt a ciszteint, és ezek a fehérjék is rendelkeznek PPIáz aktivitással. Mutagenézis kísérletek több ciklofilin és parvulin esetén azt igazolják, hogy a ciszteint nem tartalmazó mutáns is megtartja enzimátikus aktivitását (Bailey et al., 2008; Fanghänel & Fischer, 2004; Mueller et al., 2011; Zydowsky et al., 1992). Továbbá mind  $C^{13}$  NMR kísérletek, mind másodlagos kinetikus izotópeffektus kísérletek egyértelműen kizárták, hogy a karbonil szén tetraédres átmeneti állapotot képezne a reakció során (Barman & Hamelberg, 2014; Mercedes-Camacho et al., 2013; Rosen et al., 1990).

Az alternatív mechanizmusok kizárásán kívül egyes bizonyítékok közvetlenül is alátámasztják a csavart amid mechanizmust. Már 1990-ben észrevették, hogy az FK506 molekula – az FKBP-k kötőpartnere – szerkezetileg hasonlít a csavart amidra (Rosen et al., 1990). Később megmutatták, hogy a PPIáz enzimek nagyobb affinitással kötik peptidil-prolil peptidkötés átmeneti állapotát, mint akár a cisz, akár a transz konformációt (S. Fischer et al., 1993; Tork Ladani et al., 2015). Ezt a jelenséget preferált átmenetiállapot-kötésnek nevezték. A csavart átmeneti állapot megvalósulhat az N-, illetve C-terminális aminosav rotációjával is, ám molekuladinamikai szimulációkkal megmutatták, hogy az N-terminális aminosav rotációja a valószínűbb, és ez a feltételezés összhangban van mind a röntgenkristallográfiai, mind az NMR spektroszkópiai megfigyelésekkel (Trzesniak & van Gunsteren, 2006). Szintén molekuladinamikai szimulációkkal, valamint elméleti kémiai számításokkal azt is megmutatták, hogy a ciklofilin A enzim kötősebe elektrosztatikus kölcsönhatásokkal stabilizálja az átmeneti állapotot (Camilloni et al., 2014), és az ebből adódó mechanizmusnak az *elektrosztatikus fogantyú* elnevezést javasolták.

A parvulinokban található két konzervált hisztidinről (Pin1-ben H59 és H157) sokáig azt feltételezték, hogy szerepet játszanak a katalitikus mechanizmusban (6.C ábra). Kiderült azonban, hogy a H59L/H158L kettős mutáns biológiailag aktív és így a mutáns enzimet hordozó élesztő is életképes marad (Bailey et al., 2008). Mivel Pin1-ben a C113 mutációjával is erre jutottak, ebből azt a következtetést vonták le, hogy mind a ciszteinnek, mind hisztidineknek inkább szerkezeti, semmit katalitikus szerepet tulajdonítható. Ezt igazolja a hPar14 humán parvulinval végzett NMR kísérlet is, amelyben a ligandummal való kölcsönhatáskor a hisztidinek kémiai eltolódása csak igen csekély

mértékben változott meg, vagyis egyik hisztidin sem lép közvetlen kölcsönhatásba a ligandummal (Terada et al., 2001).

A parvulin fehérjecsaldában öt konzervált aminosavból álló hidrogénkötés-hálózatot azonosítottak az aktív centrum közelében, amelynek pontos szerepe a katalízisben nem tisztázott (Barman & Hamelberg, 2014; Mueller et al., 2011; Ranganathan et al., 1997). E hálózat részét képezik a korábban katalitikusan funkcionálisnak gondolt hisztidin és cisztein aminosavak is (H-C/D-S-S/T-H) (6.C ábra). A hidrogénkötés-hálózatot kialakító aminosavak mutációja bizonyos esetekben, de nem törvényszerűen csökkenti az enzim katalitikus aktivitását, ám egyik sem lép közvetlen kölcsönhatásba a ligandummal. Ugyanakkor észrevették, hogy Pin1-ben a cisztein protonáltsági állapota modulálja a hidrogénkötés-hálózatban résztvevő aminosav-oldalláncok orientációját (Barman & Hamelberg, 2014). Mindez azt valószínűsíti, hogy ezeknek az aminosavaknak a szerkezet stabilizálásában, valamint a fehérje dinamikájának szabályozásában van fontos szerepe.

A PPIázok belső dinamikája szoros összefüggésben van a mechanizmussal. A ciklofilin A és a Pin1 enzimeken végzett NMR relaxációs kísérletekkel megmutatták, hogy a katalízis közben megfigyelhető belső mozgások már a szabad fehérjében is jelen vannak, és a konformációs kicserélődés időskálája ( $R_{2,obs} = 1200 \text{ s}^{-1}$ ) egy nagyságrendbe esik a mért katalitikus sebességi állandóval ( $k_{cat} = 3300 \text{ s}^{-1}$ ) (Eisenmesser et al., 2005; Labeikovskiy et al., 2007). Ez arra utal, hogy a konformációs fluktuáció olyan jellemzője a fehérjének, amely nélkül annak katalitikus aktivitása nem magyarázható.

A PPIáz enzimek belső dinamikájának vizsgálatakor különös figyelem irányult a Pin1-ben található katalitikus PPIáz domén és az N-terminális WW domén közötti kölcsönhatásra, ugyanis a WW domén jelenléte szükséges a Pin1 biológiai aktivitásához (P. J. Lu et al., 1999). Számos arra utaló jelet találtak, hogy a WW domén nem pusztán a kötőpartner felismerésében játszik szerepet, ahogy azt korábban feltételezték, hanem allosztérikus módon képes szabályozni a katalitikus domén dinamikáját, és ezáltal magát a katalízist is (Olsson et al., 2016). A Pin1 I28A mutációja anélkül, hogy befolyásolná a két domén szerkezetét, gátolja az interdomén kölcsönhatást, és ezáltal megváltoztatja mindkét domén konformereloszlását és egyúttal a teljes enzim biológiai aktivitását is (Wilson et al., 2013).

Az interdomén kölcsönhatás erősödése pozitívan szabályozza a katalitikus domén kötési affinitását. Érdekes módon bizonyos eredmények azt támasztják alá, hogy a WW domén ligandumkötése megnöveli az interdomén dinamikát, lecsökkenti a két domén közötti kölcsönhatás mértékét és ezáltal negatívan szabályozza a katalitikus domén aktivitását (X. Wang et al., 2015). Más eredmények szerint viszont ennek az ellenkezője történik: a ligandum kötődése a WW doménhez megnöveli a domének közötti kölcsönhatást és így a katalitikus aktivitást is (Namanja et al., 2011). Ezt az ellentmondást Cavalli és munkatársai azzal a megfigyeléssel tudták feloldani, amely szerint a WW doménnek két állapota – egy alapállapota és egy gerjesztett állapota – létezik, és a különböző ligandumok különböző állapotokhoz kötődnek nagyobb affinitással, ezáltal pedig más allosztérikus szabályozó mechanizmust tudnak kiváltani (Olsson et al., 2016).

Peng és munkatársai a Pin1 oldalláncainak gyors (ps-ns időskálájú) dinamikáját vizsgálva egy hidrofób vezetékét azonosították, amely összeköti a kötőzsebet az interdomén interfésszel (Namanja et al., 2007). A vezeték mentén található hidrofób oldalláncok flexibilitása a ligandumkötéskor lecsökken. Később azt is felismerték, hogy a transz konformációjú inhibitor gyengébb választ vált ki a vezetékben, mint a cisz konformer, vagyis a dinamikában megfigyelhető változás sztereoselektív (Namanja et al., 2011). Zhou és munkatársai ezeket az eredményeket felhasználva és tovább vizsgálva arra jutottak, hogy az allostérikus interdomén szabályozásban valójában két útvonal vesz részt, amely közül az egyik már az apo fehérjében is működik, ám a második csak azután kapcsol be, hogy a ligandum hozzákötődött a WW doménhez (Guo et al., 2015). A második útvonal egyirányú: a választ csak a WW domén állapotának megváltozása váltja ki, és a katalitikus domén kötőzsebének átrendeződéséhez vezet.

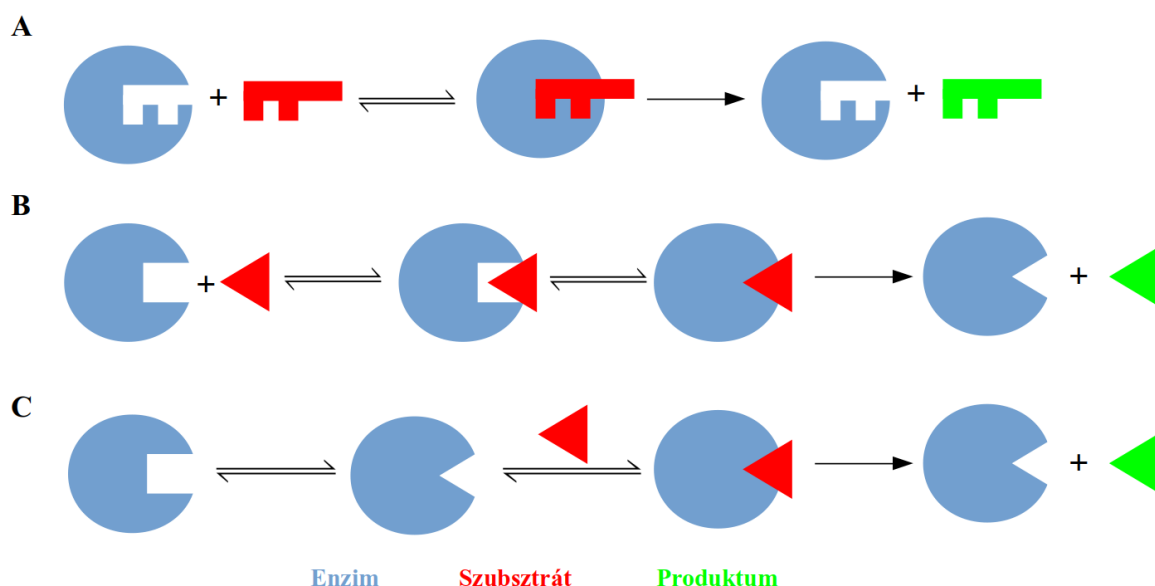
Az eddig felsorolt eredmények alapján kirajzolódó összkép szerint a parvulin-típusú peptidil-prolil cisz-transz izomerázok katalitikus mechanizmusában az egész fehérje belső dinamikája szerepet játszik. A katalízist számos hatás befolyásolja egyszerre: a hidrogénkötés-hálózat, a konzervált hisztidinek, valamint az N-terminális WW doménnel való kölcsönhatás. Ugyanakkor ezeknek a hatásoknak a pontos szerepe, illetve egymásra gyakorolt hatása nem tisztázott.

## 2.3 A fehérjemozgások jellemzése

### 2.3.1 A fehérjék belső dinamikája

A fehérjék biológiai működésének sokáig uralkodó magyarázata volt a szekvencia-szerkezet-funkció paradigma, amely szerint egy fehérje szekvenciája meghatározza a háromdimenziós szerkezetét, szerkezete pedig meghatározza a funkcióját. Szemléletesen fejezi ki e paradigmát az Emil Fischer-féle kulcs-zár modell: a fehérje csak azzal a kötőpartnerrel képes kölcsönhatást kialakítani, amely pontosan beleillik a kötőzsebébe, akár a kulcs a zárba (8. ábra). Ha pedig egy fehérje elveszíti natív harmadlagos szerkezetét, akkor szükségszerűen “elromlik”, azaz megszűnik megfelelően ellátni biológiai funkcióját. Ennek tipikus példája a fehérjeaggregáció, amikor vízoldhatatlan fibrillumok keletkeznek, de a természetes szerkezettől való ennél jóval kisebb eltérés is végzetes lehet a fehérje működésére nézve.

Az elmúlt néhány évtizedben általánosan elfogadottá vált, hogy a fehérjék dinamikus molekulák és belső mozgásaik sok esetben meghatározó szerepet játszanak biológiai funkciójukban is. Ennek egyik legáltalánosabb esete az allosztérikus szabályozás, amelynek során valamilyen külső hatásra – legtöbbször egy ligandum kötődésekor – megváltozik a fehérje konformációja, ami a kötőzsebtől távol lévő pozícióban is változást idézhet elő és így szabályozni tudja a további kötőpartnerekkel kialakított kölcsönhatás mértékét. Ez a jelenség azonban könnyen megmagyarázható a szerkezet-funkció paradigma egyszerű kiterjesztésével: mindössze annyit kell feltételeznünk, hogy egy adott fehérjéhez nem csak egy, hanem néhány konformáció tartozik.



8. ábra: Enzimreakciók lehetséges mechanizmusának sematikus ábrázolása. **A:** Az Emil Fischer-féle kulcs-zár modell. **B:** Indukált illeszkedés. **C:** Konformerszelekció.

A rendezetlen fehérjék vizsgálata azonban gyökeresen átalakította a biológiai funkcióról való felfogásunkat (Tompa, 2012b). Ezek a molekulák egyáltalán nem rendelkeznek stabil háromdimenziós szerkezettel, és ez a hagyományos kulcs-zár modell alapján kizárná, hogy stabil kölcsönhatást

alakítsanak ki. A tapasztalat ezzel szemben azt mutatja, hogy a rendezetlen fehérjék az esetek túlnyomó többségében fontos biológiai funkciót látnak el.

A szerkezet és a dinamika szoros összefüggésben van egy fehérjemolekula konformációs energiafelületével. A hagyományos képnek egy tölcészerű energiafelület felel meg, amelynek egyetlen, jól definiált minimumpontja van, vagy esetleg további néhány, amelyek a hibás feltekeredéshez tartoznak, és amelyeket a natív körülmények között a fehérjemolekula elkerül a feltekeredés során (Boehr et al., 2009). A dinamikus fehérjék energiafelületének azonban számos lokális minimuma van, amelyek energiaszintje egymással összemérhető, és az őket elválasztó potenciálgát sem túl nagy ahhoz, hogy a termális fluktuáció ne tegye lehetővé a dinamikus egyensúly létrejöttét.

A dinamikus fehérjék ligandumkötésének magyarázatára is új modellt kellett bevezetni, amelyek a kulcs-zár modell kiegészítései. A fehérjét nem egy stabil konformáció, hanem lehetséges konformációk adott eloszlása jellemzi, így a ligandummal való kölcsönhatáskor sem az "egyetlen", stabil konformációban történik változás, hanem a konformerek eloszlásában. Ez kétféle mechanizmus szerint történhet. A konformerszelekció esetén a ligandum a már meglévő, valamilyen arányban jelen lévő konformerek közül választja ki azt, amelyikkel energetikailag a legkedvezőbb kölcsönhatást tudja kialakítani. Ezzel szemben az indukált illeszkedés során a konformációs változást, illetve a konformereeloszlás megváltozását maga a ligandum kötődése váltja ki (8.B-C ábra) (Boehr et al., 2009). Számos konkrét esetben a fehérje-ligandum kölcsönhatást e két mechanizmusból kiindulva, vagy ezek ötvözésével lehet megmagyarázni.

A fehérjék natív biológiai funkcióját sok esetben nem pusztán a lehetséges konformerek közötti dinamikus egyensúly fenntartása biztosítja, hanem maguk a belső mozgások. Még ha a fehérje ténylegesen jellemezhető is egy stabil szerkezettel, az átlagszerkezet körüli fluktuációnak meghatározó szerepe lehet. A ligandummal való kölcsönhatás során megváltozhat egy-egy, a kötőzsebtől akár távoli régió dinamikája, ami közvetlen hatással lehet a kötési entrópiára, és ezáltal magára a szelektivitásra vagy az affinitásra (Kay, 1998). A fehérjék belső dinamikája tehát alapvetően meghatározza azok biológiai funkcióját.

### 2.3.2 NMR-rel mérhető, gyors dinamikát tükröző paraméterek

A fehérjeszerkezet-vizsgálati módszerek közül kiemelkedik a mágneses magrezonancia (NMR) spektroszkópia, amely egyfelől oldatfázisban, azaz a natívhoz kellően közel lévő körülmények között tudja mérni a molekulákat, emellett pedig széleskörű szerkezeti és dinamikai információt szolgáltat. Ugyanakkor az NMR-rel megfigyelhető és közvetlenül mérhető információkból közel sem triviális olyan paramétereket kinyernünk, amelyek szemléletes képet adnak a fehérje belső dinamikájáról. Ebben a fejezetben röviden kitérünk e mérési módszerek metodológiájára és matematikai hátterére.

A fehérjék belső mozgásaira szándékosan szokás többes számban hivatkozni, ugyanis a fehérjemolekula különböző részei más-más időskálán mutatnak dinamikát. Ezek mintegy 14

nagyságrendet fednek le, és a különböző mozgásokra jellemző időskálák sok esetben (legalább részben) átfednek egymással (2. táblázat).

Belső mozgás	Belső korrelációs idő ( $\tau_c$ )
Oldallánc rotáció	0,1 ps - 10 ps
Gerincdinamika	1 ps - 10 ns
Rotációs korrelációs idő ( $\tau_c$ )	1 ns - 10 ns
Katalízis	1 $\mu$ s - 1 ms
Hurok- és hajtú régiók záródása	0,1 $\mu$ s - 10 ms
Másodlagos szerkezeti elemek kialakulása	10 ns - 1 ms
Fehérjefeltekeredés	1 ms - 1 h
Fehérjeaggregáció	1 s - 1 év

2. táblázat: Egy fehérjemolekula jellegzetes belső mozgásai és az azokhoz tartozó belső korrelációs idők ( $\tau_c$ ) időskálája. A táblázat tartalmazza a fehérjemolekulák (mérettől függő) rotációs korrelációs idejének jellemző nagyságrendjét ( $\tau_c$ ) (Gáspári & Perczel, 2010).

A belső mozgások vizsgálatához fontos referenciapont a fehérjemolekulára jellemző  $\tau_c$  rotációs korrelációs idő, amely azt az időt jelzi, amíg a fehérje 1 radiánt elfordul (Gáspári & Perczel, 2010). Ez függ a fehérje méretétől, de a külső körülményektől is, mint például a hőmérséklet, a pH és az oldat viszkozitása, de általában ns-os nagyságrendbe esik. A fehérjék belső mozgásait jellemzően gyors (ps-ns) és lassú ( $\mu$ s-ms) mozgásokra szokás kategorizálni aszerint, hogy a rotációs korrelációs időnél gyorsabb vagy lassabb időskálán történnek.

A gyors mozgások közé tartozik az oldalláncok rotációja és a gerincdinamika. Ezek közvetett megfigyelésére a legelterjedtebb megoldás az amid nitrogénekhez tartozó transzverzális és longitudinális  $^{15}\text{N}$ -relaxációs idők ( $T_1$ ,  $T_2$ ), valamint a heteronukleáris ( $^1\text{H}$ - $^{15}\text{N}$ ) NOE-k (nukleáris Overhauser-effektus) mérése. A lassú mozgások megfigyelésére leggyakrabban reziduális dipoláris csatolást (RDC) vagy a relaxációs diszperziót szokás mérni (Lange et al., 2008; Markwick et al., 2008).

E dolgozat fókuszában a fehérjék gyors dinamikája áll. Anélkül, hogy külön kitérnénk az erre alkalmas NMR-kísérletek technikai magvalósítására, áttekintjük, hogy a közvetlen kísérleti adatokból miként lehet a gyors fehérjemozgásokra vonatkozó szemléletes információt kinyerni.

A relaxációs adatokból dinamikai információk kinyerésére alkalmas, leggyakrabban használt eljárás a Lipari-Szabó-féle modellfüggetlen formalizmus (Chen et al., 2007; Gáspári & Perczel, 2010; Lipari & Szabo, 1982; J. W. Peng, 2012; L. Zhang et al., 2012). Az elnevezés arra utal, hogy a módszer nem él semmilyen előfeltevéssel a rotációs-diffúziós tenzorban található relaxációs vektorok orientációját illetően. A relaxációs vektornak molekuláris bukdácsolásból (*molecular tumbling*) adódó fluktuációjának leírására a  $C(t)$  korrelációs függvényt használjuk. Ez egyfajta memóriafüggvény, amely kifejezi, hogy a vektor orientációja  $0$  és  $t$  időpillanatban mennyire korrelál egymással – alakja ennél fogva valamilyen lecsengő görbét vesz fel.

A Lipari-Szabó formalizmus alapfeltevése, hogy a gyors és lassú mozgások függetlenek egymástól. Így a korrelációs függvény két tényezőre bontható:

$$C(t) = C_o(t)C_e(t) \quad (1)$$

ahol

$$C_o(t) = e^{-t/\tau_c} \quad (2)$$

fejezi ki a teljes (*overall*) molekula lassú, rotációs-diffúziós mozgását ( $\tau_c$  a rotációs korrelációs idő), valamint

$$C_e(t) = S^2 + (1 - S^2) \cdot e^{-t/\tau_e} \quad (3)$$

pedig a belső mozgásokat fejezi ki (Lipari & Szabo, 1982). Itt  $\tau_e$  az effektív belső mozgások korrelációs idejét jelöli,  $S^2$  pedig az ún. általános rendparamétert, amelyre később részletesen is kitérünk. A fenti két egyenletből következik, hogy

$$\tau^{-1} = \tau_c^{-1} + \tau_e^{-1} \quad (4)$$

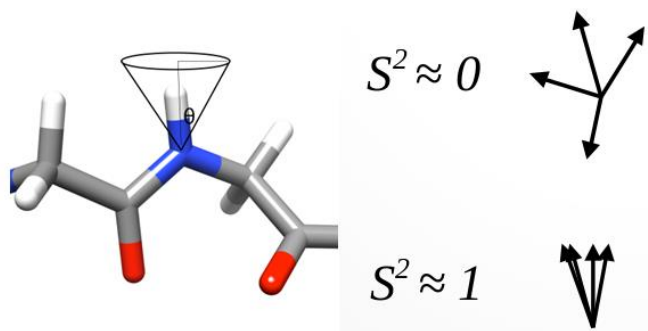
ahol  $\tau$  az adott relaxációs vektor által ténylegesen érzékelt korrelációs idő.

A korrelációs függvény Fourier-transzformáltja az ún. spektrális sűrűségfüggvény:

$$J(\omega) = \frac{2}{5} \left( \frac{\tau_c}{1 + (\omega\tau_c)^2} \right) = \frac{2}{5} \left( \frac{S^2\tau_c}{1 + (\omega\tau_c)^2} + \frac{(1 - S^2)\tau}{1 + (\omega\tau)^2} \right) \quad (5)$$

amely kifejezi, hogy mekkora valószínűséggel találunk a molekulában  $\omega$  frekvenciával fluktuáló elemet. A spektrális sűrűség leképezés (*spectral density mapping*) módszere összefüggést teremt a mért relaxációs idők, valamint a spektrális sűrűségfüggvény között (Jeffrey W. Peng & Wagner, 1992). Valójában a relaxációs idők a következő frekvenciákon mintavételezik a spektrális sűrűségfüggvényt: 0,  $\omega_N$ ,  $\omega_H$ ,  $\omega_N + \omega_H$  és  $\omega_N - \omega_H$ , ahol  $\omega_N$  és  $\omega_H$  az amid nitrogén- és hidrogénatomok Larmor-frekvenciája az egyes aminosavakban.

A Lipari-Szabó-formalizmus során összességében a mért relaxációs-időket illesztik a spektrális-sűrűségfüggvény egyenletére. A módszer eredménye a teljes molekulára vett rotációs-korrelációs idő ( $\tau_c$ ), az egyes amid N-H kötésvektorok fluktuációjának frekvenciáját kifejező effektív korrelációs idő ( $\tau_e$ ), valamint annak amplitúdóját kifejező általános rendparaméter ( $S^2$ ). Ez fejezi ki, hogy a relaxációs vektor mozgása mennyire korlátozott, vagy szemléletes képpel élve: mekkora az amid N-H kötésvektor által bejárt kúp szélessége.  $S^2 = 1$  esetén egyáltalán nincs belső mozgás, míg  $S^2 = 0$  esetén a kötésvektor egyforma valószínűséggel mutat a tér bármely irányába (9. ábra).



9. ábra: Az amid N-H kötésvektorok gyors dinamikája az általános rendparaméterekkel ( $S^2$ ) kifejezve:  $S^2 = 1$  esetén egyáltalán nincs belső mozgás, míg  $S^2 = 0$  esetén a kötésvektor egyforma valószínűséggel mutat a tér bármely irányába.



Számos változata létezik a Lipari-Szabó-formalizmusnak (Gáspári & Perczel, 2010). A  $^{13}\text{C}$  relaxációs idők mérésével például hasonló információhoz lehet jutni a metil oldalláncokra vonatkozó gyors dinamikát illetően. Ezenkívül bizonyos esetekben a mérési adatok megfelelő illesztéséhez a modell kiterjesztésére is szükség van, amelyben további paramétereket vezetnek be: például a lassú időskálán történő konformációs kicserélődést leíró  $R_{ex}$ -et. De lehetőségünk nyílik arra is, hogy a belső mozgásokat tovább szeparáljuk jellemző időskálájuk szerint  $>100$  ps és  $<100$  ps mozgásokra. Ezeket a  $S_f^2$  (*fast*) és  $S_s^2$  (*slow*) rendparaméterek írják le (az elnevezés megtévesztő, ugyanis mindkét mozgás gyorsabb a rotációs korrelációs időnél). Ezek az általános rendparaméterrel az alábbi összefüggésben állnak:

$$S^2 = S_f^2 S_s^2 \quad (6)$$

Mint láttuk, a Lipari-Szabó-formalizmus által nyert általános rendparaméterek közvetlen és szemléletes képet adnak a fehérjegerinc vagy -oldalláncok gyors mozgásairól. Anélkül, hogy részletesen kitérnénk rájuk, említést érdemelnek azonban további NMR-kísérletek, amelyek különböző időskálájú belső mozgásokról szolgáltatnak információt. Ide tartoznak a reziduális dipoláris csatolás (RDC) (ps- $\mu$ s), a J-csatolás elemzése ( $\mu$ s-s), H/D kicserélődés (ms-s) vagy a relaxációs diszperziót mérő módszerek, például a CPMG (Carr-Purcell-Meiboom-Gill) módszer ( $\mu$ s-ms) (Gáspári & Perczel, 2010; J. W. Peng, 2012; L. Zhang et al., 2012).

Mindezek mellett pedig azokat az információkat, amelyeket hagyományosan a fehérje szerkezetére vonatkoztattak (pl. kémiai eltolódás, homonukleáris NOE-k), szintén fel lehet használni olyan szerkezeti sokaságok létrehozására, amelyek tükrözik a fehérje belső dinamikáját. Erre részletesen a következő fejezetben térünk ki.

### 2.3.3 Dinamikus fehérjeszerkezeti sokaságok

A fehérjék dinamikus volta szükségessé teszi, hogy finomítsuk a róluk alkotott modelljeinket. A szerkezet-funkció paradigmát kielégítően tükrözi egy statikus modell, amely a fehérjemolekula minden atomjának háromdimenziós koordinátáját tartalmazza. Ilyen szerkezetek vannak a fehérje-adatbázisban (PDB) is, amely a legszélesebb körben ismert és használt, kísérletesen meghatározott fehérjeszerkezeteket tartalmazó nyilvános adatbázis (Berman et al., 2000).

Fontos hangsúlyozni, hogy a merev fehérjeszerkezeteket tartalmazó modellek nem “rosszak”, hiszen mindegyik modell csak annyira lehet jó vagy rossz, amennyire eszközként tudjuk használni a kísérleti tapasztalatok megmagyarázására vagy egy elmélet szemléletes illusztrálására. Az teszi szükségessé egy új modell bevezetését, hogy a meglévő már nem képes megmagyarázni bizonyos kísérleti megfigyeléseket. Azonban még ilyenkor sem a teljes régi modell elvetése a megfelelő eljárás, hanem annak kiegészítése, hogy így összhangban legyen az újabb kísérleti eredmények magyarázatára alkotott elmélettel.

Egy merev fehérjeszerkezeti modell önmagában nem tükröz semmiféle belső mozgást: nem olvasható ki belőle sem a fehérjemolekulának az átlagszerkezet körüli fluktuációja, sem pedig a bejárt konformációs tér. A PDB-szerkezetek adatai között némi kivételt képez a röntgenkristallográfiával meghatározott kristályszerkezetben található B-faktor, ami az egyes atomi pozíciók bizonytalanságát tükrözi, és ez legalább részben összefüggésbe hozható az atomi fluktuációkkal. Ám a kristályszerkezetek eleve nem natív körülmények között készülnek, és ezért nem várhatjuk el, hogy a kimerevített állapotba kényszerített fehérjemolekula a természetes belső dinamikája szerint viselkedjen.

Ezzel szemben egy NMR-kísérlet elvégzésekor mintegy  $10^{16}$ - $10^{17}$  fehérjemolekula található a mintacsőben, amelyek az oldatfázisban eleve belső dinamikát mutatnak (Gáspári & Perczel, 2010). Ennek megfelelően az NMR-rel megmért paraméterek egyszerre tükröznek sokaság- és időátlagot. Sőt, valójában minden kísérletes NMR-paraméter szükségszerűen dinamikai információt is hordoz a fehérjéről – még akkor is, ha az elvégzett kísérletek többségénél nem ezt az információt keressük, és ezért kihasználatlanul maradnak.

A hagyományos szerkezetmeghatározás során egy olyan stabil, háromdimenziós fehérjeszerkezetet keresünk, amely megfelel a kísérletesen megmért paramétereknek. Az NMR esetében leginkább a homonukleáris  $^1\text{H}$ - $^1\text{H}$  NOE-kat szokták használni, amelynek a magok közötti távolságtól való jól ismert függése praktikus eszközzé teszi a fehérjék háromdimenziós szerkezetének meghatározásához. A “egyedi szerkezetfinomítás” (*single structure refinement*) néven ismert eljárás során egy molekuladinamikai szimuláció segítségével a fehérjemolekulát egy olyan konformációba igyekszünk hozni, amely egyszerre felel meg a fehérjeszerkezetekről (kötéshosszakról, kötésszögekről, amidcsíkokról, stb.) való a-priori ismereteinknek, valamint a kísérletesen meghatározott  $^1\text{H}$ - $^1\text{H}$  távolságoknak. Az esetek túlnyomó részében azonban ilyen szerkezet nem létezik.

Általános tapasztalat az NMR-rel foglalkozó fehérjeszerkezet-kutatók körében, hogy a végső szerkezet nagyon gyakran nem felel meg az összes kísérletes NOE-nak. A ma már gyakorlatilag fekete dobozként használt szerkezetfinomító szoftverek fel vannak készülve arra, hogy a bemenetként használt  $^1\text{H}$ - $^1\text{H}$  távolságokat nem tudják mind rákényszeríteni a szerkezetre, ezért valamilyen algoritmus szerint ezek közül néhányat a szimuláció során szép lassan elhagynak, hogy a többi teljesülhessen. Az számít jó NMR-szerkezetnek, amelyben a lehető legkevesebb kísérletes NOE sérül.

A PDB-ben lévő NMR-szerkezetek is több konformert tartalmaznak, ám ezek nem szerkezeti sokaságok. Az NMR-es szerkezetfinomítás során ugyanis minden egyes szerkezetre rákényszerítjük az összes mért paramétert. Általában az eljárás során számos szerkezet hoznak létre, és az adatbázisba csak a legjobb néhányat töltik fel, vagyis azokat, amelyekben a legkevesebb kísérletes paraméter sérül – ám ez a bizonyos “legkevesebb” még a legjobb NMR-szerkezetek esetén sem nulla.

A fehérjeszerkezeti sokaságok nem egy, hanem számos háromdimenziós szerkezetet tartalmaznak. Külön-külön egyik szerkezettől sem várjuk el, hogy megfeleljen a kísérletes paramétereknek, hanem az egész sokaság kell, hogy tükrözze azokat. Ezt a legtöbb – bár nem mindegyik – paraméter esetén úgy alkalmazzuk, hogy a sokaságban lévő egyes szerkezetekre visszaszámolt

paraméterek átlagát igyekszünk megfeleltetni a kísérletes értékeknek, ezzel is tükrözve, hogy a mért értékek minden esetben sokaság- és időátlagra vonatkoznak.

### 2.3.4 A szerkezeti sokaságok előállítása mint matematikai probléma

A kísérletes paramétereknek megfelelő szerkezeti sokaságok generálása matematikailag egy alulhatározott probléma: egyrészt a rendszernek sokkal több szabadsági foka van, mint a mért paraméterek száma, másrészt pedig – a sokaság méretére tett megkötés hiányában – elméletileg is végtelen sok olyan szerkezeti sokaság létezik, amelyek megfelelnek a kísérleti paramétereknek, és még ha létezik is ilyen megkötés, gyakorlatilag akkor is végtelen számú sokaságot elő lehet állítani.

Másrészt amiatt is gyakran kritizálják szerkezeti sokaságok előállításának módszereit – különösen a külső kényszerfeltételekkel megkötött MD szimulációt –, hogy egy mesterséges erőhatást visz be a modellbe. Ahogy bármely modell, úgy a molekuladinamikai szimulációkhoz felhasznált – nem mellesleg az idővel egyre szofisztikáltabb – erők célja, hogy minél pontosabban leírják a modellezett molekulák *valóságos* viselkedését. Így a kritikusok szerint a külső kényszerfeltételek szükségessége valójában a rendelkezésünkre álló erők hiányosságának burkolt beismerése.

Az egyik elvárás, amit a szerkezeti sokaságokat generáló módszerek matematikai szabatoságával kapcsolatban támasztani lehet, hogy feleljenek meg az entrópiamaximum elvének: a lehetséges, a-priori adatokat tartalmazó valószínűségi eloszlások közül az a legjobb, amelyben maximális a rendszer által hordozott információs entrópia (Shannon-entrópia) (Jaynes, 1957a, 1957b). Ez az elv a fizika történetében először kapcsolatot teremt a statisztikus mechanika és az információelmélet között: korábban ugyanis az entrópiát pusztán a termodinamikai elméletek validációjára használták, ezután viszont az entrópiamaximum elve az új elméletek kiindulási pontjává vált (Cesari et al., 2018).

A szerkezeti sokaságokkal foglalkozó kutatók komoly erőfeszítéseket tettek annak érdekében, hogy a sokaságok előállítása megfeleljen az entrópiamaximum elvének. A legmegfelelőbb valószínűségi eloszlást kereső sokaságelőkészítő-eljárások esetén ez egy gyakorlatilag megkerülhetetlen elv, és több tanulmány is bemutatott olyan algoritmust, ahol megfelelően teljesül (Bonomi et al., 2015; Hummer & Köfinger, 2015). Az egyik tanulmány fő konklúziója, hogy az alkalmazott külső paraméterek száma lineárisan kell, hogy skálázzódjon a sokaság méretével (Hummer & Köfinger, 2015). Mások a konformerszelekciós algoritmusokról bizonyították be, hogy megfelelnek az entrópiamaximum elvének (Beauchamp et al., 2014; Berlin et al., 2013; Bonomi et al., 2015; Leung et al., 2016; Sanchez-Martinez & Crehuet, 2014).

## 2.4 A molekuladinamikai szimuláció

A szimuláció mint tudományos módszer hidat képez kísérlet és elmélet között (Michael P. Allen, 2004). A számítási kapacitás ugrásszerű növekedésének hála, a számítógép mára egy virtuális laboratóriummá növekedett, amelyben a szimulációk futtatásakor a szó legszorosabb értelmében vett kísérleteket végzünk.

A fizikai törvények elegáns, zárt egyenlet formájában történő kifejezése gyakran analitikusan nem megoldható matematikai problémákhoz vezet (Rapaport, 2004). Ilyenkor elkerülhetetlen, hogy egy elmélet alkalmazhatósága érdekében az egyenletek analitikus megoldását numerikus módszerekkel helyettesítsük.

Molekuláris rendszerek modellezésének két legelterjedtebb módszere a Monte-Carlo- és a molekuladinamikai (MD) szimuláció. Az előbbi egy sztochasztikus eljárás, melynek során a szimuláció minden egyes lépését egy véletlenszám határozza meg: lényegében egy halmaz (pl. lehetséges molekuláris konformációk) véletlenszerű mintavételezéséről van szó. Az MD ezzel szemben determinisztikus, amelynek elméleti alapját a newtoni soktest-probléma csatolt differenciálegyenlet-rendszerének analitikus megoldása adja. Mindkét módszer széles körben alkalmazott fehérjék belső dinamikájának vizsgálatára.

Az MD egy klasszikus fizikai megközelítés: a molekulák atomjait töltéssel rendelkező tömegpontoknak tekintjük, melyeket az elektrosztatikus potenciálból, valamint a molekula szerkezetéből származó erők mozgatnak. Nem vesszük figyelembe a molekula szubatomi részeit (elektronok, atompályák, molekulapályák), és az elemi részecskék kvantum viselkedését is figyelmen kívül hagyjuk. A kvantummechanikai megközelítést klasszikus fizikai leírással felváltva “durvább”, kevésbé szofisztikált képet kapunk a vizsgált rendszerről, azonban az elméleti kémiai módszerek igen magas számítási igénye praktikusán lehetetlenné teszi, hogy akár egy kisebb méretű fehérjemolekulát is a kvantummechanika eszköztárával modellezzünk.

Az MD szimulációban bármely  $m_i$  tömegű  $\mathbf{r}_i$  atom mozgásának alapegyenlete:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i \quad (7)$$

Az atomra ható  $\mathbf{f}_i$  erőt a potenciális energia parciális differenciálásával számítjuk ki:

$$\mathbf{f}_i = -\frac{\partial}{\partial \mathbf{r}_i} U \quad (8)$$

ahol  $U(\mathbf{r}^N)$  a potenciálisenergia-függvény, és  $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  jelöli az összes,  $3N$  atomi koordinátát ( $N$  az atomok száma).

A potenciális energiát empirikus megközelítéssel írjuk le, melyre több különböző modell is létezik. A legelterjedtebb gyakorlat szerint a kölcsönhatásokat intramolekuláris és téren át ható kölcsönhatások összegére bontjuk:

$$U = U_{intra\text{mol}} + U_{tér} \quad (9)$$

Az intramolekuláris kölcsönhatások a kötéshosszak nyújtásából, a kötésszögek hajlításából és a torziós szögek változásából származnak:

$$U_{intramol} = \frac{1}{2} \sum_{kötések} k_{ij}^r (r_{ij} - r_{eq})^2 + \frac{1}{2} \sum_{kötésszögek} k_{ijk}^\theta (\theta_{ijk} - \theta_{eq})^2 + \frac{1}{2} \sum_{torziós\ szögek} \sum_m k_{ijkl}^{\Phi, m} (1 + \cos(m\Phi_{ijkl} - \gamma_m)) \quad (10)$$

ahol  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$  két atom közötti aktuális távolság, melynek egyensúlyi értéke  $r_{eq}$ . Hasonlóan,  $\theta_{ijk}$  három atom által bezárt szög,  $\theta_{eq}$  egyensúlyi értékkel, valamint  $\Phi_{ijkl}$  négy atom által meghatározott torziós szög, melynek viszonyítási értéke  $\gamma_m$ .  $k$ -val jelöljük az egyes intramolekuláris kölcsönhatások súlyfaktorát.

A téren át ható kölcsönhatásokat, külső elektromos tér hiányában az egyes, egymástól  $r$  távolságra lévő atompárok közötti kölcsönhatással szokás közelíteni:

$$U_{tér} = \sum_{i,j} u^{Coulomb}(r) + u^{LJ}(r) \quad (11)$$

Töltések vagy parciális töltések esetén elektrosztatikus kölcsönhatás lép fel, melyet a Coulomb-potenciál ír le:

$$u^{Coulomb}(r) = \frac{Q_1 Q_2}{4\pi\epsilon_0 r} \quad (12)$$

ahol  $Q_1$  és  $Q_2$  az atomok parciális töltése,  $\epsilon_0$  pedig a vákuum permittivitása. Atompárok között fellép a van der Waals-kölcsönhatás, amelyet leggyakrabban a Lenard-Jones-potenciállal szokás közelíteni:

$$u^{LJ}(r) = 4\epsilon \left[ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right] \quad (13)$$

amelynek paraméterei  $\epsilon$ , a potenciálgödör mélysége, valamint  $\sigma$ , a kölcsönhatás hatótávolsága.

Az atomok mozgását tehát egy differenciálegyenlet-rendszer írja le, melynek nincs analitikus megoldása. Számos numerikus megoldás létezik azonban, és ezek között máig a legnépszerűbbek az elsők között megjelent sebesség-Verlet-algoritmus (*velocity Verlet algorithm*), illetve az azzal algebrailag ekvivalens leapfrog-algoritmus. Az elnevezés (*leapfrog* = bakugrás) arra utal, hogy az algoritmus a hely- és sebesség-adatokat egymáshoz képest fél időlépésnyivel ( $\Delta t$ ) eltolva adja meg.

A sebesség-Verlet-módszer alapegyenletei:

$$\begin{aligned} \vec{x}(t + \Delta t) &= \vec{x}(t) + \vec{v}(t)\Delta t + \frac{1}{2}\vec{a}(t)\Delta t^2 \\ \vec{v}(t + \Delta t) &= \vec{v}(t) + \frac{\vec{a}(t) + \vec{a}(t + \Delta t)}{2}\Delta t \end{aligned} \quad (14)$$

ahol  $\vec{x}(t)$ ,  $\vec{v}(t)$  és  $\vec{a}(t)$  az atom pozíciója, sebessége és gyorsulása az adott időpillanatban, míg a leapfrog-módszer egyenletei:

$$v_{i+\frac{1}{2}} = v_{i-\frac{1}{2}} + a_i \Delta t \quad (15)$$

$$x_{i+1} = x_i + v_{i+\frac{1}{2}}\Delta t$$

de a pozíció és sebesség itt is kifejezhetők az időlépés egész számú többszörösének pillanataiban:

$$\begin{aligned} x_{i+1} &= x_i + v_i\Delta t + \frac{1}{2}a_i\Delta t^2 \\ v_{i+1} &= v_i + \frac{1}{2}(a_i + a_{i+1})\Delta t \end{aligned} \tag{16}$$

ahol  $x_i$ ,  $x_{i+\frac{1}{2}}$  és  $x_{i+1}$  az atomi pozíció az  $i$ -edik,  $i$  és  $i+1$ -edik időlépésben, hasonló jelöléssel a sebességre és a gyorsulásra.

A molekuladinamikai szimulációk során, különösen a biomolekulák modellezésekor indokolt a kötéstávolságok állandó értéken tartása, ugyanis a kovalens kötések vibrációs frekvenciája jóval gyorsabb időskálán mérhető, mint a molekula egyéb konformációs változásai. A kötéshosszakat az empirikusan meghatározott értékre kényszerítjük. Ezt algebrai eljárással, az esetek többségében a Lagrange-multiplikátorok módszerével hajtjuk végre. Így lecsökkentjük a molekula belső szabadsági fokainak a számát, és gyakorlatilag a potenciálisenergia-függvény változóinak a számát is. A kötéshosszakra alkalmazott kényszerfeltételekre kifejlesztett első algoritmus a SHAKE volt (Ryckaert et al., 1977), melyet követett a valamivel gyorsabb LINCS (Hess et al., 1997).

Fontos kiemelni, hogy bár magyarul ezeket a megkötéseket is szokás “kényszerfeltételnek” nevezni, mégsem azonos módon alkalmazzuk őket, mint a szintén kényszerfeltételként hivatkozott kísérleti paramétereket. Az előbbieket ugyanis algebrai módszerekkel kezeljük és a változók számát csökkentjük, míg az utóbbiakat egy külső potenciál segítségével kényszerítjük rá a molekulára, de elviekben nem szüntetjük meg a molekula szabadsági fokát. Az eredeti, angol nyelvű kifejezés jobban tükrözi ezt a különbséget: az előbbieket *constraint*, az utóbbiakat *restraint* néven hivatkozzák.

Hogy az amúgy sem csekély számítási igényt csökkentjük, a szimulációk során távolság-limitet (*cutoff*) alkalmazunk a kölcsönhatások kezelésére. Az egymástól ennél nagyobb távolságra lévő atompárok közötti kölcsönhatási energiát elhanyagoljuk.

A szimulációk egy fontos kérdése a szimulációs doboz mérete, ez határozza meg ugyanis a doboz falával kölcsönható molekulák számát, ami egy mesterséges hatásként jelentkezik. A számítási idővel történő takarékoság ugyanakkor azt diktálja, hogy minél kisebb szimulációs dobozt válasszunk, amelybe kevesebb oldószermolekula fér. Azonban még ha nagyon sok, akár  $10^6$  atomot szimulálunk is, a felülettel érintkező atomok aránya eléri a 6%-ot, ami egy számottevő faktort jelent. E probléma kezelésére alkalmazzák a periodikus határfeltételeket. Ennek során a szimulációs doboznak lényegében nincsen fala: amelyik molekula translációs mozgással kilép a dobozból, az a másik, szemközti oldalon belép. Ugyanez a megfontolás alkalmazható kölcsönhatásokra is, ez azonban biomolekulák esetén szükségessé teszi a doboz méretének gondos megválasztását, hogy a szimulált részecske elég távol legyen a doboz bármely falától, ezzel kizárva, hogy önmagával alakítson ki kölcsönhatást. A térkihasználást növelendő, a kocka vagy téglatest alakú szimulációs dobozok helyett szokás rombododekaédert vagy csonkolt oktaédert is alkalmazni.

### 3 Célkitűzés

Munkám célja a PSD-95 fehérje N-terminális PDZ1-2 tandemje, valamint a parvulin-típusú PPIáz fehérjék belső dinamikája és biológiai funkciója közötti összefüggés jobb megértése volt. A parvulinok vizsgálatának megkezdésekor már rendelkezésemre álltak a laboratóriumunkban korábban létrehozott dinamikus fehérjeszerkezeti sokaságok. Ezért az első, célként kitűzött feladat csak a PDZ tandemre vonatkozott:

*Olyan dinamikus fehérjeszerkezeti sokaságok előállítása kísérletes külső paraméterekkel megkötött molekuladinamikai szimulációval, amelyek tükrözik a vizsgált fehérje gyors (ps-ns) időskálán történő belső mozgásait.*

Az általam, illetve a korábban előállított szerkezeti sokaságok elemzésével elérendő célt a következő két pont foglalja össze:

*A leghangsúlyosabb gyors belső mozgások azonosítása főkomponens elemzéssel.*

*Atomi szintű modell megalkotása a vizsgált fehérjék biológiai szerepe és belső mozgásai közötti összefüggés magyarázatára.*

A két vizsgált fehérjecsalád biológiai funkciója kapcsán az alábbi célokat fogalmaztam meg:

*A ligandummal való kölcsönhatás és a szupramoduláris dinamika közötti összefüggés vizsgálata PDZ tandemben.*

*Az intradomén belső mozgások és a szupramoduláris dinamika közötti összefüggés vizsgálata PDZ tandemben.*

*A vizsgált parvulin-típusú rotamázok dinamikai viselkedésének összehasonlítása, és a különbségeket okozó faktorok azonosítása.*

## 4 Alkalmazott módszerek

### 4.1 Külső kényszerfeltételekkel megkötött molekuladinamikai szimuláció

#### 4.1.1 Kényszerfeltételként alkalmazott kísérletes paraméterek

Munkám során háromféle kísérletes paraméteret alkalmaztam: a sokaságok előállításához  $^1\text{H}$ - $^1\text{H}$  NOE-kat és  $S^2$  rendparamétereket, a sokaságok validációjához pedig  $S^2$  rendparamétereket és kémiai eltolódásokat.

A szerkezeti sokaságok létrehozása során adott egy sor kísérletesen meghatározott paraméter (esetemben  $^1\text{H}$ - $^1\text{H}$  NOE-k és  $S^2$  rendparaméterek), és keressük azt a sokaságot, amelyik a legjobban tükrözi a mért értékeket. Ehhez léteznie kell egy olyan

$$X_{calc}^k = X(\mathbf{r}^k) \quad (17)$$

függvénynek, amely kapcsolatot teremt egy fehérjeszerkezetben lévő  $k$ -adik atom vagy aminosav adott atomjainak (pl. amid N és H)  $\mathbf{r}^k$  pozíciói és az ún. visszszámolt (*back-calculated*)  $X_{calc}^k$  paraméter között. Sokaságok esetén ez a függvény értelemszerűen az összes szerkezet atomi koordinátájától függ:

$$X_{calc}^k = X(\mathbf{r}_1^k, \mathbf{r}_2^k, \dots, \mathbf{r}_N^k) \quad (18)$$

ahol  $\mathbf{r}_1^k, \mathbf{r}_2^k, \dots, \mathbf{r}_N^k$  rendre a  $k$ -adik atom vagy aminosav atomjainak az 1., 2., ...  $N$ -edik replikában lévő pozícióit jelöli ( $N$  a sokaság mérete). Általában (bár nem mindig) ez nem más, mint az egyes replikákra visszszámolt paraméter átlaga:

$$X_{calc}^k = \frac{1}{N} \sum_{i=1}^N X(\mathbf{r}_i^k) \quad (19)$$

A következőkben bemutatom a három – a szimulációk vagy a validálás során – felhasznált paraméter alkalmazását.

1.  $^1\text{H}$ - $^1\text{H}$  homonukleáris NOE-k. A homonukleáris NOE-k mérése az NMR-rel történő szerkezetmeghatározás rutin eljárása. Egymástól 5-6 Å távolságra lévő protonok adnak a spektrumban detektálható keresztcsúcsot, amelyből a kalibráció után proton-proton távolság adatokat nyerhetünk. Bár a fenti megközelítésből az következne, hogy a kísérletes NOE-knak való megfelelést egyszerűen két atomi pozíció közötti távolsággal lehet ellenőrizni, ennek ellenére közel sem egyértelmű azok alkalmazása sem a szerkezetmeghatározásban, sem pedig szerkezeti sokaságok generálásában vagy validálásában (Ángyán & Gáspári, 2013).

A keresztcsúcs intenzitása a két atommag közötti távolság negatív hatodik hatványával arányos. A fehérjemolekula fluktuációja miatt szinte lehetetlen egyetlen olyan szerkezetet találni, amelyben mindegyik kísérletesen meghatározott NOE teljesül. Belső dinamikát tükröző



szerkezeti sokaságok esetén ezért a visszazámolt proton-proton távolságot a negatív hatodik hatványok szerint átlagolják

$$r_{calc}^{-6} = \frac{1}{N} \sum r^{-6} \quad (20)$$

ahol  $N$  a sokaság méretét jelöli. Az ilyen módon előállított és validált sokaságokban többnyire drasztikusan lecsökken a nem teljesülő kísérletes NOE-k száma.

Az NOE-kból származtatott távolságok voltak a legelőször alkalmazott külső kényszerfeltételek MD szimulációk során: mind az idő-, mind a sokaságtárolás igen elterjedt módszer (Bonvin et al., 1994; Torda et al., 1990). Az előbbi esetben a szimuláció minden lépésekor egy bizonyos időablakban megjelenő szerkezetekre átlagolják a proton-proton távolságokat, az utóbbi esetben pedig több, párhuzamos replikát szimulálnak. Én az itt bemutatott szimulációk mindegyikében sokaságtárolást alkalmaztam.

2. Az  $S^2$  értékek olyan kísérletes paraméterek, amelyeket nem lehet egyedi replikákra alkalmazni, hiszen szemléletesen egy kötésvektor (pl. amid N-H vagy oldallánc Ca-C $\beta$  vektor) széttartását írják le, és ezt csak sokaságalapú reprezentációval lehet tükrözni. A szerkezeti sokaságból a következő képlettel lehet visszazámolni a rendparamétert (Best & Vendruscolo, 2004; Chen et al., 2007):

$$S_{ij}^2 = \frac{3}{2} (\langle \hat{x}_{ij}^2 \rangle^2 + \langle \hat{y}_{ij}^2 \rangle^2 + \langle \hat{z}_{ij}^2 \rangle^2 + 2\langle \hat{x}_{ij}\hat{y}_{ij} \rangle^2 + 2\langle \hat{y}_{ij}\hat{z}_{ij} \rangle^2 + 2\langle \hat{x}_{ij}\hat{z}_{ij} \rangle^2) - \frac{1}{2} \quad (21)$$

ahol  $\hat{x}_{ij}$ ,  $\hat{y}_{ij}$  és  $\hat{z}_{ij}$   $i$  és  $j$  atom közötti egységvektor komponensei, a  $\langle \hat{x}_{ij}^2 \rangle$  jelölés pedig a mennyiségek átlagát jelenti. Sok esetben, pl. molekuladinamikai szimulációk során a kötéshosszakat állandó értéken tartjuk, így a fenti kifejezés a következő alakra egyszerűsödik:

$$S_{ij}^2 = \frac{3}{2(r_{ij}^{min})^4} (\langle x_{ij}^2 \rangle^2 + \langle y_{ij}^2 \rangle^2 + \langle z_{ij}^2 \rangle^2 + 2\langle x_{ij}y_{ij} \rangle^2 + 2\langle y_{ij}z_{ij} \rangle^2 + 2\langle x_{ij}z_{ij} \rangle^2) - \frac{1}{2} \quad (22)$$

ahol  $x_{ij}$ ,  $y_{ij}$  és  $z_{ij}$   $i$  és  $j$  atom közötti kötésvektor komponensei,  $r_{ij}^{min}$  pedig a kötéshossz.

3. Kémiai eltolódás (*chemical shift* - CS). A kémiai eltolódás elviekben megfelelő információt hordoz, hogy belőle a harmadlagos szerkezetet rekonstruálni lehessen (Shen et al., 2008), de ez gyakorlatban nem egy triviális feladat. A SHIFTX, és a későbbi SHIFTX2 programcsomagok egy olyan algoritmust implementálnak, amellyel fehérjék szerkezetéből lehet visszazámolni a kémiai eltolódásaikat (Han et al., 2011; Neal et al., 2003), ám ezek az algoritmusok sokáig nem voltak alkalmazhatóak MD szimulációkban, mert a formula nem differenciálható. Nem sokkal később kifejlesztettek differenciálható kémiaieltolódás-kifejezéseket, amelyek lehetővé tették, hogy az MD programcsomagok azokat is implementálják (Camilloni et al., 2013; Robustelli et al., 2010).

Én az itt bemutatott elemzésekben csak a sokaságok validációra használtam a kísérletes kémiai eltolódásokat. Általában egy adott atomtípusra (pl. Ca, N, NH, C) visszazámolt kémiai eltolódás korrelációját szokták megadni az aminosavszám függvényében a mért adatokhoz

képezt. Szerkezeti sokaság esetén a kémiai eltolódások adott atomra vett átlagát szokás megadni a 19. egyenlet szerint – én magam is ezt a módszert választottam. Általános tapasztalat, hogy az ilyen módon meghatározott, átlagos kémiai eltolódások korrelációja magasabb értéket ér el, mint az egyes replikákra számolt korrelációk átlaga – ez szintén szemléletes bizonyítéka a sokaságalapú megközelítés hatékonyságának.

#### 4.1.2 A kényszerfeltételek alkalmazása

A legtöbb molekuladinamikai programcsomagnak részét képezi számos kísérletes paraméter külső kényszerfeltételként történő alkalmazhatósága. Az NOE-k és  $S^2$  rendparaméterek együttes alkalmazása során felmerülő nehézségek kezelésére a GROMACS 4.5.5 molekuladinamikai programcsomagnak (Hess et al., 2008) a laboratóriumunkban fejlesztett kiterjesztését használtuk (Fizil et al., 2015).

A szerkezeti sokaságok generálásához szükség van egy energiafüggvényre, amelynek értéke a visszszámolt és a kísérletes paraméterek közötti eltérés nagyságától függ. Molekuladinamikai szimulációban van egy további feltétel: az energiafüggvénynek deriválhatónak kell lennie az atomi koordináták szerint, ugyanis így lehet kiszámolni a szimuláció során az egyes atomokra ható erőt.

Energiafüggvényként egy ún. félharmonikus potenciált alkalmaztam (Best & Vendruscolo 2004, Ángyán & Gáspári 2013):

$$E = \begin{cases} \alpha M(\rho - \rho_{min})^2 & \rho > \rho_{min} \\ 0 & \rho \leq \rho_{min} \end{cases} \quad (23)$$

ahol  $\alpha$  egy megfelelően megválasztott súlyfaktort,  $M$  pedig a sokaság méretét jelöli.  $\rho$  az alábbi képlettel definiált, ún. előrehaladási változó:

$$\rho = \frac{1}{N} \sum_k (X_{calc}^k - X_{exp}^k)^2 \quad (24)$$

$\rho_{min}$  az előrehaladási változónak a szimuláció során aktuálisan elért legalacsonyabb értéke.

A kényszerfeltételekkel megkötött molekuladinamikai szimuláció során az erőtérből származó energiafüggvényt ( $E_{FF}$ , ld. 9. egyenlet) kiegészítjük az imént bemutatott büntető energiattal ( $E_{restr}$ ):

$$E_{tot} = E_{FF} + E_{restr} \quad (25)$$

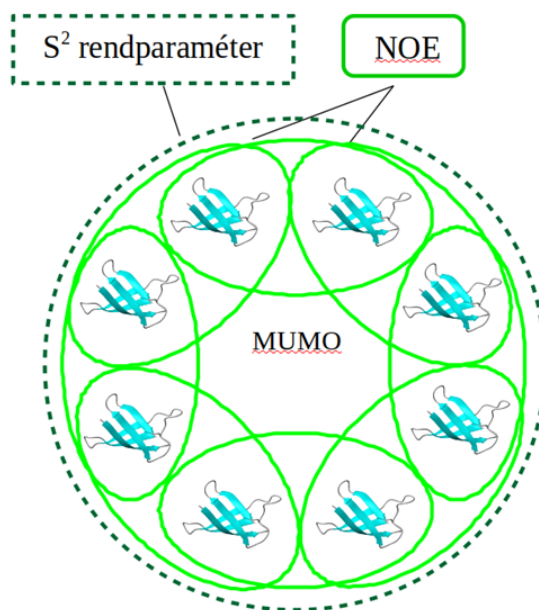
és az így kapott potenciális energiából számítjuk ki a szimuláció minden egyes lépésekor az atomokra eső erőt (Cavalli et al., 2013; Lange et al., 2008; Lindorff-Larsen et al., 2005).

#### 4.1.3 A MUMO-protokoll

Külső kényszerfeltételekkel megkötött MD szimuláció során az NOE-k és az  $S^2$  rendparaméterek együttes alkalmazása nem triviális, melynek megoldására a MUMO (*minimal under-restraining, minimal over-restraining*) algoritmust használtam.

E két paraméter külön-külön történő alkalmazása már bevett gyakorlat volt (Best & Vendruscolo, 2004; Bonvin et al., 1994), amikor Vendruscolo és munkatársai megmutatták, hogy az alul- és

túlillesztés elkerüléséhez ideális replikák száma nem azonos a két esetben: az NOE-kat párosával, a rendparamétereket pedig nagyobb, tipikusan 8 replikát tartalmazó alsokaságokra érdemes átlagolni. A MUMO protokollban (Richter et al., 2007) ezért a rendparamétereket a teljes alsokaságra, az NOE-kat pedig párosával átlagolják, de minden replika fele-fele arányban két párnak is tagja (10. ábra).



10. ábra: A MUMO-algoritmus (minimal under-restraining, minimal over-restraining) lényegének sematikus ábrázolása. Az algoritmus során több, kényszerfeltételekkel megkötött párhuzamos replikát futtatunk. A  $S^2$  rendparaméter-megkötéseket mindegyik replikára egyformán alkalmazzuk, még az NOE távolság-kényszerfeltételeket párosával átlagolva olyan módon, hogy minden egyes replikára a két-két szomszédjával vett átlagot határozzuk meg (Richter et al., 2007).

#### 4.1.4 Lokális illesztés

Többdoménes fehérjék belső dinamikája több, különböző időskálán történik: rendszerint a doméneken belüli gyors gerincmozgások függetlennek tekinthetők a domének egymáshoz viszonyított, lassú elmozdulásától (ld. 2.3.2. fejezet). Ezt feltétlenül figyelembe kell venni az  $S^2$  rendparaméterek külső kényszerfeltételként történő alkalmazásánál, mivel ezek minden esetben csak a gyors, ps-ns időskálájú belső mozgásokat tükrözik. Az  $S^2$ -megkötéseket tartalmazó MD szimulációk eredeti algoritmusai erre nem alkalmasak, mivel abban minden egyes szimuláció lépés során a párhuzamosan szimulált replikákat egy közös referencia-szerkezetre illesztjük (Best & Vendruscolo, 2004).

Hogy figyelembe vegyük a domének közötti lassú elmozdulást, ún. lokális illesztést kell alkalmaznunk. Ennek során az egyes doménekre külön-külön meghatározunk egy-egy referencia-szerkezetet, és a szimuláció során az illesztést, valamint a domének gerincdinamikájából származó  $S^2$ -értékek visszaszámolását és az ebből származó erőhatás számítását is külön végezzük el a doménekre. Ez az eljárás összhangban van a PSD-95 fehérje PDZ1-2 tandemjére mért  $S^2$  értékekkel, ahol a két doménre meghatározott eltérő korrelációs idő bizonyítja a domének független reorientációját (W. Wang et al., 2009) (ld. 2.1.4. fejezet)

A lokális illesztési algoritmust Zajácz-Epresi Nóra implementálta a GROMACS molekuladinamikai programcsomagnak a laboratóriumunkban fejlesztett kiegészítésébe, és az e dolgozat alapjául szolgáló tanulmányban került először publikálásra a PSD-95 fehérje PDZ1-2 tandemjének szimulációján bemutatva (Kovács et al., 2020). Az eljárás lehetővé teszi, hogy minden egyes  $S^2$  paraméter esetében külön meghatározzuk azt az atomcsoportot, amelyhez történő illesztéssel az  $S^2$  paraméterek szerkezetből való visszaszámolása történik. A doménekre történő alkalmazásnál értelemszerűen az 1. domén minden aminosavja esetében az  $S^2$  értékhez rendelt atomcsoport az 1. domén, míg a 2. domén  $S^2$  értékei esetében a 2. domén gerincatomjai. Az  $S^2$  értékek és az atomcsoportok megadása a topológia fájlban (.top) lehetséges.

#### 4.1.5 A szimulációk felparaméterezése

A PDZ tandemről készült szerkezeti sokaságok előállításához a MUMO protokollt használtam, az  $S^2$  paraméterek lokális illesztésével kiegészítve: a rendparamétereket 8 párhuzamos replikára alkalmaztam, a távolság-kényszerfeltételeket pedig párosával (10. ábra). A távolság- és  $S^2$ -kényszerfeltételek erőállandóit – számos előzetes futtatás tapasztalatai alapján – rendre  $8 \cdot 10^4$ -re és  $10^8$ -ra állítottam (ld. 23. egyenlet). A molekuladinamikai szimuláció természeténél fogva determinisztikus, ezért hogy a 8 párhuzamos szimuláció egymástól különböző trajektóriát járjon be, a kezdeti sebességeket véletlenszerűen rendeltem hozzá az atomokhoz.

A kötéshosszak állandó értéken tartásához a LINCS algoritmust használtam (Hess et al., 1997). A szimulációkat 300 K hőmérsékleten futtattam, a Berendsen-termostát alkalmazásával (Berendsen et al., 1984). A fehérjemolekulákra optimalizált AMBER99SB-ILDN erőteret alkalmaztam (Lindorff-Larsen et al., 2010). A téglatest alakú szimulációs doboz mérete úgy lett megválasztva, hogy a kiindulási fehérjeszerkezet – amelyben a két domén egymástól távol, kinyújtott interdomén orientációban helyezkedik el – bármely atomja legalább 15 Å távolságra legyen a doboz szélétől. Explicit oldószerként a TIP3P (Jorgensen et al., 1983), implicit oldószerként pedig a GBSA (Genheden & Ryde, 2015) oldószermodellt használtam. Explicit oldószer használata esetén a szimulációs dobozba pontosan annyi iont ( $\text{Na}^+$  kationt vagy  $\text{Cl}^-$  aniont) helyeztem, hogy rendszer semleges legyen: ez mind a szabad, mind a komplex forma esetén 4 kationt jelentett. A szimulációkat 20 ns-ig futtattam 1 fs időlépésekkel. A trajektóriát minden 100 ps-ban mintavételeztem, ám az első 2000 ps mintavételeit a hőmérséklet és nyomás ekvibráció céljából elhagytam. Mivel minden szimulációban 8 párhuzamos replikát alkalmaztam és mintavételeztem, így ez a protokoll sokaságonként 1448 replikát eredményezett.

A szimulációk futtatását három lépés előzte meg: energiaminimalizáció, és az NVT-, valamint NPT-ekvibráció. Az energiaminimalizáció konjugált gradiens módszerrel történt, 0,01 nm-es lépéstávolsággal, minden 200. lépésben egy gradiens módszer lépés beiktatásával, 200 kJ/mol/nm legnagyobb erőhatásig. Explicit oldószerben az energiaminimalizációt megismételtem a doboz oldószerrel történő feltöltése után is, ám ebben az esetben a legnagyobb erő 2000 kJ/mol/nm volt, és a 200 lépésenként beiktatott gradiens módszert elhagytam. Az oldószer NVT- és NPT-ekvibrációja a

fehérje atomjainak rögzítése mellett történt, 500-500 ps-ig, 1 fs-os időlépéssel. Mindkét esetben a Berendsen-termostátot (Berendsen et al., 1984), az NPT ekvilibráció esetén pedig Parrinello-Rahman-barosztátot (Parrinello & Rahman, 1980) is alkalmaztam.

## 4.2 Dinamikus szerkezeti sokaságok validálása

Az előállított szerkezeti sokaság validációjához be kell vezetnünk egy jósági mérőszámot, amely kifejezi az összes kísérletes  $X_{exp}^k$  és visszszámolt  $X_{calc}^k$  paraméter közötti eltérés mértékét. Az összes paraméterre használható, és gyakran alkalmazott mérőszám a Pearson-féle korrelációs koefficiens, mely egy adott szerkezetre vagy a sokaság adott replikájára a következő képlettel fejezhető ki:

$$R^2 = \frac{\sum_k (X_{calc}^k - \bar{X}_{calc}^k)(X_{exp}^k - \bar{X}_{exp}^k)}{\sqrt{\sum_k (X_{calc}^k - \bar{X}_{calc}^k)^2} \sqrt{\sum_k (X_{exp}^k - \bar{X}_{exp}^k)^2}} \quad (26)$$

ahol  $X_{calc}^k$  és  $X_{exp}^k$  a  $k$ -adik aminosavra meghatározott visszszámolt és kísérletes paraméter,  $\bar{X}_{calc}^k$  és  $\bar{X}_{exp}^k$  pedig ugyanezen paraméterek átlagos értékeik. A szummázás az aminosavakra történik.

A korrelációs koefficiens szignifikanciáját Student-féle t-próbastatisztika elvégzésével igazoljuk:

$$t = \frac{R}{\sqrt{(1 - R^2)(n - 2)}} \quad (27)$$

ahol  $R$  a Pearson-féle korrelációs együttható,  $n$  pedig a mintaelemszám. A  $t$  érték Student-féle t-eloszlást követ, melynek adott  $p$  szignifikancia szintre meghatározott  $t_p$  értékeit táblázatok tartalmazzák a mintaelemszám függvényében. Tetszőlegesen választott  $p$  szignifikancia annak valószínűségét fejezi ki, hogy  $t_p > t$ , ahol  $t$  a próbastatisztika elvégzésével adódó érték. Szemléletes képpel kifejezve  $p$  annak valószínűsége, hogy tévedünk, ha elutasítjuk a nullhipotézist, vagyis azt a feltevést, hogy a két változó eloszlása független egymástól. Általánosan igaz, hogy nagyobb mintaelemszámmal könnyebb bizonyítani a két változó közötti szignifikáns korrelációt.

A kémiai eltolódások visszszámolásához a SHIFTX2 programcsomagot használtam (Han et al., 2011), amely szerkezeti sokaságokra is használható: átlagolja az egyedi replikákra számolt értékeket. Az  $S^2$  rendparamétereket a parvulinok vizsgálata során a CoNSEnsX szerverrel (Dudola et al., 2017), a PDZ tandem vizsgálata során pedig a következő fejezetben bemutatott ProDitor, saját fejlesztésű programcsomaggal számoltam. Mind a Pearson-korrelációs együttható, mind pedig a t-próbából adódó  $t$ -érték és  $p$  szignifikancia szint kiszámolásához az Apache Commons<sup>1</sup> java könyvtárat használtam, melyet szintén a ProDitor programcsomagban implementáltam.

<sup>1</sup> <https://commons.apache.org/proper/commons-math/>

### 4.3 Fehérjeszerkezeti adatfájlok elemzése a ProDitor saját fejlesztésű programcsomaggal

Fehérjeszerkezeti – PDB – fájlok szerkesztésére és vizualizációjára számos szoftver áll rendelkezésre, ám gyakran előfordul, hogy ezek dokumentációja pontatlan vagy nem teljes, illetve az alkalmazott algoritmusok nem módosíthatóak vagy nem lehetséges a paraméterek felhasználói beállítása. Az alkalmazott módszerek reprodukálhatóságának érdekében egy saját programcsomagot fejlesztettem a fehérjeszerkezeti sokaságokat tartalmazó PDB fájlok beolvasására, szerkesztésére és alapvető elemzési műveletek elvégzésére, melyet ProDitor (**Protein Editor**) néven kereszteltem el<sup>2</sup>. A ProDitor program saját műveletek elvégzésén kívül interfészt képez egy sor további, az elemzés során alkalmazott szoftverhez is, melyeknek megérti és szerkeszteni képes kimeneti és bemeneti fájljait, és bizonyos esetben ezeket képes a háttérben futtatni. Ezek közé tartozik a GROMACS 4.5.5, a DSSP, a SHIFTX2, a ProDy és a gnuplot.

#### 4.3.1 RMSD-számolás

Egy szerkezeti sokaság adott részének vagy egészének konformációs kompaktságát a globális és lokális RMSD-értékekkel jellemeztem. Két, azonos atomokat tartalmazó  $I$  és  $J$  szerkezet közötti RMSD-t a következő egyenlettel számoljuk ki:

$$RMSD_{I,J} = \sqrt{\frac{1}{n} \sum_{k=1}^n |\mathbf{r}_k^I - \mathbf{r}_k^J|^2} \quad (28)$$

ahol  $\mathbf{r}_k^I$  és  $\mathbf{r}_k^J$  a  $k$ -adik atom pozíciója  $I$  és  $J$  szerkezetekben,  $n$  pedig a számoláshoz felhasznált atomok száma. Fehérjékre legtöbbször gerinckonformáció diverzitását szokás meghatározni, ezért az RMSD számolásakor – ellenkező értelmű utalás hiányában – minden esetben csak a peptidgerinc N, C $\alpha$  és C atomjait vettem figyelembe.

Egy szerkezeti sokaság globális RMSD értéke nem más, mint az összes szerkezetnek az átlagos atomi pozíciókat tartalmazó szerkezethez viszonyított RMSD értékeinek átlaga:

$$RMSD = \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{n} \sum_{k=1}^n |\mathbf{r}_k^i - \bar{\mathbf{r}}_k|^2} \quad (29)$$

ahol  $N$  a sokaság mérete,  $\mathbf{r}_k^i$  a  $k$ -adik atom pozíciója az  $i$ -edik replikában,  $\bar{\mathbf{r}}_k$  pedig a  $k$ -adik atom átlagos pozíciója.

<sup>2</sup> A ProDitor teljes forráskódja elérhető a következő linken: <http://users.itk.ppke.hu/~kovbe14/downloads/>

A fentiekhez hasonlóan definiálhatjuk a lokális RMSD-t, amely a fehérje egy bizonyos régiójára, jellemzően egy-egy aminosavra meghatározott RMSD. Én az aminosavakra vett lokális RMSD számolásakor minden esetben az adott aminosavnak az N, Ca és C atomjait használtam fel.

Mind a globális, mind a lokális RMSD számolást megelőzi a szerkezetek illesztése, ami garantálja az algoritmus reprodukálhatóságát.

### 4.3.2 Sokaságok illesztése

Két, azonos atomokat tartalmazó fehérjeszerkezet összehasonlításakor elengedhetetlen a szerkezetek egymáshoz való illesztése. A ProDitor programcsomag ehhez a Kabsch-algoritmus implementációját használja fel (Kabsch, 1976b, 1976a). Ennek során azt a rotációs mátrixot keressük, amely egy adott ponthalmazt egy másik ponthalmazhoz képest olyan orientációba forgat, hogy így a két ponthalmaz közötti RMSD minimális legyen.

Első lépésként mindkét ponthalmaz geometriai középpontját az origóba mozdítjuk. Ezt követően felírjuk a két ponthalmaz koordinátáit tartalmazó  $N \times 3$  méretű  $P$  és  $Q$  mátrixokat, ahol  $N$  a ponthalmaz mérete. Ezekből definiáljuk a  $H$  kovariancia mátrixot:

$$H = P^T Q \quad (30)$$

Az  $R$  optimális rotációs mátrix végül a következő formulából adódik:

$$R = V \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d \end{pmatrix} U^T \quad (31)$$

ahol  $V$  és  $U$  mátrixok  $H$  szinguláris érték felbontásával (SVD – *singular value decomposition*) jönnek létre:

$$H = USV^T \quad (32)$$

$d$  pedig a következő mátrix determinánsa:

$$d = \det(VU^T) \quad (33)$$

Egy sokaság illesztésekor a ProDitor programcsomag először kiszámolja minden atom átlagos pozícióját, majd ehhez az ponthalmazhoz illeszti az egyes szerkezeteket, és végül kiszámolja az egész sokaság RMSD-jét. Ezt a ciklust iteratíván addig folytatja, amíg az RMSD már nem változik számottevően.

## 4.4 A leghangsúlyosabb belső mozgások azonosítása főkomponens analízissel

Egy globuláris fehérjéről készült dinamikus fehérjeszerkezeti sokaság legtöbbször az átlagos konformáció körüli fluktuációt reprezentálja, ez a dinamika azonban számos belső mozgás kombinációjaként adódik. Ezek azonosítása és dekompozíciója igen hasznos a dinamikát alapvetően meghatározó atomi szintű kölcsönhatások azonosításakor.

Egy  $N$  atomot tartalmazó fehérjét felfoghatjuk a  $3N$  dimenziós tér egy pontjának, amennyiben a teljes szerkezetét az atomjainak Descartes-koordinátájával határozzuk meg. Ilyen értelemben a belső mozgások egy  $3N$  dimenziós vektor mentén történő elmozdulások lennének. Ez a reprezentáció azonban indokolatlanul bonyolult, hiszen a tapasztalat szerint legtöbbször egy-egy jól definiált régió, akár néhány atom korrelált elmozdulása is értelmes információt hordoz szerkezeti biológiai szempontból. Ezért rendkívül hasznosak a dimenziócsökkentő algoritmusok, amelyek közül fehérjeszerkezeti sokaságok elemzésekor a leginkább bevett módszer a főkomponens-elemzés (PCA) (Amadei et al., 1993; A. Bakan & Bahar, 2009; Kitao & Go, 1999).

A főkomponens-elemzés során a leghangsúlyosabb belső mozgások és amplitúdójuk a  $C_\alpha$  kovarianciamátrix sajátvektoraiként és sajátértékeiként adódnak. Az  $A$  kovarianciamátrix elemei

$$A_{i,j} = \langle (\mathbf{r}_i(t) - \bar{\mathbf{r}}_i)(\mathbf{r}_j(t) - \bar{\mathbf{r}}_j) \rangle \quad (34)$$

ahol  $\mathbf{r}_i(t)$  és  $\mathbf{r}_j(t)$  az időfüggő atomi Descartes-koordináták,  $\bar{\mathbf{r}}_i$  és  $\bar{\mathbf{r}}_j$  pedig  $i$  és  $j$  atomok időátlagolt Descartes-koordinátái. A kovarianciamátrix sajátvektorait *főkomponenseknek* (PC) hívjuk, és fontos jellemzőjük, hogy egymásra ortogonálisak. Ennek praktikus következménye, hogy az így azonosított belső mozgások szükségszerűen egymással nem korreláló mozgások lehetnek. A főkomponensek a  $3N$  dimenziós tér vektorai, amelyek koordinátáira vetítve a fehérje atomjainak átlagos pozícióját könnyen vizualizálhatóak a megfelelő belső mozgások.

Az algoritmus egy másik fontos jellemzője, hogy a főkomponensek a sajátértékek szerinti csökkenő sorrendben adódnak, vagyis az első főkomponens reprezentálja a fehérje leghangsúlyosabb belső mozgását, és így tovább. Tapasztalat szerint az első néhány, legfeljebb 10 főkomponens magában foglalja a szerkezetileg releváns összes belső mozgást. A sajátértékek négyzetének százalékos megoszlásával könnyen kifejezhető, hogy egy adott belső mozgás a teljes variancia mekkora hányadát teszi ki.

Bizonyos esetekben relevánsak lehetnek az oldalláncok mozgásai is, az itt bemutatott főkomponens-elemzések mindegyikénél csak a peptidgerinc dinamikáját vizsgáltuk. Tekintettel arra, hogy az amidkötésre szigorú geometriai feltételek vonatkoznak, esetünkben elegendő, ha a főkomponens-elemzéskor csak az aminosavak  $C_\alpha$  atomjait vesszük figyelembe. Ezen az elven működik a munkám során használt ProDy programcsomag (Ahmet Bakan et al., 2011, 2014).

## 4.5 Kísérletes $^1\text{H}$ - $^1\text{H}$ NOE-kból származó távolság-kényszerfeltételek sztereoszimmetrizálása

Általános tapasztalat, hogy a NOESY spektrumok asszignációja során a geminális hidrogénekhez tartozó kémiai eltolódásokat könnyű felcserélni. Mivel az ezekből származó távolság-kényszerfeltételek többnyire sztereospecifikusak, a hordozott információ gyakran hamissá vagy megbízhatatlanná válik.



Alapelvként az  $r^{-6}$  átlagolást használtam (20. egyenlet), és minden kényszerfeltétel sztereospecifikus jellegét megszüntettem. Ezáltal a lista redundanciája nőhet, így ezt a redundanciát is megszüntettem. Például ha egy adott aminosav HB1 és HB2 atomja és egy X atom között két külön kényszerfeltétel volt az eredeti listában, ezeket én a (HB1 vagy HB2) – X kényszerfeltétellel helyettesítettem, a két eredeti távolság közül a nagyobbikat felhasználva. Az eljárás célja természetesen az esetlegesen hibás sztereokémiájú, pl. HB1 – X helyett HB2 – X formában megadott kényszerfeltételek felhasználhatóságának javítása. Az említett lépéseket a csoportunkban kidolgozott Perl szkriptekkel valósítottam meg.

## 4.6 Felhasznált szoftverek

Munkám során harmadik fél által fejlesztett és publikált szoftvereket is felhasználtam, amelyeket rövid magyarázattal ellátva felsorolok.

Fehérjeszerkezetekről készült ábrákhoz a Chimera és a ChimeraX programokat (Goddard et al., 2018; Pettersen et al., 2004), kémiai képletek rajzolásához a MarvinSketch programcsomagot (<https://www.chemaxon.com>), grafikonok és diagramok létrehozásához a gnuplot programot, valamint a főkomponensek dinamikus vizualizációjához a VMD programot (Humphrey et al., 1996) használtam. A főkomponens-elemzést a ProDy-val végeztem (Ahmet Bakan et al., 2011, 2014). A szimulációkat a GROMACS 4.5.5-tel futtattam (Hess et al., 2008), a kémiai eltolódásokat a SHIFTX2-vel határoztam meg (Han et al., 2011) az előállított sokaságokból. A kísérletes szekvenciákhoz az UniProt-ban (Bateman, 2019), a kísérletes szerkezetekhez pedig a PDB-ben (Berman et al., 2000) fértem hozzá. A többszörös szerkezetillesztést a Mammoth-MULT-tal (Lupyan et al., 2005), a többszörös szekvenciaillesztést a Clustal Omega-val (Sievers et al., 2011), a domének azonosítását a BLAST (Altschul et al., 1990) és a SMART (Letunic & Bork, 2018) algoritmusaival végeztem el. A szerkezeti sokaságok validációjához a CoNSEnsX<sup>+</sup> webszervert használtam (Dudola et al., 2017). A pKa értékek predikcióját a PropKa programmal végeztem (Søndergaard et al., 2011).

## 5 A PSD-95 PDZ1-2 tandemjének ligandumfüggő intra- és interdomén dinamikai vizsgálata

### 5.1 A szerkezeti sokaságok előállítása

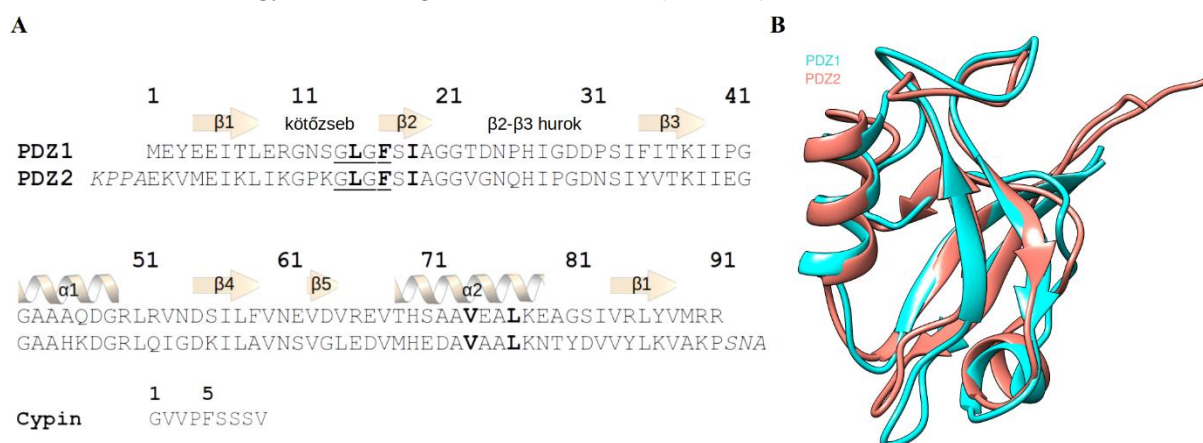
#### 5.1.1 Kiindulási adatok előkészítése

A vad-típusú PSD-95 fehérje 724 aminosavból áll és 5 domént tartalmaz: 3 PDZ domént, egy SH3 és egy GK domént (F3. táblázat). Megjegyzendő, hogy a pontos doménhatárok megállapítása különböző algoritmusok alapján történhet, ezért az egyes adatbázisban eltérő lehet. Én a BLAST keresőmotorja mögött működő SMART adatbázis fehérjedoménekre vonatkozó adatait vettem irányadónak.

A manuálisan kurált UniProt adatbázisban fellelhető információk alapján a PSD-95 fehérje három homológjának szekvenciája áll rendelkezésünkre, melyek emberből, patkányból és egérből származnak. E három homológ teljes szekvencia-illesztését az F4. ábra tartalmazza, a doménhatároknak az UniProtban szereplő feltüntetésével. Az eltérés minimális: a humán PSD-95 három aminosavban tér el az egérben található homológtól, és egyetlen aminosavban a patkányban található homológtól. Ez az utóbbi eltérés csak az SH3 domént érinti. Az egérből származó PSD-95-ről nincsen elérhető kísérletes szerkezeti információ (F5. táblázat), és a patkány, illetve humán PDZ1-2 tandem szekvenciája nem különbözik, tehát ezek közül bármely szerkezet felhasználható további vizsgálathoz.

A kiindulási szerkezet kiválasztásához összegyűjtöttem a PDB-ben fellelhető kísérletesen meghatározott szerkezeteket a PDZ1 és a PDZ2 doménről (F5. táblázat). Molekuladinamikai szimulációkban történő felhasználhatóság szempontjából előnyben részesítjük az NMR-es szerkezeteket, és azok közül is azokat, amelyek a teljes PDZ1-2 tandemet tartalmazzák és a lehető legtöbb, dinamikai információt hordozó NMR-es paraméterrel rendelkeznek. Ez alapján magától értetődő a Wang és mtsai (W. Wang et al., 2009) által meghatározott, 2KA9 PDB kóddal jelölt szerkezet kiválasztása. Ez a szerkezet tartalmazza az F3. táblázatban bemutatott számozás szerinti teljes PDZ1-2 tandemet, egy 3 aminosav hosszú, PDZ2-höz tartozó farokrégiót (vagyis az eredeti számozás szerinti 61-249. aminosavakat), valamint ligandumként a két PDZ domén egyik kötőpartnerének: a cypin-nek 9 C-terminális aminosavját. A PDB fájl 20 modellt tartalmaz, melyek sem az interdomén orientációban, sem az intradomén konformációban nem térnek el jelentősen egymástól. Kiindulási szerkezetnek a szerzők által reprezentatívnak megjelölt 1. konformert használtam. A továbbiakban átveszem a 2KA9-es PDB szerkezetben használt aminosav-számozást, amely a PDZ1 domén első aminosavját jelöli meg 1-esként. Ez alapján a doménhatárok: 1-91 a PDZ1, és 91-186 a PDZ2 doménre.

A PSD-95 fehérje PDZ1 és PDZ2 doménje 52%-os szekvencia-azonosságot mutat egymással, szerkezetileg pedig teljesen homológok: az őket felépítő aminosavak és a másodlagos szerkezeti elemek között kölcsönösen egyértelmű megfelelés mutatkozik (11. ábra).



11. ábra: A PSD-95 fehérje PDZ1 és PDZ2 doménjének szekvencia- és szerkezetillesztése. **A:** a PDZ1 és PDZ2 domén szekvenciaillesztése. A PDZ2 N-terminálisa (linker régió) és farokrégiója dőlt betűvel van jelölve. A GLGF motívumhoz tartozó aminosavak aláhúzással vannak kiemelve. A hidrofób kötőseb aminosavjai félkövérrel vannak szedve. Alul a 2KA9 PDB szerkezetben lévő cypin ligandum C-terminális szekvenciája látható. **B:** A PDZ1 és PDZ2 domén szerkezetillesztése.

A 2KA9 szerkezethez tartozó kísérletes adatok közül a  $^1\text{H}$ - $^1\text{H}$  NOE-k letölthetők a PDB adatbázisból. A jobb reprodukálhatóság érdekében a sztereospecifikus távolságokat sztereoszimmetrizáltam, és az így keletkező duplikátumokat eltávolítottam. E művel végén 3169 távolság-kényszerfeltétel maradt a szerkezethez tartozó eredeti 3281  $^1\text{H}$ - $^1\text{H}$  NOE-ből. Ahogy az a 3. táblázatban is látható, ezeknek a legnagyobb részét a PDZ1, illetve a PDZ2 doménhez tartozó intradomén kényszerfeltételek teszik ki. Külön figyelmet érdemel, hogy a listában egyetlen interdomén  $^1\text{H}$ - $^1\text{H}$  NOE sem szerepel: ez összhangban van a cikk szerzőinek azon megállapításával, amely szerint a komplexált PDZ1-2 tandemben jelentősen megnövekedett interdomén dinamika figyelhető meg. A PDZ domének és a ligandumok közötti NOE-kat leszámítva a listában 1148 aminosavon belüli, 742 szomszédos aminosavak közötti és 1256, szekvenciában egymástól távol lévő aminosav közötti NOE található.

	PDZ1	linker	PDZ2	farokrégió	ligandum1	ligandum2
1. PDZ1	1455					
2. linker	9	34				
3. PDZ2	0	7	1602			
4. farokrégió	0	4	17	18		
5. ligandum1	11	0	0	0	0	
6. ligandum2	0	0	12	0	0	0

3. táblázat: A 2KA9 szerkezethez tartozó sztereoszimmetrizált  $^1\text{H}$ - $^1\text{H}$  távolság-kényszerfeltételek megoszlása a PDZ domének, valamint a ligandumok között.

A sztereoszimmetrizált NOE-kat megvizsgáltam a PDB-ben elérhető 20 szerkezetre, és azt találtam, hogy mindössze 3 távolság-kényszerfeltétel sérül, legfeljebb 0.67 Å-mel. (Megjegyzendő, hogy a sztereoszimmetrizálást követő, sokaság-alapú átlagolásnak jóval tágabb a rossz minőségű NOE-

kal szemben mutatott tűrőképessége, mint a replika-alapú távolság-megkötésnek, ahogy azt az imént említett fejezetben már részletesen is kifejtettem.) Ezek közül egy a PDZ1 domén  $\beta 1$  és  $\beta 6$  számai között, egy másik szintén a PDZ1 domén  $\alpha 2$  hélixében, egy pedig a PDZ2 domén  $\alpha 1$  hélixében található. Alapos szemrevételezés és megfontolás után úgy döntöttem, ezeknek az NOE-sérüléseknek az alacsony száma, csekély mértéke és elhelyezkedése a fehérjében nem indokolja, hogy kihagyjam őket a szimulációból.

A tanulmányban publikált  $^1\text{H}$  és  $^{15}\text{N}$  kémiai eltolódásokat, valamint az  $S^2$ , azaz az amid N-H általános rendparamétereket a szerzők személyes kérésre a rendelkezésemre bocsátották (W. Wang et al., 2009). A rendparamétereket 2 frekvencián, 600 MHz-en és 800 MHz-en is meghatározták. Az adatsorok teljességét figyelembe véve a 600 MHz-en felvett adatokból származó amid rendparaméterek használata mellett döntöttem. Ebből összesen 159 állt rendelkezésre, amelyekből azonban az 4.1.4. fejezetben bemutatott lokális illesztési algoritmus működésénél fogva csak azokat használhattam fel, amelyek a doménhatárokon belüli aminosavakhoz tartoznak: így összesen a PDZ1 és a PDZ2 doménre rendre 79, illetve 80  $S^2$  maradt. A szerzők által elküldött adatsor tartalmazott továbbá 170 amid N és H kémiai eltolódást, melyeket kényszerfeltételként nem, hanem csak az előállított sokaságok validációjához használtam fel.

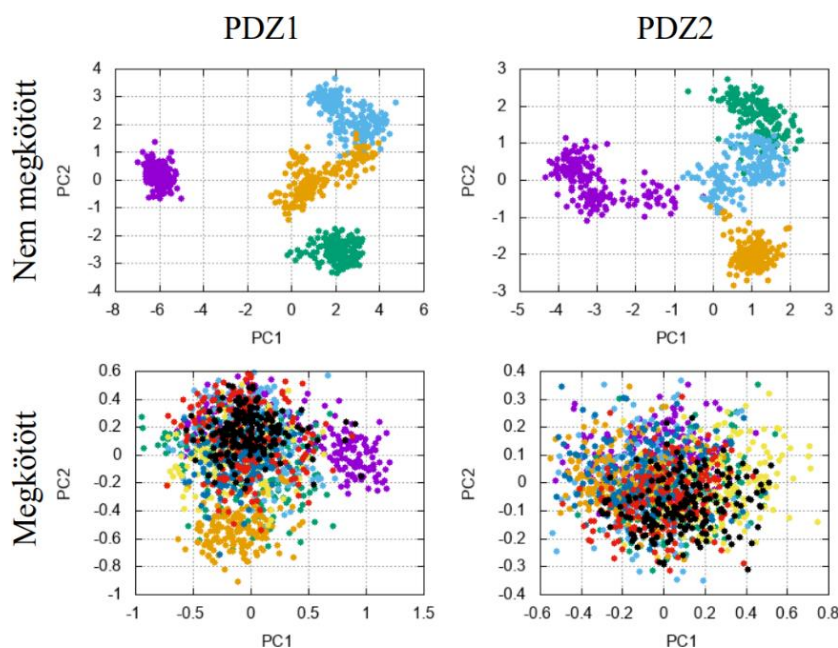
### 5.1.2 A molekuladinamikai szimulációk futtatása

A szerkezeti sokaság előállítását több tesztszimuláció előzte meg. Ezek során meggyőződtem arról, hogy a kiindulási szerkezet ép és valóban energiaminimalizálható. Ellenőriztem, hogy a megfigyelhető mennyiségek – a hőmérséklet, a nyomás, az energia, a kényszerfeltételekből származó erőhatások – reális értékeket vesznek fel, és a kezdeti szakaszt leszámítva a fluktuációjuk nagyjából egyenletes. Cél volt, hogy a kényszerfeltételek erőállandóját olyan magasra tudjam állítani, amely még nem torzítja el a szerkezetet, ám kellő erővel kényszeríti a molekulát a kísérletes paramétereknek megfelelő dinamikára. Rendre azt tapasztaltam, hogy a LINCS algoritmus elég érzékeny ahhoz, hogy a kényszerfeltételekhez beállított irreálisan magas erőállandók esetén figyelmeztet vagy hibaüzenetet ad, mert ilyenkor többnyire előfordul, hogy valamely torziós szög kiugróan magas sebességgel kezd rotálódni. Az implicit oldószermodellben futtatott szimulációk nagyjából egy nagyságrenddel gyorsabbak, mint az explicit oldószermodellben futtatottak, ezért kiválóan alkalmasak voltak próbaszimulációk végzésére és az előzetes eredmények kiértékelésére. Implicit oldószerben lehetőségem volt a többinél jóval hosszabb, 100 ns-os próbaszimuláció futtatására is, amely során meggyőződtem arról, hogy a fehérjemolekula által több rövid szimuláció során lefedett konformációs tér nem különbözik lényegesen egyetlen hosszú szimulációtól.

A tesztszimulációk rávilágítottak a távolság-kényszerfeltételek használatának szükségességére. A tapasztalat alapján ugyanis azok a párhuzamos trajektóriák, amelyek szimulációikor nem használtak távolság-kényszerfeltételeket, látszólag divergáltak egymástól, és egy-egy olyan konformáció körül

koncentrálódtak, amelyek egymástól jól elkülönülnek a kétdimenziós PCA ábrázoláson. Távolság-kényszerfeltételek alkalmazása során ez az effektus megszűnt (12. ábra).

Bár a szimulációk során a cél az adott rendszer megfigyelt állapotának lehető leghitelesebb reprodukciója, az elérhető kísérletes paraméterek szűkös száma olykor kompromisszumokra kényszerítheti a számítógépes biológust. Wang és mtsai. például a PDZ1-2 tandemnek csak a komplexált állapotára vonatkozó eredményeit publikálták: a szabad fehérjéről nem sikerült elég adatot gyűjteniük a “gyenge jelhomogenitás és az amidjelek jelentős részének eltűnése” miatt (W. Wang et al., 2009). Azonban azért, hogy a ligandum kötődésének a fehérje belső dinamikájára gyakorolt hatását vizsgálni tudjam, úgy döntöttem, a kísérletes adatok egy részét a szabad forma előállításakor és validálásakor is felhasználom. Az NOE-kra vonatkozó, imént felvázolt effektus indokoltá tette, hogy ugyanazokat az intradomén NOE-kat alkalmazzam távolság-kényszerfeltételként a szabad formára, amelyeket a komplex állapotban mértek meg, ugyanis ezek a PDZ domének szerkezeti stabilitását szolgálják. Hasonlóan elfogadhatónak láttam, hogy a kémiai eltolódás adatokat egyaránt felhasználjam a szabad és komplex formáról készült szerkezeti sokaság validációjához, ugyanis a peptiddel való titráláskor tapasztalt szerkezeti változás elhanyagolhatóan, a kémiaieltolódás-változás pedig minimálisnak mutatkozott (W. Wang et al., 2009). (Megjegyzendő, hogy – amint azt a következő fejezetben számszerűen is bemutatom – ezt a döntés a sokaságok validációja is igazolta utólag.)



12. ábra: Intradomén távolság-kényszerfeltételekkel (NOE) megkötött és megkötések nélkül futtatott PDZ1-2 tandem tesztszimulációk összehasonlítása. A különböző színek a párhuzamos trajektóriákat reprezentálják. A nem megkötött sokaság 4, a megkötött sokaság 8 párhuzamos replikából készült. A megkötött sokaságokban az NOE kényszerfeltételeket párosával alkalmaztam (ld. 10. ábra.) A szimulációk 20 ns-ig futottak, a kiértékeléskor az első 2 ns elhagyásával. A főkomponenseket a szűkített doménhatárookra (4-88, illetve 99-183 aminosav) határoztam meg.

Az alkalmazott rendparaméterek – különösen a korábban tárgyalt lokális illesztési algoritmussal alkalmazva – csak az intradomén gyors mozgásokra hatnak kényszerfeltételként: a két domén relatív reorientációjára ilyen módon nem vonatkozik megkötés. Mivel azonban a ligandumkötésnek éppen a

belső dinamikára gyakorolt hatását szerettem volna vizsgálni, a szabad állapotra nem alkalmaztam az – amúgy komplexált formára meghatározott – rendparamétereket.

Az elvégzett tesztszimulációk eredményeképpen három, explicit oldószermodellben előállított sokaságot választottam ki, melyeket további, részletes analízisnek vettem alá: egyet a szabad, kettőt pedig a komplex formáról. Mindhárom sokaság előállításakor alkalmaztam a kísérletes  $^1\text{H}$ - $^1\text{H}$  NOE-kat távolság-kényszerfeltételként, továbbá az egyik komplexált sokaság előállításánál a lokálisan illesztett  $S^2$  paramétereket is felhasználtam. Az egyszerűség kedvéért ezekre a sokaságra a továbbiakban *szabad*, *komplex* és *megkötött komplex* formaként fogok hivatkozni, ahol a megkötöttség az  $S^2$  kényszerfeltételek alkalmazására vonatkozik. A három sokaság uniójára mint *teljes sokaság*, valamint a két nem megkötött sokaság, illetve a két komplex sokaság uniójára mint *nem megkötött sokaság*, illetve *egyesített komplex sokaság* fogok hivatkozni (4. táblázat).

	<b>Szabad sokaság</b>	<b>Komplex sokaság</b>	<b>Megkötött komplex sokaság</b>
<b>Összevont sokaságok:</b>			
<i>Teljes sokaság</i>	<b>X</b>	<b>X</b>	<b>X</b>
<i>Nem megkötött sokaság</i>	<b>X</b>	<b>X</b>	
<i>Egyesített komplex sokaság</i>		<b>X</b>	<b>X</b>

4. táblázat: A három előállított PDZ tandem-sokaság (oszlopok), valamint az azok kombinációjával alkotott további sokaságok (sorok), és ezek elnevezése.

## 5.2 A szerkezeti sokaságok elemzése

### 5.2.1 A szerkezeti sokaságok validálása

Az előállított sokaságokból kivágtam a PDZ1 és PDZ2 domént, és a fehérjegerinc N, Ca és C atomjaira meghatároztam a globális RMSD értékeket (5. táblázat). Ezek azt igazolják, hogy mindkét domén stabil háromdimenziós szerkezettel rendelkezik. A lokális RMSD értékek mutatnak némi fluktuációt a szekvencia mentén: az N- és C-terminális farokrégiókra, valamint a  $\beta$ 1- $\beta$ 2 és a  $\beta$ 2- $\beta$ 3 hurokrégiókban magasabbak. Ez azonban összefügg az egyes aminosavakra alkalmazott távoli távolságkényszerfeltételek számával (F2. ábra). Épp emiatt elfogadott, hogy a lokális RMSD értékekből nem szabad a fehérje egyes régióinak belső dinamikájára következtetni.

Sokaság	Sokaság mérete	Kémiai eltolódás korrel.		S <sup>2</sup> korrel.	RMSD (Å)	
		N	H		PDZ1	PDZ2
<i>Eredeti PDB</i>	20	0,83	0,41	0,10	0,57 ± 0,01	0,62 ± 0,04
<i>Szabad</i>	1448	0,85	0,55	0,22	0,73 ± 0,01	0,66 ± 0,01
<i>Komplex</i>	1448	0,85	0,55	0,16	0,75 ± 0,02	0,58 ± 0,01
<i>Megkötött komplex</i>	1448	0,85	0,55	0,88	0,67 ± 0,01	0,63 ± 0,02

5. táblázat: Az előállított PDZ1-2 tandem sokaságok validációja a kísérletes paramétereknek való megfeleléssel. Az RMSD értékeket a kiindulási szerkezet (PDB: 2KA9) számozása szerinti 1-91 és 96-186 aminosavakra határoztuk meg a PDZ1 és a PDZ2 domén esetén, csak a peptidgerinc Ca, N és C atomjait felhasználva. A korrelációt a visszszámolt kémiai eltolódással és S<sup>2</sup> rendparaméterekkel 4.2. fejezetben bemutatottak szerint végeztem.

Meghatároztam a három szerkezeti sokaságból visszszámolt amid N és H kémiai eltolódások Pearson-korrelációját a kísérletes adatokkal (5. táblázat). Tekintettel arra, hogy a kémiai eltolódásokat nem használtam fel a sokaságok előállításánál, a jelentős javulás a H kémiai eltolódásban, valamint az enyhe javulás a N kémiai eltolódásában az eredeti PDB sokasághoz képest arra utal, hogy a létrehozott sokaságok elfogadhatóan tükrözik a kísérletes értékeket.

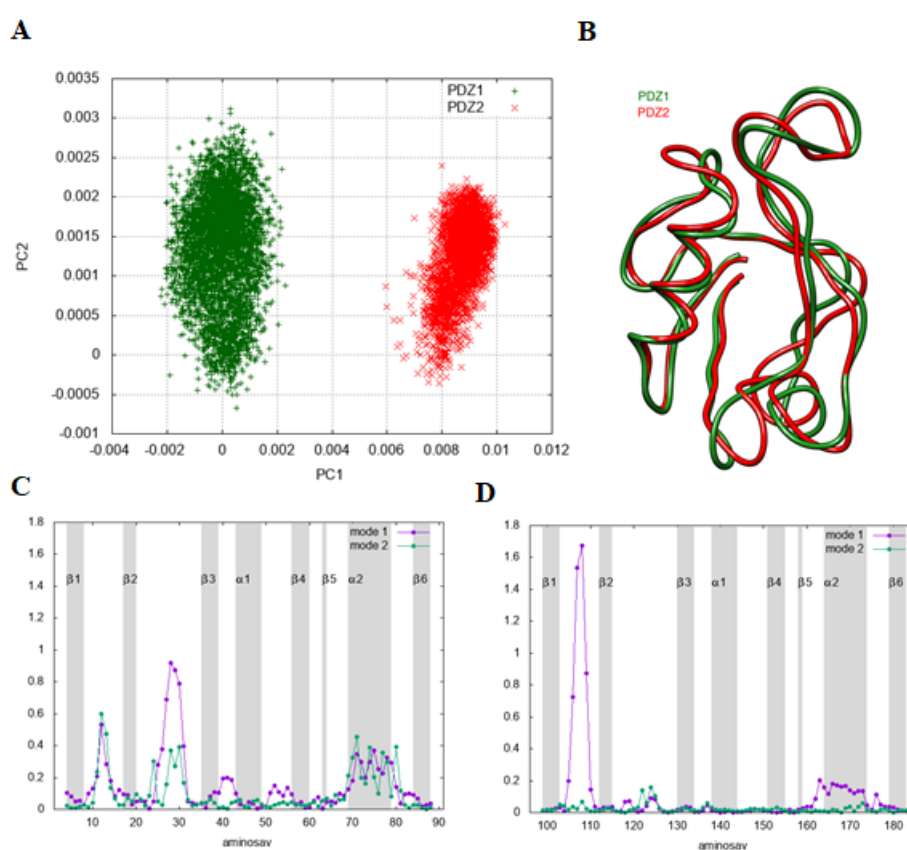
A visszszámolt rendparamétereket – a lokális illesztés elvének megfelelően – külön-külön határoztam meg a PDZ1 és a PDZ2 doménekre, majd az így kapott értékeket egy adathalmazként kezeltem. Az így visszszámolt S<sup>2</sup> értékek korrelációja ugrásszerűen megnövekszik a megkötött komplex sokaságban, ami érthető, hiszen csak ebben a sokaságban alkalmaztam azokat a kényszerfeltételeket. Ugyanakkor figyelemre méltó, hogy az S<sup>2</sup>-kényszerfeltételek alkalmazása nem rontotta el a kémiai eltolódásoknak való megfelelést.

### 5.2.2 Ligandumfüggő intradomén mozgások

A leghangsúlyosabb intradomén mozgások meghatározását főkomponens-analízissel végeztem el. Mindkét domén N- és C-terminális farokrégiója jelentős fluktuációt mutat, amely azonban nem releváns az intradomén dinamika szempontjából, ezért a domének 3-3 terminális aminosavját ebből az

elemzésből elhagytam. Így a rövidített doménhatárok a következők szerint alakultak: 4-88 a PDZ1, és 99-183 a PDZ2 doménre (doménenként 85 aminosav).

Bár a két PDZ domén nagyfokú szerkezeti és szekvenciális hasonlóságot mutat, a belső dinamikájuk jelentősen eltér egymástól. Ezt a legjobban a teljes sokaságból származó PDZ1 és PDZ2 doméneken elvégzett főkomponens-elemzés mutatja be: a két domén adatpontjai jól láthatóan elválnak egymástól az első főkomponens mentén (13.A ábra). A teljes sokaságra átlagolt, Ca atomokat tartalmazó PDZ1 és PDZ2 szerkezetek egymásra illesztésekor egyértelműen látszik, hogy maga az átlagos szerkezet, amely körül a gyors mozgások fluktuálnak, kis mértékben eltér a hurokrégiókban (13.B ábra). Emellett az is igaz, hogy maguk a belső mozgások is eltérnek, mint irányukra, mind amplitúdójukra nézve (13.C,D ábra).



13. ábra: A PDZ1 és PDZ2 domén belső mozgásainak összevetése. **A:** A teljes sokaságban szereplő PDZ1 és PDZ2 doménről készült főkomponens-analízis. **B:** Az egymásra illesztett átlagos PDZ1 és PDZ2 szerkezetek (csak a Ca atomok). **C-D:** Az első két főkomponensből számolt atomi elmozdulások amplitúdója az aminosav sorszám függvényében ábrázolva a PDZ1, illetve PDZ2 doménre.

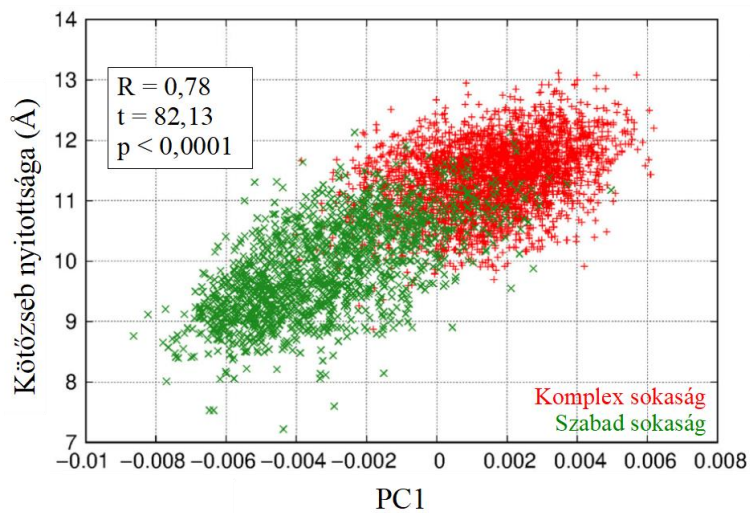
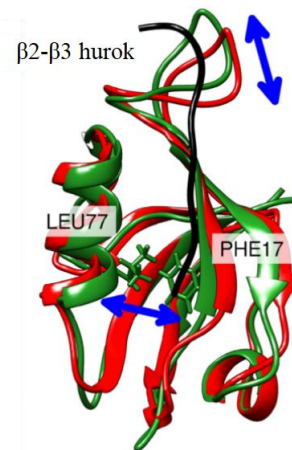
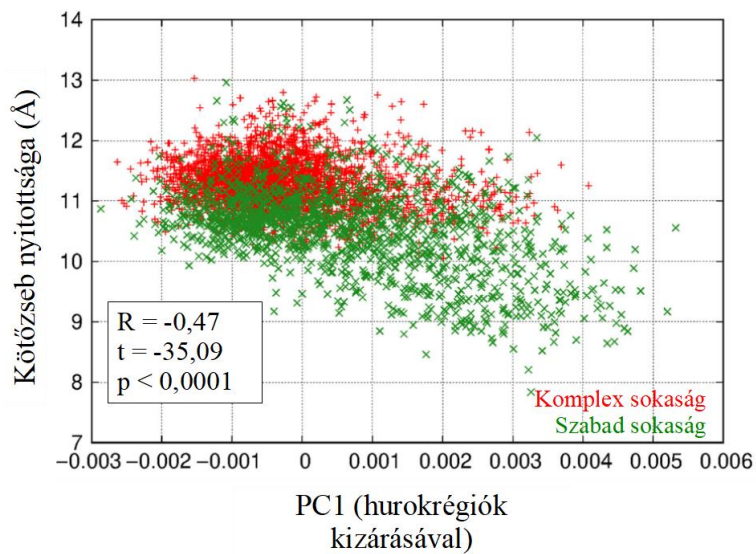
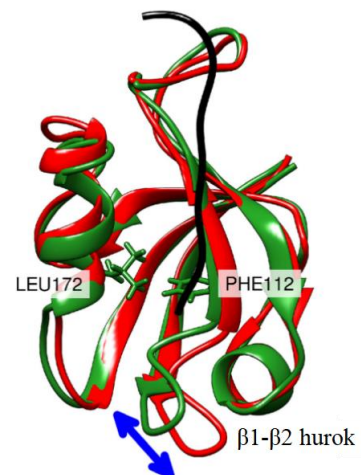
Annak ellenére, hogy az  $S^2$ -kényszerfeltételek alkalmazása ugrásszerű korrelációt eredményezett, a megkötött sokaságban az intradomén változások nem mutattak látványos változást a nem megkötött sokaságokhoz képest (F6. ábra). A PCA-ábrán látható, hogy a megkötött és nem megkötött komplex sokaságok nagyjából azonos régiót fednek le mind az első, mind a második főkomponens mentén, valamint a leghangsúlyosabb főkomponensek közötti átfedés is viszonylag magas.



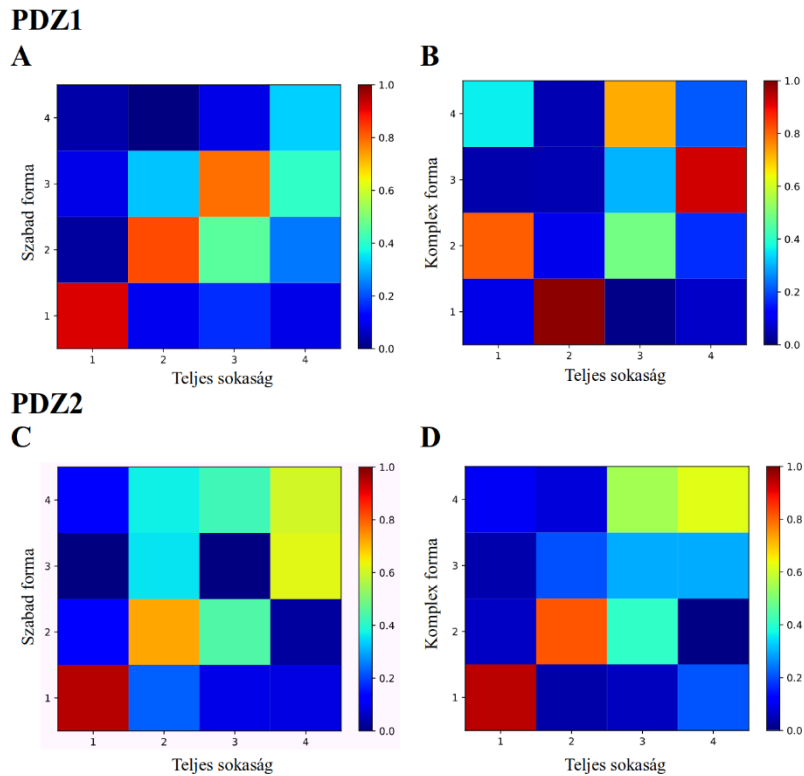
A PDZ1 domén első és második főkomponense egyaránt a flexibilis  $\beta$ 1- $\beta$ 2 és  $\beta$ 2- $\beta$ 3 hurokrégió, valamint az  $\alpha$ 2 hélix elmozdulását reprezentálja, azonban nem azonos irányban. Az első főkomponensben az  $\alpha$ 2 hélix eltávolodik a  $\beta$ 1 száltól. Ennek jelentősége abban áll, hogy a  $\beta$ 1 szál tövében található a PDZ domének karakterisztikus GLGF-motívuma, valamint a karboxilát-kötőzseb. Az első főkomponens tehát a kötőzseb nyílását és becsukódását reprezentálja. Ennek számszerűsítésére alkalmas a PDZ1-ben a Leu77 és a Phe17, PDZ2-ben pedig a Leu172 és Phe112 aminosavak Ca atomjai közötti távolság, amelyek a kötőzsebet egy-egy oldalról körülvevő,  $\alpha$ 2 hélixben és  $\beta$ 1 szálaban találhatóak, és a C-terminális valin hidrofób oldalláncának kötésében játszanak szerepet.

Az első főkomponens a PDZ1-ben szignifikáns korrelációt mutat a kötőzseb nyitottságával ( $R = 0,78, p < 0,0001$ ). Ugyanakkor szembeűnő, hogy a szabad és a komplex szerkezetek más régiót fednek le (14. ábra): a komplex állapotban a kötőzseb inkább nyitottabb, a szabad állapotban zártabb. Ez a megfigyelés egyfelől összhangban van azzal az intuitív képpel, amely szerint a ligandum csak a nyitott kötőzsebbel tud kölcsönhatást kialakítani, másrészt pedig fontos bizonyítéka annak a feltételezések, hogy ligandumkötéskor megváltozik PDZ domén konformációs eloszlása. Megjegyzendő, hogy a kinyíló-becsukódó mozgás a komplex sokaságban is megjelenik, de kevésbé hangsúlyos: míg a szabad sokaságban az első, a komplex sokaságban csak a második főkomponens mutat magas átfedést a teljes sokaság első főkomponensével (15. ábra). Emellett a komplex sokaság második főkomponense nem is mutat olyan magas korrelációt a kötőzseb nyitottságával ( $R = 0,39$ ). A magas átfedés sokkal inkább tulajdonítható a laza  $\beta$ 1- $\beta$ 2 és  $\beta$ 2- $\beta$ 3 hurokrégiók elmozdulásának.

Az imént bemutatott jelenség a PDZ2 doménben is megfigyelhető, de kevésbé hangsúlyos. A teljes sokaság első főkomponense csak közepes korrelációt mutat a PDZ2 kötőzsebének a nyitottságával ( $R = 0,41$ ), ugyanis az első két főkomponensben a  $\beta$ 1- $\beta$ 2 és  $\beta$ 2- $\beta$ 3 hurokrégiók (105-109 és 117-127 aminosavak) mozgása dominál. A főkomponens-análízis megismétlése e régiók kizárása után továbbra sem túl magas, de valamivel jobb korrelációt eredményezett ( $R = -0,47$ ), ám a szignifikancia szint ( $p < 0,0001$ ) meggyőzően a PDZ2 domén esetén is meggyőzően igazolja a első főkomponens és a kötőzseb nyitottsága közötti összefűggést. A PDZ1-hez hasonló a sokaságok eloszlása: a komplex szerkezetek a nyitottabb állapot körül sűrűsödnek (14. ábra). A PDZ2 doménre számolt átfedések az egyes sokaságok főkomponensi között szintén összhangban van ezzel a konklúzióval (15. ábra).

**A PDZ1****B****C PDZ2****D**

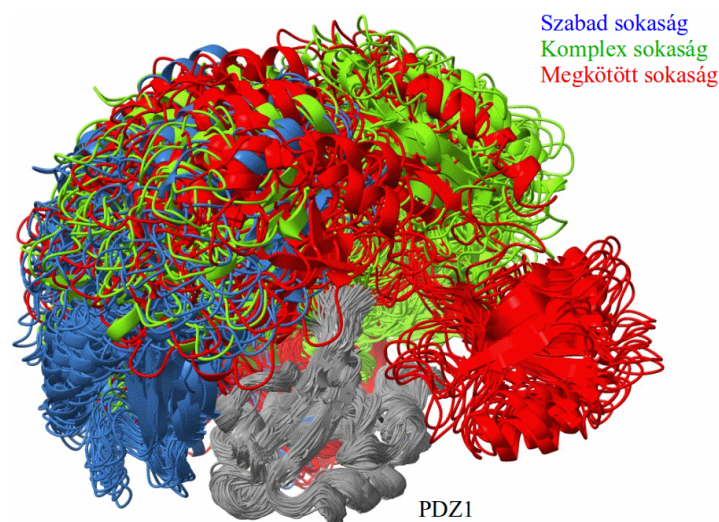
14. ábra: Intradomén mozgások vizsgálata a PDZ1-2 tandemben. A főkomponens elemzést a szűkített PDZ1 doménre (4-88 aminosav), illetve a szűkített PDZ2 doménre (99-183 aminosav) végeztem el, ez utóbbiból kihagyva a flexibilis hurokrégiókat (105-109 és 117-127 aminosav). A kötőzseb nyitottságát a Leu77 és Phe17, valamint a Leu172 és Phe112 aminosavak Ca atomjai közötti távolsággal mértem. **A,C:** A szabad és a komplex sokaság eloszlása az első főkomponens és a kötőzseb nyitottság mentén, feltüntetve a korrelációt. **B,D:** Az első főkomponens által reprezentált mozgás két extrém konformációja, a legjobban elmozduló régiók jelölésével. A ligandumot fekete szalag jelöli.



15. ábra: A szabad (A,C) és a komplex (B,D) sokaság főkomponenseinek átfedése a teljes sokaság főkomponenseivel. A főkomponens-analízist az intradomén mozgásokra, a szűkített doménhatárokat (4-88, illetve 99-183 aminosav) felhasználva határoztam meg.

### 5.2.3 Ligandumfüggő interdomén mozgások

Az interdomén mozgások vizsgálatához mindegyik sokaság PDZ1 doménjét egy közös templátra illeszttem, amelyek, tekintettel a domén stabil, globuláris szerkezetére, nagyon jól illeszkednek egymáshoz ( $C\alpha$  RMSD = 0,81 Å). Ugyanakkor a PDZ2 domének ezáltal a térben nagyon szétszórtnan helyezkednek el ( $C\alpha$  RMSD = 22,93 Å) (16. ábra).

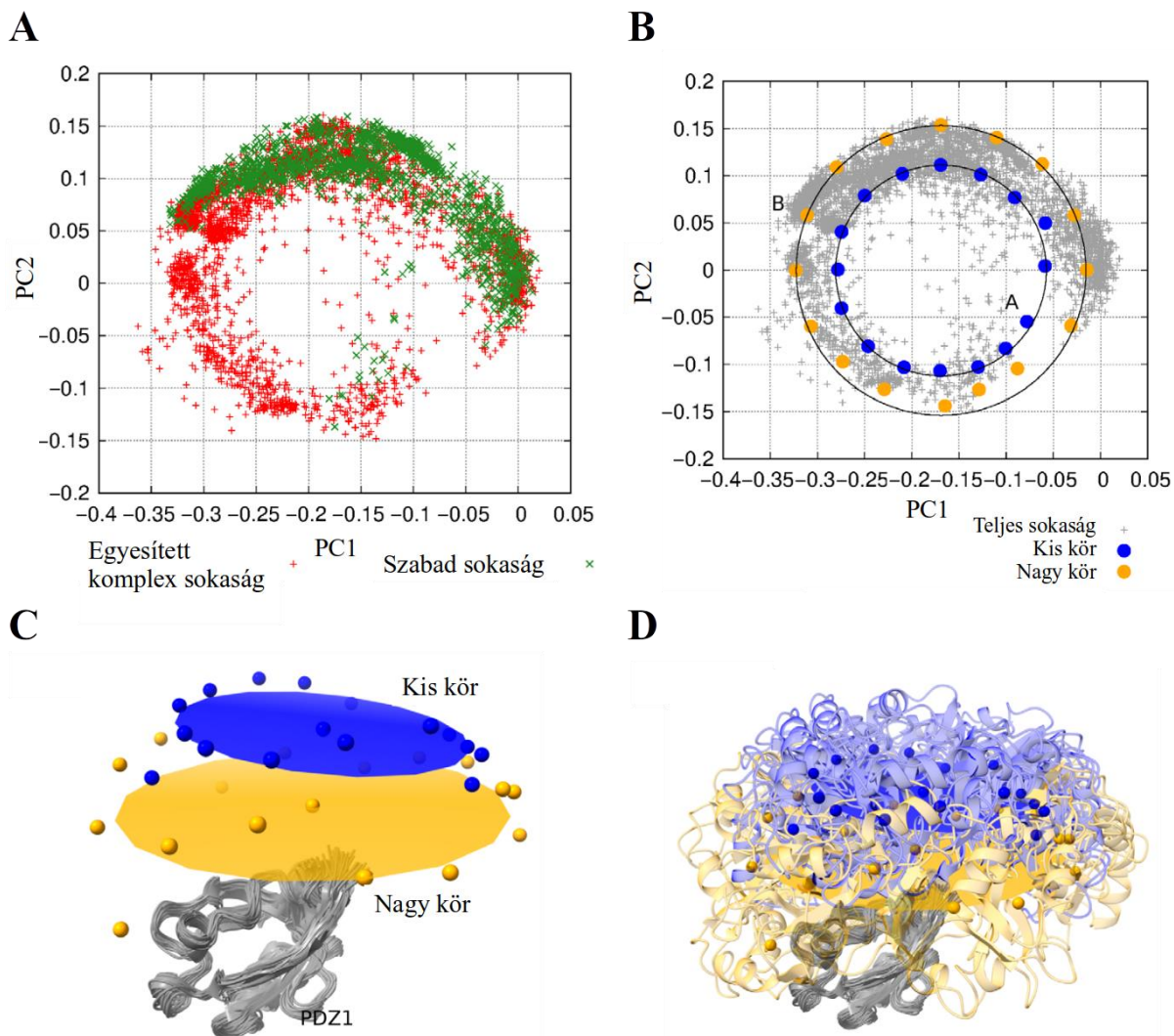


16 ábra: Mindhárom generált PDZ1-2 tandem sokaság, egy közös PDZ1 templátra illesztve. Az rajzolószoftver számítási kapacitásának végeessége miatt az ábrán csak minden tizedik replika szerepel.

Mivel az interdomén mozgások amplitúdója jóval nagyobb, mint az intradomén mozgásoké, ez az elrendezés kifejezetten alkalmas a két domén reorientációjának vizsgálatára. Ehhez egy olyan főkomponens-analízist végeztem el a PDZ2 domén rövidített szakaszán (99-183 aminosav), amely során a teljes sokaság közös PDZ1 templátra való illesztése után nem mozdítottam el a szerkezeteket. (Alapesetben a főkomponens-elemzés első lépése a szerkezetek egymáshoz történő illesztése – most azonban ezt a lépést elhagytam.) Az ilyen módon meghatározott főkomponensek tehát a PDZ2 domén elmozdulását reprezentálják a PDZ1 doménhez képest. A továbbiakban a teljes sokaságnak az így elvégzett főkomponens-elemzésből kapott első két főkomponens mentén történő ábrázolását *interdomén PCA-ábrának* fogom hívni (17.A ábra).

Szembetűnő, hogy az interdomén PCA-ábrán a pontok nagyjából egy kör mentén helyezkednek el (17.A ábra). Megállapítottam, hogy a körív mentén történő elmozdulás a háromdimenziós térben a PDZ2 domén körkörös elmozdulásának felel meg a PDZ1 domén körül, míg a sugárirányban történő elmozdulás a két domén egymáshoz való közeledését-távolodását reprezentálja. Ennek számszerűsítésére a következő eljárást végeztem el. Meghatároztam a PCA-ábra pontjaira legjobban illeszkedő kör sugarát, és az ábrára két kört rajzoltam ugyanazzal a középponttal, ám az eredeti sugár 0,8-, illetve 1,1-szeresével. Ezt a két kört 16 ekvidisztáns pontban mintavételeztem, és megkerestem a mintavételezési pontokhoz legközelebb eső pontot a PCA-ábrán (17.B ábra). Végül kiszámoltam az ezeknek a pontoknak megfelelő PDZ2 domének Ca atomjainak tömegközéppontját, és ezekre a pontokra is illesztettem egy-egy kört a háromdimenziós térben (17.C,D ábra). Így gyakorlatilag bebizonyítottam, hogy az interdomén PCA-ábrán mutatkozó kör megfeleltethető egy körnek a háromdimenziós térben.

A PDZ2 domén körkörös mozgása lehetővé teszi az interdomén torziós szög értékének számszerű definícióját. Ehhez a kétdimenziós PCA-ábrára legjobban illeszkedő körből indultam ki. Ezt szintén 16 ponton mintavételeztem, és a mintavételi pontokhoz legközelebb eső adatpontoknak megfelelő szerkezetből ismét kiszámoltam a PDZ2 domén tömegközéppontját – az imént bemutatott módszerrel –, majd pedig ezekre a pontokra is meghatároztam a legjobban illeszkedő kört a háromdimenziós térben. E kör középpontját és normálvektorát kiszámoltam a PDZ1 domén atomjaihoz képest: egészen pontosan a Phe17, Ile57 és Val84 Ca atomjai által meghatározott koordináta-rendszerben. Azért ezt a három atomot választottam, mert a másodlagos szerkezetnek robosztus építőelemei (a  $\beta 2$ ,  $\beta 3$  és  $\beta 6$  szál részei), és alacsony a lokális RMSD-jük. Ezáltal lehetővé válik, hogy bármely tetszőleges PDZ1 szerkezethez hozzárendeljünk egy háromdimenziós kört, amely mentén a PDZ2 elmozdulását várjuk. Az interdomén torziós szög végül a PDZ2 tömegközéppontjának és a PDZ1 domén Phe17 aminosav Ca atomjának a körre való projekciójából, vagyis e két pontnak a kör középpontjával bezárt szögeként adódik (18.A ábra).



17. ábra: **A:** A PDZ1 domének közös templátra való illesztését követő, PDZ2 domének rövidített szakaszán (99-183 aminosav) elvégzett főkomponens-elemzés ábrázolása az első két főkomponens mentén. **B:** A pontokra legjobban illeszkedő kör sugarának 0,8, illetve 1,1-szeres nagyítása 16 ekvidisztáns pontban mintavételezve, a mintavételezési pontokhoz legközelebb eső PCA-pontok kiemelésével. **C, D:** A kiemelt PCA-pontoknak megfelelő PDZ szerkezetek Ca atomjainak tömegközéppontja térben ábrázolva, és az azokra illesztett háromdimenziós kör.

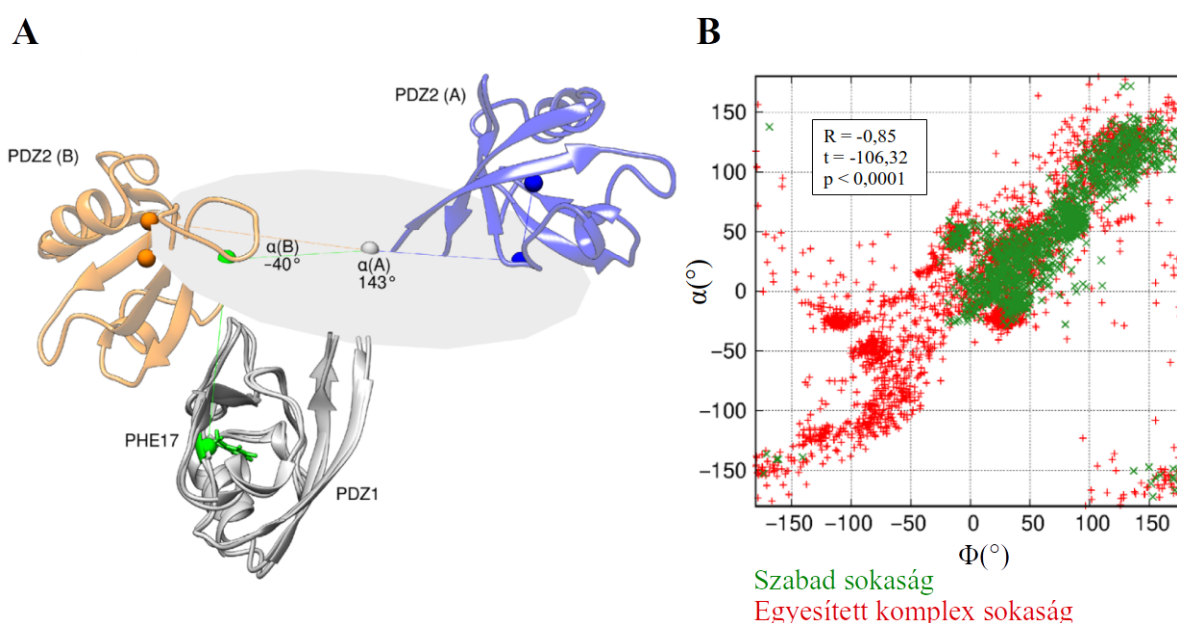
Az imént bemutatott módszerrel meghatározott interdomén torziós szög nagyszerűen korrelál a kétdimenziós PCA-ábra pontjaiból meghatározott polárkoordináták szögkomponensével ( $R = 0,98$ ), ami szemléletesen bizonyítja kétdimenziós ábrán és a háromdimenziós térben megfigyelhető körök egymásnak való megfelelését.

Az interdomén torziós szögre az eredeti tanulmány szerzői, Wang és mtsai. is adtak egy definíciót (W. Wang et al., 2009). Az ő megközelítésükben ezt a szöget a két doménhez kötődő ligandum 3 C-terminális aminosavjának tömegközéppontja és a két PDZ domén tömegközéppontja határozza meg: egészen pontosan a  $tkp_{lig1}-tkp_{PDZ1}-tkp_{PDZ2}-tkp_{lig2}$  pontok által definiált torziós szöggként adódik, ahol  $tkp_{lig}$  az ligandum atomjainak,  $tkp_{PDZ}$  pedig a PDZ domén atomjainak a tömegközéppontja. Megállapítottam, hogy az általam definiált torziós szög ( $\alpha$ ), és a Wang és mtsai. általi definícióból reprodukált torziós



szög ( $\phi$ ) egy fáziskorrekciót követően szintén nagyon jól korrelálnak egymással ( $R = 0,85$ ,  $p < 0,0001$ ) (18.B ábra).

Szembevetendő, hogy a ligandum jelenléte milyen nagy mértékben befolyásolja az interdomén mozgásokat: a PCA-ábra tanúsága szerint a komplex sokaság által lefedett konformációs tér jóval nagyobb, mint a szabad sokaság által lefedett (17.A ábra). Ugyanezt a következtetést lehet levonni az interdomén torziós szögekből: a komplex szerkezetek interdomén torziós szöge jóval nagyobb tartományban változik, mint a szabad szerkezeteké (18.B ábra). Ennek a felismerésnek igen nagy jelentősége van. Azt bizonyítja ugyanis, hogy sikerült olyan szerkezeti sokaságokat létrehoznom, amelyek egyfelől megfelelnek a kísérletes paramétereknek (5. táblázat), másfelől viszont tükrözik Wang és mtsainak fő konklúzióját, azaz hogy a ligandumkötés jelentős interdomén mozgásokat indukál.



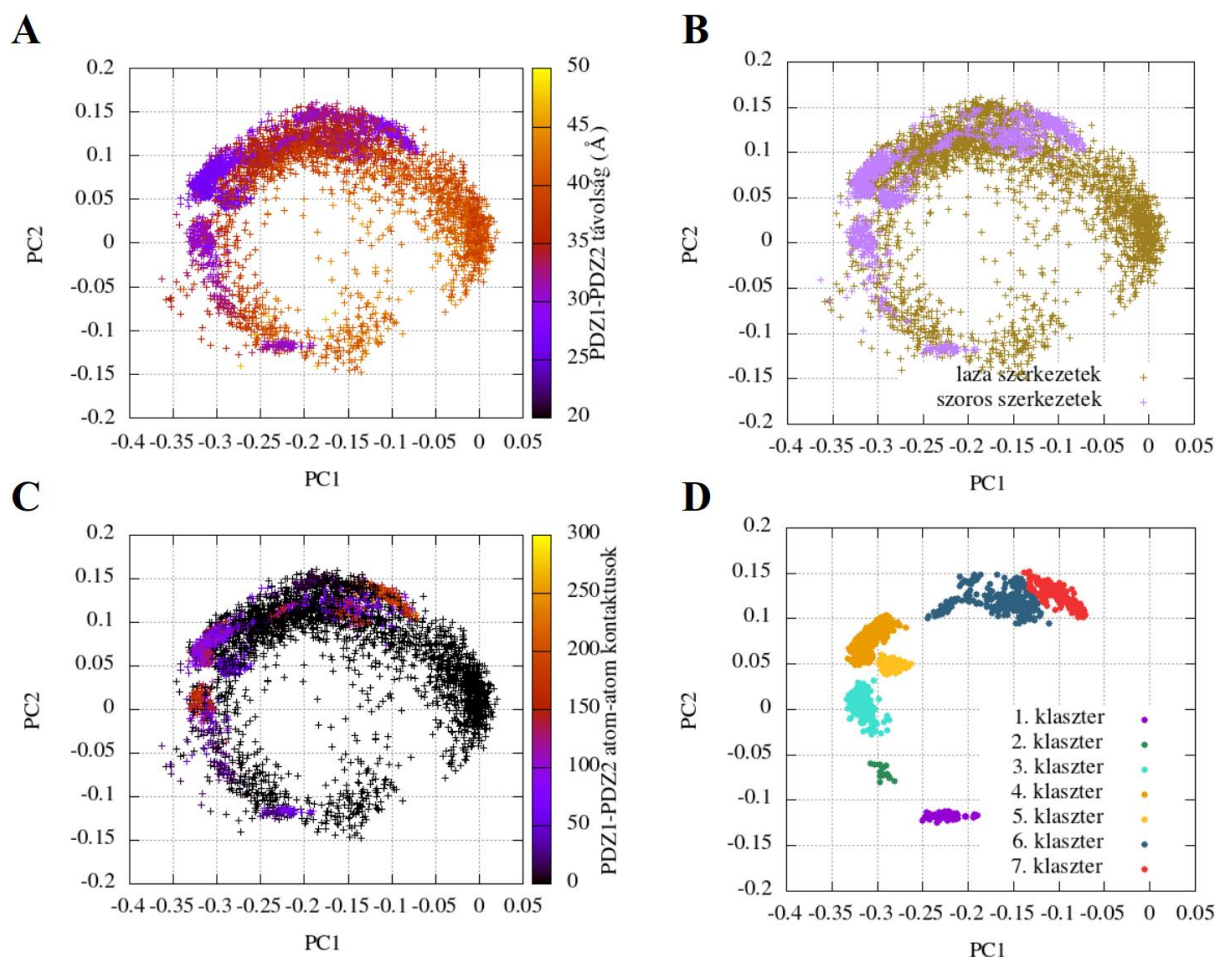
18. ábra: interdomén szögek ábrázolás a PDZ1-2 tandemben. **A:** az interdomén szög meghatározása szemléletesen ábrázolva (ld. YYY fejezet). **B:** az itt bemutatott ( $\alpha$ ), valamint a Wang és mtsai. által publikált ( $\Phi$ ) módszerrel számolt interdomén szögek korrelációja az általunk előállított sokaságra meghatározva.

Ahogy azt az intradomén mozgások esetén is megállapítottuk, az  $S^2$ -megkötéseknek nincs szembevetendő hatása az interdomén dinamikára. Az PCA-ábrán nincs olyan régió, amelyet kizárólag a megkötött sokaság fedne le (F7. ábra). Ellenkezőleg: a nem megkötött komplex sokaság mutat némi eltérést a szabad és a megkötött komplex sokaság által bejárt interdomén konformációs tértől. Ez felvethetné a kétséget a nem megkötött komplex sokaság hitelessége felől, ám ne felejtjük el, hogy a szabad szerkezetre nem álltak rendelkezésre kísérletes rendparaméterek, amelyeket kényszerfeltételként alkalmazhattam volna, így félrevezető kizárólag valamely sokaság által lefedett vagy nem lefedett régió nagyságáról messzemenő következtetéseket levonni.

## 5.2.4 Szoros illeszkedésű interdomén orientációk klaszterezése

Ahogy azt az előző fejezetben bemutattam, a PDZ1 domének közös templátra való illesztését követően a főkomponens-elemzést a PDZ2 doménon elvégezve képet kaphatunk a lehetséges interdomén orientációkról és a két domén reorientációja során bejárt konformációs térről. A teljes sokaság egyes replikáit reprezentáló pontok ábrázolása az első két főkomponens által definiált kétdimenziós térben egy nagyjából kör alakú eloszlást eredményez (17. ábra) (ld. 5.2.3. fejezet).

Szembetűnő, hogy a teljes sokaság pontjai nem egyenletesen oszlanak el a PCA-ábrázoláson látható kör mentén: egy bizonyos interdomén torziósszög-tartomány sűrűbben lefedett. Megvizsgálva a két PDZ domén tömegközéppontja közötti távolságot, az is feltűnik, hogy éppen az a régió tartalmaz több szerkezetet, amely a rövidebb domén-domén távolságot lehetővé tevő interdomén orientációt tartalmazza (19.A ábra). Ezzel összefüggésbe hozható az a tény is, hogy a PDZ1 domén nem teljesen a közepén helyezkedik el az interdomén torziós szög definíciójához meghatározott körnek, hanem enyhén elhajlik az egyik irányba (17.C ábra). Ez az elhajlás kedvez az interdomén kölcsönhatás kialakulásának is az adott tartományban.



19. ábra: PDZ1-2 tandem szerkezetek kategorizálása az interdomén kontaktusok és távolság szerint. Az ábrák az interdomén főkomponensek szerint vannak ábrázolva (ld. szöveg), és a **D** panel kivételével a teljes sokaságot tartalmazzák. **A**: a teljes sokaság, az interdomén távolság szerint színezve. **B**: Szoros és laza szerkezetek eloszlása az interdomén PCA-ábrán. **C**: A teljes sokaság, az atom-atom kontaktusok száma szerint színezve. **D**: A szoros szerkezetekből előállított klaszterek.

Az interdomén atom-atom kontaktusok számát azoknak a nehézatompároknak a számával definiáltam, amelyek legfeljebb 5 Å távolságra vannak egymástól (19.C ábra). Világosan látszik, hogy a sok interdomén atom kontaktot lehetővé tevő interdomén orientációk éppen ugyanazok, amelyek egyúttal a közeli domén-domén távolságot is megengedték.

Hogy képet kapjak a két PDZ domén közötti kölcsönhatás és az interdomén dinamika közötti összefüggésről, különválasztottam azokat a szerkezeteket, amelyekben az interdomén kontaktusok száma legalább 20 volt. Ezeket szoros szerkezeteknek, a többit laza szerkezetnek neveztem (19.B ábra). A szoros szerkezetek aránya teljes sokaságban 37%-nak, a szabad és a komplex sokaságban pedig 32 és 40%-nak adódott.

Az szoros szerkezetek viszonylag jól elkülönülő szigeteket alkotnak a kétdimenziós PCA-ábrán. Felhasználva a PCA-ábrára illeszkedő körből meghatározott  $r$  és  $\phi$  polárkoordinátákat, a szoros szerkezeteket alávetttem a legtávolabbi szomszéd elvét követő, "teljes láncmódszer" elnevezésű (*complete linkage*) hierarchikus klaszterezési eljárásnak, melynek eredményeképpen ezek a szerkezetek 5 klaszterre különültek el. Szemrevételezés után ezek közül 2 klasztert további két-két részre bontottam: így végül összesen 7 klaszter adódott (19.D ábra). Néhány kiugró adatpontot manuálisan eltávolítottam. A klaszterek méretét, a doménekre lebontott globális RMSD-értékeit és a kísérletes kémiai eltolódás értékeknek való megfelelését az F8. táblázat tartalmazza.

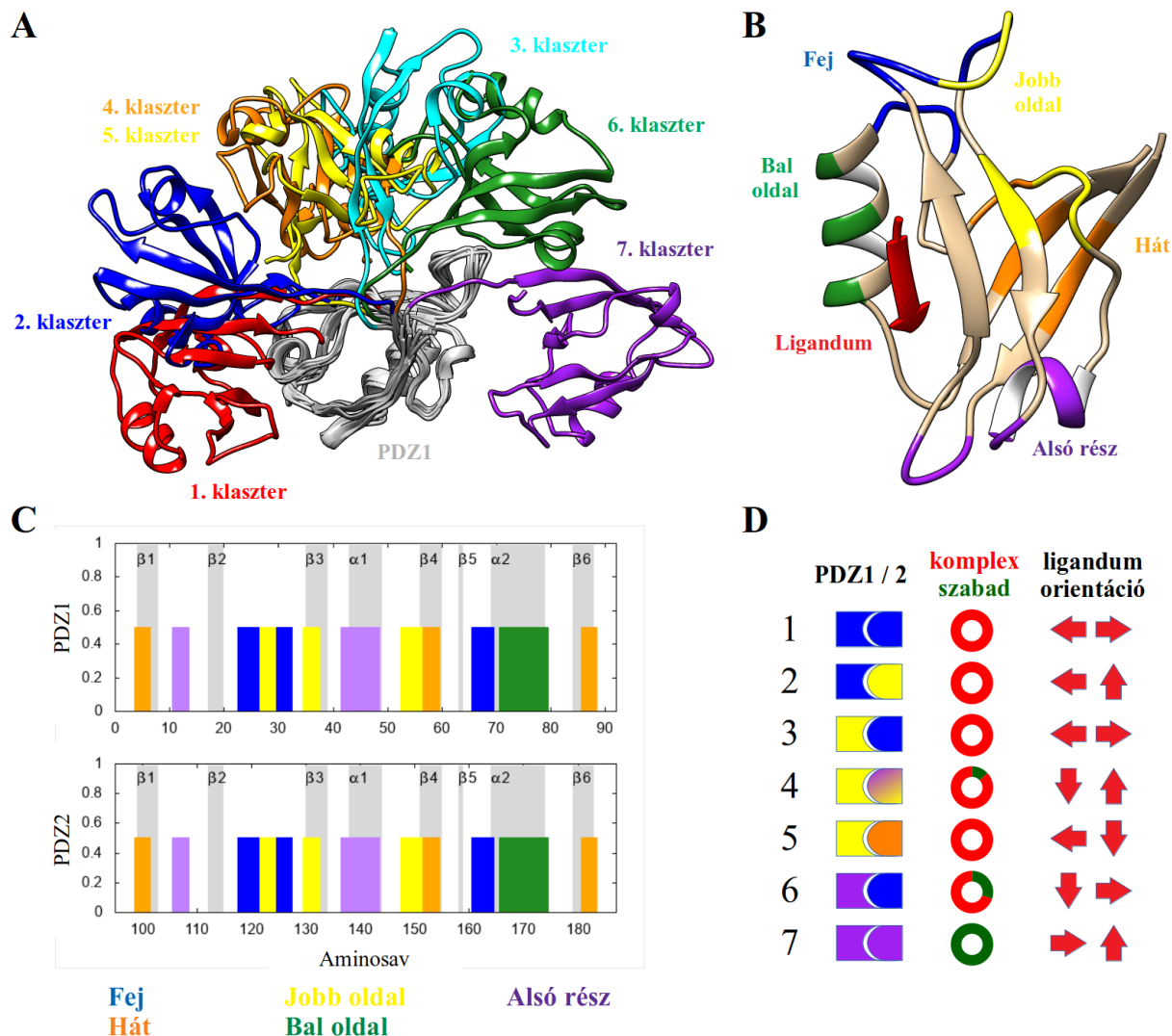
Az így létrehozott klaszterek egy-egy jól definiált szoros illeszkedésű interdomén-orientációt reprezentálnak. Minden klaszterből kiválasztottam a reprezentatív szerkezetet, vagyis amelynek a legkisebb globális RMSD-je volt az átlagos gerinckonformációhoz képest (20.A ábra). Hogy egyszerű legyen meghatározni az egyes klaszterek által favorizált interdomén-interfészt, a PDZ domén felületét 5 részre osztottam, melyeket a ligandumkötő zseb helyzetéhez viszonyítva jobb és bal oldalnak, fejnek, hátnak és alsó résznek neveztem el (20.B,C ábra). A 20.D ábrán látható, hogy a 7 klaszter mindegyike egy-egy egyedi kombináción keresztül alakítja ki az interdomén-interfészt, ugyanakkor az interdomén-orientációra érvényes korlátozások miatt nem alakulhat ki bármilyen interfész-páros.

Mivel a PSD-95 egy membrán-asszociált fehérje, a PDZ1-2 tandem szupramoduláris szerkezetének jelentősége abban áll, hogy a két domén orientációja nagyban meghatározza a lehetséges kölcsönható partnerek elhelyezkedését a sejten belül. A PDZ1-2 tandem kísérletesen meghatározott szerkezetei e tekintetben egymással legalább részben ellentmondó eredményeket hoztak. Ezért megvizsgáltam a klaszterekben a két PDZ doménhez kötődő ligandum relatív orientációját, melyhez a 4 C-terminális aminosav Ca pozícióját használtam fel.

Szembetűnő, hogy a különböző interdomén-orientációk egyúttal különböző ligandum-orientációt tesznek lehetővé (20.D ábra, F9. táblázat). Az egy irányba mutató ligandumok lehetővé tennék a membránreceptorok szinergikus kötését, ahogy azt Long és mtsai. feltételezték az egyik első PDZ1-2 tandem szerkezet alapján (J. F. Long et al., 2003). Azonban a két domén orientációja az általam meghatározott klaszterek egyikében sem párhuzamos egymással, és egyúttal merőleges az interdomén tengelyhez képest, ami lehetővé tenné ezt a fajta kölcsönhatást. A legtöbb klaszterben tapasztalt



ligandum-orientáció inkább egy olyan elrendezést preferál, amelyben az egyik domén membránreceptorhoz, a másik pedig egy intracelluláris partnerhez kötődik. Ne felejtjük azonban el, hogy a klaszterek mind szoros szerkezetek tartalmazznak, amelyek a teljes interdomén konformációs térnek csak egy részét fedik le. Ha figyelembe vesszük a teljes sokaság által lehetővé tett összes interdomén-orientációt, sok olyat találunk közöttük, amelyben a két domén egymással nincs kontaktusban, és a két ligandum párhuzamos egymással (F10. ábra).



20. ábra: PDZ1-2 tandem klaszterek jellemzése az interdomén interfész szempontjából. **A:** Az egyes klaszterek reprezentatív szerkezetei, közös PDZ1 templátra illesztve. **B:** A PDZ domén felületének felosztása öt különböző régióra. **C:** A PDZ domén felületének öt különböző régiója az aminosavsáv függvényében ábrázolva. Külön jelölve vannak a PDZ domének másodlagos szerkezeti elemei. **D:** Az interdomén interfész kialakulása az egyes klaszterekben, az imént definiált régiók szerint színezve. Az ábrán szerepel a szabad és komplex szerkezetek aránya az egyes klaszterekben, illetve a két ligandum egymáshoz viszonyított orientációja.

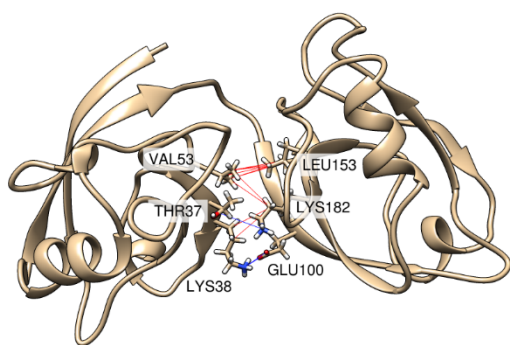
### 5.2.5 Atomi szintű kölcsönhatások a szoros illeszkedésű klaszterekben

Az egyes klaszterek által kialakított interdomén-interfész szemléletes ábrázolásán túl (20. ábra) további elemzéseket is elvégeztem a két PDZ domén, illetve az egyes PDZ domének és ligandumjaik között létrejövő atom-atom kontaktusokról. Külön vizsgáltam a hidrogénkötéseket és a hidrofób

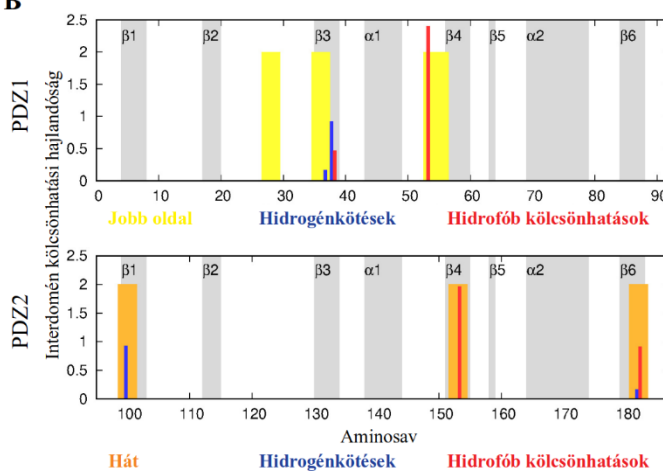
kölcsönhatásokat. Előbbieket a szög-távolság kritérium alapján határoztam meg (Baker & Hubbard, 1984; Xu et al., 1997), utóbbiakat pedig az egymáshoz közeli, legfeljebb 5 Å távolságra lévő interdomén atom-atom kontaktként azonosítottam a következő aminosavak oldalláncának bármely szénatomja között: Ala, Ile, Leu, Met, Phe, Pro, Thr, Trp, Val, Arg, Lys, vagy a Gly Ca atomja. Meghatároztam azt is, hogy a PDZ domének egyes aminosavjai mennyire hajlamosak részt venni valamilyen interdomén kölcsönhatásban: ezt az adott aminosav valamely atomja által kialakított interdomén hidrogénkötés vagy interdomén hidrofób kölcsönhatások számának a sokaság vagy klaszter méretéhez viszonyított arányként definiáltam, és *interdomén kölcsönhatási hajlandóságnak* neveztem. Megjegyzendő, hogy az interdomén kölcsönhatási hajlandóság 1-nél nagyobb értéket is felvehet egy klaszterben, ha az adott aminosav egynél több atomja is gyakran alakít ki interdomén hidrogénkötést vagy hidrofób kölcsönhatást.

Az atomi szintű interdomén kölcsönhatásokat az 5. klaszter példáján a 21. ábra mutatja be, de az F11. ábra tartalmazza az összes klaszterról hasonló koncepcióval készült ábrázolást. Ahogy az elvárható, az interdomén kölcsönhatást kialakító aminosavak belesznek az előző fejezetben definiált tartományokba, amelyek között létrejön az interdomén interfész. Az F12. táblázat azon interdomén hidrogénkötések és hidrofób kölcsönhatások teljes listáját tartalmazza, amelyek legalább egy klaszterben a szerkezetek legalább 10%-ában megjelennek. Szembetűnő, hogy a lehetséges kölcsönhatásoknak mindegyik klaszter egy-egy teljesen egyedi kombinációját alakítja ki, és viszonylag csekély azoknak a kölcsönhatásoknak a száma, amelyek egynél több klaszterben is számottevő arányban megjelennek.

**A**



**B**



21. ábra: Atomi szintű interdomén kölcsönhatások az 5. klaszter példáján bemutatva. **A:** Az interdomén hidrogénkötések (kék) és hidrofób kölcsönhatások (piros), kiemelve az azokat kialakító aminosavakat. **B:** Az interdomén kölcsönhatási hajlandóság (ld. szöveg) az aminosavsám függvényében ábrázolva. Külön jelöltük azokat a PDZ régiókat, amelyek az 5. klaszterben az interfész kialakításában részt vesznek (ld. 20. ábra).

Szintén meghatároztam a két PDZ domén és ligandumjai között kialakuló hidrogénkötéseket, és kiszámoltam ezek előfordulási arányát az egyes klaszterekben (F13. táblázat). A korábban publikált eredményekkel összhangban három olyan hidrogénkötést találtam, amelyek kiugróan magas arányban vannak jelen a szerkezetekben: ezek a hidrofób kötőzseb Phe és Ile aminosavjai, valamint a ligandum

0 és -2 aminosavjai között jönnek létre, és főszerepet játszanak a ligandum stabilizálásában (3. ábra). Ugyanakkor feltűnő, hogy számos más, kisebb arányban megjelenő hidrogénkötés is kialakul, ám ezek előfordulása szintén klaszter-specifikus.

Ahogy azt a PDZ domén szakirodalmának áttekintésekor bemutattam, több tanulmány is arra a következtetésre jutott, hogy a PDZ domének ligandumspecificitását nem lehet kizárólag a 0 és -2 aminosavak változatosságával megmagyarázni: a PDZ doménnek a szűken vett kötőzseben kívüli régiói, valamint a ligandum felső aminosavjai közötti kölcsönhatások meghatározó szerepet játszhatnak a ligandumspecificitás és a belső dinamika szabályozásában (Fuentes et al., 2004; Mostarda et al., 2012; Tochio et al., 2000). Az imént bemutatott eredmények alátámasztják ezeket a feltételezéseket és egyúttal arra engednek következtetni, hogy a ligandum-domén és domén-domén kölcsönhatások, valamint az interdomén orientáció egymással összefüggésben vannak és minden bizonnyal kölcsönös szabályozó szerepet töltenek be.

### 5.2.6 A klaszterekben található interdomén orientációk összevetése kísérletesen meghatározott szerkezetekkel

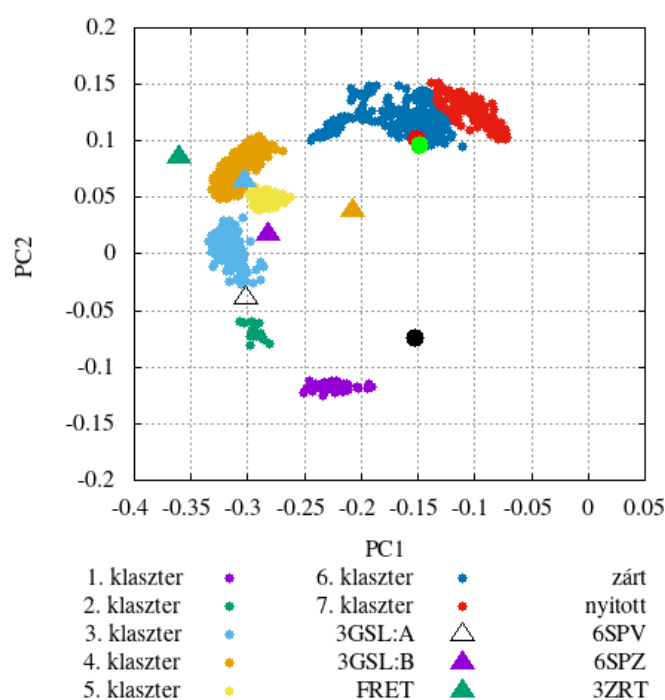
A PDZ1-2 tandem szupramolduláris szerkezetének meghatározására tett erőfeszítések több, egymásnak részben ellentmondó kísérletes szerkezetet eredményeztek, melyeket az irodalmi bevezetésben részletesen áttekintettünk (4. ábra). A fő különbség ezek között szerkezetek között a két PDZ domén egymáshoz viszonyított orientációja, és ebből következően a két kötőpartner lehetséges elhelyezkedése a posztszinaptikus sejtben a PDZ tandemhez képest.

Az elérhető kísérletes szerkezeteket összevettem az általam meghatározott klaszterekkel és a 5.2.3. fejezetben bemutatott interdomén főkomponens-elemzést is kiterjesztettem ezekre a szerkezetekre. Ennek eredményét mutatja be a (22. ábra).

Az összevetésbe bevett kísérletes szerkezetek a következők voltak. Négy kristályszerkezet, melyek PDB kódjai 3ZRT (Bach et al., 2012), 3GSL (Sainlos et al., 2010) és 6SPV a szabad, valamint 6SPZ a komplexált állapotra (Ahmad Rodzli et al., 2019), (a 3GSL kóddal jelzett kristályszerkezetben két molekula található az aszimmetrikus egységben, melyeket a PDB-ben lévő fehérjeláncokra alkalmazott nevezéktan szerint 3GSL:A és 3GSL:B-vel jelöltem), továbbá a Förster-féle rezonancia-energiaátadás módszerével meghatározott szerkezet (FRET) (J. J. J. McCann et al., 2011), valamint ez utóbbi szerkezet diszkrét molekuladinamikai szimulációval (DMD) történő modellezéséből származó két további, nyitott és zárt szerkezet (Yanez Orozco et al., 2018).

A kristályszerkezetek közül a 3GSL-hez tartozó két fehérjelánc a 3-as klaszter két, egymással szemközti szélén helyezkedik el. Ennek megfelelően mindkét szerkezetben a 3-as klaszterre jellemző interdomén interfész alakul ki a PDZ1 jobb oldala és a PDZ2 feje között: a két PDZ domén szorosan illeszkedik egymáshoz. A 6SPV és a 6SPZ azonos körülmények között készült kristályokból keletkeztek, ám az előbbi a PDZ1-2 tandem szabad, az utóbbi pedig komplexált formájáról.

Szupramoduláris szerkezetüket tekintve igen hasonlóak egymáshoz: mindkét forma a 6-os klaszter részét képezi, amit megerősít a PDZ1 alsó része és a PDZ2 feje között kialakuló interfész. Itt is megfigyelhető a két domén szoros illeszkedése. Egészen más a helyzet azonban a 3ZRT esetében: ez a szerkezet nem tartozik egyik klaszterhez sem, ugyanis a két domén nem szorosan illeszkedik, hanem a tandem sokkal inkább egy kinyújtózott interdomén konformációt vesz fel. Megjegyzendő, hogy ennek a szerkezetnek a minősége a legrosszabb a kristályszerkezetek közül mind a felbontás (3,4 Å), mind pedig az elektronsűrűségi térképen nem látható atomok számát illetően. Azonban az aszimmetrikus egységben lévő 4 fehérjelánc szupramoduláris szerkezete szinte teljesen pontosan megegyezik, és amúgy egy kristályszerkezettől elvárható, hogy a két PDZ domén relatív orientációját még egy viszonylag rosszabb felbontás esetén is meg lehessen állapítani.



22. ábra: Az elérhető kísérletes PDZ1-2 tandem szerkezetek és az általunk előállított klaszterek közös interdomén PCA-ábrán ábrázolva.

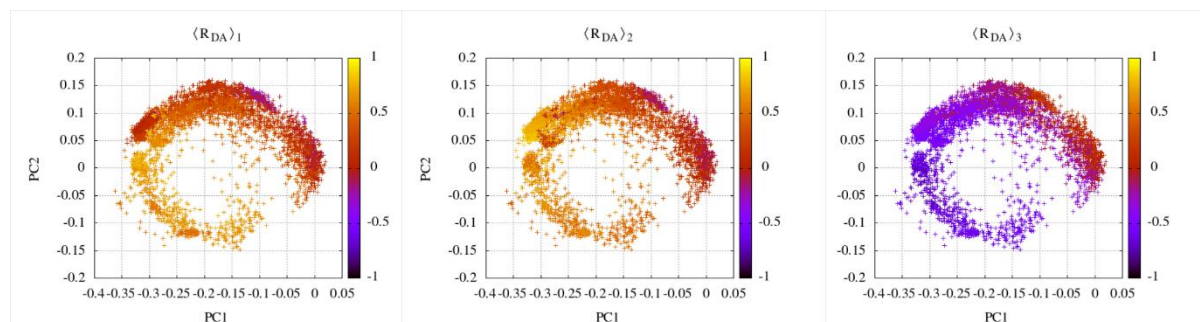
Figyelemre méltó, hogy a három különböző módszerrel meghatározott kristályszerkezet egymástól igen eltérő szupramoduláris szerkezetet mutat. A 6SPV és a 6SPZ szerkezetek esetén még a ligandum jelenléte sem eredményezett jelentős változást az interdomén orientációban: az azonos kristályosítási körülmények, illetve a szabad formából készült kristályok használata oltókristályként a komplex forma kristályosítása során szinte teljesen azonos interdomén konformációt eredményezett. Mindez arra enged következtetni, hogy nem lehet egyetlen, stabil szerkezettel jellemezni a PDZ1-2 tandem interdomén orientációját, hanem lehetséges konformációk sokaságáról beszélhetünk, amelyek eloszlása nagyban függ a körülményektől.

Az interdomén PCA-ábra tanúsága szerint a FRET-szerkezet kívül esik az itt lefedett konformációs régióban: ez alapján úgy tűnik, hogy a mi sokaságaink nem tartalmazzák ezt a lehetséges

interdomén orientációt. Azonban ezt a szerkezetet némi távolságtartással kezelhetjük. Néhány évvel később maguk a szerzők kezdtek el újra foglalkozni vele (Yanez Orozco et al., 2018), és a FRET kísérlet során gyűjtött adatokat diszkrét molekuladinamikai szimulációk segítségével újra feldolgozták, és nem egyetlen szerkezetet, hanem lehetséges konformációk sokaságát határozták meg. Ezek közül két, viszonylag sűrűn betöltött lokális minimum bontakozott ki, melyeket nyitott és zárt konformációnak hívtak, és amelyeket már tartalmazza a mi sokaságunk által bejárt konformációs tér. Ahogy a neve is mutatja, a nyitott konformerben nincs számottevő interdomén kontaktus, ezért az egyik klaszterhez sem tartozik. A zárt szerkezet viszont mind a PCA-ábra alapján, mind pedig a két PDZ domén jobb oldala között létrejövő interfész alapján a 4. klaszter egyértelmű tagja.

A sokaságainkat összevettem továbbá a FRET-elemzés során kapott kísérletes adatokkal is. A FRET kísérlet során a két PDZ domén egyes aminosavjaihoz fluoreszcens csoportokat kötöttek, és az ezek közötti energiatranszfert mérték, amely az elmélet értelmében összefüggésben van a kölcsönható csoportok közötti távolsággal. A nyers mérési eredményekből egy matematikai modell segítségével végül összesen három, egyenként 10-10 átlagos csoportközi távolságot tartalmazó adatkészletet határoztak meg. Ezek közül a két legnagyobb súllyal megjelenő adatkészlet megfelel a nyitott és a zárt konformációnak, a harmadikról pedig megállapították, hogy egy viszonylag nagy, jól lefedett konformációs térnek megfelelő átlagos szerkezetet reprezentál (Yanez Orozco et al., 2018).

Mindegyik szerkezetre kiszámoltam azt a 10 interdomén távolságot, amelyek a FRET-elemzésben szereplő csoportközi távolságnak felelnek meg, és meghatároztam azok korrelációját a kísérletből származó adatokkal. Ezek a távolságok nem lehetnek teljesen azonosak, én ugyanis az érintett aminosavak Ca atomjai közötti távolságot mértem, nem pedig az aminosavakhoz kötött fluoreszcens csoportok közötti távolságot, ám a tendenciák még az így kapott korrelációkban is jól látszanak. Megállapítható, hogy azok a régiók, amelyek a leginkább korrelálnak az első, második és harmadik átlagos csoportközi távolságokat tartalmazó adatkészlettel, viszonylag jól elkülönülnek az interdomén-PCA ábrán (23. ábra).



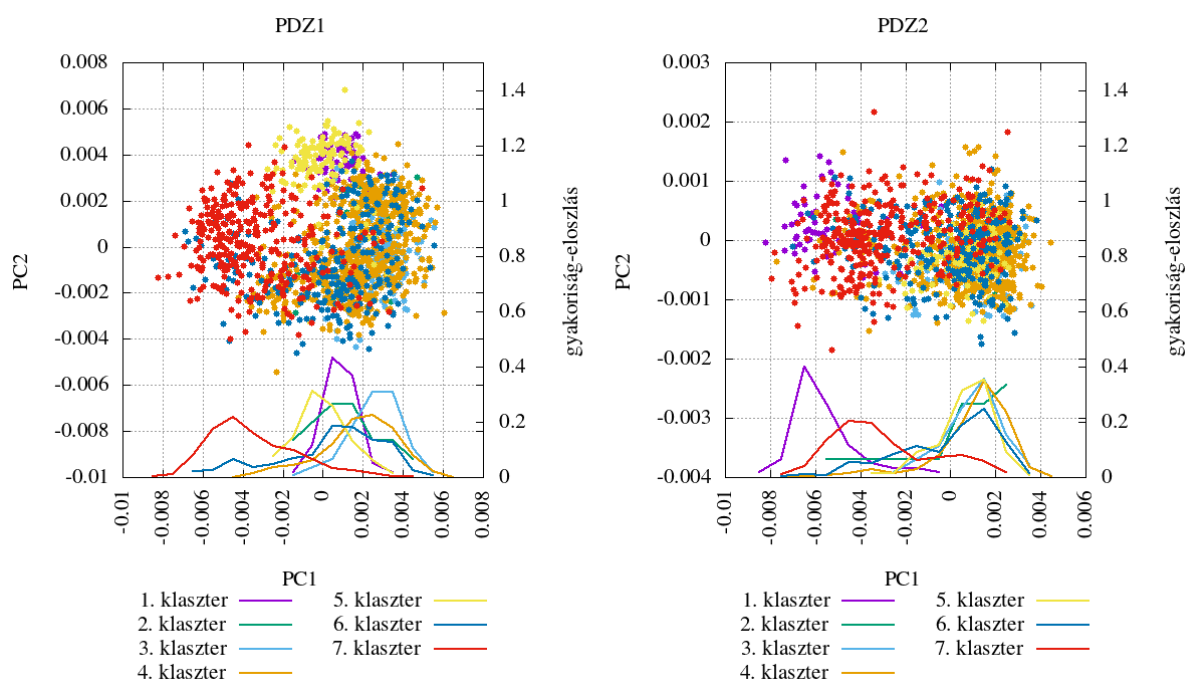
23. ábra: Az általunk előállított teljes sokaság és a kísérletes FRET adatok közötti korreláció. Az általunk előállított sokaság egyes szerkezeteiből visszszámoltam azt a 10 interdomén atom-atom távolságot, amelyek a FRET-mérések során adódtak (Yanez Orozco 2018), és meghatároztam az adatpontok korrelációját a finomítás során előállt három adatkészlettel ( $\langle R_{DA} \rangle_1$ ,  $\langle R_{DA} \rangle_2$  és  $\langle R_{DA} \rangle_3$ ).

Még szemléletesebb mutatja ezt be az egyes klaszterekre, valamint a kísérletes szerkezetekre számolt korreláció a FRET adatokkal (F14. táblázat). A FRET-elemzés szerinti legnagyobb és második

legnagyobb arányban populált állapot, azaz a nyitott és a zárt konformáció, egyértelműen megfeleltethető az általunk létrehozott 3. és 4. klaszternek. Vegyük észre, hogy a mi elemzésünk alapján a nyitott állapot nem esik ugyan bele egyik klaszter által lefedett konformációs térbe sem, ám az interdomén orientációja alapján határozottan a 3. klaszterre hasonlít, amennyiben a PDZ1 domén jobb oldala és a PDZ2 domén feje vannak egymáshoz a legközelebb. Különös figyelmet érdemel a FRET-elemzés szerinti legkevésbé populált állapot, amely a szerzők állítása szerint nem feleltethető meg egyetlen stabil szerkezetnek, hanem lehetséges interdomén konformációk eloszlásaként lehet értelmezni. Ez az állapot a 7. klaszterrel mutat valamekkora – hanem is kiugró, de a többinél magasabb – korrelációt, ám az interdomén PCA-ábránk alapján, a szerzők konklúziójával összhangban, a konformációs tér viszonylag nagy részét lefedi.

### 5.2.7 Az intra- és interdomén mozgások összefüggésének vizsgálata

Az egyes PDZ doménekre elvégzett főkomponens-elemzéssel meghatároztam a leghangsúlyosabb intradomén mozgásokat (ld. 5.2.2. fejezet), az interdomén dinamika elemzése során pedig 7 klaszterre különítettem el a lehetséges szoros illeszkedésű szerkezeteket (ld. 5.2.4. fejezet). Ha csak klasztereknek megfelelő adatpontokat ábrázoljuk az első két intradomén főkomponens szerint, az tűnik fel, hogy az egyes klaszterekhez tartozó szerkezetek eloszlása nem egyenletes az első főkomponens mentén (24. ábra). A PDZ1 domén esetében az 1., a PDZ2 domén esetében pedig az 1. és a 7. klaszter szeparálódik leginkább a többi sokaságtól.



24. ábra: Az intra- és interdomén mozgások közötti összefüggés vizsgálata a PDZ1-2 tandemben. Az ábrán az általunk előállított klaszterek adatpontjai vannak ábrázolva a PDZ1- és PDZ2-intradomén mozgásaira meghatározott főkomponensek szerint, valamint az egyes klaszterek eloszlása az első főkomponens mentén.



Ahogy a 5.2.2. fejezetben bemutattam, a PDZ1 doménben az első főkomponens mindkét ligandum esetén a hidrofób kötőzseb kinyíló-becsukódó mozgását, valamint a  $\beta$ 1- $\beta$ 2 és  $\beta$ 2- $\beta$ 3 hurokrégiók elhajlását reprezentálja. Ezzel szemben a PDZ2 doménben a kötőzseb kinyíló-becsukódó mozgását elnyomja a  $\beta$ 1- $\beta$ 2 hurokrégió mozgása. Az 5.2.4. fejezetben definiált kategorizálás alapján a  $\beta$ 1- $\beta$ 2 hurok a PDZ domén aljához, a  $\beta$ 2- $\beta$ 3 pedig részben a PDZ domén fejéhez és jobb oldalához tartozik. A 20.D ábra alapján ezeknek a régióknak valamelyike egy kivétellel mindegyik klaszterben szerepet játszik az interdomén interfész kialakításában. Ez a megállapítás azt bizonyítja, hogy az intradomén mozgások és az interdomén orientáció egymással összefüggenek.

Az interdomén orientáció alapján előállított klaszterek egyaránt tartalmazhatnak szabad és komplex szerkezeteket is. A PDZ1 domén esetén az első főkomponenstől elkülönülő klaszter csak szabad szerkezeteket tartalmaz, a PDZ2 domén esetén szeparálódó két klaszter azonban – az 1. és a 7. – egyike kizárólag szabad, a másik pedig kizárólag komplex szerkezeteket tartalmaz.

A PDZ domének és a ligandumok között meghatározott hidrogénkötés-mintázatok arról tanúskodnak, hogy számos lehetséges hidrogénkötés jöhet létre a ligandum felső aminosavjai és a PDZ domén  $\beta$ 2- $\beta$ 3 hurokrégiója között (F13. táblázat). Ezeknek a hidrogénkötéseknek azonban az egyes klaszterenként más-más eloszlását figyelhetjük meg.

Mindebből arra következtethetünk, hogy a PDZ domének ligandumkötése, intradomén dinamikája és szupramoduláris reorientációja között egy rendkívül összetett, számos apró atomi szintű kölcsönhatás által modulált szabályozó mechanizmus áll fenn. A ligandum megváltoztatja a sokaság első intradomén főkomponens mentén történő eloszlását, amely közvetlen hatással van a  $\beta$ 1- $\beta$ 2 és a  $\beta$ 2- $\beta$ 3 hurokrégiók dinamikájára – az utóbbira a közvetlen kölcsönhatások által is. E hurokrégiók pedig, amint láttuk, főszerepet töltenek be a két domén közötti interfész kialakításában, ezáltal szabályozva a lehetséges intradomén orientációt. Ez a szabályozó szerep az ellenkező irányban is megvalósulhat: bizonyos interdomén orientációk a szabad, mások pedig a komplex állapotot favorizálják.

### 5.3 Diszkusszió – Első téziscsoport

A PSD-95 fehérje igen sokrétű biológiai funkciót lát el: szerepet játszik a posztzinaptikus denzitásban található membránreceptorok klaszterezésében, a jelátviteli folyamatokban adapterfehérjeként viselkedik, valamint szabályozza a hosszútávú potenciáció és depresszió kiváltásához szükséges küszöbpotenciált (Kim & Sheng, 2004; H.-J. Lee & Zheng, 2010; Sheng & Sala, 2001). Ennek megfelelően számos szinaptikus fehérjéről bizonyosodott be, hogy a PSD-95 kötőpartnere, és ezek egy jelentős része az N-terminális PDZ1-2 tandemmel alakít ki kölcsönhatást.

A PSD-95 kötőpartnerek sokfélesége zavarba ejtő. Nem világos, hogy milyen módon diszkriminál a lehetséges ligandumok között és hogy hogyan szabályozódik a PDZ tandem szupramoduláris reorientációja, amely lehetővé teszi, hogy a posztzinaptikus denzitas különböző rétegeiben lokalizálódó partnerekkel alakítson ki kölcsönhatást.

Mindemellett egyre részletesebb képünk alakult ki a PDZ1-2 tandem ligandumspecificitásáról. Világossá vált, hogy az eddig kanonikusnak vélt, három osztályba sorolható kötési mód nem magyarázza kellő részletességgel sok PDZ domén finomhangolt affinitását a többféle lehetséges ligandum felé. A GLGF-kötőmotívumon és a C-terminális hidrofób aminosavat (legtöbbször valint) stabilizáló hidrofób zseben kívül egyre nagyobb jelentőséget tulajdonítanak a  $\beta$ 2- $\beta$ 3 huroknak, amely a ligandum felsőbb aminosavjaival alakít ki kölcsönhatást (Mostarda et al., 2012). Erről a régióról az is bebizonyosodott, hogy mind a gyors, mind a lassú időskálán lecsökken a belső dinamikája a ligandumkötés hatására (Fuentes et al., 2004).

Általánosságban igaz, hogy a PDZ tandem funkciója eltér az azt alkotó domének funkciójának egyszerű összegétől, de a PSD-95 PDZ1-2 tandemje esetén a két domén kölcsönhatásának mértéke, amely a köztük lévő szinergiára utalna, alig detektálható. NMR-es kísérletekkel viszont megállapították, hogy a komplexált PDZ1-2 tandem sokkal nagyobb interdomén dinamikát mutat, mint a szabad forma (W. Wang et al., 2009). Ezenkívül több tanulmány igyekezte meghatározni a PDZ1-2 tandem szupramoduláris szerkezetét, ám ezek egymásnak részben ellentmondó következtetésekre jutottak mind a két domén relatív orientációját, mind pedig a kötőpartnereknek a fehérjéhez képesti lehetséges elhelyezkedését tekintve. Mindezek alapján nyitott kérdés marad, hogy melyek a szupramoduláris szerkezetet stabilizáló, meghatározó interdomén kölcsönhatások, és ezek milyen összefüggésben vannak a kötőpartner jelenlétével.

A munkám során  $^1\text{H}$ - $^1\text{H}$  NOE távolság- és  $\text{S}^2$  rendparaméter-kényszerfeltételekkel megkötött molekuladinamikai szimulációkat futtattam, hogy ezáltal a PDZ1-2 tandem kísérletesen megfigyelt belső mozgásait tükröző szerkezeti sokaságot állítsak elő. Mind a távolság-, mind a rendparaméter-kényszerfeltételek csak a gyors, ps-ns időskálájú intradomén mozgásokra nézve jelentettek megkötést: a két PDZ domén egymáshoz képesti, lassabb időskálán történő elmozdulását semmi sem gátolta. Ezt az  $\text{S}^2$  rendparaméterek lokális illesztési algoritmusai lehetővé tették, amelyben az egyes doménekhez tartozó paramétereket mindegyik szimulációs lépésben külön-külön kényszerítjük rá a doménekre. A



kísérletesen megmért adatok között sem szerepelt egyetlen interdomén  $^1\text{H}$ - $^1\text{H}$  NOE: a távolságkényszerfeltételek a szimuláció során csupán a domének szerkezeti stabilitását szolgálták.

*1.1. tézis: Előállítottam a PSD-95 PDZ1-2 tandemjének kísérletileg mért dinamikai viselkedését tükröző szerkezeti sokaságot.*

Az  $\text{S}^2$ -megkötések alkalmazásával a kísérletes paramétereknek való megfelelés ugrásszerű növekedését sikerült elérni anélkül, hogy eközben az amid N és H kémiai eltolódásoknak való megfelelés romlott volna az eredeti sokasághoz képest (5. táblázat). A sokaságokra meghatározott intradomén globális RMSD értékek megfelelnek egy globuláris, megfelelően feltekeredett fehérjétől elvárhatóknak.

A felhasznált kényszerfeltételek csak a domének belső mozgására nézve jelentettek megkötést. A két domén egymáshoz képesti relatív elmozdulását az alkalmazott lokális illesztési algoritmus nem korlátozza, ezáltal lehetőség nyílt a szupramoduláris dinamika vizsgálatára is.

*1.2. tézis: Főkomponens elemzés segítségével megállapítottam, hogy a ligandumkötés hatására megváltozik a PDZ domén kötőzsebének kinyíló-becsukódó mozgása.*

Az intradomén mozgások vizsgálata során bebizonyosodott, hogy a ligandum kötése befolyásolja a sokaság eloszlását az első főkomponens mentén, amely mindkét doménben a hidrofób kötőzsebkinyílását-becsukódását és a  $\beta 1$ - $\beta 2$  hurok mozgását, a PDZ1 doménben pedig ezenkívül a  $\beta 2$ - $\beta 3$  hurok mozgását reprezentálja (14. ábra). Ez a mozgás egyformán jelen van a szabad és a komplexált PDZ doménekben, de a komplex sokaság szignifikánsan kisebb régiót jár be e mozgás mentén, mint a szabad sokaság. Az effektus valamivel hangsúlyosabban van jelen a PDZ1, mint a PDZ2 doménben. Ez az eredmény összhangban van azzal a megállapítással, amely szerint ligandumkötéskor a  $\beta 2$ - $\beta 3$  hurok megmerevedése figyelhető meg mind a gyors, mind a lassú időskálán (Fuentes et al., 2004).

*1.3. tézis: Az általam előállított PDZ1-2 sokaságok lefedik a teljes szupramoduláris konformációs teret, amelyben a komplexált forma korlátozott dinamikára képes a szabad formához képest. A lehetséges szoros illeszkedésű szupramoduláris orientációkat 7 klaszterre osztottam, amelyek egyedi interdomén interfészt alakítanak ki.*

A közös PDZ1 templátra illesztett sokaság PDZ2 doménje térben nagyon széttartóvá válik. Az ilyen módon rögzített PDZ2 sokaságon elvégzett főkomponens-elemzés információt ad a két domén relatív mozgásáról. Az első két főkomponens átszámítható polárkoordinátákká, amelynek szögkomponense az interdomén torziós szögnek, sugárkomponense a két domén távolságának feleltethető meg (17. ábra).

A ligandum jelenléte jelentősen megnöveli a sokaság által bejárt interdomén konformációs teret is, ezzel gyakorlatilag lehetőséget adva egy nagyobb amplitúdójú szupramoduláris átrendeződésre. Ezzel lényegében sikerült egy atomi szintű modellt alkotnom az elemzés kiindulópontjául szolgáló kísérletes megfigyelésnek, azaz hogy a komplexált PDZ1-2 tandem jelentősen megnövekedett interdomén dinamikát mutat a szabad formához képest.

Azok a szupramoduláris konformerek, amelyek között jelentős interdomén kölcsönhatás alakul ki, 7 klaszterre oszthatók, amelyek egymástól jól elkülönülő régiót fednek le az első két interdomén főkomponens mentén (19.D ábra). Mivel ezekben más a két domén egymáshoz viszonyított orientációja, ezért más az interdomén kölcsönhatási interfész is. A PDZ1 doménnek 3, a PDZ2 doménnek 4 régiója vesz részt a kölcsönhatás kialakításában a különböző klaszterekben (20. ábra). A matematikailag lehetséges interfész-kombinációknak csak egy része valósulhat meg a – részben a linker rövidegéből adódó – sztérikus és geometriai megkötések miatt.

Annak ellenére, hogy a rendparaméterek kényszerfeltételként történő alkalmazása ugrásszerű növekedést eredményezett a kísérletes paramétereknek való megfelelésben, meglepő módon a megkötött komplex sokaság sem az interdomén mozgásokban, sem a lehetséges interdomén orientációkban nem mutatott jelentős újdonságot a nem megkötött sokaságokhoz képest. Ez a megállapítás arra készítet, hogy megvizsgáljuk a rendparaméter-megkötések hozzáadott értékét a szimuláció során.

Még ha az intra- és interdomén mozgások irányán és amplitúdóján nem is eredményeznek szembetűnő változást a rendparaméter-megkötések, a sokaság eloszlására azonban határozottan befolyást gyakorol, ami pedig, ahogy az elemzés során láttuk, a különböző időskálájú belső mozgások és a ligandumkötés közötti finomhangolt szabályozó mechanizmusnál fogva nagyon is jelentős hatással lehet a PDZ tandem viselkedésére. Vegyük észre azt is, hogy az interdomén PCA-ábra azon régiói, melyeket nem fed le sem a szabad, sem a megkötött komplex sokaság, két klasztert is tartalmaznak, és ezek közül az egyikre kísérletes szerkezetek is vannak (F7. ábra, 22. ábra). És végezetül azt se feledjük, hogy a szabad szerkezetre nincsenek elérhető rendparaméterek, így valójában nem tudhatjuk, hogy egy  $S^2$ -megkötött szabad sokaság vajon mekkora interdomén konformációs teret járna be.

*1.4. tézis: Az előállított sokaságot összevettem a nyilvános adatbázisban elérhető kísérletesen meghatározott PDZ1-2 szerkezetekkel. A lefedett konformációs tér magában foglalja a kísérletes szerkezeteket. Az szoros illeszkedésű konformerekből készült klaszterek valószínűsítik a még nem megfigyelt, de geometriai szempontból lehetséges szupramoduláris szerkezetek létezését.*

Bármely szimuláció hitelességének végső próbája a kísérletesen igazolt eredményekkel való összevetetőség. Esetünkben kifejezetten komoly jelentősége van a modelljeink és az elérhető kísérletes PDZ1-2 tandem szerkezetek közötti megfelelésnek. Elmondható, hogy olyan sokaságot hoztam létre, amely magába foglalja a korábbi, egymásnak látszólag ellentmondó szerkezetek szinte mindegyikét: a

FRET-szerkezet kivételével (amelyen egy későbbi tanulmányban további finomításokat végeztek) mindegyik kísérletes szerkezetet tartalmazza a sokaságaink által bejárt interdomén konformációs tér, valamint a szoros illeszkedésű kísérletes szerkezetek mindegyike hozzárendelhető az általunk meghatározott 7 klaszter valamelyikéhez (22. ábra).

Különös figyelmet érdemel az egyik kísérletes PDZ1-2 kristályszerkezet, amely az itt bemutatott elemzés publikálása után született meg (PDB: 6SPV a szabad, 6SPZ a komplex formára). Bár a szupramoduláris konformációja nem egyezett meg semelyik korábbi kísérletes szerkezettel, az előállított sokaságnak mégis megfelelt: mind az interdomén orientáció, mind az interdomén interfész alapján a 6. klaszter tagjai lett. Hasonló megfontolással valószínűsíthető, hogy a PDZ1-2 tandemnek más elérhető szupramoduláris szerkezetei is vannak, amelyeket a sokaságban tapasztalt kisebb súlyuk miatt nehéz kísérletesen is megfigyelni (pl. 1. és 7. klaszter), ám a dinamikai egyensúlyban betöltött szerepüknél fogva biológiailag is releváns formák lehetnek.

*1.5. tézis: A sokaság elemzéséből levont következtetések alapján javaslatot tettem a PDZ1-2 tandem kötőpartnereinek a szupramoduláris dinamikára gyakorolt szabályozó mechanizmusára. A kötőpartnereknek a  $\beta$ 2- $\beta$ 3 hurokrégióval kialakuló lokális kölcsönhatások által modulálják az interdomén orientációt, amely régió a lehetséges szupramoduláris konformerek többségében szerepet játszik az interdomén interfész kialakításában.*

A PDZ1-2 tandem viselkedésével kapcsolatos legégetőbb kérdés a két domén egymással kialakított kölcsönhatása, a tandem szupramoduláris szerkezete és a ligandumkötés közötti összefüggés mechanizmusa. A sokaságok elemzésekor azonosítottam a két PDZ domén azon régióit, amelyek a leggyakrabban vesznek részt az interdomén interfész kialakításában. Szembetűnő, hogy ezek jelentős részben egybeesnek a fehérje azon régióival, amelyek belső dinamikájára leginkább hatással van a ligandummal létrejövő kölcsönhatás: a  $\beta$ 1- $\beta$ 2 és  $\beta$ 2- $\beta$ 3 hurok a PDZ1, valamint a  $\beta$ 1- $\beta$ 2 hurok a PDZ2 doménben. Megvizsgálva a szoros illeszkedésű klasztereket, azt tapasztaljuk, hogy egy kivétellel mindegyik klaszterben e régiók valamelyike részt vesz az interdomén interfész kialakításában.

Feltűnő, hogy az interdomén és intradomén főkomponensek egymástól nem függetlenek. A szoros illeszkedésű klaszterek, amelyek implicit módon magukban hordozzák az interdomén főkomponensektől való függést, nem egyenletesen oszlanak el az első intradomén főkomponens mentén (24. ábra).

Az egyes klaszterekben más-más hidrogénkötés-mintázat jön létre a ligandum felső (-3 – -7) aminosavjai és a PDZ domén  $\beta$ 2- $\beta$ 3 hurokrégiója között (F13. táblázat). Mivel pedig a klaszterek különböző interdomén-orientációkat képviselnek, ezzel a megállapítással összefüggést teremtettem a szupramoduláris szerkezet és a ligandummal kialakított kölcsönhatás között.

Több tanulmány alátámasztotta, hogy a kanonikus karboxilát kötőzseben túl a  $\beta$ 2- $\beta$ 3 hurok komoly szerepet játszik a ligandumspecifititás finomhangolásában. Mivel a PDZ1-2 tandem biológiai

szerepével kapcsolatos egyik legnagyobb kérdés, hogy hogyan tud diszkriminálni a különböző, térben is máshol elhelyezkedő kötőpartnerei között, az itt bemutatott modell erre ismagyarázatul szolgálhat. A sokaságainkból levont következtetés alapján ugyanis megállapítható, hogy a ligandum jelenléte modulálja az interdomén dinamikát. Ebből azonban indirekt módon arra is következtethetünk, hogy a domének orientációja is – a  $\beta$ 2- $\beta$ 3 hurokrégió keresztül – modulálni tudja a ligandumaffinitást. Ez a kép összhangban lenne azzal a felismeréssel is, amely szerint a PDZ1-2 tandem a PSD-95 fehérjén belül egy önálló szerkezeti egységet alkot, és a PDZ2 és SH3 domén közötti, kb. 60 aminosav hosszú linkernek köszönhetően független mozgásra képes (J. J. J. McCann et al., 2011).

A szimulációkból természetesen csak óvatosan szabad következtetni a PSD-95 PDZ1-2 tandem *in vivo* viselkedésére. Modelljeink egyrészt csak egy rövid, C-terminális peptidszegmenst tartalmaztak ligandumként, míg a PSD-95 valóságos kötőpartnerei jóval nagyobb fehérjék, és a méretük komoly sztérikus akadályt jelenthet a PDZ tandem szupramoduláris flexibilitására nézve. Igaz, sok esetben kötőpartnerek C-terminális a PDZ kötő motívumai gyakran egy rendezetlen rész végén helyezkednek el, ami csökkenti a két PDZ domén relatív orientációjának kritikus voltát. Ez a megállapítás viszont csak akkor érvényes, ha a terminális rendezetlen régió nem túl rövid, illetve érvényét veszti azoknál a kötőpartnereknél, amelyekben a PDZ kötő motívum nem terminális, hanem belső szekvencia.

Összességében a modelljeink alapján azt valószínűsíthetjük, hogy a ligandumkötés, a domének közötti kölcsönhatás és a szupramoduláris dinamika között egy összetett, számos atomi kölcsönhatás által finomhangolt szabályozó szerep áll fenn. A ligandumszabályozott szupramoduláris flexibilitás mindezek alapján a PSD-95 fehérje biológiai funkciójának integráns része, és a PDZ1-2 tandem példája kiindulópontul szolgálhat más, hasonló rendszerek vizsgálatához.

## 6 A parvulin-típusú cisz-transz izomerázok belső mozgásainak vizsgálata

### 6.1 A felhasznált szerkezeti sokaságok

A rotamázok belső mozgásainak vizsgálatához szerzőtársam, Dr. Czajlik András futtatott külső kényszerfeltételekkel megkötött molekuladinamikai szimulációkat, és az így létrehozott sokaságokon én végeztem el a további elemzéseket (Czajlik et al., 2017). Így a munkafolyamat első fázisára – a kiindulási szerkezetek és kényszerfeltételek kiválasztására, valamint a szimulációk futtatására – csak röviden térek ki, hogy aztán részletesen bemutassam a szerkezeti sokaságok elemzését és az azokból levont következtetéseket.

Három olyan parvulin-típusú cisz-transz izomerázt választottunk ki, amelyekhez rendelkezésre álltak  $^1\text{H}$ - $^1\text{H}$  NOE-k és az amidkötésekre mért  $S^2$  általános rendparaméterek: az SaPrsA-t a *Staphylococcus aureus*ból, a TbPin1-et a *Trypanosoma brucei*ből és a CsPinA-t a *Cenarchaeum symbiosum*ból (25. ábra). E három fehérjéhez tartozó, általunk felhasznált kísérletes szerkezet PDB kódja rendre 2JZV (Heikkinen et al., 2009), 2LJ4 (Sun et al., 2012) és 2RQS (Jaremko et al., 2011). A SaPrsA és a CsPinA nem Pin1-típusú parvulinok, vagyis nem tartalmaznak WW-domént és a katalitikus mechanizmusuk foszforilációfüggetlen. A TbPin1 ezzel szemben azok közé a Pin1-típusú parvulinok közé tartozik, amelyek mechanizmusa foszforilációfüggő ugyan, de a Pin1-gyel ellentétben nem tartalmaznak WW-domént.

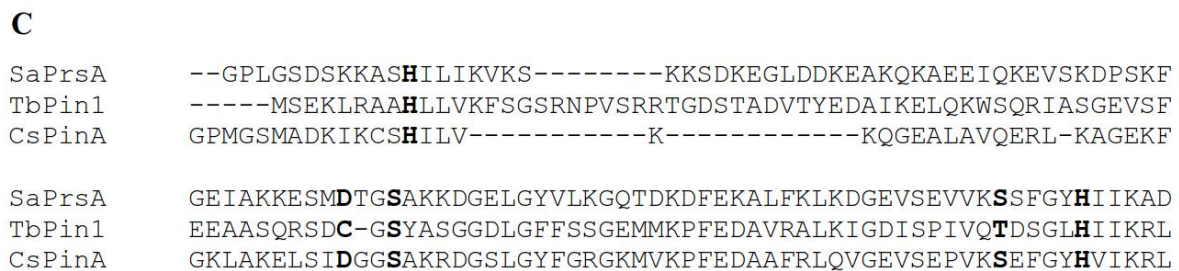
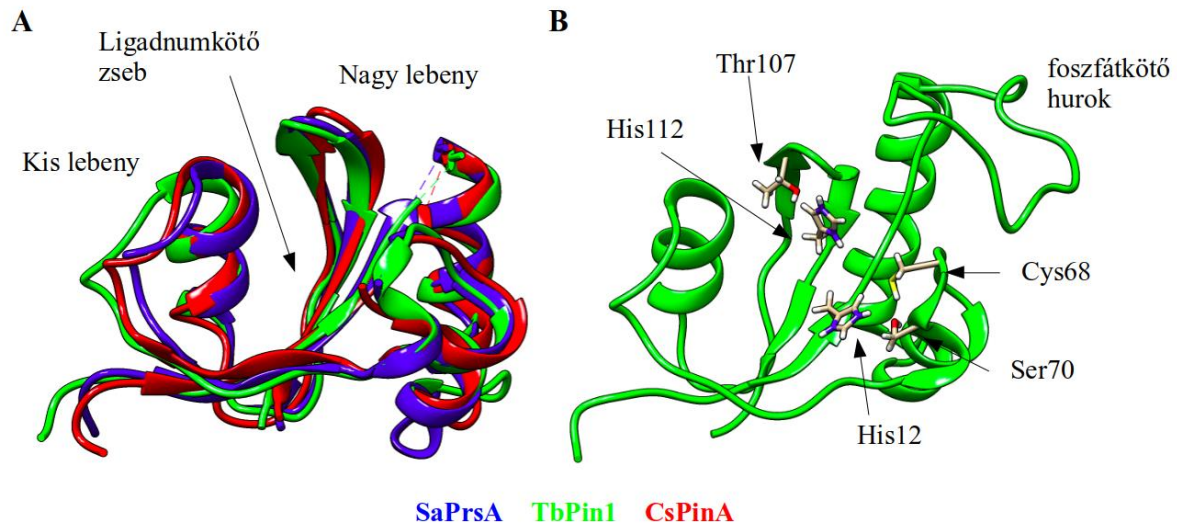
A vizsgálandó parvulinok kiválasztásánál két szempont játszott szerepet. Egyrészt olyan szerkezetet kerestünk, amelyekre rendelkezésre állnak jó minőségű, megbízható kísérletes dinamikai paraméterek – konkrétan  $^1\text{H}$ - $^1\text{H}$  NOE-k és gerinc  $S^2$  rendparaméterek – másrészt pedig elvárás volt, hogy a vizsgált parvulinok viszonylag különbözőek legyenek, hogy az összehasonlító elemzés felderítse az eltérő ligandumaffinitás és a belső dinamika közötti összefüggéseket.

Az NOE-k mint távolság-kényszerfeltételek a PDB adatbázisban nyilvánosan hozzáférhetők, míg a publikált gerinc-rendparamétereket a szerzők személyesen bocsátották a rendelkezésünkre. A felhasznált távolság-kényszerfeltételeket előszűrésnek vetettük alá: a sztereospecifikus távolságokat szimmetrizáltuk, és az így keletkező duplikátumokat eltávolítottuk. Azokat a távolság-kényszerfeltételeket, amelyek a kiindulási PDB sokaságban legalább 0,5 Å-mel sérültek, elhagytuk. Így végül a 2JZV, 2LJ4 és 2RQS szerkezetre rendre 1398, 1458 és 1132 távolság-kényszerfeltételt alkalmaztunk.

A távolság- és rendparaméter-kényszerfeltételeket együttesen a MUMO-protokollnak megfelelően alkalmaztuk: a rendparamétereket 8 párhuzamos replikára, az NOE-kat pedig párosával átlagoltuk (10. ábra). A szimulációkhoz az AMBER99SB erőteret és a TIP3P explicit oldószermodellt

használtuk. 8 párhuzamos replikát 10 ns-ig szimuláltunk, és az első 1 ns-ot elhagytuk a hőmérséklet- és nyomás-ekvilibráció céljából. A trajektóriákat minden 100 ps-ban mintavételeztük: így a létrejövő szerkezeti sokaságok egyenként 728 replikából álltak.

Mindhárom szerkezetre létrehoztunk egy-egy kényszerfeltételekkel megkötött sokaságot, valamint összehasonlításképpen egy-egy nem megkötött sokaságot is, amelyben pontosan ugyanazokat a kiindulási beállításokat alkalmaztuk, de a rendparaméter- és távolság-kényszerfeltételek erőállandóit 0-ra állítottuk.



25. ábra: Az elemzéshez felhasznált SaPrsA, TbPin1 és CsPinA parvulin-típusú cisz-transz izomerázok egymáshoz illesztett szerkezete és szekvenciája. Az illesztés a MAMMOTH-MULT (Lupyan et al 2005) programmal készült. **A:** A három fehérje reprezentatív szerkezete a  $C_{\alpha}$  atomok pozíciója szerint egymáshoz illesztve. A szerkezetek csak a konszenzus aminosavakat tartalmazzák. **B:** A hidrogénkötés-hálózatot alkotó 5 aminosav elhelyezkedése a TbPin1 szerkezetén bemutatva. A foszfátkötő hurok külön van jelölve ezen az ábrán, ugyanis ezt a régiót a három ábrázolt fehérje közül csak a TbPin1 tartalmazza. **C:** A három fehérje szekvencia-illesztése, külön kiemelve a hidrogénkötés-hálózatot alkotó aminosavakat.

## 6.2 A szerkezeti sokaságok elemzése

### 6.2.1 A kísérleti adatoknak való megfelelés

A szimulációk eredményeképpen tehát a SaPrsA, TbPin1 és CsPinA parvulin-típusú cisz-transz izomeráz fehérjékre egy-egy megkötött és nem megkötött sokaságot kaptunk, amelyeket összevettem az eredeti, PDB-ben elérhető kísérletes sokasággal is.

A peptidgerincre meghatározott globális RMSD értékek arról tanúskodnak, hogy mindegyik fehérje kompakt, jól definiált globuláris szerkezettel rendelkezik (6. táblázat). Feltűnő tendencia, hogy a flexibilis farokrégiók kizárásával meghatározott RMSD értékek mindegyik fehérje esetén az eredeti sokaság, megkötött sokaság és nem megkötött sokaság sorrendjét követve növekszenek, amely egyúttal a konformációs flexibilitás tendenciózus növekedésére is utal. Ez nem meglepő, ismerve a sokaságok előállítására használt eljárásokat. Az eredeti, PDB-ből nyert sokaság a *single structure refinement* módszerével készült, amely során mindegyik távolság-kényszerfeltételt rákényszerítik minden egyes replikára. Ezzel szemben az általunk előállított, megkötött sokaságban a valamivel megengedőbb, sokaság-átlagoláson alapuló megkötést alkalmaztuk.

Sokaság	Sokaság mérete	RMSD Å		S <sup>2</sup> korrel.	Kémiai eltolódás korrel.	
		Teljes molekula	Merev régió	Teljes molekula	Ca	Ha
<b>SaPrsA</b>						
eredeti	25	2,01 ± 0,65	0,75 ± 0,22	0,59	0,98	0,93
nem megkötött	728	2,09 ± 0,57	1,43 ± 0,38	0,58	0,98	0,92
megkötött	728	1,89 ± 0,65	1,12 ± 0,20	0,93	0,99	0,94
<b>TbPin1</b>						
eredeti	20	0,68 ± 0,18	0,43 ± 0,08	0,38	0,96	0,84
nem megkötött	728	2,12 ± 0,49	1,93 ± 0,46	0,16	0,97	0,86
megkötött	728	1,71 ± 0,52	1,58 ± 0,53	0,92	0,97	0,86
<b>CsPinA</b>						
eredeti	20	2,16 ± 0,72	0,64 ± 0,17	0,11	0,97	0,66
nem megkötött	728	3,35 ± 1,00	2,35 ± 0,80	0,23	0,98	0,72
megkötött	728	2,44 ± 0,72	1,21 ± 0,26	0,78*	0,98	0,75

6. táblázat: Az előállított parvulin-sokaságok gerinc RMSD-je és megfelelése a kísérletes paramétereknek. Az eredeti sokaságok a 2JZV, 2LJ4 és 2RQS PDB kóddal jelölt szerkezetek voltak sorrendben az SaPrsA, TbPin1 és CsPinA fehérjékre. Az RMSD értékek meghatározásakor a peptidgerinc mindegyik atomját figyelembe vettük. A merev régiókra meghatározott RMSD értékek számolásakor elhagytuk a fehérjék viszonylag nagy flexibilitást mutató farokrégióját, így a fennmaradó régiók a 6-111, 3-115 és 6-97 aminosavakat tartalmazzák sorrendben az SaPrsA, TbPin1 és CsPinA fehérjékre. Az RMSD értékeket a MOLMOL programmal, a kémiaieltolódás- és S<sup>2</sup>-korrelációkat a CoNSEnsX programmal számoltuk.

\*A CsPinA fehérje esetén a Gly97 aminosavra visszszámolt S<sup>2</sup> értéket elhagytuk a korreláció számolásakor.

A sokaságok hitelessége mellett szól, hogy bár kémiai eltolódásokat nem alkalmaztunk kényszerfeltételként, a kémiaieltolódás-korrelációk az általunk előállított sokaságban nem rosszabbak,

mint az eredeti sokaságban, sőt, a CsPinA esetében szignifikáns javulást tapasztalhatunk. Az  $S^2$ -megkötések alkalmazásával az  $S^2$ -korrelációk ugrásszerű növekedését tapasztalhatjuk. A CsPinA esetében a Gly97 aminosavra visszazámolt  $S^2$ -értéket elhagytuk a korreláció meghatározásakor, ugyanis ennél az aminosavnál egy konformációs elhajlást tapasztaltunk, ami két alternatív konformációt eredményezett a végső sokaságban. Azt valószínűleg a legfeljebb 0,5 Å-mel sérülő NOE-megkötések szigorú alkalmazása okozza: ezek elhagyásakor az effektus csökkenthető. Mindenesetre, mivel ez az aminosav nem szerepel a konszenzus szekvenciában, az elemzéseink további részét nem érinti.

## 6.2.2 A leghangsúlyosabb belső mozgások

Annak érdekében, hogy össze tudjam hasonlítani a három vizsgált fehérje belső mozgásait, egymáshoz illesztettem őket a többszörös szerkezetillesztés (*multiple structure alignment*) módszerrel, a MAMMOTH-MULT programcsomagot használva. A konszenzus szekvencia 89 aminosavat tartalmaz, és magába foglalja a kötőzsebet és az azt körülvevő lebenyek túlnyomó részét (25. ábra).

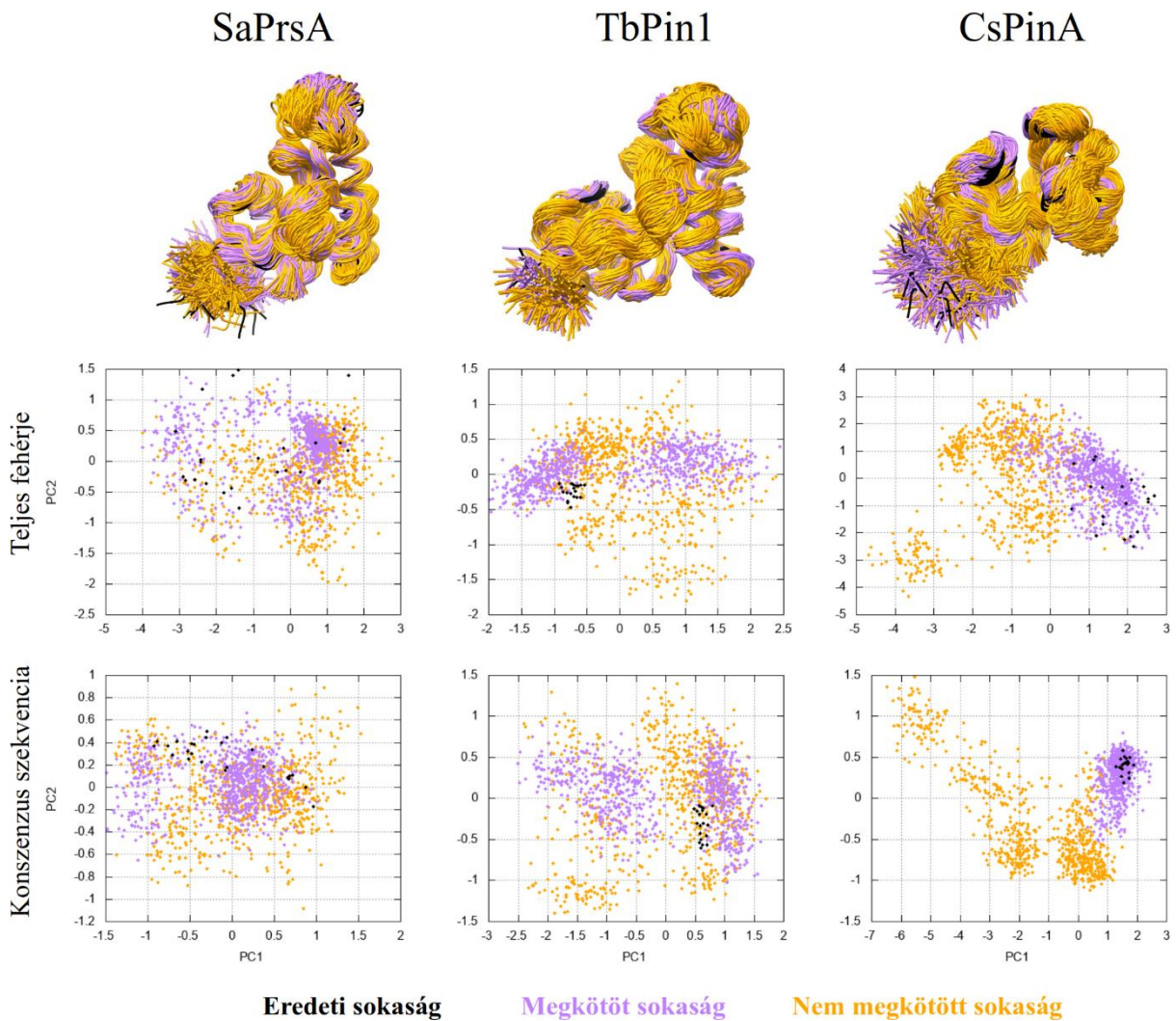
Elsőként elvégeztem egy-egy főkomponens-elemzést a három fehérjéhez tartozó eredeti, megkötött és nem megkötött sokaság összevonásával létrehozott sokaságon (26. ábra). Az adatpontok első két főkomponens mentén történő eloszlásából megerősíthetjük a globális RMSD-kból levont következtetést: a megkötött sokaságok szűkebb konformációs teret fednek le, mint a nem megkötött sokaságok. A főkomponens-ábrákon az is látszik, hogy (CsPinA esetét kivéve) a nem megkötött sokaságok által lefedett konformációs tér tartalmazza a megkötött sokaságokat, amelyek pedig tartalmazzák az eredeti sokaságokat. Ez a tendencia még feltűnőbb a csak a konszenzus szekvencia aminosavjaira elvégzett főkomponens-elemzés eredményét mutató ábrázolásokon.

A három különböző parvulin belső mozgásainak összehasonlításához a megkötött sokaságok konszenzus szekvenciájához tartozó aminosavjain egy újabb főkomponens-elemzést végeztem el (27. ábra). A három fehérje jól szeparálódik az első két főkomponens mentén, melyek a megkötött sokaságok esetén a teljes variancia 44 és 25%-át, a nem megkötött sokaságok esetén pedig 36 és 24%-át teszik ki.

A megkötött és nem megkötött sokaságokból számolt főkomponensek egymással alacsony átfedést mutatnak, leszámítva a megkötött sokaság első és a nem megkötött sokaság második főkomponensét (28.A ábra). Ez a főkomponens a kis és a nagy lebeny egymástól való eltávolodását és közeledését, egyfajta *légző mozgást* reprezentál, amely valójában éppen a kötőzseb kinyílásának és becsukódásának felel meg (28.B ábra). A kötőzseb nyitottságát intuitívan a két lebeny tetején lévő két aminosav  $C_\alpha$  atomjának távolságával lehet számszerűsíteni, amely a megkötött sokaságok esetén figyelemreméltóan jól korrelál az első főkomponenssel ( $R = 0,82$ ,  $p < 0,0001$ ). Világosan látszik, hogy a három fehérje más-más régiót fed le nemcsak az első főkomponens, hanem a kötőzseb nyitottságának megfelelő dimenzióban is: az SaPrsA a leginkább zárt, a CsPinA valamivel nyitottabb, és a TbPin1, amelyik a legszélesebb tartományt öleli át az ábrán, egyúttal a legnyitottabb (28.C ábra). A légző



mozgás tehát a TbPin1 fehérjében a legdominánsabb, és valamivel kisebb mértékben ugyan, de jelen van a másik két fehérjében is.

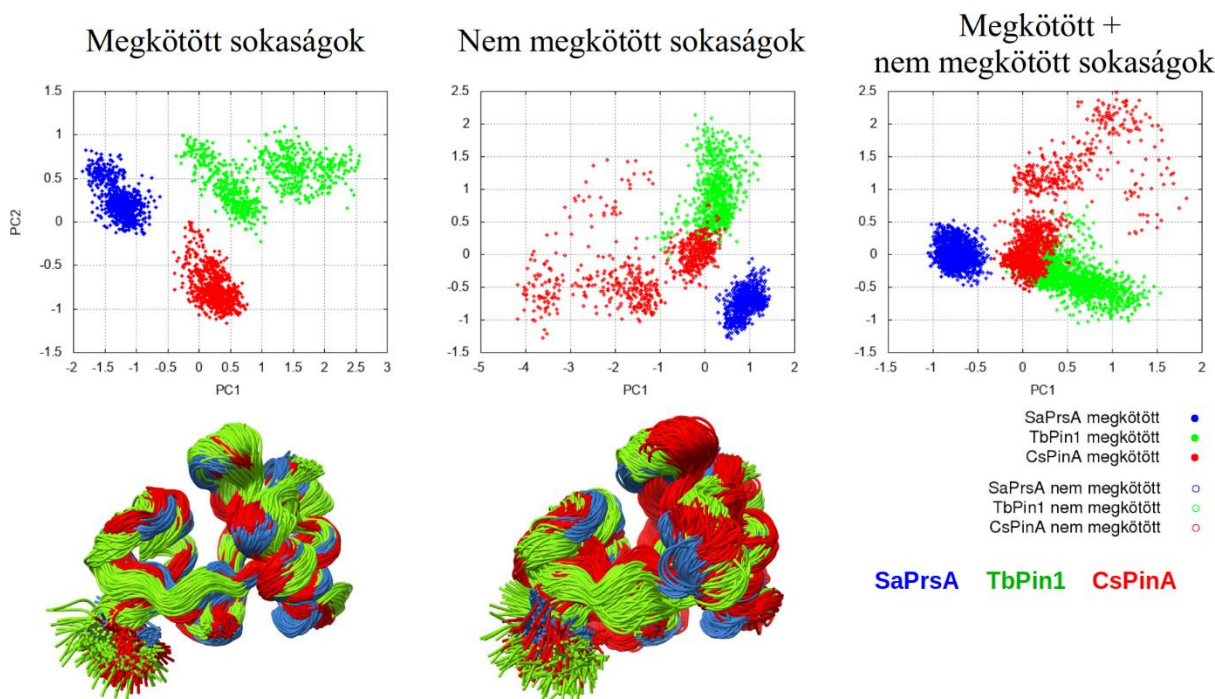


26. ábra: A három vizsgált fehérjéből készült sokaságok egymásra illetve (felső sor), illetve a megkötött, nem megkötött és az eredeti sokaságokat tartalmazó összevont sokaságokon elvégzett főkomponens-elemzés ábrázolása. Külön ábrák mutatják a teljes fehérjéken (középső sor), illetve csak a konszenzus szekvenciákon elvégzett (alsó sor) főkomponens-elemzés eredményeit.

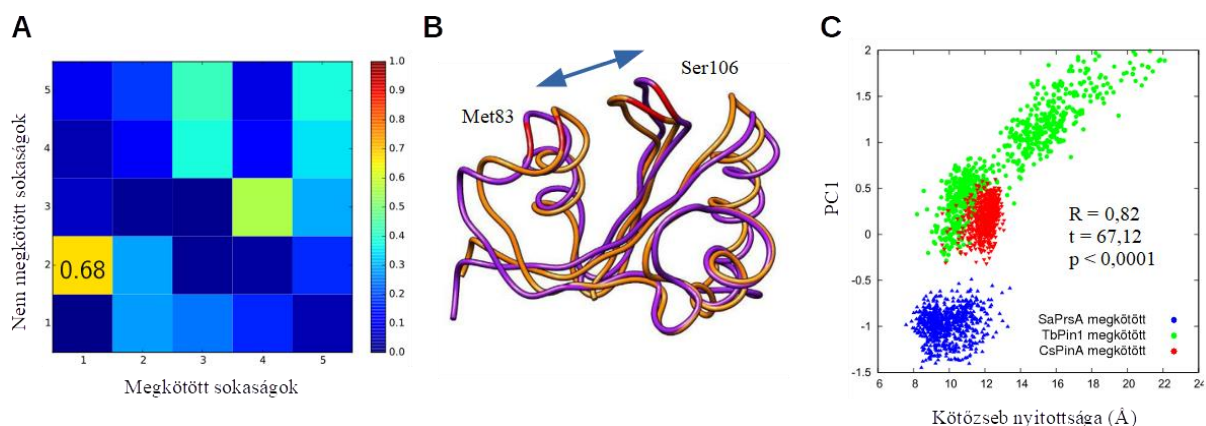
Az általános rendparaméterek a fehérje gyors (ps-ns) mozgásait reprezentálják. A létrehozott sokaságoktól elvárható, hogy egy, a natív állapothoz közeli, átlagos szerkezet körüli gyors fluktuációt tükrözzenek, amelyről a főkomponens-elemzés ad szemléletes képet. Azonban TbPin1 által széles tartományban mintavételezett légző mozgásától azt várnánk, hogy ennél valamivel lassabb időskálán történik.

Ezt a látszólagos ellentmondást az a feltevés oldja fel, amely szerint az  $S^2$ -kényszerfeltételek elsősorban nem a mozgások amplitúdójára, hanem azok irányára jelentenek megkötést. Az RMSD-értékek is azt mutatják, hogy a TbPin1 semmivel sem flexibilisebb a másik két fehérjénél (6. táblázat). A főkomponens-elemzés eredményeit is figyelembe véve azt mondhatjuk, hogy a TbPin1-ben a

konformációs fluktuáció a másik két fehérjéhez képest hasonló mértékű, de más irányban történik, ami más főkomponenseket eredményez.



27. ábra: A három különböző parvulin fehérjét tartalmazó egyesített sokaságok, és az azokra elvégzett főkomponens-analízis eredménye. Az elemzéshez csak a konszenzus szekvencia aminosavjait használtuk fel (ld. 25. ábra). A főkomponens-elemzést elvégeztük csak a megkötött, csak a nem megkötött, valamint a megkötött és nem megkötött sokaságokból álló kombinált sokaságra.



28. ábra: A megkötött parvulin-sokaságokra meghatározott első főkomponens és a kötőszob nyitottsága közötti összefüggés. **A:** A megkötött és nem megkötött sokaságok első öt főkomponense közötti átfedés, külön feltüntetve az egyik kiugró értéket. **B:** Az első főkomponens által reprezentált leghangsúlyosabb mozgás. A két szerkezet a két extrém konformációnak felel meg. Piros színnel van jelölve az a két aminosav, amelyeket a kötőszob nyitottságának definíciójához felhasználtam. **C:** A három sokaság eloszlása a kötőszob nyitottsága és az első főkomponens szerint, feltüntetve a korrelációt is.

Ne felejtjük el, hogy a légző mozgást reprezentáló, első főkomponens valójában a három megkötött sokaságot tartalmazó, kombinált sokaság első főkomponenseként adódott, így nem elvárható, hogy mindegyik fehérjében ez a bizonyos légző mozgás, még ha jelen is van, a leghangsúlyosabb mozgások között legyen. Ezt igazolják az egyes megkötött sokaságok főkomponensei és a kombinált

sokaság főkomponensei közötti átfedések (28. B ábra). A CsPinA-ban és a SaPrsA-ban a légző mozgás nem jelenik meg észrevehetően az első négy főkomponens között. A TbPin1-ben is csak az első főkomponenssel mutat némi átfedést, de ebbe a főkomponensbe más mozgások is belekeverednek.

A létrehozott sokaságok valójában nem bizonyítják, hogy a légző mozgás ténylegesen, a szimulációban is mintavételezett gyors időskálán történik. Egy plauzibilis értelmezés, hogy a sokaságok a gyors mozgások által lefedett konformációs tér maximumát mintavételezik, de a lassú mozgások szerinti más-más állapot körül. Mindenesetre a kísérletes  $S^2$ -eknek való jó megfelelés mindenképp megerősíti azt a feltevést, hogy ezekben a fehérjékben jelen van a légző mozgás. A TbPin1 ráadásul foszforilált ligandumok izomerizációját katalizálja, ezért a kötőzseb viszonylag nagy amplitúdójú kinyíló-becsukódó mozgásának ténylegesen lehet szerkezeti biológiai relevanciája. Ezt erősíti az a megfigyelés, hogy a TbPin1 foszfátkötő hurokrégiójában néhány aminosavnál NMR-es relaxációs kísérletekkel kémiai kicserélődést figyeltek meg, igaz, a lassabb időskálán (Sun et al., 2012).

### 6.2.3 Összevetés kísérletes parvulin szerkezetekkel

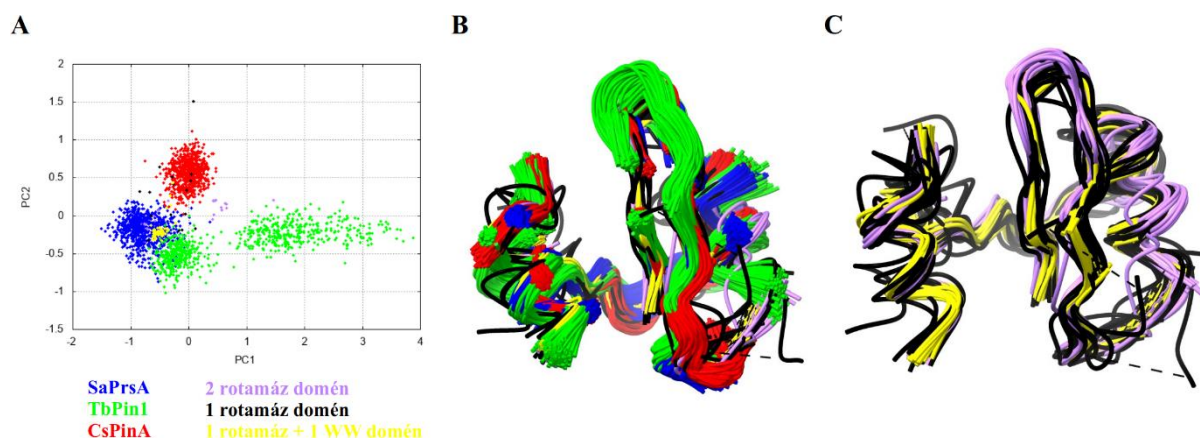
Az előállított sokaságokat összevettem száz darab, PDB-ből elérhető kísérletes parvulin-típusú rotamáz-szerkezettel. A kiválasztott domének listája tartalmazta a három vizsgált fehérje (SaPrsA, TbPin1 és CsPinA) rotamáz doménjét is. Az elemzésbe bevett szerkezeteket három csoportra osztottam aszerint, hogy a teljes fehérje pontosan egy rotamáz, két rotamáz vagy pedig egy rotamáz és egy WW domént tartalmaz. A felhasznált szerkezetek közül így némelyek ugyanannak a fehérjeláncnak részei, mások pedig egy adott PDB-szerkezet különböző fehérjeláncához tartoznak. Ilyen módon összesen 16 fehérjéről készült 78 PDB-szerkezetet használtam fel (F16. táblázat, F17. táblázat).

Első lépésként elvégeztem a száz parvulin domén többszörös (F18. ábra). A konszenzus-szekvencia 53 aminosavat tartalmaz, amelyet a továbbiakban szűkített konszenzusnak nevezek. Érdekes módon ennek a listának nem képezi részét az N-terminushoz közelebbi konzervált hisztidin, annak ellenére, hogy az irodalom szerkezetileg és funkcionálisan jelentős szerepet tulajdonít neki.

Ezt követően főkomponens-elemzést végeztem a három általunk előállított megkötött sokaságból, valamint a 100 kísérletes rotamáz doménből álló szerkezet 53 aminosavat tartalmazó szűkített konszenzus régióján (29. ábra). Az általunk előállított sokaságok jóval nagyobb konformációs teret fednek le, mint a PDB-ből nyert kísérletes szerkezetek. Ez első közelítésben meglepő lehet, hiszen a mi sokaságaink mindössze három különböző parvulin domént tartalmaznak, szemben a kísérletes szerkezetek széles változatosságával. Ugyanakkor a kristályszerkezet zsúfolt elrendezésbe kényszeríti a molekulákat, ami némi magyarázattal szolgál a kisebb konformációs variabilitásra.

Az első főkomponens, akárcsak a megkötött sokaságokra elvégzett elemzés esetén, a kötőzseb kinyíló-becsukódó mozgásának felel meg. A kísérletes sokaságok közül egyedül a két rotamáz domént tartalmazó fehérjék szerkezetei mutatnak jelentős eloszlást e főkomponens mentén. Az egyetlen ilyen fehérje, amelyről mindkét rotamáz domént tartalmazó kísérletes szerkezet készült, az 1M5Y, amely 4

ekvivalens láncot tartalmaz. Az N-terminális rotamáz domének mind pozitív, a C-terminál domének pedig mind negatív értéket vesznek fel az első főkomponens mentén. Erre az a magyarázat, hogy az N-terminális domént a nagy lebeny kiterjesztése veszi körül, amely a kötőzsebet valamelyest nyitottabb konformációba kényszeríti.



29. ábra: A három általunk előállított megkötött sokaságból, valamint a 100 kísérletes rotamáz doménből álló szerkezet 53 aminosavat tartalmazó szűkített konszenzus régiójának ábrázolása és az így létrejött sokaságon végzett főkomponens-elemzése. Az ábrán külön színnel vannak jelölve azon fehérjékből származó rotamáz domének, amelyek pontosan egy rotamáz, két rotamáz, vagy pedig egy rotamáz és egy WW domént tartalmaznak. **A:** Az adatpontok ábrázolása az első két főkomponens mentén. **B:** A három általunk előállított sokaság és a 100 kísérletes rotamáz egymásra illesztett szűkített konszenzus régiója. **C:** A 100 kísérletes rotamáz egymásra illesztett szűkített konszenzus régiója.

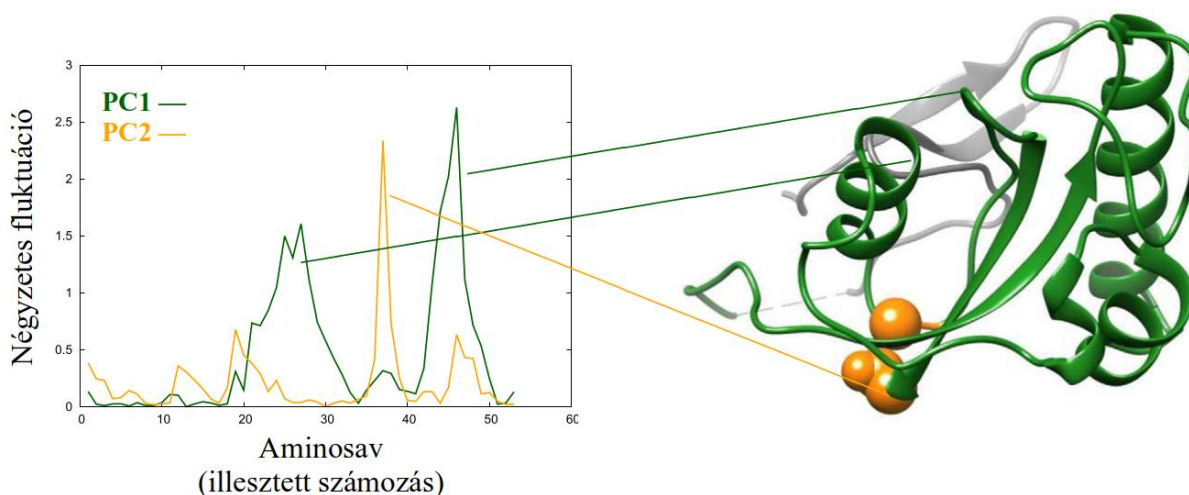
Ebben az elemzésben az általunk előállított megkötött sokaságok jelentős szeparációt mutatnak a második főkomponens mentén is (29. ábra). Ez a mozgás a kötőzseb alsó régiójának az elmozdulását reprezentálja, ahol a kis és nagy lebeny kapcsolódik egymáshoz. Ez valójában a légző mozgást lehetővé könyökrégió, hiszen ehhez kapcsolódva tud a két lebeny közeledni egymáshoz vagy távolodni egymástól. A második főkomponens épp ott mutatja a maximális fluktuációt, ahol az első főkomponens a két lebeny fluktuációjához képest jóval alacsonyabb értéket vesz fel: a Pin1 számozása szerint a Gly114 aminosavnál, vagy a konszenzus számozás szerinti 37. aminosavnál (30. ábra).

A 29.A ábra tanúsága szerint a kísérletes rotamáz szerkezeteknek is viszonylag nagy az eloszlása a második főkomponens mentén. Ebből praktikusán azt a következtetést vonhatjuk le, hogy ezek a konformációk egy lassabb időskálán történő mozgás mentén mintavételezett állapotot tükröznek. Ugyanakkor ebből az elemzésből nem állapítható meg, hogy ez egy minden parvulin-típusú rotamázra jellemző mozgás-e, valamint hogy ténylegesen hattással van-e a katalízisre.

Camilloni és Vendruscolo 2015-ös munkájából (Camilloni & Vendruscolo, 2015) kiindulva meghatároztam a különböző parvulin-szerkezetekben az elektrosztatikus térnek a ligandum prolinjának amidisíkjára merőleges komponensét. A tanulmány szerzői arra a következtetésre jutottak, hogy ez az elektrosztatikus erő fogantyúként stabilizálja a csavart amid átmeneti állapotot. Én nem találtam korrelációt a kötőzseb nyitottsága és az elektrosztatikus tér Z-komponense között a katalizált prolin amidisíkjában. Igaz, Camilloni és mtsai jóval pontosabb, kvantumkémiai számolásokkal határozták meg



az elektromos tér nagyságát, én viszont mindössze klasszikus fizikai megközelítést alkalmaztam, az alkalmazott molekuladinamikai erőter által az atomokra helyezett parciális töltéseket felhasználva.



30. ábra: Az első két főkomponens négyzetes fluktuációja a szűkített konszenzus aminosavsorszám függvényében. A főkomponens-elemzést a három általunk előállított megkötött sokaságból, valamint a 100 kísérletes rotamáz doménből álló szerkezet 53 aminosavat tartalmazó szűkített konszenzus régióján végeztük el. A jobb oldali ábrán a Pin1 fehérje azon régiói vannak jelölve, amelyek az első két főkomponens által reprezentált belső mozgásban a legnagyobb fluktuációt mutatják.

A könyökrégió egyúttal az a része a parvulin doméneknek, amelyik közvetlen kapcsolatban van a WW doménnel. A Pin1 esetén például mind a WW domént nélkülöző mutánsok, mind pedig a WW doménnel kölcsönhatást megakadályozó pontmutánsok szignifikánsan más aktivitást mutatnak, mint a vad típusú Pin1. Ezzel összhangban van az is, hogy a mi elemzésünkben is jól szeparálódnak a WW domént tartalmazó, illetve nem tartalmazó fehérjék a második főkomponens mentén, és ez a módus éppen a könyökrégió mozgásának felel, amely közvetlen kölcsönhatásban van a WW doménnel. Ezekből az észrevételekből azt a meggyőző következtetést lehet levonni, hogy a könyökrégió, akár mutációkon, akár a WW doménnel való kölcsönhatásán keresztül, modulálja a légzőmozgás amplitúdóját, és ezáltal a fehérje enzimatis aktivitását is.

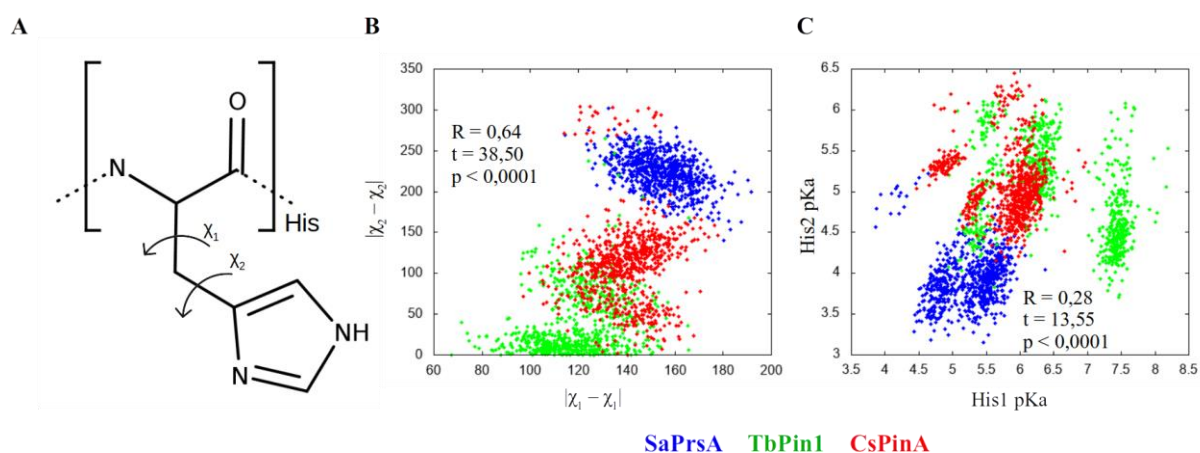
#### 6.2.4 A konzervált hisztidinek és a hidrogénkötés-hálózat vizsgálata

A parvulin-típusú rotamáz doménekben található két konzervált hisztidin, amelyek egy 5 aminosavból kialakított hidrogénkötés-hálózat részei (25.B ábra). A katalízisben betöltött lehetséges szerepük vizsgálatának céljából kiszámoltam e két hisztidin oldalláncának  $\chi_1$  és  $\chi_2$  torziós szögeit, valamint a PropKa predikciós programmal meghatároztam az imidazol-gyűrű pKa-értékét.

Egymás függvényében ábrázoltam két hisztidin  $\chi_1$  és  $\chi_2$  torziós szögei közötti különbség abszolút értékét, valamint a két hisztidin pKa-értékeit (31. ábra). Az oldallánc-orientációk viszonylag magas korrelációt mutatnak ( $R = 0,64$ ), a két hisztidin pKa értékei ennél jóval alacsonyabbat ( $R = 0,28$ ). Azonban a meghatározott Pearson-korrelációs koefficiens mindkét esetben szignifikáns ( $p < 0,0001$ ) és ez kellő bizonyítékul szolgál arra, hogy a két hisztidin viselkedése összefügg egymással, és az is

valószínűsíthető, hogy a belső dinamika és a katalízis szabályozásában is egymással együttműködve vesznek részt.

A 31.B és 31.C ábrákon azt is megfigyelhetjük, hogy a három vizsgált fehérjéből előállított sokaság valamelyest elszeparálódik, akár csak az első két főkomponens által reprezentált leghangúlyosabb belső mozgások mentén. Jelen esetben ez azonban nem csak az eltérő kiindulási konformációval magyarázható, hanem a vizsgált hisztidinek protonáltsági állapotával (32. ábra). A TbPin1 és a CsPinA fehérjék kísérletes szerkezeteiben az imidazolgyűrű teljesen protonált állapotban volt. Ezzel szemben a SaPrsA esetén a szerzők részletesen megvizsgálták a hisztidin protonáltságot, amelyet a PDB-be feltöltött kísérletes szerkezet is megőrzött.

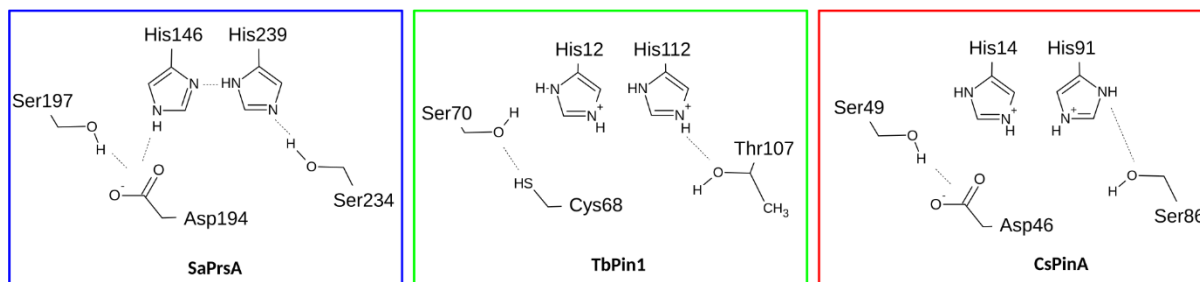


31. ábra: A két konzervált hisztidin orientációja és pKa értéke. Az ábrázolt adatpontok a három vizsgált fehérjére előállított megkötött sokaságokhoz tartoznak. A: A hisztidin aminosav oldalláncának orientációja  $\chi_1$  és  $\chi_2$  torziós szögekkel definiálva. B: A két hisztidinhez tartozó  $\chi_1$  és  $\chi_2$  torziós szögek közötti különbségek abszolút értékei, egymás függvényében ábrázolva. C: A két hisztidin pKa értéke, egymás függvényében ábrázolva. A predikció a PropKa programmal készült.

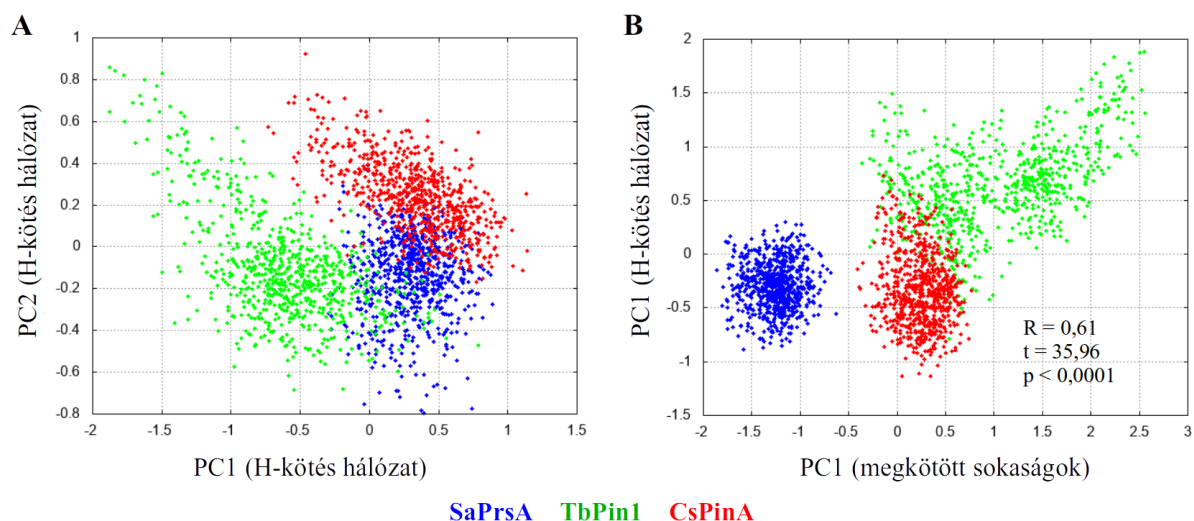
Az eltérő protonáltsági állapot más-más hidrogénkötés-mintázat kialakítását teszi lehetővé a hisztidinek számára, amely közvetlen hatással van az oldalláncok konformációs szabadságára is. A szimulációkhoz felhasznált kiindulási szerkezetek közül egyedül a SaPrsA fehérjében volt teljes a hidrogénkötés-hálózat, a másik két esetben az N-terminális hisztidin nem tud hidrogénkötést kialakítani. Ezzel összhangban ebben a fehérjében a legkisebb a két hisztidin oldallánc-torziósszögeinek variabilitása (31.B ábra).

Hogy képet kapjak a hidrogénkötés-hálózat hatásáról a fehérje belső dinamikájára, elvégeztem egy főkomponens-elemzést azon az öt aminosavon, amely érintett a hálózatban (33. ábra). Figyelemre méltó, hogy mind az öt aminosav a nagy lebenyben van. Az első főkomponens a hálózat két végén található Ser/Thr és Cys/Asp relatív elmozdulását, a második főkomponens pedig a két hisztidin oldalláncainak reorientációját reprezentálja. Az első főkomponens közepes, de szignifikáns korrelációt mutat ( $R = 0,61$ ,  $p < 0,0001$ ) az összes megkötött sokaságra elvégzett elemzésből adódó első főkomponenssel, ami a légzőmozgást reprezentálja (33.B ábra). Ez nem meglepő annak fényében, hogy a hidrogénkötés-hálózatot alkotó aminosavak a konszenzus-szekvencia részei, tehát valójában a

megkötött sokaságok részhalmazát képezik. Azonban az első főkomponens által reprezentált mozgásban a két lebeny relatív elmozdulása dominál, míg a hidrogénkötés-hálózatra meghatározott főkomponensek a nagy lebeny belső mozgását reprezentálják. Mindezen eredmények alapján csábító azzal a feltételezéssel élni, hogy a hidrogénkötés dinamikus kialakulása-felbomlása modulálja a nagy lebeny  $\beta$ -szálainak mozgását, megakadályozva a nagy amplitúdójú konformációs átrendeződéseket.



32. ábra: A három vizsgált fehérjéhez tartozó kiindulási szerkezetekben található két konzervált hisztidin protonáltsági állapota, és a protonáltsági állapot függvényében létrejövő hidrogénkötés-hálózat a nagy lebeny 5 érintett aminosav-oldallánca között (ld. 25. ábra).



33. ábra: A hidrogénkötés-hálózatban érintett 5 aminosavra elvégzett főkomponens-elemzés. Az érintett aminosavak a konszenzus-szekvencia szerint: His9, Cys/Asp40, Ser42, Thr/Ser79 és His84 (ld. 25. ábra). Az ábrázolt adatpontok a három vizsgált fehérjére előállított megkötött sokaságokhoz tartoznak. **A:** Az adatpontok eloszlása az 5, hidrogénkötés-hálózatban érintett aminosavra meghatározott első és második főkomponens mentén. **B:** A hidrogénkötés-hálózatra meghatározott első főkomponens és a három megkötött sokaságra meghatározott első főkomponens egymás függvényében ábrázolva, a korreláció feltüntetésével.

### 6.3 Diszkusszió – Második téziscsoport

A peptidil-prolil cisz-transz izomerázok vagy rotamázok a peptidláncban egy prolin előtti peptidkötés cisz-transz izomerizációját katalizálják (Hanes, 2015). A PPIáz enzimek minden egyes sejtalkotóban jelen vannak és igen sokféle biológiai folyamatban vesznek részt: például a fehérjefeltékeredésben, a sejtciklus szabályozásában, az apoptózisban vagy a transzkripcióban. Számos betegség kezeléséhez terápiás célpontot jelenthetnek, ugyanis sok, kórosan elváltozott sejtben túlexpresszálódnak, akár olyannyira, hogy a betegség markerének tekinthetők. Mindezek a biológiai funkciók a PPIáz enzimeket a sejtbiológia egyik központi szereplőjévé teszik (Göthel & Marahiel, 1999; K. P. Lu et al., 2007).

Bár egy viszonylag egyszerű kémiai átalakulást katalizálnak, az enzim működésének pontos mechanizmusa nem ismert. Általánosan elfogadott, hogy a katalízis során nem bomlik fel a peptidkötés, hanem a reakció egy csavart átmeneti állapotton keresztül játszódik le.

Számos jel arra utal, hogy a parvulin-típusú rotamázok szerkezetében található két konzervált hisztidin, illetve az öt aminosav által – beleértve e két hisztidint, továbbá két szerint és egy ciszteint is – kialakított hidrogénkötés-hálózatnak fontos szerep jut a katalízisben. Kiderült azonban, hogy a kettős hisztidin-mutáns és a cisztein pontmutáns is katalitikusan aktív marad, valamint hogy a ligandumkötéskor a hisztidinek kémiai eltolódása csak minimálisan változik meg, tehát nem lépnek közvetlen kölcsönhatásba a ligandummal (Bailey et al., 2008; Terada et al., 2001). Ezek az eredmények azt sejtetik, hogy mind a konzervált hisztidineknek, mind pedig a hidrogénkötés-hálózatot kialakító aminosavaknak inkább szerkezeti és dinamika szerepük van, és nem vesznek részt közvetlenül a katalitikus mechanizmusban.

A Pin1-típusú parvulinok foszforilációfüggő mechanizmust követnek: a prolin előtti aminosav csak foszforilált treonin vagy szerin lehet. Ezek a parvulinok a katalitikus domén mellett tartalmaznak egy N-terminális WW domént, amelynek szerepe szintén nem teljesen felderített. Mivel azok a pontmutánsok, amelyekben a mutáció megbontja a két domén közötti kölcsönhatást, inaktívak, arra következtettek, hogy a WW domén allostérikus módon szabályozza a katalitikus domén dinamikáját (Olsson et al., 2016).

Hogy felderítsük a parvulin-típusú cisz-transz izomerázok belső mozgásai és katalitikus mechanizmusa közötti lehetséges összefüggéseket, valamint a hidrogénkötés-hálózat és a WW doménnel kialakított kölcsönhatás szerepét, dinamikus fehérjeszerkezeti sokaságokat állítottunk elő a SaPrsA, TbPin1 és a CsPinA fehérjéről. A sokaságok előállításához a MUMO-protokollnak megfelelően kísérletes  $S^2$  rendparaméter- és  $^1\text{H}$ - $^1\text{H}$  NOE távolság-megkötéseket alkalmaztunk külső kényszerfeltételként, de mindhárom fehérjéhez előállítottunk egy-egy nem megkötött sokaságot is.

Az sokaságok elemzésekor megállapítottam, hogy jó egyezést mutatnak a kísérletes dinamikai paraméterekkel. Külön kiemelendő, hogy a kényszerfeltételek alkalmazásával úgy tudtunk javulást elérni a  $S^2$  rendparamétereknek való megfelelésben, hogy közben nem rontottuk el a kémiai eltolódásoknak való megfelelést, amelyeket viszont nem használtunk fel a sokaságok előállításához,



csak a validációjához. Ez alapján elmondható, hogy olyan szerkezeti sokaságokat sikerült előállítanunk, amelyek jól tükrözik a gyors (ps-ns) időskálán történő belső mozgásokat.

*2.1. tézis: Főkomponens elemzés segítségével megállapítottam, hogy a parvulin-típusú rotamázokra jellemző a kötőzseb kinyíló-becsukódó mozgása. E mozgás mentén a három vizsgált fehérje más-más állapotot mintavételez.*

A parvulin-sokaságok elemzésekor adódó első főkomponens által reprezentált leghangsúlyosabb belső mozgás megfeleltethető a kötőzseb kinyíló-becsukódó mozgásának, amelyet légző mozgásnak hívtam. A három fehérjéről előállított sokaság más-más régiót fed le az első főkomponens mentén (28. ábra).

A légzőmozgás egyaránt jelen van a nem megkötött sokaságokban is, ám az ezeken végzett főkomponens-elemzés alapján ez csak a második leghangsúlyosabb belső mozgás.

A három vizsgált fehérje közül a TbPin1 fedi le a legnagyobb régiót az első főkomponens mentén, ami arra utal, hogy ez a mozgás leginkább ebben a fehérjében van jelen. Figyelembe kell vennünk azonban, hogy a szimulációk a fehérjék gyors belső mozgásait tudják mintavételezni, ezenkívül a kényszerfeltételként alkalmazott  $S^2$ -megkötések is a gyors dinamikára vonatkoznak. Ugyanakkor a katalízis időállandójával egy nagyságrendbe eső konformációs kicserélődés ennél lassabb időskálán volt csak megfigyelhető.

A légző mozgás csak a három megkötött sokaság uniójából adódó sokaságban volt a leghangsúlyosabb. Ez arra utal, hogy valóban jelen van a parvulinokban, ám a gyors dinamikát tükröző sokaságok ezt a mozgást nem tudják önmagukban tükrözni, hanem a lassú légzőmozgás szerinti más-más állapotot mintavételeznek.

*2.2. tézis: Az előállított sokaságokat összevettem nyilvános adatbázisban elérhető kísérletes szerkezetekkel. A konszenzus szekvencián elvégzett főkomponens elemzés segítségével megállapítottam, hogy a parvulin-típusú rotamázokra szintén jellemző a kis és nagy lebenyt összekötő könyökrégió elmozdulása, amely egyúttal a WW doménnel rendelkező parvulinokban a két domén közötti kölcsönhatásért felelős.*

Az általunk előállított sokaság és további 100, kísérletes rotamáz szerkezet összevonásával készült sokaság elemzésekor az derült ki, hogy a második leghangsúlyosabb belső mozgás a légzőmozgást lehetővé tévő könyökrégió (a nagy és kis lebeny csatlakozási pontjának) elmozdulását reprezentálja (29-30. ábra). E felismerésnek jelentősége, hogy épp ez a régió alakít ki közvetlen kölcsönhatást a szomszédos WW doménnel. Ezzel összhangban van az a kísérletes megfigyelés, hogy a Pin1 esetén a könyökrégióban található, I28A mutáció anélkül csökkenti le az enzimikus aktivitást, hogy közben befolyásolná bármely domén szerkezetét (Wilson et al., 2013).

A három előállított sokaság az ilyen módon adódó második főkomponens mentén is viszonylag jól szeparálódik, tehát nemcsak a légzőmozgás, hanem a könyökrégió elmozdulása mentén is más-más állapotot mintavételeznek. Ebből az következik, hogy e két fontos mozgás egymással összefüggésben van.

A WW domént tartalmazó és nem tartalmazó parvulinok valamelyest szeparálódnak a könyökrégió mozgását reprezentáló főkomponens mentén. Ez a megfigyelés összhangban van azzal a feltételezéssel, hogy a WW domén, amennyiben jelen van, allostérikus módon szabályozza a katalitikus domén dinamikáját. Az itt elvégzett elemzések magyarázatot adnak a WW domén szerepére a katalitikus mechanizmusban: a könyökrégió kontrollálásán keresztül modulálják a kötőzseb kinyílását.

*2.3. tézis: Megállapítottam, hogy a hidrogénkötés-hálózat protonáltsági állapota befolyásolja nagy lebeny belső dinamikáját a parvulin-típusú rotamázokban.*

A két konzervált hisztidin, amely része a hidrogénkötés-hálózatnak, egymással valamelyest összefüggő módon reorientálódik, ahogy erre mind a prediktált pKa értékekből, mind pedig az oldallánc-torziósszögek változásából következtettem (31. ábra). Ezenkívül észrevettem, hogy az öt, hidrogénkötés-hálózatot kialakító aminosavon elvégzett főkomponens-elemzés során adódó leghangsúlyosabb belső mozgás korrelál a légzőmozgással (33. ábra). A korreláció önmagában természetesen nem jelent ok-okozati összefüggést. Azonban az általunk előállított sokaságok észrevehetően szeparálódnak mind e főkomponens mentén, mind pedig a hisztidinek oldallánc-orientációja szerint, amely megmagyarázható az eltérő protonáltsági állapotokkal. A SaPrsA fehérjében a hisztidinek egyszeresen protonáltak, ami a teljes hidrogénkötés-hálózat létrejöttét lehetővé teszi, a másik két fehérjében azonban a hisztidinek kétszeres protonációja a hálózatnak csak részletes létrejöttét engedi meg.

Az eredményeink publikálását követően megjelent egy tanulmány, amely NMR-relaxációs és H/D kicserélődés kísérletekkel szintén arra a konklúzióra jutott, hogy a hidrogénkötés-hálózat állapota modulálja a Pin1 katalitikus aktivitását. Ezek a kísérletes eredmények megerősítik az általunk előállított sokaságok és az azokból levont következtetések validitását (J. Wang et al., 2017).

*2.4. tézis: A szerkezeti sokaságok elemzéséből levont következtetések alapján egy általános modellt javasoltam a parvulin-típusú rotamázok szabályozására. A kötőzseb kinyíló-becsukódó mozgásának dinamikáját és amplitúdóját egyrészt a hidrogénkötés-hálózat protonáltsági állapota, másrészt a WW domén (amennyiben jelen van) a könyökrégió mozgásán keresztül modulálja. A különböző parvulinok közötti eltérő aktivitás a kötőzseb eltérően modulált kinyíló-becsukódó dinamikájával magyarázható.*

A bemutatott eredmények a parvulin-típusú rotamázokban a léggő mozgás amplitúdója és dinamikája határozza meg a ligandum-szelektivitást és szabályozza a katalízist. A léggő mozgást pedig két további faktor összehangolt hatása modulálja: nevezetesen a WW doménnel való kölcsönhatás és a nagy lebenyben található hidrogénkötés-hálózat állapota. Elképzelhető, hogy ez a két hatás nem független egymástól, bár ennek teszteléséhez, valamint e szabályozó mechanizmusoknak a Pin1-ben azonosított hidrofób vezetékkel (Namanja et al., 2007) (ld. 2.2.2. fejezet) való összefüggésének vizsgálatához további, gondosan megtervezett *in vitro* kísérletekre van szükség.

Ez a modell, bár szükségszerű egyszerűsítéseket tartalmaz, ezért óvatosan szabad csak belőle következtetéseket levonni, mégis mindenképpen biztató abból a szempontból, hogy összhangban van számos korábbi eredménnyel. A Pin1 molekuladinamikai szimulációi is azt mutatták, hogy a kötősebet körülvevő hurkok mozgása a katalitikus és WW domén közötti kölcsönhatástól, valamint a mind a katalitikus, mind a WW doménnek a ligandummal kialakított kölcsönhatásától is függ (Guo et al., 2015; Olsson et al., 2016). A hidrogénkötés-hálózatról tett megállapításaink megfeleltethetők a Barman and Hamelberg által javasolt dinamikus hidrogénkötésnek, amely meghatározza a nagy lebeny dinamikáját (Barman & Hamelberg, 2014). A mienkéhez hasonló eredményekre jutottak az egyik FKBP-típusú rotamáz vizsgálatakor: a kötősebet dinamikája itt is meghatározó szerepet tölt be a ligandumkötésben és a katalízisben (Quistgaard et al., 2016). Mindezek alapján valószínűsíthető, hogy ez egy általános mechanizmus a fehérjék működésének finomhangolására.

## 7 Konklúzió és kitekintés

A számítási kapacitás növekedésével, valamint az egyre szofisztikáltabb elméleti megközelítések folyamatos fejlődésével a számítógépes szimuláció mint tudományos módszer mára a kísérletes megfigyelés egyenrangú társává nőtte ki magát. Ez különösen igaz olyan bonyolult rendszerek vizsgálatakor, mint egy fehérjemolekula, amelynek belső dinamikáját még az erre leginkább alkalmas NMR-es kísérletekkel is csak közvetve lehet megfigyelni. Egy figyelmesen megtervezett és felparaméterezett molekuladinamikai szimuláció gyakorlatilag egy önálló kísérletnek tekinthető, amely nélkül igen nehéz még a legalapvetőbb NMR-es mérési eredményeket is értelmezni. Ugyanakkor egy szimuláció érvényességét és az abból levont következtetések jogosságát csakis egy valós kísérlettel lehet megerősíteni. Mérés és szimuláció tehát egymást kiegészítő eszközök lehetnek egy valós rendszer vizsgálatához, melyek között egy iteratív módon fejlesztett elméleti modell teremt kapcsolatot.

Azok a fehérjék, amelyek belső dinamikája meghatározó szerepet tölt be a biológiai funkciójuk szabályozásában, csak sokaságalapú megközelítéssel vizsgálhatók. A számtalan, különböző időskálán történő lokális vagy globális fluktuáció, az allosztérikus szabályozás, a kötőpartner hatása a populáció eloszlására vagy egy belső mozgás amplitúdójának és irányának megváltozására: ezek mind olyan jelenségek, amelyeket lehetetlen egyszerre figyelembe venni a hagyományos szerkezet-funkció paradigma keretein belül. Ennek következménye, hogy a dinamikus fehérjeszerkezeti sokaságokból levont következtetések a fehérje működésének atomi szintű mechanizmusára nézve szükségszerűen más nyelvi eszköztárral élnek. Konformáció, konformációs kicserélődés, egy-egy jól definiált kölcsönhatás kialakulása és felbomlása helyett eloszlásokról, fluktuációkról, belső mozgások amplitúdójának megváltozásáról kapunk információt.

Ebben a dolgozatban két olyan fehérjemolekula molekuladinamikai szimulációval történő vizsgálatát mutattam be, amelyekre számos kísérletesen meghatározott paraméter áll rendelkezésre, ám dinamikus jellegük miatt ezek önmagukban nehezen értelmezhetők. Az itt bemutatott eredmények jó példát szolgáltatnak arra, hogy a számítógépes szimuláció tényleges többletinformációt ad a rendszer viselkedéséről, a mérési eredményekből közvetlenül levonható következtetéseken túl.

Mind a PSD-95 N-terminális PDZ1-2 tandemjének, mind a parvulin-típusú peptidil-prolil cisz-transz izomerázoknak vizsgálata során külső kényszerfeltételekkel megkötött molekuladinamikai szimulációkkal állítottunk elő dinamikus szerkezeti sokaságokat. Kényszerfeltételekként a peptidgerince meghatározott  $S^2$  rendparamétereket, valamint  $^1\text{H}$ - $^1\text{H}$  NOE távolság-megkötéseket alkalmaztunk, melyek a molekula gyors (ns-ps) mozgásait tükrözik. A szimuláció során a vizsgált molekulákat mintegy rákényszerítjük, hogy tükrözzék a kísérletesen megfigyelt paramétereket. Ennek eredménye azonban jóval több annál, hogy az előállított sokaság jó megfelelést mutat a felhasznált paramétereknek: az így megjelenő belső mozgások vizsgálatkor alapvető biológiai szabályozó mechanizmusokra bukkanhatunk.

A két rendszer intradomén mozgásainak vizsgálata során egy közös elv látszik kibontakozni: a fehérjemolekula leghangsúlyosabb mozgásai között mindkét esetben megjelenik egy olyan fluktuáció, amely a ligandumkötő zseb kinyílását és becsukódását eredményezi, és amely mozgás természete (a mozgásban résztvevő régiók fluktuációjának mértéke és iránya) megváltozik a kötőpartnerrel kialakított kölcsönhatás során. Ezenkívül mindkét fehérje esetén azt láttuk, hogy ez a kinyíló-becsukódó mozgás közvetlen összefüggésben van a molekula egyéb, biológiailag releváns szerkezeti vagy dinamikai állapotával. Az eredményeink alapján arra következtethetünk, hogy a kötőzseb dinamikájának modulációja általános szabályozó mechanizmus lehet globuláris fehérjékben.

A PDZ1-2 tandem belső dinamikája magában foglal egy eggyel magasabb szinten, és ennél fogva más időskálán történő fluktuációt: a szupramoduláris reorientációt. A molekuladinamikai szimuláció ezért egy újfajta megközelítéssel történt: a gyors mozgásokra vonatkozó  $S^2$ -megkötéseket lokális illesztési algoritmussal külön-külön alkalmaztuk az egyes doménekre, és ezzel nem befolyásoltuk az interdomén dinamikát. Az előállított szerkezeti sokaságokból egyértelmű összefüggést figyeltünk meg az intra- és interdomén mozgások között. A kötőpartnerrel kialakított lokális kölcsönhatások a kötőzseb dinamikájának modulációján keresztül hatással voltak a két domén közötti interfészre, és ezáltal a szupramoduláris dinamikára is. Ez a mechanizmus szintén általános modellként szolgálhat többdoménű rendszerek biológiai funkciójának vizsgálatához.

A számítógépes szimulációkból levont következtetések, bármennyire is intuitívak, nem nélkülözhetik a kísérletes megfigyeléseknek való megfelelést. Ezért mindkét vizsgált rendszerről előállított sokaságot összevetettük elérhető kísérletes szerkezetekkel. Az eredményeink hitelességét erősíti, hogy a sokaságaink által lefedett konformációs tér magában foglalja ezeket a szerkezeteket. A már meglévő kísérletes szerkezetek eloszlása és pontos elhelyezkedése a konformációs térben egy észrevehető mintázat mentén történik: a parvulinok esetén ez a kötőzseb mozgását lehetővé tevő könyökrégió fluktuációjában, a PDZ tandem esetén pedig a pontos interdomén orientációban mutatkozik meg. Ezek a mintázatok arra is utalhatnak, hogy a sokaságainknak prediktív ereje van – bár ennek a feltételezésnek az igazolásához nagyobb mintán elvégzett vizsgálatok szükségesek. Kiemelendő azonban, hogy mindkét munkánk publikálása után születtek új kísérletes eredmények, és ezek összhangban voltak az általunk előállított sokaságokkal. Ezekben a tanulmányokban bizonyítást nyert a parvulinokban lévő hidrogénkötés-hálózat szerepe a kötőzseb dinamikájának modulációjában (J. Wang et al., 2017), valamint megjelent két új kísérletes PDZ tandem szerkezet (Ahmad Rodzli et al., 2019), amelyeket korábban nem figyeltek meg, ám amelyek létezését valószínűsíteni lehetett a szoros illeszkedésű PDZ tandem sokaságaink alapján.

A két vizsgált rendszerről előállított szerkezeti sokaságok elemzése során valóban olyan következtetésekre jutottunk, amelyeket pusztán a mérési eredményekből nem lehetett volna levonni. Az itt bemutatott kutatási módszer és az elért eredmények jó példái lehetnek a különböző tudományterületek közötti egymásra utaltságának, és az ebből fakadó gyümölcsöző együttműködésnek.

## 8 Irodalomjegyzék

- Ahmad Rodzli, N., Lockhart-Cairns, M., Levy, C. W., Chipperfield, J., Bird, L., Baldock, C., & Prince, S. M. (2019). How the dual PDZ domain from Postsynaptic density protein 95 clusters ion channels and receptors. Running title: Clustering by PDZ1-2 of PSD-95. *BioRxiv*, 775726. <https://doi.org/10.1101/775726>
- Alderson, T. R., Lee, J. H., Charlier, C., Ying, J., & Bax, A. (2018). Propensity for cis-Proline Formation in Unfolded Proteins. *ChemBioChem*, 19(1), 37–42. <https://doi.org/10.1002/cbic.201700548>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Amadei, A., Linssen, A. B. M., & Berendsen, H. J. C. (1993). Essential dynamics of proteins. *Proteins: Structure, Function, and Genetics*, 17(4), 412–425. <https://doi.org/10.1002/prot.340170408>
- Ángyán, A. F., & Gáspári, Z. (2013). Ensemble-based interpretations of NMR structural data to describe protein internal dynamics. *Molecules (Basel, Switzerland)*, 18(9), 10548–10567. <https://doi.org/10.3390/molecules180910548>
- Ayala, G., Wang, D., Wulf, G., Frolov, A., Li, R., Sowadski, J., Wheeler, T. M., Lu, K. P., & Bao, L. (2003). The prolyl isomerase Pin1 is a novel prognostic marker in human prostate cancer. *Cancer Research*, 63(19), 6244–6251. <http://www.ncbi.nlm.nih.gov/pubmed/14559810>
- Bach, A., Clausen, B. H., Møller, M., Vestergaard, B., Chi, C. N., Round, A., Sørensen, P. L., Nissen, K. B., Kastrop, J. S., Gajhede, M., Jemth, P., Kristensen, A. S., Lundström, P., Lambertsen, K. L., & Strømgaard, K. (2012). A high-affinity, dimeric inhibitor of PSD-95 bivalently interacts with PDZ1-2 and protects against ischemic brain damage. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), 3317–3322. <https://doi.org/10.1073/pnas.1113761109>
- Bailey, M. L., Shilton, B. H., Brandl, C. J., & Litchfield, D. W. (2008). The dual histidine motif in the active site of Pin1 has a structural rather than catalytic role. *Biochemistry*, 47(44), 11481–11489. <https://doi.org/10.1021/bi800964q>
- Bakan, A., & Bahar, I. (2009). The intrinsic dynamics of enzymes plays a dominant role in determining the structural changes induced upon inhibitor binding. *Proceedings of the National Academy of Sciences*, 106(34), 14349–14354. <https://doi.org/10.1073/pnas.0904214106>
- Bakan, Ahmet, Dutta, A., Mao, W., Liu, Y., Chennubhotla, C., Lezon, T. R., & Bahar, I. (2014). Evol and ProDy for bridging protein sequence evolution and structural dynamics. *Bioinformatics*, 30(18), 2681–2683. <https://doi.org/10.1093/bioinformatics/btu336>
- Bakan, Ahmet, Meireles, L. M., & Bahar, I. (2011). ProDy: Protein dynamics inferred from theory and experiments. *Bioinformatics*, 27(11), 1575–1577. <https://doi.org/10.1093/bioinformatics/btr168>
- Baker, E. N., & Hubbard, R. E. (1984). Hydrogen bonding in globular proteins. *Progress in Biophysics and Molecular Biology*, 44(2), 97–179. <http://www.ncbi.nlm.nih.gov/pubmed/6385134>
- Balastik, M., Lim, J., Pastorino, L., & Lu, K. P. (2007). Pin1 in Alzheimer's disease: multiple

- substrates, one regulatory mechanism? *Biochimica et Biophysica Acta*, 1772(4), 422–429. <https://doi.org/10.1016/j.bbadis.2007.01.006>
- Bao, L., Kimzey, A., Sauter, G., Sowadski, J. M., Lu, K. P., & Wang, D.-G. (2004). Prevalent overexpression of prolyl isomerase Pin1 in human cancers. *The American Journal of Pathology*, 164(5), 1727–1737. [https://doi.org/10.1016/S0002-9440\(10\)63731-5](https://doi.org/10.1016/S0002-9440(10)63731-5)
- Barman, A., & Hamelberg, D. (2014). Cysteine-mediated dynamic hydrogen-bonding network in the active site of Pin1. *Biochemistry*, 53(23), 3839–3850. <https://doi.org/10.1021/bi5000977>
- Bateman, A. (2019). UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Research*, 47(D1), D506–D515. <https://doi.org/10.1093/nar/gky1049>
- Beauchamp, K. A., Pande, V. S., & Das, R. (2014). Bayesian energy landscape tilting: Towards concordant models of molecular ensembles. *Biophysical Journal*, 106(6), 1381–1390. <https://doi.org/10.1016/j.bpj.2014.02.009>
- Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8), 3684–3690. <https://doi.org/10.1063/1.448118>
- Berlin, K., Castañeda, C. A., Schneidman-Duhovny, D., Sali, A., Nava-Tudela, A., & Fushman, D. (2013). Recovering a representative conformational ensemble from underdetermined macromolecular structural data. *Journal of the American Chemical Society*, 135(44), 16595–16609. <https://doi.org/10.1021/ja4083717>
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28(1), 235–242. <https://doi.org/10.1093/nar/28.1.235>
- Best, R. B., & Vendruscolo, M. (2004). Determination of protein structures consistent with NMR order parameters. *Journal of the American Chemical Society*, 126(26), 8090–8091. <https://doi.org/10.1021/ja0396955>
- Boehr, D. D., Nussinov, R., & Wright, P. E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. *Nature Chemical Biology*, 5(11), 789–796. <https://doi.org/10.1038/nchembio.232>
- Bonomi, M., Camilloni, C., Cavalli, A., & Vendruscolo, M. (2015). MetaInference: A bayesian inference method for heterogeneous systems. *Science Advances*, January, 22. <https://doi.org/10.1126/sciadv.1501177>
- Bonvin, A. M. J. J., Boelens, R., & Kaptein, R. (1994). Time- and ensemble-averaged direct NOE restraints. *Journal of Biomolecular NMR*, 4(1), 143–149. <https://doi.org/10.1007/BF00178343>
- Butterfield, D. A., Abdul, H. M., Opii, W., Newman, S. F., Joshi, G., Ansari, M. A., & Sultana, R. (2006). Pin1 in Alzheimer's disease. *Journal of Neurochemistry*, 98(6), 1697–1706. <https://doi.org/10.1111/j.1471-4159.2006.03995.x>
- Camilloni, C., Cavalli, A., & Vendruscolo, M. (2013). Assessment of the use of NMR chemical shifts as replica-averaged structural restraints in molecular dynamics simulations to characterize the dynamics of proteins. *Journal of Physical Chemistry B*, 117(6), 1838–1843. <https://doi.org/10.1021/jp3106666>
- Camilloni, C., Sahakyan, A. B., Holliday, M. J., Isern, N. G., Zhang, F., Eisenmesser, E. Z., &

- Vendruscolo, M. (2014). Cyclophilin A catalyzes proline isomerization by an electrostatic handle mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(28), 10203–10208. <https://doi.org/10.1073/pnas.1404220111>
- Camilloni, C., & Vendruscolo, M. (2015). A tensor-free method for the structural and dynamical refinement of proteins using residual dipolar couplings. *Journal of Physical Chemistry B*, *119*(3), 653–661. <https://doi.org/10.1021/jp5021824>
- Cavalli, A., Camilloni, C., & Vendruscolo, M. (2013). Molecular dynamics simulations with replica-averaged structural restraints generate structural ensembles according to the maximum entropy principle. *Journal of Chemical Physics*, *138*(9). <https://doi.org/10.1063/1.4793625>
- Cesari, A., Reißer, S., & Bussi, G. (2018). Using the maximum entropy principle to combine simulations and solution experiments. *Computation*, *6*(1), 1–26. <https://doi.org/10.3390/computation6010015>
- Chen, Y., Campbell, S. L., & Dokholyan, N. V. (2007). Deciphering Protein Dynamics from NMR Data Using Explicit Structure Sampling and Selection. *Biophysical Journal*, *93*(7), 2300–2306. <https://doi.org/10.1529/biophysj.107.104174>
- Cho, K. O., Hunt, C. A., & Kennedy, M. B. (1992). The rat brain postsynaptic density fraction contains a homolog of the drosophila discs-large tumor suppressor protein. *Neuron*, *9*(5), 929–942. [https://doi.org/10.1016/0896-6273\(92\)90245-9](https://doi.org/10.1016/0896-6273(92)90245-9)
- Cierpicki, T., Bushweller, J. H., & Derewenda, Z. S. (2005). Probing the supramodular architecture of a multidomain protein: The structure of syntenin in solution. *Structure*, *13*(2), 319–327. <https://doi.org/10.1016/j.str.2004.12.014>
- Coley, A. A., & Gao, W.-J. (2018). PSD95: A synaptic protein implicated in schizophrenia or autism? *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *82*, 187–194. <https://doi.org/10.1016/j.pnpbp.2017.11.016>
- Collins, M. O., Husi, H., Yu, L., Brandon, J. M., Anderson, C. N. G., Blackstock, W. P., Choudhary, J. S., & Grant, S. G. N. (2006). Molecular characterization and comparison of the components and multiprotein complexes in the postsynaptic proteome. *Journal of Neurochemistry*, *97*(m), 16–23. <https://doi.org/10.1111/j.1471-4159.2005.03507.x>
- Craven, & Brecht. (1998). PDZ proteins organize synaptic signaling pathways. *Cell*, *93*(4), 495–498. [https://doi.org/10.1016/s0092-8674\(00\)81179-4](https://doi.org/10.1016/s0092-8674(00)81179-4)
- Craveur, P., Joseph, A. P., Poulain, P., De Brevern, A. G., & Rebehmed, J. (2013). Cis-trans isomerization of omega dihedrals in proteins. *Amino Acids*, *45*(2), 279–289. <https://doi.org/10.1007/s00726-013-1511-3>
- Czajlik, A., Kovács, B., Permi, P., & Gáspári, Z. (2017). Fine-tuning the extent and dynamics of binding cleft opening as a potential general regulatory mechanism in parvulin-type peptidyl prolyl isomerases. *Scientific Reports*, *7*. <https://doi.org/10.1038/srep44504>
- Davis, T. L., Walker, J. R., Campagna-Slater, V., Finerty, P. J., Paramanathan, R., Bernstein, G., MacKenzie, F., Tempel, W., Ouyang, H., Lee, W. H., Eisenmesser, E. Z., Dhe-Paganon, S., & Dhe-Paganon, S. (2010). Structural and biochemical characterization of the human cyclophilin family of peptidyl-prolyl isomerases. *PLoS Biology*, *8*(7), e1000439. <https://doi.org/10.1371/journal.pbio.1000439>
- Denesyuk, A. I., Vihinen, M., Lundell, J., Zavyalov, V. P., & Korpela, T. (1993). Structural Similarity



- of the Binding Sites of Cyclophilin A-Cyclosporine A and FKBP-FK506 Systems. *Biochemical and Biophysical Research Communications*, 192(2), 912–917. <https://doi.org/10.1006/BBRC.1993.1502>
- Dhulesia, A., Gsponer, J., & Vendruscolo, M. (2008). Mapping of two networks of residues that exhibit structural and dynamical changes upon binding in a PDZ domain protein. *Journal of the American Chemical Society*, 130(28), 8931–8939. <https://doi.org/10.1021/ja0752080>
- Doyle, D. a., Lee, A., Lewis, J., Kim, E., Sheng, M., & MacKinnon, R. (1996). Crystal structures of a complexed and peptide-free membrane protein- binding domain: Molecular basis of peptide recognition by PDZ. *Cell*, 85(7), 1067–1076. [https://doi.org/10.1016/S0092-8674\(00\)81307-0](https://doi.org/10.1016/S0092-8674(00)81307-0)
- Dudola, D., Kovács, B., & Gáspári, Z. (2017). CoNSEnsX+ Webserver for the Analysis of Protein Structural Ensembles Reflecting Experimentally Determined Internal Dynamics. *Journal of Chemical Information and Modeling*, 57(8), 1728–1734. <https://doi.org/10.1021/acs.jcim.7b00066>
- Eberhardt, E. S., Loh, S. N., Hinck, A. P., & Raines, R. T. (1992). Solvent Effects on the Energetics of Prolyl Peptide Bond Isomerization. *Journal of the American Chemical Society*, 114(13), 5437–5439. <https://doi.org/10.1021/ja00039a072>
- Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E., & Kern, D. (2005). Intrinsic dynamics of an enzyme underlies catalysis. *Nature*, 438(7064), 117–121. <https://doi.org/10.1038/nature04105>
- Ernst, A., Appleton, B. a., Ivarsson, Y., Zhang, Y., Gfeller, D., Wiesmann, C., & Sidhu, S. S. (2014). A Structural Portrait of the PDZ Domain Family. *Journal of Molecular Biology*, 426(21), 3509–3519. <https://doi.org/10.1016/j.jmb.2014.08.012>
- Fanghänel, J., & Fischer, G. (2004). Insights into the catalytic mechanism of peptidyl prolyl cis/trans isomerases. *Frontiers in Bioscience : A Journal and Virtual Library*, 9, 3453–3478. <http://www.ncbi.nlm.nih.gov/pubmed/15353370>
- Fanning, A. S., & Anderson, J. M. (1999). PDZ domains: fundamental building blocks in the organization of protein complexes at the plasma membrane. *The Journal of Clinical Investigation*, 103(6), 767–772. <https://doi.org/10.1172/JCI6509>
- Feng, W., Shi, Y., Li, M., & Zhang, M. (2003). Tandem PDZ repeats in glutamate receptor-interacting proteins have a novel mode of PDZ domain-mediated target binding. *Nature Structural Biology*, 10(11), 972–978. <https://doi.org/10.1038/nsb992>
- Feng, W., & Zhang, M. (2009). Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nature Reviews. Neuroscience*, 10(2), 87–99. <https://doi.org/10.1038/nrn2540>
- Feng, Z., Chen, X., Zeng, M., & Zhang, M. (2019). Phase separation as a mechanism for assembling dynamic postsynaptic density signalling complexes. *Current Opinion in Neurobiology*, 57, 1–8. <https://doi.org/10.1016/j.conb.2018.12.001>
- Fischer, G., & Aumüller, T. (2003). Regulation of peptide bond cis/trans isomerization by enzyme catalysis and its implication in physiological processes. *Reviews of Physiology, Biochemistry and Pharmacology*, 105–150. <https://doi.org/10.1007/s10254-003-0011-3>
- Fischer, G, Bang, H., & Mech, C. (1984). [Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides]. *Biomedica Biochimica Acta*,

43(10), 1101–1111. <http://www.ncbi.nlm.nih.gov/pubmed/6395866>

Fischer, Gunter, Tradler, T., & Zarnt, T. (1998). The mode of action of peptidyl prolyl cis/trans isomerases in vivo: binding vs. catalysis. *FEBS Letters*, 426(1), 17–20. [https://doi.org/10.1016/S0014-5793\(98\)00242-7](https://doi.org/10.1016/S0014-5793(98)00242-7)

Fischer, S., Michnick, S., & Karplus, M. (1993). A Mechanism for Rotamase Catalysis by the FK506 Binding Protein (FKBP). *Biochemistry*, 32(50), 13830–13837. <https://doi.org/10.1021/bi00213a011>

Fizil, Á., Gáspári, Z., Barna, T., Marx, F., & Batta, G. (2015). “Invisible” conformers of an antifungal disulfide protein revealed by constrained cold and heat unfolding, CEST-NMR experiments, and molecular dynamics calculations. *Chemistry - A European Journal*, 21(13), 5136–5144. <https://doi.org/10.1002/chem.201404879>

Fuentes, E. J., Der, C. J., & Lee, A. L. (2004). Ligand-dependent Dynamics and Intramolecular Signaling in a PDZ Domain. *Journal of Molecular Biology*, 335(4), 1105–1115. <https://doi.org/10.1016/j.jmb.2003.11.010>

Fuentes, E. J., Gilmore, S. a., Mauldin, R. V., & Lee, A. L. (2006). Evaluation of Energetic and Dynamic Coupling Networks in a PDZ Domain Protein. *Journal of Molecular Biology*, 364(3), 337–351. <https://doi.org/10.1016/j.jmb.2006.08.076>

Gardoni, F., Marcello, E., & Di Luca, M. (2009). Postsynaptic density-membrane associated guanylate kinase proteins (PSD-MAGUKs) and their role in CNS disorders. *Neuroscience*, 158(1), 324–333. <https://doi.org/10.1016/j.neuroscience.2008.07.068>

Gáspári, Z., & Perczel, A. (2010). Protein Dynamics as Reported by NMR. *Annual Reports on NMR Spectroscopy*, 71(C), 35–75. <https://doi.org/10.1016/B978-0-08-089054-8.00002-2>

Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. In *Expert Opinion on Drug Discovery* (Vol. 10, Issue 5, pp. 449–461). Informa Healthcare. <https://doi.org/10.1517/17460441.2015.1032936>

Ghosh, A., Saminathan, H., Kanthasamy, A., Anantharam, V., Jin, H., Sondarva, G., Harischandra, D. S., Qian, Z., Rana, A., & Kanthasamy, A. G. (2013). The Peptidyl-prolyl Isomerase Pin1 Up-regulation and Proapoptotic Function in Dopaminergic Neurons. *Journal of Biological Chemistry*, 288(30), 21955–21971. <https://doi.org/10.1074/jbc.M112.444224>

Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., & Ferrin, T. E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science*, 27(1), 14–25. <https://doi.org/10.1002/pro.3235>

Göthel, S. F., & Marahiel, M. A. (1999). Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. In *Cellular and Molecular Life Sciences* (Vol. 55, Issue 3, pp. 423–436). <https://doi.org/10.1007/s000180050299>

Guo, J., Pang, X., & Zhou, H.-X. X. (2015). Two pathways mediate interdomain allosteric regulation in Pin1. *Structure*, 23(1), 237–247. <https://doi.org/10.1016/j.str.2014.11.009>

Han, B., Liu, Y., Ginzinger, S. W., & Wishart, D. S. (2011). SHIFTX2: Significantly improved protein chemical shift prediction. *Journal of Biomolecular NMR*, 50(1), 43–57. <https://doi.org/10.1007/s10858-011-9478-4>

Hanes, S. D. (2015). Prolyl isomerases in gene transcription. *Biochimica et Biophysica Acta*,

1850(10), 2017–2034. <https://doi.org/10.1016/j.bbagen.2014.10.028>

- Harris, B. Z., & Lim, W. a. (2001). Mechanism and role of PDZ domains in signaling complex assembly. *Journal of Cell Science*, 114(Pt 18), 3219–3231. <https://doi.org/10.1083/JCB.142.2.545>
- Heikkinen, O., Seppala, R., Tossavainen, H., Heikkinen, S., Koskela, H., Permi, P., & Kilpeläinen, I. (2009). Solution structure of the parvulin-type PPIase domain of *Staphylococcus aureus* PrsA-- implications for the catalytic mechanism of parvulins. *BMC Structural Biology*, 9(1), 17. <https://doi.org/10.1186/1472-6807-9-17>
- Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A Linear Constraint Solver for molecular simulations. *Journal of Computational Chemistry*, 18(12), 1463–1472. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H)
- Hess, B., Kutzner, C., Van Der Spoel, D., & Lindahl, E. (2008). GRGMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4(3), 435–447. <https://doi.org/10.1021/ct700301q>
- Hummer, G., & Köfinger, J. (2015). Bayesian ensemble refinement by replica simulations and reweighting. *Journal of Chemical Physics*, 143(24). <https://doi.org/10.1063/1.4937786>
- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33–38, 27–28. <http://www.ncbi.nlm.nih.gov/pubmed/8744570>
- Hung, A. Y., & Sheng, M. (2002). PDZ domains: Structural modules for protein complex assembly. *Journal of Biological Chemistry*, 277(8), 5699–5702. <https://doi.org/10.1074/jbc.R100065200>
- Imamura, F., Maeda, S., Doi, T., & Fujiyoshi, Y. (2002). Ligand Binding of the Second PDZ Domain Regulates Clustering of PSD-95 with the Kv1.4 Potassium Channel. *Journal of Biological Chemistry*, 277(5), 3640–3646. <https://doi.org/10.1074/jbc.M106940200>
- Jakob, R. P., Zoldak, G., Aumuller, T., & Schmid, F. X. (2009). Chaperone domains convert prolyl isomerases into generic catalysts of protein folding. *Proceedings of the National Academy of Sciences*, 106(48), 20282–20287. <https://doi.org/10.1073/pnas.0909544106>
- Jaremko, Ł., Jaremko, M., Elfaki, I., Mueller, J. W., Ejchart, A., Bayer, P., & Zhukov, I. (2011). Structure and dynamics of the first archaeal parvulin reveal a new functionally important loop in parvulin-type prolyl isomerases. *Journal of Biological Chemistry*, 286(8), 6554–6565. <https://doi.org/10.1074/jbc.M110.160713>
- Jaynes, E. T. (1957a). Information theory and statistical mechanics. *Physical Review*, 106(4), 620–630. <https://doi.org/10.1103/PhysRev.106.620>
- Jaynes, E. T. (1957b). Information theory and statistical mechanics. II. *Physical Review*, 108(2), 171–190. <https://doi.org/10.1103/PhysRev.108.171>
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2), 926–935. <https://doi.org/10.1063/1.445869>
- Kabsch, W. (1976a). A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, 34(5), 922–923. <https://doi.org/10.1107/S0567739478001680>
- Kabsch, W. (1976b). A solution for the best rotation to relate two sets of vectors. *Acta*

- Crystallographica Section A*, 32(5), 922–923. <https://doi.org/10.1107/S0567739476001873>
- Kang, B. S., Cooper, D. R., Jelen, F., Devedjiev, Y., Derewenda, U., Dauter, Z., Otlewski, J., & Derewenda, Z. S. (2003). PDZ tandem of human syntenin: Crystal structure and functional properties. *Structure*, 11(4), 459–468. [https://doi.org/10.1016/S0969-2126\(03\)00052-2](https://doi.org/10.1016/S0969-2126(03)00052-2)
- Kay, L. E. (1998). Protein dynamics from NMR. *Biochemistry and Cell Biology = Biochimie et Biologie Cellulaire*, 76(2–3), 145–152. <https://doi.org/10.1139/o98-024>
- Ke, H. (1992). Similarities and differences between human cyclophilin A and other beta-barrel structures. Structural refinement at 1.63 Å resolution. *Journal of Molecular Biology*, 228(2), 539–550. [https://doi.org/10.1016/0022-2836\(92\)90841-7](https://doi.org/10.1016/0022-2836(92)90841-7)
- Kennedy, M. B. (2000). Signal-processing machines at the postsynaptic density. *Science (New York, N.Y.)*, 290(5492), 750–754. <https://doi.org/10.1126/science.290.5492.750>
- Kennedy, Mary B. (1997). The postsynaptic density at glutamatergic synapses. *Trends in Neurosciences*, 20(6), 264–268. [https://doi.org/10.1016/s0166-2236\(96\)01033-8](https://doi.org/10.1016/s0166-2236(96)01033-8)
- Kim, E., & Sheng, M. (2004). PDZ domain proteins of synapses. *Nature Reviews Neuroscience*, 5(10), 771–781. <https://doi.org/10.1038/nrn1517>
- Kim, E., & Sheng, M. (2009). The postsynaptic density. *Current Biology : CB*, 19(17), R723-4. <https://doi.org/10.1016/j.cub.2009.07.047>
- Kitao, A., & Go, N. (1999). Investigating protein dynamics in collective coordinate space. *Current Opinion in Structural Biology*, 9(2), 164–169. [https://doi.org/10.1016/S0959-440X\(99\)80023-2](https://doi.org/10.1016/S0959-440X(99)80023-2)
- Kofron, J. L., Kuzmic, P., Kishore, V., Colon-Bonilla, E., & Rich, D. H. (1991). Determination of kinetic constants for peptidyl prolyl cis-trans isomerases by an improved spectrophotometric assay. *Biochemistry*, 30(25), 6127–6134. <https://doi.org/10.1021/bi00239a007>
- Kovács, B., Zajác-Epresi, N., & Gáspári, Z. (2020). Ligand-dependent intra- and interdomain motions in the PDZ12 tandem regulate binding interfaces in postsynaptic density protein-95. *FEBS Letters*, 594(5), 887–902. <https://doi.org/10.1002/1873-3468.13626>
- Kozlov, G., Banville, D., Gehring, K., & Ekiel, I. (2002). Solution structure of the PDZ2 domain from cytosolic human phosphatase hPTP1E complexed with a peptide reveals contribution of the  $\beta$ 2- $\beta$ 3 loop to PDZ domain-ligand interactions. *Journal of Molecular Biology*, 320(4), 813–820. [https://doi.org/10.1016/S0022-2836\(02\)00544-2](https://doi.org/10.1016/S0022-2836(02)00544-2)
- Kwan, J., Sczaniecka, A., Heidary Arash, E., Nguyen, L., Chen, C.-C., Ratkovic, S., Klezovitch, O., Attisano, L., McNeill, H., Emili, A., & Vasioukhin, V. (2016). DLG5 connects cell polarity and Hippo signaling protein networks by linking PAR-1 with MST1/2. *Genes & Development*, 30(24), 2696–2709. <https://doi.org/10.1101/gad.284539.116>
- Labeikovsky, W., Eisenmesser, E. Z., Bosco, D. A., & Kern, D. (2007). Structure and Dynamics of Pin1 During Catalysis by NMR. *Journal of Molecular Biology*, 367(5), 1370–1381. <https://doi.org/10.1016/j.jmb.2007.01.049>
- Lang, K., Schmid, F. X., & Fischer, G. (1987). Catalysis of protein folding by prolyl isomerase. *Nature*, 329(6136), 268–270. <https://doi.org/10.1038/329268a0>
- Lange, O. F., Lakomek, N. A., Farès, C., Schröder, G. F., Walter, K. F. A., Becker, S., Meiler, J., Grubmüller, H., Griesinger, C., & De Groot, B. L. (2008). Recognition dynamics up to

- microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science*, 320(5882), 1471–1475. <https://doi.org/10.1126/science.1157092>
- Lau, C. G., & Zukin, R. S. (2007). NMDA receptor trafficking in synaptic plasticity and neuropsychiatric disorders. *Nature Reviews Neuroscience*, 8(6), 413–426. <https://doi.org/10.1038/nrn2153>
- Lee, H.-J., & Zheng, J. J. (2010). PDZ domains and their binding partners: structure, specificity, and modification. *Cell Communication and Signaling : CCS*, 8, 8. <https://doi.org/10.1186/1478-811X-8-8>
- Lee, T. H., Pastorino, L., & Lu, K. P. (2011). Peptidyl-prolyl cis-trans isomerase Pin1 in ageing, cancer and Alzheimer disease. *Expert Reviews in Molecular Medicine*, 13(June), 1–26. <https://doi.org/10.1017/S1462399411001906>
- Letunic, I., & Bork, P. (2018). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research*, 46(D1), D493–D496. <https://doi.org/10.1093/nar/gkx922>
- Leung, H. T. A., Bignucolo, O., Aregger, R., Dames, S. A., Mazur, A., Bernèche, S., & Grzesiek, S. (2016). A Rigorous and Efficient Method to Reweight Very Large Conformational Ensembles Using Average Experimental Data and to Determine Their Relative Information Content. *Journal of Chemical Theory and Computation*, 12(1), 383–394. <https://doi.org/10.1021/acs.jctc.5b00759>
- Lindorff-Larsen, K., Best, R. B., DePristo, M. A., Dobson, C. M., & Vendruscolo, M. (2005). Simultaneous determination of protein structure and dynamics. *Nature*, 433(7022), 128–132. <https://doi.org/10.1038/nature03199>
- Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., & Shaw, D. E. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function and Bioinformatics*, 78(8), 1950–1958. <https://doi.org/10.1002/prot.22711>
- Lipari, G., & Szabo, A. (1982). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *Journal of the American Chemical Society*, 104(17), 4546–4559. <https://doi.org/10.1021/ja00381a009>
- Liu, W., Wen, W., Wei, Z., Yu, J., Ye, F., Liu, C. H., Hardie, R. C., & Zhang, M. (2011). The INAD scaffold is a dynamic, redox-regulated modulator of signaling in the *Drosophila* eye. *Cell*, 145(7), 1088–1101. <https://doi.org/10.1016/j.cell.2011.05.015>
- Long, J.-F., Feng, W., Wang, R., Chan, L.-N., Ip, F. C. F., Xia, J., Ip, N. Y., & Zhang, M. (2005). Autoinhibition of X11/Mint scaffold proteins revealed by the closed conformation of the PDZ tandem. *Nature Structural & Molecular Biology*, 12(8), 722–728. <https://doi.org/10.1038/nsmb958>
- Long, J. F., Tochio, H., Wang, P., Fan, J. S., Sala, C., Niethammer, M., Sheng, M., & Zhang, M. (2003). Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95. *Journal of Molecular Biology*, 327(1), 203–214. [https://doi.org/10.1016/S0022-2836\(03\)00113-X](https://doi.org/10.1016/S0022-2836(03)00113-X)
- Long, J., Wei, Z., Feng, W., Yu, C., Zhao, Y. X., & Zhang, M. (2008). Supramodular Nature of GRIP1 Revealed by the Structure of Its PDZ12 Tandem in Complex with the Carboxyl Tail of Fras1. *Journal of Molecular Biology*, 375(5), 1457–1468. <https://doi.org/10.1016/j.jmb.2007.11.088>

- Lu, K. P., Finn, G., Lee, T. H., & Nicholson, L. K. (2007). Prolyl cis-trans isomerization as a molecular timer. *Nature Chemical Biology*, 3(10), 619–629.  
<https://doi.org/10.1038/nchembio.2007.35>
- Lu, K. P., Suizu, F., Zhou, X. Z., Finn, G., Lam, P., & Wulf, G. (2006). Targeting carcinogenesis: A role for the prolyl isomerase Pin1? *Molecular Carcinogenesis*, 45(6), 397–402.  
<https://doi.org/10.1002/mc.20216>
- Lu, M., Zhang, L., Sartippour, M. R., Norris, A. J., & Brooks, M. N. (2006). EG-1 interacts with c-Src and activates its signaling pathway. *International Journal of Oncology*, 29(4), 1013–1018.
- Lu, P. J., Zhou, X. Z., Shen, M., & Lu, K. P. (1999). Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science*, 283(5406), 1325–1328.  
<https://doi.org/10.1126/science.283.5406.1325>
- Lu, Z., & Hunter, T. (2014). Prolyl isomerase Pin1 in cancer. *Cell Research*, 24(9), 1–17.  
<https://doi.org/10.1038/cr.2014.109>
- Luck, K., Charbonnier, S., & Travé, G. (2012). The emerging contribution of sequence context to the specificity of protein interactions mediated by PDZ domains. *FEBS Letters*, 586(17), 2648–2661. <https://doi.org/10.1016/j.febslet.2012.03.056>
- Lupyan, D., Leo-Macias, A., & Ortiz, A. R. (2005). A new progressive-iterative algorithm for multiple structure alignment. *Bioinformatics*, 21(15), 3255–3263.  
<https://doi.org/10.1093/bioinformatics/bti527>
- Markwick, P. R. L., Malliavin, T., & Nilges, M. (2008). Structural Biology by NMR: Structure, Dynamics, and Interactions. *PLoS Computational Biology*, 4(9), e1000168.  
<https://doi.org/10.1371/journal.pcbi.1000168>
- Matena, A., Rehic, E., Hönig, D., Kamba, B., & Bayer, P. (2018). Structure and function of the human parvulins Pin1 and Par14/17. *Biological Chemistry*, 399(2), 101–125.  
<https://doi.org/10.1515/hsz-2017-0137>
- McCann, J. J. J., Zheng, L., Chiantia, S., & Bowen, M. E. E. (2011). Domain Orientation in the N-Terminal PDZ Tandem from PSD-95 Is Maintained in the Full-Length Protein. *Structure*, 19(6), 810–820. <https://doi.org/10.1016/j.str.2011.02.017>
- McCann, J. J., Zheng, L., Rohrbeck, D., Felekyan, S., Kuhnemuth, R., Sutton, R. B., Seidel, C. a. M., & Bowen, M. E. (2012). Supertertiary structure of the synaptic MAGuK scaffold proteins is conserved. *Proceedings of the National Academy of Sciences*, 109(39), 15775–15780.  
<https://doi.org/10.1073/pnas.1200254109>
- Mercedes-Camacho, A. Y., Mullins, A. B., Mason, M. D., Xu, G. G., Mahoney, B. J., Wang, X., Peng, J. W., & Eitzkorn, F. A. (2013). Kinetic isotope effects support the twisted amide mechanism of Pin1 peptidyl-prolyl isomerase. *Biochemistry*, 52(44), 7707–7713.  
<https://doi.org/10.1021/bi400700b>
- Michael P. Allen. (2004). Introduction to Molecular Dynamics Simulation. In N. Attig, K. Binder, H. Grubmüller, & K. Kremer (Eds.), *Computational Soft Matter: From Synthetic Polymers to Proteins, Lecture Notes* (pp. 1–28). John von Neumann Institute for Computing.
- Migaud, M., Charlesworth, P., Dempster, M., Webster, L. C., Watabe, A. M., Makhinson, M., He, Y., Ramsay, M. F., Morris, R. G., Morrison, J. H., O'Dell, T. J., & Grant, S. G. (1998). Enhanced long-term potentiation and impaired learning in mice with mutant postsynaptic density-95

- protein. *Nature*, 396(6710), 433–439. <https://doi.org/10.1038/24790>
- Morais Cabral, J. H., Petosa, C., Sutcliffe, M. J., Raza, S., Byron, O., Poy, F., Marfatia, S. M., Chishti, A. H., & Liddington, R. C. (1996). Crystal structure of a PDZ domain. *Nature*, 382(6592), 649–652. <https://doi.org/10.1038/382649a0>
- Mostarda, S., Gfeller, D., & Rao, F. (2012). Beyond the binding site: The role of the  $\beta 2$  -  $\beta 3$  loop and extra-domain structures in PDZ domains. *PLoS Computational Biology*, 8(3). <https://doi.org/10.1371/journal.pcbi.1002429>
- Mueller, J. W., Link, N. M., Matena, A., Hoppstock, L., Ruppel, A., Bayer, P., & Blankenfeldt, W. (2011). Crystallographic proof for an extended hydrogen-bonding network in small prolyl isomerases. *Journal of the American Chemical Society*, 133(50), 20096–20099. <https://doi.org/10.1021/ja2086195>
- Namanja, A. T., Peng, T., Zintsmaster, J. S., Elson, A. C., Shakour, M. G., & Peng, J. W. (2007). Substrate Recognition Reduces Side-Chain Flexibility for Conserved Hydrophobic Residues in Human Pin1. *Structure*, 15(3), 313–327. <https://doi.org/10.1016/j.str.2007.01.014>
- Namanja, A. T., Wang, X. J., Xu, B., Mercedes-Camacho, A. Y., Wilson, K. A., Etzkorn, F. A., & Peng, J. W. (2011). Stereospecific gating of functional motions in Pin1. *Proceedings of the National Academy of Sciences of the United States of America*, 108(30), 12289–12294. <https://doi.org/10.1073/pnas.1019382108>
- Nath, P. R., & Isakov, N. (2015). Insights into peptidyl-prolyl cis-trans isomerase structure and function in immunocytes. In *Immunology Letters* (Vol. 163, Issue 1, pp. 120–131). Elsevier B.V. <https://doi.org/10.1016/j.imlet.2014.11.002>
- Neal, S., Nip, A. M., Zhang, H., & Wishart, D. S. (2003). Rapid and accurate calculation of protein 1 H, 13 C and 15 N chemical shifts. 215–240.
- Neff, R. A., Gomez-Varela, D., Fernandes, C. C., & Berg, D. K. (2009). Postsynaptic scaffolds for nicotinic receptors on neurons. *Acta Pharmacologica Sinica*, 30(6), 694–701. <https://doi.org/10.1038/aps.2009.52>
- Nguyen, K., Iskandar, M., & Rabenstein, D. L. (2010). Kinetics and equilibria of cis/trans isomerization of secondary amide peptide bonds in linear and cyclic peptides. *Journal of Physical Chemistry B*, 114(9), 3387–3392. <https://doi.org/10.1021/jp1000286>
- Nourry, C., Grant, S. G. N., & Borg, J. (2003). PDZ domain proteins: plug and play! *Science's STKE : Signal Transduction Knowledge Environment*, 2003(179), RE7. <https://doi.org/10.1126/stke.2003.179.re7>
- Olsson, S., Strotz, D., Vögeli, B., Riek, R., & Cavalli, A. (2016). The Dynamic Basis for Signal Propagation in Human Pin1-WW. *Structure (London, England : 1993)*, 24(9), 1464–1475. <https://doi.org/10.1016/j.str.2016.06.013>
- Pal, D., & Chakrabarti, P. (1999). Cis peptide bonds in proteins: Residues involved, their conformations, interactions and locations. *Journal of Molecular Biology*, 294(1), 271–288. <https://doi.org/10.1006/jmbi.1999.3217>
- Parrinello, M., & Rahman, A. (1980). Crystal structure and pair potentials: A molecular-dynamics study. *Physical Review Letters*, 45(14), 1196–1199. <https://doi.org/10.1103/PhysRevLett.45.1196>

- Peng, J. W. (2012). Exposing the moving parts of proteins with NMR spectroscopy. *Journal of Physical Chemistry Letters*, 3(8), 1039–1051. <https://doi.org/10.1021/jz3002103>
- Peng, Jeffrey W., & Wagner, G. (1992). Mapping of spectral density functions using heteronuclear NMR relaxation measurements. *Journal of Magnetic Resonance (1969)*, 98(2), 308–332. [https://doi.org/10.1016/0022-2364\(92\)90135-T](https://doi.org/10.1016/0022-2364(92)90135-T)
- Petit, C. M., Zhang, J., Sapienza, P. J., Fuentes, E. J., & Lee, A. L. (2009). Hidden dynamic allostery in a PDZ domain. *Proceedings of the National Academy of Sciences of the United States of America*, 106(43), 18249–18254. <https://doi.org/10.1073/pnas.0904492106>
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612. <https://doi.org/10.1002/jcc.20084>
- Quistgaard, E. M., Weininger, U., Ural-Blimke, Y., Modig, K., Nordlund, P., Akke, M., & Löw, C. (2016). Molecular insights into substrate recognition and catalytic mechanism of the chaperone and FKBP peptidyl-prolyl isomerase SlyD. *BMC Biology*, 14(1), 82. <https://doi.org/10.1186/s12915-016-0300-3>
- Rahfeld, J. U., Rücknagel, K. P., Schelbert, B., Ludwig, B., Hacker, J., Mann, K., & Fischer, G. (1994). Confirmation of the existence of a third family among peptidyl-prolyl cis/trans isomerases Amino acid sequence and recombinant production of parvulin. *FEBS Letters*, 352(2), 180–184. [https://doi.org/10.1016/0014-5793\(94\)00932-5](https://doi.org/10.1016/0014-5793(94)00932-5)
- Rahfeld, J. U., Schierhorn, A., Mann, K., & Fischer, G. (1994). A novel peptidyl-prolyl cis/trans isomerase from Escherichia coli. *FEBS Letters*, 343(1), 65–69. [https://doi.org/10.1016/0014-5793\(94\)80608-X](https://doi.org/10.1016/0014-5793(94)80608-X)
- Ranganathan, R., Lu, K. P., Hunter, T., & Noel, J. P. (1997). Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell*, 89(6), 875–886. [https://doi.org/10.1016/S0092-8674\(00\)80273-1](https://doi.org/10.1016/S0092-8674(00)80273-1)
- Rapaport, D. C. (2004). The Art of Molecular Dynamics Simulation. In *The Art of Molecular Dynamics Simulation* (pp. 1–10). Cambridge University Press. <https://doi.org/10.1017/cbo9780511816581>
- Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., & Fischer, G. (1998). Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *Journal of Molecular Biology*, 279(2), 449–460. <https://doi.org/10.1006/jmbi.1998.1770>
- Richter, B., Gsponer, J., Várnai, P., Salvatella, X., & Vendruscolo, M. (2007). The MUMO (minimal under-restraining minimal over-restraining) method for the determination of native state ensembles of proteins. *Journal of Biomolecular NMR*, 37(2), 117–135. <https://doi.org/10.1007/s10858-006-9117-7>
- Roberts, S., Delury, C., & Marsh, E. (2012). The PDZ protein discs-large (DLG): The “Jekyll and Hyde” of the epithelial polarity proteins. *FEBS Journal*, 279(19), 3549–3558. <https://doi.org/10.1111/j.1742-4658.2012.08729.x>
- Robustelli, P., Kohlhoff, K., Cavalli, A., & Vendruscolo, M. (2010). Using NMR chemical shifts as structural restraints in molecular dynamics simulations of proteins. *Structure*, 18(8), 923–933. <https://doi.org/10.1016/j.str.2010.04.016>



- Rosen, M. K., Standaert, R. F., Galat, A., Nakatsuka, M., & Schreiber, S. L. (1990). Inhibition of FKBP rotamase activity by immunosuppressant FK506: Twisted amide surrogate. *Science*, 248(4957), 863–866. <https://doi.org/10.1126/science.1693013>
- Rudrabhatla, P., Zheng, Y.-L., Amin, N. D., Kesavapany, S., Albers, W., & Pant, H. C. (2008). Pin1-dependent prolyl isomerization modulates the stress-induced phosphorylation of high molecular weight neurofilament protein. *The Journal of Biological Chemistry*, 283(39), 26737–26747. <https://doi.org/10.1074/jbc.M801633200>
- Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3), 327–341. [https://doi.org/10.1016/0021-9991\(77\)90098-5](https://doi.org/10.1016/0021-9991(77)90098-5)
- Ryo, A., Togo, T., Nakai, T., Hirai, A., Nishi, M., Yamaguchi, A., Suzuki, K., Hirayasu, Y., Kobayashi, H., Perrem, K., Liou, Y.-C., & Aoki, I. (2006). Prolyl-isomerase Pin1 accumulates in lewy bodies of parkinson disease and facilitates formation of alpha-synuclein inclusions. *The Journal of Biological Chemistry*, 281(7), 4117–4125. <https://doi.org/10.1074/jbc.M507026200>
- Sainlos, M., Tigaret, C., Poujol, C., Olivier, N. B., Bard, L., Breillat, C., Thiolon, K., Choquet, D., & Imperiali, B. (2010). Biomimetic divalent ligands for the acute disruption of synaptic AMPAR stabilization. *Nature Chemical Biology*, 7(2), 81. <https://doi.org/10.1038/nchembio.498>
- Sanchez-Martinez, M., & Crehuet, R. (2014). Application of the maximum entropy principle to determine ensembles of intrinsically disordered proteins from residual dipolar couplings. *Physical Chemistry Chemical Physics*, 16(47), 26030–26039. <https://doi.org/10.1039/c4cp03114h>
- Schiene-Fischer, C., & Yu, C. (2001). Receptor accessory folding helper enzymes: The functional role of peptidyl prolyl cis/trans isomerases. *FEBS Letters*, 495(1–2), 1–6. [https://doi.org/10.1016/S0014-5793\(01\)02326-2](https://doi.org/10.1016/S0014-5793(01)02326-2)
- Schmid, F. X., Grafl, R., Wrba, A., & Beintema, J. J. (1986). Role of proline peptide bond isomerization in unfolding and refolding of ribonuclease. *Proceedings of the National Academy of Sciences of the United States of America*, 83(4), 872–876. <https://doi.org/10.1073/pnas.83.4.872>
- Sekerina, E., Rahfeld, J. U., Müller, J., Fanghänel, J., Rascher, C., Fischer, G., & Bayer, P. (2000). NMR solution structure of hPar14 reveals similarity to the peptidyl prolyl cis/trans isomerase domain of the mitotic regulator hPin1 but indicates a different functionality of the protein. *Journal of Molecular Biology*, 301(4), 1003–1017. <https://doi.org/10.1006/JMBI.2000.4013>
- Shaw, P. E. (2002). Peptidyl-prolyl isomerases: A new twist to transcription. *EMBO Reports*, 3(6), 521–526. <https://doi.org/10.1093/embo-reports/kvf118>
- Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J. M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K. K., Lemak, A., Ignatchenko, A., Arrowsmith, C. H., Szyperski, T., Montelione, G. T., Baker, D., & Bax, A. (2008). Consistent blind protein structure generation from NMR chemical shift data. *Proceedings of the National Academy of Sciences of the United States of America*, 105(12), 4685–4690. <https://doi.org/10.1073/pnas.0800256105>
- Sheng, M. (2001). The postsynaptic NMDA-receptor-PSD-95 signaling complex in excitatory synapses of the brain. *Journal of Cell Science*, 114(7), 1251. [www.biologists.com/jcs](http://www.biologists.com/jcs)
- Sheng, M., & Hoogenraad, C. C. (2007). The Postsynaptic Architecture of Excitatory Synapses: A More Quantitative View. *Annual Review of Biochemistry*, 76(1), 823–847.

<https://doi.org/10.1146/annurev.biochem.76.060805.160029>

- Sheng, M., & Kim, E. (2011). The Postsynaptic Organization of Synapses. *Cold Spring Harbor Perspectives in Biology*, 3(12), a005678–a005678. <https://doi.org/10.1101/cshperspect.a005678>
- Sheng, M., & Sala, C. (2001). PDZ Domains and the Organization of Supramolecular Complexes. *Annual Review of Neuroscience*, 24(1), 1–29. <https://doi.org/10.1146/annurev.neuro.24.1.1>
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1), 539. <https://doi.org/10.1038/msb.2011.75>
- Somarelli, J. A., Lee, S. Y., Skolnick, J., & Herrera, R. J. (2008). Structure-based classification of 45 FK506-binding proteins. *Proteins*, 72(1), 197–208. <https://doi.org/10.1002/prot.21908>
- Søndergaard, C. R., Olsson, M. H. M., Rostkowski, M., & Jensen, J. H. (2011). Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pK<sub>a</sub> values. *Journal of Chemical Theory and Computation*, 7(7), 2284–2295. <https://doi.org/10.1021/ct200133y>
- Sun, L., Wu, X., Peng, Y., Goh, J. Y., Liou, Y.-C. C., Lin, D., & Zhao, Y. (2012). Solution structural analysis of the single-domain parvulin TbPin1. *PLoS ONE*, 7(8), e43017. <https://doi.org/10.1371/journal.pone.0043017>
- Terada, T., Shirouzu, M., Fukumori, Y., Fujimori, F., Ito, Y., Kigawa, T., Yokoyama, S., & Uchida, T. (2001). Solution structure of the human parvulin-like peptidyl prolyl cis/trans isomerase, hPar14. *Journal of Molecular Biology*, 305(4), 917–926. <https://doi.org/10.1006/jmbi.2000.4293>
- Theuerkorn, M., Fischer, G., & Schiene-Fischer, C. (2011). Prolyl cis/trans isomerase signalling pathways in cancer. In *Current Opinion in Pharmacology* (Vol. 11, Issue 4, pp. 281–287). Elsevier Ltd. <https://doi.org/10.1016/j.coph.2011.03.007>
- Tochio, H., Hung, F., Li, M., Brecht, D. S., & Zhang, M. (2000). Solution structure and backbone dynamics of the second PDZ domain of postsynaptic density-95. *Journal of Molecular Biology*, 295(2), 225–237. <https://doi.org/10.1006/jmbi.1999.3350>
- Tompa, P. (2012a). On the supertertiary structure of proteins. *Nature Chemical Biology*, 8(7), 597–600. <https://doi.org/10.1038/nchembio.1009>
- Tompa, P. (2012b). Intrinsically disordered proteins: A 10-year recap. In *Trends in Biochemical Sciences* (Vol. 37, Issue 12, pp. 509–516). <https://doi.org/10.1016/j.tibs.2012.08.004>
- Tonikian, R., Zhang, Y., Sazinsky, S. L., Currell, B., Yeh, J. H., Reva, B., Held, H. A., Appleton, B. A., Evangelista, M., Wu, Y., Xin, X., Chan, A. C., Seshagiri, S., Lasky, L. A., Sander, C., Boone, C., Bader, G. D., & Sidhu, S. S. (2008). A specificity map for the PDZ domain family. *PLoS Biology*, 6(9), 2043–2059. <https://doi.org/10.1371/journal.pbio.0060239>
- Torda, A. E., Scheek, R. M., & van Gunsteren, W. F. (1990). Time-averaged Nuclear Overhauser Effect Distance Restraints Applied to Tendimistat. *J.Mol.Biol.*, 214, 223–235.
- Tork Ladani, S., Souffrant, M. G., Barman, A., & Hamelberg, D. (2015). Computational perspective and evaluation of plausible catalytic mechanisms of peptidyl-prolyl cis–trans isomerases. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1850(10), 1994–2004. <https://doi.org/10.1016/j.bbagen.2014.12.023>

- Trzesniak, D., & van Gunsteren, W. F. (2006). Catalytic mechanism of cyclophilin as observed in molecular dynamics simulations: Pathway prediction and reconciliation of X-ray crystallographic and NMR solution data. *Protein Science*, *15*(11), 2544–2551. <https://doi.org/10.1110/ps.062356406>
- Wang, J., Kawasaki, R., Uewaki, J., Rashid, A., Tochio, N., & Tate, S. (2017). Dynamic Allostery Modulates Catalytic Activity by Modifying the Hydrogen Bonding Network in the Catalytic Site of Human Pin1. *Molecules*, *22*(6), 992. <https://doi.org/10.3390/molecules22060992>
- Wang, N. X., Lee, H.-J., & Zheng, J. J. (2008). Therapeutic use of PDZ protein-protein interaction antagonism. *Drug News & Perspectives*, *21*(3), 137–141. <http://www.ncbi.nlm.nih.gov/pubmed/18560611>
- Wang, W., Weng, J., Zhang, X., Liu, M., & Zhang, M. (2009). Creating conformational entropy by increasing interdomain mobility in ligand binding regulation: a revisit to N-terminal tandem PDZ domains of PSD-95. *Journal of the American Chemical Society*, *131*(2), 787–796. <https://doi.org/10.1021/ja8076022>
- Wang, X., Mahoney, B. J., Zhang, M., Zintsmaster, J. S., & Peng, J. W. (2015). Negative Regulation of Peptidyl-Prolyl Isomerase Activity by Interdomain Contact in Human Pin1. *Structure (London, England : 1993)*, *23*(12), 2224–2233. <https://doi.org/10.1016/j.str.2015.08.019>
- Wedemeyer, W. J., Welker, E., & Scheraga, H. A. (2002). Proline cis-trans isomerization and protein folding. *Biochemistry*, *41*(50), 14637–14644. <https://doi.org/10.1021/bi020574b>
- Weiss, M. S., Jabs, A., & Hilgenfeld, R. (1998). Peptide bonds revisited. *Nature Structural and Molecular Biology*, *5*(8), 676. <https://doi.org/10.1038/1368>
- Wilson, K. a., Bouchard, J. J., & Peng, J. W. (2013). Interdomain interactions support interdomain communication in human pin1. *Biochemistry*, *52*(40), 6968–6981. <https://doi.org/10.1021/bi401057x>
- Woods, D. F., & Bryant, P. J. (1991). The discs-large tumor suppressor gene of Drosophila encodes a guanylate kinase homolog localized at septate junctions. *Cell*, *66*(3), 451–464. [https://doi.org/10.1016/0092-8674\(81\)90009-X](https://doi.org/10.1016/0092-8674(81)90009-X)
- Wulf, G. M., Ryo, A., Wulf, G. G., Lee, S. W., Niu, T., Petkova, V., & Lu, K. P. (2001). Pin1 is overexpressed in breast cancer and cooperates with Ras signaling in increasing the transcriptional activity of c-Jun towards cyclin D1. *The EMBO Journal*, *20*(13), 3459–3472. <https://doi.org/10.1093/emboj/20.13.3459>
- Xu, D., Tsai, C. J., & Nussinov, R. (1997). Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Engineering*, *10*(9), 999–1012. <https://doi.org/10.1093/protein/10.9.999>
- Yanez Orozco, I. S., Mindlin, F. A., Ma, J., Wang, B., Levesque, B., Spencer, M., Rezaei Adariani, S., Hamilton, G., Ding, F., Bowen, M. E., & Sanabria, H. (2018). Identifying weak interdomain interactions that stabilize the supertertiary structure of the N-terminal tandem PDZ domains of PSD-95. *Nature Communications*, *9*(1), 3724. <https://doi.org/10.1038/s41467-018-06133-0>
- Yao, J. L., Kops, O., Lu, P. J., & Lu, K. P. (2001). Functional conservation of phosphorylation-specific prolyl isomerases in plants. *The Journal of Biological Chemistry*, *276*(17), 13517–13523. <https://doi.org/10.1074/jbc.M007006200>
- Ye, F., & Zhang, M. (2013). Structures and target recognition modes of PDZ domains: recurring themes and emerging pictures. *The Biochemical Journal*, *455*, 1–14.

<https://doi.org/10.1042/BJ20130783>

- Zeng, M., Chen, X., Guan, D., Xu, J., Wu, H., Tong, P., & Zhang, M. (2018). Reconstituted Postsynaptic Density as a Molecular Platform for Understanding Synapse Formation and Plasticity. *Cell*, *174*(5), 1172-1187.e16. <https://doi.org/10.1016/j.cell.2018.06.047>
- Zeng, M., Shang, Y., Araki, Y., Guo, T., Haganir, R. L., & Zhang, M. (2016). Phase Transition in Postsynaptic Densities Underlies Formation of Synaptic Complexes and Synaptic Plasticity. *Cell*, *166*(5), 1163-1175.e12. <https://doi.org/10.1016/j.cell.2016.07.008>
- Zhang, L., Bouguet-Bonnet, S., & Buck, M. (2012). Combining NMR and molecular dynamics studies for insights into the allostery of small GTPase-protein interactions. *Methods in Molecular Biology (Clifton, N.J.)*, *796*(2), 235–259. [https://doi.org/10.1007/978-1-61779-334-9\\_13](https://doi.org/10.1007/978-1-61779-334-9_13)
- Zhang, Q., Fan, J. S., & Zhang, M. (2001). Interdomain Chaperoning between PSD-95, Dlg, and Zo-1 (PDZ) Domains of Glutamate Receptor-interacting Proteins. *Journal of Biological Chemistry*, *276*(46), 43216–43220. <https://doi.org/10.1074/jbc.M105996200>
- Ziff, E. B. (1997). Enlightening the postsynaptic density. In *Neuron* (Vol. 19, Issue 6, pp. 1163–1174). [https://doi.org/10.1016/S0896-6273\(00\)80409-2](https://doi.org/10.1016/S0896-6273(00)80409-2)
- Zydowsky, L. D., Etzkorn, F. A., Chang, H. Y., Ferguson, S. B., Stolz, L. A., Ho, S. I., & Walsh, C. T. (1992). Active site mutants of human cyclophilin A separate peptidyl-prolyl isomerase activity from cyclosporin A binding and calcineurin inhibition. *Protein Science*, *1*(9), 1092–1099. <https://doi.org/10.1002/pro.5560010903>

## 9 Függelék

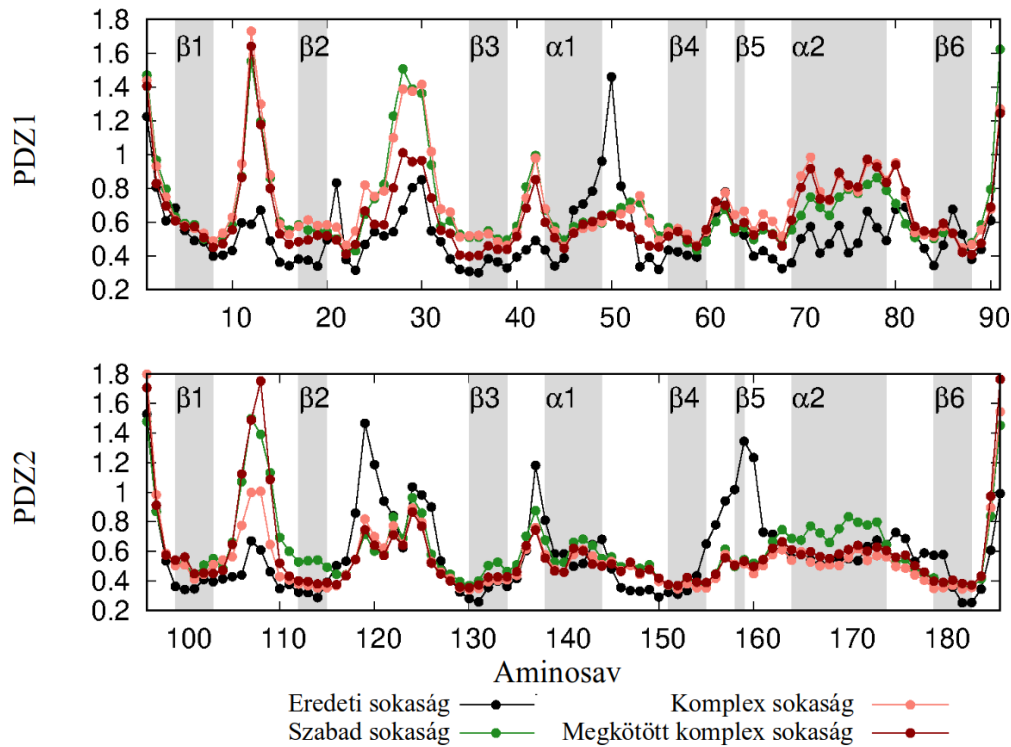
*F1. táblázat: A PSD-t felépítő fehérjék kategóriái és legfontosabb képviselőik. A második oszlop az egyes kategóriák százalékos előfordulási arányát tartalmazza, a harmadik oszlopban pedig a nagyobb, jól mérhető számban fellelhető fehérjék neve mellett lévő szám azok abszolút gyakoriságát jelzi egy átlagos méretű PSD-ben (Sheng & Hoogenraad, 2007; Sheng & Kim, 2011).*

Kategória funkció szerint	Előfordulási arány	Legfontosabb példák és átlagos abszolút gyakoriságuk
membránreceptorok és ioncsatornák	6%	AMPA-receptor ( <b>60</b> ) NMDA-receptor ( <b>60</b> ) mGluR1/5 receptor-alegységek ( <b>20</b> )
vázfehérjék	6%	GKAP ( <b>150</b> ) Homer ( <b>60</b> ) Lin-7 MAGUK (pl. PSD-95) ( <b>360</b> ) Shank ( <b>150</b> ) AKAP ( <b>20</b> )
transzlációs molekuláris gépezet	6%	
sejtadhéziós molekulák	7%	$\alpha$ -Catenin $\beta$ -Catenin ( <b>40</b> ) $\delta$ -Catenin N-Cadherin ( <b>30</b> ) Densin-180 NCAM-140 Neurofascin Neurologin
dajkafehérjék	2%	HSP40
citoszkeleton (aktin)	12%	Actin $\alpha$ -Actinin $\alpha$ -Adducin Ankyrins Arp2/3 complex Contactin Drebrin LIM Neurabin $\alpha/\beta$ -Spectrins
citoszkeleton (egyéb)	4%	MAP1A MAP1B MAP2 Septin ( <b>100</b> ) Tubulin
mitokondrium fehérjéi	6%	
kináz és foszfatáz szabályozó fehérjék	11%	CaMKII $\alpha/\beta/\delta$ ( <b>5600</b> ) Casein kinase 2 PKC $\gamma$ ( <b>20</b> ) Protein phosphatase PP1 Protein phosphatase PP2A
membrán trafficking	5%	AP1 AP2 Clathrin NSF

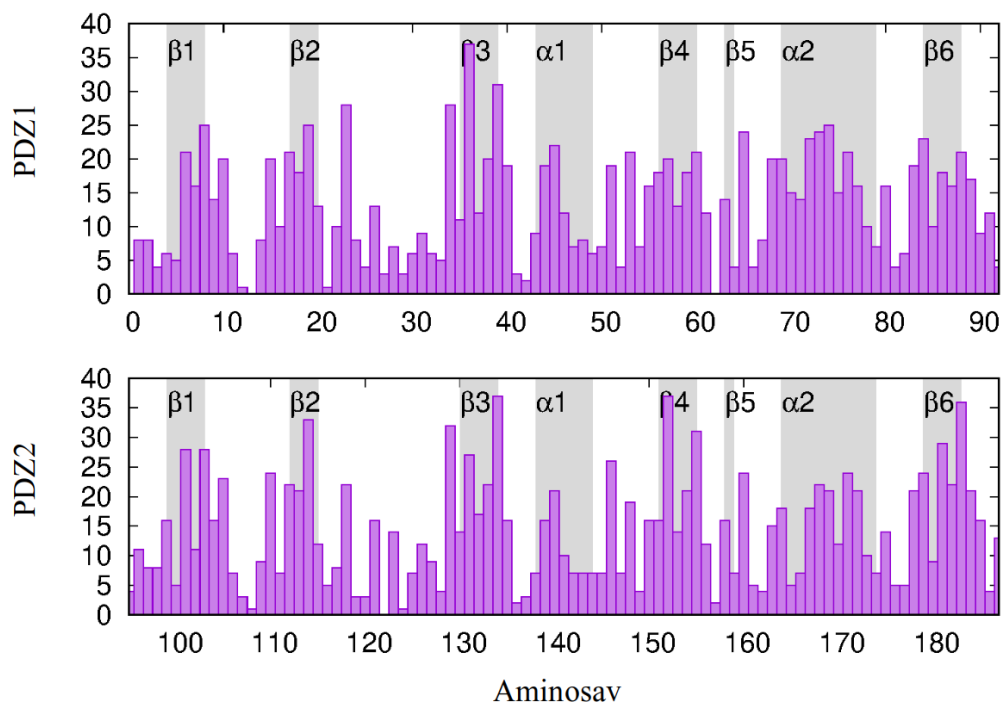
		SNAP-25 IP Synapsin Syntaxin BP
metabolikus fehérjék	7%	
GTPáz szabályozó fehérjék	8%	ARF3 cAMP-GEFII (RapGEF) GIT1 (ArfGAP) Heterotrimeric G proteins Kalirin (RhoGEF) PIKE-L (ArfGAP) Ras SPAR (RapGAP) SynGAP (RasGAP)
motorfehérjék	4%	Dynein Myosin IIb Myosin V Myosin VI
egyéb	15%	

F2. ábra: **A:** Lokális RMSD-k az aminosav sorszám függvényében, az eredeti PDB sokaságra, valamint az általunk előállított szabad, komplex és megkötött komplex sokaságra. A lokális RMSD számolásához a peptidgerinc N, C és Ca atomjait vettem figyelembe (ld. 4.3.1 fejezet). **B:** Távoli NOE távolságkényszerfeltételek száma aminosavanként. Azon aminosavak közötti NOE-kényszerfeltételeket tekintetem távolinká, amelyek legalább 2 távolságra vannak egymástól a szekvenciában.

**A**



**B**



F3. táblázat: A PSD-95 fehérje felépítése. A doménhatárok kijelölése a BLAST (Altschul et al., 1990) keresőalgorithmus által is használt SMART (Letunic & Bork, 2018) adatbázis alapján történt.

<b>Domén</b>	<b>Hosszúság</b>	<b>Doménhatárok</b>
PDZ1	91	62-152
PDZ2	91	157-247
PDZ3	85	310-394
SH3	62	429-490
GK	167	544-710



F4. ábra: Az egérből, patkányból és emberből származó PSD-95 teljes szekvencia-illesztése a Clustal Omega programmal. Külön jelölve vannak a doménhatárok, valamint azok az aminosavak, amelyekben különbözik a szekvencia.

CLUSTAL O(1.2.4) multiple sequence alignment

```

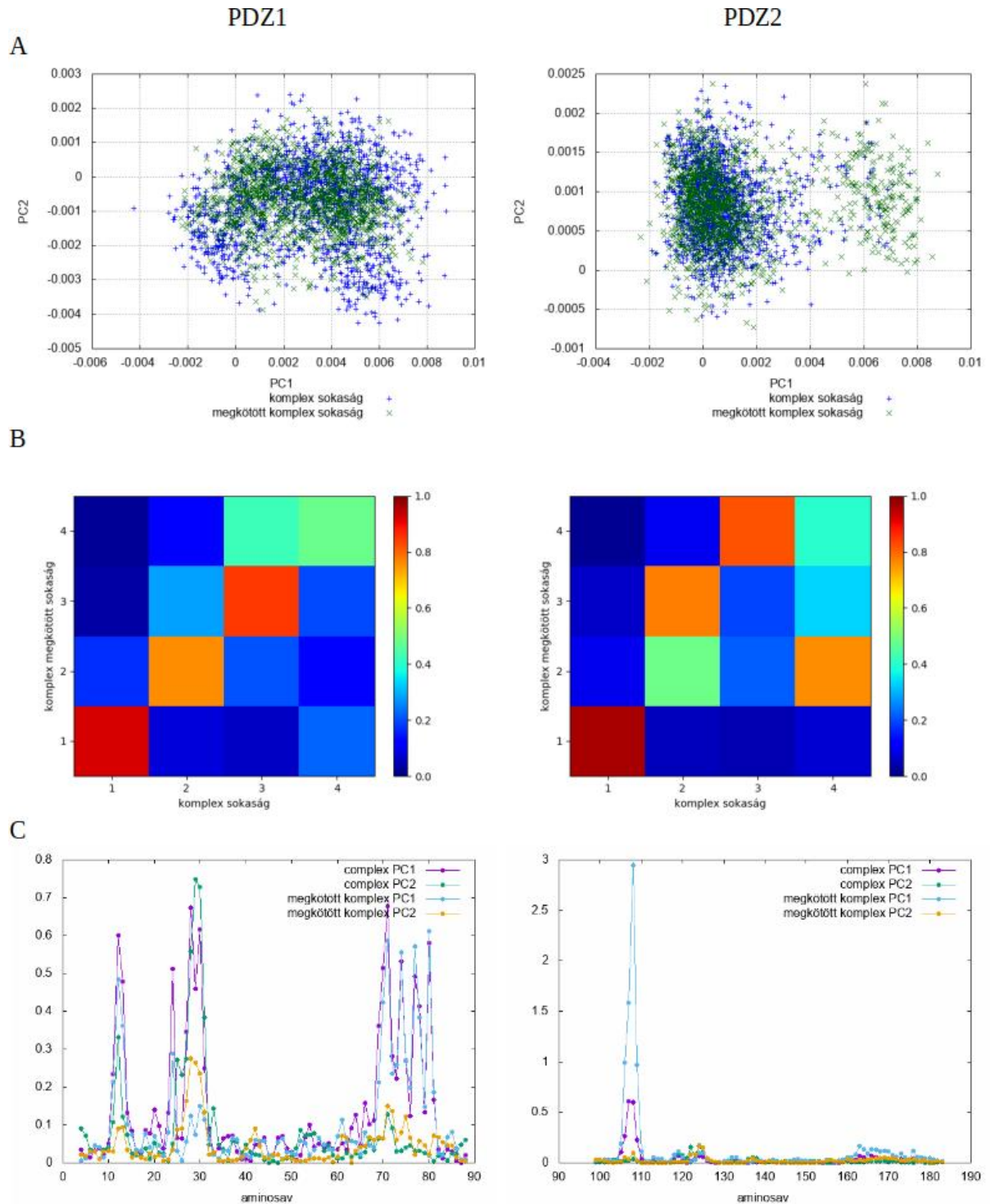
sp|Q62108|DLG4_MOUSE      MDCLCIVTTKKYRYQDEDTPPLEHSPAHLPNQANSPPVIVNTDTLEAPGYELQVNGTEGE 60
sp|P31016|DLG4_RAT       MDCLCIVTTKKYRYQDEDTPPLEHSPAHLPNQANSPPVIVNTDTLEAPGYELQVNGTEGE 60
sp|P78352|DLG4_HUMAN     MDCLCIVTTKKYRYQDEDTPPLEHSPAHLPNQANSPPVIVNTDTLEAPGYELQVNGTEGE 60
*****
                        PDZ1
sp|Q62108|DLG4_MOUSE      MEYEEITLERGNSGLGFSIAGGTDNPHIGDDPSIFITKIIIPGGAAAQDGRRLVNDLILFV 120
sp|P31016|DLG4_RAT       MEYEEITLERGNSGLGFSIAGGTDNPHIGDDPSIFITKIIIPGGAAAQDGRRLVNDLILFV 120
sp|P78352|DLG4_HUMAN     MEYEEITLERGNSGLGFSIAGGTDNPHIGDDPSIFITKIIIPGGAAAQDGRRLVNDLILFV 120
*****
                        PDZ2
sp|Q62108|DLG4_MOUSE      NEVDVREVTHTSAAVEALKEAGSIVRLYVMRRKPPAEKIEIKLIKPKGLGFSIAGGVGN 180
sp|P31016|DLG4_RAT       NEVDVREVTHTSAAVEALKEAGSIVRLYVMRRKPPAEKVMIEIKLIKPKGLGFSIAGGVGN 180
sp|P78352|DLG4_HUMAN     NEVDVREVTHTSAAVEALKEAGSIVRLYVMRRKPPAEKVMIEIKLIKPKGLGFSIAGGVGN 180
*****
                        PDZ3
sp|Q62108|DLG4_MOUSE      QHIPGDNSIYVTKIIEGGAHKGRLQIGDKILAVNSVGLQEDVMHEDAVAALKNTYDVVY 240
sp|P31016|DLG4_RAT       QHIPGDNSIYVTKIIEGGAHKGRLQIGDKILAVNSVGLQEDVMHEDAVAALKNTYDVVY 240
sp|P78352|DLG4_HUMAN     QHIPGDNSIYVTKIIEGGAHKGRLQIGDKILAVNSVGLQEDVMHEDAVAALKNTYDVVY 240
*****
sp|Q62108|DLG4_MOUSE      LKVAKPSNAYLSDSYAPPDITTSYSQHLNEISHSSYLGTDYPTAMTPTSPRRYSPVAKD 300
sp|P31016|DLG4_RAT       LKVAKPSNAYLSDSYAPPDITTSYSQHLNEISHSSYLGTDYPTAMTPTSPRRYSPVAKD 300
sp|P78352|DLG4_HUMAN     LKVAKPSNAYLSDSYAPPDITTSYSQHLNEISHSSYLGTDYPTAMTPTSPRRYSPVAKD 300
*****
                        PDZ3
sp|Q62108|DLG4_MOUSE      LLGEEDI PREPRRIV IHRGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRKGQDIL 360
sp|P31016|DLG4_RAT       LLGEEDI PREPRRIV IHRGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRKGQDIL 360
sp|P78352|DLG4_HUMAN     LLGEEDI PREPRRIV IHRGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRKGQDIL 360
*****
sp|Q62108|DLG4_MOUSE      SVNGVDLRNASHEQAAIALKNAGQTVTIIAQYKPEEYSRFEAKIHDLEQLMNSLGSST 420
sp|P31016|DLG4_RAT       SVNGVDLRNASHEQAAIALKNAGQTVTIIAQYKPEEYSRFEAKIHDLEQLMNSLGSST 420
sp|P78352|DLG4_HUMAN     SVNGVDLRNASHEQAAIALKNAGQTVTIIAQYKPEEYSRFEAKIHDLEQLMNSLGSST 420
*****
                        SH3
sp|Q62108|DLG4_MOUSE      ASLRSNPKRGFYIRALFDYDKTKDCGFLSQALSFRFGDVLHVIDASDEEWWQARRVHSDS 480
sp|P31016|DLG4_RAT       ASLRSNPKRGFYIRALFDYDKTKDCGFLSQALSFRFGDVLHVIDAGDEEWWQARRVHSDS 480
sp|P78352|DLG4_HUMAN     ASLRSNPKRGFYIRALFDYDKTKDCGFLSQALSFRFGDVLHVIDASDEEWWQARRVHSDS 480
*****
                        GK
sp|Q62108|DLG4_MOUSE      ETDDIGFIPSKRRVERREWSRLKAKDWGSSSGSQGREDSVLSYETVTVQMEVHYARPIIIL 540
sp|P31016|DLG4_RAT       ETDDIGFIPSKRRVERREWSRLKAKDWGSSSGSQGREDSVLSYETVTVQMEVHYARPIIIL 540
sp|P78352|DLG4_HUMAN     ETDDIGFIPSKRRVERREWSRLKAKDWGSSSGSQGREDSVLSYETVTVQMEVHYARPIIIL 540
*****
sp|Q62108|DLG4_MOUSE      GPTKDRANDDLLSEFPDKFGSCVPHTTRPKREYIDGRDYHFVSSREKMEKDIQAHKFIE 600
sp|P31016|DLG4_RAT       GPTKDRANDDLLSEFPDKFGSCVPHTTRPKREYIDGRDYHFVSSREKMEKDIQAHKFIE 600
sp|P78352|DLG4_HUMAN     GPTKDRANDDLLSEFPDKFGSCVPHTTRPKREYIDGRDYHFVSSREKMEKDIQAHKFIE 600
*****
sp|Q62108|DLG4_MOUSE      AGQYNHSHLYGTSVQSVREVAEQGKHCILDVSANAVRRLQAAHLHPAIAIFIRPRSLNVLE 660
sp|P31016|DLG4_RAT       AGQYNHSHLYGTSVQSVREVAEQGKHCILDVSANAVRRLQAAHLHPAIAIFIRPRSLNVLE 660
sp|P78352|DLG4_HUMAN     AGQYNHSHLYGTSVQSVREVAEQGKHCILDVSANAVRRLQAAHLHPAIAIFIRPRSLNVLE 660
*****
sp|Q62108|DLG4_MOUSE      INKRITEEQARKAFDRATKLEQEFTECFSAIVEGDSFEEIYHKVQRVIEDLSGPYIWVPA 720
sp|P31016|DLG4_RAT       INKRITEEQARKAFDRATKLEQEFTECFSAIVEGDSFEEIYHKVQRVIEDLSGPYIWVPA 720
sp|P78352|DLG4_HUMAN     INKRITEEQARKAFDRATKLEQEFTECFSAIVEGDSFEEIYHKVQRVIEDLSGPYIWVPA 720
*****
sp|Q62108|DLG4_MOUSE      RERL      724
sp|P31016|DLG4_RAT       RERL      724
sp|P78352|DLG4_HUMAN     RERL      724
****

```

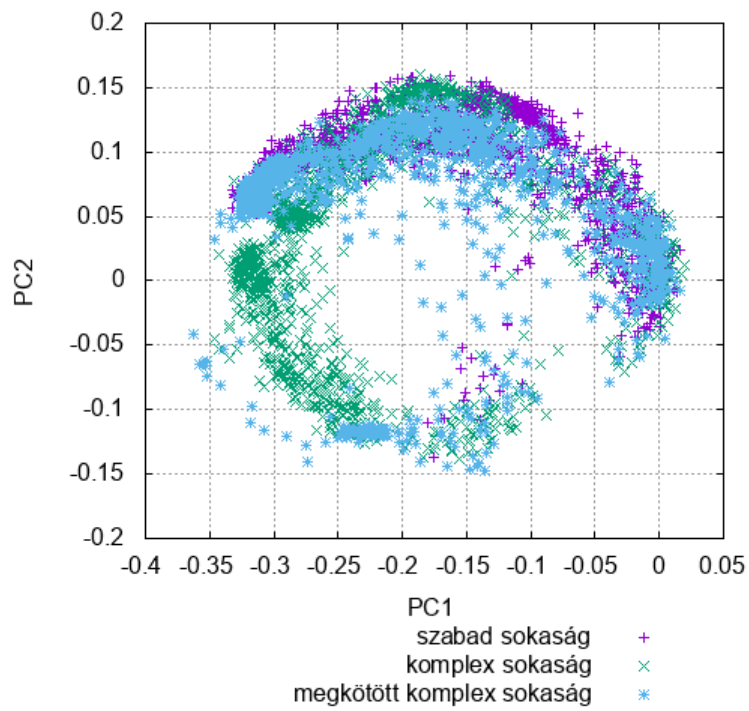
F5. táblázat: A PSD-95 fehérje PDZ1 és PDZ2 doménjéről készült, PDB adatbázisban elérhető kísérletes szerkezetek, valamint a hozzáférhető NMR-s kísérletes adatok (kémiai eltolódás:  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ , rendparaméter:  $S^2$ , valamint  $^1\text{H}$ - $^1\text{H}$  NOE-k). A nyilvánosan nem elérhető adatok zárójelben szerepelnek.

PDB kód	Organizmus	Módszer	Felbontás	Tartomány	Forma	Elérhető NMR-es adatok
1KEF	Homo sapiens	NMR		PDZ1	Szabad	$^1\text{H}$ , $^{13}\text{C}$ , $^{15}\text{N}$
3ZRT	Homo sapiens	Röntgenkristallográfia	3.40 Å	PDZ1-2	Szabad	
6SPV	Homo sapiens	Röntgenkristallográfia	2.04 Å	PDZ1-2	Szabad	
6SPZ	Homo sapiens	Röntgenkristallográfia	2.08 Å	PDZ1-2	Komplex	
1IU0	Rattus norvegicus	NMR		PDZ1	Szabad	$^1\text{H}$ , $^{13}\text{C}$ , $^{15}\text{N}$
1IU2	Rattus norvegicus	NMR		PDZ1	Szabad	$^1\text{H}$ , $^{13}\text{C}$ , $^{15}\text{N}$
1QLC	Rattus norvegicus	NMR		PDZ2	Szabad	$^1\text{H}$ - $^1\text{H}$ NOE
1RGR	Rattus norvegicus	NMR		PDZ1	Komplex	$^1\text{H}$ , $^{13}\text{C}$ , $^{15}\text{N}$
2KA9	Rattus norvegicus	NMR		PDZ1-2	Komplex	( $^1\text{H}$ , $^{15}\text{N}$ ), $^1\text{H}$ - $^1\text{H}$ NOE, ( $S^2$ )
2XKX	Rattus norvegicus	Elektronmikroszkóp	22.90 Å	Teljes fehérje	Szabad	
3GSL	Rattus norvegicus	Röntgenkristallográfia	2.05 Å	PDZ1-2	Szabad	

F6. ábra: A megkötött komplex, és a nem megkötött komplex sokaság intradomén mozgásainak összehasonlítása a PDZ1 és PDZ2 doménekben. **A:** A PDZ1 és PDZ2 doménekre elvégzett főkomponens-analízis ábrázolása. **B:** A nem megkötött és a megkötött komplex sokaság interdomén főkomponenseinek átfedése. **C:** A megkötött komplex és a nem megkötött komplex sokaság első két főkomponense által reprezentált négyzetes fluktuációk az aminosav sorszám függvényében.



F7. ábra: A három előállított PDZ1-2 tandem sokaságra elvégzett interdomén főkomponens-analízis (ld. 5.2.3. fejezet).



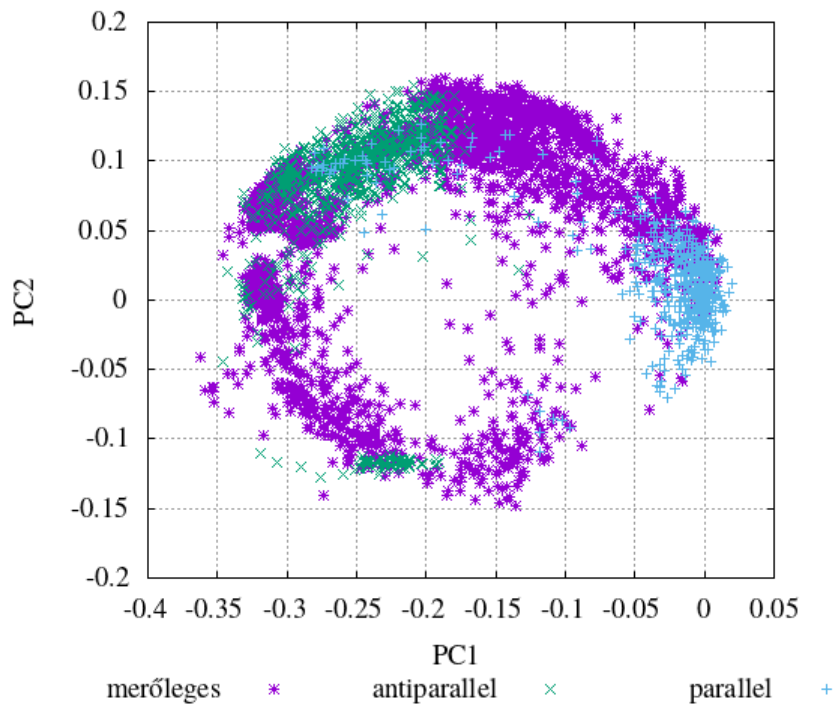
F8. táblázat: A PDZ1-2 tandemről készített klaszterek mérete (konformerek száma), a visszszámolt kémiai eltolódások korrelációja és a globális RMSD értékek a peptidgerinc atomjaira meghatározva.

Klaszter	Méret	Kémiai eltolódás korrel.		RMSD (Å)	
		N	H	PDZ1	PDZ2
1#	60	0,85	0,56	0,48 ± 0,00	0,55 ± 0,01
2#	15	0,85	0,55	0,72 ± 0,01	0,51 ± 0,01
3#	152	0,84	0,54	0,55 ± 0,02	0,54 ± 0,01
4#	631	0,84	0,54	0,60 ± 0,02	0,54 ± 0,01
5#	131	0,84	0,55	0,67 ± 0,02	0,55 ± 0,02
6#	250	0,84	0,55	0,79 ± 0,02	0,63 ± 0,01
7#	317	0,84	0,53	0,68 ± 0,02	0,61 ± 0,01

F9. táblázat: A PDZ1-2 tandemről készített klaszterek mérete (konformerek száma), a komplex forma aránya, az interfészt kialakító régiók (20.B ábra), valamint a ligandumok pozíciója: orientációja egymáshoz, illetve az interdomén tengelyhez viszonyítva, és elhelyezkedésük (ld. F11. ábra).

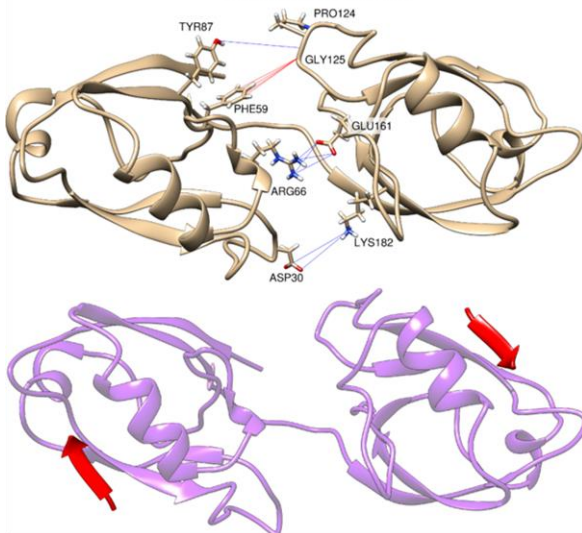
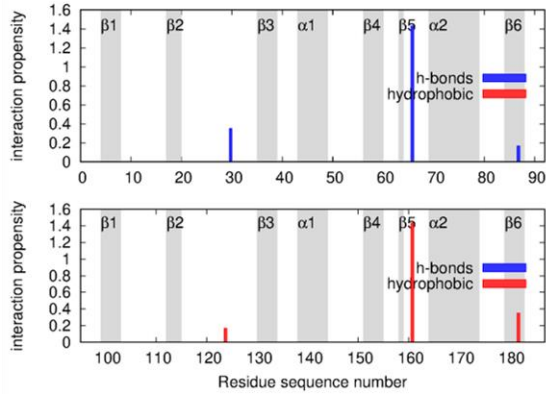
Klaszter	Méret	Komplex forma (%)	Interfész régió		Ligandumok orientációja egymáshoz viszonyítva	Ligandumok orientációja az interdomén tengelyhez viszonyítva	Ligandumok elhelyezkedése
			PDZ1	PDZ2			
1#		100%	Fej	Fej	Antiparallel	Parallel	Ellentétes oldal
2#	60	100%	Fej	Jobb oldal	Merőleges	-	Azonos oldal
3#	15	100%	Jobb oldal	Fej	Antiparallel	Parallel	Azonos oldal
4#	152	87%	Jobb oldal	Alsó rész (Jobb oldal)	Antiparallel	Merőleges	Azonos oldal
5#	631	100%	Jobb oldal	Hát	Merőleges	-	Ellentétes oldal
6#	131	70%	Alsó rész	Alsó rész	Merőleges	-	Ellentétes oldal
7#	250	4%	Alsó rész	Alsó rész	Merőleges	-	Ellentétes oldal

F10. ábra: A lehetséges ligandum-orientációk eloszlása a PDZ1-2 tandemben az interdomén PCA-ábrán. A ligandum-orientációkat a  $-2$  aminosav Ca atomjából a  $0$  aminosav Ca atomjába mutató vektorok által bezárt szög alapján definiáltam, a következő elv szerint.  $<45^\circ$ : *parallel*,  $>45^\circ$  és  $<135^\circ$ : *merőleges*,  $>135^\circ$ : *antiparallel*.

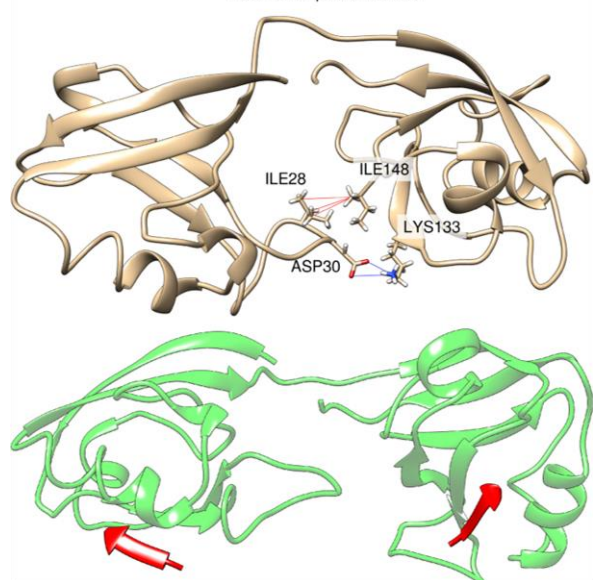
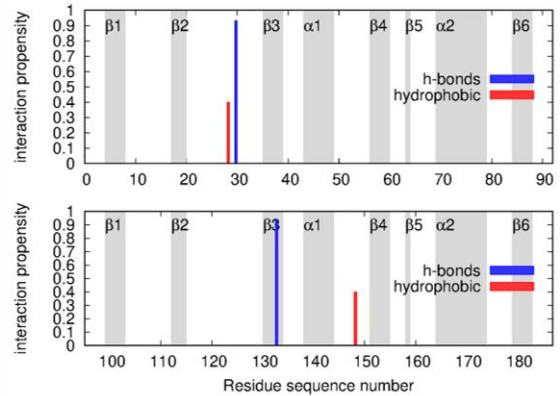


F11. ábra: A PDZ1-2 tandemről készített klaszterek ábrázolásai. Főnt: az interdomén kölcsönhatási hajlandóság minden egyes aminosavra (ld. 5.2.5. fejezet). A hidrogénkötések kék, a hidrofób kölcsönhatások piros színnel vannak jelölve. Középen: minden egyes klaszter reprezentatív szerkezete, külön jelölve az interdomén hidrogénkötéseket kékkel, valamint az interdomén hidrofób kölcsönhatásokat pirossal. Alul: minden egyes klaszter reprezentatív szerkezete, a ligandum 4 C-terminális aminosavját pirossal jelölve. Mindegyik ábrán a PDZ1 domén bal, a PDZ2 domén jobb oldalon van.

## Cluster 1

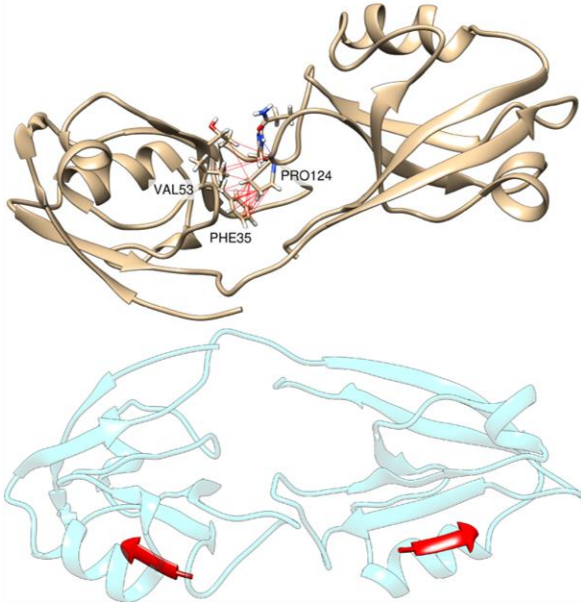
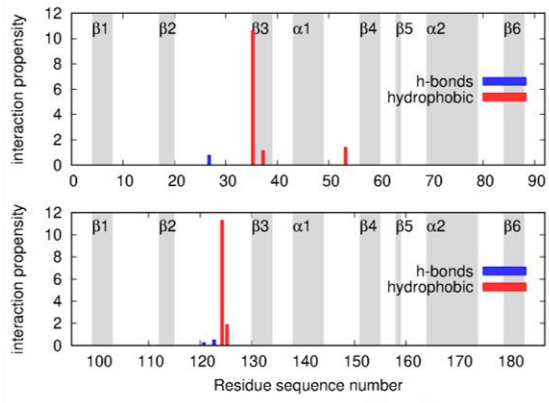


## Cluster 2

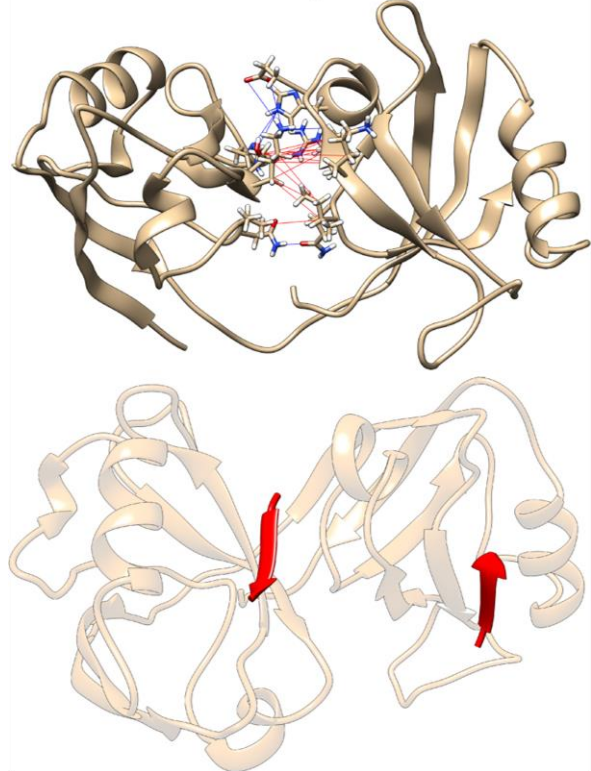
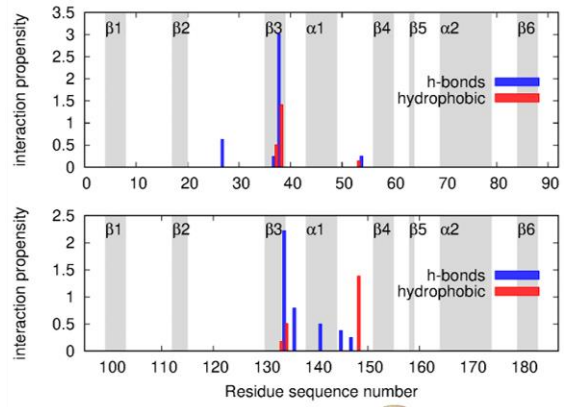




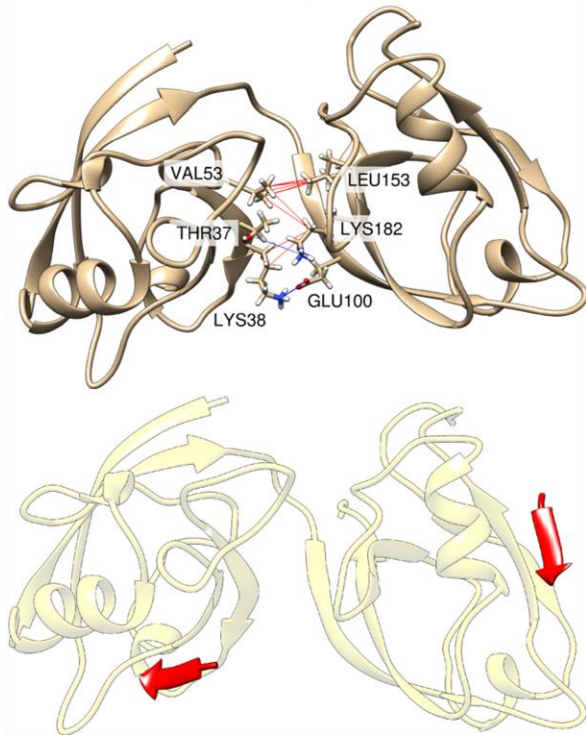
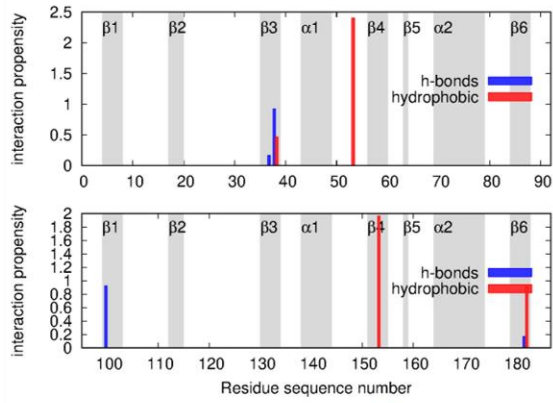
## Cluster 3



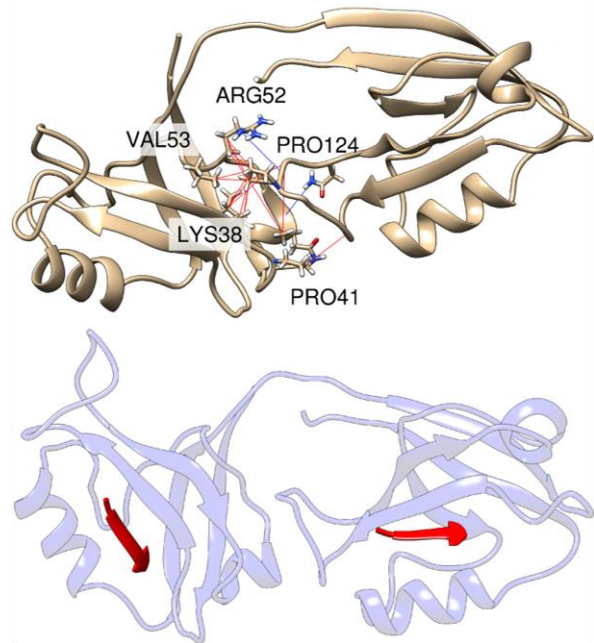
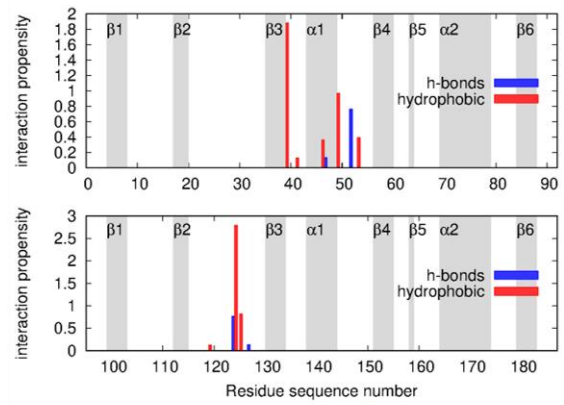
## Cluster 4



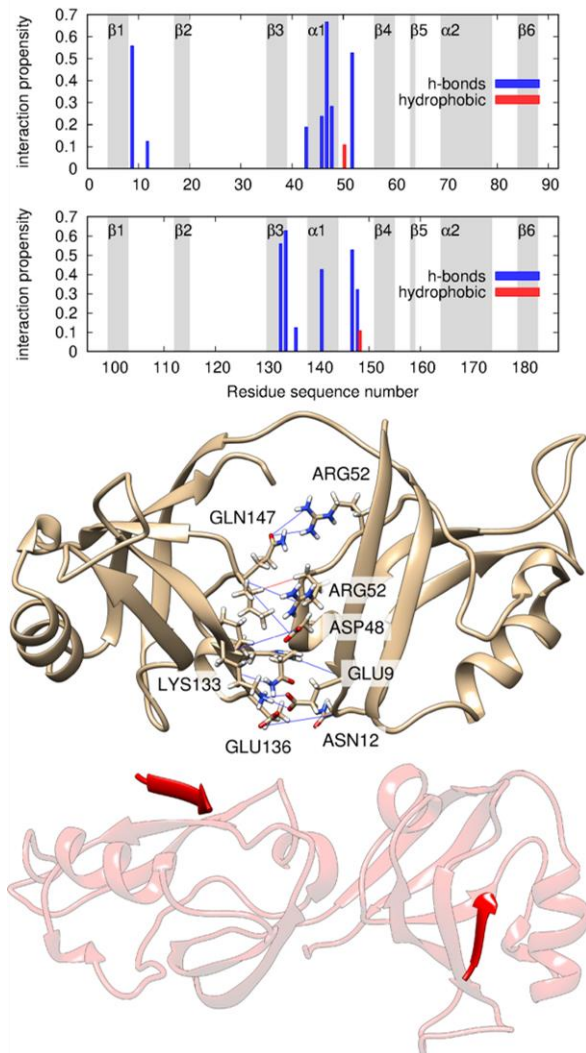
## Cluster 5



## Cluster 6



## Cluster 7



F12. táblázat: Az interdomén hidrogénkötések és hidrofób kölcsönhatások előfordulási aránya az egyes PDZ1-2 tandem klaszterekben. A legalább 10%-os előfordulási arányok zöld háttérrel vannak kiemelve. Csak olyan kölcsönhatások vannak feltüntetve a táblázatban, amelyek valamelyik klaszterben legalább 10%-ban megjelennek.

Hidrogénkötések								
PDZ1 atom	PDZ2 atom	cl1	cl2	cl3	cl4	cl5	cl6	cl7
ASN12@N	GLU136@OE2	0%	0%	0%	0%	0%	0%	12%
HIS27@ND1	HIS141@NE2	0%	0%	0%	26%	0%	0%	0%
HIS27@NE2	GLN121@OE1	0%	0%	26%	0%	0%	0%	0%
HIS27@NE2	ILE123@O	0%	0%	50%	0%	0%	0%	0%
HIS27@O	ARG145@NH1	0%	0%	0%	38%	0%	0%	0%
ASP30@OD1	LYS133@NZ	0%	13%	2%	0%	0%	0%	0%
ASP30@OD1	LYS182@NZ	20%	0%	0%	0%	0%	0%	0%
ASP30@OD2	LYS133@NZ	0%	80%	7%	0%	0%	0%	0%
ASP30@OD2	LYS182@NZ	15%	0%	0%	0%	0%	0%	0%
THR37@O	LYS182@NZ	0%	0%	0%	0%	17%	0%	0%
LYS38@NZ	GLU100@OE1	0%	0%	0%	1%	45%	0%	0%
LYS38@NZ	GLU100@OE2	0%	0%	0%	0%	47%	0%	0%
LYS38@NZ	ILE134@O	0%	0%	0%	55%	0%	0%	0%
LYS38@NZ	GLU136@OE1	0%	0%	0%	39%	0%	0%	0%
LYS38@NZ	GLU136@OE2	0%	0%	0%	40%	0%	0%	0%
GLY43@O	HIS141@NE2	0%	0%	0%	0%	0%	8%	19%
ALA46@O	HIS141@NE2	0%	0%	0%	0%	0%	0%	24%
GLN47@NE2	ILE134@O	0%	0%	0%	0%	0%	6%	51%
GLN47@O	ASN127@ND2	0%	0%	0%	0%	0%	14%	1%
GLN47@O	ILE148@N	0%	0%	0%	0%	0%	1%	16%
ASP48@O	ILE148@N	0%	0%	0%	0%	0%	0%	16%
ASP48@OD2	ILE134@N	0%	0%	0%	0%	0%	0%	12%
ARG52@NE	PRO124@O	0%	0%	0%	0%	0%	11%	0%
ARG52@NH1	GLN147@OE1	0%	0%	0%	0%	0%	2%	19%
ARG52@NH2	GLN147@OE1	0%	0%	0%	1%	0%	4%	34%
ASN54@ND2	GLN147@OE1	0%	0%	0%	25%	0%	0%	0%
ARG66@NH1	GLU161@OE1	45%	0%	0%	0%	0%	0%	0%
ARG66@NH1	GLU161@OE2	32%	0%	0%	0%	0%	0%	0%
ARG66@NH2	GLU161@OE1	32%	0%	0%	0%	0%	0%	0%
ARG66@NH2	GLU161@OE2	35%	0%	0%	0%	0%	0%	0%

TYR87@OH	PRO124@O	17%	0%	0%	0%	0%	0%	0%
GLU9@OE1	LYS133@NZ	0%	0%	0%	0%	0%	0%	32%
GLU9@OE2	LYS133@NZ	0%	0%	0%	0%	0%	0%	24%
<b>Hidrofób kölcsönhatások</b>								
<b>PDZ1 atom</b>	<b>PDZ2 atom</b>	<b>cl1</b>	<b>cl2</b>	<b>cl3</b>	<b>cl4</b>	<b>cl5</b>	<b>cl6</b>	<b>cl7</b>
ILE28@CB	ILE148@CG2	0%	13%	1%	0%	0%	0%	0%
ILE28@CG1	ILE148@CG2	0%	13%	3%	0%	0%	0%	0%
ILE28@CG2	ILE148@CG2	0%	13%	5%	0%	0%	0%	0%
PHE35@CB	PRO124@CB	0%	0%	86%	0%	0%	0%	0%
PHE35@CB	PRO124@CG	0%	0%	74%	0%	0%	0%	0%
PHE35@CD1	PRO124@CB	0%	0%	80%	0%	0%	0%	0%
PHE35@CD1	PRO124@CD	0%	0%	18%	0%	0%	0%	0%
PHE35@CD1	PRO124@CG	0%	0%	87%	0%	0%	0%	0%
PHE35@CD2	PRO124@CB	0%	0%	63%	0%	0%	0%	0%
PHE35@CD2	PRO124@CD	0%	0%	18%	0%	0%	0%	0%
PHE35@CD2	PRO124@CG	0%	0%	78%	0%	0%	0%	0%
PHE35@CE1	PRO124@CB	0%	0%	49%	0%	0%	0%	0%
PHE35@CE1	PRO124@CD	0%	0%	17%	0%	0%	0%	0%
PHE35@CE1	PRO124@CG	0%	0%	81%	0%	0%	0%	0%
PHE35@CE2	PRO124@CB	0%	0%	34%	0%	0%	0%	0%
PHE35@CE2	PRO124@CD	0%	0%	13%	0%	0%	0%	0%
PHE35@CE2	PRO124@CG	0%	0%	64%	0%	0%	0%	0%
PHE35@CG	PRO124@CB	0%	0%	84%	0%	0%	0%	0%
PHE35@CG	PRO124@CD	0%	0%	20%	0%	0%	0%	0%
PHE35@CG	PRO124@CG	0%	0%	86%	0%	0%	0%	0%
PHE35@CZ	PRO124@CB	0%	0%	28%	0%	0%	0%	0%
PHE35@CZ	PRO124@CD	0%	0%	16%	0%	0%	0%	0%
PHE35@CZ	PRO124@CG	0%	0%	64%	0%	0%	0%	0%
THR37@CG2	PRO124@CB	0%	0%	41%	0%	0%	0%	0%
THR37@CG2	GLY125@CA	0%	0%	74%	1%	0%	0%	0%
THR37@CG2	ILE134@CB	0%	0%	0%	17%	0%	0%	0%
THR37@CG2	ILE134@CG1	0%	0%	0%	12%	0%	0%	0%
THR37@CG2	ILE134@CG2	0%	0%	0%	22%	0%	0%	0%
LYS38@CB	ILE134@CB	0%	0%	0%	14%	0%	0%	0%

LYS38@CB	ILE148@CB	0%	0%	0%	23%	0%	0%	0%
LYS38@CB	ILE148@CG1	0%	0%	0%	61%	0%	0%	0%
LYS38@CB	LYS182@CD	0%	0%	0%	0%	26%	0%	0%
LYS38@CD	LYS133@CG	0%	0%	0%	18%	0%	0%	0%
LYS38@CD	ILE134@CB	0%	0%	0%	39%	0%	0%	0%
LYS38@CD	ILE134@CG2	0%	0%	0%	24%	0%	0%	0%
LYS38@CD	ILE148@CG1	0%	0%	0%	14%	0%	0%	0%
LYS38@CG	ILE134@CB	0%	0%	0%	43%	0%	0%	0%
LYS38@CG	ILE134@CG1	0%	0%	0%	21%	0%	0%	0%
LYS38@CG	ILE134@CG2	0%	0%	0%	26%	0%	0%	0%
LYS38@CG	ILE148@CG1	0%	0%	0%	27%	0%	0%	0%
LYS38@CG	LYS182@CD	0%	0%	0%	0%	21%	0%	0%
ILE39@CB	PRO124@CB	0%	0%	0%	0%	0%	41%	0%
ILE39@CB	PRO124@CG	0%	0%	0%	0%	0%	23%	0%
ILE39@CG1	PRO124@CB	0%	0%	0%	0%	0%	44%	0%
ILE39@CG1	PRO124@CD	0%	0%	0%	0%	0%	11%	0%
ILE39@CG1	PRO124@CG	0%	0%	0%	0%	0%	25%	0%
ILE39@CG2	PRO124@CB	0%	0%	0%	0%	0%	31%	1%
ILE39@CG2	PRO124@CG	0%	0%	0%	0%	0%	14%	0%
PRO41@CB	GLY119@CA	0%	0%	0%	0%	0%	13%	0%
ALA46@CB	PRO124@CB	0%	0%	0%	0%	0%	25%	1%
ALA46@CB	GLY125@CA	0%	0%	0%	0%	0%	11%	0%
GLY49@CA	PRO124@CB	0%	0%	0%	0%	0%	26%	1%
GLY49@CA	GLY125@CA	0%	0%	0%	0%	0%	70%	2%
ARG50@CB	ILE148@CB	0%	0%	0%	0%	0%	0%	11%
ARG52@CB	PRO124@CB	0%	0%	0%	0%	0%	16%	0%
ARG52@CB	PRO124@CG	0%	0%	0%	0%	0%	21%	0%
ARG52@CG	PRO124@CB	0%	0%	0%	0%	0%	11%	0%
ARG52@CG	PRO124@CG	0%	0%	0%	0%	0%	18%	0%
VAL53@CB	GLY125@CA	0%	0%	35%	0%	0%	0%	0%
VAL53@CB	LEU153@CD1	0%	0%	0%	0%	22%	0%	0%
VAL53@CG1	PRO124@CB	0%	0%	1%	0%	0%	24%	0%
VAL53@CG1	PRO124@CG	0%	0%	1%	0%	0%	15%	0%
VAL53@CG1	LEU153@CD1	0%	0%	0%	0%	43%	0%	0%
VAL53@CG1	LEU153@CD2	0%	0%	0%	0%	11%	0%	0%

VAL53@CG1	LYS182@CD	0%	0%	0%	0%	18%	0%	0%
VAL53@CG2	PRO124@CB	0%	0%	26%	0%	0%	8%	0%
VAL53@CG2	GLY125@CA	0%	0%	78%	0%	0%	0%	0%
VAL53@CG2	ILE148@CB	0%	0%	0%	14%	0%	0%	0%
VAL53@CG2	LEU153@CD1	0%	0%	0%	0%	92%	0%	0%
VAL53@CG2	LEU153@CD2	0%	0%	0%	0%	14%	0%	0%
VAL53@CG2	LEU153@CG	0%	0%	0%	0%	15%	0%	0%
VAL53@CG2	LYS182@CD	0%	0%	0%	0%	27%	0%	0%
PHE59@CD2	GLY125@CA	15%	0%	0%	0%	0%	0%	0%
PHE59@CE1	GLY125@CA	15%	0%	0%	0%	0%	0%	0%
PHE59@CE2	GLY125@CA	32%	0%	0%	0%	0%	0%	0%
PHE59@CZ	GLY125@CA	33%	0%	0%	0%	0%	0%	0%



F13. táblázat: A PDZ domén is a ligandum között létrejövő hidrogénkötések az egyes klaszterekben.

Nagyon gyakran előforduló PDZ-ligandum hidrogénkötések (≥50% mindegyik klaszterbe)												
Viszonylag gyakran előforduló PDZ-ligandum hidrogénkötések (≥10%)												
PDZ1												
PDZ atom	ligandum atom	PDZ régió	ligandum aminosav	all	cl1	cl2	cl3	cl4	cl5	cl6	cl7	
PHE17@O	VAL198@N	β2	0	83%	87%	80%	89%	82%	78%	84%	79%	
ILE19@N	SER196@O	β2	-2	98%	100%	100%	99%	98%	96%	98%	100%	
ILE19@O	SER195@OG	β2	-3	15%	18%	0%	12%	16%	15%	18%	36%	
ILE19@O	SER196@N	β2	-2	94%	100%	100%	93%	91%	99%	94%	100%	
GLY21@N	GLN190@O	β2	-8	4%	0%	73%	10%	1%	0%	0%	0%	
GLY21@N	PHE194@O	β2	-4	22%	12%	0%	3%	40%	3%	9%	36%	
ASP24@O	VAL191@N	β2-β3	-7	1%	0%	0%	0%	0%	0%	2%	21%	
ASN25@ND2	VAL192@O	β2-β3	-6	21%	0%	0%	1%	42%	0%	2%	0%	
ASN25@ND2	PRO193@O	β2-β3	-5	1%	17%	0%	0%	0%	0%	0%	0%	
ASN25@OD1	VAL191@N	β2-β3	-7	3%	0%	13%	1%	0%	0%	11%	29%	
ASN25@OD1	VAL192@N	β2-β3	-6	2%	0%	0%	1%	2%	0%	5%	50%	
HIS27@ND1	VAL192@N	β2-β3	-6	1%	0%	20%	5%	0%	0%	0%	0%	
HIS27@O	GLN190@NE2	β2-β3	-8	2%	0%	0%	11%	0%	2%	1%	0%	
HIS27@O	VAL191@N	β2-β3	-7	8%	0%	0%	0%	0%	63%	1%	0%	
HIS27@O	VAL192@N	β2-β3	-6	3%	0%	7%	0%	0%	6%	14%	0%	
THR37@OG1	SER195@OG	β3	-3	14%	2%	67%	19%	2%	46%	17%	29%	
HIS70@NE2	VAL192@O	α2	-6	4%	42%	0%	0%	3%	0%	0%	0%	
HIS70@NE2	SER196@OG	α2	-2	7%	5%	33%	14%	7%	5%	5%	0%	
PDZ2												
PDZ atom	ligand atom			all	cl1	cl2	cl3	cl4	cl5	cl6	cl7	
PHE112@O	SER206@OG	β2	-1	14%	5%	27%	13%	16%	18%	9%	7%	
PHE112@O	VAL207@N	β2	0	53%	88%	60%	59%	47%	57%	52%	43%	
SER113@OG	SER204@O	β2	-3	1%	0%	13%	2%	0%	0%	1%	0%	
SER113@OG	SER205@O	β2	-2	14%	2%	20%	11%	15%	12%	14%	21%	
SER113@OG	SER206@OG	β2	-1	8%	28%	20%	5%	5%	9%	10%	7%	
ILE114@N	SER205@O	β2	-2	95%	98%	87%	93%	95%	98%	94%	100%	
ILE114@O	SER204@OG	β2	-3	16%	20%	7%	11%	19%	18%	12%	7%	

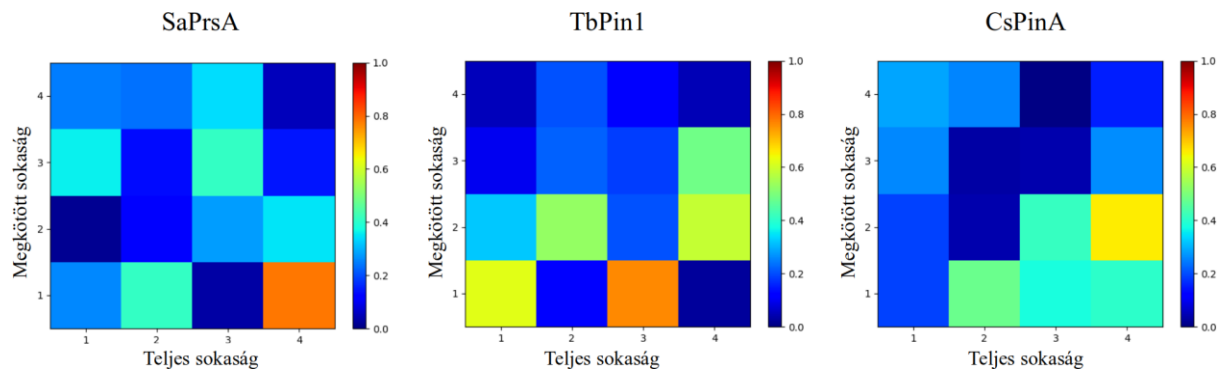


ILE114@O	SER205@N	$\beta$ 2	-2	76%	83%	40%	80%	77%	86%	70%	50%
ASN120@ND2	VAL201@O	$\beta$ 2- $\beta$ 3	-6	6%	0%	0%	0%	11%	5%	0%	0%
ASN120@ND2	PRO202@O	$\beta$ 2- $\beta$ 3	-5	5%	5%	0%	0%	1%	34%	0%	0%
ASN120@ND2	PHE203@O	$\beta$ 2- $\beta$ 3	-4	8%	62%	0%	14%	3%	12%	1%	0%
ASN120@OD1	VAL201@N	$\beta$ 2- $\beta$ 3	-6	6%	0%	0%	0%	7%	2%	13%	0%
ASN120@OD1	PHE203@N	$\beta$ 2- $\beta$ 3	-4	1%	20%	0%	0%	0%	1%	0%	0%
HIS122@NE2	VAL201@N	$\beta$ 2- $\beta$ 3	-6	1%	0%	0%	0%	0%	1%	5%	14%
HIS165@NE2	PRO202@O	$\alpha$ 2	-5	1%	12%	0%	0%	0%	2%	0%	0%
HIS165@NE2	SER205@OG	$\alpha$ 2	-2	5%	48%	0%	13%	1%	0%	2%	0%
GLU166@OE1	VAL200@N	$\alpha$ 2	-7	1%	0%	13%	3%	0%	0%	0%	0%
GLU166@OE1	VAL201@N	$\alpha$ 2	-6	1%	0%	13%	3%	0%	0%	0%	0%
GLU166@OE2	GLN199@NE2	$\alpha$ 2	-8	2%	3%	0%	0%	0%	10%	3%	14%
GLU166@OE2	VAL200@N	$\alpha$ 2	-7	1%	0%	40%	2%	0%	0%	0%	0%
GLU166@OE2	VAL201@N	$\alpha$ 2	-6	1%	0%	33%	2%	0%	0%	0%	0%
LYS173@NZ	SER206@O	$\alpha$ 2	-1	5%	0%	7%	6%	5%	5%	3%	14%

F14. táblázat: Egyedi konformerek és a kísérletes FRET-adatok közötti korreláció. Az egyes konformerek a 7 klaszter reprezentatív szerkezetei, valamint a kísérletes szerkezete. A FRET-adatsorok 10-10 interdomén atom-atom távolságot tartalmaznak (Yanez Orozco et al., 2018) (ld. 5.2.6. fejezet).

Konformer	$\langle R_{DA} \rangle_1$ korrel.	$\langle R_{DA} \rangle_2$ korrel.	$\langle R_{DA} \rangle_3$ korrel.
<i>Klaszter reprezentatív</i>			
klaszter 1	0.51	0.60	-0.60
klaszter 2	0.56	0.65	-0.68
klaszter 3	0.76	0.54	-0.65
klaszter 4	0.13	0.78	-0.39
klaszter 5	0.62	0.23	-0.48
klaszter 6	0.48	0.25	-0.27
klaszter 7	-0.17	-0.09	0.23
<i>Kísérletes szerkezetek</i>			
FRET	0.60	0.71	-0.60
nyitott	0.85	0.22	-0.41
zárt	0.48	0.78	-0.65
3GSL:A	0.74	0.51	-0.65
3GSL:B	0.82	0.45	-0.61

F15. ábra: Az egyes parvulin-típusú fehérjékből készült megkötött sokaságok leghangúlyosabb belső mozgásainak átfedése a mindhárom sokaságot tartalmazó, kombinált sokaság mozgásaival. A főkomponens-elemzést a konszenzus szekvencia aminosavjaira végeztük el.



F16. táblázat: A kísérletes parvulin-típusú rotamáz-doméneket tartalmazó teljes fehérjék listája. A táblázatban külön szerepel az egyes fehérjék egyedi Uniprot ([uniprot.org](http://uniprot.org)) azonosítója, teljese neve, valamint hogy a teljes fehérje összesen egy (rot) vagy két (rot + rot) rotamáz domént, illetve egy rotamáz és egy WW domént (rot + WW) tartalmaz.

Uniprot	Szerkezet	Fehérje angol neve
A9A2C7	rot	PpiC-type peptidyl-prolyl cis-trans isomerase
D0ZJR9	rot + rot	Chaperone SurA
O74049	rot	Parvulin-like peptidyl-prolyl isomerase
P0A9L5	rot	Peptidyl-prolyl cis-trans isomerase C
P0ABZ6	rot + rot	Chaperone SurA
P0ADY1	rot	Peptidyl-prolyl cis-trans isomerase D
P24327	rot + rot	Foldase protein PrsA
P60747	rot	Foldase protein PrsA
Q0PAS1	rot	Putative peptidyl-prolyl cis-trans isomerase Cbf2
Q13526	rot + WW	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1
Q2G9Z0	rot	Chaperone SurA
Q57XM6	rot	Uncharacterized protein
Q57YG1	rot	Peptidyl-prolyl cis-trans isomerase
Q72NB3	rot	PpiC domain-containing protein
Q9C475	rot + WW	Peptidyl-prolyl cis-trans isomerase
Q9SL42	rot	Peptidyl-prolyl cis-trans isomerase Pin1
Q9Y237	rot	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 4

F17. táblázat: Az elemzéshez felhasznált, parvulin-típusú rotamáz doméneket tartalmazó kísérletes szerkezetek PDB kódjai. A szerkezetek mellett fel van tüntetve, hogy a teljes fehérje összesen egy (rot) vagy két (rot + rot) rotamáz domént, illetve egy rotamáz és egy WW domént (rot + WW) tartalmaz.

PDB kód	szerkezet	PDB kód	szerkezet	PDB kód	szerkezet
1EQ3	rot	2RUC	rot + WW	3KAD	rot + WW
1F8A	rot + WW	2RUD	rot + WW	3KAF	rot + WW
1FJD	rot	2XP3	rot + WW	3KAG	rot + WW
1J6Y	rot	2XP4	rot + WW	3KAH	rot + WW
1JNS	rot	2XP5	rot + WW	3KAI	rot + WW
1JNT	rot	2XP6	rot + WW	3KCE	rot + WW
1M5Y	rot + rot	2XP7	rot + WW	3NRK	rot
1NMV	rot + WW	2XP8	rot + WW	3NTP	rot + WW
1NMW	rot + WW	2XP9	rot + WW	3ODK	rot + WW
1PIN	rot + WW	2XPA	rot + WW	3OOB	rot + WW
1YW5	rot + WW	2XPB	rot + WW	3RFW	rot
1ZCN	rot + WW	2ZQS	rot + WW	3TC5	rot + WW
1ZK6	rot + rot	2ZQT	rot + WW	3TCZ	rot + WW
2F21	rot + WW	2ZQU	rot + WW	3TDB	rot + WW
2ITK	rot + WW	2ZQV	rot + WW	3UI4	rot
2JZV	rot	2ZR4	rot + WW	3UI5	rot
2KGJ	rot	2ZR5	rot + WW	3UI6	rot
2LJ4	rot	2ZR6	rot + WW	3WH0	rot + WW
2M08	rot	3GPK	rot	4G2P	rot + rot
2M1I	rot	3I6C	rot + WW	4QIB	rot + WW
2MNT	rot	3IK8	rot + WW	4TNS	rot + WW
2PV1	rot + rot	3IKD	rot + WW	4TYO	rot + WW
2PV2	rot + rot	3IKG	rot + WW	4U84	rot + WW
2PV3	rot + rot	3JYJ	rot + WW	4U85	rot + WW
2Q5A	rot + WW	3KAB	rot + WW	4U86	rot + WW
2RQS	rot	3KAC	rot + WW	4WO7	rot + rot



3wh0AR1 -----K-----ITRTKEEAELELINGYIQIKS-----GE-E  
3ui4AR1 -----KHGKIMEAMEKLKS-----G--M  
3kacAR1 -----RPSSWRQE-----ITRTQEAELELINGYIQIKS-----GE-E  
2xp7AR1 -----TKEEAELELINGYIQIKS-----GE-E  
3gpkBR1 -----A-----TEENKQPVFANAEKIVEQLKQ-----G--G  
3tdbAR1 -----SSWRQE-----ITRTKEEAELELINGYIQIKS-----GE-E  
2rucAR1 HLLVKHSQS---RRPSSWRQE-----ITRTKEEAELELINGYIQIKS-----GE-E  
1jnsAR1 -----HILVKEEKLALDLEQIKN-----GA  
2f21AR1 -----RRPSSWRQ-----EKIT-RTKEEAELELINGYIQIKS-----GE-E  
2pv3BR1 -----PLPE---N-----P-----TSDQVNEAESQARAIVDQARN-----  
1m5yCR1 -----DQVNEAESQARAIVDQARN-----G--A  
3oobAR1 -----EEAELELINGYIQIKS-----GE-E  
3tc5AR1 -----TKEEAELELINGYIQIKS-----GE-E  
2pv2BR1 HILIPLP-E---N-----P-----TSDQVNEAESQARAIVDQARN-----G--A  
2xp3AR1 -----TKEEAELELINGYIQIKS-----GE-E  
3tczAR1 -----SSWRQE-----ITRTKEEAELELINGYIQIKS-----GE-E  
3kabAR1 -----TKEEAELELINGYIQIKS-----GE-E  
1m5yBR2 -----QIAADIKS-----GK-T  
2xp4AR1 -----EEAELELINGYIQIKS-----GE-E  
4u86AR1 -----KIKS-----GE-E  
4tnsBR1 -----S-----GE-E  
4g2pAR1 -----LLK-----P-----SPINDQQARLKLEEIAADIKS-----GK-T  
4tyoAR1 --RPSSWRQE-----Q-----ITRTQEAELELINGYIQIKS-----GE-E  
4tnsAR1 -----GE-E  
4u85AR1 -----KIKS-----GE-E  
4g2pBR1 -----LLK-----P-----SPINDQQARLKLEEIAADIKS-----GK-T  
4tyoBR1 SRRPSSWRQE-----Q-----ITRTQEAELELINGYIQIKS-----GE-E  
4qibAR1 -----WRQE-----K-----ITRTKEEAELELINGYIQIKS-----GE-E  
4wo7AR1 -----DKKTAEVEKLLKK-----G--E  
4wo7BR1 -----KKTAEVEKLLKK-----G--E  
4u84AR1 -----KIKS-----GE-E  
  
2pv2CR1 DFGKLAIAHSAD-Q-QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
2jzvAR1 KFGEIAKKESMDTG-SAKKDGELGYVL---KG-QTKDFEKALFKLKDQ-----EVS-E-  
2xp5AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1m5yBR1 DFGKLAIAHSADQ--QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
3odkAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3nrkAR1 SFALIAGSPRNDPA-LRARRGMVEWISSFDLY-KYSKITATIAAPLNG-----GVS--E  
1jntAR1 DFGKLAKKHSICPS--GKRGGDLGEFR---QG-QMVPAFDKVVFSCPVL-----EPT-G-  
21j4AR1 SFEEAASQFSDC-S-SYASGGDLGFFS---SG-EMMKPFEDAVRALKIG-----DIS-P-  
3jyjAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3ui6AR1 RFNEVAAQYSED---KARQGGDLGWMT---RG-SMVGFQEAALPVSGMDKPVFT-DP  
3kadAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMAKPFEDASFALRTG-----EMS-G-  
1zcnAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2zr4AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1nmwAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3kahAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2rqsAR1 KFGKLAKELSIDGG-SAKRDGSLGYFG---RG-KMVKPFEDAAFRLQVG-----EVS-E-  
1m5yDR2 TFAAAAKEFSQDPG-SANQGGDLGWAT---PD-IFDPAFRDALTRLNKG-----QMS-A-  
2xpbAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1yw5AR1 KLSELANTESDC-S-SHDRGGDLGFFS---KG-QMQPFEEAALNLHVG-----EVS-N-  
3ui5AR1 RFNEVAAQYSED---KARQGGDLGWMT---RG-SMVGFQEAALPVSGMDKPVFT-DP  
2m08AR1 KFGKLAKELSIDSG-SAKKGNLGYFT---KG-MMVKPFEDAAFRLQVG-----EVS-E-  
2mntAR1 EFVQVVRDFSEC-G-SAKRDGDLGMVE---SG-TYTEGFDTVAFLSKSG-----EVS-A-  
2xp8AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3ikgBR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3jyjBR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1zk6AR1 KFEDLAKEYSTDS--SASKGGDLGWFA---KEGQMDFTFSKAAFKLKTG-----EVS-D-  
2kgjAR1 DFAALAKEKSADII-SARNGGDMGWLE---DA-TIPELKNAG-LKEKG-----QLS-G-  
3gpkAR1 SFVAYARQYSEA-S-TAAVGGDLGWLR---LA-QLPTELATTAAS-GPG-----QLA-G-  
3ikdAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2q5aAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2pv3AR1 DFGKLAIAHSAD-Q-QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
3rfwAR1 KFSELAKEKSIDPG-SKNQGGELGWFD---QS-TMVKPFDDAAAFALKNG-----TIT-TT  
1pinAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3ik8AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3kceAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3ikdBR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2xpaAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3kaiAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2zr6AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1m5yDR1 DFGKLAIAHSADQ--QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
2pv2DR1 DFGKLAIAHSAD-Q-QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
2pv2AR1 DFGKLAIAHSAD-Q-QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
2itkAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-

1f8aBR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3kagAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2zqvAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2xp9AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1eq3AR1 RFNEVAAQYSED---KARQGGDLGWMT---RG-SMVGFQEAAFALPVSGMDKPVFTDP-  
2zqtAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QAQKPFEDASFALRTG-----EMS-G-  
3ikgAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1m5yAR2 TFAAAAKEFSQDPG-SANQGGDLGWAT---PD-IFDPAFRDALTRLNKG-----QMS-A-  
3kafAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMAKPFEDASFALRTG-----EMS-G-  
2zquAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2zqsAR1 DFESLASQFSDA-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1j6yAR1 NFEVAVTRVSDC-S-SAKRGGDLGSFG---RG-QMQKPFEEATYALKVG-----DIS-D-  
2xp6AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMAKPFEDASFALRTG-----EMS-G-  
2rudAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3kacBR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3ik8BR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3ntpAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2ml1AR1 KFGKLAKELSIDGG-SAKRDGSLGYFG---RG-KMVKPFEDAAFRLQVG-----EVS-E-  
2pv1AR1 DFGKLAIAHSAD-Q-QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
1nmvAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1m5yAR1 DFGKLAIAHSADQ--QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
3i6cAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1fjdAR1 RFNEVAAQYSED---KARQGGDLGWMT---RG-SMVGFQEAAFALPVSGMDKPVFTDP-  
1m5yCR2 TFAAAAKEFSQDPG-SANQGGDLGWAT---PD-IFDPAFRDALTRLNKG-----QMS-A-  
3i6cBR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3wh0AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3ui4AR1 RFNEVAAQYSED---KARQGGDLGWMT---RG-SMVGFQEAAFALPVSGMDKPVFTDP-  
3kacAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2xp7AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3gpkBR1 SFVAYARQYSEA-S-TAAVGGDLGWIR---LA-QLPTELATAAAS-GPG-----QLA-G-  
3tdbAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2rucAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1jnsAR1 DFGKLAKKHSICPS--GKRGGDLGEFR---QG-QMVPAFDKVVFSCPVL-----EPT-G-  
2f21AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2pv3BR1 DFGKLAIAHSAD-Q-QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
1m5yCR1 DFGKLAIAHSADQ--QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
3oobAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3tc5AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
2pv2BR1 DFGKLAIAHSAD-Q-QALNGGQMGWGR---IQ-ELPGIFAQALSTAKKG-----DIV-G-  
2xp3AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3tczAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
3kabAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
1m5yBR2 TFAAAAKEFSQDPG-SANQGGDLGWAT---PD-IFDPAFRDALTRLNKG-----QMS-A-  
2xp4AR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
4u86AR1 DFESLASQFSD--SSAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
4tnsBR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
4g2pAR1 TFAAAAKEYSQD-PGSANQGGDLGWAT---PD-IFDPAFRDALTKLHKG-----QIS-A-  
4tyoAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
4tnsAR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
4u85AR1 DFESLASQFSD--SSAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
4g2pBR1 TFAAAAKEYSQD-PGSANQGGDLGWAT---PD-IFDPAFRDALTKLHKG-----QIS-A-  
4tyoBR1 DFESLASQFSDC-S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
4qibAR1 DFESLASQFSD--S-SAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-  
4wo7AR1 KFEDLAKEYSTD-S-SASKGGDLGWFA--KE-GQDETFSKAAFKLKTG-----EVS-D-  
4wo7BR1 KFEDLAKEYSTD-S-SASKGGDLGWFA--KE-GQDETFSKAAFKLKTG-----EVS-D-  
4u84AR1 DFESLASQFSD--SSAKARGDLGAFS---RG-QMQKPFEDASFALRTG-----EMS-G-

2pv2CR1 PIRS-GVGFHILKVND-----  
2jzvAR1 VVKS-SFGYHI IKAD-----  
2xp5AR1 PVFT-DSGIHII LRTE-----  
1m5yBR1 PIRS-GVGFHILKVNDLRTEVHARHILL-----  
3odkAR1 PVFT-DSGIHII LRTE-----  
3nrkAR1 VFRDERKRYCILKIEGKRPTMENLR-----  
1jntAR1 PLHT-QFGYHI IKV-----  
21j4AR1 IVQT-DSGLHI IKR-----  
3jyjAR1 PVFT-DSGIHII LRTE-----  
3ui6AR1 PVKT-KFGYHI IMVEGRK-----  
2zr5AR1 PVFT-DSGIHII LRTE-----  
3kadAR1 PVFT-DSGIHII LRTE-----  
1zcnAR1 PVFT-DSGIHII LRTE-----  
2zr4AR1 PVFT-DSGIHII LRTE-----  
1nmwAR1 PVFT-DSGIHII LRTE-----  
3kahAR1 PVFT-DSGIHII LRTE-----  
2rqsAR1 PVKS-EFGYHVIKR-----  
1m5yDR2 PVHS-SFGWHLIELL-DTRNVDKRAYRM-----  
2xpbAR1 PVFT-DSGIHII LRTE-----



1yw5AR1 IIET-NSGVHILQRT-----  
3ui5AR1 PVKT-KFGYHIIMVEGRK-----  
2m08AR1 PIKS-EFGYHI IKR-----  
2mntAR1 PVET-ELGVHLIYRV-----  
2xp8AR1 PVFT-DSGIHILRTE-----  
3ikgBR1 PVFT-DSGIHILRTE-----  
3jyjBR1 PVFT-DSGIHILRTE-----  
1zk6AR1 PVKT-QYGYHI IKKTEE-----  
2kgjAR1 VIKS-SVGFLIVRLDD-----  
3gpkAR1 PVEI-RGGFSILYLIDKREGH-----  
3ikdAR1 PVFT-DSGIHILRTE-----  
2q5aAR1 PVFT-DSGIHILRTE-----  
2pv3AR1 PIRS-GVGFHILKVNDLAAQKD-----  
3rfwAR1 PVKT-NFGYHVIL-----  
1pinAR1 PVFT-DSGIHILRTE-----  
3ik8AR1 PVFT-DSGIHILRTE-----  
3kceAR1 PVFT-DSGIHILRTE-----  
3ikdBR1 PVFT-DSGIHILRTE-----  
2xpaAR1 PVFT-DSGIHILRTE-----  
3kaiAR1 PVFT-DSGIHILRTE-----  
2zr6AR1 PVFT-DSGIHILRTE-----  
1m5yDR1 PIRS-GVGFHILKVNDLRGESKNISVTE-----  
2pv2DR1 PIRS-GVGFHILKVND-----  
2pv2AR1 PIRS-GVGFHILKVND-----  
2itkAR1 PVFT-DSGIHILRTE-----  
1f8aBR1 PVFT-DSGIHILRTE-----  
3kagAR1 PVFT-DSGIHILRTE-----  
2zqvAR1 PVFT-DSGIHILRTE-----  
2xp9AR1 PVFT-DSGIHILRTE-----  
1eq3AR1 PVKT-KFGYHIIMVE-----  
2zqtAR1 PVFT-DSGIHILRTE-----  
3ikgAR1 PVFT-DSGIHILRTE-----  
1m5yAR2 PVHS-SFGWHLIELL-DTRNVD-RAYRM-----  
3kafAR1 PVFT-DSGIHILRTE-----  
2zquAR1 PVFT-DSGIHILRTE-----  
2zqsAR1 PVFT-DSGIHILRTE-----  
1j6yAR1 IVDT-DSGVHI IKRTA-----  
2xp6AR1 PVFT-DSGIHILRTE-----  
2rudAR1 PVFT-DSGIHILRT-----  
3kacBR1 PVFT-DSGIHILRTE-----  
3ik8BR1 PVFT-DSGIHILRTE-----  
3ntpAR1 PVFT-DSGIHILRTE-----  
2mliAR1 PVKS-EFGYHVIKR-----  
2pv1AR1 PIRS-GVGFHILKVND-----  
1nmvAR1 PVFT-DSGIHILRT-----  
1m5yAR1 PIRS-GVGFHILKVNDLRGESKNISVT-----  
3i6cAR1 PVFT-DSGIHILRTE-----  
1fjdAR1 PVKT-KFGYHIIMVE-----  
1m5yCR2 PVHS-SFGWHLIELL-DTRN---RAYRMLMNRKFSEEAAS  
3i6cBR1 PVFT-DSGIHILRTE-----  
3wh0AR1 PVFT-DSGIHILRTE-----  
3ui4AR1 PVKT-KFGYHIIMVEGRK-----  
3kacAR1 PVFT-DSGIHILRTE-----  
2xp7AR1 PVFT-DSGIHILRTE-----  
3gpkBR1 PVEI-RGGFSILYLIDKR-----  
3tdbAR1 PVFT-DSGIHILRTE-----  
2rucAR1 PVFT-DSGIHILRT-----  
1jnsAR1 PLHT-QFGYHI IKV-----  
2f21AR1 PVFT-DSGIHILRTE-----  
2pv3BR1 PIRS-GVGFHILKVNDLAAQKD-----  
1m5yCR1 PIRS-GVGFHILKVNDL-TEVHARHILLK-----  
3oobAR1 PVFT-DSGIHILRTE-----  
3tc5AR1 PVFT-DSGIHILRTE-----  
2pv2BR1 PIRS-GVGFHILKVND-----  
2xp3AR1 PVFT-DSGIHILRTE-----  
3tczAR1 PVFT-DSGIHILRTE-----  
3kabAR1 PVFT-DSGIHILRTE-----  
1m5yBR2 PVHS-SFGWHLIELL-DTRN---RAYRMLMNRKFSEEAAS  
2xp4AR1 PVFT-DSGIHILRTE-----  
4u86AR1 PVFT-DSGIHILRTE-----  
4tnsBR1 PVFT-DSGIHILRTE-----  
4g2pAR1 PVHS-SFGWHLIELLDTRKV-----  
4tyoAR1 PVFT-DSGIHILRTE-----  
4tnsAR1 PVFT-DSGIHILRTE-----  
4u85AR1 PVFT-DSGIHILRTE-----  
4g2pBR1 PVHS-SFGWHLIELLDTRK-----

4tyoBR1 PVFT-DSGIHIILRTE-----  
4qibAR1 PVFT-DSGIHIILRTE-----  
4wo7AR1 PVKT-QYGYHIKKTEERGKY-----  
4wo7BR1 PVKT-QYGYHIKKTEERGKYD-----  
4u84AR1 PVFT-DSGIHIILRTE-----