

KÍSÉRLETILEG MEGHATÁROZOTT, BELSŐ DINAMIKÁT TÜKRÖZŐ FEHÉRJESZERKEZETI SOKASÁGOK: ELŐÁLLÍTÁS, ELEMZÉS ÉS BIOLÓGIAI RELEVANCIA

GÁSPÁRI ZOLTÁN

Az utóbbi évtizedekben nyilvánvalóvá vált, hogy a fehérjék térszerkezete mellett belső dinamikájuk is kulcsfontosságú szerepet játszik biológiai funkciójuk meghatározásában és finomhangolásában. A belső mozgások atomi szintű reprezentációjára olyan új modellek kerültek előtérbe, amelyek ugyanannak a molekulának nagyszámú különböző térszerkezeteit tartalmazzák, amelyek közötti különbségek nem elsősorban bizonytalanságból adódnak, hanem azt biztosítják, hogy a szerkezetek összessége megfeleltethető legyen egyes, a belső dinamikára vonatkozó kísérleti paramétereknek. Alább röviden áttekintem az ilyen térszerkezeti sokaságok előállításának főbb módszereit és azok validálásának legfontosabb alapelveit. Végül két, saját kutatásainkból vett példán bemutatom a koncepció alkalmazhatóságát.

1. Bevezetés

A fehérjék az élet alapvető molekulái közé tartoznak. A molekuláris biológia régi alaptétele szerint a fehérjék szekvenciája, azaz az őket felépítő aminosavak sorrendje meghatározza térszerkezetüket, azaz atomjaik egymáshoz képesti térbeli elhelyezkedését. Az utóbbi évtizedekben ez annyival egészült ki, hogy a szekvencia nem csupán a térszerkezetet, hanem annak időbeli átalakulásait, azaz dinamikáját is meghatározza. A jól meghatározott háromdimenziós szerkezettel bíró fehérjék mellett ez igaz az ún. funkcionálisan rendezetlen fehérjékre is, amelyek biológiai feladatukat úgy látják el, hogy nem rendelkeznek egyetlen jól körülírható térszerkezettel. Ezek a fehérjék igen dinamikusak, rövid idő alatt nagyon változatos térbeli elrendeződéseket vehetnek fel, szerkezetük mégsem teljesen véletlenszerű.

Az utóbbi évtizedben minden korábbinál nyilvánvalóbbá vált, hogy a fehérjék belső dinamikája, azaz a térszerkezet időbeli megváltozása jelentős szerepet játszik a biológiai funkció betöltésében. A fehérjék belső mozgásai között vannak, amelyek jellemzően néhány ezred, vagy akár milliomod másodperc alatt játszódnak le, és vannak, amelyek akár 10^{-9} másodpercnél is jóval gyorsabban. A gyorsabb mozgásokat tipikusan a molekuláris folyamatok entrópiájához való hozzájárulásuk, a

lassabb, milliszekundum-mikroszekundum időskálájú átrendeződéseket pedig a kémiai reakciók segítéséhez (katalízis), egyéb molekulákhoz való kötéshez szükséges alakváltozások miatt tartjuk fontosnak[7], bár a határvonalak természetesen nem minden esetben élesek. A jelenleg népszerű elmélet, a konformerszelekció szerint egy adott partner, pl. gyógyszermolekula a folytonosan változó szerkezetű fehérjemolekulák közül azokkal létesít kapcsolatot, amelyeknek alakja (és ezzel együtt fizikai-kémiai tulajdonságai, pl. elektromos töltéseinek helyzete) éppen megfelelő a kölcsönhatás kialakításához. Ezáltal a szerkezetváltozásnak megfelelő kémiai egyensúly eltolódik a fehérje-partner komplex irányába, így végül a fehérjék nagy része kötött állapotba kerül.

A fehérjék szerkezetét az őket alkotó atomok 3D koordinátaival adjuk meg. A belső dinamika leírásához több ilyen koordinátakészletet használunk, ez a szerkezeti sokaság, a gyakorlatban ponthalmazok összessége. A különböző ponthalmazokban (szerkezetekben) ugyanazon atom pozícióját leíró koordináták közötti különbségek tükrözik az adott atom mozgékonyágát úgy, hogy közben minden teljes ponthalmaz egy-egy kémiai is reális molekulaszerkezetet ír le.

2. Dinamikus fehérjeszerkezeti sokaságok előállítása és elemzése

A fehérjék minden egyes szerkezetéhez rendelhetünk egy, az atomok egymás közötti helyzetéből származó potenciális energiát. Ideális esetben, ha ismernénk minden lehetséges szerkezetet és azok energiáját, akkor a Boltzmann-eloszlás alapján meg tudnánk határozni az egyes szerkezetek gyakoriságát, és ezáltal a fehérje megvalósuló szerkezeteinek és azok egymáshoz viszonyított gyakoriságának teljes leírását tudnánk nyújtani. A gyakorlatban ez több okból sem megvalósítható. A lehetséges szerkezetek száma csillagászati, egy 100 aminosavból álló fehérje esetében óvatos becslés szerint is nagyságrendben 3^{100} térszerkezetet kellene előállítani és megvizsgálni. Másrészt a potenciális energia pontos kiszámítása is nagyon erőforrásigényes. Ezért jelentősen egyszerűsített fizikai modellek segítségével, az ismert fehérjeszerkezetekre jellemző tulajdonságokat felhasználva próbáljuk meg a várhatóan leggyakoribb, legfontosabb szerkezeteket és azok egymás közötti átmeneteit leírni. Ezen leírás pontosságát növeli, ha megköveteljük, hogy a szerkezetek összessége feleljen meg a kísérletileg meghatározott paramétereknek. Ennek legfontosabb feltétele, hogy ezek a paraméterek a kémiai szerkezet ismeretében becsülhetőek, „visszaszámolhatóak” legyenek, azaz legyen egyértelmű kapcsolat a térbeli szerkezet és az adott paraméter között. A kísérleti paraméterek lehetnek olyanok, amelyek egyenként, minden szerkezetre külön kiszámíthatóak és a szerkezetek összességére vonatkozóan megfelelő módon átlagolnunk kell őket, illetve olyanok, amelyek csak több szerkezet, azaz ponthalmaz együttesére értelmezhetőek. Az elsőre példa az ún. nukleáris Overhauser-effektusból származtatható, a hidrogénatomok közötti távolságokra vonatkozó paraméter, mely minden szerke-

zetre kiszámítható adott atompár esetére, a sokaságra átlagolni pedig nemlineáris módon szükséges (1. egyenlet):

$$r_{avg} = \left(\sum_{i=1}^N r_i^{-6} \right)^{-1/6}, \quad (1)$$

ahol r_i az i . szerkezetben mért távolság, N a szerkezetek száma, r_{avg} pedig a kísérleti adattal összevethető, a szerkezetek sokaságára vonatkozó átlag.

Ezzel szemben az ún. S^2 rendparaméter csak több szerkezeten értelmezhető, egyetlen szerkezeten értéke mindig 1 (2. egyenlet, [4]).

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

ahol \hat{x}_{ij} az i és j atomok mint háromdimenziós pontok közötti vektor irányával megegyező irányú egységvektor x komponense egy szerkezetben, $\langle \hat{x}_{ij} \rangle$ pedig ezek átlaga az összes szerkezetben.

A gyakorlatban a dinamikus sokaságok előállítására két fő megközelítés létezik [2]: a kísérleti adatokkal mint megkötésekkel kombinált, sokaság alapú, ún. molekuladinamikai szimuláció, illetve a nagy szerkezeti sokaságból a paramétereknek való megfelelés alapján válogató szelekciós eljárások. Előbbi megközelítésnél egyszerűen több fehérjeszerkezetben modellezük az atomok elmozdulását olyan módon, hogy az egyidejűleg megvalósuló atomi elrendeződések megfeleljenek a kísérleti paramétereknek. A szelekciós eljárások során nagyszámú, kémiaiailag reális lehetséges atomelrendeződést generálunk, és ezek közül választunk ki szerkezeteket úgy, hogy azok összessége minél jobban teljesítse a mért paramétereket. A szelekciós eljárások lehetnek determinisztikusak vagy sztochasztikusak.

A kapott sokaságok minőségének ellenőrzése, validálása során két szempontot kell érvényesítenünk [1], [8]. Egyrészt elvárás, hogy a szerkezetcsalád minden tagja geometriailag reális legyen, azaz az atomok mint pontok között definiált távolságok és szögek fizikailag és kémiaiilag megfelelőek legyenek, pl. az egymáshoz kémiai kötással kapcsolódó atomok ne legyenek egymástól túl távol, az egymáshoz nem kapcsolódóak túl közel stb. Másrészt természetesen kívánatos, hogy a szerkezetek összessége megfeleljen az adott molekulára mért kísérleti adatoknak, azoknak is, amelyeket felhasználtunk a szerkezetek előállításakor, és azoknak is, amelyeket nem. Ez utóbbi a keresztvalidáció, melynek fontos jelentősége van annak megítélésében, hogy a kapott szerkezeti sokaságok ténylegesen mennyire jó modelljei a valóságban megvalósuló molekuláris szerkezeteknek és elmozdulásoknak, és felhasználhatóak-e biokémiai jelenségek értelmezésére.

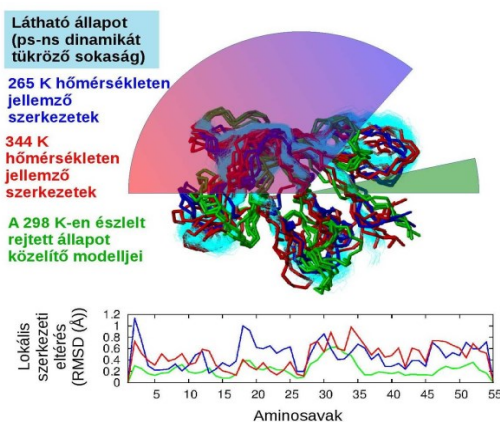
3. Alkalmazások

3.1. A PAF gombaellenes fehérje

A PAF (*Penicillium* antifungal protein) egy ígéretes gombaellenes fehérje, mely szelektíven pusztítja az *Aspergillus* nem fajait. Működési mechanizmusáról ugyanakkor igen keveset tudunk, kötőpartnere nem ismeretes. Erről az 55 aminosavas fehérjéről a Debreceni Egyetemen Batta Gyula vezetésével átfogó NMR-vizsgálatok készültek, amelyek arra utalnak, hogy az NMR-spektroszkópiával közvetlenül észlelhető forma csupán a molekulák mintegy 70%-át képviseli, a maradék kb. 30% a mérések számára „láthatatlan” „sötét anyag”. A molekula szerkezete mind hűtés, mind melegítés hatására megbomlik. Mindemellett sikerült egy, a mérhető formával kémiai egyensúlyban lévő, közvetlenül nem észlelhető állapot detektálása, amely azonban önmagában csupán a „sötét anyag” elenyésző részéért lehet felelős. Kísérleti adatokból származó megkötéseket alkalmazó molekuladinamikai számításokkal [6] előállítottuk az észlelhető állapot gyors, 10^{-12} – 10^{-9} másodperces időskálájú dinamikáját leíró térszerkezeti sokaságot, valamint egy megfelelően változatos szerkezethalmazból kiindulva sztochasztikus szelekcióval feltérképeztük a hideg és meleg kitekeredett állapotra jellemző szerkezetek közötti különbségeket. Ezen felül modelleztünk néhány olyan szerkezetet, amelyek együttesen, közelítőleg leírják a fő formával egyensúlyt tartó „rejtett” állapotot. Elmondható, hogy bár a fehérje nem megy át drasztikus szerkezetváltozáson, egyes jól körülhatárolható szerkezeti részei jelentősen és jellegzetes módon különböznek az egyes állapotokban. Ezt az ún. lokális RMSD (root mean square deviation), az egyes atomi pozíciók eltéréseit jellemző mennyiség is jól mutatja (3. egyenlet, 1. ábra). Összességében egy olyan modell állítható fel, ahol számos, közvetlenül nem észlelhető forma tart egyensúlyt a közvetlenül mérhető állapottal. Ezek funkcionális szerepe jelenleg kérdéses [5] (1. ábra).

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N d_i^2}, \quad (3)$$

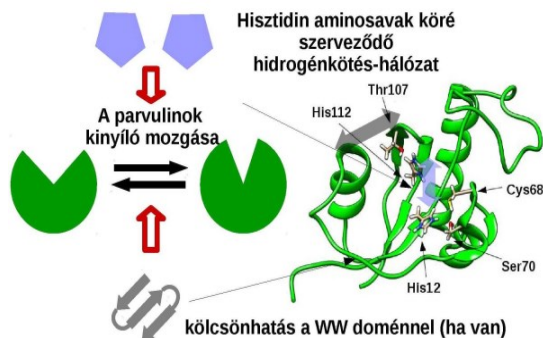
ahol d_i a vizsgált atom távolsága az N darab szerkezetre vonatkozó átlagos pozíciójától az i . szerkezetben.



1. ábra. A PAF gombaellenes fehérje állapotai. A diagramon a lokális RMSD látható, késsel a hideg, pirossal a meleg kitekeredett forma és a szobahőmérsékleten domináns látható állapot különbsége, zölddel a hideg és meleg kitekeredett formák közötti eltérés.

3.2. Parvulin típusú peptidil-prolil izomerázok

Egy újabb, saját kutatásainkból vett példa az ún. parvulin típusú peptidil-prolil izomerázok (fehérjemolekulákban adott molekularészletnél lejátszódó szerkezeti átalakulást segítő fehérjék) összehasonlító vizsgálatából származik. Ezen izomerázok közül egyesek egyrészt segítik más fehérjék szerkezetének megfelelő kialakulását a sejten belül, mások fontos genetikai szabályozó folyamatokban vesznek részt. Munkánk során a molekulacsalád három különböző tagjának gyors, $10^{-12} - 10^{-9}$ másodperces időskálájú dinamikát tükröző szerkezeti sokaságait állítottuk elő, majd elvégeztük ezek összehasonlító elemzését. Megállapítottuk, hogy a vizsgált fehérjék közötti legfontosabb különbség a szerkezet „kinyíló” mozgásának mértékében van. Ez a kinyíló mozgás azon régiót érinti, ahol a parvulinok a partnerfehérjéket megkötik, és azok szerkezeti átalakulását, izomerizációját katalizálják, segítik. Részletesebb vizsgálatunk, többek között a szerkezeti sokaságok főkomponens-elemzése és korábbi irodalmi adatok alapján felállítottunk egy olyan modellt, amely szerint ezen kinyíló mozgás jellemzői hozzájárulnak annak meghatározásához, hogy az egyes parvulinok mely partnerfehérjéket ismerik fel és azok átalakulását milyen hatékonyan segítik. Maga a kinyíló mozgás többféle módon befolyásolható, vagy egyes, minden parvulinban megtalálható aminosavak közötti atomi kölcsönhatások, nevezetesen hidrogénkötések adott parvulinra jellemző létrejöttével, vagy olyan, csak bizonyos parvulinokban meglévő további molekularészlet, ún. WW domén segítségével, melyek szintén felismerik a partner molekulákat [3] (2. ábra).



2. ábra. A parvulinok kötőszobének kinyíló mozgása és az azt szabályozó faktorkok, balra vázlatosan, jobb oldalon egy konkrét parvulinon bemutatva, kiemelve a konzervált hisztidin aminosavak köré szerveződő hidrogénkötés-hálózatot.

Köszönetnyilvánítás

A kutatásokat az NKFI Hivatal támogatta az 104198. számú pályázat segítségével.

Hivatkozások

- [1] ÁNGYÁN, A. F., SZAPPANOS, B., PERCZEL, A., AND GÁSPÁRI, Z.: *CoNSEnsX: an ensemble view of protein structures and NMR-derived experimental data*, BMC Struct. Biol., Vol. **10**, p. 39 (2010). DOI: [10.1186/1472-6807-10-39](https://doi.org/10.1186/1472-6807-10-39)
- [2] ÁNGYÁN, A. F., AND GÁSPÁRI, Z.: *Ensemble-based interpretations of NMR structural data to describe protein internal dynamics*, Molecules, Vol. **18** No. **9**, pp. 10548-10567 (2013). DOI: [10.3390/molecules180910548](https://doi.org/10.3390/molecules180910548)
- [3] CZAJLIK, A., KOVÁCS, B., PERMI P., AND GÁSPÁRI, Z.: *Fine-tuning the extent and dynamics of binding cleft opening as a potential general regulatory mechanism in parvulin-type peptidyl prolyl isomerases*, Sci. Rep. Vol. **7**. art. No. **44504** (2017). DOI: [10.1038/srep44504](https://doi.org/10.1038/srep44504)
- [4] HENRY, E. R., AND SZABÓ, A.: *Influence of vibrational motion on solid state line shapes and NMR relaxation*, J. Chem. Phys., Vol. **82** No. **11**, pp. 4753-4761 (1985). DOI: [10.1063/1.448692](https://doi.org/10.1063/1.448692)
- [5] FIZIL, Á., GÁSPÁRI, Z., BARNÁ, T., MARX F., AND BATTÁ, G.: *"Invisible" conformers of an antifungal disulfide protein revealed by constrained cold and heat unfolding, CEST-NMR experiments and molecular dynamics calculations*, Chem. Eur. J., Vol. **21** No. **13**, pp. 5136-5144 (2015). DOI: [10.1002/chem.201404879](https://doi.org/10.1002/chem.201404879)

- [6] RICHTER, B., GSPONER, J., VÁRNAI, P., SALVATELLA X., AND VENDRUSCOLO, M.: *The MUMO (minimal under-restraining minimal over-restraining) method for the determination of native state ensembles of proteins*, J. Biomol. NMR, Vol. **37**, pp. 117-135 (2007). DOI: [10.1007/s10858-006-9117-7](https://doi.org/10.1007/s10858-006-9117-7)
- [7] VÉRTÉSSY, B., AND OROSZ, F.: *From “fluctuation fit” to “conformational selection”*: evolution, rediscovery, and integration of a concept, BioEssays, Vol. **33** No **1**, pp. 30-34 (2011). DOI: [10.1002/bies.201000068](https://doi.org/10.1002/bies.201000068)
- [8] VRANKEN, W. F.: *NMR structure validation in relation to dynamics and structure determination*, Prog. Nucl. Magn. Reson. Spectrosc., Vol. **82**, pp. 27-38 (2014). DOI: [10.1016/j.pnmrs.2014.08.001](https://doi.org/10.1016/j.pnmrs.2014.08.001)



Gáspári Zoltán 1976-ban született. PhD fokozatát 2004-ben szerezte meg az Eötvös Loránd Tudományegyetemen, szerkezeti biológia szakterületen Perczel András témavezetésével. 2018-ban habilitált a Pázmány Péter Katolikus Egyetemen. 2011-ig tanársegéd, illetve adjunktus az ELTE TTK-n, jelenleg a PPKE ITK docense. Szakterülete a fehérjemolekulák tér szerkezetének és belső dinamikájának vizsgálata NMR-spektroszkópiai és számítási módszerekkel, elsősorban a szinaptikus jelátvitelben részt vevő egyes molekulákra koncentrálva. Pályája során rövid kutatói ösztöndíjak segítségével látogatta meg a Helsinki és a Cambridge-i Egyetemet, valamint két alkalommal nyerte el a Bolyai János Kutatói Ösztöndíjat.

GÁSPÁRI ZOLTÁN

Pázmány Péter Katolikus Egyetem
 Információs Technológiai és Bionikai Kar
 1083 Budapest, Práter u. 50./A
gaspari.zoltan@itk.ppke.hu

PROTEIN STRUCTURAL ENSEMBLES REFLECTING INTERNAL DYNAMICS: GENERATION, EVALUATION AND BIOMEDICAL RELEVANCE

ZOLTÁN GÁSPÁRI

Protein molecules are dynamic by nature and their internal dynamics is considered as a key factor determining and fine-tuning their mode of action. Motions in the microsecond-millisecond regime are generally considered as governing molecular recognition by changing the shape and surface properties of proteins, whereas dynamics on the picosecond-nanosecond time scale is

primarily thought to contribute to conformational entropy. Getting a detailed, atomic-level description of these motions is a nontrivial task. Currently, the best general approach is to combine molecular mechanics with experimental data obtained for the given protein. The resulting representation is an ensemble of protein conformations whose diversity reflects the motions of the protein occurring at a given time scale[1]. Such ensembles can be generated by molecular dynamics calculations directly restrained with experimental data or by selecting a suitable subset of a large pre-calculated pool of conformations. The resulting ensemble should be carefully evaluated for compliance with experimental parameters with special emphasis on those not included in the generation of the ensemble (cross-validation). Only then can the ensemble be analyzed in terms of biological relevance, whether it can be used to get closer to the understanding of mechanistic details of protein action. Application of the methods on the antifungal protein PAF and parvulin-type peptidyl-proline isomerases are summarized briefly.

Keywords: applied mathematics, molecular dynamics, protein structure.

Mathematics Subject Classification (2000): 92E10,70E99,74E99.