

DOKTORI ÉRTEKEZÉS

Hudáky Ilona

2009

FEHÉRJESZAKASZOK ÉS PEPTIDEK TÉRSZERKEZETI ELEMZÉSE

Doktori értekezés

HUDÁKY ILONA

Témavezető: Dr. Perczel András D.Sc. tanszékvezető egyetemi tanár

Kémia Doktori Iskola

Vezető: Dr. Inzelt György D.Sc. egyetemi tanár

Elméleti és fizikai kémia, anyagszerkezetkutatás program

Vezető: Dr. Surján Péter D.Sc. egyetemi tanár

Eötvös Loránd Tudományegyetem

Kémiai Intézet

2009

TARTALOMJEGYZÉK

KÖSZÖNETNYILVÁNÍTÁS	3
KÖZLEMÉNYEK JEGYZÉKE	4
RÖVIDÍTÉSEK JEGYZÉKE	6
BEVEZETŐ	7
Peptidszakaszok térszerkezeti nevezékrendszere	7
Peptidmodellek konformációanalízise gázfázisban és oldószer modellezésével	9
A további vizsgált peptidmodellek térszerkezeti és biológiai jelentősége	10
<i>Prolin</i>	10
<i>Prolil-prolin</i>	11
<i>Vicinális diszulfidhidas ciszteinil-cisztein</i>	16
MÓDSZEREK	18
Peptidek konformációs kódja és nevezékrendszere	18
Konformerek kvantumkémiai megközelítése	25
Fehérjék aminosavainak természetes konformációs összetétele és a modellpeptid becült konformációs összetétele	26
Pszedorotációs paraméterek számítása prolin pirrolidingyűrűjére	27
β -Kanyarok osztályozása	30
Hidrogénhid-kötések jellemzése	31
EREDMÉNYEK	33
Alanin diamid stabilitásvizsgálata az oldószer modellezésével	33
<i>Ramachandran-felshínek és konformerstabilitás eltérő elméleti szinteken</i>	33
<i>Az eltérő elméleti szintű Ramachandran-felshínek összevetése</i>	37
<i>Frekvenciaszámítások</i>	38
<i>Konformerek számított stabilitásadatainak korrelációja fehérjék alaninjainak konformációs változatosságával</i>	40
<i>Az oldószer hatásának modellezése</i>	42

Prolin	43
<i>Kísérleti adatok statisztikai elemzése</i>	43
<i>Ab initio számítások</i>	48
<i>Összevetés</i>	52
Prolil-prolin	55
<i>Ab initio, DFT és PCM számítások</i>	55
<i>Peptidszekvenciák statisztikai elemzése</i>	74
<i>Fehérjeszekvenciák statisztikai elemzése</i>	76
<i>Összevetés</i>	84
<u><i>Hosszított szerkezetek</i></u>	86
<u><i>β-kanyarok</i></u>	87
<u><i>Gyűrűállás</i></u>	89
Vicinális diszulfidhidas ciszteinil-cisztein	90
<i>A 4,5-ditiaheptano-7-laktám gyűrűmodell ab initio és DFT számításai</i>	90
<i>A HCO-ox-[L-Cys-L-Cys]-NH₂ peptidmodellre vonatkozó ab initio és DFT számítások</i>	93
<i>Az Xxx-Cys-Cys -Yyy és Xxx-ox-[Cys-Cys]-Yyy peptidszekvenciák adatbázisokból nyert térszerkezeti adatainak elemzése</i>	98
<i>Összevetés</i>	102
ÖSSZEFOGLALÁS	107
Peptidek konformációs nevezékrendszere	107
Alanin diamid stabilitásvizsgálata az oldószer modellezésével	107
Prolin	108
Prolil-prolin	108
Vicinális diszulfidhidas ciszteinil-cisztein	109
IRODALMI HIVATKOZÁSOK	111

KÖSZÖNETNYILVÁNÍTÁS

Szeretetteljes aggódó figyelmét köszönöm gimnáziumi kémia tanárnőmnek, Dr. Nagy Zsigmondnének, aki a kémia felé fordította érdeklődésemet és felkészített a tanulmányi versenyekre, hiszen nélküle életem fontos találkozásai elmaradtak volna.

Hálámat fejezem ki témavezetőmnek, Dr. Perczel Andrásnak, hogy az évek hosszú során változó élethelyzeteimben a kutatói munkám feltételeit megteremtette és azt irányította. Dr. Csizmadia G. Imrénének köszönöm a tudomány iránti lelkesedés lankadatlan sugárzását és az első publikációm születése körüli „bábáskodást”. Hálás vagyok Dr. Gráf Lászlónak, hogy egy évig az MTA-ELTE Biotechnológiai Kutatócsoport tagjaként dolgozhattam.

Dr. Farkas Ödönnek köszönöm a kémiai számításokról folytatott számos sokrétű és hasznos beszélgetést, valamint tanácsait a Gaussian programcsomag használatával kapcsolatban. Köszönöm Dr. Gáspári Zoltánnak a szekvenciák kigyűjtését fehérjeszerkezeti adatbankokból és a vicinális diszulfidhidak elemzése terén nyújtott segítségét. Hálával tartozom Dr. Jákli Imrénének a munkám eredményes elvégzéséhez szükséges számítástechnikai háttér kiépítéséért, Dr. Harmat Veronikának pedig a peptidszekvenciák kismolekulás adatbázisból történő kigyűjtésében nyújtott segítségéért. A családi munkalétkörért hálás köszönet illeti minden kedves munkatársamat, így különösen Füzéry Annát, Dr. Czajlik Andrást, Dr. Beke Tamást, Láng Andrást, Pálfi Villót, Kiss Róbertet és Dr. Bodor Andreát.

Végezetül köszönöm férjemnek, Dr. Hudáky Péternek, mindkettőnk szüleinek, valamint Márton, Katinka, Benedek és Vince nevű gyermekeinknek, hogy a doktori iskolai feladataim végzését segítették, támogatták és elviselték.

KÖZLEMÉNYEK JEGYZÉKE

(impakt faktor a megjelenés évében / független idéző / összes idéző 2008.12.7-ig)

A doktori dolgozat alapját képező közlemények

- (I) **Hudáky I**, Baldoni HA, Perczel A. Peptide Models XXXVIII. Proline Conformers from X-ray crystallographic database and from ab initio computations. *J Mol Struct (THEOCHEM)* 2002;582:233-249. (1.014 / 11 / 19)
- (II) **Hudáky I**, Perczel A. Ab initio analysis of the conformational changes of the prolyl-proline model peptide. *J Mol Struct (THEOCHEM)* 2003;630:135-140. (1.027 / 9 / 11)
- (III) **Hudáky I**, Kiss R, Perczel A. A nomenclature of peptide conformers. *J Mol Struct (THEOCHEM)* 2004;675:177-183. (1.007 / 7 / 9)
- (IV) **Hudáky I**; Gáspári Z; Carugo O; Čemažar M; Pongor S; Perczel A. Vicinal disulfide bridge conformers by experimental methods and by ab initio and DFT molecular computations. *Proteins-Structure Function and Bioinformatics* 2004;55(1):152-168. (4.429 / 5 / 8)
- (V) **Hudáky I**; Hudáky P; Perczel A. Solvation model induced structural changes in peptides. A quantum chemical study on Ramachandran surfaces and conformers of alanine diamide using the polarizable continuum model. *J Comp Chem* 2004;25(12):1522-1531. (3.168 / 10 / 11)
- (VI) **Hudáky I**; Perczel A. Prolylproline unit in model peptides and in fragments from databases. *Proteins-Structure Function and Bioinformatics* 2008;70:1389-1407.

A doktori dolgozat témájához kapcsolódó további közlemények

- (VII) **Hudáky I**, Perczel A. Ab initio conformational analysis of the prolyl-proline peptides and their potential to modify the main-chain fold of proteins. In *Peptides 2000 (Proceedings of the Twenty-Sixth European Peptide Symposium)*, ed. by Martinez J and Fehrentz JA, EDK, Paris, France, 2001;417-418.
- (VIII) **Hudáky I**, Perczel A. Ab initio conformational analysis of the cis-trans isomerization of the prolyl-proline peptide. In *Peptides 2002 (Proceedings of the Twenty-Seventh European Peptide Symposium)*, ed. by Benedetti E and Pedone C, Edizioni Ziino, Napoli, Italy, 2003;850-851.

- (IX) Carugo O; Čemažar M; Zahariev S; **Hudáky I**; Gáspári Z; Perczel A; Pongor S. Vicinal disulfide turns. *Protein Engineering* 2003;16(9):637-639.
- (X) Gáspári Z; **Hudáky I**; Czajlik A; Perczel A. Is there an excuse for the non-conformist? Notes on the calculated energies, atom-atom contacts and natural abundance of the different conformers of alanine in proteins. *J Mol Struct (THEOCHEM)* 2004;675:141-148.

Egyéb közlemények

- (XI) Hudáky P; **Hudáky I**; Perczel A. Peptide models XXXV. Protonated and deprotonated N-Formyl-L-histidinamide: an ab initio study on side-chain potential energy surfaces of all major backbone conformers. *J Mol Struct (THEOCHEM)* 2002;583:199-213.
- (XII) Bágyi I; Balogh B; Czajlik A; Élias O; Gáspári Z; Gergely V; **Hudáky I**; Hudáky P; Kalászi A; Károlyházy L; Keserű K; Kiss R; Krajsovszky G; Láng B; Nagy T; Rác A; Szentesi A; Tábi T; Tapolcsányi P; Vaik J; Koo JCP; Chass GA; Farkas Ö; Perczel A; Mátyus P. Generation and analysis of the conformational potential energy surfaces of N-acetyl-N-methyl-L-alanine-N'-methylamide. An exploratory ab initio study. *J Mol Struct (THEOCHEM)* 2003;625:121-136.
- (XIII) Lam JSW, Koo JCP, **Hudáky I**, Varro A, Papp JG, Penke B, Csizmadia IG. Predicting the conformational preferences of N-acetyl-4-hydroxy-L-proline-N'-methylamide from the proline residue. *J Mol Struct (THEOCHEM)* 2003;666-667:285-289.

RÖVIDÍTÉSEK JEGYZÉKE

Boc–	(CH ₃) ₃ C–O–CO–, terc-butiloxikarbonil-csoport
CSD	Cambridge Structural Database
DFT	Density Functional Theory
For	HCO–, formilcsoport
Hyp	hidroxiprolin
I-PCM	Isodensity Polarizable Continuum Model
PCM	Polarizable Continuum Model
PDB	Protein Data Bank
pro	D-prolin
RHF	restricted Hartree-Fock
SCIPCM	Self-Consistent Isodensity Polarizable Continuum Model
TS	átmeneti állapot (transition state)
Xxx, Yyy	tetszőleges aminosavak

BEVEZETŐ

Doktori munkám során olyan biológiai szempontból is fontos fehérje- és peptidszakaszok adatbázisokban közzétett térszerkezetének elemzését és kvantumkémiai modellezését végeztem el, melyek gyűrűs oldalláncaik által „mozgásukban korlátozottak”. A választott modellek biológiai jelentőségét példázza jelenlétük kötőszöveink kollagénjében vagy az ingerületátvitelben szerepet játszó receptorokban. Vizsgálataim szorosan támaszkodnak a kutatócsoportunk által néhány aminosav diamidján, ill. a formil-alanil-alaninamidon elvégzett konformációs elemzésekre, ill. az aminosavak gerincállásának jellemzésére bevezetett nevezéktanra. Minden esetben az aminosavak természetben gyakoribb L konfigurációját tanulmányoztam, ezt hangsúlyozom is a modellvegyületek nevében, de elhanyagolom kiírását a fehérjék szekvenciáiban. A modellpeptidek konformációs térének egészét vizsgáltam, így a minimumok összegyűjtésekor tekintettel voltam az oldallác eltérő állásainak és a *cis* amidnak a lehetőségeire is. Az ilyen alkalmazott kvantumkémiai számításokon alapuló kutatások képesek egy-egy „építőelem” részletes térszerkezeti jellemzésére, különböző elméleti szintű megközelítések alkalmazhatóságának tesztelésére, de az eltérő kísérleti módszerek megbízhatóságának összevetésére is.

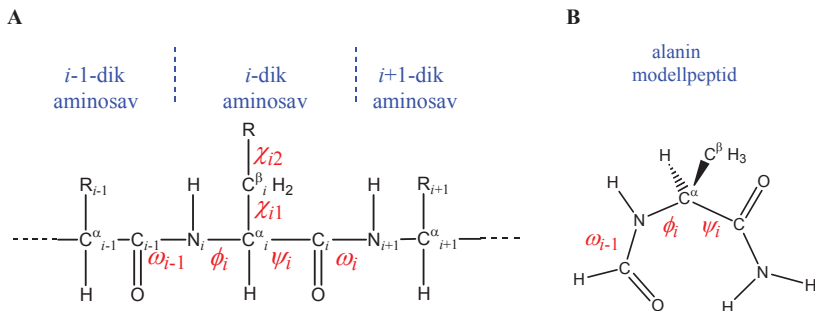
A doktori értekezés alapjául szolgáló hat publikációm 2002 és 2008 között jelent meg. A peptidek térszerkezeti nevezékkendszerét taglaló cikk¹ arra kívánta felhívni a figyelmet, hogy az elnevezések egységesítése segítené a különböző tudományterületek közötti együttműködést. A prolin,² a prolil-prolin³ és a vicinális diszulfidhidas ciszteinil-cisztein⁴ konformereit ismertető közleményeimben adatbázisból nyert kísérleti adatokat vettem össze az első esetben csak gázfázisú *ab initio* adatokkal, a két későbbiben már oldószeri is modellező számításokkal. Az oldószer figyelembevételének módját és hasznosságát az alanin példáján tanulmányoztam és 2004-ben publikáltam.⁵ A korábban megjelentetett munkáimat^{2,6} azonban a doktori értekezésben sem egészítettem ki oldószermodelles adatokkal.

Peptidszakaszok térszerkezeti nevezékkiszere

Szükségessé vált a peptidek olyan térszerkezeti nevezékkiszereének kidolgozása, amely egyaránt alkalmas egyszerű diamidok és hosszabb fehérjeláncok megjelölésére. Térszerkezeti adatok kezelése mindennapos tevékenységgé vált a biológiában, a biokémiában,

a kémia kísérleti és elméleti ágaiban. A gyors és tömör, ám ugyanakkor átlátható térszerkezeti leírásokhoz ezen tudományok mindegyikében kifejlődött egy-egy sajátos szakzsargon, melyek bár szilárd alapokkal rendelkeznek a megfelelő szakirodalomban, egymástól mégis a kelletténél jobban eltérnek. Kizárólag olyan alternatív nevezék jósolható életképesnek, amely átlátható, magától értetődő, szükség szerint fejleszthető, kiterjeszhető vagy akár egyszerűsíthető is. Egy ideális nómenklatúra elősegíthetné a fehérjék és peptidek változatos megközelítéseinek együttműködését.

Javaslatot teszünk egy olyan nevezékrendszerre, mely véleményünk szerint eleget tesz az egyszerűség és az egyértelműség követelményeinek. Itt és a továbbiakban is az aminosav kifejezés N- és C-terminálisán is amiddá alakított α -aminokarbonsavat jelent függetlenül attól, hogy kísérleti vagy elméleti módszerekkel tanulmányozták. A rövidség kedvéért peptidnek tekintünk egy tetszőleges hosszúságú peptidláncot a diamidtól, az oligopeptiden át a fehérjék polipeptidláncáig. A gerincet a ϕ és a ψ torziók határozzák meg (ld. **B.1. Ábra**), habár tágabb értelemben az ω szög is hatással van a gerincszerkezetre. Hasonlóan ahhoz, ahogy a peptidek kémiai azonosítására, primer szerkezetük megadására az N-terminálistól a C-terminálisig felsoroljuk az aminosavakat egybetűs vagy hárombetűs kóddal, esetleg teljes névvel, a térszerkezet ugyancsak jelölhető az egyes aminosavak egyedi konformációs kódjának láncolatával. A nevezéktan kialakításakor figyelembe vettük a vonatkozó IUPAC ajánlásokat.⁷



B.1. Ábra Polipeptid és peptidmodell térszerkezeti legfontosabb torziósögei. **A:** Fehérje vagy peptid szekvenciájának i -dik aminosavát jellemző torziósögek: ω_{i-1} ($C^{\alpha}_{i-1}-C_{i-1}-N_i-C^{\alpha}_i$) a peptidkötést jellemzi, ϕ_i ($C_{i-1}-N_i-C^{\alpha}_i-C_i$) és ψ_i ($N_i-C^{\alpha}_i-C_i-N_{i+1}$) a gerincet, míg χ_{i1} ($N_i-C^{\alpha}_i-C_i-C^{\beta}_i$) és a további torziók az oldallánc állását írják le. **B:** A HCO-Ala-NH₂ modellpeptidet jellemző torziósögek. Az ábra hangsúlyozza az L konfigurációjú aminosav C^α kiralitáscentruma körüli térállást.

Peptidmodellek konformációanalízise gázfázisban és oldószer modellezésével

A fehérjeszerkezetek leírásában a ϕ és a ψ torzószög, valamint az általuk definiált Ramachandran-térkép fontossága már az 1960-as években ismertté vált.^{8,9} Azóta sem hanyatlak az érdeklődés és kísérletileg meghatározható Ramachandran-diagramok és a számított Ramachandran-felületek összevetése iránt.¹⁰ Miközben egy fehérje másodlagos szerkezetét jól jellemzi az, hogy egyes szakaszainak aminosavai a Ramachandran-térkép mely területén helyezkednek el, addig az aminosavmodellek kvantumkémiai módszerekkel kiszámítható Ramachandran-felületeinek alacsony energiájú völgyei megmutatják a lehetséges konformerek geometriáját és várható gyakoriságuk becslésének alapjául is szolgálnak. Az aminosav diamid modellek,¹¹ pl. HCO-Xxx-NH_2 és $\text{H}_3\text{CCO-Xxx-NHCH}_3$, széles körben elterjedtek a peptidek és fehérjék építőelemeinek jellemzésére. Az alanin diamid¹² különös figyelemben részesül, mivel ez a királis aminosavak legegyszerűbb modellvegyülete. A HCO-L-Ala-NH_2 gázfázisú Ramachandran-felszíneinek és kritikus pontjaik átfogó elemzése jelentős korrelációt mutatott egyes eltérő elméleti szintek között.¹³

Az aminosav diamid modellek gázfázisban számított adatait terhelő, a kísérletileg meghatározott szerkezetű fehérjék megfelelő adatainak tükrében fellépő legnagyobb probléma az, hogy ilyen kis modellek rendszerint nem adnak számot két kísérletileg rendkívül gyakori konformerről, nevezetesen a jobbmenetes α -hélixről és a poliprolin II-ről. Az előfordulási adatok közepes korrelációt mutatnak a számított energiaértékekkel.¹⁴ Jobb az egyezés olyankor, ha (a) a modellvegyület néhány aminosavegységből álló oligopeptid,¹⁵ (b) a kvantumkémiai számítás figyelembe veszi az oldószerhatást,¹⁶ vagy e két feltétel mindegyike teljesül.^{17,18}

Megrajzolták a HCO-L-Ala-NH_2 modellpeptid potenciális energiefelületét változatos oldószerre olyan módon, hogy az RHF/6-31G(d) elméleti szinten kapott geometriákra a molekulát körülölelő oldószer polarizálható közegként figyelembe vevő PCM ill. I-PCM módszerrel¹⁹ számoltak energiát („single point” energia) RHF/6-311++G(2d,2p) elméleti szinten.²⁰ Jelentős előrelépés a gázfázisú számításokhoz képest az, hogy a PCM és I-PCM „single point” felszínek az α -hélix és a poliprolin II régiókhöz is (α_L ill. ϵ_L) alacsony energiát rendelnek, így jelentősen javítva az egyezést a fehérjékben fellelhető alaninok konformációs gyakoriságával. Másrészt viszont a PCM és I-PCM „single point” Ramachandran-felszínek 19 ill. 17 energiaminimum létezését jósolják. Ilyen nagy gyakoriság a gázfázisú számításoknál sem tapasztalható,^{12,13} és helyzetük sem igazolható olyan „konformációs vonzasközpontként”,

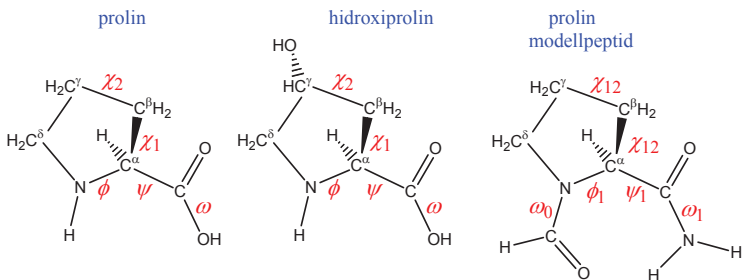
melyeket krisztallográfiai úton meghatározott szerkezetű fehérjék alaninjainak Ramachandran-térképén mutattak ki.²¹ Valódi relaxált szerkezetek igazolása a geometria optimalását és lehetőség szerint a második deriváltak vizsgálatát igényli.

A frekvencia-számításokat általánosan használják annak igazolására, hogy egy optimált szerkezet valódi minimum, tekintettel arra, hogy egy minimumnak a legalacsonyabb frekvenciája is pozitív. Továbbmenve, a kiszámított frekvenciák össze is vethetők kísérleti értékekkel, bár ekkor rendszerint skálafaktorok bevezetése válik szükségessé.^{22,23,24} DFT vibrációs vizsgálatok azonban skálázás nélkül is jó egyezést mutattak kísérleti adatokkal.^{25,26} A fehérje IR frekvenciák, különösen is az amid I sáv függése a másodlagos szerkezettől jelentős irodalommal rendelkezik.²⁷ Mindazonáltal viszonylag friss eredmény az alanin modellpeptid teljes konformációs könyvtárára elvégzett kvantumkémiai frekvenciaszámítás.²⁸

A további vizsgált peptidmodellek térszerkezeti és biológiai jelentősége

Prolin

Az α -aminosavak között a prolin és a hidroxiprolin különlegességét az adja, hogy szekunder aminocsoporttal, tehát iminocsoporttal rendelkeznek. Fehérjékben, peptidekben így a gerinchez az oldallánc nem csak az α -szénatomon, hanem a nitrogénen keresztül is kapcsolódik (ld. **B.2. Ábra**). Mindennek a szerkezeti következményei: (a) kisebb energiakülönbség a prolint/hidroxiprolint megelőző peptidcsoport *cisz* és *transz* helyzete között, (b) a gerinc ϕ torziószöge *gauche*- állásban kötött, (c) az oldallánc öttagú gyűrűben rögzül. A peptidcsoport erős hajlandóságot mutat *cisz* állás felvételére makromolekulákban prolint megelőzően: 6-9 % fehérjékben, 33 % rövidebb polipeptidekben.^{29,30} A ϕ torzió kötöttsége miatt a Ramachandran-térkép nagyobb része a prolin által nem érhető el. Míg három telített szénatom nyílt lánc esetén $3 \times 3 = 9$ oldallánckonformer létét engedné meg, a pirrolidingyűrűnél mindössze két állás különböztethető meg. A prolin és a hidroxiprolin tehát az aminosavak között nemcsak egyedi szerkezeti kötöttségekkel, hanem sajátos konformációs preferenciával is rendelkezik.



B.2. Ábra Az L-prolin, az L-hidroxiprolin és a HCO-L-Pro-NH₂ modellpeptid.

Protil-prolin

Biológiailag aktív peptidek és fehérjék gyakran tartalmazzák a protil-prolin szekvenciárészletet vagy annak hidroxilált változatát. Amint az a későbbiekben bizonyítást nyer, a **Pro–Pro** szekvencia vagy egy hosszított szerkezetet vesz fel, vagy egy β -kanyart formál. E két határozottan elkülönülő kategórián belül is változatos konformációk azonosíthatók a torziószögek alapján. A β -kanyarok osztályozása hosszú múltra tekint vissza.^{31,4} A protil-prolin I, III, VIa és VIb típusú β -kanyart képezhet. Hangsúlyozandó, hogy a prolin aminosavra nem illeszthető a legelterjedtebb nyújtott („extended”) szerkezet, amelyben ϕ_i és ψ_i torziószög egyaránt *anti* helyzetben áll (vö. **M.5/C Ábra**). Bevezettük a „hosszított” („elongated”) kifejezést.³ A **Pro–Pro** hosszított alkatainak csak egy, bár igen gyakori megvalósulása a poliprolin II, ahol az aminosavak ψ_i torziószögei rendre *anti* állásúak. A **B.1. Táblázatban** felsorolt makromolekulák sokrétű biológiai aktivitással rendelkeznek, miközben jól érzékeltetik a kettő vagy több prolinból ill. hidroxiprolinból álló szekvenciárészlet konformációs változatosságát is.

Szomszédos prolinokat tartalmaz néhány természetes eredetű, immungyengítő hatású ciklopeptid, így az antamanid,³² a ciklinopeptid A,³³ a cikloleonorinin³⁴ és a cikloleonoripeptid D.³⁵ E ciklopeptidek jellemzően 9-12 aminosavból állnak, a **Pro–Pro** részletük pedig VIa vagy VIb típusú β -kanyart képez, ahol a közbülső peptidkötés *cis* helyzetű.

B.1. Táblázat Néhány kísérleti úton meghatározott térszerkezetű $Xxx-(Pro)_n$ -Yyy szekvenciárészlet. Az első négy példa természetes ciklopeptidekből származik, a továbbiak a PDB-ben ismertett lineáris polipeptidekből. Feltüntettem a polipeptid nevét, eredetét, a szerkezet meghatározásának módját, az irodalmi hivatkozást, PDB azonosító kódot (ha van), a szekvenciát Xxx és Yyy kifejtésével (n : a közvetlenül kapcsolódó prolinok száma), a $(Pro)_n$ szakasz konformációs kódját (ld. **M.5. Ábra** az aminosavként 4 karaktert tartalmazó kód magyarázatául) és másodlagos szerkezetének hagyományos besorolását.

Polipeptid	Polipeptid eredete	Szerek. meghat. (felbontás)	Ref	PDB kód	Xxx-(Pro) _n -Yyy	(Pro) _n konformer kód	(Pro) _n peptidszakasz osztályozása
Antamid	gyilkos galóca (<i>Amanita phalloides</i>)	Röntgen	32		Phe- Pro-Pro -Phe Val- Pro-Pro -Ala	$\beta_1 + \alpha_1 -$ $\beta_1 + \alpha_1 +$	Vib β -kanyar Vib β -kanyar
Cikloleuropeptid A	lenmag	Röntgen, NMR	33		Val- Pro-Pro -Phe	$\beta_1 + \alpha_1 +$	Vla β -kanyar
Cikloleonorinin	gyöngyajak gyümölcse	NMR	34		Tyr- Pro-Pro -Tyr	β_1, α_1	Vla β -kanyar
Cikloleuropeptid D	gyöngyajak gyümölcse (<i>Leonurus heterophyllus</i>)	NMR, Röntgen	35		Ser- Pro-Pro-Pro -Tyr	$\beta_1 + \beta_1 + \alpha_1 +$	poliprolin II és más hosszított
α -Bungarotoxin	bungarkigyó mérge (<i>Bungarus multicinctus</i>)	Röntgen 2,5 Å	38	2abx	Cys- Pro-Pro -Gly His- Pro-Pro -Lys	β_1, β_1 β_1, β_1	poliprolin II poliprolin II
		NMR	37	1abt	Cys- Pro-Pro -Gly	$\alpha_1 - \beta_1 -$	nem definiált β -kanyar
		NMR	39	114w	Cys- Pro-Pro -Gly	$\alpha_1 - \beta_1 +$	nem definiált β -kanyar
		NMR	39	1ljz	Cys- Pro-Pro -Gly	$\alpha_1 - \beta_1 -$	nem definiált β -kanyar
Fasciculin I	zöld mamba kigyó mérge (<i>Dendroaspis angusticeps</i>)	Röntgen 1,9 Å	40	1fas	His- Pro-Pro -Lys Cys- Pro-Pro -Gly	$\beta_1 - \alpha_1 +$ $\beta_1 + \beta_1 +$	Vib β -kanyar poliprolin II
Toxin γ	kobraméreg (<i>Naja nigricollis</i>)	Röntgen 1,55 Å	41	1tqx	Ile- Pro-Pro -Phe	$\beta_1 - \alpha_1 +$	Vla β -kanyar
Kardiotoxin CTX IIb	kobraméreg (<i>Naja mossambica</i>)	NMR	42	2ccx	Ile- Pro-Pro -Phe	$\beta_1 + \gamma_1 +$	hosszított, γ -kanyarral
Toxin α	kobraméreg (<i>Naja nigricollis</i>)	NMR	43	1nea	Gln- Pro-Pro -Thr	$\beta_1 + \beta_1 +$	poliprolin II
Konotoxin ψ -PIIE	kúpsiga mérge (<i>C. purpurascens</i>)	NMR	44	1as5	His- Hyp-Hyp -Cys	$\beta_1 - \beta_1 -$	poliprolin II
Konotoxin ψ -PIIF	kúpsiga mérge (<i>C. purpurascens</i>)	NMR	45	1jlp	Gly- Hyp-Hyp -Cys	$\beta_1 - \beta_1 -$	poliprolin II
Konotoxin GS	kúpsiga mérge (<i>C. geographus</i>)	NMR	46	1ag7	Cys- Hyp-Hyp -Gln	$\alpha_1 - \beta_1 +$	nem definiált β -kanyar
Konotoxin TVIIA	kúpsiga mérge (<i>Conus tulipa</i>)	NMR	47	1eyo	Cys- Hyp-Hyp -Val	$\alpha_1 - \alpha_1 -$	nem definiált β -kanyar
μ -Konotoxin GIIIA	kúpsiga mérge (<i>C. geographus</i>)	NMR	48	1tgc	Thr- Hyp-Hyp -Lys	$\beta_1 + \alpha_1 +$	Vib β -kanyar
μ -Konotoxin GIIIB	kúpsiga mérge (<i>C. geographus</i>)	NMR	49	1gib	Thr- Hyp-Hyp -Arg	$\beta_1 + \alpha_1 +$	Vib β -kanyar
Trombin	emberi	Röntgen 3 Å	50	1dwb, ...	Tyr- Pro-Pro -Trp	α_1, α_1	III β -kanyar

(folytatás következik)

(B.1. Táblázat folytatása)

Polipeptid	Polipeptid eredete	Szerk. meghat. módszer	Ref	PDB* kód	Xxx-(Pro) _n -Yyy	(Pro) _n konformer kód	(Pro) _n peptidszakasz osztályozása
Uracil DNS glikoziláz	emberi	Röntgen 1,9 Å	53	1akz	Tyr- Pro-Pro-Pro -His Val-(Pro) _L -Ser	cβ₁-rβ₁+rα₁-rβ₁+rβ₁-rβ₁+rα₁-	hosszított poliprolin II és más
	herpes simplex vírus	Röntgen 1,75 Å	54		Lys- Pro-Pro-Arg Val-(Pro) _L -Ser	cβ₁+rβ₁-rβ₁-rβ₁-rα₁-	hosszított poliprolin II és más
SH3 doménhez kötő szekvenciák	<i>Escherichia coli</i>	Röntgen 1,6 Å	52	4eug	Tyr- Pro-Pro-Gln Ile- Pro-Pro-Ser	cβ₁+rβ₁-rβ₁-rα₁-	hosszított poliprolin II és más
	szintetikus	NMR	56	1klz	Ile-(Pro) _L -Gly	cβ₁+rβ₁-rβ₁-rβ₁+rβ₁+rβ₁-	hosszított poliprolin II
Titin α-Aktinin kollagénszerű modellpeptidek	szintetikus	NMR	57	1aze	Val-(Pro) _J -Val- Pro-Pro-Arg	rβ₁+rβ₁-rβ₁+rβ₁+rβ₁+rβ₁-	és poliprolin II
	szintetikus	Röntgen 1,6 Å	58	1bbz	Ser-(Pro) _L -Pro	rβ₁+rβ₁+rβ₁+rβ₁-	poliprolin II
	emberi	Röntgen 2,44 Å	59	1ya5	Ala- Pro-Pro-Asn	rβ₁+rβ₁-rβ₁-	poliprolin II
	emberi	NMR	60	1h8b	Leu- Pro-Pro-Asp	rβ₁-rα₁-rβ₁-	hosszított poliprolin II
	szintetikus	Röntgen 1,97 Å, 1,6 Å	62	1a3i, 1a3j	ismétlődő Pro-Pro-Gly	főleg rβ₁+rβ₁-rβ₁-	poliprolin II
	szintetikus	Röntgen 2 Å	63	1bkv	ismétlődő Pro-Hyp-Gly	főleg rβ₁+rβ₁-rβ₁-	poliprolin II

Az α -bungarotoxin egy háromujjú α -neurotoxin.^{36,a} A nikotinos acetil-kolin receptort modellező peptidekhez való kötődése igen kedvelt vizsgálati rendszer a ligandumfüggő ioncsatornák kutatásában. NMR-szerkezete³⁷ “több kulcsfontosságú tekintetben” is eltérni látszik röntgen szerkezetétől.³⁸ A szekvenciális különbségeken túlmenően érdekes, hogy az NMR szerint a **Pro–Pro** egy nem definiált β -kanyar, míg a 2,5 Å felbontású röntgen poliprolin II szerkezetet ad. Számos α -neurotoxint felölelt az a szekvenciális illesztés, mely erősen konzervált Cys–**Pro–Pro**–Gly szekvenciát talált szigorúan változatlan ciszteinnel.³⁹ Egy hatékony acetil-kolin észteráz inhibitor, a fasciculin 1,⁴⁰ két eltérő téralkatú **Pro–Pro** szekvenciaegységet tartalmaz. Habár a toxin γ ⁴¹ szekvenciálisan csaknem azonos a kardiotoxin CTX IIB-vel,⁴² hiszen mindössze egyetlen aminosavban térnek el, **Pro–Pro** az előbbiben VIa típusú β -kanyart alkot, míg az utóbbiban hosszított szerkezetet γ -kanyarral. A toxin α hosszított alkatú **Pro–Pro** egysége megtöri az első hurok szabályos kettős β -redőjét.⁴³

A konotoxinok tengeri kúpcsigákból izolált, diszulfidhidridokban gazdag peptidek, és az α -neurotoxinok egy további csoportját képezik.³⁶ Sokukban található hidroxiprolil-hidroxiprolin szekvenciariészlet, amely az NMR vizsgálatok tanúsága szerint változatos amidállásokat mutat. A konotoxin ψ -PIIIE⁴⁴ és ψ -PIIIF⁴⁵ 24 aminosavból álló részlegesen homológ peptidek, melyek N-terminálisán a **Hyp–Hyp** egység kizárólag *transz* amidokkal rendelkező hosszított szerkezet. A belső szekvenciális helyzetek a **Hyp–Hyp** részletet β -kanyarba hajlítják a konotoxin GS⁴⁶ és TVIIA,⁴⁷ továbbá a μ -konotoxin GIIIA⁴⁸ és GIIIB⁴⁹ esetén.

A trombinban a Tyr–**Pro–Pro**–Trp motívum a kimotripszinhez viszonyított 9 aminosavnyi inszerció része, és hozzájárul az inhibitorokkal szemben a többi szerin-proteázétól eltérő érzékenységéhez.⁵⁰ Négy megvizsgált inhibitor bármelyikének jelenlétében a prolil-prolin a trombinban III típusú β -kanyart formál.⁵¹

Szekvenciaillesztéssel nyert bizonyosságot, hogy két konzervatív **Pro–Pro** szakasz is fellelhető az emberi, a herpes simplex vírusból származó, továbbá az *Escherichia coli* eredetű uracil DNS glikozilázban (hUDG, vUDG, ill. eUDG).⁵² Az eUDG-ben⁵² mindkét szakasz

^a three-finger α -neurotoxin

egy-egy β -szalagot köt α -hélixhez. Bár a hUDG-ben⁵³ és a vUDG-ben⁵⁴ is növekszik a prolinok száma, az eUDG számozás szerinti 39-40, ill. 86-87 helyzetben sértetlenül megmarad a kétféle hosszított gerincalkat, nevezetesen a **cE₁fE₁** és a **fE₁fE₁**.

A fehérjék kölcsönhatásában részt vevő domének gyakran tartalmaznak prolinban gazdag szekvenciákat. Ilyenek pl. az SH3 ill. WW doménhez kötő peptidligandumok, melyek két vagy akár több szomszédos prolint is felsorakoztatnak.⁵⁵ A prolinban gazdag szekvenciák rendszerint poliprolin II szerkezetet vesznek fel, vagyis gerincük a **fE₁** konformáció ismétléséből áll. A hidrogénhidakban protonakceptorként működő karbonilcsoportokhoz könnyen hozzáférhető, térszerkezetileg kötött hidrofób oldalláncok társulnak. Az SH3 doménhez kötődő szekvenciák^{56,57,58} között *cisz* peptidkötést tartalmazó is előfordult, miközben a szerkezet hosszított maradt valamennyi prolinpár tekintetében.

A titin, egy gigantikus sokdoménes izomfehérje, és az α -aktinin, egy Ca^{2+} -független EF-kéz fehérje, a szarkomer esszenciális és kölcsönható összetevői. Mindkettő hosszított prolil-prolint tartalmaz.^{59,60}

A vázfehérje kollagén prolil-hidroxiprolin egységek tömegeivel rendelkezik, miközben a prolint ill. a hidroxiprolint megelőző peptidkötések *cisz-transz* izomerizációja sebességmeghatározó lépés a fehérjeszerkezet kialakulásában.⁶¹ Ismétlődő **Pro-Pro-Gly**⁶² és **Pro-Hyp-Gly**⁶³ szekvenciát tartalmazó kollagénszerű modellpeptideken végzett röntgen kristallográfiai vizsgálatok a kollagén leggyakoribb prolil-prolin konformációjaként a **fE₁+fE₁**-szerkezetet mutatták ki, tehát egy olyan poliprolin II egységet, melyben a két pirrolidingyűrű ellentétes állású.

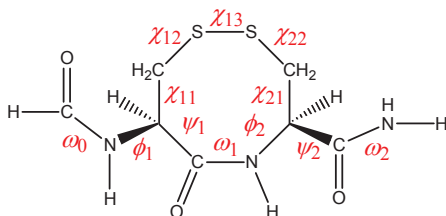
Említést érdemel továbbá, hogy a prolin rendezetlenségre hajlamosítja a fehérjét, ez az aminosav kedveli leginkább a mozgékony kanyarulatokat („hot loops”).⁶⁴ A rendezetlen, mozgékony régiók térszerkezeti jellemzése komoly kihívás, ezért rendszerint be sem kerülnek a PDB-be, vagy részlegesen megbízható adatokkal terheltek. A rendezetlen fehérjék és domének mindazonáltal számtalan változatos élettani feladatot látnak el.⁶⁵

A kvantumkémiail vizsgálatok olyan konformációs és energetikai adatokkal szolgálnak, melyek megkerülhetetlenek a makromolekulák prolil-prolin egységeinek szerkezetmeghatározásánál. A számítások a hosszított szerkezeteknek és β -kanyaroknak is többféle konformerét adják meg becült gyakoriságukkal együtt.

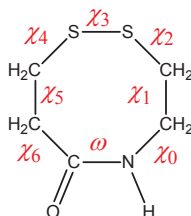
Vicinális diszulfidhidas ciszteinil-cisztein

Diszulfidhidak rendszerint szekvenciálisan távoli polipeptidrészeket kötnek össze. Így meglepő, hogy némely természetes fehérjében szomszédos cisztein oldalláncok között létesül kovalens kapcsolat és így kialakul egy nyolctagú gyűrű (ld. **B.3 Ábra**). Feltételezhető, hogy egy ilyen ritka motívum a fehérjében fontos szerkezeti vagy funkcionális szereppel bír.

A



B



B.3. Ábra A fehérjékben található vicinális diszulfidhidak modellvegyületei és jellemző torziósözegek. **A:** HCO-*ox*-[L-Cys-L-Cys]-NH₂ peptid. **B:** A 4,5-ditiaheptano-7-laktám. A torziósözegek mindegyikét gyűrűatomok definiálják. Megjegyzendő, hogy míg a laktám ω , χ_1 , χ_2 , χ_3 és χ_4 torziója rendre azonos a modellpeptid ω_1 és χ_{21} , χ_{22} , χ_{13} és χ_{12} torziósözegeivel, addig χ_5 , χ_6 és χ_0 nem azonos χ_{11} -gyel, ψ_1 -gyel és ϕ_2 -vel (vö. **B.1/A Ábra**).

A legjelentősebb példa a nikotinos acetil-kolin receptorok α alegysége, melyben az acetil-kolint kötő ligandumfüggő pentamer ionsatornákra specifikus, konzervatív diszulfidhíd részt vesz a ligandumok ill. a toxinok megkötésében. Az ilyen receptorok egyes típusai megtalálhatók az ideg-izom kapcsolatokban, valamint a pre- és posztszinaptikus idegvégződésekben. Szerepük a nikotinfüggőségben és olyan betegségekben, mint a Parkinson- és az Alzheimer-kór,⁶⁶ a nikotinos acetil-kolin receptorokat a gyógyszertervezés fontos célmolekulájává tette.

Az *Alcaligenes* észteráz 713 egy α/β hidroláz (1qlw),⁶⁷ melyben a vicinális diszulfidről feltételezik, hogy hozzájárul a β -kanyar szerkezeti stabilitásához, miközben a Cys71 gerinc nitrogénje az oxianion-üreg képzésének részese.

A kinoprotein alkohol dehidrogenázok (1h4i⁶⁸, 1h4j⁶⁹, 1g72⁷⁰, 4aah⁷¹, 1kb0⁷², 1kv9⁷³, 1flg⁷⁴) pirrolokinolin-kinon proszтетikus csoportot tartalmaznak. Ehhez térben közel található a konzervált *ox*-[Cys-Cys] egység, és habár közvetlenül nem vesz részt a katalízisben, a ciszteinek bármelyikének kicserélése az enzimaktivitás elvesztését eredményezi.^{69,75}

Az atrakotoxin J-ACTX-Hv1c esetében is irányított mutagenézis vizsgálatok bizonyították, hogy a biológiai aktivitáshoz elengedhetetlen a vicinális diszulfidhíd.⁷⁶ A kétszeres mutáns (Cys13Ser, Cys14Ser) toxin megtartja natívhoz hasonló konformációját,⁷⁷ tehát maga az S–S-kötés a társ-molekulák által felismert részlet kialakításában játszhat szerepet.

Az alginát liáz (1hv6⁷⁸, 1qaz⁷⁹) esetében az *ox*-[Cys–Cys] egység a második hurok (loop2) oldószerrel érintkező részén helyezkedik el. A diszulfid közel található több konzervált, az aktivitásért felelős aminosavhoz is, így szerepet játszhat az aktív centrum konformációjának megtartásában.

A karboxipeptidáz T (1obr⁸⁰) Cys155–Cys156 hídja szokatlan a karboxipeptidázok között. Feltételezett szerepe a kapcsolódó felszíni hurok merevségének növelése, hogy az ellenállóbb legyen a proteolízissel szemben.

Az *Anguilla anguilla* agglutininban (1k12⁸¹) a vicinális diszulfid érintkezik az oldószerrel és feltételezik, hogy a molekulák közötti nem specifikus kapcsolatok kialakításában, továbbá a ligandkötésért felelős hurkok konformációjának és merevségének biztosításában játszik szerepet.

Az emberi hepcidin két legjellemzőbb formája, a hepcidin-20 és a hepcidin-25 diszulfidhídja két β -redőzött szalagot összekötő hurokban található.⁸² Meglepő, hogy a vicinális diszulfid az NMR-mérések időskáláján kicserélődési folyamatokban vesz részt és a molekula legmozgékonyabb szakaszának mutatkozik.

A PDB-ben talált egyetlen nem natív *ox*-[Cys–Cys] szekvenanciaegység az α -konotoxin GI mesterséges diszulfidhíd izomerje (1xgc⁸³). A Cys2 és Cys3 diszulfidos összekötése legalább két domináns és három kisebb jelentőségű konformert eredményez szemben az eredeti molekula jól meghatározott egységes szerkezetével. Az izomernek az a konformerje, melynek a szerkezetét sikeresen kiszámították, jelentősen eltér a vad típusú toxinétól, tekintettel a gerincre ill. a kénatomokra vonatkozó 4,2 Å-ös ill. 7 Å-ös RMSD értékekre. Érdekes módon ez az izomer mégis tizedakkora biológiai aktivitást mutat, mint a natív forma, miközben a többi lehetséges nem natív izomer teljesen inaktív.

A palmitoil fehérje tioészteráz (1ei9⁸⁴, 1eh5⁸⁴, 1exw⁸⁵) és az α -L-arabináz (1gyd, 1gye és 1gyh⁸⁶) esetében a vicinális diszulfidhidak biológiai szerepét nem sikerült kimutatni.

MÓDSZEREK

Peptidek konformációs kódja és nevezérendszere

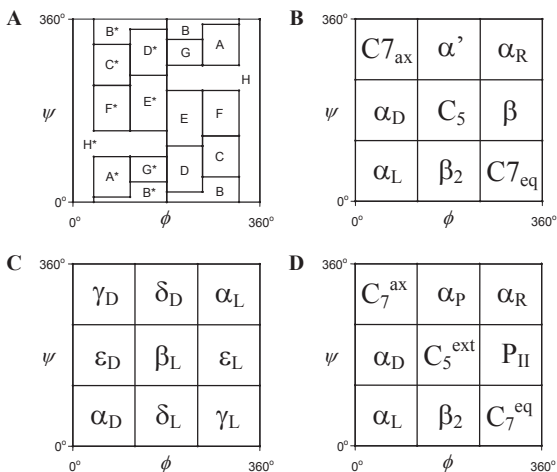
Egy molekula valamely konformációja a molekula konformációs terének egy pontja. Ezt a konformációs teret a geometria független szabadsági fokai (pl. kötэшossz, kötэшsszög, torziósszög), mint megannyi egymásra merőleges dimenzió feszíti ki. A molekula térbeli szerkezetének módosulásával együtt változó valamely fizikai tulajdonságát (pl. energiáját, entrópiáját), tekinthetjük olyan függvénynek, melynek az értelmezési tartománya a többdimenziós konformációs tér. Ha az energiát egy, két vagy több dimenzió felett értelmezzük, rendre görbéről, felszínről ill. hiperfelszínről beszélünk. A konformerek az energiafüggvény lokális minimumai, a molekula nyugalmi konformációi. Megjegyzendő, hogy a többatomos molekulák szabadsági fokainak többféle kombinációja is alkalmas a konformációs tér belső koordinátáinak megadására. A kiválasztásnál rendszerint a kémiai konvenciókra támaszkodunk, miközben alternatívákat is elfogadhatnánk.

Térszerkezeti szempontból elegendő azokra a szabadsági fokokra összpontosítanunk, amelyek mentén egynél több minimum lehetséges, vagyis valószínűsíthető legalább két olyan konformer létezése, melyek e dimenzióban jelentősen eltérnek, miközben a konformációs tér többi kijelölt koordinátája mentén alig különböznek egymástól. Kötэшossz és kötэшsszög aligha lehet megkülönböztető jegy, de a peptidek torziósszögei között is léteznek összefüggések. Ezért a nevezérendszerben csak a „vezérkoordinátákra” vagyunk tekintettel, nevezetesen az egyes aminosavak ω_{i-1} , ϕ_i , ψ_i , χ_{ij} torziósszögeire (ld. **B.1. Ábra**), összhangban a konvenciókkal, de elismerve az alternatívák alkalmasságát is. A konformációs elemzés célja a modellvegyület kiválasztott vizsgálati módszer szerinti konformációs könyvtárának előállítása, vagyis fellelhető összes nyugalmi szerkezetének megtalálása és jellemzése. A nevezérendszer pedig a konformerek ésszerű, „magától értetődő” címkézése.

A térszerkezet pontos megadásához n atomos molekula esetén (ha $n > 2$) $3n-6$ koordináta szükséges, melyből a belső koordinátákat alkalmazó Z-mátrix esetén $n-3$ a torziósszög.^{87,88} Egy konformációs nevezérendszernek csak kevés koordinátával kell törődnie, mert a szerkezeti információk nagyobb részét (az atomok kovalens kapcsolatát, a konfigurációt) már a kémiai név önmagában meghatározza. Például egy L-alanil egység 10 atomot ad egy peptidhez, de konformációja már 3 torziósszöggel jellemezhető, a többi adat kiszámolható egy kiválasztott kvantummechanikai elméleti szinten (ld. **B.1/B Ábra**).

Egy peptid sokdimenziós konformációs tere a peptidet felépítő aminosavak konformációs tereinek, mint megannyi egymástól nem független alternék a kombinációja. Ezek az alternék jelentős hasonlóságot mutatnak egymással, hiszen a természetes aminosavak csak oldalláncaikban különböznek. Peptidmodellek korábbi vizsgálatai alapján állíthatjuk, hogy a konformációs tér minimumainak száma nem haladhatja meg az egyes független konformációs szabadsági fokok (pl. ω_{i-1} , ϕ_i , ψ_i , χ_{ij}) mentén lehetséges minimumok aritmetikai szorzatát.⁸⁹ Gyakorlati szempontból hasznosnak tűnik továbbá, hogy a peptidek konformációs kódját az egyes összetevő aminosavak lokális konformációs kódjának összefűzésével adjuk meg az N-terminálistól a C terminális felé haladva.

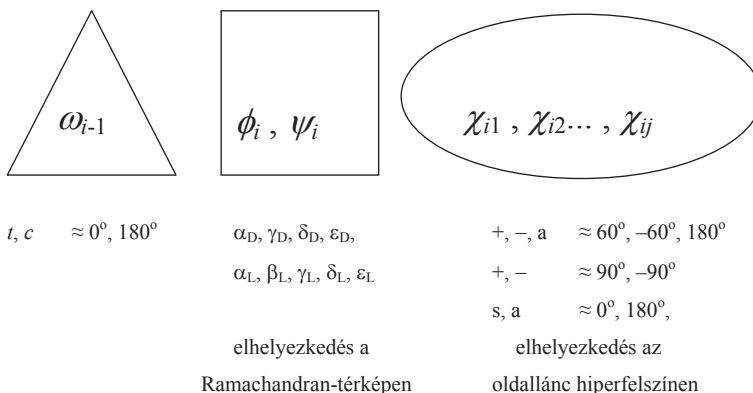
A peptidek konformációs szempontból érdekes torziószögei két vagy három minimummal rendelkezhetnek.⁹⁰ A peptidkötés (amidkötés) *cisz* vagy *transz* (ekkor $\omega_{i-1} \approx 0^\circ$ ill. $\omega_{i-1} \approx 180^\circ$). A gerincállás célszerűen címkézhető a ϕ_i és ψ_i paraméterekkel meghatározott Ramachandran-térkép régiói alapján (ld. **M.1. Ábra**). Az oldallánc torziószögei (χ_{i1} , χ_{i2} , stb.) rendszerint *gauche*- (*g*-, kb. -60°), *gauche*+ (*g*+, kb. 60°) vagy *anti* (*a*, kb. 180°) állást vehetnek fel. Telítetlen atomok hatására a lehetőségek azonban leszűkülhetnek egy-egy párra, pl. pozitív (+, kb. 90°) és negatív (-, kb. -90°), vagy *szin* (*s*, kb. 0°) és *anti* (*a*, kb. 180°) párra. A nevezérendszernek az ezen torziók mentén észlelhető változatokat kell visszatükörznie.



M.1. Ábra Aminosav diamidegységek gerinckonformereinek négyféle jelölése a Ramachandran-térképen való elhelyezkedésük szerint: **A**⁹¹, **B**⁹², **C**⁹³, **D**⁹⁴.

A konformációs kód alapvetően minőségi leírása a szerkezetnek, bár hordoz bizonyos mennyiségi jellemzőket is. Fontos megjegyezni, hogy a peptidszekvenciák háromdimenziós szerkezete összehasonlítható konformációs kódjuk összevetésével (összefésülés vagy alignment). A lizozim és öt pontmutációjának gerinckonformációját az ACAP program (Amino acid conformational analyses of proteins) segítségével sikerült összevetni.⁹⁵ CD4 molekula estében hasonló megközelítést alkalmaztak a molekuladinamikával nyert szerkezetek jellemzésére (ún. asszignáció).⁹⁶

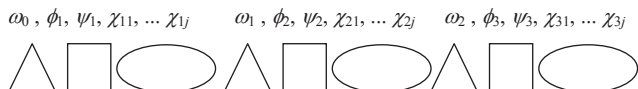
A fent említett két esetben, az alanin kvantumkémiai módszerekkel nyert konformereit használták mintaként, ún. templátként az aminosavgerinc osztályozásához. Lehetséges azonban a gerinckonformerek csoportosítása a Ramachandran-térkép $120^\circ \times 120^\circ$ -os metszetei alapján is (ld. **M.1/C Ábra**). A két módszer nem vezet azonos eredményre, mert a templátok nem a régiók („vízgyűjtőterületek”) középpontjára esnek.



M.2. Ábra Egyetlen aminosav diamidegység konformációs kódjának képzése általánosan, ill. konkrét karakterek használatával. Megjegyzendő, hogy az aminosavval szorosabb térszerkezeti kapcsolatban áll a megelőző peptidkötés (ω_{i-1}), mint a rákövetkező (ω_i), nevezetesen a prolint az egyéb aminosavaknál jelentősen gyakrabban előzi meg *cisz* állású amid.

A kódon belül érdemes eltérő karaktereket alkalmazni (a) a peptidkötések, (b) a $[\phi, \psi]$ párok és (c) az oldalláncok leírására, mert akkor egy-egy típus elhagyása a gyakorlatban különösebb magyarázat nélkül sem okoz félreértést. A jelölések (a) *cisz* ill. *transz* amidra dőlt latin *c* ill. *t*; (b) a gerincre a Ramachandran-térkép szisztematikus felosztását követő, pl. a

kutatócsoportunkban az **M.1/C Ábra** szerint görög α , β , γ vagy δ betű és alsó indexbe tett latin $_L$ ill. $_D$ kombinációja, (c) az oldallánca pedig a megfelelő torzószőgek térállását tükröző $+$, $-$, a és kivételes esetekben s betűk sora (ld. **M.2** és **M.3. Ábra**). Elhagyjuk a kódból az oldallánca vonatkozó jeleket, ha csak a peptidgerincet vizsgáljuk. Ha minden amid *transz* állású, feleslegessé válik a t kiírása. Így pl. egy 4 aminosavegységnyi jobbmenetes hélix kódja $\alpha_L \alpha_L \alpha_L \alpha_L$ vagy egyszerűbben $\alpha_L \alpha_L \alpha_L$.



M.3. Ábra Peptidlánca vonatkozó konformációs kód képzésének vázlata.

A *cisz* peptidkötés a *transz*-tól könnyen, egyértelműen megkülönböztethető. A c és t jelölés lehetséges alternatívája a geometriai izomériára a IUPAC által is szorgalmazott E és Z ($\omega_j \approx 0^\circ$ ill. $\omega_j \approx 180^\circ$), mindazonáltal a peptidekre vonatkozó szakirodalom ezt kerüli.

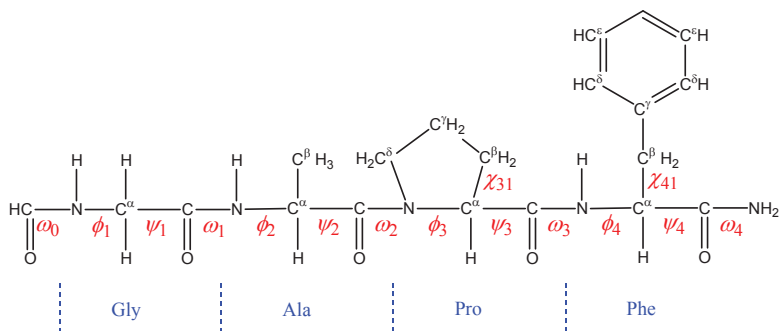
A Ramachandran-térkép felosztására számtalan javaslat ismert. Karplus 12 régiót különböztet⁹⁷ meg, melyeket a következőképpen jelöl: α_L , α_R , β_S , β_P , γ , γ' , δ_L , δ_R , ϵ , ϵ' , ϵ'' és ζ . Ez a felosztás hasonló Efimovéhoz.⁹⁸ Ezeket megelőzően javasolták Zimmerman és munkatársai⁹¹ 16 terület elkülönítését: A, B, C, D, E, F, G, H, A*, B*, C*, D*, E*, F*, G* és H*. Richardson és Richardson⁹⁹, Rooman¹⁰⁰, Thornton¹⁰¹ és munkatársaik szintén tettek javaslatot a gerinckonformerek csoportosítására. E megközelítések mindegyike fehérjék röntgenszerkezetein alapul.

A királis C^α központhoz geminálisan kapcsolódó két amidcsoport $E=E(\phi, \psi)$ energiatárcsájának topológiáján alapul kutatócsoportunk jelölési sémája.⁹³ Feltételezzük, hogy a felszínen legfeljebb $3 \times 3 = 9$ energiaminimum jelenik meg (ld. **M.1/C Ábra**). A görög betűk közül α az α -hélixre (vö. $(\alpha_L)_n$), β a β -redőzött rétegekre (vö. $(\beta_L)_n$), míg γ a γ -kanyarokra utal. Ez a felosztás tükrözi a Ramachandran-felszín királis voltát, nevezetesen, hogy az L-aminosavak az L-völgy konformereit kedvelik (α_L , β_L , γ_L , δ_L , ϵ_L), míg a D-aminosavak a D-völgyet részesítik előnyben (α_D , β_D , γ_D , δ_D , ϵ_D).

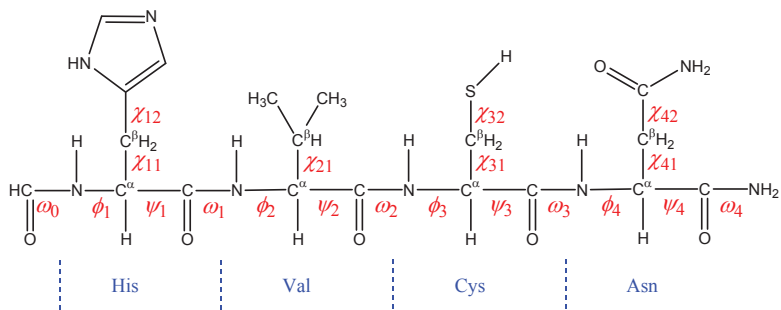
Oldalláncok esetében eltekinthetünk egy-egy torzió vizsgálatától (ld. **M.4. Ábra**), ha (a) állása korábban megadott torziókódból egyértelműen következik (pl. prolin χ_{ij} , ha $j \neq 1$), (b) kémiailag azonos rotamerekkel rendelkezik (pl. alanin χ_{11} , fenilalanin χ_{12}), (c) a módszerrel

meg nem határozott helyzetű hidrogénatomot jellemez (pl. tirozin χ_2 , ha a fenolos OH protonja nem különböztet meg konformereket). A prolin nevezéktana (ld. **M.5. Ábra**) a pszeudorotáció tárgyalása után válik érthetőbbé.

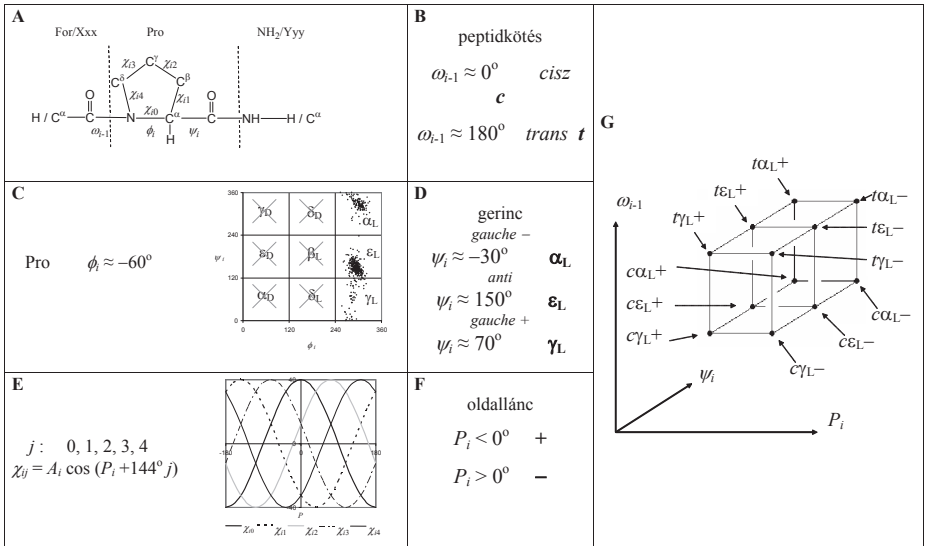
A



B



M.4. Ábra Kiemelt konformációs jelentőséggel bíró torziószögek egy-egy hipotetikus modellpeptid aminosavain szemléltetve. **A:** HCO-Gly-Ala-Pro-Phe-NH₂. **B:** HCO-His-Val-Cys-Asn-NH₂.



M.5. Ábra A prolin térszerkezeti nevezékntana. (folytatás a következő oldalon)

(M.5. Ábra folytatása)

Minden prolin egységhez egy négykaraktes kód tartozik. Tetszőleges n egész szám esetén, a **(Pro) $_n$** szekvenca konformációs kódja hasonlóképpen $4n$ karakterből áll. **A:** Kiválasztott torziószőgek a HCO–L-Pro–NH₂ modellpeptidben, valamint peptidek és fehérjék Xxx–Pro–Yyy szekvenciárszleteiben. A gerinc torzióit (ω_{i-1} , ϕ és ψ_i) gerincatomokkal adjuk meg (C ^{α} , karbonil C és amid N, ill. helyettesítő H megfelelő négyesével). Az oldallánc χ_{ij} torzióit a pirrolidingyűrű N, C ^{α} , C ^{β} , C ^{γ} és C ^{δ} atomjai definiálják. Így ϕ és χ_{i0} nem azonosak, de nem is függetlenek egymástól. **B:** A ω_{i-1} torzió hozzájárulása a nevezékhez. **C:** Gerinctérszerkezetek címkézése az aminosavnak a Ramachandran-térképen található helyzete alapján.⁹³ A Ramachandran-térkép felosztható 9 ideális vízgyűjtő területre, amelyekből a prolin csak három érhet el. **D:** A ψ_i torziószőg hatása a konformációs kódra. **E:** Az oldallánc χ_{ij} torzióiból kiszámítható a pirrolidingyűrű két pszeudorotációs paramétere (A_i és P_i). A pszeudorotációs amplitúdó, A_i , a χ_{ij} értékeknek egy felső korlátja, míg a fázis, P_i , ezen torziók hányadosait jellemzi. **F:** A pszeudorotációs fázisszőg tükrözödése a prolin térszerkezeti kódjában. Megjegyzendő, hogy a –, ill. + jel mindig ellentétes P_i előjelével, azonban rendszerint megegyezik χ_{i1} előjelével. A – és + gyűrűállásokra az irodalomban elterjedt az Up és Down, ill. az *anti* és *syn* megkülönböztetés. **G:** Egyetlen L-prolin egység ideálisan létező 12 térszerkezete, valamint ezek elhelyezkedése a független ω_{i-1} , P_i és ψ_i változók által kifeszített konformációs térben.

Konformerek kvantumkémiai megközelítése

A kvantumkémiai számításokat¹⁰² a GAUSSIAN98¹⁰³ ill. a GAUSSIAN03¹⁰⁴ programcsomaggal végeztem, az optimalásokat a GDIIS¹⁰⁵ kereső algoritmus segítette. Egy energiaszámítás a molekula atommagjait a kezdeti geometriában rögzíti és az elektronszerkezetet a kiválasztott módszerrel a kiválasztott bázison közelítve adja meg az energiát („single point”). Egy optimalás energiaszámítások sorozata, a nyugalmi állapotot még el nem ért molekulában ható erők mentén apránként változtatva geometriát keresi a legmélyebb energiát. Az RHF módszer (restricted Hartree-Fock, megszorított Hartree-Fock) a többatomos molekulák szerkezetét hullámfüggvények segítségével írja le, iteráció útján keresi azt a determináns alakú hullámfüggvényt, amely a rendszer legmélyebb energiát adja. A DFT módszer (Density Functional Theory, sűrűségfunkcionál-elmélet, pl. B3LYP) a többatomos molekulák szerkezetét elektronsűrűség-funkcionálokkal közelíti. A molekula teljes elektronszerkezetét olyan egyelektron-függvények szorzatának tekintjük, melyeket az egyes atomokhoz rendelt bázisfüggvények kombinációjával állítunk elő. Az általam használt bázisok Gauss típusúak, pl. a 3-21G bázis az atomok belső pályáit három Gauss-függvénnyel írja le, míg a vegyértékhéjat két olyan pályával, melyek kettő ill. egy primitív Gauss-függvényből állnak. A bázisokat gyakran kiegészítik polarizációs és diffúz hullámfüggvények, előbbieket az atommagok körüli aszimmetria leírását segítik (jelük pl. *, d, p) utóbbiak az atommagoktól távoli térrészek leírását segítik (jelük pl. +, ++ a bázis nevében). Az oldószerhatást PCM módszerrel vettem figyelembe.¹⁹

A számítások az egyes térszerkezetek teljes elektronikus energiáját (E) Hartree egységben adják meg. A globális minimum az adott modellvegyületnek azon téralkata, amely a választott elméleti szinten a legalacsonyabb energiával rendelkezik, tehát a legnagyobb stabilitású és várhatóan a leggyakoribb konformere. A többi konformer relatív energiáját (ΔE) a globális minimumra vonatkoztattam és kcal.mol⁻¹ egységbe számoltam át:

$$\Delta E = \frac{E - E_{glob.min.}}{627,51 \text{ kcal.mol}^{-1} . \text{Hartree}^{-1}} \quad (1)$$

A Boltzmann-eloszlás elvét alkalmazva, a globális minimumra vonatkoztathatjuk a relatív becsült gyakoriságot is a következő egyenlet segítségével:

$$p_{rel} = 100\% e^{-\frac{\Delta E}{RT}} \quad (2)$$

ahol $e=2,718282$ és 300 K hőmérsékleten $RT=0,595371\text{ kcal.mol}^{-1}$. Ez azt jelenti, hogy ilyen hőmérsékleten a 0, ill. 0,5, 1, 2 vagy 3 kcal.mol^{-1} relatív energiával rendelkező téralkatok a leggyakoribb konformerhez képest azonos, ill. 43, 19, 3 és 1% várható gyakoriságúak. Amennyiben (a) minden alacsony energiájú konformert megtaláltuk és (b) azok reverzibilis úton átalakulhatnak egymásba, akkor az egyes konformerek abszolút várható gyakoriságát kiszámítva:

$$p = \frac{P_{rel}}{\sum_{konformerek} P_{rel}} \quad (3)$$

ezek összessége megmutatja a modell konformációs összetételét.

Fehérjék aminosavainak természetes konformációs összetétele és a modellpeptid becslt konformációs összetétele

Statistikailag elemezhető kísérleti adatok forrásául három adatbázist használtam: (1) kisméretű szintetikus és izolált peptidok röntgenszerkezeteire vonatkozóan a Cambridge-i Szerkezeti Adatbázist (CSD, Cambridge Structural Database¹⁰⁶), (2) fehérjék és polipeptidek NMR-rel ill. röntgendiffrakcióval meghatározott szerkezeteivel kapcsolatosan pedig a teljes Fehérjeszerkezeti Adatbankot (PDB, Protein Data Bank¹⁰⁷, Rel. 103, 2003. januári frissítés) és (3) a jó felbontású, nem homológ fehérjeszerkezetekből álló PDB SELECT¹⁰⁸ gyűjteményt (2002. áprilisi frissítés). A geometriai és a gyakorisági adatokat összevettem a modellvegyületek kvantumkémiai számításainak vonatkozó eredményeivel.

Fehérjék alanin aminosavának természetes konformációs összetételét egy korábban kidolgozott módszer¹⁴ szerint számoltuk ki. A kísérleti adatok forrásául a PDB SELECT-et választottuk. Az 5393 alanin között 1767 belső és 3626 felszíni volt. E három csoportból összesen 24 természetes konformer alcsoportot származtattunk a következő eljárással: egy kísérletileg meghatározott $[\phi, \psi]$ párt a választott elméleti szinten egy adott konformerhez tartozónak tekintünk, ha a Ramachandran-térképen a számított $[\phi, \psi]$ értékpont köré rajzolt 15° vagy 30° sugarú körön belül helyezkedik el. Ha valamely minimumok egymás közvetlen közelében helyezkednek el (tehát $(\Delta\phi^2 + \Delta\psi^2)^{1/2} < 30^\circ$ ill. 60°), akkor az átfedések elkerülése érdekében csak a legalacsonyabb energiájú minimumot vettük figyelembe. A fehérjékben talált alanin konformerek természetes gyakoriságának logaritmusát és a különböző elméleti szinteken kiszámolt relatív energiák közötti korreláció mennyiségi jellemzésére R^2 -et választottuk R -rel szemben.

Pszeudorotációs paraméterek számítása prolin pirrolidingyűrűjére

A prolin pirrolidingyűrűje nem síkalkatú, hanem redőzött. Torziószögei így nullától jelentősen eltérhetnek, egymástól azonban nem függetlenek. A pszeudorotáció elmélete hosszú történettel rendelkezik, így számtalan molekulára alkalmazták, következésképpen többféle felírása is létezik. A jelen dolgozatban következő összefüggést alkalmazom:

$$\chi_{ij} = A_i \cos(P_i + 144^\circ j) \quad (4)$$

ahol A_i és P_i dimenziója is fok, és A_i mindig pozitív. A szekvenciában i -dik aminosavként elhelyezkedő prolin A_i amplitúdóját és a P_i fázisszögét a **M.5/A ábrán** definiált öt χ_{ij} torzióból olyan módon származtatjuk, hogy bármely 0 és 4 közötti j egész számra teljesüljön a fenti összefüggés. Így az A amplitúdó az oldallánc öt torziószögének egy felső korlátját képezi, hiszen a trigonometrikus függvény definíciójából következően:

$$-1 \leq \cos(P_i + 144^\circ j) \leq 1 \quad (5)$$

amit a pozitív A_i -vel szorozva:

$$-A_i \leq A_i \cos(P_i + 144^\circ j) \leq A_i \quad (6)$$

$$-A_i \leq \chi_{ij} \leq A_i \quad (7)$$

A P fázisszög pedig az oldallánc torziószögeinek arányait jellemzi:

$$\frac{\chi_{ij}}{\chi_{ik}} = \frac{A_i \cos(P_i + 144^\circ j)}{A_i \cos(P_i + 144^\circ k)} = \frac{\cos(P_i + 144^\circ j)}{\cos(P_i + 144^\circ k)} \quad (8)$$

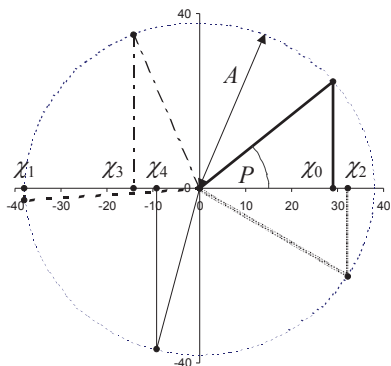
amint azt a **M.6. Ábra** szemlélteti.

A (4) kifejezés azonos a következővel:

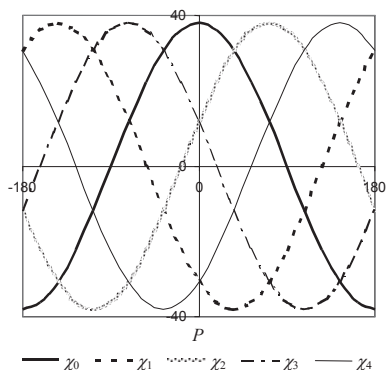
$$\chi_{ij} = A_i \cos\left(P_i + \frac{4\pi}{5} j\right) \quad (9)$$

amennyiben P_i mértékegysége radián és nem fok.

A



B



M.6. Ábra Egy tetszőleges öttagú gyűrű χ_0 , χ_1 , χ_2 , χ_3 és χ_4 torziószögei és a pszeudorotációs paraméterek (A és P) közötti összefüggés. **A** Minden χ_j érték az A hosszúságú $P+144^\circ j$ szögben álló vektornak a vízszintes tengelyen képzett vetülete. **B** A χ_j torziószögek értékei P függvényében, ha $A=38^\circ$. Az A pszeudorotációs amplitúdó az öt χ_j értéknek egy felső korlátja, míg a P fázisszög pedig arányait jellemzi.

A (4) kifejezésben a koszinusz függvény választása egyszerű fázispreferencia, mivel a prolin kedvező energiájú konformereiben a pszeudorotációs fázis $+90^\circ$ és -90° közelébe esik, és így ábrázolása kényelmesen megvalósítható a -180° és 180° közötti szakaszon. Szinuszos függvényt választva

$$\chi_{ij} = A_i \sin(P''_i + 144^\circ j) \quad (10)$$

a pszeudorotációs fázis egyszerűen eltolódik: $P''_i = P + 90^\circ$.

Figyelembe véve a bármely azonos dimenziójú P_i -re és X -re teljesülő trigonometrikus kifejezést:

$$\cos(P_i + X) = \cos P_i \cos X - \sin P_i \sin X \quad (11)$$

a pirrolidingyűrű torziószögei a következő összeget adják:

$$\sum_{j=0}^4 \chi_{ij} = A_i \sum_{j=0}^4 \cos(P_i + 144^\circ j) = A_i \left[\cos P_i \sum_{j=0}^4 \cos(144^\circ j) - \sin P_i \sum_{j=0}^4 \sin(144^\circ j) \right] = 0 \quad (12)$$

Az öt χ_{ij} összege nulla, mivel 0° , 144° , 288° , 432° és 576° koszinuszai és szinusza is nullává összegződnek.

A prolinban a pirrolidingyűrű aszimmetrikusan szubsztituált, ezért nem várható, hogy akár a C₂ akár a C_S szimmetria, tehát a forgástengely vagy a siktükör szerint szimmetrikus csavart ill. boríték alkat megvalósuljon. A gyűrű öt torziójára valójában csak egy viszonylag kis hibával illeszthető a mindössze két származtatott paramétert tartalmazó pszeudorotációs kifejezés. A χ_{ij} értékeknek a számításokban használt digitális pontossága, vagyis a figyelembe vett tizedesjegyek száma, tovább rontja az illesztést. Kicsi A_i érték, vagyis lapos gyűrű esetén különösen bizonytalaná válhat P_i . Így az összeg végül csak közelíti a nullát:

$$\sum_{i=0}^4 \chi_i \approx 0 \quad (13)$$

Amennyiben az öt χ_{ij} összege -2° és $+2^\circ$ közé esik, a gyűrű a pszeudorotációs kifejezésre illeszkedőnek tekinthető.

Figyelembe véve a bármely azonos dimenziójú P_i -re és X -re teljesülő trigonometrikus kifejezést:

$$\cos(P_i + X) - \cos(P_i - X) = -2 \sin P_i \sin X \quad (14)$$

megállapíthatjuk, hogy:

$$\begin{aligned} & \chi_{i1} - \chi_{i2} + \chi_{i3} - \chi_{i4} = \\ & = A_i [\cos(P_i + 144^\circ) - \cos(P_i + 288^\circ) + \cos(P_i + 432^\circ) - \cos(P_i + 576^\circ)] = \\ & = A_i [\cos(P_i + 144^\circ) - \cos(P_i - 72^\circ) + \cos(P_i + 72^\circ) - \cos(P_i - 144^\circ)] = \\ & = -2A_i \sin P_i (\sin 144^\circ + \sin 72^\circ) \end{aligned} \quad (15)$$

Az A_i és P_i pszeudorotációs paramétereknek a χ_{i0} , χ_{i1} , χ_{i2} , χ_{i3} és χ_{i4} torziószögekből történő számítására a következő algoritmust vezettem be:

1. lépés
$$A_i \sin P_i = \frac{\chi_{i1} - \chi_{i2} + \chi_{i3} - \chi_{i4}}{-2(\sin 144^\circ + \sin 72^\circ)} \quad (16)$$

Figyelembe véve, hogy a $\chi_{i0} = A_i \cos P_i$ (melyet a (4) kifejezésbe $j=0$ -t helyettesítve kapunk), valamint a következő trigonometriai egyenlőséget

$$\sin^2 P_i + \cos^2 P_i = 1 \quad (17)$$

a számítás az alábbiak szerint halad:

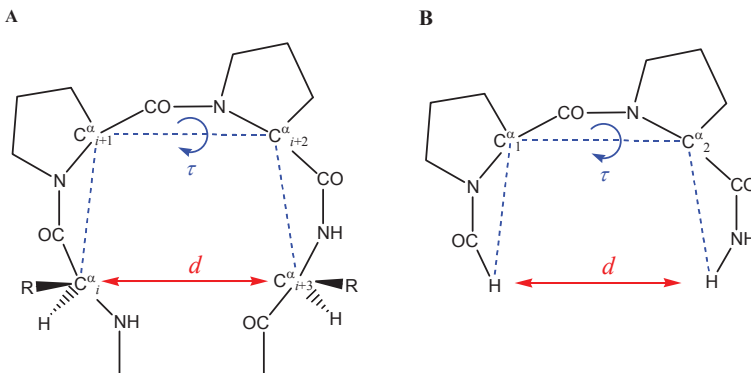
2. lépés $A_i = \sqrt{(A_i \sin P_i)^2 + \chi_{i0}^2}$ ahol $A_i \geq 0$ (18)

3. lépés $\left\{ \begin{array}{l} \text{ha } A_i \sin P_i \geq 0 \\ \text{ha } A_i \sin P_i < 0 \end{array} \right. \begin{array}{l} P_i = \frac{180^\circ}{\pi} \arccos\left(\frac{\chi_{i0}}{A_i}\right) \\ P_i = -\frac{180^\circ}{\pi} \arccos\left(\frac{\chi_{i0}}{A_i}\right) \end{array}$ (19)

A $180^\circ/\pi$ együttható elhanyagolható, amennyiben az árkuszkoszinusz függvény értékészlete fokokban értendő radián helyett. Az A_i paraméter mindig pozitív, míg P_i mindig -180° és $+180^\circ$ közé esik.

β -Kanyarok osztályozása

A β -kanyarokat az **M.1. Táblázat** alapján soroltam be a klasszikus típusokba. A β -kanyarok τ és d paramétereit (ld. **M.7. Ábra**) jellegzetesen különböztetik a hosszított szerkezetekétől.



M.7. Ábra Dipeptid szekvenciaegységek térszerkezetét átfogóan jellemző τ torziószög és a d távolság definíciója a prolin-prolin példáján szemléltetve. **A:** Fehérje/peptid **Xxx-Pro-Pro-Yyy** szekvenciarészletében a τ torziószöget négy szomszédos C^α atom ($C^\alpha_i - C^\alpha_{i+1} - C^\alpha_{i+2} - C^\alpha_{i+3}$) adja meg fokokban mérve. A d távolság a C^α_i és C^α_{i+3} atomok távolsága Å egységben. **B:** A **HCO-Pro-Pro-NH₂** modellpeptidben a C^α_i és C^α_{i+3} atomokat hidrogén helyettesíti, tehát τ a $H-C^\alpha_1-C^\alpha_2-H$ torziószög, míg d a két helyettesítő hidrogénatom távolsága.

β-Kanyar típus	Javasolt szerkezeti kód ¹⁰⁹	Ideális torziószög (°)			
		ϕ_{i+1}	ψ_{i+1}	ϕ_{i+2}	ψ_{i+2}
I	$t\alpha_L t\delta_L, t\alpha_L t\alpha_L, t\alpha_L t\gamma_L$	-60	-30	-90	0
I'	$t\alpha_D t\delta_D, t\alpha_D t\alpha_D, t\alpha_D t\gamma_D$	60	30	90	0
II	$t\epsilon_L t\alpha_D, t\epsilon_L t\gamma_D, t\epsilon_L t\delta_D$	-60	120	80	0
II'	$t\epsilon_D t\alpha_L, t\epsilon_D t\gamma_L, t\epsilon_D t\delta_L$	60	-120	-80	0
III	$t\alpha_L t\alpha_L$	-60	-30	-60	-30
III'	$t\alpha_D t\alpha_D$	60	30	60	30
IV	nem egyértelműen definiált	I-III' β-kanyar típusokhoz hasonló, de egy vagy két torziószög 40°-nál erősebben eltér az ideális értéktől ¹¹⁰			
V	$t\gamma_L t\gamma_D$	-80	80	80	-80
V'	$t\gamma_D t\gamma_L$	80	-80	-80	80
VIa *	$t\epsilon_L c\alpha_L, t\epsilon_L c\gamma_L, t\epsilon_L c\delta_L$	-60	120	-90	0
VIb *	$t\beta_L c\alpha_L, t\beta_L c\gamma_L, t\beta_L c\delta_L,$ $t\beta_L c\epsilon_L$	-120	120	-60	0 ¹¹¹ vagy 150 ¹¹²
VII ¹¹³	$t\epsilon_L t\alpha_L, t\epsilon_L t\gamma_L, t\epsilon_L t\delta_L$	-90	140	-65	0
VIII	$t\alpha_L t\beta_L$	-60	-30	-120	120

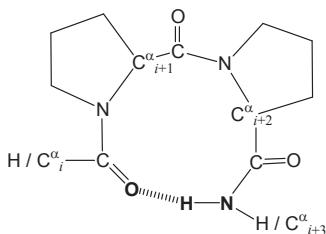
M.1. Táblázat A leggyakoribb β-kanyar típusok. Az első oszlop a fehérjék és peptidok kísérleti vizsgálatában elterjedt hagyományos jelölést tartalmazza, a második a több lehetséges konformer szisztematikus elnevezését, a további oszlopok pedig a jellemző ϕ és ψ torziószögeket.

* *cisz* peptidkötés az $i+1$. és az $i+2$. aminosav között.

Hidrogénhid-kötések jellemzése

A hidrogénhid elektronegativ heteroatom (pl. oxigén) nemkötő elektronpárja és egy másikhoz (pl. nitrogénhez) kapcsolódó, ezért elektronban szegény hidrogénatom között kialakuló, viszonylag erős másodrendű kötés. A periodikus másodlagos szerkezetek (pl. α -hélix, β -redő) stabilitásának alapja egy jól meghatározható, ugyancsak periodikus hidrogénhid-rendszer. A molekula energiáját csökkenti, azaz szerkezetét stabilizálja. A kvantumkémiai számítások eredményessége szempontjából kulcsfontosságú, hogy a modellvegyületet a választott módszerrel vizsgálva elhanyagolunk-e a szekvencia más részletével vagy az oldószer molekuláival képzett hidrogénhid-kötést. Egyetlen aminosav

diamidja (ld. **B.1. Ábra**) csak γ -kanyart vagy inverz γ -kanyart kialakítva képezhet hidrogénhidat (ϵ_D ill ϵ_L konformer). Két aminosavból felépülő peptidmodellek azonban már kellő méretűek ahhoz, hogy β -kanyarok stabilitásához hozzájáruló hidrogénhidakat tartalmazzanak. A prolin-prolin jellemzésénél a hidrogénhidakat két csoportba osztottam (ld. **M.8. Ábra**).

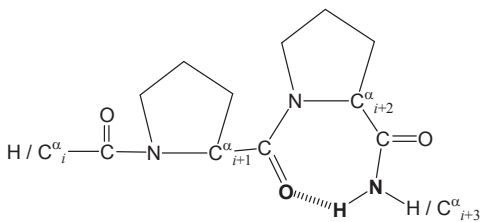


A típus

$(i+3) \rightarrow i$

10-tagú pszeudogyűrű

változatos β -kanyartípusok



B típus

$(i+3) \rightarrow (i+1)$

7-tagú pszeudogyűrű

inverz γ -kanyar

M.8. Ábra A **Pro–Pro** szekvencia némely konformerében fellelhető belső hidrogénhidak sematikus ábrája. Definiáló atomok vastagon kiemelve. Az oxigénatom Å-ben mért távolsága a hidrogéntől $d(OH)$, a protondonor nitrogéntől pedig $d(ON)$. A három atom szögét ($\alpha(OHN)$) fokban adom meg.

EREDMÉNYEK

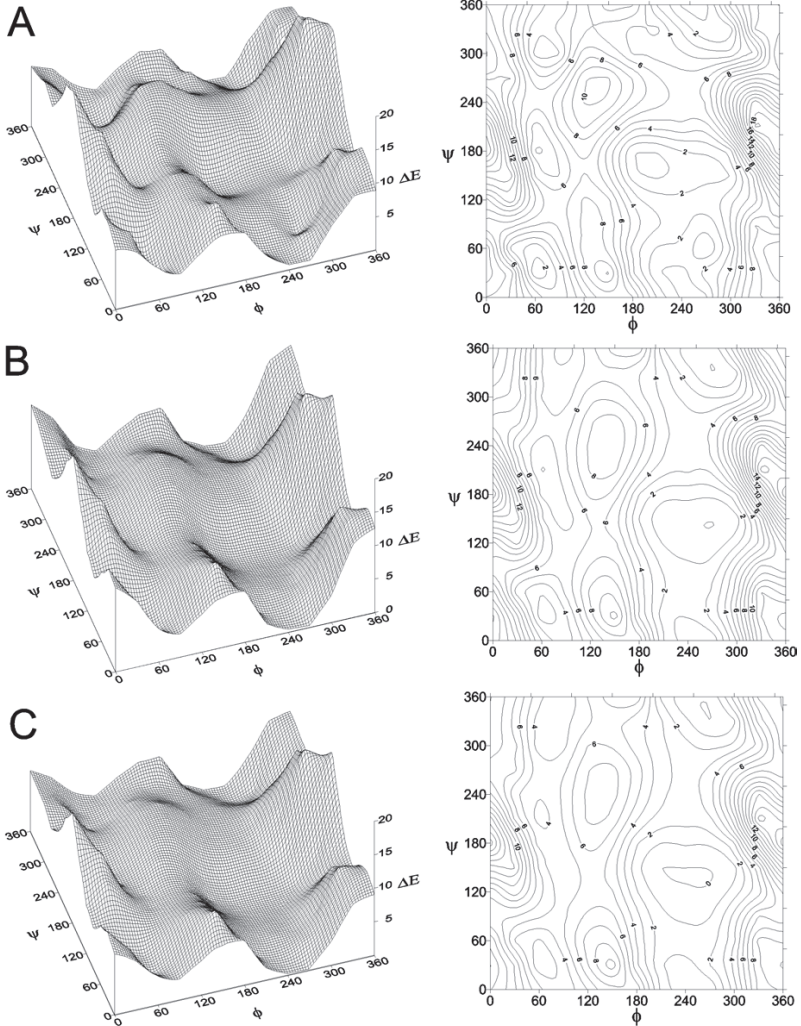
Alanin diamid stabilitásvizsgálata az oldószer modellezésével

Kristályrácsban és oldószeres közegben vizsgált polipeptidek térszerkezetének modellezése gázfázisú kvantumkémiai számításokkal jellegzetes hibákkal terhelt. Megvizsgáltam ezért a PCM számítások alkalmasságát az alanin diamid példáján.

Ramachandran-felület és konformerstabilitás eltérő elméleti szinteken

A HCO-L-Ala-NH₂ modellpeptid RHF/PCM/3-21G, RHF/PCM/6-31+G(d) és B3LYP/PCM/6-31+G(d) elméleti szinten nyert Ramachandran-felületét (ld. **E.1.1. Ábra**) a ψ paraméter mentén egy magas hegygerinc jellemzi a $\phi \approx 0^\circ$ és egy alacsonyabb a $\phi \approx 150^\circ$ értéknél. Az α_L , β_L , ϵ_L , δ_L és γ_L régiókat felölelő (ld. **M.1/C Ábra**) L-völgy, határozottan mélyebb fekvésű a D-típusú konformereket tartalmazó D-völgynél ($\phi \approx 60^\circ$).

Az RHF/PCM/3-21G, RHF/PCM/6-31+G(d) és B3LYP/PCM/6-31+G(d) elméleti szintű felület elemzése alapján várható konformerek száma rendre 12, 7 ill. 7 (ld. **E.1.1. Táblázat**). Az optimált minimumok száma a fenti és a B3LYP/PCM/6-31++G(d,p) elméleti szinten rendre 7, 8, 10 és 7 (ld. **E.1.2. Táblázat**). A B3LYP/PCM/6-31+G(d) felület alacsony energiájú területének újraszámítása a ϕ és a ψ paraméter mentén a korábbi 30°-os helyett 15°-os növekményekkel sem eredményezett további konformert. A B3LYP/PCM/6-31+G(d) szintű többszörös minimumok (ϵ_{D1} és II, α_{L1} és II, β_{L1} , II és III) mindegyike megmaradt „tight” (tehát 30-szoros) konvergencia kritérium mellett is, kivéve a β_{LII} -t, mely β_{LIII} -má alakult.



E.1.1. Ábra A HCO-L-Ala-NH₂ modellpeptid PCM-mel számított Ramachandran-felületének **A:** RHF/PCM/3-21G, **B:** RHF/PCM/6-31+G(d), és **C:** B3LYP/PCM/6-31+G(d) elméleti szinten számolva.

E.1.1. Táblázat Lehetséges energiaminimumok a HCO-L-Ala-NH₂ modellpeptid eltérő elméleti szinteken PCM-mel kiszámított Ramachandran-felszínei alapján.

RHF/3-21G												
ϕ	270	270	240	210	180	90	90	60	60	60	0	0
ψ	60	0	30	150	330	150	0	300	180	30	270	90
ΔE	0,00	0,77	0,83	0,84	4,76	4,96	3,24	2,63	3,79	1,11	9,76	7,32
RHF/6-31+G(d)												
ϕ	270	270	60	90	60	90	0					
ψ	330	150	30	0	210	150	270					
ΔE	0,00	0,04	2,29	3,45	3,88	4,52	8,45					
B3LYP/6-31+G(d)												
ϕ	270	240	60	90	60	60	90					
ψ	330	150	30	0	300	210	150					
ΔE	0,00	0,01	2,17	3,06	3,35	3,58	4,17					

Az RHF/PCM/3-21G elméleti szinten kapott eredmények nagyon hasonlítanak a gázfázisú adatokra.¹³ Sem az α -hélixet felépítő α_L konformer, sem a poliprolin II-t alkotó ϵ_L nem képez valódi minimumot. A másik hét konformációs terület (ld. **M.1/C Ábra**) mindegyike éppen egy konformert ad. A globális minimum az inverz γ -kanyar, vagyis a γ_L konformer. Megállapítható tehát, hogy a fehérjékben fellelhető alanin aminosavak leírásában az RHF/PCM/3-21G elméleti szint nem mutat semmilyen előnyt sem a gázfázisú számításokhoz képest.

Minden magasabb szintű PCM számítás megegyezik a következőkben: (i) a globális minimum az α_L régióban található, (ii) nincs γ_L , δ_D ill. δ_L konformer (iii) egyes konformációs területeken a konformerek többszöröződése észlelhető (pl. α_L , β_L vagy ϵ_D , de sohasem α_D vagy γ_D). Érdekes, hogy a két ϵ_D konformer már a vonatkozó 30°-os felbontású térképeken is látszik, míg az α_L és a β_L konformerek multiplikációjára csak nagyobb felbontás (pl. a ϕ torzió mentén 5°-os) vet fényt. Az alacsony energiájú területek kiterjedtsége az α_L , a β_L (és részben az ϵ_L) konformerek körül jól összeegyeztethető az alanin ezen konformereinek fehérjékben tapasztalt kiemelkedő természetes gyakoriságával.

E.1.2. Táblázat A HCO-L-Ala-NH₂ modellpeptid eltérő elméleti szinteken PCM alkalmazása mellett optimált konformerei.

RHF/PCM/3-21G						
Konformer	ΔE	ω_b	ϕ	ψ	ω_l	
α_D	0,83	176,65	58,47	40,81	177,66	
β_L	0,34	-180,00	-165,15	166,34	177,01	
δ_D	4,30	176,96	-174,03	-46,64	-177,82	
δ_L^a	0,69	-175,49	-117,68	13,35	177,96	
ε_D	3,57	-175,22	69,43	171,58	179,55	
γ_D	1,58	173,72	73,99	-54,34	-178,39	
γ_L	0,00	-174,82	-86,25	65,46	-178,33	

RHF/PCM/6-31+G(d)						
Konformer	ΔE	ΔG^c	ω_b	ϕ	ψ	ω_l
α_D	2,11	3,04	175,09	62,15	37,34	176,15
α_L	0,00	0,71	-176,46	-84,18	-20,32	-177,78
β_L I	0,37	0,53	-178,71	-140,70	146,70	176,57
β_L II	0,34	0,00	-179,58	-126,68	144,14	176,43
ε_D I	4,29	4,75	-175,77	74,34	172,37	-175,48
ε_D II	3,93	4,68	-179,29	59,62	-145,95	-176,55
ε_L	0,27	0,50	177,71	-80,32	147,42	176,59
γ_D	3,85	5,13	174,95	76,40	-49,09	-177,15

B3LYP/PCM/6-31+G(d)						
Konformer	ΔE	ΔG^c	ω_b	ϕ	ψ	ω_l
α_D	2,19	3,53	175,41	62,50	36,28	177,07
α_L I	0,00	1,02	-176,19	-87,03	-16,76	-178,41
α_L II	0,18	1,21	-176,38	-91,27	-13,78	-178,91
β_L I	0,24	0,85	-179,20	-140,42	144,68	176,17
β_L II ^b	0,24	0,00	-179,94	-127,21	140,24	176,40
β_L III	0,23	0,61	-179,54	-130,55	141,26	176,33
ε_D I	4,26	5,14	-176,80	75,43	171,75	-176,21
ε_D II	3,80	4,94	179,94	59,08	-143,00	-177,52
ε_L	0,26	0,72	179,08	-101,88	138,38	176,47
γ_D	2,84	4,92	174,50	72,84	-50,98	-177,88

B3LYP/PCM/6-311++G(d,p)						
Konformer	ΔE	ω_b	ϕ	ψ	ω_l	
α_D	1,99	175,43	62,53	36,15	177,07	
α_L I	0,00	-176,07	-87,96	-16,17	-178,53	
α_L II	0,10	-176,10	-94,18	-11,23	-179,44	
β_L I	0,14	-179,58	-138,61	145,94	176,00	
β_L II	0,10	-179,79	-129,16	142,49	176,13	
ε_D	3,52	-179,85	60,58	-146,72	-176,34	
γ_D	2,78	174,63	72,84	-51,84	-178,12	

^a valójában a γ_L régióban; ^b ennek a konformációnak a természetét megkérdőjelezhető, mivel a legalacsonyabb frekvenciája mindössze 8 cm⁻¹ (ld. **E.1.5. Táblázat**) és szorosabb konvergencia kritérium mellett végzett optimálás hatására átalakult β_L III konformerré; ^c frekvenciaszámításokból származtatott adat.

Globális minimumok Hartree egységben: γ_L , $E_h = -412,500859$ (RHF/PCM/3-21G); α_L , $E_h = -414,831704$ (RHF/PCM/6-31+G(d)); α_L I, $E_h = -417,272104$ (B3LYP/PCM/6-31+G(d)); α_L I, $E_h = -417,388282$ (B3LYP/PCM/6-311++G(d,p)). Az oldószer-korrekcióval nyert szabadentalpia globális minimuma β_L II: $G_h = -414,727902$ (RHF/PCM/6-311++G(d,p)), $G_h = -417,179583$ (B3LYP/PCM/6-311++G(d,p)).

Mind a 19 konformert, melyeket RHF/PCM/6-311++G(2d,2p)//RHF/6-31G(d) számítások jósoltak²⁰ teljes optimalásnak vettem alá RHF/PCM/6-31+G(d) elméleti szinten (kiindulási geometriák és esetleges migrációk tekintetében ld. **E.1.3. Táblázat**). Ezen az úton is eljutottam a nyolc RHF/PCM/6-31+G(d) minimumhoz, míg a többi 11 jósolt konformer léte nem igazolódott.

E.1.3. Táblázat A HCO-L-Ala-NH₂ modellpeptid RHF/PCM/6-311++G(2d,2p)//RHF/6-31+G(d) elméleti szintű számítással azonosított konformereinek optimalása és esetleges migrációja RHF/PCM/6-31+G(d) elméleti szinten.

Konformer ^a	$\Delta E_{\text{jósolt}}^a$	$\phi_{\text{jósolt}}^a$	$\psi_{\text{jósolt}}^a$	konformer ^b	ΔE^b	migráció ^b
C5	0,0	-150	150	β_L I	0,37	
β_1	0,4	-120	150	β_L II	0,34	
B	0,4	-120	15			α_L
α_R	0,7	-75	-30	α_L	0,00	
PII	1,2	-60	150	ϵ_L	0,27	
PII	1,3	-90	180			ϵ_L
C7 _{eq}	1,5	-90	75			ϵ_L
α_L	2,2	75	30	α_D	2,11	
2 ₁	2,6	-165	75			ϵ_L
H	2,7	-90	-150			ϵ_L
H	3,6	-150	-135			β_L I
α_D	3,9	60	180	ϵ_D I	4,29	
C7 _{ax}	4,0	75	-60	γ_D	3,85	
α_D	4,3	60	-120			ϵ_D II
α_D	4,4	60	-150	ϵ_D II	3,93	
α_D	4,5	75	135			ϵ_D I
H	4,7	-90	-105			α_L
X	4,9	165	-15			α_L
α'	5,0	-135	-75			α_L

^a Elnevezés és adatok Iwaoka szerint.²⁰

^b Elnevezés és adatok az E.1.2. **Táblázat** szerint.

Az eltérő elméleti szintű Ramachandran-felszínek összetevése

A korrelációs együtthatók és a lineáris regressziós adatok (ld. **E.1.4. Táblázat**) jól jellemzik az egyes elméleti szinteken kapott PCM energiák összevethetőségét. Az RHF/PCM/3-21G, RHF/PCM/6-31+G(d) és B3LYP/PCM/6-31+G(d) elméleti szintek jelentős korrelációt mutatnak. Ha a pontoknak csak azt a felét vesszük figyelembe, melyeknél a B3LYP/PCM/6-31+G(d) szinten $\Delta E < 5 \text{ kcal.mol}^{-1}$, az RHF és a B3LYP módszerek közötti korreláció a 6-31+G(d) bázison szinte változatlan (R^2 0,9838-ról 0,9671-re csökken). Ugyanezen pontokat tekintve az RHF/PCM/3-21G energiák összevethetősége bármely magasabb szint tekintetében határozottan csökken.

E.1.4. Táblázat A HCO-L-Ala-NH₂ modellpeptid eltérő elméleti szinten PCM-mel kiszámított Ramachandran-felcsíneinek korrelációja. Az energiefelcsíneket a ϕ és a ψ paraméter mentén 30°-onként optimált szerkezetek energiájával jellemeztük.

x	RHF/3-21G	RHF/3-21G	RHF/6-31+G(d)
y	RHF/6-31+G(d)	B3LYP/6-31+G(d)	B3LYP/6-31+G(d)

A:

m	0,9299	0,8068	0,8448
b	0,0318	-0,1200	-0,0142
R^2	0,8826	0,9160	0,9838

B:

m	0,7724	0,6950	0,8544
b	0,4041	0,1521	-0,0669
R^2	0,6018	0,6456	0,9671

A: Mind a 144 pont összevetve.

B: Csak az a 79 pont összehasonlítva, ahol, $\Delta E < 5 \text{ kcal}\cdot\text{mol}^{-1}$ a B3LYP/PCM/6-31+G(d) elméleti szinten.

x és y : korreláltatott ΔE adatok a jelölt elméleti szinteken.

m és b : Az $y=mx+b$ lineáris regressziós egyenlet paraméterei.

R^2 : x és y korrelációja.

Frekvenciaszámítások

Frekvenciaszámításokat végeztem az RHF/PCM/6-31+G(d) és a B3LYP/PCM/6-31+G(d) konformerekre azonos elméleti szinteken. Néhány frekvenciát (hullámszámot²⁴) az **E.1.5. Táblázat** mutat be. Habár a "legalacsonyabb frekvencia" minden szerkezet esetében pozitív, $\beta_{\text{L}}\text{II}$ konformernél e két elméleti szinten mindössze 14 ill. 8 cm^{-1} , megkérdőjelezve, hogy e szerkezet valódi minimum volna. Valóban, a szorosabb (tight) konvergencia kritérium alkalmazása a B3LYP/PCM/6-31+G(d) szintű optimálás során a $\beta_{\text{L}}\text{II}$ szerkezet $\beta_{\text{L}}\text{III}$ -má alakulását okozta.

A HCO-L-Ala-NH₂ modellpeptid 42 normálrezgése közül kijelöltem az amid A, az amid I és az amid II frekvenciákat. Az amid A rezgések adják a legmagasabb frekvenciákat: A 40. frekvencia NH nyújtás, míg a 41. és a 42. szimmetrikus és antiszimmetrikus NH₂ nyújtás. Legnagyobb intenzitással az amid I frekvencia rendelkezik, mely főleg CO nyújtás, de az amidsík egyéb nyújtásaival és hajlításaival kombinálódik. Több konformerben (pl. γ_{D} RHF ill. α_{D} B3LYP módszerrel, továbbá minden α_{L} típusú konformer) a 33. vibráció egyértelműen hozzárendelhető a -CONH₂ atomcsoporthoz, míg a 34. HCONH- molekularészlethez (ld. **B.1/A Ábra**). Más konformerekben e két csoport CO nyújtásai nem válnak szét. Az ϵ_{L} és az $\epsilon_{\text{N}}\text{I}$, továbbá a β_{L} típusú konformerekben mindkét frekvencia azonos mértékben okoz rezgést a két amidcsoporton. A 31. és a 32. frekvencia amid II rezgés, főként

NH ill. NH₂ síkbeli hajlítás. Az amid III frekvenciák kiválasztását más rezgésekkel való erős kombinálódásuk akadályozta.

E.1.5. Táblázat A HCO-L-Ala-NH₂ modellpeptid RHF/PCM/6-31+G(d) és B3LYP/PCM/6-31+G(d) elméleti szinten azonosított konformereinek rezgési adatai.

Vibrációtípus.		első	amid II			amid I		amid A	
N ^o		1	31	32	33	34	40	41	42
RHF/PCM/6-31+G(d)									
α_D	Frek	69	1678	1759	1824	1839	3610	3684	3808
	<i>int</i>	7	333	200	747	908	281	211	218
α_L	Frek	44	1691	1761	1816	1835	3607	3679	3808
	<i>int</i>	17	569	205	881	809	256	221	228
β_L I	Frek	32	1708	1767	1810	1837	3630	3660	3777
	<i>int</i>	10	473	179	1597	238	280	236	222
β_L II	Frek	14	1709	1766	1810	1836	3616	3662	3779
	<i>int</i>	13	507	179	1593	239	277	232	221
ϵ_D I	Frek	34	1693	1774	1828	1852	3620	3676	3798
	<i>int</i>	12	387	181	612	1071	255	210	196
ϵ_D II	Frek	49	1682	1774	1815	1835	3622	3691	3863
	<i>int</i>	43	366	157	1091	637	289	203	186
ϵ_L	Frek	39	1700	1766	1822	1839	3619	3664	3780
	<i>int</i>	16	502	178	1293	447	259	225	216
γ_D	Frek	52	1691	1768	1817	1832	3602	3688	3841
	<i>int</i>	19	350	351	1081	516	278	182	306
B3LYP/PCM/6-31+G(d)									
α_D	Frek	60	1535	1608	1681	1695	3324	3428	3541
	<i>int</i>	7	196	149	696	772	248	184	180
α_L I	Frek	37	1541	1606	1679	1690	3302	3424	3544
	<i>int</i>	17	382	166	794	757	237	190	189
α_L II	Frek	32	1547	1606	1678	1690	3308	3425	3547
	<i>int</i>	18	388	167	806	741	227	189	188
β_L I	Frek	29	1553	1618	1671	1695	3340	3403	3512
	<i>int</i>	9	239	155	1399	270	245	217	184
β_L II	Frek	8	1554	1617	1671	1694	3326	3405	3514
	<i>int</i>	12	249	146	1456	232	243	211	182
β_L III	Frek	22	1555	1618	1671	1694	3328	3403	3512
	<i>int</i>	12	248	149	1441	241	244	214	183
ϵ_D I	Frek	28	1549	1625	1683	1707	3334	3419	3531
	<i>int</i>	13	241	150	508	944	229	188	161
ϵ_D II	Frek	42	1540	1623	1676	1692	3337	3436	3599
	<i>int</i>	41	201	118	963	558	257	178	146
ϵ_L	Frek	26	1553	1617	1675	1694	3319	3405	3514
	<i>int</i>	17	286	131	1413	236	237	207	180
γ_D	Frek	60	1546	1613	1675	1686	3312	3411	3528
	<i>int</i>	13	161	310	1172	309	256	183	359

Frek: Skálázatlan frekvenciák (hullámszámok) cm⁻¹ egységben, *int*: intenzitás km.mól⁻¹ egységben. N^o: A rezgés sorszáma a modellpeptid 42 fundamentális rezgésének sorában.

Az **E.1.5. Táblázat** a frekvenciákat skálázás nélkül adja meg, mivel a gázfázisú adatokkal nyert skálafaktorok alkalmazhatósága PCM számításokra nem bizonyított. Meg kell jegyezni azonban, hogy az RHF módszerrel kapott adatok sokkal nagyobbak, mint (a) azonos bázison a B3LYP módszer eredményei, továbbá (b) néhány modellrendszerre, pl. N-metil-acetamidra,²⁵ ill. fehérjékre,²⁷ az irodalomban fellelhető kísérleti adatok. Megállapítottam, hogy az RHF/PCM/6-31+G(d) amid A, amid I és amid II frekvenciák 0,9241-del történő szorzása a vonatkozó B3LYP/PCM/6-31+G(d) adatokra igen jó becslést nyújt ($R^2=0,9998$). Tehát e két elméleti szint skálafaktorainak hányadosa 0,9241, ami összevethető a skálafaktorok gázfázisra már megalapozott értékeinek arányával (pl. $f_{\text{RHF/6-31G(d)}/f_{\text{B3LYP/6-31G(d)}}=0,9312$, $f_{\text{RHF/6-31+G(d)}/f_{\text{B3LYP/6-31G(d)}}=0,9330$).²² A fehérje másodlagos szerkezetek relatív IR frekvenciáit²⁷ nem tükrözik a számolt tendenciák. Feltételezhető, hogy erre a célra a HCO-L-Ala-NH₂ még túl rövid modell, nem tartalmazhatja az egyes téralkatokat stabilizáló hidrogénhidak mindegyikét.

A frekvenciaszámításokkal megkapjuk a konformerek szabadentalpiáját (ld. ΔG értékek az **E.1.2. Táblázatban**). E paraméter alapján a β_{LII} konformer lesz a globális minimum mindkét elméleti szinten. Az L-völgy valamennyi minimuma 1,5 kcal.mol⁻¹-nál kedvezőbb, míg a D-völgyé egységesen 3 kcal.mol⁻¹ feletti.

Konformerek számított stabilitásadatainak korrelációja fehérjék alaninjainak konformációs változatosságával

A PDB SELECT¹⁰⁸ fehérjéinek (2002. áprilisi frissítés) alanin aminosavairól gyűjtött kísérleti adatokat összevettem a PCM számítások eredményeivel egy korábban kidolgozott eljárás szerint¹⁴ (ld. **MÓDSZEREK**). Jelentős átfedések elkerülése érdekében az RHF/PCM/6-31+G(d) elméleti szinten a β_{LI} konformert; a B3LYP/PCM/6-31+G(d) szinten az α_{LII} , a β_{LI} és II konformereket; míg a B3LYP/PCM/6-311++G(d,p) szinten az α_{LII} és a β_{LI} konformert hagytuk figyelmen kívül.

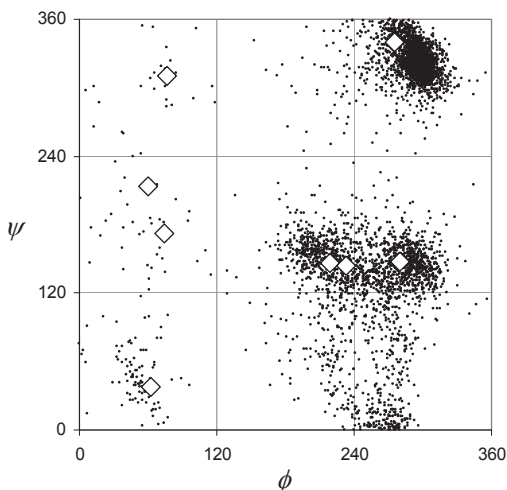
A reprezentatív minimumok köré rajzolt 15° sugarú körökkel mindössze a kísérleti [ϕ, ψ] párok 4-14% jelölhető meg (ld. **E.1.6. Táblázat**). Az elfogadás sugarát 30°-ra növelve az asszignáció 12-55%-ra növekszik. E tekintetben az RHF/PCM/6-31+G(d) kiválósága a többi elméleti szint felett arra vezethető vissza, hogy itt az ϵ_{L} konformer valódi minimum, mindamelltt kevésbé negatív ϕ torziószöggel rendelkezik (−80°, ld. **E.1.2. Ábra**) mint B3LYP/PCM/6-31+G(d) szinten ($\phi = -102^\circ$).

E.1.6. Táblázat Korreláció a fehérjékben talált alanin konformerek természetes gyakoriságának logaritmusával és a különböző elméleti szinteken kiszámolt relatív energiák között.

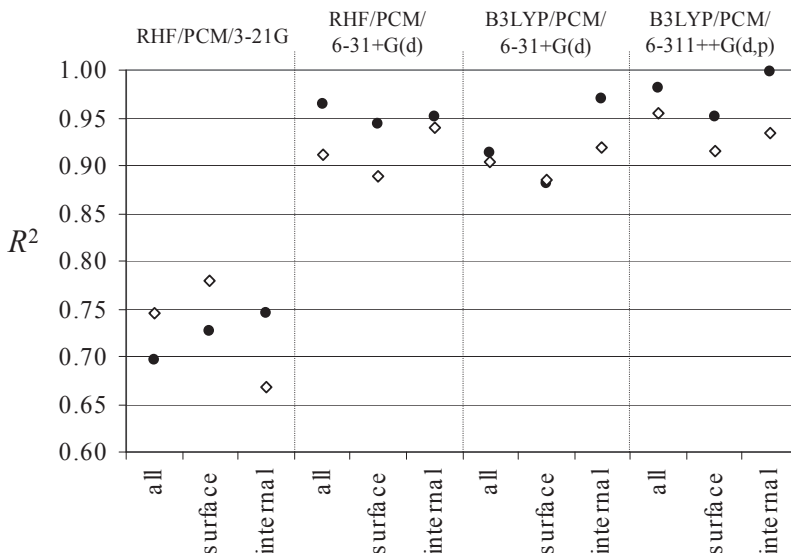
elméleti szint	alanin	R^2 (15°)	% besorolt (15°)	R^2 (30°)	% besorolt (30°)
RHF/PCM/3-21G	mind	0,7447	5	0,6968	12
RHF/PCM/3-21G	felszíni	0,7788	4	0,7265	12
RHF/PCM/3-21G	belső	0,6681	5	0,7449	13
RHF/PCM/6-31+G(d)	mind	0,9118	14	0,9641	54
RHF/PCM/6-31+G(d)	felszíni	0,8889	13	0,9433	54
RHF/PCM/6-31+G(d)	belső	0,9400	14	0,9508	55
B3LYP/PCM/6-31+G(d)	mind	0,9035	11	0,9132	39
B3LYP/PCM/6-31+G(d)	felszíni	0,8858	12	0,8802	40
B3LYP/PCM/6-31+G(d)	belső	0,9197	10	0,9697	38
B3LYP/PCM/6-311++G(d,p)	mind	0,9551	8	0,9812	31
B3LYP/PCM/6-311++G(d,p)	felszíni	0,9160	8	0,9517	31
B3LYP/PCM/6-311++G(d,p)	belső	0,9336	7	0,9977	31

% besorolt: a kísérletileg meghatározott $[\phi, \psi]$ párok azon százaléka, mely valamely kiszámolt minimum köré rajzolt 15° ill. 30° sugarú körbe esik a Ramachandran-térképen.

A kísérleti konformációs gyakoriságok logaritmusával jól korrelál a számított ΔE minden elméleti szinten, az RHF/PCM/3-21G kivételével (ld. **E.1.6. Táblázat** és **E.1.3. Ábra**). A fehérjék belsejében található alaninok erősebb vonzódást mutatnak a Ramachandran-térkép „mély minimumaihoz” (α_L , β_L és ϵ_L), mint a felszíni aminosavak, mely utóbbiak gyakran vesznek fel ϵ_D vagy γ_D konformációt is.¹⁴ Egyszerű számolási módszertünk azonban nem ad megbízható alapot a belső és a felszíni alaninok korrelációs különbségének mélyebb elemzésére.



E.1.2. Ábra Kísérleti és számított $[\phi, \psi]$ párok a Ramachandran-felszínen. Pontok: PDB SELECT-ből kigyűjtött 5393 alanin adatai; üres rombusz: az RHF/PCM/6-31+G(d) elméleti szinten optimált konformerek adatai.



E.1.3. Ábra Korreláció az alanin konformerek fehérjékben talált természetes gyakoriságának logaritmusával és a különböző elméleti szinteken kiszámolt relatív energiák között. Egy kísérletileg meghatározott $[\phi, \psi]$ pár (all: mind, surface: felszíni, internal: belső) akkor sorolható egy adott konformerhez, ha a kiszámolt minimum köré rajzolt 15° (üres rombusz) ill. 30° (teli kör) sugarú körbe esik a Ramachandran-térképen.

Az oldószer hatásának modellezése

A $\text{CH}_3\text{CHO-L-Ala-NHCH}_3$ modellpeptid konformereit vizes közegben B3LYP/6-31G(d) elméleti szinten azonosították (1) Onsager modellel, (2) SCIPCM-mel, (3) négy explicit vízmolekulával oldószeres közeg nélkül, végül (4) Onsager modell és négy explicit vízmolekula alkalmazásával.¹¹⁴ Az első két megközelítés a gázfázisú számításokhoz hasonló eredményt adott, mivel a γ_L konformer lett a globális minimum. A másik kettőnél γ_L eltűnt és a globális minimum ϵ_L ill. α_L lett. Az RHF/PCM/6-31+G(d) és B3LYP/PCM/6-31+G(d) eredményekkel szemben, ahol az L-völgy minimumai határozottan kedvezőbb energiájúak a D-völgy konformereinél, explicit vízmolekulákkal β_L relatív energiája eléri az $1,9 \text{ kcal.mol}^{-1}$ értéket, miközben α_D -é mindössze $1,2 \text{ kcal.mol}^{-1}$. Az **E.1.2. Ábra** tanúsága szerint a fehérjék alaninjai számára a fő konformációs gócpontok α_L , ϵ_L és β_L , míg α_D sokkal kevésbé kedvező szerkezet. Így a fehérjék alanin aminosavait jobban modellezzük az RHF/PCM/6-31+G(d) és B3LYP/PCM/6-31+G(d) eredmények, mint az explicit vízmolekulákkal végzett számítások

gázfázisban ill. Onsager modellel. Azt azonban, hogy az explicit vízmolekulák figyelembevétele a PCM módszerrel együtt további fejlődést jelentene-e, nem vizsgáltam.

Prolin

Kísérleti adatok statisztikai elemzése

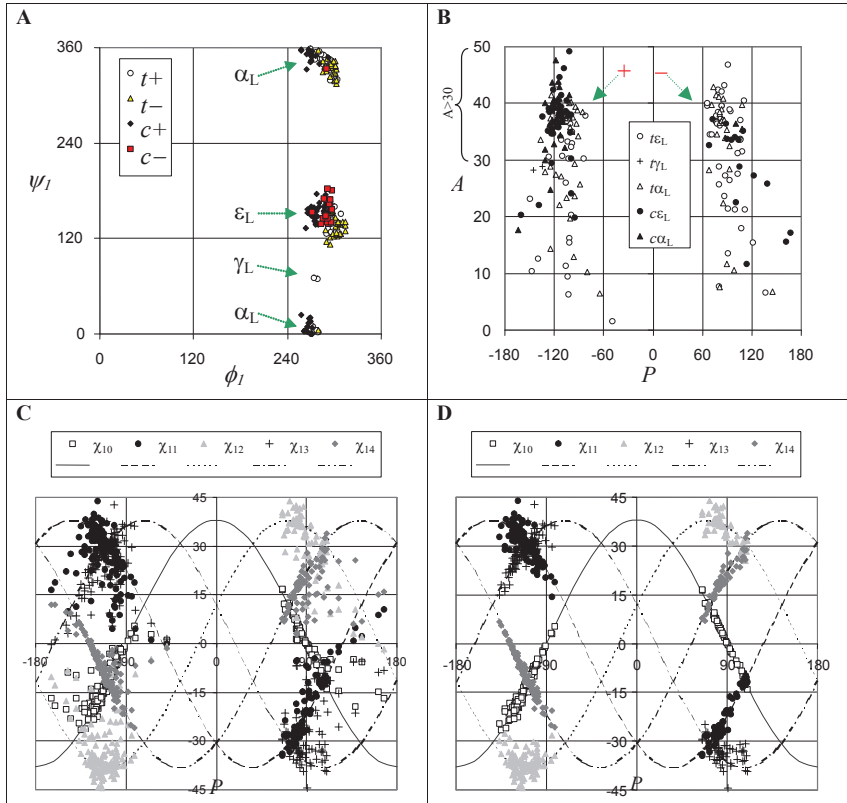
A Cambridge-i Szerkezeti Adatbázisban (CSD, 2001. áprilisban) 106 olyan peptidet találtunk, amely legalább egy Xxx–Pro–Yyy szekvenciát tartalmazott. Összesen 214 Xxx–Pro–Yyy szekvenciát^b gyűjtöttünk ki. A szekvenciákat a prolinok térszerkezete alapján csoportokba soroltam és kóddal jelöltem (ld. **M.5. Ábra**). A *cisz* és a *transz* konformáció megkülönböztetése egymástól egyszerű feladatnak mutatkozott csakúgy, mint a pozitív és a negatív gyűrűállásoké. Az ϵ_L gerinckonformációnak α_L -tól, ill. γ_L -tól való elkülönítésére azonban a $\psi = 120^\circ$ és $\psi = 0^\circ$ határértékek alkalmatlannak bizonyultak, mivel az ϵ_L típusú konformerek ψ értékei között olyan alacsonyok is előfordultak, mint 112° , míg az α_L típusúaké elérhette a 24° -ot is (ld. **E.2.1/A. Ábra**).

Az ideálisan várható 12 konformer közül 9-et találtam meg. Az **E.2.1. Táblázat** összefoglalja a 214 szekvencia mindegyikének figyelembevételével származtatott (ezen túl CSD(all) halmazként hivatkozott) konformerek előfordulását és geometriai adatait. Az ω_1 , ϕ_2 , A , P és ψ_2 paraméterek az egyedi szekvenciák vonatkozó értékeinek átlagai. A fenti paraméterek szórását a következő általános képlettel számítottam:

$$\sqrt{\frac{n \sum x^2 - (\sum x)^2}{n(n-1)}} \quad (20)$$

^b A peptidek azonosító kódja a CSD-ben: ACTDGU10, ALPRAL10, ANTAHC10, ANTAML10, BEHTEN, BIHTUH10, BIHXUL10, BINJIR, BIPHIR10, BRAXGU, BUBLUF, BUFTOL, BUNYEO10, BUYXOI, CACNOJ10, CALCAT, CALCIB, CAZSEB, CEJSUF, CEWCIO10, CGLPGL, CGPGAP10, CGPSAQ, CINCYED, COJRIC10, COXJII, COXJOO, COXKAB, CPRGCA20, CYBGPP, CYCAPP, CYDGPAA, DABVIL, DASXIE, DETWUU, DEWFEQ, DOLJAP, DUPFAV, DUPKEE, DUTLAF10, DUVGOQ10, DUYTIA, FIVSAE, FUDWIK, GICHOP, GIDJAE, GIDNUC, GIPKAR10, GUFXOU, HEBHUR, HEBKUU, JEHMAK, JINGAO, JUJHUR, JUXHAL, KEFWUN, KIVDIC, KUFRAE, LACSUD, LEBDEB, LEJWUS, LENKIY, LINCOA, NEHYAA, NEHYAA01, NEHYAA02, NEHYAA03, NIGNUM, NUWNEY, PAANTD, PAANTD01, PALJOO, PALJUH, PAPGAP, PAPRVA, POHMUU, PROGLY20, PVANSB, PVANTS, PVPALM, PVPALM10, SEFTIG, SIMVIT, SOWGOA, SUKKAK, SUMNOD, TALVAD, TISKOV, TOSWED, VEDJIX, VENXER, VIGSIN, VOMLEO, VORZEH, VOWZUC, VPGVGB10, WEXPOE, WEXPUK, YAGKOG, YAGKOG10, YUWXUJ, ZAJPAB, ZAZQIA, ZEVBEH, ZOHMIS, ZORRED.

A χ_{20} , χ_{21} , χ_{22} , χ_{23} és χ_{24} torziók jellemző értékeit A és P átlagainak ismeretében a **(4) Egyenlet** segítségével számoltam ki. Azoknak a konformációknak a törlésével, amelyekben valamelyik A érték 30° -nál kisebb, egy második adathalmazt származtattam, melyre ezentúl CSD($A>30$) halmazként hivatkozom (ld. **E.2.2. Táblázat**).



E.2.1. Ábra Kísérletileg meghatározott Xxx–Pro–Yyy szekvenciák a Cambridge-i Szerkezeti Adatbázisból, CSD(all). A geometriai paraméter dimenziója fok. **A:** Ramachandran-térkép. Jól megfigyelhető, hogy számos ϵ_L és α_L konformer az γ_L vízgyűjtőterület ideális határán túla kerül. **B:** Pseudorotációs térkép. A diagram felső része, ahol tehát A legalább 30° , kijelöli a CSD($A>30$) adathalmazt. **C, D:** Az oldallánc torziósögei a pseudorotációs fázis függvényében. A görbék az ideális χ_{1j} értékeket mutatják $A=38^\circ$ esetére számítva (ld. **(3.1) Egyenlet**). **C:** CSD(all). **D:** CSD($A>30$).

E.2.1. Táblázat A CSD-ből gyűjtött 214 Xxx–Pro–Yyy szekvencia jellemző adatai (a továbbiakban CSD(all)-ként hivatkozva).

Konf.	<i>n</i>	<i>p</i>	ω_0	ϕ_1		A		P		χ_{10}	χ_{11}	χ_{12}	χ_{13}	χ_{14}	ψ_1		
<i>t</i> _{E_L} ⁺	26	12	177	±9	-66	±7	25	±11	-106	±20	-7	20	-25	21	-9	148	±12
<i>r</i> _L ⁺	2	1	179	±1	-84	±3	29	±0	-139	±8	-22	28	-24	11	7	70	±1
<i>t</i> _{α_L} ⁺	34	16	-174	±5	-75	±10	31	±10	-106	±16	-8	24	-31	26	-11	-13	±14
<i>t</i> _{E_L} ⁻	39	18	176	±6	-60	±7	32	±10	90	±16	0	-18	30	-30	19	140	±14
<i>t</i> _{α_L} ⁻	26	12	-176	±3	-65	±7	32	±11	88	±15	1	-20	31	-30	18	-25	±11
<i>c</i> _{E_L} ⁺	32	15	8	±12	-83	±8	36	±7	-116	±14	-16	32	-36	26	-6	153	±11
<i>c</i> _{α_L} ⁺	39	18	8	±7	-92	±6	38	±5	-118	±10	-18	34	-37	26	-5	-1	±12
<i>c</i> _{E_L} ⁻	15	7	-3	±12	-70	±6	28	±8	109	±29	-9	-8	23	-28	23	155	±15
<i>c</i> _{α_L} ⁻	1	0	-9		-70		37		100		-6	-16	32	-36	26	-26	

Konf.: konformer a javasolt nevezéktan szerint jelölve (ld. **M.S. Ábra**); *n*: az adott konformációjú szekvenciák száma; *p*: konformer gyakorisága százalékban; ω_0 , ϕ_1 , *A*, *P* és ψ_1 : átlagértékek szórással; χ_{10} , χ_{11} , χ_{12} , χ_{13} és χ_{14} : átlagolt *A* és *P* értékekből számítva. A torziósögek (ω_0 , ϕ_1 , ψ_1 és az öt χ_{1j} torzió), valamint a pszeudotorációs paraméterek (*A* és *P*) dimenziója fok.

E.2.2. Táblázat Jellemző adatok 152 jó redőzöttséget mutató Xxx–Pro–Yyy szekvenciára vonatkozóan (a továbbiakban CSD(A>30)-ként hivatkozva). A konformereket az **E.2.1. Táblázat** alapadatait képező 214 szekvenciából A>30 megszorítással szűrtem ki.

Konf.	<i>n</i>	<i>p</i>	ω_0	ϕ_1		A		P		χ_{10}	χ_{11}	χ_{12}	χ_{13}	χ_{14}	ψ_1		
<i>t</i> _{E_L} ⁺	15	10	177	±9	-67	±8	34	±3	-103	±11	-8	26	-34	29	-13	147	±14
<i>t</i> _{α_L} ⁺	20	13	-174	±4	-76	±9	38	±2	-107	±15	-11	30	-38	31	-12	-13	±15
<i>t</i> _{E_L} ⁻	25	16	176	±7	-57	±7	38	±4	86	±14	3	-24	37	-35	20	138	±12
<i>t</i> _{α_L} ⁻	19	13	-176	±4	-64	±7	38	±3	85	±10	3	-25	37	-35	19	-24	±11
<i>c</i> _{E_L} ⁺	27	18	8	±12	-83	±7	38	±4	-114	±10	-16	33	-38	28	-8	153	±10
<i>c</i> _{α_L} ⁺	37	24	9	±7	-92	±5	38	±3	-117	±7	-17	34	-38	27	-6	-1	±12
<i>c</i> _{E_L} ⁻	8	5	-1	±14	-69	±4	35	±2	91	±15	-1	-20	33	-33	21	156	±18
<i>c</i> _{α_L} ⁻	1	1	-9		-70		37		100		-6	-16	32	-36	26	-26	

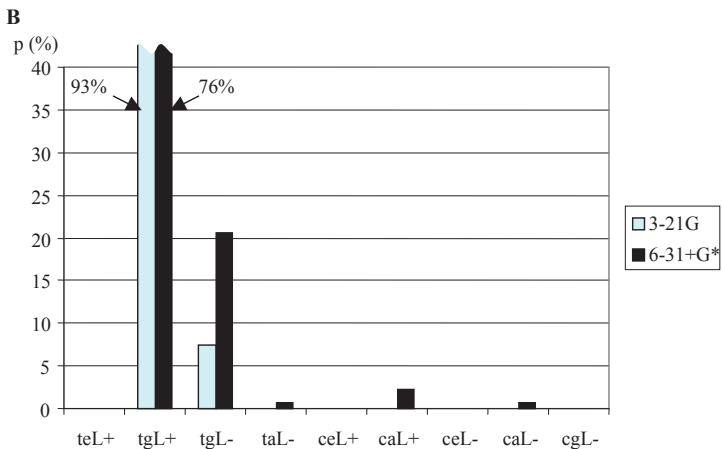
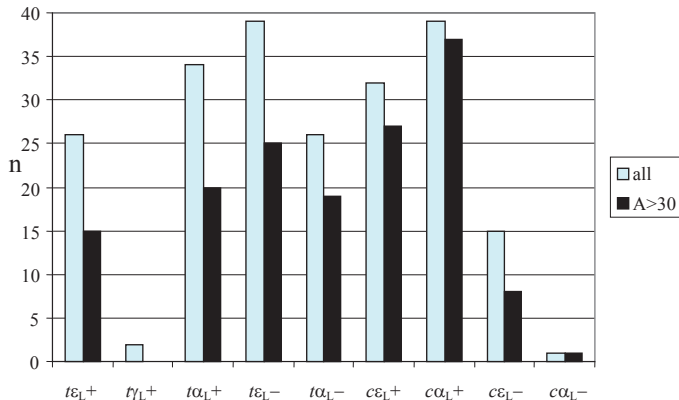
(ld. **E.2.1. Táblázat**)

Meglepően sok *cisz* konformer található: a CSD(all) 41 %-a, ill. a CSD(A>30) 48 %-a. Az ideálisan várható négy γ_L típusú konformer közül csak egy fordult elő (γ_{L+} , a CSD(all) két szekvenciájában). Az ϵ_L és α_L típusú konformerek mindegyike szerepel mindkét halmazban, de a negatív gyűrűállás *cisz* peptidkötés után ($c\epsilon_{L-}$ és $c\alpha_{L-}$) ritka. Általánosan megfigyelhető, hogy a *cisz* peptidkötés a *transz* peptidkötés utáni értékeikhez képest csökkenti ϕ_2 -t, növeli A -t (kivévelt képez az ϵ_{L-}), növeli P abszolút értékét, és növeli ψ_2 -t (kivévelt képez az α_{L-}). A gyűrűállások nem befolyásolják jelentősen a *transz* ω_1 értékeket, míg a pozitív gyűrűk ($P < 0$) a negatív állásúakhoz képest csökkentik a *cisz* ω_1 értékét, növelik ϕ_2 -t, és csökkentik A -t és ψ_2 -t is. (ld. **E.2.1/A** és **B. Ábra**)

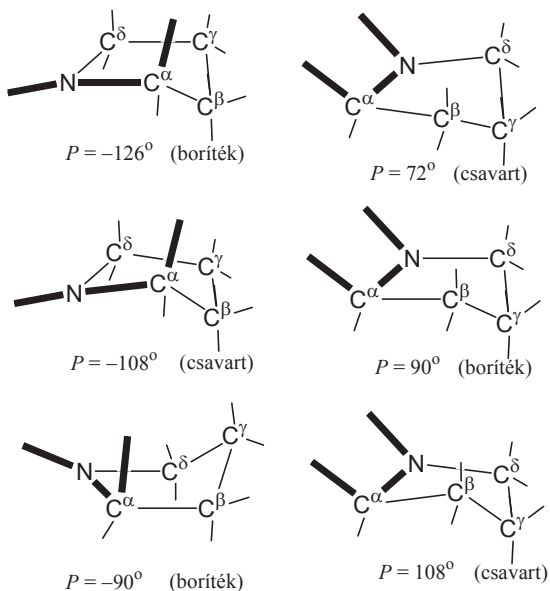
A CSD(all) és CSD(A>30) halmazok összevetése alapján, amint az az **E.2.1.** és **E.2.2. Táblázatból** kitűnik, az A>30 kikötés eltérően hat a *cisz* és a *transz* peptidkötésekre, nevezetesen több *transz* konformer tűnik el, közöttük az összes γ_{L+} . Az ω_1 , ϕ_2 és ψ_2 paraméterek és szórásaik nem mutatnak jellegzetes változást. Az A>30 kikötés értelemszerűen növeli A értékét és csökkenti a szórását. Kevésbé nyilvánvaló, hogy P abszolút értékét és szórását is csökkenti. A konformerek gyakoriságát az **E.2.2/A. Ábra** szemlélteti.

A CSD(A>30) halmaztól a torzítatlan gyűrűjú prolinok jobb jellemzését várjuk. Míg a CSD(all)-ban az egyes prolinok -164° -tól -50° -ig és 65° -tól 167° -ig veszik fel P értékeit, addig a CSD(A>30)-ban P mindig a -137° - -82° és 65° - 110° intervallumon belül található. Lényegében tehát a negatív és a pozitív P értékek szórása egyaránt megfelelődik. Az **E.2.1/C** és **D Ábra** megmutatja a CSD(all) és a CSD(A>30) halmazban az egyes prolinok öt χ értékének illeszkedését a **(4) Egyenlettel** számított χ görbékre, ahol P -360° -tól 360° -ig terjedő értékei mellett A mindvégig önkényesen 38° .^c A kísérleti χ értékek elhagyják az idealizált görbéket, hiszen A is szóródik 38° körül. A pirrolidingyűrű P -126° -os, -90° -os és 90° -os értéke mellett boríték konformációjú, míg -108° , 72° , és 108° esetén csavartnak tekinthető (ld. **E.2.3. Ábra**).

^c A görbék az öt χ érték idealizált változását mutatják a pozitív és a negatív gyűrűállások között P paraméter mentén lehetséges mindkét irányú átmenet során, miközben A mindvégig önkényesen 38° .



E.2.2. Ábra A prolin konformációs összetétele kísérleti és elméleti megközelítésben. **A:** Kísérletileg meghatározott Xxx-Pro-Yyy szekvenciák a Cambridge-i Szerkezeti Adatbázisból: CSD(all) ill. CSD(A>30) (214 ill. 152 prolin) **B:** A HCO-L-Pro-NH₂ modellpeptid konformációs megoszlása % egységben, *ab initio* számításokkal alapján.



E.2.3. Ábra A prolin pirrolidingyűrűjének fontosabb boríték és csavart konformációi.

Ab initio számítások

A 12 ideális konformer közül 9-et találtam RHF/3-21G elméleti szinten, mivel a $\tau\alpha_1+$, $\tau\epsilon_1-$, $\epsilon\gamma_1+$ nem adott nyugalmi szerkezetet. AZ RHF/3-21G minimumok további optimálása RHF/6-31+G(d) elméleti szinten 7 konformert adott, $\tau\epsilon_1+$ és $\epsilon\gamma_1-$ más minimumokba konvergált. Az optimálásra eltűnő minimumok egyike, $\epsilon\gamma_1-$, RHF/3-21G elméleti szinten a legmagasabb energiájú konformer. A másik, $\tau\epsilon_1+$, pedig a modellpeptid konformációs hiperfelszínén a globális minimum, $\tau\gamma_1+$, igen közeli szomszédja. A **E.2.4.** és **E.2.5. Táblázat** bemutatja a konformerek jellegzetes térszerkezeti adatait (a továbbiakban 3-21G és 6-31+G* halmazként hivatkozva).

E.2.4. Táblázat A HCO-L-Pro-NH₂ modellpeptid RHF/3-21G elméleti szintű *ab initio* számítással azonosított konformerjeinek jellemző adatai (a továbbiakban 3-21G-ként hivatkozva).

Konf.	<i>E</i>	ΔE	<i>p_{rel}</i>	<i>P</i>	ω_b	ϕ_1	χ_{10}	χ_{11}	χ_{12}	χ_{13}	χ_{14}	<i>A</i>	<i>P</i>	ψ_1	ω_1
<i>r</i> _L +	-488,943851	6,59	0	0,00	175,80	-70,94	-15,55	33,66	-39,76	30,11	-9,20	39,79	-113,00	150,58	178,27
<i>r</i> _L +	-488,954348	0,00	100	92,59	-173,04	-83,28	-15,22	33,51	-39,98	30,28	-9,56	39,84	-112,46	67,96	-178,84
<i>r</i> _L -	-488,951951	1,50	8	7,40	-171,78	-82,74	-12,74	-11,63	30,27	-37,01	31,69	38,13	109,52	68,73	-178,84
<i>t</i> _L -	-488,941003	8,37	0	0,00	-171,38	-74,53	4,55	-26,72	38,49	-35,02	19,38	39,13	83,32	-23,41	-179,27
<i>c</i> _L +	-488,941638	7,98	0	0,00	-7,69	-62,42	-13,41	32,75	-40,21	31,83	-11,64	40,14	-109,52	172,38	178,54
<i>c</i> _L +	-488,945374	5,63	0	0,01	10,28	-96,40	-17,13	34,55	-39,65	29,22	-7,59	39,93	-115,40	2,06	179,97
<i>c</i> _L -	-488,940219	8,87	0	0,00	-3,72	-51,15	6,18	-27,40	38,19	-33,81	17,57	38,51	80,76	178,56	179,31
<i>c</i> _L -	-488,937928	10,30	0	0,00	2,02	-59,86	4,44	-26,19	37,98	-34,28	19,04	38,43	83,37	101,18	-178,18
<i>c</i> _L -	-488,943170	7,01	0	0,00	9,35	-67,48	10,00	-30,04	39,00	-32,32	14,12	38,83	75,08	-26,09	-176,85

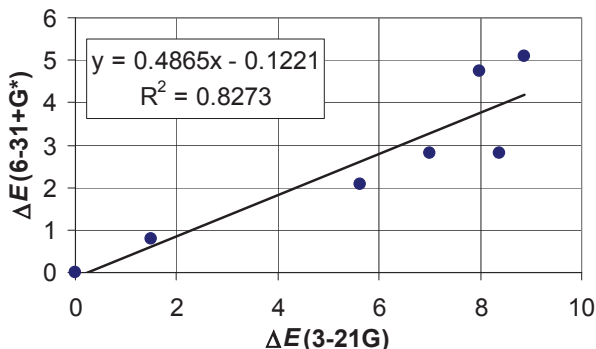
Konf: konformer a javasolt nevezéktan szerint jelölve (ld. **M.S. Ábra**); *E*: Energia Hartree egységben; ΔE : relatív energia kcal.mol⁻¹ egységben; *A* geometriai paraméterek (ω_b , ϕ_1 , ψ_1 , χ_{1j} , *A* és *P*) dimenziója fok. *A*, *P*: pszeudorotációs amplitúdó és fáziszög az ω_{1j} torziósögből kiszámítva; *p_{rel}*: a konformer becslült gyakorisága a ***r*_L+** konformerhez képest 300 K-en, % egységben; *p* a konformer abszolút gyakorisága 300 K hőmérsékleten, %-ban megadva.

E.2.5. Táblázat A HCO-L-Pro-NH₂ modellpeptid RHF/6-31+G(d) elméleti szintű *ab initio* számítással nyert konformerjeinek jellemző adatai (a továbbiakban 6-31+G*-ként hivatkozva).

Konf.	<i>E</i>	ΔE	<i>p_{rel}</i>	<i>P</i>	ω_b	ϕ_1	χ_{10}	χ_{11}	χ_{12}	χ_{13}	χ_{14}	<i>A</i>	<i>P</i>	ψ_1	ω_1
<i>r</i> _L +	-491,702066	0,00	100	76,29	-171,92	-85,21	-13,72	31,14	-37,41	28,60	-9,40	37,24	-111,62	72,12	-174,94
<i>r</i> _L -	-491,700827	0,78	27	20,68	-170,01	-84,48	-8,85	-15,34	32,60	-37,22	29,34	38,24	103,38	77,98	-172,34
<i>t</i> _L -	-491,697597	2,80	1	0,69	-166,90	-79,54	0,37	-23,20	36,65	-35,90	22,55	38,44	89,45	-10,69	171,66
<i>c</i> _L +	-491,694496	4,75	0	0,03	-1,88	-71,89	-16,27	31,90	-36,28	26,32	-6,24	36,55	-116,43	159,71	179,71
<i>c</i> _L +	-491,698744	2,08	3	2,30	10,08	-91,89	-14,88	31,81	-37,37	28,14	-8,30	37,41	-113,44	-3,36	178,49
<i>c</i> _L -	-491,693965	5,08	0	0,01	-2,34	-52,99	4,44	-25,46	36,66	-33,42	18,39	37,28	83,16	163,59	179,61
<i>c</i> _L -	-491,697603	2,80	1	0,69	9,27	-74,18	4,33	-25,50	36,86	-33,70	18,57	37,50	83,37	-19,22	-177,68

(ld. **E.2.4. Táblázat**)

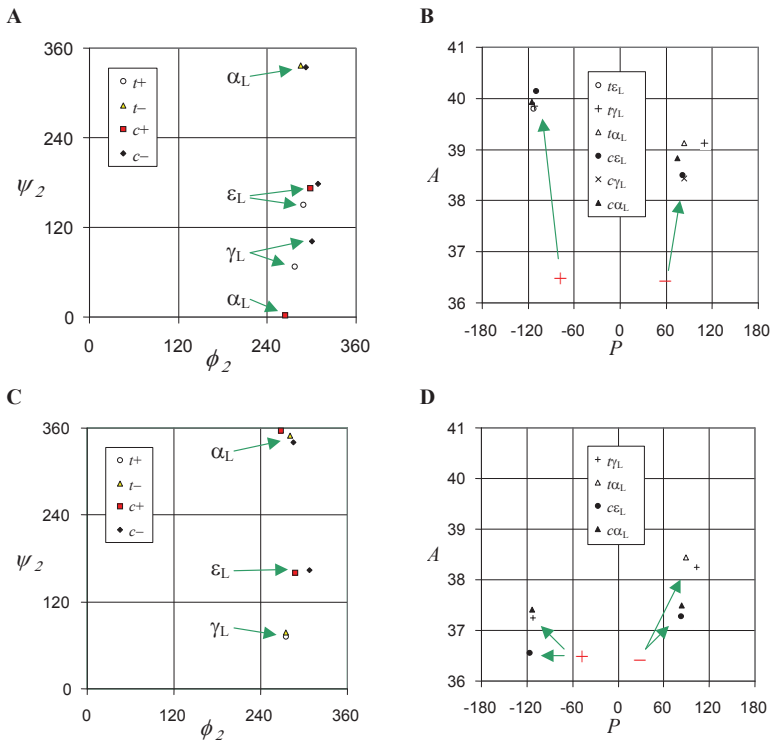
Mindkét halmaz globális minimuma $\tau_{\text{L}+}$. A globális minimumhoz viszonyított energiák korrelációját **E.2.4. Ábra** mutatja. A magasabb szinten a ΔE értékek kb. 50 %-kal csökkennek, így a várható gyakoriságok különbségei kevésbé szembeötlők. A **E.2.2/B. Ábra** szemlélteti a konformereknek a **(18) Egyenlet** segítségével becsült előfordulási gyakoriságát.



E.2.4. Ábra A HCO-L-Pro-NH₂ modellpeptid RHF/6-31+G(d) ill. RHF/3-21G elméleti szinten számolt relatív energiáinak korrelációja. ΔE dimenziója kcal.mol⁻¹.

A magasabb elméleti szint a konformerek geometriájának szisztematikus változásához vezet. Átlagosan az mondható, hogy ω_1 2,0°-ot növekszik, ϕ_2 3,2°-ot csökken, A 1,7°-ot csökken, P 0,9°-ot növekszik, miközben ψ_2 viszonylag nagy változásai (4°-15°) nullává átlagolódnak (ld. **E.2.5. Ábra**). Pozitív gyűrűállás mellett nagyobb A csökkenése: míg a 3-21G elméleti szint a pozitív állású gyűrűknél ad nagyobb amplitúdót, addig a 6-31+G(d) elméleti szint a negatívoknál (vö. **E.2.4/B** és **D Ábra**). A magasabb elméleti szint növeli a *transz* konformerek ψ_2 értékét, de csökkenti a *cisz* konformerekét (kivével $\alpha_{\text{L}-}$).

Baldoni és munkatársai a HCO-L-Pro-NH₂ modellpeptid csak *transz* peptidkötéssel rendelkező alkatait vizsgálva¹¹⁵ ismertették már az RHF/3-21G elméleti szint négy *transz* konformerét, de további optimalásaik a 6-31G(d) RHF elméleti szinten csak a hidrogénhiddal rendelkező $\tau_{\text{L}+}$ és $\tau_{\text{L}-}$ konformereket hagyták meg. Ezzel szemben RHF/6-31+G(d) elméleti szinten sikerült optimálnom a $\alpha_{\text{L}-}$ konformert, és csak $\tau_{\text{L}+}$ konvergált más szerkezeté.



E.2.5. Ábra A HCO-L-Pro-NH₂ modellpeptid *ab initio* konformereinek geometriai jellemzői. ϕ_2 , ψ_2 , P és A mértékegysége fok. **A**: RHF/3-21G konformerek helyzete a Ramachandran-térképen. **B**: RHF/3-21G konformerek helyzete a pseudorotációs térképen. **C**: RHF/6-31+G(d) konformerek helyzete a Ramachandran-térképen. **D**: RHF/6-31+G(d) konformerek helyzete a pseudorotációs térképen.

Kang a CH₃CO-L-Pro-NHCH₃ modellpeptid ECEPP/3 force field módszerrel kapott alacsony energiájú konformereit RHF/3-21G elméleti szinten optimalva 8 konformert kapott.¹¹⁶ Ezek közül 7 megfeleltethető a HCO-L-Pro-NH₂ modell általam megtalált $t\gamma_L+$, $t\gamma_L-$, $t\alpha_L-$, $c\epsilon_L+$, $c\epsilon_L-$, $c\alpha_L+$, ill. $c\alpha_L-$ konformerének, de míg az előbbi tanulmány $t\alpha_L+$ konformert ismertet, $t\epsilon_L+$ -t nem ad. A két vizsgálat mindegyikében megtalált konformerek globális minimumhoz, tehát $t\gamma_L+$ -hoz viszonyított energiái igen hasonlóak.

Összevetés

Jóllehet a CSD(all) és a 3-21G halmaz egyaránt kilenc minimumot ad, nem pontosan ugyanazokat a konformereket. Az egyes ideális konformerek meglétét ill. hiányát a különböző módszerekkel nyert halmazokban az **E.2.6. Ábra** foglalja össze.

		CSD				<i>ab initio</i>			
		–	+			–	+		
t r a n s z	}	ϵ_L	all A>30	all A>30	ϵ_L	 	3-21G 	}	t r a n s z
		γ_L	 	 	γ_L	3-21G 6-31+G*	3-21G 6-31+G*		
		α_L	all A>30	all A>30	α_L	3-21G 6-31+G*	 		
c i s z	}	ϵ_L	all A>30	all A>30	ϵ_L	3-21G 6-31+G*	3-21G 6-31+G*	}	c i s z
		γ_L	 	 	γ_L	3-21G 	 		
		α_L	all A>30	all A>30	α_L	3-21G 6-31+G*	3-21G 6-31+G*		
		–	+			–	+		

E.2.6. Ábra Az egyes ideális konformerek megléte ill. hiánya a prolin különböző módszerekkel meghatározott halmazaiban. CSD: A Cambridge-i Szerkezeti Adatbázisból kigyűjtött szekvenciák; *ab initio*: A HCO–L–Pro–NH₂ modellpeptid *ab initio* számítással nyert konformerei. Két-két halmaz képviseli mindegyik módszert: CSD(all) és CSD(A>30) a kísérleti, míg 3-21G és 6-31+G* a kvantumkémiai megközelítést.

A tizenkét ideális konformer közül hét az adatbázisban is megtalálható és optimálható is legalább az alacsonyabb elméleti szinten. Egyetlen ideális szerkezet, nevezetesen a $\mathbf{c}\gamma_{\mathbf{L}+}$, hiányzik a CSD-ből is és a számítás során is más alkattá konvergál. Ez összességében nyolc átfedést ad a kísérleti és a kvantumkémiai eredmények között. Csak kísérletileg található $\mathbf{t}\epsilon_{\mathbf{L}-}$ és $\mathbf{t}\alpha_{\mathbf{L}+}$, míg kizárólag *ab initio* megközelítéssel $\mathbf{t}\gamma_{\mathbf{L}-}$ és $\mathbf{c}\gamma_{\mathbf{L}-}$ konformer. E négy eltérés azzal magyarázható, hogy az *ab initio* módszer túlhangsúlyozza a $\gamma_{\mathbf{L}}$ konformerek, különösen a lokális hidrogénhidat tartalmazó $\mathbf{t}\gamma_{\mathbf{L}}$ szerkezetek szerepét, míg alábecsüli az $\epsilon_{\mathbf{L}}$ és az $\alpha_{\mathbf{L}}$ konformerek fontosságát, mely szerkezeti elemek csak más aminosavakkal tudnak hidrogénhidakat kialakítani. Mivel a CSD(A>30) $\mathbf{t}\gamma_{\mathbf{L}+}$, míg a 6-31+G* halmaz $\mathbf{t}\epsilon_{\mathbf{L}+}$ és $\mathbf{c}\gamma_{\mathbf{L}-}$ konformert sem tartalmaz, a kísérleti és az elméleti halmazok összevetésében az átfedések és eltérések aránya 7/5-re változik.

A különböző megközelítéssel nyert prolinkonformerek kvantitatív összevetését az **E.2.6. Táblázat** elemzi. Kizárólag a ϕ_2 , A , P és ψ_2 torziószögek különbségeinek figyelembevételével bevezettem egy összegző paramétert:

$$Dev = \sqrt{\Delta\phi_2^2 + \Delta A^2 + \Delta P^2 + \Delta\psi_2^2} \quad (21)$$

A konformerek kiválasztott geometriai jellemzőinek eltérése a kísérleti és az elméleti halmazok között rendszerint a kísérleti értékek szórásához hasonló mértékű. A legjobb illeszkedés a CSD(A>30) és a 6-31+G* adja. A {CSD(A>30)}-{6-31+G*} típusú különbségek abszolút értékei jelentősek a ϕ_2 (akár 16°), csekélyek az A (legfeljebb 3°), és átlagosan 9°-nyi a P , 8°-nyi a ψ_2 paraméter esetében. Az eltérés előjele csak egy esetben jósolható meg, nevezetesen ψ_2 értéke mintegy 6,3°-kal nagyobb a 6-31+G* halmazban.

Mind a négy halmaz között különösen jó illeszkedés tapasztalható a $\mathbf{c}\alpha_{\mathbf{L}+}$ konformerre. Ez a szerkezet igen gyakori a kísérleti halmazokban, míg a kvantumkémiai úton számított energiája a *cisz* típusú konformerek között a legalacsonyabb.

E.2.6. Táblázat Az eltérő módszerekkel kapott azonos kódú prolinkonformerek néhány geometriai paraméterének különbségei és a kísérleti adatok szórása.

Hivatkozott módszer ^a	Konf.	$\Delta\phi_1$	ΔA	ΔP	$\Delta\psi_1$	<i>Dev</i>
{CSD(all)}	$t\epsilon_{L+}$	6,51	11,24	19,82	11,63	26,40
{CSD(A>30)}	$t\epsilon_{L+}$	7,72	2,71	10,84	13,56	19,20
{CSD(A>30)}- {CSD(all)}	$t\epsilon_{L+}$	-1,88	8,44	2,73	-0,16	9,07
{CSD(all)}- {3-21G}	$t\epsilon_{L+}$	5,35	-14,36	6,94	-2,98	17,09
{CSD(A>30)}- {3-21G}	$t\epsilon_{L+}$	3,47	-5,92	9,67	-3,14	12,27
{CSD(all)}	$r\gamma_{L+}$	2,73	0,36	7,76	0,55	8,25
{CSD(all)}- {3-21G}	$r\gamma_{L+}$	-0,82	-11,31	-26,88	2,03	29,24
{CSD(all)}- {6-31+G*}	$r\gamma_{L+}$	1,11	-8,70	-27,72	-2,13	29,15
{6-31+G*}- {3-21G}	$r\gamma_{L+}$	-1,93	-2,60	0,84	4,16	5,34
{CSD(all)}	$t\alpha_{L+}$	9,84	9,68	16,31	13,70	25,38
{CSD(A>30)}	$t\alpha_{L+}$	9,21	2,40	14,72	14,73	22,90
{CSD(A>30)}- {CSD(all)}	$t\alpha_{L+}$	-1,21	6,81	-1,56	0,55	7,11
{CSD(all)}	$t\epsilon_{L-}$	7,06	9,95	15,67	13,81	24,19
{CSD(A>30)}	$t\epsilon_{L-}$	6,57	3,65	13,72	12,43	19,98
{CSD(A>30)}- {CSD(all)}	$t\epsilon_{L-}$	2,65	6,30	-4,11	-2,32	8,30
{6-31+G*}- {3-21G}	$r\gamma_{L-}$	-1,74	0,11	-6,14	9,25	11,24
{CSD(all)}	$t\alpha_{L-}$	7,20	10,95	15,15	10,99	22,84
{CSD(A>30)}	$t\alpha_{L-}$	6,61	2,78	9,71	10,60	16,07
{CSD(A>30)}- {CSD(all)}	$t\alpha_{L-}$	0,54	5,72	-2,56	0,67	6,33
{CSD(all)}- {3-21G}	$t\alpha_{L-}$	9,73	-7,07	4,27	-1,46	12,84
{CSD(A>30)}- {3-21G}	$t\alpha_{L-}$	10,27	-1,35	1,71	-0,79	10,52
{CSD(all)}- {6-31+G*}	$t\alpha_{L-}$	14,74	-6,38	-1,85	-14,18	21,50
{CSD(A>30)}- {6-31+G*}	$t\alpha_{L-}$	15,28	-0,66	-4,42	-13,51	20,88
{6-31+G*}- {3-21G}	$t\alpha_{L-}$	-5,01	-0,69	6,13	12,72	15,00
{CSD(all)}	$c\epsilon_{L+}$	8,08	6,83	13,58	10,54	20,18
{CSD(A>30)}	$c\epsilon_{L+}$	7,36	3,89	9,55	9,89	16,08
{CSD(A>30)}- {CSD(all)}	$c\epsilon_{L+}$	-0,41	2,39	1,46	0,16	2,84
{CSD(all)}- {3-21G}	$c\epsilon_{L+}$	-20,59	-4,04	-6,19	-19,45	29,27
{CSD(A>30)}- {3-21G}	$c\epsilon_{L+}$	-20,99	-1,65	-4,73	-19,28	28,94
{CSD(all)}- {6-31+G*}	$c\epsilon_{L+}$	-11,12	-0,46	0,72	-6,78	13,05
{CSD(A>30)}- {6-31+G*}	$c\epsilon_{L+}$	-11,52	1,94	2,18	-6,61	13,60
{6-31+G*}- {3-21G}	$c\epsilon_{L+}$	-9,47	-3,58	-6,91	-12,67	17,63

(folytatás következik)

(az E.2.6. Táblázat folytatása)

{CSD(all)}	$c\alpha_{L+}$	6,32	5,19	10,09	12,06	17,73
{CSD(A>30)}	$c\alpha_{L+}$	5,09	3,46	6,78	12,23	15,28
{CSD(A>30)}- {CSD(all)}	$c\alpha_{L+}$	-0,73	0,89	1,41	0,39	1,86
{CSD(all)}- {3-21G}	$c\alpha_{L+}$	4,76	-2,34	-2,90	-3,04	6,77
{CSD(A>30)}- {3-21G}	$c\alpha_{L+}$	4,03	-1,45	-1,49	-2,65	5,25
{CSD(all)}- {6-31+G*}	$c\alpha_{L+}$	0,25	0,18	-4,86	2,38	5,42
{CSD(A>30)}- {6-31+G*}	$c\alpha_{L+}$	-0,48	1,07	-3,45	2,77	4,58
{6-31+G*}- {3-21G}	$c\alpha_{L+}$	4,51	-2,53	1,96	-5,42	7,74
{CSD(all)}	$c\epsilon_{L-}$	6,50	8,14	28,61	15,21	34,03
{CSD(A>30)}	$c\epsilon_{L-}$	4,36	1,56	14,71	17,75	23,52
{CSD(A>30)}- {CSD(all)}	$c\epsilon_{L-}$	1,41	6,18	-18,23	1,13	19,33
{CSD(all)}- {3-21G}	$c\epsilon_{L-}$	-19,09	-10,15	28,31	-23,72	42,80
{CSD(A>30)}- {3-21G}	$c\epsilon_{L-}$	-17,68	-3,96	10,09	-22,60	30,67
{CSD(all)}- {6-31+G*}	$c\epsilon_{L-}$	-17,25	-8,93	25,92	-8,75	33,55
{CSD(A>30)}- {6-31+G*}	$c\epsilon_{L-}$	-15,84	-2,74	7,69	-7,63	19,38
{6-31+G*}- {3-21G}	$c\epsilon_{L-}$	-1,84	-1,22	2,40	-14,97	15,32
{CSD(A>30)}- {CSD(all)}	$c\alpha_{L-}$	0,00	0,00	0,00	0,00	0,00
{CSD(all)}- {3-21G}	$c\alpha_{L-}$	-2,87	-2,32	25,18	0,34	25,45
{CSD(A>30)}- {3-21G}	$c\alpha_{L-}$	-2,87	-2,32	25,18	0,34	25,45
{CSD(all)}- {6-31+G*}	$c\alpha_{L-}$	3,83	-0,99	16,89	-6,53	18,53
{CSD(A>30)}- {6-31+G*}	$c\alpha_{L-}$	3,83	-0,99	16,89	-6,53	18,53
{6-31+G*}- {3-21G}	$c\alpha_{L-}$	-6,70	-1,34	8,29	6,87	12,75

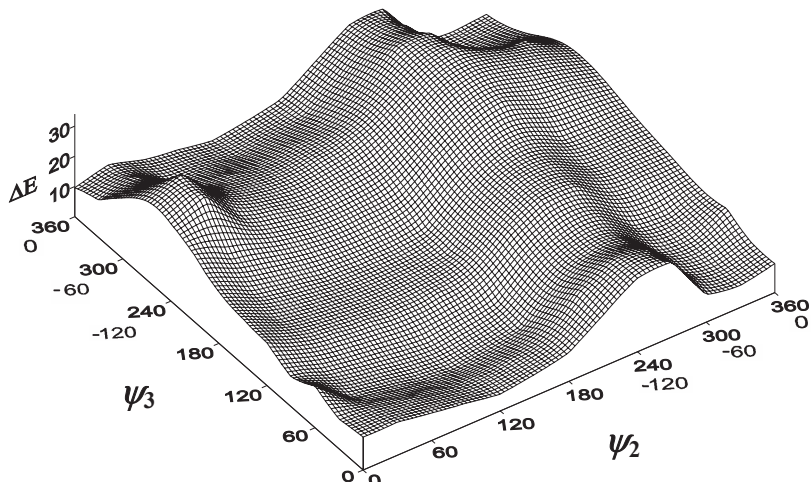
^a{CSD(all)} ill. {CSD(A>30)}: a választott paraméter szórása a kísérletileg meghatározott szerkezetekben, mindig pozitív; egyébként: a választott paraméter két módszerrel kapott értékének különbsége. Ld. E.2.1-3.4. Táblázat.

Profil-prolin

Ab initio, DFT és PCM számítások

A HCO—L-Pro—L-Pro—NH₂ modellpeptidnek a peptidkötések *transz* állása mellett RHF/3-21G elméleti szinten kiszámított ψ_2 - ψ_3 potenciálisenergia-felszínét (E.3.1. Ábra) jellemző nagy energiájú területek (i) egy hegygerinc $\psi_2 \approx 270^\circ$ mentén (ii) és egy erre merőleges hegygerinc $\psi_3 \approx 270^\circ$ mentén. A globális minimum $\psi_2 \approx 120^\circ$ és $\psi_2 \approx 75^\circ$ környezetébe várható, míg egy lokális minimum előre jelezhető $\psi_2 \approx -45^\circ$ és $\psi_2 \approx 15^\circ$ közelébe is. A ψ_2 - ψ_3 potenciálisenergia-felszín 30^o-os növekmények esetében nem vezet a *transz* peptidkötésű HCO—L-Pro—L-Pro—NH₂ modell többi minimumának nyomára, noha ψ_2 és ψ_3 egyaránt három minimumot ad (ld. M.5/C Ábra), és ezért 3x3=9 ideális gerinckonformerre kell tekintettel lenni. Így a 6 független szabadságfokkal ($\omega_1, \psi_2, P_2, \omega_2, \psi_3$

és P_3) rendelkező HCO—L-Pro—L-Pro—NH₂ peptidmodell hiperterének rendszerezett felosztása alterekre két paraméter (pl. ψ_2 - ψ_3 , P_2 - P_3 , stb.) kitüntetésével, majd ezek mentén a potenciálisenergia-felület számítása csak igen drágán szolgáltatná az energiaminimumok teljes vagy jellemző gyűjteményét, vagyis sok-sok felszín kis növekményű előállítását útján.



E.3.1. Ábra A HCO—L-Pro—L-Pro—NH₂ modellpeptid RHF/3-21G elméleti szintű ψ_2 - ψ_3 potenciálisenergia-felületének valamennyi amidkötés *transz* állása mellett. ΔE relatív energia a globális minimum felett kcal.mol⁻¹ egységben. ψ_2 és ψ_3 egysége fok. A ψ torziószögek esetében a topológiai 0-360° intervallum előnyösebb az elterjedt -180+180° szakasszal szemben, mert így az ideális minimumok a térkép széleitől távolabb kerülnek (vö. M.5/C. Ábra).

A HCO—L-Pro—L-Pro—NH₂ peptidmodell ideális minimumaira RHF/3-21G elméleti szinten végeztem kereső optimalásokat. Összesen 51 konformert találtam (ld. **E.3.1. Táblázat**). Ezek közül 23 felépíti a modell csak *transz* peptidkötést tartalmazó teljes konformációs alkönyvtárát, megadva így hét gerinckonformert ($\tau\epsilon_L\tau\epsilon_L$, $\tau\epsilon_L\tau\gamma_L$, $\tau\epsilon_L\tau\alpha_L$, $\tau\gamma_L\tau\epsilon_L$, $\tau\gamma_L\tau\gamma_L$, $\tau\alpha_L\tau\gamma_L$, $\tau\alpha_L\tau\alpha_L$) a fent említett ψ_2 - ψ_3 potenciálisenergia-felszínen várt kilenc ideális minimumból. Az 51 minimumból 29 felépíti a csak pozitív gyűrűállású modellpeptid teljes konformációs alkönyvtárát. A HCO—L-Pro—L-Pro—NH₂ egy vagy két *cisz* kötést tartalmazó konformációs alkönyvtárai hiányosak abban az értelemben, hogy némely gerinckonformernél megelégedtem csupán a pozitív gyűrűállású szerkezet kiszámításával.

Az 51 konformer további optimalása RHF/6-31+G(d) elméleti szinten 35 minimumot eredményezett (ld. **E.3.2. Táblázat**), míg 16 más szerkezetté konvergált. Minden RHF/6-31+G(d) minimumra energiaszámítást végeztem B3LYP/6-311++G(d,p) elméleti szinten (ld. **E.3.2. Táblázat** harmadik oszlopa). Néhány minimumot a B3LYP/6-311++G(d,p) elméleti szinten optimaltam is (ld. **E.3.3. Táblázat**). Jelentős korrelációt mutat ($R^2=0,9994$) az azonos konformerhez tartozó $\Delta E_{\text{B3LYP/6-311++G(d,p)}/\text{RHF/6-31+G(d)}}$ energiák (mint x) és a $\Delta E_{\text{B3LYP/6-311++G(d,p)}}$ energiák (mint y) között 13 adatpontra illeszkedő lineáris regressziós egyenlet ($y = 1,0299x - 0,0292$). A gázfázisú számítások különböző elméleti szintjei rendre a $\tau\epsilon_L+\tau\gamma_L+$ vagy a $\tau\epsilon_L-\tau\gamma_L+$ szerkezetet adták globális minimumként, melyek mindegyike egy nyújtott szerkezet inverz γ -kanyarral és egymástól csak az első pirrolidingyűrű állásában különböznek.

A *transz* peptidkötések minden elméleti szinten könnyen megkülönböztethetők a *cisz*-től csakúgy, mint a pozitív gyűrűállás a negatívtól. A gerinckonformációk azonban ψ -nek nem az ideális -60° , 60° és 180° értékei körül csoportosulnak, ezért $\psi = 40^\circ$ ill. $\psi = 100^\circ$ küszöbértékeket választottam a γ_L gerinckonformáció megkülönböztetésére α_L -tól ill. ϵ_L -től. Az RHF/3-21G elméleti szinten a $\tau\alpha_L+\tau\alpha_L+$, a $\tau\alpha_L+\tau\gamma_L-$, $\tau\gamma_L+\tau\alpha_L+$ és a $\tau\alpha_L+\tau\alpha_L+$ ideális konformer megkettőződött (ld. **E.3.1. Táblázat**). Az RHF/6-31+G(d) elméleti szinten azonban az alkalmazott nevezékrendszer már egyértelmű, vagyis a kettőzött minimumok itt azonos szerkezetekké konvergáltak.

E.3.1. Táblázat A HCO-L-Pro-L-Pro-NH₂ modellpeptid RHF/3-21G elméleti szintű *ab initio* számításokkal azonosított konformerei.

Konformer	ΔE	ϕ_1	ϕ_2	ψ_2	A_2	P_2	ω_2	ϕ_3	ψ_3	A_3	P_3	ω_3	τ	d
$\alpha_1+\alpha_4+$ (I)	9,13	-173	-70,7	-6,6	36,7	-101,3	170	-102,4	12,2	39,5	-133,0	178	51,8	5,2
$\alpha_1+\alpha_4+$ (II)	9,41	-174	-49,4	-40,7	33,6	-59,7	-178	-94,9	13,6	40,5	-120,9	177	41,0	5,1
$\alpha_1+\alpha_4-$	10,20	-173	-71,9	-1,8	37,4	-102,0	164	-68,2	-21,9	40,3	82,9	-178	72,4	5,5
$\alpha_1-\alpha_4+$	7,23	-173	-53,4	-35,5	37,9	80,1	-179	-97,7	15,3	39,9	-129,6	177	41,9	5,1
$\alpha_1-\alpha_4-$	8,87	-173	-54,3	-30,7	38,8	74,1	177	-65,5	-21,9	39,9	82,9	-179	61,4	5,4
$\alpha_1+\tau_{11}+$	10,28	-174	-54,3	-48,1	32,1	-60,6	-175	-85,1	70,0	40,4	-115,7	-175	60,4	5,6
$\alpha_1+\tau_{11}-$ (I)	12,08	-175	-51,9	-46,5	30,5	-50,2	-177	-83,4	72,7	39,4	119,6	-174	63,3	5,6
$\alpha_1+\tau_{11}-$ (II)	10,85	-172	-89,7	1,8	40,2	-113,2	-180	-82,9	72,8	39,0	117,6	-178	104,5	7,2
$\alpha_1-\tau_{11}+$	7,96	-174	-61,1	-42,3	38,1	78,5	-170	-86,7	69,5	39,7	-121,8	-176	68,0	6,1
$\alpha_1-\tau_{11}-$	9,83	-174	-56,7	-42,4	37,9	73,4	-177	-83,1	72,9	39,1	118,8	-175	66,6	5,8
$\tau_{11}+\tau_{11}+$	3,50	-167	-120,4	66,1	40,7	-143,7	-179	-82,0	70,3	41,8	-106,9	-179	165,8	8,3
$\tau_{11}+\alpha_4+$	8,47	-167	-119,3	61,4	40,9	-143,7	173	-64,7	153,0	41,6	-102,2	178	170,6	8,9
$\alpha_1+\alpha_4+$	5,83	-180	-77,5	124,2	38,3	-111,3	174	-105,4	9,4	40,6	-120,4	178	-175,3	8,4
$\alpha_1+\alpha_4-$	7,27	-179	-82,1	120,9	38,0	-114,5	179	-68,2	-30,4	39,1	73,2	-178	-163,0	8,2
$\alpha_1-\alpha_4+$	6,90	-177	-87,1	116,1	41,0	119,4	-171	-102,8	9,7	40,5	-121,9	178	-177,7	8,4
$\alpha_1-\alpha_4-$	8,32	-176	-92,5	112,1	41,1	123,5	-177	-66,8	-31,4	39,0	75,9	178	-166,7	8,2
$\alpha_1+\tau_{11}+$	0,00	-178	-81,5	114,0	37,5	-112,6	-172	-88,0	68,9	39,5	-115,6	-179	-147,2	8,1
$\alpha_1+\tau_{11}-$	2,02	-179	-81,7	112,1	37,3	-112,2	-178	-83,8	72,1	38,3	108,6	-179	-148,5	8,1
$\alpha_1-\tau_{11}+$	0,90	-177	-89,0	109,6	40,6	120,6	-170	-87,4	68,7	39,5	-116,5	-179	-148,7	8,1
$\alpha_1-\tau_{11}-$	2,95	-176	-92,5	107,2	40,8	123,3	-176	-82,7	72,3	37,8	108,7	-179	-150,8	8,1
$\alpha_1+\alpha_4+$	6,21	-178	-82,4	110,6	37,9	-113,5	-180	-83,8	159,3	39,2	-123,3	178	-152,9	8,9
$\alpha_1+\alpha_4-$	6,91	179	-75,1	127,9	38,2	-107,8	163	-61,4	163,2	39,1	81,7	179	-132,8	8,7
$\alpha_1-\alpha_4+$	7,23	-177	-89,8	106,0	40,9	122,1	-178	-83,2	159,1	39,2	-125,1	178	-154,4	8,8

(folytatás következik)

(az E.3.1. Táblázat folytatása)

Konformer	ΔE	ω_1	ϕ_1	ψ_2	A_2	P_2	ω_2	ϕ_2	ψ_3	A_3	P_3	ω_3	τ	d
$\alpha_1+\alpha_4+$	13,30	-174	-44,7	-53,1	34,6	-55,3	4	-95,8	-15,9	39,0	-116,4	-177	-155,5	7,1
$\alpha_1-\alpha_4+$	11,86	-172	-62,6	-41,0	38,8	83,5	12	-90,2	-23,8	38,1	-109,5	-176	-134,8	6,6
$\alpha_1+\alpha_3+$	16,29	-173	-53,1	-46,1	35,2	-56,7	-3	-86,6	174,3	40,7	-129,2	177	-112,1	6,5
$\eta_1+\alpha_4+$ (I)	7,85	-169	-114,3	69,2	39,4	-134,9	11	-112,8	26,9	41,2	-139,0	179	-40,6	3,9
$\eta_1+\alpha_4+$ (II)	7,30	-167	-118,5	46,9	41,0	-138,6	2	-89,9	-17,8	38,6	-114,6	-176	-65,2	4,3
$\eta_1+\alpha_3+$	4,54	-165	-119,9	55,7	42,9	-152,3	-5	-85,5	164,9	40,1	-127,3	175	-17,8	3,4
$\alpha_1+\alpha_4+$	0,41	174	-60,3	145,1	40,2	-104,8	9	-91,6	21,6	40,5	-118,1	178	36,1	5,2
$\alpha_1-\alpha_4+$	1,90	176	-57,0	141,4	39,7	94,5	12	-92,6	19,1	40,6	-117,2	178	35,3	5,2
$\alpha_1+\alpha_3+$	12,13	178	-92,5	159,4	39,7	-138,4	-11	-73,7	136,8	39,2	-118,1	-178	45,5	5,9
$\alpha_4+\eta_1+$	14,22	7	-88,7	2,0	39,1	-114,4	-176	-99,0	6,6	41,1	-119,0	178	38,9	4,9
$\alpha_4-\eta_1+$	13,63	4	-60,0	-33,5	37,7	77,6	-170	-100,9	6,0	40,8	-121,8	178	12,9	4,2
$\alpha_4+\eta_1+$	9,80	12	-50,5	-49,1	32,2	-56,3	-177	-85,1	68,7	40,4	-113,5	-179	32,6	3,8
$\alpha_4-\eta_1+$	7,43	6	-60,0	-39,8	37,9	76,8	-170	-87,3	69,6	39,7	-118,9	-179	41,2	4,5
$\alpha_4+\eta_2+$	12,91	16	-41,9	-51,9	36,7	-45,2	172	-65,5	161,4	40,9	-102,4	178	44,0	4,7
$\alpha_4-\eta_2+$	11,15	11	-44,1	-46,7	39,4	58,9	170	-65,3	161,9	41,1	-101,5	178	45,6	4,8
$\alpha_1+\eta_1+$	5,75	14	-122,1	69,7	39,7	-146,7	-174	-84,8	70,2	40,7	-113,0	-179	135,6	7,2
$\alpha_1+\eta_2+$	10,65	13	-119,7	60,5	40,4	-144,1	176	-70,0	155,3	40,1	-112,7	178	132,8	7,8
$\alpha_1+\eta_1+$	9,41	-7	-66,4	165,0	39,9	-111,7	-177	-85,0	70,8	40,1	-112,2	-179	-122,3	6,8
$\alpha_1+\eta_2+$	7,51	-6	-65,6	163,9	39,0	-106,7	164	-70,6	161,8	39,4	-116,7	178	-127,8	7,6
$\alpha_4+\alpha_1+$ (I)	14,41	11	-94,6	-5,6	37,4	-120,2	-1	-90,9	-16,6	38,9	-120,5	-176	-151,9	6,6
$\alpha_4+\alpha_1+$ (II)	13,13	11	-55,0	-59,3	28,9	-58,4	6	-97,5	-13,6	39,3	-118,7	-177	165,8	6,8
$\alpha_4+\alpha_1+$	12,30	8	-69,0	-44,8	38,5	85,4	12	-92,6	-20,6	38,6	-113,1	-176	-173,0	6,7
$\alpha_4+\alpha_2+$	14,25	11	-52,8	-44,4	32,7	-58,3	-10	-82,8	175,0	40,8	-132,5	177	-151,4	6,7
$\alpha_4-\alpha_2+$	11,91	4	-63,5	-29,6	37,9	80,1	0	-78,2	169,2	40,8	-118,1	176	-126,3	6,1
$\alpha_1+\alpha_4+$	11,30	10	-113,4	80,5	37,4	-141,3	15	-98,5	-3,9	40,8	-117,0	179	-49,8	3,4
$\alpha_1+\alpha_3+$	11,22	16	-127,1	49,2	42,1	-149,4	-9	-80,1	164,9	40,4	-125,1	-179	-62,4	4,3
$\alpha_1+\alpha_4+$	9,25	-5	-74,3	157,7	39,8	-117,0	4	-84,6	-3,1	40,0	-109,5	-178	15,2	3,7
$\alpha_1+\alpha_3+$	9,65	-4	-68,6	152,0	39,5	-111,3	-17	-61,0	176,6	40,2	-112,5	179	43,5	4,4

ΔE : relatív energia, kcal.mol⁻¹, a globális minimum, $\alpha_1+\eta_1+$, $E_h = -809,913664$ Hartree, felett. Minden geometriai paraméter fokban, kiv.: d , mely Å egységben. A, P : pszeudorotációs amplitúdó és fáziszög az öt χ_i torziószögből kiszámítva.

E.3.2. Táblázat A HCO–L-Pro–L-Pro–NH₂modellpeptid RHF/6–31+G(d) elméleti szintű *ab initio* számításokkal megtalált konformerei.

Konformer	ΔE^a	ΔE^b	ω_1	ϕ_2	ψ_2	A_2	P_2	ω_2	ϕ_3	ψ_3	A_3	P_3	ω_3	τ	d
$t\alpha_1+t\alpha_1+$	6,38	6,16	-170	-55,5	-36,6	34,3	-63,6	-178	-87,1	-1,2	38,0	-115,5	173	47,7	5,4
$t\alpha_1+t\alpha_1-$	6,13	6,29	-172	-66,1	-17,5	33,2	-88,8	173	-70,7	-20,7	38,7	87,5	176	65,7	5,6
$t\alpha_1-t\alpha_1+$	4,28	4,29	-170	-57,4	-34,1	38,0	77,0	-178	-89,1	0,7	37,6	-121,8	172	48,3	5,5
$t\alpha_1-t\alpha_1-$	4,00	4,53	-171	-59,0	-29,7	38,0	76,3	178	-70,3	-19,9	38,5	86,6	175	60,1	5,5
$t\alpha_1+r\eta_1+$	6,58	5,68	-169	-85,6	-10,0	36,9	-107,8	-173	-86,7	74,6	37,7	-122,7	-173	96,3	7,2
$t\alpha_1+r\eta_1-$	8,40	7,00	-171	-76,6	-17,3	34,8	-98,6	180	-81,6	83,5	38,3	109,5	-170	78,9	6,8
$t\alpha_1-r\eta_1+$	5,27	4,84	-170	-67,2	-35,1	37,7	80,4	-170	-88,2	74,7	36,7	-122,8	-171	73,3	6,5
$t\alpha_1-r\eta_1-$	6,56	5,59	-170	-65,1	-34,2	37,7	77,2	-178	-81,9	82,3	38,0	107,8	-169	74,5	6,3
$t\epsilon_1+t\alpha_1+$	2,71	3,64	178	-68,1	133,5	34,9	-109,3	-171	-91,9	-0,7	38,9	-109,0	169	-149,6	8,2
$t\epsilon_1+t\alpha_1-$	3,49	4,36	178	-69,4	130,5	34,7	-108,3	-178	-73,3	-24,0	38,1	77,1	173	-147,8	8,1
$t\epsilon_1-t\alpha_1+$	2,76	3,29	-180	-64,0	131,4	38,1	95,7	-171	-92,0	1,6	39,1	-109,3	169	-150,0	8,2
$t\epsilon_1-t\alpha_1-$	3,60	4,23	-180	-64,6	130,0	38,2	95,6	-178	-73,3	-23,1	38,2	76,8	173	-147,0	8,1
$t\epsilon_1+r\eta_1+$	0,00	0,13	178	-67,9	137,7	34,7	-109,1	-174	-87,9	76,5	37,5	-109,8	-175	-123,4	7,9
$t\epsilon_1+r\eta_1-$	1,69	1,25	179	-68,7	132,3	34,5	-106,6	177	-78,0	94,4	38,0	91,3	-171	-122,5	7,9
$t\epsilon_1-r\eta_1+$	0,12	0,00	-180	-64,2	134,2	37,9	95,9	-174	-88,1	76,7	37,5	-109,4	-174	-126,0	7,9
$t\epsilon_1-r\eta_1-$	1,83	1,09	-180	-64,5	131,1	38,1	94,7	178	-78,9	93,0	38,0	91,7	-171	-123,3	7,9
$t\epsilon_1+t\epsilon_1+$	1,26	2,95	177	-70,2	148,8	35,5	-114,7	179	-72,6	135,3	35,3	-106,6	-173	-99,1	8,0
$t\epsilon_1+t\epsilon_1-$	1,50	2,55	179	-69,6	133,4	34,7	-108,3	174	-62,1	132,4	37,8	78,2	-172	-108,8	8,0
$t\alpha_1+c\alpha_1+$	7,41	7,36	-164	-112,9	30,2	39,8	-126,4	8	-91,4	-18,9	36,8	-111,5	-178	-73,2	4,7
$t\alpha_1-c\alpha_1+$	7,63	8,19	-165	-81,6	-14,1	38,4	86,4	21	-86,8	-17,8	37,1	-101,8	-171	-98,5	5,5
$r\eta_1+c\epsilon_1+$	5,93	5,07	-164	-125,3	65,3	40,0	-160,8	-2	-91,5	159,9	37,5	-132,9	-177	-11,4	3,7
$t\epsilon_1+c\alpha_1+$	0,01	0,24	176	-62,5	143,3	35,7	-108,3	11	-90,2	1,8	39,1	-115,8	172	31,5	5,3
$t\epsilon_1-c\alpha_1+$	0,11	0,08	178	-59,6	140,5	38,1	95,9	12	-90,5	0,3	39,0	-115,5	173	29,9	5,3

(folytatás következik)

(az E.3.2. Táblázat folytatása)

Konformer	ΔE^a	ΔE^b	ω_1	ϕ_2	ψ_2	A_2	P_2	ω_2	ϕ_2	ψ_2	A_3	P_3	ω_3	τ	d
$c\alpha_1-t\alpha_1+$	7,66	7,60	6	-65,4	-36,7	36,6	81,0	-167	-90,2	-1,9	38,6	-113,3	170	21,1	4,5
$c\alpha_1+r_{L+}$	6,24	5,20	9	-85,8	-11,7	35,7	-112,4	-174	-87,4	76,5	37,4	-118,2	-175	60,0	5,5
$c\alpha_1-r_{L+}$	5,20	4,23	7	-64,2	-36,7	36,7	80,1	-171	-88,3	74,8	37,1	-116,9	-176	41,8	4,7
$c\alpha_1-t\epsilon_1+$	6,47	7,01	8	-61,1	-43,9	37,3	77,1	-178	-70,4	137,8	35,8	-104,9	-173	51,6	5,4
$c\epsilon_1+r_{L+}$	2,53	2,37	-3	-72,0	162,9	36,2	-118,1	-176	-84,7	79,0	37,6	-110,7	-175	-125,0	7,0
$c\epsilon_1+t\epsilon_1+$	3,27	4,38	-2	-72,6	158,0	36,0	-117,4	176	-69,0	139,6	35,3	-110,3	-173	-119,0	7,3
$c\alpha_1+c\alpha_1+$	10,17	9,74	12	-66,3	-42,9	30,4	-79,8	-5	-91,7	-5,4	38,1	-128,4	-177	173,8	6,9
$c\alpha_1-c\alpha_1+$	8,47	8,01	8	-58,6	-48,8	38,0	74,0	-2	-93,0	-13,2	37,3	-125,7	179	171,2	6,8
$c\alpha_1+c\epsilon_1+$	8,89	8,21	9	-98,8	13,3	39,0	-117,6	3	-80,2	158,9	37,5	-116,0	-176	-84,4	4,9
$c\alpha_1-c\epsilon_1+$	8,83	8,59	7	-73,5	-15,8	38,4	79,6	11	-73,8	159,9	37,6	-107,0	-176	-98,4	5,1
$c\epsilon_1+c\alpha_1+$	5,07	5,55	-2	-76,0	156,3	36,0	-119,6	5	-86,0	-10,2	37,1	-111,3	-176	13,4	4,0
$c\epsilon_1+c\epsilon_1+$	7,07	7,72	-2	-72,5	158,8	36,2	-118,5	-7	-71,3	165,2	37,4	-119,4	173	52,1	5,0

^a optimalmás RHF/6-31+G(d) elméleti szinten.

^b energiaszámítás B3LYP/6-311++G(d,p)/RHF/6-31+G(d) elméleti szinten.

ΔE relatív energia, kcal.mol⁻¹, a globális minimum, $t\alpha_1+r_{L+}$, $E_h = -814,458611^a$; és $t\alpha_1-r_{L+}$, $E_h = -819,615470^b$ Hartree, felett. Minden geometriai paraméter fokban, kiv.: d Å egységben. A , P : pszeudorotációs amplitúdó és fáziszög az ω χ_i torziószögből kiszámítva.

E.3.3. Táblázat A HCO-L-Pro-L-Pro-NH₂ modellpeptid B3LYP/6-311++G(d,p) elméleti szintű DFT számításokkal megtalált konformerei.

Konformer	ΔE	ϕ_1	ϕ_2	ψ_2	A_2	P_2	ϕ_2	ϕ_3	ψ_2	A_3	P_3	ϕ_2	τ	d
$t\alpha_L+t\alpha_L-$	6,42	-174	-62,8	-17,8	34,0	-90,1	172	-72,8	-16,4	38,47	91,04	178	64,2	5,5
$t\epsilon_L+r\eta_L+$	0,13	-179	-72,8	123,8	35,4	-107,4	-173	-86,1	70,4	37,20	-111,36	-177	-135,4	8,1
$t\epsilon_L+r\eta_L-$	1,06	-179	-74,4	123,0	35,1	-107,7	-180	-79,9	79,9	37,13	98,24	-174	-134,1	8,1
$t\epsilon_L-r\eta_L+$	0,00	-177	-69,8	124,0	38,0	96,1	-172	-85,6	70,6	37,42	-109,71	-177	-134,6	8,0
$t\epsilon_L-r\eta_L-$	0,96	-177	-72,8	123,0	38,2	98,1	-179	-79,7	80,6	37,12	97,29	-174	-133,8	8,1
$t\alpha_L-c\alpha_L+$	8,34	-164	-85,2	-7,0	38,4	93,6	17	-88,2	-13,3	37,31	-105,83	-175	-95,9	5,4
$r\eta_L+c\epsilon_L+$	5,20	-166	-121,7	74,4	38,8	-162,3	-1	-99,4	149,8	37,36	-139,39	-178	-7,5	4,0
$t\epsilon_L+c\alpha_L+$	0,30	176	-61,5	140,4	36,7	-104,4	10	-92,4	12,2	39,17	-119,80	174	30,3	5,3
$t\epsilon_L-c\alpha_L+$	0,13	179	-59,8	137,8	38,1	93,7	11	-93,6	11,5	39,12	-120,29	174	27,6	5,2
$c\alpha_L-r\eta_L+$	4,41	7	-62,9	-40,2	36,9	80,6	-171	-85,1	70,9	37,63	-114,09	-177	41,1	4,7
$c\epsilon_L+r\eta_L+$	2,48	-1	-76,2	158,3	36,4	-118,4	-175	-82,9	71,5	37,85	-111,38	-177	-129,3	7,1
$c\alpha_L+c\alpha_L+$	9,99	10	-62,1	-44,2	30,8	-80,0	-3	-92,8	-5,0	37,87	-127,63	-179	174,6	6,8
$c\alpha_L-c\alpha_L+$	8,31	9	-58,1	-49,1	38,0	74,5	1	-94,0	-10,7	37,43	-124,54	179	173,1	6,8

ΔE : relatív energia, kcal.mol⁻¹, a globális minimum, $t\epsilon_L-r\eta_L+$, $E_h = -819,620188$ Hartree, felett. Minden geometriai paraméter fokban, kiv.: d Å egységben. A, P : pszeudorotációs amplitúdó és fáziszög az öt χ_i torziószögből kiszámítva.

Az RHF/6-31+G(d) szinten kapott 35 minimum mindegyikét optimaltám az oldószert polarizálható közegeként figyelembe vevő kvantumkémiai módszerrel (PCM) B3LYP/6-31+G(d) elméleti szinten. Az alacsony energiájú minimumok esetében mind a négy gyűrűállást igyekeztem megkeresni (ld. **E.3.4. Táblázat**). Az oldószermódel alkalmazásának legjelentősebb hatása a gázfázisú számításokhoz képest az, hogy a globális minimum ($\mathbf{t\epsilon_1-t\epsilon_1+}$) nem tartalmaz inverz γ -kanyart, hanem egy valódi poliprolin II szerkezet. A $\mathbf{t\epsilon_1\gamma_1}$ gerinckonformernek csak két gyűrűállását találtam meg ($++$ és $-+$), a másik kettő ($+-$ és $--$) $\mathbf{t\epsilon_1\epsilon_1}$ gerinckonformerré alakult. E migrációnak köszönhetően fedeztem fel a $\mathbf{t\epsilon_1+t\epsilon_1-}$ konformer megkettőződését. Az energiaminimumok többszöröződését PCM módszer alkalmazásakor már korábban is tapasztaltam HCO-L-Ala-NH₂ modellpeptidnél,⁵ de a prolinnál részletesen nem vizsgáltam. Habár a két $\mathbf{t\epsilon_1+t\epsilon_1-}$ konformer rendkívül hasonló geometriai adatokkal és energiával rendelkezik szorosabb (tight) optimalási kritérium esetében sem szűnt meg kettőződésük. $\mathbf{t\epsilon_1-t\epsilon_1-}$ létezik B3LYP/PCM/6-31+G(d) szinten, de gázfázisban nem található. Ezzel szemben $\mathbf{t\epsilon_1+t\alpha_1+}$ valódi minimum RHF/6-31+G(d) szinten, de PCM módszer mellett eltűnik.

A klasszikus hidrogénhidak a HCO—L-Pro—L-Pro—NH₂ peptidben kétfélék lehetnek (ld. **M.8. Ábra**), mivel a protonakceptorként szereplő karbonil oxigént tartalmazhatja egyrészt a formilcsoport (A típus), másrészt a két prolin közötti peptidkötés (B típus), míg a protondonor mindig az NH₂-csoportnak a gerinchez viszonyítva *cisz* helyzetű hidrogénatomja. Néhány hidrogénkötést tartalmazó konformerben egy-egy peptidkötés *cisz* állású lehet (ld. **E.3.5. Táblázat**), azonban két *cisz* peptidkötés már megakadályozza a hidrogénhid-kötés kialakítását.

Az A típusú hidrogénkötést tartalmazó konformerekben, a formilcsoport és az aminocsoport, melyek a peptidmodell szekvenciálisan ellentétes végein helyezkednek el, térben közel kerülnek egymáshoz. Az ilyen hidrogénkötés az első peptidcsoportra (ω_1) *transz* állást kényszeríti, míg a második prolinra α_1 gerinckonformációt. Ezek a konformerek (i) $\mathbf{t\alpha_1\alpha_1}$ vagy (ii) $\mathbf{t\epsilon_1\alpha_1}$ gerincűek lehetnek tetszőleges gyűrűállás mellett. Ebből következően a két prolin közötti peptidkötés *cisz-transz* izomerizációja megvalósulhat az A típusú hidrogénkötés megtartása közben.⁶ A $\mathbf{t\alpha_1\alpha_1}$ gerinckonformáció I vagy III típusú β -kanyar, míg a $\mathbf{t\epsilon_1\alpha_1}$ VIa típusú β -kanyar (ld. **M.1. Táblázat**).

E.3.4. Táblázat A HCO-L-Pro-L-Pro-NH₂ modellpeptid B3LYP/PCM/6-31+G(d) elméleti szintű számításokkal megtalált konformerei.

Konformer	ΔE	ϕ_1	ϕ_2	ψ_2	A_2	P_2	ω_2	ϕ_3	ψ_3	A_3	P_3	ω_3	τ	d
$t\alpha_L+t\alpha_L+$	3,62	-174	-48,4	-44,5	34,1	-63,5	-178	-79,3	-6,5	37,9	-105,8	177	46,9	5,3
$t\alpha_L+t\alpha_L-$	3,30	-174	-57,8	-27,7	32,1	-79,7	175	-68,0	-20,0	39,0	88,4	-180	61,4	5,4
$t\alpha_L-t\alpha_L+$	1,69	-172	-53,3	-40,2	38,5	77,5	-179	-81,7	-3,9	37,3	-112,0	177	47,2	5,3
$t\alpha_L-t\alpha_L-$	1,03	-173	-55,6	-33,9	38,5	78,2	177	-67,5	-19,5	38,7	87,5	180	58,1	5,4
$t\alpha_L+r\gamma_L+$	5,76	-172	-78,7	-20,2	35,1	-105,7	-170	-86,0	66,6	37,7	-119,9	-178	89,6	7,0
$t\alpha_L+r\gamma_L-$	6,53	-174	-72,4	-20,8	33,9	-98,6	-180	-79,2	88,6	38,3	103,7	-174	91,9	6,9
$t\alpha_L-r\gamma_L+$	3,96	-171	-67,6	-39,9	38,1	85,4	-168	-86,3	66,4	37,3	-117,2	-178	73,2	6,6
$t\alpha_L-r\gamma_L-$	4,63	-172	-63,9	-36,1	38,4	81,1	-177	-78,7	85,5	37,9	101,9	-173	80,3	6,6
$t\epsilon_L+t\alpha_L-$	1,62	-179	-74,7	165,6	37,0	-109,5	-177	-66,1	-34,3	37,9	85,0	-175	-111,2	7,9
$t\epsilon_L-t\alpha_L+$	0,56	-175	-67,5	150,2	38,4	90,9	-174	-79,6	-22,6	37,3	-102,6	-178	-134,9	8,3
$t\epsilon_L-t\alpha_L-$	1,15	-176	-65,7	145,0	38,5	88,2	-179	-68,1	-32,2	37,9	83,4	-175	-134,8	8,2
$t\epsilon_L+r\gamma_L+$	1,56	-179	-72,6	159,0	36,6	-106,7	-176	-84,0	65,3	37,9	-113,4	-179	-101,7	7,7
$t\epsilon_L-r\gamma_L+$	1,06	-175	-69,0	151,7	38,4	91,8	-176	-83,8	65,3	38,0	-113,1	-179	-110,4	7,8
$t\epsilon_L+t\epsilon_L+$	0,42	-179	-72,0	156,7	36,6	-106,1	-180	-73,6	150,6	36,2	-108,5	176	-94,7	8,5
$t\epsilon_L+t\epsilon_L-$ (I)	0,54	-179	-73,4	162,1	37,0	-107,8	178	-61,1	145,0	37,9	86,9	178	-78,3	8,1
$t\epsilon_L+t\epsilon_L-$ (II)	0,48	-179	-73,8	162,1	37,0	-108,4	178	-61,4	145,1	37,9	87,1	178	-78,4	8,1
$t\epsilon_L-t\epsilon_L+$	0,00	-175	-69,2	151,5	38,3	92,8	-180	-72,8	150,0	36,3	-107,5	176	-100,2	8,6
$t\epsilon_L-t\epsilon_L-$	0,12	-175	-71,4	157,3	38,4	95,1	177	-60,3	145,7	37,8	86,9	178	-84,6	8,2
$t\alpha_L-c\alpha_L+$	5,16	-170	-64,6	-41,5	38,9	82,5	4	-91,1	-12,0	37,8	-122,3	-179	-139,1	7,0
$t\gamma_L+c\alpha_L+$	5,98	-164	-119,6	44,6	39,8	-146,8	6	-96,1	-9,7	38,0	-121,1	-179	-63,2	4,5
$t\gamma_L+c\epsilon_L+$	6,30	-164	-121,2	47,9	40,1	-153,7	1	-91,0	165,9	37,9	-125,6	176	-22,5	3,7
$t\epsilon_L+c\alpha_L+$	1,05	178	-59,7	141,7	36,3	-93,8	9	-90,6	11,8	39,2	-118,8	174	28,9	5,2
$t\epsilon_L+c\alpha_L-$	2,84	-180	-71,3	159,4	36,5	-106,0	3	-75,9	-22,3	37,8	92,1	-179	28,1	5,9
$t\epsilon_L-c\alpha_L+$	0,45	-178	-58,8	139,6	38,6	85,9	9	-91,6	11,7	39,2	-119,7	174	28,5	5,2
$t\epsilon_L-c\alpha_L-$	1,97	-179	-56,7	140,7	38,7	83,9	7	-79,4	-4,7	38,5	96,2	177	29,2	5,1

(folytatás következik)

(az E.3.4. Táblázat folytatása)

Konformer	ΔE	ω_1	ϕ_2	ψ_2	A_2	P_2	ω_2	ϕ_3	ψ_3	A_3	P_3	ω_3	τ	d
$ca_{11}-ra_{11}+$	4,17	6	-66,2	-39,9	37,2	83,7	-168	-79,0	-25,4	36,9	-99,9	-177	17,3	4,6
$ca_{11}+r_{11}+$	5,49	8	-84,1	-16,3	34,8	-114,7	-174	-86,3	65,9	37,7	-116,5	-179	55,0	5,3
$ca_{11}-r_{11}+$	4,43	5	-62,9	-40,5	36,9	82,5	-171	-86,0	68,5	37,4	-115,0	-178	39,5	4,7
$ca_{11}-re_{11}+$	3,17	5	-60,0	-46,4	37,1	79,9	-176	-72,8	151,6	36,5	-104,7	176	45,6	5,9
$ce_{11}+r_{11}+$	1,41	0	-78,8	162,6	36,3	-117,9	-175	-84,3	64,4	38,0	-112,4	-179	-128,9	7,1
$ce_{11}+re_{11}+$	0,16	0	-78,4	161,3	36,3	-118,4	-180	-73,0	150,3	36,3	-106,8	177	-120,7	7,7
$ce_{11}+re_{11}-$	0,25	0	-76,9	168,5	36,5	-117,3	180	-63,0	145,4	38,0	88,4	178	-103,3	7,4
$ce_{11}-re_{11}+$	0,29	-1	-58,8	154,8	37,4	85,0	180	-72,3	150,8	36,3	-107,3	176	-115,8	7,3
$ce_{11}-re_{11}-$	0,39	-1	-60,1	162,9	37,2	86,5	178	-61,7	145,6	38,0	87,7	177	-99,9	7,0
$ca_{11}+ca_{11}+$	7,05	9	-72,7	-33,0	31,3	-98,8	-3	-92,9	-5,7	38,6	-130,0	-179	-175,7	6,9
$ca_{11}-ca_{11}+$	5,61	6	-54,6	-49,5	38,2	73,3	0	-93,5	-8,7	38,3	-126,2	-180	173,0	6,8
$ca_{11}+ce_{11}+$	7,19	15	-105,2	16,7	39,7	-123,4	3	-83,3	163,6	37,8	-116,8	177	-63,1	5,1
$ca_{11}-ce_{11}+$	5,50	8	-59,0	-42,7	37,4	75,9	-4	-84,8	160,6	37,7	-128,4	174	-141,5	6,7
$ce_{11}+ca_{11}+$	1,66	1	-79,7	160,9	36,2	-117,4	4	-86,5	-10,9	37,5	-114,6	-179	13,4	4,2
$ce_{11}+ce_{11}+$	1,87	0	-78,7	164,2	36,4	-119,2	-3	-78,8	160,2	36,8	-118,4	175	52,6	5,3

ΔE : relatív energia, kcal.mol⁻¹, a globális minimum, $re_{11}-re_{11}+$, $E_0 = -819,443497$ Hartree, felett. Minden geometriai paraméter fokban, kiv.: d Å egységben. A, P : pszeudorotációs amplitúdó és fázisszög az öt χ_{ij} torziószögből kiszámítva.

E.3.5a. Táblázat Hidrogénhid-kötések jellemzői a HCO–L-Pro–L-Pro–NH₂ modellpeptid RHF/3–21G elméleti szinten azonosított konformereiben.

Konformer	típus	$d(\text{ON})$	$d(\text{OH})$	$a(\text{OHN})$	ΔE
$t\alpha_L+t\alpha_L+$ (I)	A	3,09	2,12	163,5	9,13
$t\alpha_L+t\alpha_L+$ (II)	A	2,96	1,98	167,1	9,41
$t\alpha_L+t\alpha_L-$	A	3,08	2,09	168,3	10,20
$t\alpha_L-t\alpha_L+$	A	3,00	2,02	166,3	7,23
$t\alpha_L-t\alpha_L-$	A	2,99	2,00	169,1	8,87
$t\alpha_L+t\gamma_L+$	B	2,84	1,98	141,7	10,28
$t\alpha_L+t\gamma_L-$	B	2,84	1,98	141,6	12,08
$t\alpha_L-t\gamma_L+$	B	2,84	1,98	141,9	7,96
$t\alpha_L-t\gamma_L-$	B	2,83	1,98	141,2	9,83
$t\gamma_L+t\gamma_L+$	B	2,79	1,94	140,8	3,50
$t\varepsilon_L+t\gamma_L+$	B	2,84	1,98	142,1	0,00
$t\varepsilon_L+t\gamma_L-$	B	2,80	1,95	140,6	2,02
$t\varepsilon_L-t\gamma_L+$	B	2,84	1,98	142,4	0,90
$t\varepsilon_L-t\gamma_L-$	B	2,80	1,95	140,7	2,95
$t\varepsilon_L+c\alpha_L+$	A	2,89	1,94	156,6	0,41
$t\varepsilon_L-c\alpha_L+$	A	2,88	1,93	156,8	1,90
$c\alpha_L+t\gamma_L+$	B	2,82	1,98	139,9	9,80
$c\alpha_L-t\gamma_L+$	B	2,83	1,98	140,7	7,43
$c\gamma_L+t\gamma_L+$	B	2,81	1,97	140,7	5,75
$c\varepsilon_L+t\gamma_L+$	B	2,84	1,99	140,1	9,41

A hidrogénhid típusok besorolása az **M.8. Ábra** szerint. Atomtávolságok, $d(\text{ON})$ és $d(\text{OH})$, Å-ben, $a(\text{OHN})$ kötésszög fokban, ΔE relatív energia kcal.mol⁻¹ egységben.

E.3.5b. Táblázat Hidrogénhid-kötések jellemzői a HCO–L-Pro–L-Pro–NH₂ modellpeptid RHF/6–31+G(d) elméleti szinten megtalált konformereiben.

Konformer	típus	$d(\text{ON})$	$d(\text{OH})$	$a(\text{OHN})$	ΔE
$t\alpha_L+t\alpha_L+$	A	3,20	2,21	168,5	6,38
$t\alpha_L+t\alpha_L-$	A	3,26	2,26	171,9	6,13
$t\alpha_L-t\alpha_L+$	A	3,24	2,26	167,9	4,28
$t\alpha_L-t\alpha_L-$	A	3,20	2,21	172,0	4,00
$t\alpha_L+t\gamma_L+$	B	2,91	2,09	138,9	6,58
$t\alpha_L+t\gamma_L-$	B	2,91	2,10	137,0	8,40
$t\alpha_L-t\gamma_L+$	B	2,96	2,12	139,6	5,27
$t\alpha_L-t\gamma_L-$	B	2,93	2,11	138,3	6,56
$t\varepsilon_L+t\gamma_L+$	B	2,94	2,12	138,1	0,00
$t\varepsilon_L+t\gamma_L-$	B	2,94	2,19	130,6	1,69
$t\varepsilon_L-t\gamma_L+$	B	2,94	2,12	138,2	0,12
$t\varepsilon_L-t\gamma_L-$	B	2,93	2,18	131,6	1,83
$t\varepsilon_L+c\alpha_L+$	A	3,08	2,19	147,9	0,01
$t\varepsilon_L-c\alpha_L+$	A	3,06	2,16	148,8	0,11
$c\alpha_L+t\gamma_L+$	B	2,92	2,11	136,8	6,24
$c\alpha_L-t\gamma_L+$	B	2,94	2,13	136,9	5,20
$c\varepsilon_L+t\gamma_L+$	B	2,94	2,14	136,2	2,53

(ld. E.3.5a. Táblázat)

E.3.5c. Táblázat Hidrogénhid-kötések jellemzői a HCO–L-Pro–L-Pro–NH₂ modellpeptid B3LYP/6–311++G(d,p) elméleti szintű konformereiben.

Konformer	típus	<i>d</i> (ON)	<i>d</i> (OH)	<i>a</i> (OHN)	ΔE
<i>t</i> α_L + <i>t</i> α_L –	A	3,11	2,11	169,0	6,42
<i>t</i> ϵ_L + <i>t</i> γ_L +	B	2,88	2,00	142,8	0,13
<i>t</i> ϵ_L + <i>t</i> γ_L –	B	2,84	1,99	140,4	1,06
<i>t</i> ϵ_L – <i>t</i> γ_L +	B	2,87	2,00	143,0	0,00
<i>t</i> ϵ_L – <i>t</i> γ_L –	B	2,85	1,99	140,0	0,96
<i>t</i> γ_L + <i>c</i> ϵ_L +	A	2,99	2,17	137,1	5,20
<i>t</i> ϵ_L + <i>c</i> α_L +	A	2,94	1,99	154,4	0,30
<i>t</i> ϵ_L – <i>c</i> α_L +	A	2,93	1,98	154,9	0,13
<i>c</i> α_L – <i>t</i> γ_L +	B	2,87	2,01	141,4	4,41
<i>c</i> ϵ_L + <i>t</i> γ_L +	B	2,87	2,00	143,0	2,48

(ld. E.3.5a. Táblázat)

E.3.5d. Táblázat Hidrogénhid-kötések jellemzői a HCO–L-Pro–L-Pro–NH₂ modellpeptid B3LYP/PCM/6–31+G(d) elméleti szinten optimált konformereiben.

Konformer	típus	kanyartípus	<i>d</i> (ON)	<i>d</i> (OH)	<i>a</i> (OHN)	ΔE
<i>t</i> α_L + <i>t</i> α_L +	A	I β	3,02	2,03	164,6	3,62
<i>t</i> α_L + <i>t</i> α_L –	A	III β	3,04	2,04	167,3	3,30
<i>t</i> α_L – <i>t</i> α_L +	A	I β	3,06	2,07	164,6	1,69
<i>t</i> α_L – <i>t</i> α_L –	A	III β	3,03	2,03	166,6	1,03
<i>t</i> α_L + <i>t</i> γ_L +	B	γ	2,87	2,00	142,4	5,76
<i>t</i> α_L + <i>t</i> γ_L –	B	γ	2,92	2,16	129,8	6,53
<i>t</i> α_L – <i>t</i> γ_L +	B	γ	2,89	2,01	143,1	3,96
<i>t</i> α_L – <i>t</i> γ_L –	B	γ	2,90	2,10	133,8	4,63
<i>t</i> ϵ_L + <i>t</i> γ_L +	B	γ	2,89	2,00	143,4	1,56
<i>t</i> ϵ_L – <i>t</i> γ_L +	B	γ	2,88	1,99	143,7	1,06
<i>t</i> ϵ_L + <i>c</i> α_L +	A	VIa β	2,96	2,01	154,2	1,05
<i>t</i> ϵ_L – <i>c</i> α_L +	A	VIa β	2,95	2,00	154,6	0,45
<i>t</i> ϵ_L – <i>c</i> α_L –	A	VIa β	2,96	2,03	149,6	1,97
<i>c</i> α_L + <i>t</i> γ_L +	B	γ	2,88	2,01	141,7	5,49
<i>c</i> α_L – <i>t</i> γ_L +	B	γ	2,90	2,03	141,2	4,43
<i>c</i> ϵ_L + <i>t</i> γ_L +	B	γ	2,89	2,01	143,5	1,41

(ld. E.3.5a. Táblázat)

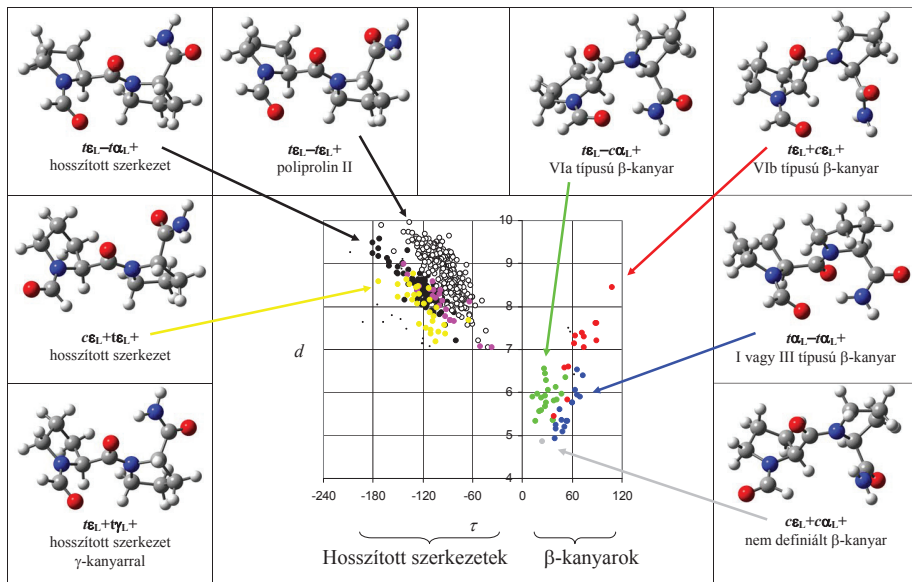
A B típusú hidrogénkötést tartalmazó konformerekben a terminális aminosoprot a két prolin közötti karbonilsoprotot közelíti meg. Ez azokban HCO–L-Pro–L-Pro–NH₂ konformerekben valósul meg, ahol a második prolin térszerkezete $t\gamma_L$, vagyis inverz γ -kanyar, bármilyen gyűrűállás esetén. Így a formilsoprot *cisz-transz* izomerizációja nincs hatással a B típusú hidrogénkötésre.

A megtalált konformerek hidrogénkötéseit jellemző geometriai paraméterek, továbbá a relatív energiák és populációs adatok (ld. E.2.2/B Ábra) elemzése azt tükrözi, hogy az egyes

elméleti szintek eltérő arányban hasznosítják a hidrogénkötés stabilizáló hatását Az RHF/3-21G elméleti szintű szerkezetek stabilitását legfőképpen a hidrogénkötések adják, konformációs összetétele kizárólag a $\tau_{\mathbf{E}_1, \mathbf{F}_1}$ és a $\tau_{\mathbf{E}_1, \mathbf{C}\alpha_1}$ változatai jelentősek (70 és 30%). Hidrogénkötést nem tartalmazó konformer csak az RHF/6-31+G(d) elméleti szinten jelentkezik számottevő mennyiségben ($\tau_{\mathbf{E}_1, \mathbf{E}_1}$ 5%).

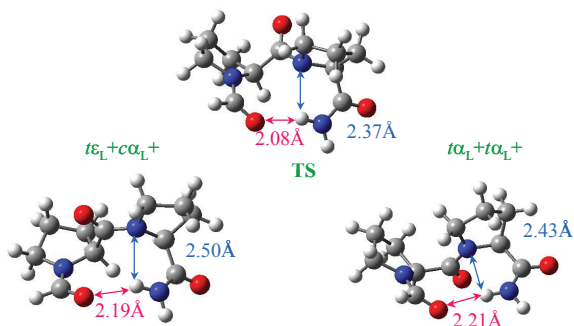
Peptidekben és fehérjékben előforduló **Pro–Pro** szekvenciaegység hatását a teljes szerkezetre a τ és a d paraméterekkel összegezhetjük: β -kanyarokban mindkettő kis értékeket vesz fel (kb. 30° ill. 5 \AA), míg hosszított szerkezetekben nagy értékeket (kb. -140° ill. 8 \AA , ld. **E.3.1-4. Táblázat**). A gázfázisú számítások mindegyikében a globális minimum egy hosszított szerkezet γ -kanyarral ($\tau_{\mathbf{E}_1, \mathbf{F}_1}$, τ -123° és -151° között, és d $8,0 \text{ \AA}$ körül, ld. **E.3.2. Ábra** a konformerek illusztrálására, megjegyezve, hogy a kísérleti τ és d értékek némileg eltérnek a számítottaktól), 1 kcal.mol^{-1} -nál kisebb relatív energiája pedig a VIa típusú β -kanyarnak van ($\tau_{\mathbf{E}_1, \mathbf{C}\alpha_1}$, τ 28° és 36° között, és d $5,2 \text{ \AA}$ körül). A PCM globális minimum egy másik hosszított szerkezet, nevezetesen a poliprolin II ($\tau_{\mathbf{E}_1, -\mathbf{E}_1+}$, $\tau = -100^\circ$ és $d = 8,6 \text{ \AA}$), amelyben nincs lokális hidrogénkötés. $0,6 \text{ kcal.mol}^{-1}$ -nál alacsonyabb energiájú a többi gyűrűváltozata is ($\tau_{\mathbf{E}_1, \mathbf{E}_1}$), továbbá azok a konformerek, melyekben az első peptidkötés *cisz* állású ($\mathbf{C}\mathbf{E}_1, \mathbf{E}_1$). A $0,6 \text{ kcal.mol}^{-1}$ -os küszöbérték alatti még a hosszított $\tau_{\mathbf{E}_1, -\mathbf{C}\alpha_1+}$ ($\tau = -135^\circ$ és $d = 8,3 \text{ \AA}$) és egy VIa típusú β -kanyar ($\tau_{\mathbf{E}_1, -\mathbf{C}\alpha_1+}$, $\tau = 28,5^\circ$ és $d = 5,2 \text{ \AA}$). Meglepően alacsony energiájú (2 kcal.mol^{-1} alatti) némely két *cisz* peptidkötést is magában foglaló szerkezet: a $\mathbf{C}\mathbf{E}_1, +\mathbf{C}\alpha_1+$ és a $\mathbf{C}\mathbf{E}_1, +\mathbf{E}_1+$ ($\tau = 13^\circ$ ill. 53° , $d = 4,2 \text{ \AA}$ ill. $5,3 \text{ \AA}$). Mindkettő β -kanyarnak tekintendő, bár a klasszikus kanyarok között ezeket még nem osztályozták. Ezzel szemben, a gázfázisú számítások 5 kcal.mol^{-1} alatti relatív energiát nem rendeltek egyetlen *cc* peptidkötés-mintázatú konformerhez sem.

Tanulmányoztam a $\text{HCO-L-Pro-L-Pro-NH}_2$ modellvegyület két prolinja közötti peptidkötés *cisz-transz* izomerizációját az alacsonyabb energiájú átmeneti állapotban (TS) keresztül, melyben az ω_2 (ill. ζ_2) torziószög közel 90° -os. Bizonyítottam, hogy a TS valóban a $\tau_{\mathbf{E}_1, +\mathbf{C}\alpha_1+}$ és a $\tau_{\mathbf{C}\alpha_1, +\mathbf{C}\alpha_1+}$ minimum között található (ld. **E.3.3-9. Ábrák**). E nyugalmi szerkezetek egyike sem azonos a modell globális minimumával (ld. **E.3.6. Táblázat**). A *cisz-transz* izomerizáció összetett folyamat. A C–N kötés körüli elforduláson kívül, melyhez a nitrogénatom piramidalizációja társul, más torziószögek is megváltozhatnak.



E.3.2. Ábra (ld. következő oldal)

E.3.2. Ábra (folytatás) A legfontosabb **Pro–Pro** gerinckonformerek a HCO–L–Pro–L–Pro–NH₂ modellpeptiden bemutatva, valamint a PDB SELECT-ből nyert **Pro–Pro** konformerek d - τ diagramja. A pontok néhány kevésbé gyakori konformerhez tartoznak (konformációs kód feltüntetése nélkül) Megjegyzendő, hogy egy adott konformer kísérletileg meghatározott τ és d értékei és az egyes elméleti szinteken számított adatok közötti több-kevesebb eltérés mutatkozik. (Ld. **M.5. Ábra** a konformációs kód kifejtéséhez és **M.7. Ábra** a τ és a d paraméterek meghatározásához.)



E.3.3. Ábra A For-L-Pro-L-Pro–NH₂ modellpeptid kedvezőbb energiájú *cis-transz* átmeneti állapota a vonatkozó minimumokkal. A hidrogénhid-kötéseket hidrogén-heteroatom távolságok jellemzik.

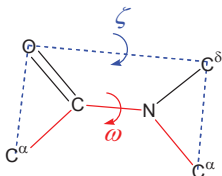
E.3.6. Táblázat A For-L-Pro-L-Pro-NH₂ modellpeptid kedvezőbb energiájú *cis*-*trans* átmeneti állapota (TS) a vonatkozó minimumokkal. A minimumok és az átmeneti állapot relatív energiái változatos elméleti szinteken.

ΔE^a	$t\epsilon_L + c\alpha_L +$	TS	$t\alpha_L + t\alpha_L +$
RHF/3-21G	0,41	16,89	9,41
RHF/6-31+G(d)	0,01	16,40	6,38
B3LYP/6-311++G(d,p)//RHF/6-31+G(d)	0,24	18,49	6,16
B3LYP/6-311++G(2d,2p)//RHF/6-31+G(d)	0,32	18,51	6,09

^a ΔE relatív energia kcal.mol⁻¹ egységben, az elméleti szintnek megfelelő globális minimumhoz viszonyítva:

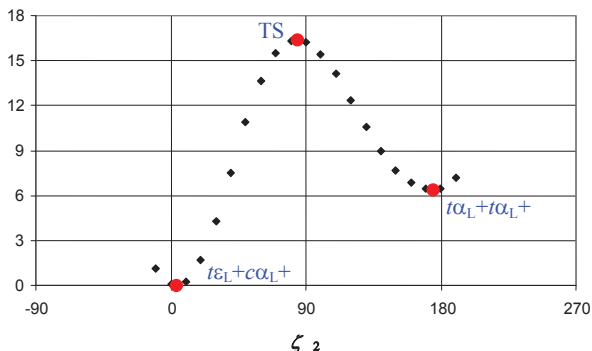
RHF/3-21G	$t\epsilon_L + t\gamma_L +$	$E_h = -809,913664$ Hartree
RHF/6-31+G(d)	$t\epsilon_L + t\gamma_L +$	$E_h = -814,458611$ Hartree
B3LYP/6-311++G(d,p)//RHF/6-31+G(d)	$t\epsilon_L - t\gamma_L +$	$E_h = -819,615470$ Hartree
B3LYP/6-311++G(2d,2p)//RHF/6-31+G(d)	$t\epsilon_L - t\gamma_L +$	$E_h = -819,640875$ Hartree

A ζ_2 paraméter alkalmazása az ω_2 torzióval szemben (ld. **E.3.4. Ábra**) azért javasolt, mert nem szerepel benne a piramidalizálódó (az amidcsoport síkjából kiemelkedő) nitrogénatom.¹¹⁷ A reakciót számítása során ζ_2 -t 10°-onként léptettem (ld. **E.3.5. Ábra**). Az ω_2 valódi és a ζ_2 képzetes torziószög egymással szoros korrelációt mutatnak (ld. **E.3.6. Ábra**). A $t\epsilon_L + c\alpha_L +$ és $t\alpha_L + t\alpha_L +$ közötti útvonalon során, a ψ_2 torziószög 143°-ról -37°-re csökken (*antiból gauche*- lesz), míg ψ_3 csaknem változatlan (ld. **E.3.7. Ábra**). Miközben a pszeudorotációs amplitúdók (A_2 és A_3) csak kissé módosulnak, a fázisszögek (P_2 és P_3) jelentős változáson mennek át (ld. **E.3.8.** és **E.3.9. Ábra**). Fontos megjegyezni, hogy a fázisszögek 36°-nyi változása képes pl. a pirrolidingyűrű egyik borítékalkatát egy másikba átvinni, ilyenkor egy másik atom távolodik el a többi négy atom síkjából (vö. **M.5/E Ábra**). A pirrolidingyűrűnek a P paraméter mentén két szomszédos borítékalkata között félfőlön mindig kialakul egy csavart konformációja (ld. **E.3.9. Ábra** példája).



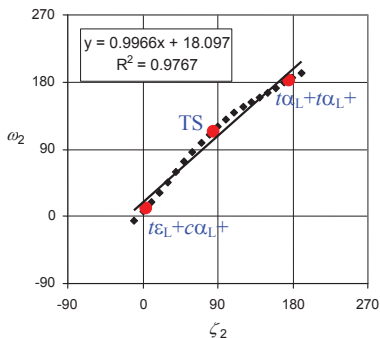
E.3.4. Ábra A peptidkötést jellemző valódi és képzetes torziószög. Az ω torziót a C^α, karbonil C, N és a következő C^α jelöli ki. A ζ képzetes torziószöveget a C^α, a karbonil O, pirrolidingyűrű C^δ atomja és a következő C^α definiálja.

A minimumokat és a TS-t is egy hidrogénhid-kötés stabilizálja, mely a formil oxigén és az egyik terminális amid hidrogén között alakul ki (ld. **M.8. Ábra** A típus). A TS-ben ugyanez a hidrogénatom a második pirrolidingyűrű piramidalizálódott nitrogénjének is a közelébe kerül (ld. **E.3.3. Ábra**).

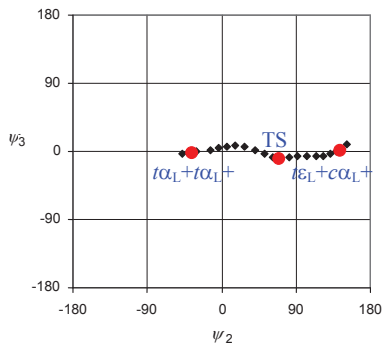


E.3.5. Ábra A *cis-transz* reakcióút energiadiagramja (RHF/6-31+G(d) elméleti szinten). Relatív energia (ΔE kcal.mol⁻¹ egységben) a globális minimumhoz ($t\epsilon_1+t\gamma_1+$) viszonyítva. Minimumok és TS körrel kiemelve és címkézve.

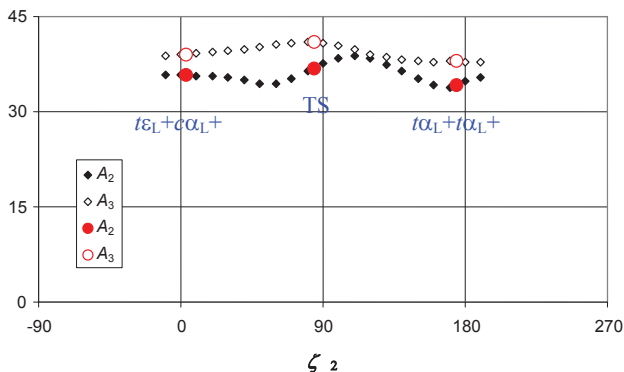
A minimumok és a TS RHF/6-31+G(d) elméleti szintű szerkezeteire DFT módszerrel is számoltam energiákat (ld. **E.3.6. Táblázat**). A nyugalmi szerkezeteknek a globális minimumhoz viszonyított energiái alig változtak a polarizációs függvények és a diffúz függvények növekvő száma mellett. A TS relatív energiája azonban emelkedett mind a 6-311++G(2d,2p), mind pedig a 6-311++G(2d,2p) bázis alkalmazásakor. A *cis-transz* izomerizáció általános képe nem változik jelentősen sem a függvénykészlet növelésének, sem a módszer változtatásának hatására.



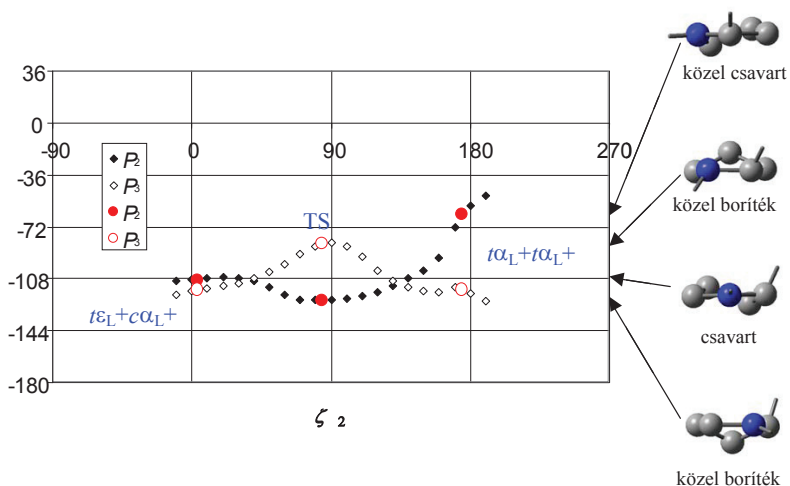
E.3.6. Ábra Az ω_2 valódi és a ζ_2 képzetes torziószög kapcsolata. Mindkét paraméter fok egységben. Minimumok és TS körrel kiemelve és címkézve.



E.3.7. Ábra A *cis-transz* izomerizáció ψ_2 - ψ_3 diagrammja. Mindkét paraméter fok egységben. Minimumok és TS körrel kiemelve és címkézve.



E.3.8. Ábra A pirrolidingyűrük pszeudorotációs amplitúdóinak (A_2 és A_3) változása a reakciút során. Mindhárom paraméter fok egységben. Minimumok és TS körül kiemelve és címkézve.



E.3.9. Ábra A pirrolidingyűrük pszeudorotációs fázisainak (P_2 és P_3) változása a reakciút során. Minden paraméter fok egységben. Minimumok és TS körül kiemelve és címkézve. $P = -72^\circ$ pirrolidin csavart konformációját jelenti, melyben N, C^α és C^β síkja felett van C^γ , míg a sík alatt C^δ . $P = -90^\circ$ boríték konformációt ad, melyben C^γ a többi négy atom síkja felett helyezkedik el. $P = -108^\circ$ csavart alakzatban a C^δ , N és C^α sík alatt található C^β , míg a sík fölött C^γ . $P = -126^\circ$ boríték konformációt jelöl, melyben C^β a többi négy atom síkja alatt látható. A gyűrűállásokat szemléltető ábrákon a peptidgerinchez tartozó egy-egy kötés megjelöli a N ill. a C^α atomot (vö. **E.2.3. Ábra** azonos gyűrűkonformációi).

Peptidszekvenciák statisztikai elemzése

A Cambridge-i Szerkezeti Adatbázisban (CSD, 2001. áprilisban) 21 különböző peptid^d tartalmazott legalább egy Xxx–**Pro–Pro**–Yyy szekvenciát. Ezek mindegyike ciklikus, és 6, 8, 9 vagy 10 aminosavegységből áll. Összesen 34 szekvenciát találtam (ld. **E.3.7. Táblázat**). Figyelembe véve e peptidek kis számát és ciklikus alkatát, konformációs adataiktól nem várhatunk többet, mint hogy kísérleti oldalról alátámasztják bizonyos **Pro–Pro** konformerek létét és szerkezeti tulajdonságait. Semmiképpen sem remélhető, hogy megmutatnák az összes minimumot, melyek nyíltláncú peptidekben és fehérjékben előfordulhatnak.

Az esetek nagy többségében (94%) legalább egy *cisz* peptidkötés szerepel, de 9%-ban a prolint megelőző peptidkötések mindegyike *cisz* állású. A leggyakoribb gerinckonformer a **t_ε1.c_αL** (67%), mely VIa típusú β-kanyar (ld. **E.3.2. Ábra**). A másik *cisz* peptidkötést tartalmazó β-kanyar típus, tehát a VIb avagy **t_ε1.c_εL**, már ritkább (18%).

Több ciklopeptidben szerepel a **Pro–Pro** szekvenciaegység kétszer. Rendszerint az ismétlődés azonos konformert is jelent (vagyis a paraméterek bár különböznek, még azonos konformert jellemeznek), de legalább azonos gerinckszerkezetet eltérő gyűrűállással. A prolil-prolin konformerek azonossága nem függ a teljes ciklopeptid szimmetrikus/aszimmetrikus voltától, de kellően hosszú peptidláncot igényel. Mindkét ciklohexapeptid, melyben a prolil-prolin ismétlődik, két különböző konformert ad. A CSD-ben BINJIR-ként címkézett ciklo(Gly–**Pro–Pro**–Gly–**Pro–Pro**)¹¹⁸ **t_α1.–t_α1+** és **c_ε1.+c_ε1+** konformert ad, míg a BUYXOI jelű ciklo(**Pro–Pro**–pro–**Pro–Pro**–pro)¹¹⁹, mely D-prolint is tartalmaz, **t_ε1.–t_ε1+** és **c_ε1.+c_ε1.–** szerkezetet hordoz. A **c_ε1.c_εL** gerinckonformer harmadik előfordulása az egyetlen további hexapeptidhez kapcsolódik (ciklo(**Pro–Pro**–Gly–Pro–Leu–Gly),¹²⁰ BUNYEO10). A **c_ε1.c_εL** gerinckszerkezetnek a hexapeptidekre jellemző előfordulása a *d* paraméter igen kicsi értékével magyarázható. A *τ* és a *d* paraméter értékei alapján a **c_ε1.c_εL** szerkezet egy ez ideig nem osztályozott β-kanyarnak tekintendő, jóllehet az 1→4 hidrogénkötés, mely rendszerint jelen van a többi β-kanyarban, itt nem alakulhat ki az első prolint megelőző peptidkötés *cisz* állása miatt.

^d A peptidek CSD azonosító kódja: ANTAHC10, ANTAML10, BINJIR, BUNYEO10, BUYXOI, DOLJAP, DUTLAF10, GIPKAR10, HEBKUU, JUJHUR, LINCOA, NEHYAA, NEHYAA01, NEHYAA02, NEHYAA03, PAANTD, PAANTD01, PVANSB, PVANTS, VEDJIX, VENXER

E.3.7. Táblázat A CSD-ből kigyűjtött peptidok **Pro–Pro** konformereinek fontosabb geometriai adatai.

Konformer	<i>n</i>	%	ω_1	ϕ_2	ψ_2	A_2	P_2	ω_2	ϕ_3	ψ_3	A_3	P_3
$t\alpha_L-t\alpha_L+$	1	3	-177	-60	-42	37	81	-176	-84	-5	37	-124
$t\epsilon_L-t\epsilon_L+$	1	3	176	-46	129	38	79	-163	-70	126	38	-82
$t\epsilon_L+c\alpha_L+$	11	32	176	-64	154	22	-111	7	-91	0	38	-118
$t\epsilon_L-c\alpha_L+$	12	35	173	-65	155	25	105	4	-90	-3	36	-116
$t\epsilon_L+c\epsilon_L+$	1	3	179	-64	139	16	-102	-4	-81	148	22	-139
$t\epsilon_L+c\epsilon_L-$	3	9	-178	-67	145	14	-85	-5	-70	143	29	98
$t\epsilon_L-c\epsilon_L-$	2	6	-179	-65	147	16	83	-10	-70	144	31	110
$c\epsilon_L+c\epsilon_L+$	1	3	8	-67	154	20	-95	5	-73	154	39	-106
$c\epsilon_L+c\epsilon_L-$	1	3	4	-83	176	42	-121	-3	-73	170	17	167
$c\epsilon_L-c\epsilon_L-$	1	3	-17	-66	170	26	138	-6	-64	157	12	114

n: az adott konformerhez tartozó L–Pro–L–Pro szekvenciaszakaszok száma; %: ugyanez százalékban. Minden torziószög a fellelhető szekvenciák nyersadataiból átlagolva és fok egységben megadva.

Fehérjeszekvenciák statisztikai elemzése

A PDB SELECT-ből gyűjtött 488 Xxx-**Pro-Pro**-Yyy szekvenciából 89-nek NMR-rel végezték el a térszerkezeti jellemzését, 203-nak legalább 2 Å felbontású röntgendiffrakcióval, 196-nak pedig ennél gyengébbel (ez rendre 18%, 42% és 40%). A különböző **Pro-Pro** konformerek gyakoriságát az **E.3.8. Táblázat** mutatja, a kiválasztott geometriai paraméterek átlagait pedig az **E.3.9. Táblázat**. Mivel az elemzést a PDB SELECT nem homolog szekvenciáiból származó **Pro-Pro** egységeken végeztem el, a **BEVEZETŐBEN** említett fehérjék és polipeptidok adatai csak részben szerepelnek köztük.

A 488 Xxx-**Pro-Pro**-Yyy szekvenciában több *cisz* peptidkötés is előfordul, de mindegyik prolint előz meg. Egyetlen fehérjét találtam, melyben egymás után hét prolin is szerepel. Egy másikban öt, négyben pedig négy közvetlenül szomszédos prolin azonosítható. Ezek rendre: (**Pro**)₇ (PDB azonosító kód: 1jvr) gerincszerkezete $f_{E1}f_{E1}f_{E1}f_{E1}f_{E1}f_{E1}f_{E1}$; (**Pro**)₅ (1f32) $f_{E1}f_{E1}f_{E1}f_{E1}f_{E1}$; (**Pro**)₄ (1eoi és 1i9s) $f_{E1}f_{E1}f_{E1}f_{E1}$, (1gcq) $c_{E1}f_{E1}f_{E1}f_{E1}$, (1hg4) $f_{E1}f_{E1}f_{E1}f_{E1}$. A 22 (**Pro**)₃ szekvenciaegység számos *cisz* peptidkötést is tartalmaz. Elemzésemben a többszörös prolinok minden egyes prolin-prolinját külön vettem figyelembe.

A peptidkötések többsége *transz* állású (ld. **E.3.8. Táblázat**). *Cisz* peptidkötés valószínűsége ω_1 ill. ω_2 helyzetben 7 ill. 8%. Megjegyzendő, hogy az NMR szerkezetek kevesebb *cisz* amidot jeleznek (3 ill. 6% a fenti pozíciókban). Az elemzésbe egyetlen *cisz* peptidkötéspár került be: Cys-**Pro-Pro**-Leu ($c_{E1}c_{E1}$, 1f0i). Az ω_1 ill. az ω_2 torziók átlaga 179° ill. 180° *transz* kötésekre, -2 ill. 0° *cisz* amidokra.

Xxx és Yyy helyzetben a különböző aminosavak eltérő mértékben kedvelték az egyes peptidkötés-mintázatokat (*cc*, *ct*, *tc*, *tt* az ω_1 és az ω_2 torzók állására vonatkoznak) (ld. **E.3.10. Táblázat** és **E.3.10. Ábra**). Feltűnő, hogy a mindössze hat Trp-**Pro-Pro**-Yyy szekvenciából három-három *ct* ill. *tt* mintázatú. A *ct* gyakori akkor is, ha Xxx tirozin, aszparaginsav vagy aszparagin (27, 22, 17%), és ritka akkor, ha Xxx leucin, prolin vagy hisztidin. Jelentős a *tc* aránya, ha Xxx glutamin (31%), de csekély akkor, ha Xxx leucin, szerin, hisztidin vagy triptofán. Ha Yyy helyzetben cisztein áll, akkor gyakori a *ct*, viszont ritka tirozin, hisztidin, metionin és triptofán előtt. A *tc* mintázat a kedvezőbb, ha Yyy helyzetben fenilalanin vagy tirozin található (40, 26%), de sohasem előzi meg a következő nyolc aminosavat: prolin, szerin, glutamin, izoleucin, aszparaginsav, cisztein, metionin, triptofán).

E.3.8. Táblázat A PDB SELECT-ből kigyűjtött **Pro–Pro** fehérjeszakaszok konformereinek természetes gyakorisága.

Konformer	fb $\leq 2\text{\AA}$ ^a	fb $> 2\text{\AA}$ ^a	NMR ^a	mind ^a	fb $\leq 2\text{\AA}$ ^b	fb $> 2\text{\AA}$ ^b	NMR ^b	mind ^b
$\alpha_{L1}+\alpha_{L1}+$	0	2	1	3	0	1	1	1
$\alpha_{L1}-\alpha_{L1}+$	1	3	0	4	0	2	0	1
$\alpha_{L1}-\alpha_{L1}-$	6	3	0	9	3	2	0	2
$\alpha_{L1}+\tau_{L1}+$	0	0	1	1	0	0	1	0
$\alpha_{L1}-\tau_{L1}-$	0	0	2	2	0	0	2	0
$\tau_{L1}+\alpha_{L1}+$	0	1	0	1	0	1	0	0
$\tau_{L1}+\tau_{L1}+$	0	0	1	1	0	0	1	0
$\tau_{L1}+\tau_{L1}-$	0	0	1	1	0	0	1	0
$\tau_{L1}-\tau_{L1}-$	0	0	1	1	0	0	1	0
$\tau_{L1}+\alpha_{L1}+$	6	7	13	26	3	4	15	5
$\tau_{L1}+\alpha_{L1}-$	11	12	1	24	5	6	1	5
$\tau_{L1}-\alpha_{L1}+$	10	11	0	21	5	6	0	4
$\tau_{L1}-\alpha_{L1}-$	15	20	3	38	7	10	3	8
$\tau_{L1}+\tau_{L1}+$	3	0	9	12	1	0	10	2
$\tau_{L1}+\tau_{L1}-$	1	0	0	1	0	0	0	0
$\tau_{L1}-\tau_{L1}+$	2	1	3	6	1	1	3	1
$\tau_{L1}-\tau_{L1}-$	0	0	2	2	0	0	2	0
$\tau_{L1}+\tau_{L1}+$	43	25	34	102	21	13	38	21
$\tau_{L1}+\tau_{L1}-$	23	23	0	46	11	12	0	9
$\tau_{L1}-\tau_{L1}+$	21	21	4	46	10	11	4	9
$\tau_{L1}-\tau_{L1}-$	27	36	5	68	13	18	6	14
$\alpha_{L1}+c\epsilon_{L1}+$	0	1	0	1	0	1	0	0
$\tau_{L1}+c\alpha_{L1}+$	6	0	0	6	3	0	0	1
$\tau_{L1}+c\alpha_{L1}-$	0	1	0	1	0	1	0	0
$\tau_{L1}-c\alpha_{L1}+$	7	2	1	10	3	1	1	2
$\tau_{L1}-c\alpha_{L1}-$	0	3	0	3	0	2	0	1
$\tau_{L1}+c\epsilon_{L1}+$	2	1	1	4	1	1	1	1
$\tau_{L1}+c\epsilon_{L1}-$	0	1	0	1	0	1	0	0
$\tau_{L1}-c\epsilon_{L1}+$	2	3	0	5	1	2	0	1
$\tau_{L1}-c\epsilon_{L1}-$	1	1	1	3	0	1	1	1
$c\epsilon_{L1}+\alpha_{L1}+$	1	1	1	3	0	1	1	1
$c\epsilon_{L1}-\alpha_{L1}+$	1	0	0	1	0	0	0	0
$c\epsilon_{L1}-\alpha_{L1}-$	0	1	0	1	0	1	0	0
$c\epsilon_{L1}+\tau_{L1}+$	0	2	1	3	0	1	1	1
$c\epsilon_{L1}-\tau_{L1}-$	1	0	0	1	0	0	0	0
$c\epsilon_{L1}+\tau_{L1}+$	4	3	2	9	2	2	2	2
$c\epsilon_{L1}+\tau_{L1}-$	6	6	1	13	3	3	1	3
$c\epsilon_{L1}-\tau_{L1}+$	1	3	0	4	0	2	0	1
$c\epsilon_{L1}-\tau_{L1}-$	1	2	0	3	0	1	0	1
$c\epsilon_{L1}+c\alpha_{L1}+$	1	0	0	1	0	0	0	0
<i>tt</i>	169	165	81	415	83	84	91	85
<i>tc</i>	18	13	3	34	9	7	3	7
<i>ct</i>	15	18	5	38	7	9	6	8
<i>cc</i>	1	0	0	1	0	0	0	0
++	66	43	64	173	33	22	72	35
+-	41	43	3	87	20	22	3	18
-+	45	44	8	97	22	22	9	20
—	51	66	14	131	25	34	16	27
all	203	196	89	488	100	100	100	100

^a konformerek száma; ^b ugyanaz százalékban; fb: röntgen szerkezetek felbontása

E.3.9a. Táblázat A PDB SELECT-ből kigyűjtött Pro-Pro fehérjeszakaszok konformereinek néhány geometriai adata.

Konformer	<i>n</i>	ϕ_1	ϕ_2	ψ_2	A_2	P_2	ϕ_2	ψ_2	A_3	P_3	τ	<i>d</i>	
$\alpha_{i-1}\alpha_i+$	3	179	-62	-41	37	-87	179	-70	-33	40	-98	49	^a 5,6
$\alpha_{i-1}\alpha_i+$	4	-179	-46	-45	46	78	178	-61	-30	42	-88	54	^a 5,5
$\alpha_{i-1}\alpha_i-$	9	-178	-49	-43	41	76	-180	-61	-29	39	84	57	^a 5,6
$\alpha_{i-1}\gamma_{i-1}+$	1	173	-68	-44	28	-90	179	-79	62	36	-119	62	^b 6,4
$\alpha_{i-1}\alpha_i-$	2	-179	-59	-51	18	109	-178	-60	-142	19	108	57	7,4
$\gamma_{i-1}\alpha_i+$	1	-173	-58	91	21	-9	178	-37	-14	37	-104	-150	8,5
$\gamma_{i-1}\alpha_i+$	1	169	-88	96	32	-129	179	-65	109	32	-105	-144	9,1
$\gamma_{i-1}\alpha_i-$	1	172	-95	80	41	-130	-163	-67	111	37	77	-145	8,9
$\gamma_{i-1}\alpha_i-$	1	-180	-104	28	32	176	180	-60	134	1	70	153	9,3
$\alpha_{i-1}\alpha_i+$	26	179	-73	143	31	-116	-178	-67	-34	30	-109	-131	8,7
$\alpha_{i-1}\alpha_i-$	24	180	-69	154	38	-105	-179	-51	-36	40	77	-108	8,2
$\alpha_{i-1}\alpha_i+$	21	180	-58	138	42	84	-180	-61	-26	39	-91	-122	8,4
$\alpha_{i-1}\alpha_i-$	38	178	-60	145	40	89	-179	-51	-41	40	77	-115	8,3
$\alpha_{i-1}\gamma_{i-1}+$	12	179	-73	146	33	-122	-179	-74	70	32	-118	-107	8,3
$\alpha_{i-1}\gamma_{i-1}+$	1	179	-76	173	47	-106	179	-76	50	41	96	-86	7,7
$\alpha_{i-1}\gamma_{i-1}+$	6	-179	-65	152	34	107	178	-72	73	32	-112	-100	8,2
$\alpha_{i-1}\gamma_{i-1}+$	2	180	-57	179	32	83	180	-52	96	32	74	-44	7,1
$\alpha_{i-1}\alpha_i+$	102	180	-70	154	34	-109	179	-70	153	33	-108	-99	9,0
$\alpha_{i-1}\alpha_i-$	46	180	-70	155	38	-105	179	-60	144	38	84	-83	8,5
$\alpha_{i-1}\alpha_i+$	46	179	-66	149	37	95	179	-68	152	37	-104	-101	8,9
$\alpha_{i-1}\alpha_i-$	68	178	-64	154	37	94	179	-59	142	37	82	-83	8,5
$\alpha_{i-1}c\alpha_i+$	1	-180	-45	-49	44	-78	1	-97	167	46	-118	-112	7,1
$\alpha_{i-1}c\alpha_i+$	6	180	-66	154	28	-97	-6	-84	7	42	-125	24	^c 5,9
$\alpha_{i-1}c\alpha_i-$	1	177	-89	165	43	-101	0	-75	2	48	87	32	^c 6,0
$\alpha_{i-1}c\alpha_i+$	10	177	-57	146	37	71	2	-86	6	40	-118	29	^c 5,8
$\alpha_{i-1}c\alpha_i-$	3	178	-55	159	45	90	0	-73	-7	39	86	47	^c 6,0
$\alpha_{i-1}c\alpha_i+$	4	180	-66	157	36	-95	3	-76	141	38	-109	69	^d 6,8
$\alpha_{i-1}c\alpha_i+$	1	180	-61	155	43	-98	0	-85	148	35	161	62	^d 7,1
$\alpha_{i-1}c\alpha_i+$	5	179	-59	158	44	85	0	-80	147	43	-111	74	^d 7,4
$\alpha_{i-1}c\alpha_i-$	3	180	-59	156	38	76	0	-68	143	38	109	77	^d 6,9

(folytatás következik)

(az E.3.9a. Táblázat folytatása)

$c_{E1}+f_{O1}+$	3	-4	-72	153	34	-116	178	-79	-18	29	-90	-169	7,7
$c_{E1}-f_{O1}+$	1	0	-62	141	45	81	-178	-70	-25	44	-115	-157	7,8
$c_{E1}-f_{O1}-$	1	2	-74	137	28	126	176	-60	-29	22	95	-168	7,6
$c_{E1}+f_{Y1}+$	3	0	-75	161	38	-111	-180	-72	54	36	-111	-129	7,3
$c_{E1}-f_{Y1}-$	1	-1	-70	162	41	91	178	-74	98	43	94	-120	7,2
$c_{E1}+f_{E1}+$	9	-7	-71	158	37	-118	-179	-76	156	37	-112	-130	8,2
$c_{E1}+f_{E1}-$	13	-1	-78	164	38	-119	179	-66	157	36	92	-122	8,2
$c_{E1}+f_{E1}+$	4	-2	-74	177	39	116	-176	-72	141	43	-99	-98	7,4
$c_{E1}-f_{E1}-$	3	0	-72	168	44	98	180	-64	144	44	86	-108	7,6
$c_{E1}+c_{O1}+$	1	-12	-72	174	37	-106	1	-91	-1	36	-130	25	^c 4,9

n : az adott konformerhez tartozó **Pro-Pro** szekvenciaszakaszok száma; %; ugyanez százalékban. Minden torziósög az fellelhető szekvenciák nyersadataiból átlagolva és fok egységben megadva, kiv.: d , melynek egysége Å.

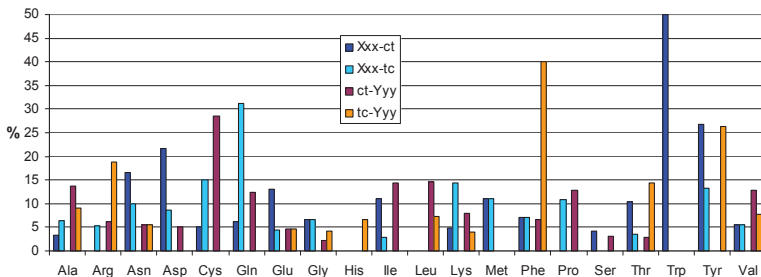
^a III típusú β -kanyar, ^b I típusú β -kanyar, ^c VIa típusú β -kanyar, ^d VIb típusú β -kanyar, ^e β -kanyarnak tekinthető τ és d értékei alapján, de ezideig nem definiált.

E.3.9b. Táblázat Pirrolidingyűrűk pszeudorotációs paramétereinek átlagértékei a PDB SELECT-ből kigyűjtött **Pro-Pro** fehérszakaszokban.

Gyűrűállás	n	A_2	P_2	A_3	P_3	χ_{20}	χ_{21}	χ_{22}	χ_{23}	χ_{24}	χ_{30}	χ_{31}	χ_{32}	χ_{33}	χ_{34}
++	173	33	-110	34	-109	-11	28	-33	26	-9	-11	28	-34	27	-10
+–	87	38	-107	39	84	-11	31	-38	31	-12	4	-26	38	-35	20
–+	97	39	90	38	-103	0	-23	37	-37	23	-8	29	-38	33	-15
—	131	38	92	38	82	-1	-21	36	-37	23	5	-26	37	-34	18

n : azon **Pro-Pro** szekvenciák száma, melyekben a pirrolidingyűrű az adott kombinációs párt képezik.

Minden geometriai paraméter fokban, A ill. P : az adott gyűrűállású konformerek súlyozott átlaga; χ_{ij} : a kötésszögnek a vonatkozó A és P pszeudorotációs paramétereiből visszászarmaztatott tipikus értéke (ld. **(4) Egyenlet**).



E.3.10. Ábra A *ct* és *tc* amidállások százalékos gyakorisága a PDB SELECT-ből gyűjtött 488 *Xxx-Pro-Pro-Yyy* fehérjeszakasz *Xxx* ill. *Yyy* helyzetű aminosavának függvényében. (Vö. E.3.10. Táblázat.)

E.3.10. Táblázat A különböző aminosavak, valamint az egyes amidállások gyakorisága a PDB SELECT-ből gyűjtött 488 *Xxx-Pro-Pro-Yyy* fehérjeszakaszban.

<i>Xxx</i>	<i>n</i>	% <i>ct</i>	% <i>tc</i>	% <i>tt</i>	<i>Yyy</i>	<i>N</i> ^o	% <i>ct</i>	% <i>tc</i>	% <i>tt</i>
Ala	31	3	6	90	Ala	22	14	9	77
Arg	19	0	5	95	Arg	16	6	19	75
Asn	30	17	10	73	Asn	18	6	6	89
Asp	23	22	9	70	Asp	20	5	0	95
Cys	*20	5	15	75	Cys	7	29	0	71
Gln	16	6	31	63	Gln	24	13	0	88
Glu	23	13	4	83	Glu	44	5	5	91
Gly	15	7	7	87	Gly	47	2	4	94
His	8	0	0	100	His	15	0	7	93
Ile	36	11	3	86	Ile	21	14	0	86
Leu	62	0	0	100	Leu	*41	15	7	76
Lys	21	5	14	81	Lys	25	8	4	88
Met	9	11	11	78	Met	7	0	0	100
Phe	28	7	7	86	Phe	15	7	40	53
Pro	37	0	11	89	Pro	39	13	0	87
Ser	24	4	0	96	Ser	33	3	0	97
Thr	29	10	3	86	Thr	35	3	14	83
Trp	6	50	0	50	Trp	1	0	0	100
Tyr	15	27	13	60	Tyr	19	0	26	74
Val	36	6	6	89	Val	39	13	8	79
all	488	8	7	85	all	488	8	7	85

n: az adott aminosav *Xxx* ill. *Yyy* helyzetű megjelenésének száma. %: az egyes amidállások százalékos gyakorisága az adott aminosav után/előtt; *ct*: ω_1 *cisz*, *tc*: ω_2 *cisz*, *tt*: ω_1 és ω_2 is *transz*. *Egyetlen *cc* amidállás található a 488 eset között, nevezetesen egy *Cys-Pro-Pro-Leu* szekvenciaszakaszban: ezt a *ct* és *tc* százalékaiba is beleszámoltam.

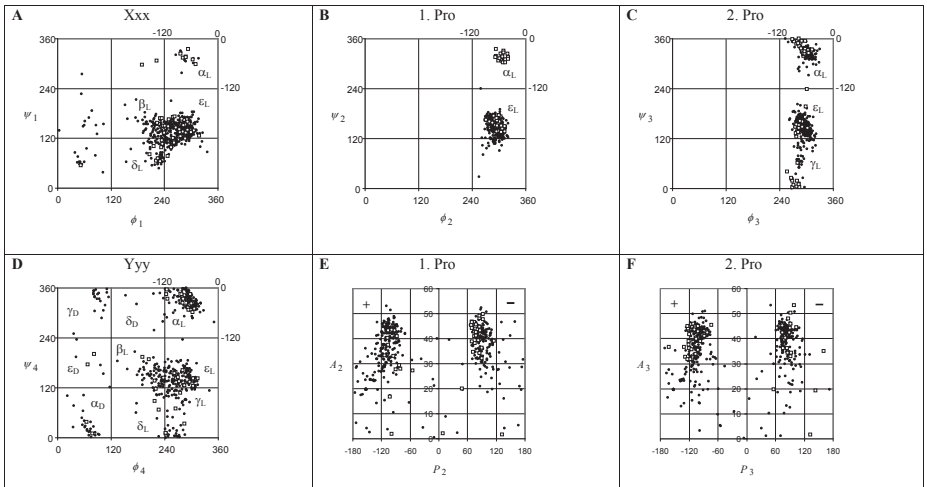
Ismeretes, hogy a prolin előtti szekvenciaszakaszok előnyben részesítik a nyújtott-hosszított szerkezeteket (β_L , ϵ_L és δ_L , vö. M.1/C Ábra) a helikállissal (α_L) szemben.¹²¹ Az *Xxx-Pro-Pro-Yyy* szekvenaciaegységekben ez igaz az *Xxx* aminosavra és az első prolinra is,

hiszen ϕ 120°-nál nagyobb és ψ 40° és 240° közötti ezeknek az aminosavaknak 92 és 96%-ában (ld. **E.3.11. Ábra**). A második prolin és Yyy már nem prolin előtti, így Ramachandran-térképeik ugyanezen régiója csak az aminosavak 69 és 50%-át tartalmazza. Xxx és Yyy Ramachandran-térképeinek (**E.3.4/A** és **D Ábra**) összevetéséből kitűnik, hogy a δ_L régiót (más konvenció szerint ζ régiót)¹²² kifejezetten a prolint megelőző aminosavak képesek benépesíteni. A δ_L területet a prolin gyűrűs szerkezete miatt nem képes elérni, viszont feltűnő, hogy a γ_L régiót elkerüli az Xxx és az első prolin, de sűrűsödést mutatkozik itt a második prolin és az Yyy aminosav esetében. Az α_L konformerekben a ψ torziószög 295°-nál kisebb a prolin előtti aminosavakban, míg az α_L -góc jóval 360° fölé tolódik olyan aminosavaknál, melyeket nem követ prolin.

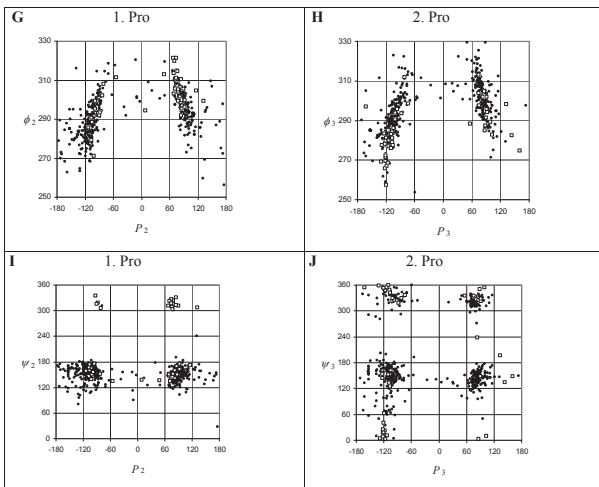
Miközben a *transz* és *cisz* peptidkötések könnyedén megjelölhetők, az ϵ_L , γ_L és α_L gerincek elválasztása nem magától értetődő (ld. **E.3.11/C Ábra**). A $\psi = 20^\circ$ ill. 100° küszöbértékeket választottam τ_{γ_L} -nek τ_{α_L} -tól ill. τ_{ϵ_L} -től való megkülönböztetésére, mivel ezeken a határértékeken túl a γ -kanyarra jellemző hidrogénkötés (B típusú, ld. **M.8. Ábra**) nem alakulhat ki. *Cisz* peptidkötés után sem alakulhat ki ez a hidrogénkötés, és $c\alpha_L$ meggyőzően el is válik $c\epsilon_L$ -től, minthogy ψ sohasem esik 41° és 124° közé. Így $c\gamma_L$ konformert nem is címkéztem. Az NMR a röntgendiffrakciónál jelentősen gyakrabban ad meg a fehérjeszerkezetekben τ_{γ_L} konformereket (ld. $\tau_{\gamma_L}\tau_{\epsilon_L}$ és $\tau_{\epsilon_L}\tau_{\gamma_L}$ gerincek az **E.3.8. Táblázatban**).

A *tt* konformerek közül $\tau_{\alpha_L}\tau_{\alpha_L}$ és $\tau_{\alpha_L}\tau_{\gamma_L}$ valódi β -kanyarnak tekintendő, nevezetesen III és I típusúnak (ld. **E.3.9a Táblázat** és **E.3.2. Ábra**), mivel τ és d értékeik kicsik (49-62°, és 5,6-6,4 Å). Ismeretes, hogy a III típusú β -kanyar a 3_{10} -hélixek építőeleme,¹²³ és a **Pro–Pro** egységek némelyike valóban hélixbe épül be, minthogy a megelőző és az utána következő aminosav is τ_{α_L} konformációjú. Némileg gyakoribb (22%) a hosszított $\tau_{\epsilon_L}\tau_{\alpha_L}$ gerinckonformer, mely gyakran kapcsol össze nyújtott vagy hosszított szerkezetet helikálissal, ti. Xxx τ_{ϵ_L} vagy τ_{β_L} , míg Yyy τ_{α_L} . A leggyakoribb (54%) $\tau_{\epsilon_L}\tau_{\epsilon_L}$ gerinckszerkezet a kollagénből ismert poliprolin-hélixben szerepelhet, ti. Xxx és Yyy is τ_{ϵ_L} .

A *tc* konformerek hosszítottak. Legtöbbjük β -kanyar, habár a d paraméter értéke rendszerint nagy. A $\tau_{\epsilon_L}c\alpha_L$ gerinc VIa típusú, míg a $\tau_{\epsilon_L}c\epsilon_L$ VIb típusú β -kanyar. A statisztikai elemzésben megjelenő egyetlen *cc* peptidkötés-mintázat ez ideig nem osztályozott β -kanyar.



E.3.11/A-F Ábra A PDB SELECT-ből gyűjtött 488 Xxx-Pro-Pro-Yyy fehérjeszakasz geometriai adatai. Telt körök: hosszított szerkezetek, üres négyzetek: β -kanyarok (vö. E.3.11. Táblázat és E.3.2. Ábra). Minden paraméter fok egységben. A, B, C, D: a négy aminosav Ramachandran-térképe topológiai 0-360° intervallumok között. A ϕ és a ψ torziószögek hagyományos -180-+180° intervallumok szerinti beosztása az ábrák jobb felső sarkánál jelenik meg. E, F: a prolinok pszeudorotációs koordinátái.



E.3.11/G-J Ábra A PDB SELECT-ből gyűjtött 488 Xxx-Pro-Pro-Yyy fehérjeszakasz geometriai adatai. A gyűrűállás és a gerinckonformáció összefüggései. Telt körök: hosszított szerkezetek, üres négyzetek: β -kanyarok (vö. E.3.11. Táblázat és E.3.2. Ábra). Minden paraméter fok egységben.

Az A pszeudorotációs amplitúdó átlaga 36° (ld. **E.3.11/E-F Ábra**), de értéke jelentősen függ a szerkezetmeghatározás módszerétől. A 2 \AA -nél jobb felbontású röntgenszerkezetekben a pszeudorotációs amplitúdó kb. 37° , a gyengébb felbontásúakban 40° , míg az NMR-rel vizsgáltakban ugyanez az érték mindössze 28° . Az A pszeudorotációs amplitúdó ++ gyűrűkombinációnál $33\text{--}34^\circ$, a sok NMR-adatnak köszönhetően, és $38\text{--}39^\circ$ az egyéb gyűrűállásoknál (ld. **E.3.9b Táblázat**). A pirrolidingyűrű gyakrabban vesz fel pozitív állást, mint negatívát (ld. **E.3.8. Táblázat**). A $\tau\alpha_L$ az egyetlen gerinckonformer, amelyik némileg jobban kedveli a negatív gyűrűállást. Ismét kitűnik, hogy az NMR-szerkezetek jelentősen ritkábban rendelnek negatív gyűrűállást a pirrolidingyűrűhöz, mint pozitívát. A P változó átlaga pozitív gyűrűállások (negatív P , pozitív χ_1) mellett -108° . Ez az érték egy csavart gyűrűt ír le, melyben a C^δ , N és C^α atomok síkja alatt található C^β , míg felette C^γ (ld. **E.2.3. Ábra**). Az ettől néhány fokkal eltérő P -értékek aszimmetrikus gyűrűt jelölnek, a szerkezet eltolódását egy-egy borítékalkat felé ($P = -90$ vagy -126°). Előbbiben C^γ a C^δ , N, C^α , és C^β atomok síkja fölé emelkedik, utóbbiban C^β a C^γ , C^δ , N és C^α atomok síkja alá süllyed. A negatív gyűrűállások (pozitív P , negatív χ_1) rendszerint egy borítékra hasonlítanak ($P = 90^\circ$), melyben C^γ a C^δ , N, C^α , és C^β atomok síkja alatt helyezkedik el. Az ettől néhány fokkal eltérő P -értékek aszimmetrikus gyűrűkhöz közelítenek ($P = 72$ és 108°). Előbbiben C^δ , N és C^α síkja felett van C^β és alatta C^γ , utóbbiban N, C^α és C^β síkja alatt látható C^γ , míg felette C^δ .

Összevetés

A (2) és (3) Egenlet alkalmazásával megadható a $\text{HCO-L-Pro-L-Pro-NH}_2$ peptidmodell konformációs összetétele az egyes elméleti szinteken. A PCM számítások szerint a leggyakoribb gerinckonformerek csökkenő rendben: $\tau\epsilon_L\tau\epsilon_L$, $c\epsilon_L\tau\epsilon_L$, $\tau\epsilon_Lc\alpha_L$, $\tau\epsilon_L\tau\alpha_L$, $\tau\epsilon_L\tau\gamma_L$, $\tau\alpha_L\tau\alpha_L$ (ld. **E.3.11. Táblázat** és **E.3.2. Ábra**). A két *cis* peptidkötést is tartalmazó szerkezetek ($c\epsilon_Lc\alpha_L$ és $c\epsilon_Lc\epsilon_L$) becsült előfordulása pedig csekély. A fehérjék prolin-prolin szekvenciaegységei pontosan ugyanezt a hat gerinckonformert öltik magukra leggyakrabban, habár némileg eltérő sorrendben: $\tau\epsilon_L\tau\epsilon_L$, $\tau\epsilon_L\tau\alpha_L$, $c\epsilon_L\tau\epsilon_L$, $\tau\epsilon_L\tau\gamma_L$, $\tau\epsilon_Lc\alpha_L$, $\tau\alpha_L\tau\alpha_L$. A 488 homológiaszűrt szekvenciaegység között csupán egyetlenben ismétlődött a *cis* peptidkötés ($c\epsilon_Lc\alpha_L$), egy rokon polipeptidben pedig, melyet a PDB SELECT nem tartalmaz, $c\epsilon_Lc\epsilon_L$ található.

E.3.11. Táblázat A prolin-prolin gerinckonformerek geometriai és természetes ill. becült gyakorisági adatai.

Gerinckonformer	Kísérletileg meghatározott geometriai paraméterek						Kísérleti konformációs összetétel (%)				Számított geometriai paraméterek (B3LYP/PCM/6-31+G(d))					Számított konformációs összetétel (%)					
	ϕ_2	ψ_2	ϕ_3	ψ_3	τ	d	$fb < 2^\circ$	$fb > 2^\circ$	NMR	mind	ϕ_2	ψ_2	ϕ_3	ψ_3	τ	d	PCM	a	b	c	d
hosszított							87,2	89,8	92,1	89,1							86,4	66,5	53,4	55,2	55,4
$\tau_{II}t\alpha_L$	-58	91	-37	-14	-150	8,5	0,0	0,5	0,0	0,2											
$\tau_{II}t\beta_L$	-96	68	-64	118	-165	9,1	0,0	0,0	3,4	0,6											
$t\beta_Lt\alpha_L$	-65	145	-57	-35	-120	8,4	20,7	25,5	19,1	22,3	-68	151	-75	-26	-132	8,2	7,8				
$t\beta_Lt\tau_L$	-69	152	-71	72	-98	8,1	3,0	0,5	15,7	4,3	-70	154	-84	65	-108	7,8	3,1	66,5	47,1	54,0	55,4
$t\beta_Lt\beta_L$	-67	153	-65	148	-92	8,8	56,2	53,6	48,3	53,7	-71	157	-67	148	-89	8,4	41,1		5,7		
$t\alpha_Lc\beta_L$	-45	-49	-97	167	-112	7,1	0,0	0,5	0,0	0,2											
$c\beta_Lt\alpha_L$	-70	147	-73	-22	-166	7,7	1,0	1,0	1,1	1,0											
$c\beta_Lt\tau_L$	-74	162	-73	65	-127	7,3	0,5	1,0	1,1	0,8	-79	163	-84	64	-129	7,1	1,2		0,7	1,3	
$c\beta_Lt\beta_L$	-75	164	-70	153	-120	8,0	5,9	7,1	3,4	5,9	-70	162	-68	148	-111	7,4	33,2				
β-kanyar							12,8	10,2	7,9	10,9							13,4	33,5	46,6	44,8	44,6
$t\alpha_Lt\alpha_L$	-51	-43	-62	-30	55	5,6	3,4	4,1	1,1	3,3	-55	-35	-71	-16	55	5,3	3,1				
$t\alpha_Lt\tau_L$	-68	-44	-79	62	62	6,4	0,0	0,0	1,1	0,2											
$t\alpha_Lt\beta_L$	-59	-51	-60	-142	57	7,4	0,0	0,0	0,0	2,2	0,4										
$t\beta_Lc\alpha_L$	-61	151	-83	4	30	5,9	6,4	3,1	1,1	4,1	-59	140	-90	34	29	5,2	9,0	33,5	46,6	44,8	44,6
$t\beta_Lc\beta_L$	-62	157	-76	144	71	7,0	2,5	3,1	2,2	2,7											
$c\beta_Lc\alpha_L$	-72	174	-91	-1	25	4,9	0,5	0,0	0,0	0,2	-80	161	-86	-11	13	4,2	0,8				
$c\beta_Lc\beta_L$ *	-74	171	-58	137	68	4,8					-79	164	-79	160	53	5,3	0,6				

A gerinckonformereket hosszított szerkezetekre és β -kanyarokra osztottam fel. Minden geometriai paraméter fokban, kiv.: d , mely \AA egységben. Konformációs összetétel (%): a gerinckonformerek százalékos előfordulása.

A kísérleti adatok alapjául PDB SELECT-ből gyűjtött 488 Xxx-Pro-Pro-Yyy fehérjeszakasz szolgált. *Feldolgozásra került a PDB-ből egy további szekvencia is (1ag7), melynek a **c β ,c β** gerinckonformerre ad kísérleti példát. $fb < 2^\circ$: legalább 2\AA felbontású röntgenszerkezetek; $res > 2^\circ$: gyengébb felbontású röntgenszerkezetek; NMR: NMR-rel meghatározott szerkezetek; mind: a PDB SELECT-ből gyűjtött 488 Pro-Pro szekvenciárészlet mindegyike.

A számított geometriai paraméterek az azonos gerinchez tartozó eltérő gyűrűállású konformerek vonatkozó B3LYP/PCM/6-31+G(d) elméleti szintű adatainak a várható gyakorisággal súlyozott átlaga. Számított konformációs összetétel: PCM: B3LYP/PCM/6-31+G(d), az összes gyűrűállás átlagában; a: RHF/3-21G, csak a ++ gyűrűállásokra tekintettel; b: RHF/6-31+G(d), csak a ++ gyűrűállásokra tekintettel; c: B3LYP/6-311++G(d,p)/RHF/6-31+G(d), csak a ++ gyűrűállásokra tekintettel; d: B3LYP/6-311++G(d,p), csak a ++ gyűrűállásokra tekintettel.

A hat kedvelt gerinckonformer leírja prolil-prolinok jelentős többségét: 96%-ot PDB SELECT összes szekvenciája, 98%-ot a nagyfelbontású röntgenszerkezetek, 97%-ot a gyengébb minőségűek, 91%-ot az NMR-rel meghatározottak közül; 72%-ot a CSD szekvenciárészleteiből; továbbá a PCM-mel becült konformációs összetétel 97%-át, és a különféle *ab initio* és DFT számításokkal vártak 98-100%-át. Figyelembe véve azt, hogy az ideálisan elképzelhető gerinckonformerek száma 36, a PCM becslés illeszkedése a kísérleti gyakorisági adatokhoz rendkívül jó. A PCM eredmények összecsengenek a nagyfelbontású röntgendiffrakció adataival, mivel a **Pro–Pro** szekvencaegységek 87%-át sorolja mindkettő a hosszított szerkezetek, és 13%-át a β -kanyarok közé.

Hosszított szerkezetek

A poliprolin II szerkezet, $\mathbf{tE_L tE_L}$ ($\tau \approx 90^\circ$ és $d \approx 8,5 \text{ \AA}$) a leggyakoribb (ld. **E.3.11/B** és **C Ábra**) a PCM populációs becslések és a kísérleti adatok szerint is. Ezt jól szemlélteti jelenléte a kollagénben,^{62,63} a titinben,⁵⁹ SH3-hoz kötődő peptidszekvenciákban,^{56,57,58} és számos toxinban (fasciculin I,⁴⁰ toxin α ,⁴³ konotoxin ψ -PIIIE⁴⁴ és ψ -PIIIF⁴⁵). Ciklopeptidben csak egyetlen hosszított **Pro–Pro** egységet találtam, és ez szintén $\mathbf{tE_L tE_L}$ (ld. **E.3.7. Táblázat**).

A $\mathbf{cE_L tE_L}$ gerinckonformer a második legstabilabb a PCM számítások alapján (33%). Kevésbé gyakori fehérjékben (6%), de jelen van például az SH3-hoz kötődő peptidekben⁵⁶ és az uracil DNS glikozilázokban.^{52,53,54} A fehérjékben a második leggyakoribb gerinckonformer a $\mathbf{tE_L t\alpha_L}$ (20%). Tartalmazzák az uracil DNS glikozilázok, de cikloleonoripeptid D³⁵ is.

Egy további hosszított szerkezet, a $\mathbf{tE_L t\gamma_L}$, természetes előfordulása mintegy 4%, de az NMR szerkezetfínomító programok 16%-ban azonosítják. Érdekes, hogy a Pro–Pro a kardiotoxin CTX Iib NMR-rel meghatározott szerkezetében⁴² $\mathbf{tE_L t\gamma_L}$ konformert ad, míg a toxin γ 1,55 \AA felbontású röntgendiffrakciós szerkezetében⁴¹ $\mathbf{tE_L c\alpha_L}$ kanyarrá hajlik, miközben e toxinok szekvenciái egyetlen aminosav kivételével azonosak. Az elméleti módszerek között is jelentős különbség mutatkozik, hiszen a $\mathbf{tE_L t\gamma_L}$ gerinckonformernek a PCM csak 3%-os populáltságot jósol, míg oldószer nélkül 47-67% a várt gyakoriság.

A valódi nyújtott szerkezet ($\tau, \omega, \phi, \psi, \approx \pm 180^\circ$, mely a β -redőket jellemzi) a Pro–Pro szekvencaegység számára nem elérhető, hiszen a ϕ torzió *gauche*- állásban rögzül. A hosszított Pro–Pro szakaszok τ értéke negatív, ezért beépülhetnek balmenetes hélixekbe, pl. balmenetes poliprolin II hélixbe.¹²⁴ A $\mathbf{tE_L tE_L}$, $\mathbf{tE_L t\gamma_L}$, $\mathbf{tE_L t\alpha_L}$ sorrendben az egyetlen jelentősen változó torzió ψ_3 , mely csökken 148 $^\circ$ -ról kb. 70 $^\circ$ -ra, majd –30 $^\circ$ -ra. Ez a változás a **Pro–Pro** egység kinyílását okozza (a τ paraméter egyre nagyobb negatív értéket vesz fel, ld.

E.3.11. Táblázat és **E.3.2. Ábra**), ennek következtében pedig egy lehetséges hélixben rendre kevesebb aminosav egység szükséges egy körbeforduláshoz, ill. egy β -redő enyhébb hajlással rendelkezik. A $\tau_{E_1}f_{E_1}$, $\tau_{E_1}f_{\gamma_1}$, $\tau_{E_1}\alpha_1$ hosszított szerkezetek kinyílásához nem társul azonos sorrendű megnyúlás. A poliprolin II szakasz, $\tau_{E_1}f_{E_1}$, a leghosszabb ($d \approx 8,5$ Å), és a γ -kanyart tartalmazó $\tau_{E_1}f_{\gamma_1}$ a legrövidebb ($d \approx 8$ Å).

A PDB SELECT hosszított szerkezeteiben τ ill. d értékei -37° és -207° , ill. 7 Å és 10 Å között változnak (ld. **E.3.2. Ábra**). Ezek a nagy intervallumok arra utalnak, hogy a hosszított **Pro–Pro** szekvenciaegységek a polipeptidekben változatos térszerkezeti elvárásokat képesek kielégíteni. Egyazon konformáción belül, a τ és a d paraméter rendszerint erős korrelációt mutat (ld. **E.3.12. Táblázat**) éppen úgy, mint amikor kinyitunk egy könyvet (növeljük τ -t) a fedő borító jobb felső sarka eltávolodik a hátsó borító alsó sarkától (növekszik d). A lineáris korrelációs paraméterek konformerenként különbözőek, tehát a τ - d párok finom szabályozása elérhető egy megfelelő gerincszerkezet kiválasztásával, ami persze energiát igényelhet.

β -kanyarok

Amint az a CSD-ből gyűjtött peptidok kis száma és ciklikus alkata miatt várható, konformációs adataik elemzése jobbra csak β -kanyarok meglétét és szerkezeti tulajdonságait tudják alátámasztani. A $\tau_{E_1}\alpha_1$, $\tau_{E_1}c_{E_1}$, $c_{E_1}c_{E_1}$, $\tau\alpha_1\alpha_1$ konformerek populáltsága rendre 67, 18, 9 és 3% a CSD-ben.

A β -kanyar típusú **Pro–Pro** gerincszerkezetek közül a $\tau_{E_1}\alpha_1$ a leggyakoribb ($\tau \approx 30^\circ$ és $d \approx 5,5$ Å) a módszertől függetlenül (PDB SELECT: 4 %, CSD: 67 %, PCM: 9 %, gázfázisú számítások: 33-47 %). Ezt a VIa típusú β -kanyart foglalja magába a ciklolinopeptid A,³³ a cikloleonorin,³⁴ és a toxin γ .⁴¹

A $\tau_{E_1}c_{E_1}$ konformer a PDB SELECT ill. a CSD szekvenciák 3 ill. 18 %-át teszi ki, miközben olyan magas energiával rendelkezik az RHF/3-21G elméleti szinten ($12,13$ kcal.mol⁻¹), hogy a további optimalásokból ki is maradt.

A $\tau\alpha_1\alpha_1$ gerinc PCM-mel becsült gyakorisága 3 %, ami meg is egyezik a PDB SELECT vonatkozó adatával. Környezetétől függően beépülhet 3_{10} -hélixbe, ill. I vagy III típusú β -kanyarnak tekinthető, amint azt a trombin⁵¹ példázza.

E.3.12. Táblázat A leggyakoribb prolin-prolin gerinckonformerek τ képzetes torziósöze és d atomtávolsága közötti korreláció.

Gerinc- konformer	n	ϕ_2	ψ_2	ϕ_3	ψ_3	τ	d	m 10^{-2}	b	R^2
hosszított										
$t\epsilon_1t\alpha_L$	109	-65 ±11	145 ±17	-57 ±12	-35 ±16	-120 ±20	8,4 ±0,4	-1,59	6,52	0,6487
$t\epsilon_1t\gamma_L$	21	-69 ±11	152 ±22	-71 ±10	72 ±18	-98 ±26	8,1 ±0,5	-1,70	6,44	0,8127
$t\epsilon_1t\epsilon_L$	262	-67 ±9	153 ±14	-65 ±10	148 ±17	-92 ±19	8,8 ±0,5	-2,19	6,75	0,6424
$c\epsilon_1t\epsilon_L$	29	-75 ±12	164 ±19	-70 ±11	153 ±16	-120 ±21	8,0 ±0,5	-1,52	6,18	0,4854
β -kanyar										
$t\alpha_Lt\alpha_L$	16	-51 ±8	-43 ±10	-62 ±8	-30 ±8	55 ±11	5,6 ±0,5	3,80	3,50	0,7677
$t\epsilon_1c\alpha_L$	20	-61 ±10	151 ±38	-83 ±9	4 ±20	30 ±12	5,9 ±0,4	0,67	5,68	0,0405
$t\epsilon_1c\epsilon_L$	13	-62 ±8	157 ±11	-76 ±7	144 ±39	71 ±21	7,0 ±0,8	3,59	4,48	0,8035

n : az adott gerinckonformer száma a PDB SELECT-ből gyűjtött 488 **Pro-Pro** fehérjeszakaszban. $\phi_2, \psi_2, \phi_3, \psi_3, \tau, d$: kísérleti adatok átlaga szórással. m, b és R^2 : az adott gerinckonformerhez sorolt n τ - d pár korrelációját jellemző $d = m \tau + b$ lineáris regressziós egyenlet paraméterei.

Az előzetes várakozásoknak megfelelően a két *cis* peptidkötéssel rendelkező $\mathbf{c\epsilon_1c\epsilon_L}$, a poliprolin I hélix konformációs építőeleme,¹²⁵ magasabb energiájú ($\Delta E=1,87 \text{ kcal.mol}^{-1}$, ld. **E.3.4. Táblázat**), és a PDB SELECT 488 szekvenciája között sem szerepel. Tartalmazza azonban a $\mathbf{c\epsilon_1c\epsilon_L}$ gerinckonformert a konotoxin GS.⁴⁶ A szerkezeti rokonságot mutató konotoxin TVIIA-ban⁴⁷ pedig a $\mathbf{c\epsilon_1c\alpha_L}$ konformer szerepel. A $\mathbf{c\epsilon_1c\epsilon_L}$ ill. a $\mathbf{c\epsilon_1c\alpha_L}$ konformer egyaránt kielégíti a β -kanyarokkal szemben támasztott kritériumokat ($\tau \approx 60^\circ$ ill. 20° , $d \approx 5 \text{ \AA}$ ill. $4,5 \text{ \AA}$), mindazonáltal a β -kanyarok között ez ideig még nem osztályozták.

Gyűrűállás

Egy-egy gerinckonformer gyűrűállásainak sorrendje csak ritkán egyezik az elméleti és a kísérleti megközelítésben, pl. a $\mathbf{t\epsilon_L-c\alpha_L+}$, a $\mathbf{t\epsilon_L+c\alpha_L+}$, a $\mathbf{t\epsilon_L-c\alpha_L-}$ és a $\mathbf{t\epsilon_L+c\alpha_L-}$ konformerek emelkedő PCM energiája jól összeegyeztethető csökkenő gyakoriságukkal a fehérjékben, továbbá $\mathbf{t\epsilon_L-c\alpha_L+}$ és $\mathbf{t\epsilon_L+c\alpha_L+}$ vonatkozó adataival a CSD peptidok között.

A kollagénben túlnyomórészt a $\mathbf{t\epsilon_L+t\epsilon_L-}$ konformer szerepel, pedig nem ez a leggyakoribb a PDB SELECT $\mathbf{t\epsilon_L/t\epsilon_L}$ gyűrűkombinációi közül, és a PCM számítások sem ehhez rendelik a legalacsonyabb energiát. Megállapítható tehát, hogy a megvalósuló gyűrűállás nem a prolin-prolin konformer sajátos tulajdonsága, hanem a kollagén triplahelikális szerkezetének következménye.

A + és – (down és up) gyűrűállásokhoz tartozó ϕ értékek jellegzetes eltérése ($-75^\circ \pm -3^\circ$ és $-60^\circ \pm -2^\circ$, ld. **E.3.11/G-H Ábra**) már korábban is ismert volt.¹²⁶ A CSD-ből gyűjtött 152 prolin adatai alapján, melyekben az *A* pszeudorotációs amplitúdó legalább 30° , a ϕ paraméter eltérése rendre 10° , 12° , 14° , 12° a $\mathbf{t\epsilon_L}$, $\mathbf{t\alpha_L}$, $\mathbf{c\epsilon_L}$, $\mathbf{c\alpha_L}$, gerinckonformer + és – gyűrűállásai között.² A pirrolidinyűrű nem merev, hanem képes olyan gerinckonformációs változásokat segíteni, mint a *cis-transz* izomerizáció.⁶ Feltehető, hogy a gyűrű redőzöttsége a polipeptid gerinckonformációs igényeitől függ, nevezetesen a + gyűrű a ϕ torzió nagyobb negatív értékeit kedveli, míg a – gyűrű a kevésbé negatívakat.

Vicinális diszulfidhidas ciszteinil-cisztein

A 4,5-ditiaheptano-7-laktám gyűrűmodell ab initio és DFT számításai

Az RHF/3-21G* elméleti szinten végrehajtott kimerítő *ab initio* vizsgálattal a 4,5-ditiaheptano-7-laktám (ld. **B.3/B Ábra**) 16 különböző konformerét találtam meg (ld. **5.1. Táblázat** és **E.4.1. Ábra**). Mindegyik szerkezet stabil a B3LYP/6-31+G(d) és a B3LYP/6-311++G(d,p) elméleti szinten is. Nyolc gyűrűben *cisz* peptidkötés található, míg a másik nyolcban *transz*. Mindegyik szerkezetnek van egy tükörképe, mely az eredetivel nem hozható fedésbe. A tükörképi konformerpárok energiái egyenlők (ti. $\Delta E^{c1} = \Delta E^{c5}$, stb.), torziószögek pedig abszolút értékben egyenlők, de előjelben különböznek ($\chi_0^{c1} = -\chi_0^{c5}$, stb.).

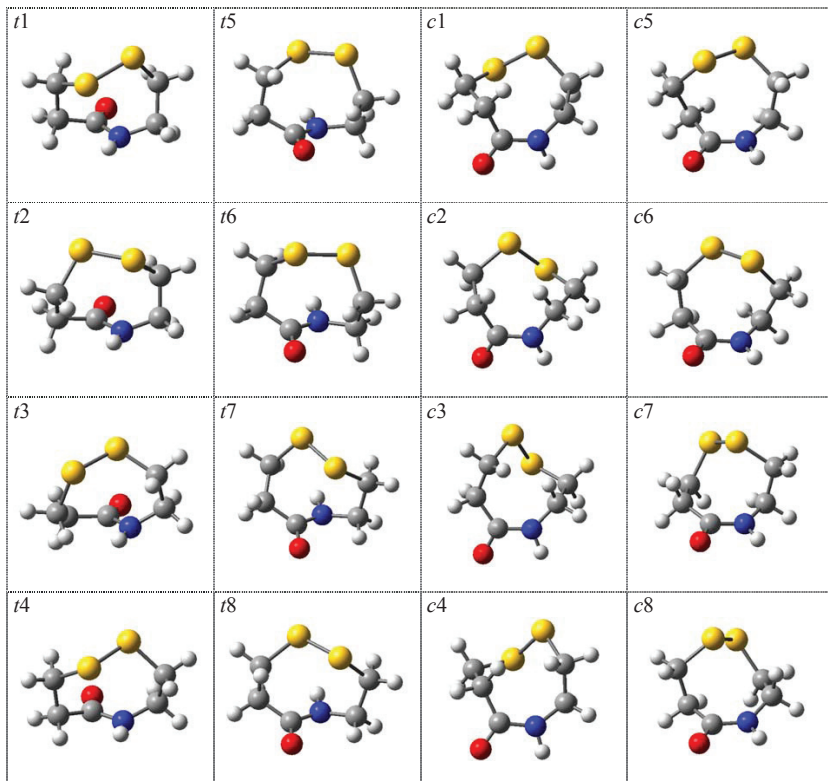
A torziószögeket az RHF/3-21G* elméleti szint is rendszerint megbízhatóan adja meg, csak néhány esetben tér el 10^0 -nál erősebben a DFT módszerek vonatkozó adataitól. A relatív energia legjelentősebb eltérése az *ab initio* és a DFT megközelítés között 2 kcal.mol⁻¹. A két DFT szint adatai jól korrelálnak, bizonyítva, hogy a B3LYP/6-311+G(d) elméleti szintű számítások elégségesnek. ΔE legnagyobb eltérése 0,24 kcal.mol⁻¹, míg a torziószögeké 1,6°. Ha csak az energia szempontjából kedvező konformereket vesszük figyelembe, az eltérések 0,14 kcal.mol⁻¹ ill. 0,8° értékre csökkennek.

A globális minimum a *c2* és *c6* tükörképi pár (utóbbi jelölhető *c2'*-nek is, hangsúlyozva, hogy *c6* tükörképe *c2*-nek) mindhárom elméleti szinten. A konformereket két csoportra osztottam a globális minimumra vonatkoztatott ΔE relatív energiáik alapján. A kedvező energiájú konformereknél (*c1*, *c2*, *c5*, *c6*, *t2*, *t4*, *t6* és *t8*) ΔE nem nagyobb, mint 2,5 kcal.mol⁻¹. A kedvezőtlen energiájú konformereknél (*c3*, *c4*, *c7*, *c8*, *t1*, *t3*, *t5* és *t7*) pedig ΔE , 5 és 10 kcal.mol⁻¹ közötti érték.

E.4.1. Táblázat A 4,5-ditiaheptano-7-laktám egyes elméleti szinteken optimált konformerei.

RHF/3-21Gd*	<i>E</i>	ΔE	χ_0	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	ω
c5 (c1')	-1113,438088	0,01	84,95	-80,19	76,75	-95,68	49,61	57,95	-107,96	0,93
c6 (c2')	-1113,438111	0,00	96,71	-52,95	-52,68	96,44	-75,83	83,47	-94,41	12,13
c7 (c3')	-1113,425905	7,66	78,43	-58,16	-55,05	78,21	24,98	-83,52	12,45	9,10
c8 (c4')	-1113,423068	9,44	-45,31	91,07	-19,31	-79,83	52,21	64,64	-94,07	25,14
t5 (t1')	-1113,428615	5,96	-45,96	-57,80	32,86	66,58	-97,69	61,97	-81,24	142,55
t6 (t2')	-1113,437700	0,26	-59,10	-48,94	91,52	-88,40	93,91	-45,95	-56,37	149,58
t7 (t3')	-1113,425481	7,93	-89,38	48,03	-90,75	81,72	20,03	-57,27	-25,51	140,11
t8 (t4')	-1113,437359	0,47	-100,58	56,06	-78,33	102,48	-86,09	64,23	-88,28	146,72
B3LYP/6-31+G(d)	<i>E</i>	ΔE	χ_0	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	ω
c5 (c1')	-1122,335805	0,25	81,79	-81,29	76,65	-92,79	48,96	56,90	-108,14	4,43
c6 (c2')	-1122,336202	0,00	91,20	-54,38	-51,10	94,29	-73,67	81,49	-97,14	18,97
c7 (c3')	-1122,326197	6,28	71,99	-58,54	-54,11	76,94	24,02	-78,03	1,88	21,66
c8 (c4')	-1122,324337	7,44	-52,55	92,23	-17,55	-79,16	51,28	64,28	-96,01	31,08
t5 (t1')	-1122,326353	6,18	-45,66	-57,37	33,26	64,66	-95,25	61,47	-80,95	139,89
t6 (t2')	-1122,332390	2,39	-60,61	-47,87	90,67	-86,58	91,20	-46,19	-53,73	145,09
t7 (t3')	-1122,323703	7,84	-87,79	48,54	-89,43	81,54	17,41	-57,24	-24,36	136,92
t8 (t4')	-1122,334458	1,09	-98,12	58,12	-78,33	99,08	-83,94	64,65	-90,20	142,73
B3LYP/6-311+G(d,p)	<i>E</i>	ΔE	χ_0	χ_1	χ_2	χ_3	χ_4	χ_5	χ_6	ω
c5 (c1')	-1122,472673	0,28	81,65	-81,75	77,13	-93,10	48,94	57,40	-108,73	4,43
c6 (c2')	-1122,473118	0,00	91,70	-54,57	-51,24	94,57	-74,28	82,22	-97,01	18,67
c7 (c3')	-1122,463055	6,32	72,02	-58,96	-53,83	76,93	24,09	-78,91	2,65	21,69
c8 (c4')	-1122,461283	7,43	-53,07	92,83	-17,27	-79,20	50,82	65,05	-95,95	30,55
t5 (t1')	-1122,463366	6,12	-46,57	-58,27	34,88	63,53	-95,91	62,34	-80,16	139,78
t6 (t2')	-1122,469531	2,25	-61,15	-48,14	90,68	-86,42	91,58	-46,84	-52,95	145,28
t7 (t3')	-1122,469996	7,61	-87,93	48,67	-89,48	81,27	18,11	-58,06	-23,91	137,01
t8 (t4')	-1122,471517	1,00	-98,76	58,21	-78,38	99,34	-84,33	64,99	-89,70	142,89

A tükörképi konformerpároknak mindig csak egyike szerepel. A c1', c2, c3, c4, t1, t2, t3 és t4 konformerek rendre tükörképei a c5, c6, c7, c8, t5, t6, t7 és t8 konformereknek. Ezt a kapcsolatot fejezik ki a zárójelbe írt alternatív kódok. A tükörképi párok torziószelei abszolút értékben megegyeznek, de előjelben különböznek. *E*: energia Hartree egységben; ΔE : relatív energia kcal.mol⁻¹ egységben, torziós szögek fokban.



E.4.1. Ábra A 4,5-ditiaheptano-7-laktám egyes konformerei. (**B.3/B Ábrának** megfelelő tájolásban.) A *c1*, *c2*, *c3*, *c4*, *t1*, *t2*, *t3* és *t4* konformerek rendre tükörképei a *c5*, *c6*, *c7*, *c8*, *t5*, *t6*, *t7* és *t8* konformereknek. A tükörképi kapcsolat hangsúlyozását elvetve, a konformerpárok kissé eltérő nézetből látszanak.

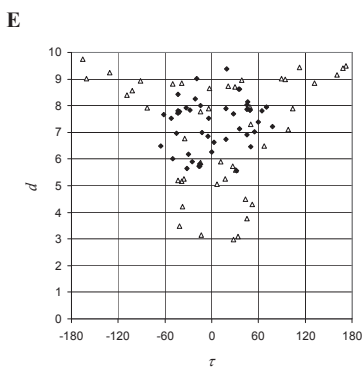
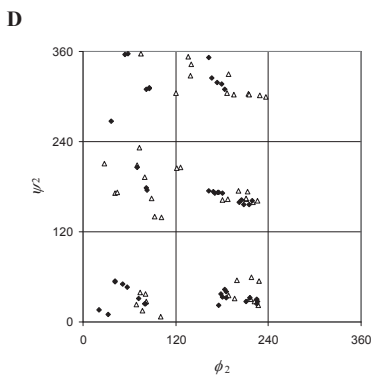
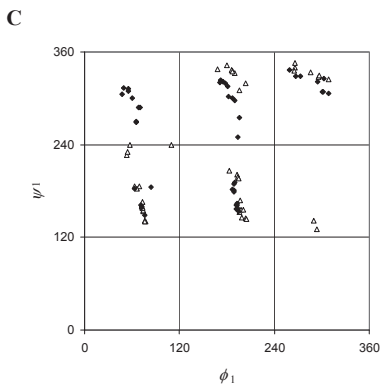
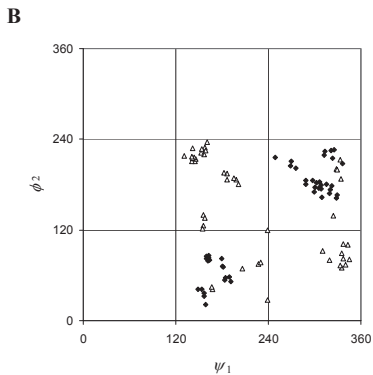
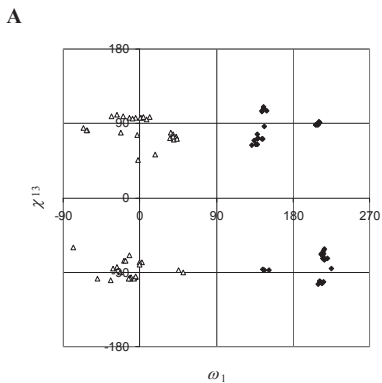
Miközben a lineáris peptidok előnyben részesítik a *transz* peptidkötést a *cisz* amiddel szemben, ez a nyolctagú laktámgyűrű könnyebben veszi fel a *cisz* kötést és feszültebb a *transz* kötéssel. Az ω paraméter átlagos eltérése az ideális 0° -tól ill. 180° -tól 12 - 19° a *cisz* amid és 35 - 39° a *transz* kötés esetében.

Gyűrű kialakításához *transz* kötés esetében az amidsík valamelyik oldalára kell kerülnie mindkét $-\text{CH}_2-\text{CH}_2-\text{S}-$ atomcsoportnak. A t_1 , t_2 , t_3 és t_4 konformerben ez pozitív χ_0 , -145° körüli ω és pozitív χ_6 révén valósul meg. Tehát ω “vezérkoordinátának” tekinthető, mely két minimummal rendelkezik (ti. -145° vagy $+145^\circ$), míg χ_0 és χ_6 függ ω -tól. χ_1 és χ_5 mutatkozik a másik két “vezérkoordinátának”. Egyik sem lehet *anti* állású, mivel egy ilyen méretű gyűrű csak egy torziószög, nevezetesen ω esetében látszik megengedni az *anti* konformáció felvételét. Így χ_1 és χ_5 is *gauche+* vagy *gauche-* lehet, és meg is valósul e lehetőségeknek mind a négy kombinációja. A maradék három torziószög (χ_2 , χ_3 és χ_4) mindegyike ω , χ_1 és χ_5 konformációjától függ. A négy energetikailag kedvező *transz* típusú konformerben (t_2 , t_4 , t_6 és t_8) a χ_1 , χ_2 , χ_3 , χ_4 és χ_5 torziószögek előjele váltakozik; $\chi_1 \approx \chi_5$ és $\chi_2 \approx \chi_4$, tehát egy kétfogású szimmetriatengely azonosítható a molekulákban, ha eltekintünk a nitrogén és a karbonil szén különbségétől.

A torziószögek előjelének váltakozását a *cisz* gyűrűk is kedvelik. Az energetikailag kedvező konformerekben csak egy-egy kötéspár található, ahol a szomszédos torziószögek helicitása azonos (ti. $\chi_4 \approx \chi_5$ c1-ben és c5-ben, míg $\chi_1 \approx \chi_2$ in c2-ben és c6-ban).

A HCO-ox-[L-Cys-L-Cys]-NH₂ peptidmodellre vonatkozó ab initio és DFT számítások

A HCO-ox-[L-Cys-L-Cys]-NH₂ peptidmodell $16 \times 3 \times 3 = 144$ ideális konformeréből (16 gyűrűkonformer, és 3-3 minimum a ϕ_1 és ψ_2 torzió mentén) 91-et találtam meg az RHF/3-21G* elméleti szinten (ld. **B.3/B Ábra**, publikáció⁴ **kiegészítő információi és E.4.2. Ábra**). A peptidmodellen belül ugyanazok az energetikailag előnyös gyűrűkonformerek, mint a 4,5-ditiaheptano-7-laktámnál, kivéve, hogy a c1 és a c2 kedvezőtlené vált, míg t5 kedvezővé. A kedvelt és kedvezőtlen konformerek szétválása nem olyan szembeötlő a peptidnél, mint a laktámnál, mivel a gyűrű energetikai szempontjai a peptidgerinc teljesen független preferenciáival kombinálódnak (ez utóbbi szempont elemzését HCO-Ala-Ala-NH₂ peptidmodellen végezték el).^{127,128}



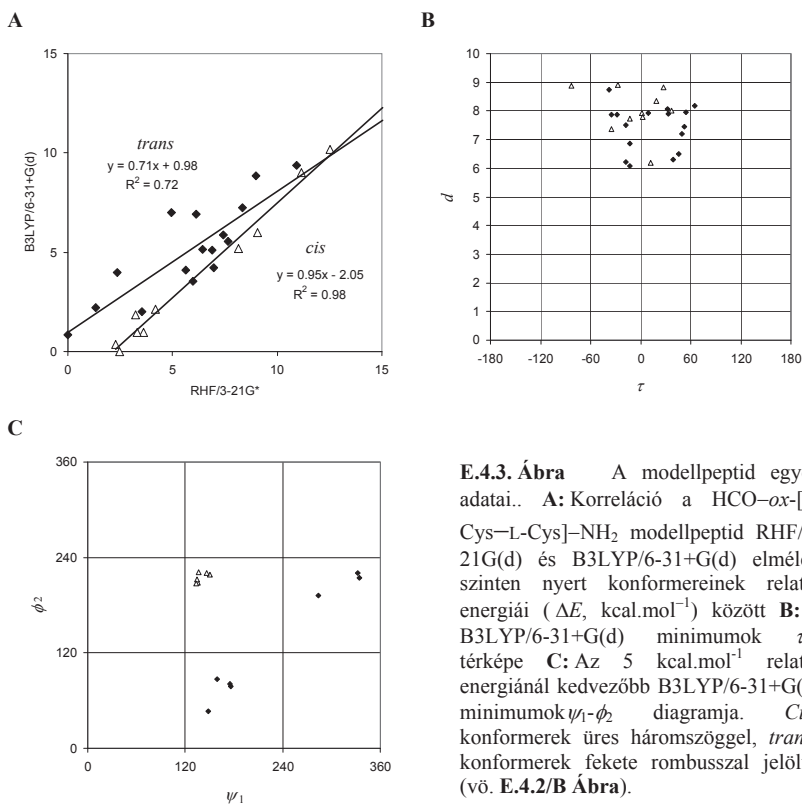
E.4.2. Ábra A HCO-ox-[L-Cys-L-Cys]-NH₂ modellpeptid B3LYP/6-31+G(d) elméleti szinten azonosított konformereinek geometriája. *Cis*z konformerek üres háromszöggel, *transz* konformerek fekete rombuszsal jelölve. **A:** ω_1 - χ_{13} térkép; **B:** ψ_1 - ϕ_2 térkép; **C:** ϕ_1 - ψ_1 térkép; **D:** ϕ_2 - ψ_2 térkép; **E:** τ - d térkép.

A legtöbb esetben, amikor az ideális minimum optimalása nem eredményezett új konformert, a kezdeti geometriának lényegében csak a gerincortozói változtak meg, miközben a gyűrűkonformer megmaradt. Mindössze egy *transz* típusú gyűrűváltozást észleltem (nevezetesen $t7 \rightarrow t8$, ti. egy ideális peptidkonformer, mely $t7$ gyűrűt tartalmaz, átkonvergált egy olyan minimummá, melynek $t8$ a gyűrűje). Mivel *transz* peptidkötés mellett a gyűrű meglehetősen feszített, a gyűrű torziószögei *transz* peptidkonformerekben nem igen térnek el a 4,5-ditiaheptano-7-laktám vonatkozó adataitól. A *cisz* gyűrűk azonban rugalmasabbnak bizonyulnak, gyűrűortozóiuk jelentősen megváltoznak a HCONH– és –CONH2 szubsztituensek hatására. Optimalás során négy nyilvánvaló esetet találtam *cisz* típusú gyűrűváltásra (két-két eset $c3 \rightarrow c8$ és $c4 \rightarrow c6$ változásra).

Az ω_1 torziószög -78° és $+51^\circ$ között változik a *cisz* kötésekre; az ideális 0° értéktől való eltérése nem nagyobb 13° -nál az alacsony energiájú konformerek esetében (ld. **E.4.2/A Ábra**). Az ideális *transz* értéktől, 180° -tól, mindig legalább 26° az eltérés. A *transz* ω_1 értékek két csoportra oszlanak: 142° és -147° köré. A ψ_1 torziószög *gauche*– vagy *anti*, míg a ϕ_2 paraméter *gauche*+ vagy *anti*; mind a négy ψ_1/ϕ_2 kombinációra van példa a *cisz* gyűrűs, de csak *anti/gauche*+ és *gauche*–/*anti* kombinációra a *transz* gyűrűs konformerekben (ld. **E.4.2/B Ábra**).

A globális minimumnál (**t6_β1_γD**) legfeljebb 7 kcal.mol^{-1} -al nagyobb relatív energiájú RHF/3-21G* konformereket és néhány kiválasztott magasabb energiájú szerkezetet tovább optimaltam B3LYP/6-31+G(d) elméleti szinten (ld. publikáció⁴ **kiegészítő információi**). A két eltérő módszerrel nyert relatív energiák jól korrelálnak *cisz* konformerekre és gyengén *transz* szerkezetekre (ld. **E.4.3/A Ábra**). A DFT módszerrel kapott globális minimum **c6_β1_β1** annak ellenére, hogy nem tartalmaz szerkezetstabilizáló hidrogénkötést, miközben az RHF globális minimumot, **t6_β1_γD**-t egy a γ -kanyar rögzíti (hidrogénkötés egy amid hidrogén és a megelőző peptidkötés karbonil oxigénje között egy héttagú pszeudogyűrűt képezve, vö. **M.8/B Ábra**).

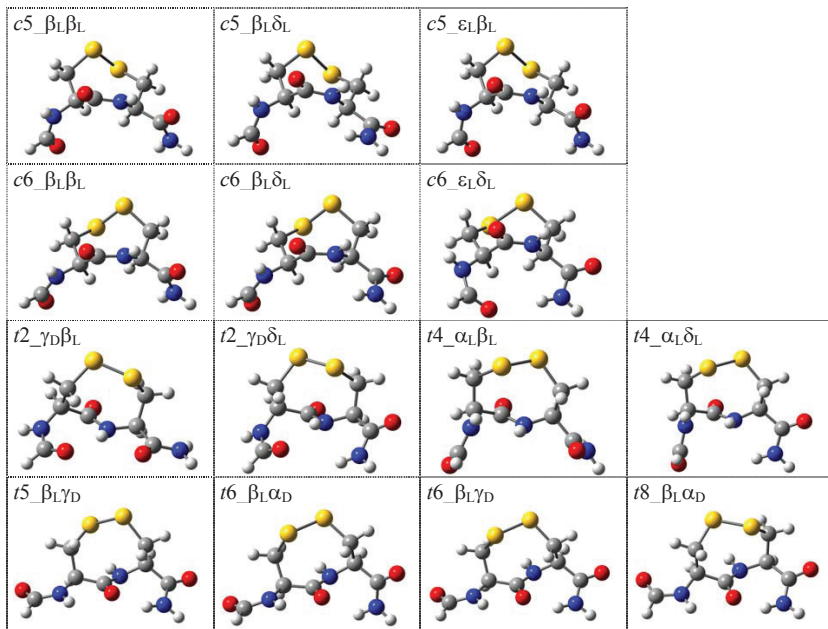
Azokat a konformereket, melyek legfeljebb 5 kcal.mol⁻¹ relatív energiával rendelkeznek akár az RHF/3-21G*, akár a B3LYP/6-31+G(d) elméleti szinten kedvező energiájú minimumnak tekintetem (ld. E.4.2. Táblázat, E.4.3/B és E.4.4. Ábra) és tovább vizsgáltam (ld. Összevetés). A $-90^\circ \leq \tau \leq 90^\circ$ kritérium alapján minden alacsony energiájú minimum β -kanyarnak minősül. A $d \leq 7$ Å kritérium azonban ritkán teljesül. A β -kanyarak hagyományos osztályozása alapján találtam I, VIa, VIb és VIII típusút. A II típusú β -kanyar ($\epsilon_1\gamma_D$ vagy $\epsilon_1\alpha_D$) szerkezetet közelíti a következő konformerek: $t5_beta_1\gamma_D$, $t6_beta_1\alpha_D$, $t6_beta_1\gamma_D$ és $t8_beta_1\alpha_D$.



E.4.2. Táblázat A HCO- α -[L-Cys-L-Cys]-NH₂ modellpeptid kedvező energiájú konformerei B3LYP/6-31+G(d) elméleti szinten.

Konformer β -kanyar	ΔE	$\Delta E'$	ϕ	ψ_1	ω_1	ϕ_2	ψ_2	χ_{11}	χ_{13}	χ_{21}	τ	d	
c5_ $\beta_1\delta_1$	n.o.	0,98	-122,40	135,25	-2,99	-151,88	149,77	173,93	-94,16	-80,45	1,10	7,94	
c5_ $\beta_1\delta_1$	Vlb	2,12	-127,64	134,93	3,62	-147,57	27,67	169,16	-94,25	-82,27	-35,23	7,38	
c5_ $\epsilon_1\beta_1$	n.o.	0,98	-114,06	133,66	-2,44	-152,49	148,84	175,13	-94,10	-80,34	2,25	7,79	
c6_ $\beta_1\beta_1$	n.o.	0,00	-154,28	150,16	8,78	-141,47	141,14	-160,11	95,42	-54,04	18,69	8,35	
c6_ $\beta_1\delta_1$	Vlb	0,38	-155,45	146,66	16,63	-139,17	18,97	-162,87	96,42	-55,29	-13,09	7,73	
c6_ $\epsilon_1\delta_1$	Vla	1,84	-80,67	136,58	14,96	-138,14	65,68	-155,83	94,15	-51,75	11,81	6,19	
t2_ $\gamma\beta_1$	n.o.	7,00	6,14	69,92	-73,38	-147,42	-176,39	162,27	177,37	87,83	46,64	-28,38	7,88
t2_ $\gamma\delta_1$	n.o.	3,97	3,11	68,04	-75,95	-145,89	-167,83	23,55	175,09	89,89	43,64	-13,23	6,09
t4_ $\alpha_1\beta_1$	VIII	3,54	2,68	-89,11	-25,95	-142,67	-145,87	113,83	58,84	-97,48	-60,37	52,02	7,45
t4_ $\alpha_1\delta_1$	I	2,02	1,16	-77,81	-27,28	-141,29	-139,07	19,69	58,28	-96,87	-62,34	38,80	6,30
t5_ $\beta_1\gamma\beta$	n.o.	4,10	3,25	-157,00	159,61	138,46	86,72	-41,74	178,21	70,65	-59,98	9,44	7,93
t6_ $\beta_1\alpha\beta$	n.o.	2,20	1,34	-159,07	175,82	144,06	77,45	22,48	75,88	-86,66	-53,20	54,32	7,95
t6_ $\beta_1\gamma\beta$	n.o.	0,86	0,00	-159,51	175,12	143,06	81,39	-39,60	76,38	-86,27	-55,76	33,40	7,89
t8_ $\beta_1\alpha\beta$	n.o.	4,23	3,38	-156,90	148,45	141,26	46,41	7,78	-178,04	104,08	50,32	-34,67	7,88

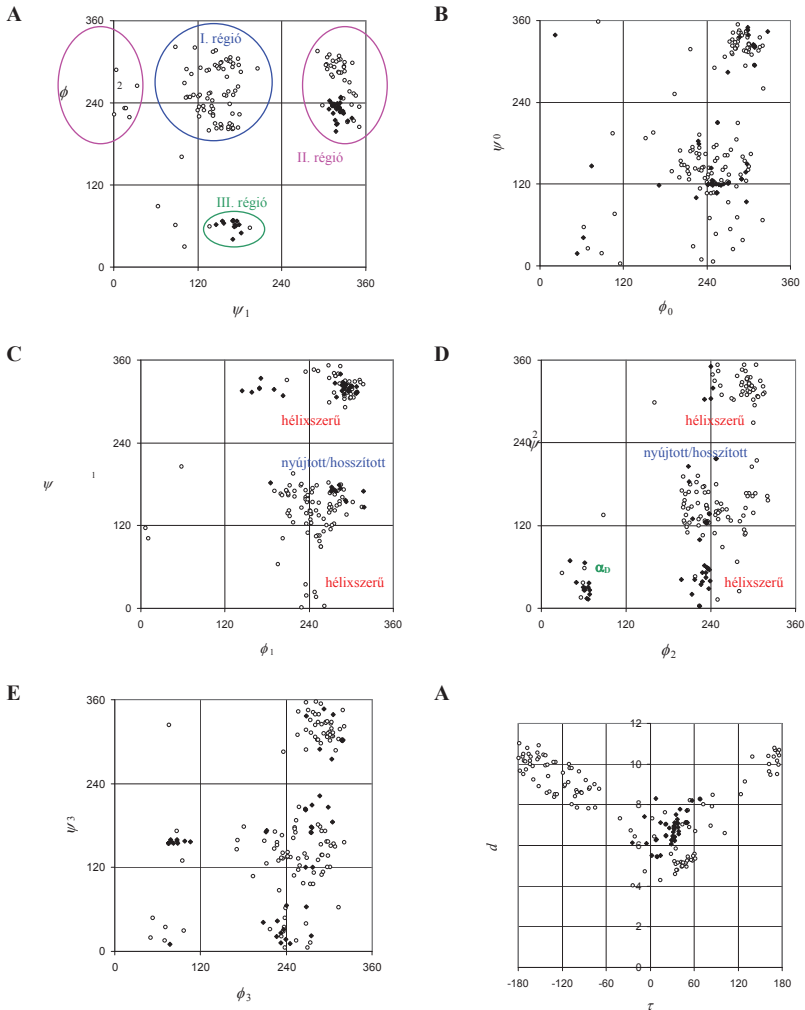
A konformer kódjának első két karaktere a laktámgyűrűt azonosítja, az utolsó négy karakter pedig a gerinckonformációt. β -kanyar típusa: ld. **M.1. Táblázat**, n.o.: ez ideig nem osztályozott. ΔE : relatív energia kcal.mol⁻¹ egységben a globális minimumhoz (c6_ $\beta_1\beta_1$, $E_h = -1459,743551$ Hartree) viszonyítva. $\Delta E'$: relatív energia kcal.mol⁻¹ a t6_ $\beta_1\gamma\beta$ minimumhoz ($E_h = -1459,742186$ Hartree) viszonyítva. Minden torzioszög fokban. d mértékegysége Å.



E.4.4. Ábra Az HCO-ox-[L-Cys-L-Cys]-NH₂ modellpeptid néhány B3LYP/6-31+G(d) elméleti szinten optimalt konformere.

Az Xxx-Cys-Cys-Yyy és Xxx-ox-[Cys-Cys]-Yyy peptidszekvenciák adatbázisokból nyert térszerkezeti adatainak elemzése

A PDB SELECT-ből kiszűrt 120 Xxx-Cys-Cys-Yyy típusú szekvenciaegységen konformációs elemzést végeztem. Mindössze 3 rendelkezett vicinális diszulfidhiddal. E nagyon ritka előfordulás miatt az Xxx-ox-[Cys-Cys]-Yyy szekvenciaegységek elemzését a teljes PDB-ből nyert adatsoron végeztem el. A két halmaz jellemző torziósögeit az **E.4.5. Ábra** szemlélteti.



E.4.5 Ábra Adatbázisból nyert szekvenciárészeklet. Xxx–Cys–Cys–Yyy a PDB SELECT-ből (üres körök) és Xxx- α -[Cys–Cys]–Yyy a PDB-ből (fekete rombuszok). Torziószögek fok egységben. **A:** ψ_1 - ϕ_2 diagram; **B:** Xxx Ramachandran-térkép, ϕ_0 - ψ_0 ; **C:** Az első cisztein Ramachandran-térképe, ϕ_1 - ψ_1 ; **D:** A második cisztein Ramachandran-térképe, ϕ_2 - ψ_2 ; **E:** Yyy Ramachandran-térkép, ϕ_3 - ψ_3 ; **F:** τ - d diagram.

A PDB SELECT Xxx–Cys–Cys–Yyy típusú szekvenciáinak 45%-ban mindkét cisztein gerince nyújtott/hosszított (többnyire β_L vagy ϵ_L , ld. **E.4.5/C-D Ábra**). A szekvenciák negyedében mindkét cisztein hélixszerű, 18 %-ában pedig mind a négy aminosav közel helikális, tehát egy jobbmertes α -hélix alakul ki. 12 %-ban egy nyújtott ciszteint egy helikális követ, és ebben az esetben rendszerint Xxx nyújtott, míg Yyy változékony konformációjú. Egy másik 12 %-ban egy hélixszerű ciszteint nyújtott követ, mely esetben sem Xxx sem Yyy gerincszerkezete nem mutat jellegzetes függést, leszámítva, hogy az amúgy is ritka konformációkat (α_D , ϵ_D , γ_D vagy δ_D) egyáltalán nem veszik fel. *Cisz* peptidkötést a két cisztein között nem találtam.

Mivel a vicinális diszulfidhíd korlátozza a ψ_1 és a ϕ_2 torziók konformációs szabadságát, érdemes ezeket a szögeket is egymás függvényeként ábrázolni (ld. **E.4.5/A Ábra**). Két jól benépesített területet határoztam meg.. Az I. régió szekvenciáiban az első cisztein nyújtott/hosszított, míg a második vagy nyújtott/hosszított vagy hélixszerű. Nagyon fontos megjegyezni, hogy a diszulfidhíddal rendelkező szekvenciák sohasem kerülnek az I. régióba. A II. régió szekvenciáiban az első cisztein hélixszerű, míg a második vagy hélixszerű vagy nyújtott/hosszított. Ez utóbbihoz sorolható több diszulfidhidas szekvencia is. A III. régiót azért jelöltem meg, mert tartalmaz néhány diszulfidhíddal rendelkező szekvenciát, nevezetesen valamelyik $t6$ vagy $t8$ gyűrűset $\epsilon_L\alpha_D$ gerincszerkezettel. A 120 Xxx–Cys–Cys–Yyy típusú szekvencia közül mindössze kettő került a III. régióba, ezekben a ciszteinek gerinckonformációja $\beta_L\alpha_D$.

A PDB-ből 48 Xxx– α -[Cys–Cys]–Yyy szekvenciaegységet sikerült gyűjteni, melyek 31 különböző fehérjéhez tartoznak (ld. **E.4.3. Táblázat**). Ezekben a ciszteinek egy esetben sem kapcsolódnak *cisz* peptidkötéssel egymáshoz. $t2$, $t4$, $t6$ és $t8$ gyűrűt rendre 2, 32, 11 és 2 szekvenciaegység tartalmazza. Egyetlen szekvencia volt, melyhez nem sikerült gyűrűt rendelnem, ebben a χ_{13} torziószög 0° körüli. Xxx rendszerint kicsi (glicin, alanin) vagy éppen kicsi hidrophil oldallánccal rendelkező aminosav (szerin, treonin). Yyy általában aszparaginsav, aszparagin, arginin vagy prolin.

E.4.3. Táblázat A PDB-ből gyűjtött Xxx-ox-[Cys-Cys]-Yyy szekvenciárszletek.

PDB	lác	No	Xxx	ox-[Cys-Cys]	Yyy	módszer	Ref.	fb.	β -kanyar	Megnevezés
1qlw	A, B	70	Gly ϵ_L	$t6_{\epsilon_1\alpha_D}$	Leu β_L	Röntgen	67	1,1	II	<i>Alcaligenes</i> észteráz 713
1gyd	B	240	Leu γ_L	$t6_{\epsilon_1\alpha_D}$	Arg δ_L	Röntgen	86	2,05	} II	α -L-arabinanáz
1gye	B	240	Leu γ_L	$t6_{\epsilon_1\alpha_D}$	Arg δ_L	Röntgen	86	2,5		
1gyh	A, B, C, D, F	240	Leu ϵ_L	$t6_{\epsilon_1\alpha_D}$	Arg γ_L	Röntgen	86	1,89	} II	<i>Anguilla anguilla</i> agglutinin
1gyh	E	240	Leu γ_L	$t6_{\epsilon_1\alpha_D}$	Arg δ_L	Röntgen	86	1,89		
1k12	A	81	Asp α_L	$t6_{\beta_1\alpha_D}$	Gly α_L	Röntgen	81	1,9	n.o.	} II
1idh	B	191	Thr α_L	$t6_{\epsilon_1\alpha_D}$	Pro α_L	NMR	130		n.o.	
1idg	B	191	Thr α_L	$t6_{\epsilon_1\alpha_D}$	Pro α_L	NMR	130		n.o.	} II
114w	B	191	Thr ϵ_L	$t2_{\delta_2\beta_L}$	Pro α_L	NMR	133		n.o.	
11jz	B	191	Thr ϵ_L	$t2_{\delta_2\beta_L}$	Pro α_L	NMR	133		n.o.	} n.o.
1kc4	B	188	Glu α_L	$t4_{\delta_2\beta_L}$	Lys γ_L	NMR	132		n.o.	
1kl8	B	188	Glu α_L	$t4_{\delta_2\beta_L}$	Lys γ_L	NMR	132		n.o.	} n.o.
11xg	B	191	Thr γ_L	$t4_{\delta_2\beta_L}$	Pro ϵ_L	NMR	131		n.o.	
11xb	B	191	Thr ϵ_D	$t4_{\delta_2\beta_L}$	Pro ϵ_L	NMR	131		n.o.	} n.o.
19b*	A, B, C, D, E	186	Ser α_L	$t4_{\alpha_1\beta_L}$	Pro α_L	Röntgen	129	2,7	VIII	
1h4i	A, C	102	Ala ϵ_L	$t4_{\alpha_1\beta_L}$	Asp ϵ_L	Röntgen	68	1,9	VIII	} VIII
1h4j	A, C, E, D	102	Ala γ_L	$t4_{\alpha_1\delta_L}$	Asp ϵ_D	Röntgen	69	3,0	I	
1g72	A, C	102	Met γ_L	$t4_{\alpha_1\delta_L}$	Asp ϵ_D	Röntgen	70	1,9	I	} I
4aah	A, C	102	Met γ_L	$t4_{\alpha_1\delta_L}$	Asp ϵ_D	Röntgen	71	2,4	I	
1kb0	A	115	Gly ϵ_L	$t4_{\alpha_1\delta_L}$	Asp ϵ_D	Röntgen	72	1,4	I	} I
1kv9	A	104	Ser δ_L	$t4_{\alpha_1\delta_L}$	Asp ϵ_D	Röntgen	73	1,9	I	
1flg	A, B	104	Pro ϵ_L	$t4_{\alpha_1\delta_L}$	Asp ϵ_D	Röntgen	74	2,6	I	} I
1ei9	A	44	Ser β_L	$t4_{\alpha_1\delta_L}$	Asn ϵ_L	Röntgen	84	2,3	I	
1eh5	A	44	Ser β_L	$t4_{\alpha_1\delta_L}$	Asn ϵ_L	Röntgen	84	2,5	I	} I
1exw	A	44	Ser β_L	nem definiált	Asn ϵ_L	Röntgen	85	2,4	I	
1hv6*	A	187	Ser α_L	$t4_{\alpha_1\delta_L}$	Asn ϵ_L	Röntgen	78	2,0	I	} I
1qaz	A	187	Ser α_L	$t4_{\alpha_1\delta_D}$	Asn ϵ_L	Röntgen	79	1,8	n.o.	
1obr	a	154	Gly α_D	$t4_{\alpha_1\delta_L}$	Gly α_D	Röntgen	80	2,3	I	} VIII
1dl0*	A	12	Ala α_L	$t4_{\alpha_1\beta_L}$	Pro α_L	NMR	77		n.o.	
1m4e	A	7	Gly α_D	$t4_{\alpha_1\delta_D}$	His ϵ_L	NMR	82		n.o.	} n.o.
1m4f	A	12	Gly γ_L	$t4_{\alpha_1\delta_D}$	His α_L	NMR	82		n.o.	
1xgc	a	1	Glu γ_D	$t4_{\delta_2\beta_D}$	Asn δ_L	NMR	83		n.o.	α -konotoxin diszulfidkötés izomer

PDB: PDB azonosító kód; * tartalmazza a PDB SELECT; No: az Xxx aminosav sorszáma; Xxx és Yyy: az Xxx ill. az Yyy aminosav minősége és gerinckonformációja (ld. M.1/C Ábra); fb. Röntgen szerkezet felbontása Å egységben; β -kanyar: típus az M.1. Táblázat szerint, n.o.: ez ideig nem osztályozott.

A szerkezetek egyik csoportját az acetyl-kolin-kötő fehérjék vagy receptorok és a származtatott polipeptidek alkotják. Az acetyl-kolin-kötő fehérjében, a gyűrű *t4* konformációjú (PDB azonosító kód: 1i9b)¹²⁹. A nikotinos acetyl-kolin receptorok α -alegységéből származtatott peptidek és toxinok (α -bungarotxin és α -kobratotxin) komplexeiben (1idh, 1idg¹³⁰, 1lxg és 1lxh¹³¹, 1kc4 és 1kl8¹³², 1l4w és 1ljz¹³³) az α -[Cys–Cys] peptidszakasz térszerkezete változatos, *t2*, *t4* és *t6* gyűrűt is tartalmazhat (ld. E.4.3. Táblázat), a vicinális diszulfid konformációs változása a toxin kötődésének hatásaként valószínűsíthető.

A kinoprotein alkohol dehidrogenázokban (1flg¹³⁴, 1h4i¹³⁵, 1h4j⁶⁹, 1g72¹³⁶, 4aah¹³⁷, 1kb0¹³⁸, 1kv9¹³⁹) Yyy mindig aszparaginsav, és egy kivételtől (1h4i) eltekintve ϵ_D gerinckonformációjú, jöllehet ez igen ritka térszerkezet a diszulfidhíd nélküli Cys–Cys egységek után (ld. E.4.5/E Ábra).

A 48 Xxx- α -[Cys–Cys]-Yyy szekvencia mindegyike kielégíti a β -kanyarok $-90^\circ \leq \tau \leq 90^\circ$ kritériumát (a τ paraméter -25° és 68° között változik, súlyozott átlaga 26°). A *d* atomtávolság az 5,4-8,3 Å intervallumba esik, átlaga of 6,8 Å. A *t6* vagy *t8* gyűrűállású konformerek (1k12 kivételével) mind $\epsilon_L\alpha_D$ gerinckonformációt vesznek fel, így II típusú β -kanyart képeznek (ld. E.4.3. Táblázat). A *t2* gyűrűs szerkezetek gerinckonformációja $\delta_D\delta_L$, ezek tehát nem klasszikus β -kanyarok. A *t4* gyűrű mellett nagy változatosságot mutat a peptidgerinc: három VIII típusú β -kanyar (*t4* $\alpha_L\beta_L$), tizenegy I típusú β -kanyar ($\alpha_L\delta_L$, egyikük nem azonosítható gyűrűkonformációval), négy pedig ez ideig nem osztályozott β -kanyar ($\delta_D\delta_L$, $\delta_D\beta_L$, $\alpha_L\delta_D$, $\delta_D\delta_D$).

Összevetés

Az α -[Cys–Cys] szekvenciaegység jelentőségét a vicinális diszulfidhíddal rendelkező fehérjék szerkezetgazó biológiai aktivitása adja. Térszerkezeti szempontból ezeket a fehérjéket röntgendiffrakcióval és NMR-spektroszkópiával vizsgálták. Részletesebb ismeretek szerzése céljából e különleges szerkezeti részletet többféle rövidebb szekvenciával modellezték, majd ezeket kísérleti vagy elméleti módszerekkel vizsgálták. Munkámmal azt kívántam bizonyítani, hogy a lehetséges legrövidebb peptidmodell, ti. a HCO- α -[L-Cys–L-Cys]-NH₂, konformációs terének kvantumkémiail vizsgálata hatékony eszközt ad az α -[Cys–Cys] szekvenciaegység konformációs változatosságának leírásához. Ez a megközelítés azonban nem képes számolni azzal a kísérleti tapasztalattal, ami pl. a prolin-prolin CSD és PDB

SELECT adatainak összevetésénél is felismerhető, hogy a hosszú fehérjeláncok a rövid peptidektől eltérő konformereket kedvelnek. Ez a probléma a konformációs összetétel becslésében jelentkezik és nem érinti a geometriai adatok pontosságát.

A vicinális diszulfidhidas fehérjék statisztikai elemzése nem adott egyetlen olyan konformert sem, melyben a gyűrű *cisz* peptidkötést tartalmaz. Rövidebb modellpeptidekben azonban kísérleti úton is kimutattak *cisz* gyűrűs konformereket, ami összecseng a fenti *ab initio* és DFT számításokkal. A Boc-*ox*-[L-Cys-L-Cys]-OCH₃ röntgendiffrakcióval meghatározott szerkezete¹⁴⁰ *c6* gyűrűt tartalmaz. A fehérjékben talált *cisz* peptidkötések alacsony aránya a rövidebb peptidmodellekre számított vagy kísérletileg meghatározott magasabb aránnyal szemben nem az *ox*-[Cys-Cys] szekvenciaegység egyedi sajátága. A szakirodalomban megjelent vélekedés szerint a kísérletileg meghatározott fehérjeszerkezetekben az aminosavszekvenciától függetlenül jóval kevesebb a *cisz* peptidkötések száma a szabadentalpia-értékek alapján jósolható gyakoriságnál.^{141,142} Tehát az, hogy természetes fehérjék 11 nem homológ családját a PDB-ben képviselő 48 *ox*-[Cys-Cys] szekvenciaegység egyikében sem található *cisz* peptidkötés, a kis számból eredő egyenetlen mintavétellel, ill. a *cisz* peptidkötéseknek szekvenciától függetlenül ritka szereplésével magyarázható.

A CH₃CO-*ox*-[L-Cys-L-Cys]-NH₂ peptidmodellt vizes közegben NMR-spekroszkópiával tanulmányozták.¹⁴³ Két *transz* és két *cisz* peptidkötésű szerkezetének konformációs aránya 47:15:29:9. A gyűrű ω_1 , ϕ_2 , χ_{11} és χ_{21} torziószögét kísérleti úton határozták meg, míg ψ_1 , χ_{12} és χ_{22} értékeit pozitív és negatív χ_{13} mellett Monte Carlo konformációs keresés alapján határozták meg. A tanulmányban közölt gyűrűkonformerek (T-, T⁻, C+ és C- az eredeti jelöléssel) egyértelműen megfeleltethetők *t4*, *t6*, *c6* és *c5* gyűrűállásoknak. A Monte Carlo konformációs keresés további alacsony energiájú térszerkezeteinek (*t2* és *t8*, avagy T+ és T⁺) megjelenését az NMR adatok nem támasztották alá. A CH₃CO-*ox*-[L-Cys-L-Cys]-NH₂ peptidmodell esetében talált mind a hat gyűrűkonformer alacsony energiájúnak mutatkozott *ab initio* és DFT vizsgálataim során is.

A fehérjékben fellelhető négy gyűrűkonformer (*t2*, *t4*, *t6* és *t8*) energetikailag kedvező szerkezetei a 4,5-ditiaheptano-7-laktámnak és a HCO-*ox*-[L-Cys-L-Cys]-NH₂ modellpeptidnek is. A **E.4.3. Táblázat** legmegbízhatóbb szerkezeteit jellemző adatok

átlaigaival jellemzem a négy gyűrűkonformer kísérletileg meghatározható torziószögeit (ld. **E.4.4. Táblázat**). A kísérleti adatok összevetése az elméleti értékekkel egyértelműen képes azonosítani a gyűrűkonformereket, ugyanazon konformerre a két különböző megközelítéssel kapott torziószögek számszerű eltérése mintegy 10-20°. A kísérletek szerint az ω_1 paraméter kevésbé tér el a *transz* peptidkötés lineáris polipeptidekben ideálisnak tekintett 180°-os értékétől, mindössze 11-21°-kal, míg a számított eltérések rendszerint 30° feletti. A legjobb korrelációt a χ_{13} torziószög mutatja, itt a két módszer különbsége csupán néhány fok.

E.4.4. Táblázat A PDB-ből gyűjtött Xxx-ox-[Cys-Cys]-Yyy szekvenciárészeklet gyűrűkonformereinek fontosabb geometriai adatai.

ring	<i>n</i>	ψ_1	ω_1	ϕ_2	χ_{11}	χ_{13}	χ_{21}	d_{S-S}
<i>t2</i> ^a	2	-45	-172	-141	135	106	11	2,05
<i>t4</i>	21	-40	-168	-131	72	-105	-41	2,05
<i>t6</i> ^b	6	172	161	62	82	-80	-35	2,09
<i>t8</i> ^a	2	158	169	52	157	109	16	2,05

n: szekvenciárészeklet száma az **E.4.3. Táblázat** alapján. Torziószögek fokban; d_{S-S} : a kénatomok távolsága Å egységben.

^a NMR-rel meghatározott szerkezet,

^b Röntgen kristallográfiával meghatározott szerkezet (felbontás: 1,1-2,0 Å).

A vicinális diszulfidhíd eredendően alkalmas β -kanyar vagy hasonló kanyarulat kialakítására, de nem illeszkedik α -hélixbe vagy β -redőbe. Egy ilyen térszerkezettel korlátozott modell kvantumkémiail vizsgálata ígéretes, a kiszámított konformerek közül számos megfeleltethető a fehérjékből nyert szekvenciaegységek valamelyikével (***t4* _{$\alpha_L\delta_L$}** , ***t4* _{$\alpha_L\beta_L$}** , ***t6* _{$\beta_L\alpha_D$}**). Néhány további kísérletileg azonosított szerkezethez olyan elméletileg számított pár rendelhető, mely lényegében csak a ϕ_1 torziószögben különbözik (***t2* _{$\delta_D\delta_L$}** → ***t2* _{$\gamma_D\delta_L$}** , ***t6* _{$\epsilon_L\alpha_D$}** → ***t6* _{$\beta_L\alpha_D$}** , ***t8* _{$\epsilon_L\alpha_D$}** → ***t8* _{$\beta_L\alpha_D$}**).

A kvantumkémiail és a kísérleti adatok összevetésén túl jogosan merül fel a kérdés, vajon a viszonylag alacsony elméleti szintű RHF/3-21G* és B3LYP/6-31+G(d) adatok mennyire képesek tükrözni egyes nagyobb számítási igényű eredményeket. Az alábbiakban több ilyen vizsgálatot is ismertetek.

Frekvenciaszámításokat végeztem B3LYP/6-31+G(d) elméleti szinten 14 kiválasztott kedvező energiájú minimumra (ld. **E.4.2. Táblázat**), és ezzel bizonyítottam, hogy mindegyik valódi minimum. Mindössze egyetlen esetben (***t8* _{$\beta_L\alpha_D$}**) találtam imaginárius frekvenciát, mely a potenciálisenergia-felszín laposságával magyarázható, hiszen szorosabb (tight) konvergencia kritérium mellett folytatott optimalás feloldotta a kritikus pont kérdésességét. A frekvenciaszámítások további termokémiai paramétereket szolgáltatottak, melyeket az egyszerű

optimalások relatív elektronikus energiáival korreláltattam (ld. **E.4.5. Táblázat**). Az új paraméterek többsége (ΔE_0 nullpont korrekcióval módosított energia; ΔE_{therm} termikus korrekcióval módosított energia; ΔH entalpia) csak jelentéktelen mértékben módosítja a konformerek viszonylagos sorrendjét, és a ΔG Gibbs szabadentalpia is jól skálázható az egyszerű elektronikus energiához (ld. **E.4.5. Táblázat** lineáris regressziós paraméterei).

E.4.5. Táblázat A HCO-ox-[L-Cys-L-Cys]-NH₂ modellpeptid B3LYP/6-31+G(d) elméleti szinteken optimalt konformereinek energiaértékei és ezek korrelációja.

Konformer	ΔE^a	ΔE_0^b	$\Delta E_{\text{therm}}^b$	ΔH^b	ΔG^b	ΔG_{PCM}^c	ΔE^d	ΔG_{PCM}^e
c5_β _L β _L	0,98	1,01	1,03	1,03	0,41	4,83	0,96	0,87
c5_β _L δ _L	2,12	2,16	2,13	2,13	1,83	3,66	2,01	1,59
c5_ε _L β _L	0,98	1,05	1,07	1,07	0,61	4,27	-	-
c6_β _L β _L	0,00	0,00	0,00	0,00	0,00	0,81	0,00	0,00
c6_β _L δ _L	0,38	0,48	0,37	0,38	0,70	0,00	0,36	0,23
c6_ε _L δ _L	1,84	2,04	2,00	2,00	1,58	6,65	1,76	-
t2_γ _D β _L	7,00	6,88	6,81	6,81	7,55	5,94	6,59	6,35
t2_γ _D δ _L	3,97	4,27	4,11	4,11	5,00	7,15	3,58	6,99
t4_α _L β _L	3,54	3,55	3,50	3,50	3,38	2,84	-	1,43
t4_α _L δ _L	2,02	2,02	1,90	1,90	2,39	2,60	1,92	1,27
t5_β _L γ _D	4,10	4,62	4,38	4,38	5,16	8,42	4,00	7,26
t6_β _L α _D	2,20	2,32	2,11	2,11	2,80	1,48	1,87	2,21
t6_β _L γ _D	0,86	1,20	0,91	0,91	1,91	2,77	0,66	3,33
t8_β _L α _D	4,23	4,78	4,45	4,45	5,24	6,47	3,97	6,63
lineáris regresszió paraméterei								
<i>m</i> (ΔE^a)	1,00	1,01	1,00	1,00	1,13	0,84	0,94	1,10
<i>b</i> (ΔE^a)	0,00	0,12	0,05	0,05	-0,02	2,09	-0,03	0,29
<i>R</i> ² (ΔE^a)	1,000	0,990	0,995	0,995	0,945	0,401	0,997	0,630
<i>m</i> (ΔG^b)	0,83	0,86	0,83	0,83	1,00	0,70	0,79	1,08
<i>b</i> (ΔG^b)	0,15	0,23	0,19	0,19	0,00	2,22	0,03	-0,10
<i>R</i> ² (ΔG^b)	0,945	0,967	0,949	0,949	1,000	0,377	0,940	0,783

Relatív energiaértékek kcal.mol⁻¹ egységben. ΔE elektronikus energia, közvetlenül az optimalásból nyerve; ΔE_0 nullpont korrekcióval módosított energia; ΔE_{therm} termikus korrekcióval módosított energia; ΔH entalpia; ΔG Gibbs szabadentalpia; ΔG_{PCM} Gibbs szabadentalpia oldószer-korrekcióval.

^a B3LYP/6-31+G(d) elméleti szintű optimalásból (ld. **E.4.2. Táblázat**)

^b B3LYP/6-31+G(d) elméleti szintű frekvenciaszámításból

^c B3LYP/PCM/6-31+G(d)//B3LYP/6-31+G(d) energiaszámítás vizes oldatra

^d a konformer B3LYP/6-311+G(d,p) elméleti szintű további optimalásából

^e B3LYP/PCM/6-31+G(d) elméleti szintű, vizes oldatra végzett optimalásból

m és *b*: paraméterek az $y=mx+b$ lineáris regressziós egyenlethez, mely az azonos oszlop energiaadataira, mint *y*-ra, és ΔE^a -ra vagy ΔG^b -re, mint *x*-re illeszkedik.

*R*²: lineáris korreláció az azonos oszlop energiaadatai és ΔE^a ill. ΔG^b között.

A peptidmodell vizes közegben való viselkedésének jellemzésére, PCM energiákat számoltam B3LYP/6-31+G(d) elméleti szinten. A konformerek viszonylagos sorrendjét megváltoztatja az oldószerhatás figyelembevétele, a globális minimum **c6_β_Lδ_L** lesz, és a gázfázisú relatív energiákkal gyenge korreláció mutatkozik. Az újonnan kapott adatok azonban nem illeszkednek jobban a kísérleti értékekhez, mint a gázfázisú optimális eredményei.

Az emelt szintű, B3LYP/6-311++G(d,p) optimálással nyert geometriai és stabilitási adatok a B3LYP/6-31+G(d) eredményekkel jól korrelálnak. A **c5_ε_Lβ_L** ill. a **t4_α_Lβ_L** konformer **c5_β_Lβ_L** ill. **t4_α_Lδ_L** szerkezeté migrált.

Attól, hogy a módszert átváltom RHF-ről B3LYP-re, és a bázist 3-21G-ről 6-31+G(d)-re, majd 6-311++G(d,p)-re a potenciálisenergia-hiperfelszín (PEHS) topológiájának alapvető megtartását várhatjuk és mindössze néhány konformer eltűnését, melyek kedvezőbb minimumokba konvergálnak.¹⁴⁴ Jelentősen megváltozhat azonban a PEHS, amikor az oldószerhatást is figyelembe vesszük a PCM alkalmazásával, és elképzelhető, hogy a konformerek gázfázisú előválogatása miatt elveszítünk egyet vagy többet a PCM PEHS alacsony energiájú minimumaiból is.

A 14 kiválasztott konformer B3LYP/6-31+G(d) elméleti szintű optimálását elvégeztem a vizes közeget PCM-mel modellezve. A **c5_ε_Lβ_L** ill. a **c6_ε_Lδ_L** konformer a **c5_β_Lβ_L** ill. **c6_β_Lδ_L** minimumba migrált. A többi tizenkét minimum esetében a PCM optimalás hatására csak apróbb térszerkezeti változások történtek: eltérések átlaga 2° a gyűrűtorzióknál, ill. 7° a ψ_1 és ϕ_2 paramétereknél, maximuma pedig 7° ill. 22°. A konformerek β -kanyar voltát jellemző d paraméter sohasem változott 0,35°-nál nagyobb mértékben. Relatív szabad entalpiák (ΔG) valamivel jobban korrelálnak a gázfázisú relatív energiákkal, mint a PCM energiaszámítások (ld. **E.4.5. Táblázat**). A PCM optimalás és a PCM energiaszámítások ΔG értékei közötti lineáris regresszió paraméterei: $m = 0,93$, $b = -0,47$ és $R^2 = 0,771$. Végeredményben tehát megállapítható, hogy a PCM energiaszámítások semmivel sem jósólják meg sikeresebben a PCM optimalás energiaértékeit, mint a gázfázisú, B3LYP/6-31+G(d) elméleti szintű számítások adatai (ez utóbbi korreláció $R^2 = 0,783$, az **E.4.5. Táblázat**ban is szerepel).

ÖSSZEFOGLALÁS

Peptidek konformációs nevezérendszere

Peptidkonformerek olyan nevezérendszerének bevezetését javasoljuk, amely egyaránt alkalmasnak bizonyult egyszerű diamidok és tetszőleges hosszúságú fehérjeláncok térszerkezeti megjelölésére. A kód utal az összetevő aminosavak téralkatára (az N-terminálistól a C-terminális felé haladva), és eltérő karaktereket alkalmaz a peptidkötések, a gerinc ($[\phi, \psi]$ párok), valamint az oldalláncok jellemzésére. Ez a nevezéktan nyitott az esetleges változtatásokra, így pl. a gerinc jelölhető a Ramachandran-térkép bármely más felosztása szerint is, miközben megtartjuk az amidok (t vagy c) és az oldallánc (+, -, a vagy s) jeleit. A nevezérendszer használható peptidszerkezetek félkvitatív leírására, tehát a konformációk megjelölésére és összevetésére akár egyazon, akár részben eltérő szekvenciák esetén. E rendszert alkalmaztam a tanulmányozott prolin, prolin-prolin és kibővítve a vicinális diszulfidhidas ciszteinil-cisztein konformerek elnevezésére.

Alanin diamid stabilitásvizsgálata az oldószer modellezésével

Kiszámoltam vizes közegre a HCO-L-Ala-NH₂ modellpeptid Ramachandran-felsőfelületét RHF/PCM/3-21G, RHF/PCM/6-31+G(d) és B3LYP/PCM/6-31+G(d) elméleti szinten. A minimumokat optimáltam mindhárom említett, továbbá a B3LYP/PCM/6-31++G(d,p) elméleti szinten. Az eredményeket a PDB SELECT-ből gyűjtött kísérleti adatokkal összevetve megállapítottam, hogy az alanin diamidon a Polarizable Continuum Model alkalmazásával végzett számítások a gázfázisú megközelítéseknél jobban modellezik a fehérjék alaninjait, mivel képesek optimált minimumot adni az α -hélix és a poliprolin II régióban is (α_L ill. β_L). A PCM ésszerű alternatívának mutatkozik az oldószerhatás figyelembevételére, ha el kívánjuk kerülni az explicit oldószer-molekuláknak a modellrendszerbe illesztésével járó számolási nehézségeket. Az RHF/PCM/6-31+G(d) és a B3LYP/PCM/6-31+G(d) elméleti szinten számolt frekvenciák jól korrelálnak egymással.

Prolin

Összeállítottam a prolin konformációs könyvtárát egyrészt a Cambridge Adatbázisból gyűjtött Xxx-**Pro**-Yyy szekvenciákra vonatkozó kísérleti adatok alapján, másrészt a HCO-L-**Pro**-NH₂ modellpeptid végzett *ab initio* számítások révén. A *cis-transz* izomerizáció és a gerinckonformáció mellett a gyűrűállás kvantitatív elemzését is elvégeztem pszeudorotációs amplitúdó, *A*, és fázisszög, *P*, számítása útján. A prolin kísérleti és kvantumkémiai módszerrel nyert adatainak minőségi és mennyiségi összevetését egyaránt elvégeztem. A konformerek kiválasztott geometriai jellemzőinek eltérése a kísérleti és az elméleti halmazok között rendszerint a kísérleti értékek szórásához hasonló mértékűnek bizonyult. A tizenkét ideális konformer közül nem pontosan ugyanazok fordulnak elő a különböző halmazokban, és az energiaszámítás alapján becsülhető gyakoriság gyengén tükrözi a CSD-ben fellelt szintetikus peptidok prolinkonformereinek gyakorisági adatait. Az eredmények jobb illeszkedése várható hosszabb modellpeptid választásától, melyben az inverz γ -kanyartól eltérő típusú hidrogénhidak is kiépülhetnek, illetve oldószermodell alkalmazásától.

Prolil-prolin

Az HCO-L-**Pro**-L-**Pro**-NH₂ peptidmodellen végzett kvantumkémiai számításokkal PCM használata mellett sikeresen jellemeztem a polipeptidekben fellelhető **Pro-Pro** szekvenciaegységeket, hiszen az elméleti és a kísérleti megközelítéssel ugyanaz a hat gerinckonformer mutatkozott legstabilabbnak a 36 ideálisan lehetséges szerkezet közül. A gyakorisági sorrend némileg eltérő, de a hat kedvező szerkezet a prolil-prolin egységek nagy többségét (72 és 100 % közötti mértékben) képes leírni. A gázfázisú számítások gyengébben teljesítettek, mivel azokhoz a gerincalkatokhoz rendeltek gyakoribb előfordulást, amelyekben belső hidrogénkötés található ($\tau_{\text{L}}\tau_{\text{L}}$ és a $\tau_{\text{L}}\alpha_{\text{L}}$). A fehérjeszerkezetek között pedig az NMR-spektroszkópiával finomítottak tűnnek kevésbé megbízhatóknak, vagyis elfogultnak a *transz* peptidkötés és a pozitív gyűrűállású pirrolidin irányába, és a röntgendiffrakcióval meghatározottaknál sokkal gyakrabban rendelnek a prolinhoz τ_{L} gerinckonformációt. Ez valószínűleg azzal magyarázható, hogy a kísérleti adatok kevés erőteljes utalást adnak a prolinok alapértelmezett jellemzőinek megváltoztatására.

A prolin-prolin szekvenciaegység fehérjékben és peptidekben jellegzetes konformációs kötöttségeket és preferenciákat okoz. A prolinokban a ϕ torziószőgek *gauche*- állásúak, és a megelőző Xxx aminosav kedveli a nyújtott-hosszított szerkezeteket. Az Xxx-**Pro-Pro**-Yyy szekvenciaegységek Xxx és Yyy helyzetben található aminosavai jelentősen befolyásolják a *cisz* peptidkötés előfordulásának valószínűségét.

A prolin-prolin kétféle téralkatot vehet fel: hosszított szerkezet alkot vagy β -kanyarrá hajlik. A fehérjékben ugyan a poliprolin II alkat ($\tau_{\mathbf{E}_1}\tau_{\mathbf{E}_L}$) a legjellemzőbb, fontos szerepet játszanak azonban más hosszított formák is, melyek némileg magasabb relatív energiával rendelkeznek és nevezékrendszerük még nem alakult ki ($\tau_{\mathbf{E}_1}\tau_{\mathbf{A}_L}$, $\mathbf{cE}_L\tau_{\mathbf{E}_L}$ és $\tau_{\mathbf{E}_L}\tau_{\mathbf{Y}_L}$). A β -kanyaroknak több típusa is elérhető a **Pro-Pro** egység számára, így a $\tau_{\mathbf{E}_L}\mathbf{cA}_L$ (VIa), $\tau_{\mathbf{A}_L}\tau_{\mathbf{A}_L}$ (I vagy III), $\tau_{\mathbf{E}_L}\mathbf{cE}_L$ (VIb), és az ez ideig nem osztályozott, magas energiájú kanyarok ($\mathbf{cE}_L\mathbf{cE}_L$ és $\mathbf{cE}_L\mathbf{cA}_L$).

A prolin-prolin olyan szekvenciaegység, ahol a lokális gerinckonformációs szempontok jól érvényesülnek a globális kooperativitással szemben. A pirrolidingyűrű ugyan konformációs kötöttségeket okoz, ugyanakkor redőzöttsége bizonyos rugalmasságot is ad a peptidgerincnek. Ez magyarázat lehet arra, hogy a prolin gyakran járul hozzá a rendezetlen fehérjék vagy domének szegmentális mozgásához. A hosszított szerkezetek 87%-os és a β -kanyarok 13%-os statisztikai gyakorisága egyazon fehérje mozgékony **Pro-Pro** szekvenciaegysége esetén az idő változó mentén valósulhat meg.

Reakcióút kiszámításával bizonyítottam a feltételezett kapcsolatot a HCO-L-Pro-L-Pro-NH₂ modellpeptid egyik *cisz-transz* átmeneti állapota és két minimum között. Az ω_2 torziószőg elfordulását a peptidgerinc ψ_2 torziójának változása és az oldallánc térszerkezetének összetett módosulása kíséri, miközben egy stabilizáló hidrogénhid-kötés folyamatosan fennmarad. A pszeudorotációs kifejezés nem csak a minimumokban és az átmeneti állapotban, hanem a reakcióút többi kiszámított pontjában is jól illeszthető volt. A két származtatott változó, *A* és *P*, az öt gyűrűtorziót 1,5°-os hibán belül képes reprodukálni.

Vicinális diszulfidhidas ciszteinil-cisztein

Az RHF/3-21G* elméleti szinten folytatott kimerítő térszerkezeti keresés és azt követő B3LYP/6-31+G(d) szintű optimalálás a $\text{HCO-ox-[L-Cys-L-Cys]-NH}_2$ peptidmodellen 14 alacsony energiájú konformert eredményezett. Az **ox-[Cys-Cys]** szekvenciaegység adatbázisból és szakirodalomból nyert kísérleti adatai jól értelmezhetők a fenti számított konformerekkel. Az **ox-[Cys-Cys]** egység a β -kanyarok változatos, így I, II, VIa, VIb és VIII, továbbá ez ideig nem osztályozott típusait képes felvenni, miközben a **Cys-Cys** szekvenciaegység vicinális diszulfidhíd nélkül nem kényszerül β -kanyar kialakítására. Bár a vicinális diszulfid fehérjékben meglehetősen ritka, előfordulásakor mégis fontos szerkezeti/funkcionális motívumnak bizonyul. Kvantumkémiaili számításaim alapján a vicinális diszulfidhíd jelentőségét az adja, hogy olyan egyébként magas energiájú β -kanyar szerkezeteket képes stabilizálni, mint a VIII típus, továbbá a *cisz* peptidkötést tartalmazó VIa és VIb típus. Fehérjékben több példa is mutatkozott VIII típusú β -kanyart képező **ox-[Cys-Cys]** szekvenciaszakaszra, VIa vagy VIb típusú β -kanyar azonban egy sincs.

IRODALMI HIVATKOZÁSOK

-
- 1 Hudáky I, Kiss R, Perczel A. A nomenclature of peptide conformers. *J Mol Struct (THEOCHEM)* 2004;675:177-183.
- 2 Hudáky I, Baldoni HA, Perczel A. Peptide Models XXXVIII. Proline Conformers from X-ray crystallographic database and from ab initio computations. *J Mol Struct (THEOCHEM)* 2002;582:233-249.
- 3 Hudáky I; Perczel A. Prolylproline unit in model peptides and in fragments from databases. *Proteins-Structure Function and Bioinformatics* 2008;70:1389-1407.
- 4 Hudáky I; Gáspári Z; Carugo O; Čemažar M; Pongor S; Perczel A. Vicinal disulfide bridge conformers by experimental methods and by ab initio and DFT molecular computations. *Proteins-Structure Function and Bioinformatics* 2004;55(1):152-168.
- 5 Hudáky I; Hudáky P; Perczel A. Solvation model induced structural changes in peptides. A quantum chemical study on Ramachandran surfaces and conformers of alanine diamide using the polarizable continuum model. *J Comp Chem* 2004;25(12):1522-1531.
- 6 Hudáky I, Perczel A. Ab initio analysis of the conformational changes of the prolylproline model peptide. *J Mol Struct (THEOCHEM)* 2003;630:135-140.
- 7 Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains approved by IUPAC-IUB Commission on Biochemical Nomenclature in 1974, World Wide Web version prepared by G.P. Moss
<http://www.chem.qmw.ac.uk/iupac/misc/pppep1.html> és további oldalak.
- 8 Sasisekharan V. Stereochemical criteria for polypeptide and protein structures. In: Ramanathan N, ed. *Collagen*. Madras, India: Wiley & Sons. 1962; pp 39-78.
- 9 Ramachandran GN, Sasisekharan V. Conformation of polypeptides and proteins. *Adv Prot Chem* 1968;23:283-438.
- 10 Karplus PA. Experimentally observed conformation-dependent geometry and hidden strain in proteins. *Prot Sci* 1996;5:1406-1420, és további hivatkozások a publikációban.
- 11 Császár AG, Perczel A. Ab initio characterization of building units in peptides and proteins. *Prog Biophys Mol Biol* 1999;71:243-309, és további hivatkozások a publikációban.
- 12 Vargas R, Garza J, Hay BP, Dixon DA. Conformational Study of the Alanine Dipeptide at the MP2 and DFT Levels. *J Phys Chem* 2002;106:3213-3218, és további hivatkozások a publikációban.
- 13 Perczel A, Farkas Ö, Jákl I, Topol IA, Csizmadia IG. Peptide models XXXIII. Extrapolation of low-level Hartree-Fock data of peptide conformation to large basis set SCF, MP2, DFT, and CCSD(T) results. The Ramachandran surface of alanine dipeptide computed at various levels of theory. *J Comput Chem* 2003;24:1026-1042.
- 14 Gáspári Z; Hudáky I; Czajlik A; Perczel A. Is there an excuse for the non-conformist? Notes on the calculated energies, atom-atom contacts and natural abundance of the different conformers of alanine in proteins. *J Mol Struct (THEOCHEM)* 2004;675:141-148.

-
- 15 Perczel A, Jákl I, Csizmadia IG. Intrinsically stable secondary structure elements of proteins: A comprehensive study of folding units of proteins by computation and analysis of data determined by X-ray crystallography. *Chem Eur J* 2003;9:2-12.
- 16 Benzi C, Improta R, Scalmani G, Barone V. Quantum Mechanical Study of the Conformational Behavior of Proline and 4R-Hydroxyproline Dipeptide Analogues in Vacuum and in Aqueous Solution. *J Comput Chem* 2002;23:341-350.
- 17 Topol IA, Burt SK, Deretey E, Tang TH, Perczel A, Rashin A, Csizmadia IG. α - and 3_{10} -Helix Interconversion: A Quantum-Chemical Study on Polyalanine Systems in the Gas Phase and in Aqueous Solvent. *J Am Chem Soc* 2001;123:6054-6060.
- 18 Improta R, Mele F, Crescenzi O, Benzi C, Barone V. Understanding the role of stereoelectronic effects in determining collagen stability. 2. A quantum mechanical/molecular mechanical study of (proline-proline-glycine)_n polypeptides. *J Am Chem Soc* 2002;124:7857-7865.
- 19 (a) Miertus S, Scrocco E, Tomasi J. Electrostatic Interaction of a Solute with a Continuum - A Direct Utilization of Ab initio Molecular Potentials for the Prevision of Solvent Effects. *Chem Phys* 1981;55:117-129, (b) Cancès MT, Mennucci B, Tomasi J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J Chem Phys* 1997;107:3032-3041, (c) Cossi M, Barone V, Mennucci B, Tomasi J. Ab initio study of ionic solutions by a polarizable continuum dielectric model. *Chem Phys Lett* 1998;286:253260, (d) Mennucci B, Tomasi J. Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries. *J Chem Phys* 1997;106:5151, (e) Amovilli C, Barone V, Cammi R, Cancès E, Cossi M, Mennucci B, Pomelli CS, Tomasi J. Recent advances in the description of solvent effects with the polarizable continuum model. *Adv Quantum Chem* 1999;32:227-261.
- 20 Iwaoka M, Okada M, Tomoda S. Solvent effects on the ϕ - ψ potential surfaces of glycine and alanine dipeptides studied by PCM and I-PCM methods. *J Mol Struct (THEOCHEM)* 2002;586:111-124.
- 21 Walther D, Cohen FE. Conformational attractors on the Ramachandran map. *Acta Cryst D* 1999;55:506-517.
- 22 Scott AP, Radom L. Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, Moller-Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors. *J Phys Chem* 1996;100:16502-16513.
- 23 Baker J, Jarzecki AA, Pulay P. Direct Scaling of Primitive Valence Force Constants: An Alternative Approach to Scaled Quantum Mechanical Force Fields. *J Phys Chem A* 1998;102:1412-1424.
- 24 Yoshida H, Takeda K, Okumura J, Ehara A, Matsuura H. A New Approach to Vibrational Analysis of Large Molecules by Density Functional Theory: Wavenumber-Linear Scaling Method. *J Phys Chem A* 2002;106:3580-3586.
- 25 Kubelka J, Kleiderling TA. Ab Initio Calculation of Amide Carbonyl Stretch Vibrational Frequencies in Solution with Modified Basis Sets. 1. *N*-Methyl Acetamide. *J Phys Chem A* 2001;105:10922-10928.
- 26 Watson TM, Hirst JD. Density Functional Theory Vibrational Frequencies of Amides and Amide Dimers. *J Phys Chem A* 2002;106:78587858-7867.

-
- 27 Barth A, Zscherp C. What vibrations tell us about proteins. *Quarterly Reviews of Biophysics* 2002;35:369-430, és további hivatkozások a publikációban.
- 28 Watson TM, Hirst JD. Density Functional Theory Vibrational Frequencies of Amides and Amide Dimers. *J Phys Chem A* 2002;106:7858-7867.
- 29 MacArthur MW, Thornton JM. Influence of proline residues on protein conformation. *J Mol Biol* 1991;218:397-412.
- 30 Milner-White EJ, Bell LH, Maccallum PH. *J Mol Biol* 1992;228:725-734.
- 31 Wilmot CM, Thornton JM. β -Turns and their distortions: a proposed new nomenclature. *Protein Engineering* 1990;3(6):479-493, és további hivatkozások a publikációban.
- 32 Karle IL. Conformation of the Lithium Ion Complex of Antamanide, a Cyclic Decapeptide and Ion Carrier, in the Crystalline State. *J Am Chem Soc* 1974;96:4000-4006.
- 33 Di Blasio B, Rossi F, Benedetti E, Pavone V, Pedone C, Temussi PA, Zanotti G, Tancredi T. Bioactive Peptides: Solid-State and Solution Conformation of Cyclolinopeptide A. *J Am Chem Soc* 1989;111:9089-9098.
- 34 Morita H, Gonda A, Takeya K, Itokawa H, Hirano T, Oka K, Shirota O. Solution State Conformation of an Immunosuppressive Cyclic Dodecapeptide, Cycloleonorinin. *Tetrahedron* 1997;53:7469-7478.
- 35 Morita H, Gonda A, Takeya K, Itokawa H, Iitaka Y. Cycloleonoripeptide D, A New Proline-Rich Cyclic Decapeptide from *Leonurus heterophyllum*. *Tetrahedron* 1997;53:1617-1626.
- 36 Nirthanan S, Gwee MCE. Three-Finger α -Neurotoxins and the Nicotinic Acetylcholine Receptor, Forty Years On. *J Pharmacol Sci* 2004;94:1-17.
- 37 Basus VJ, Song G, Hawrot E. NMR Solution Structure of an α -Bungarotoxin/Nicotinic Receptor Peptide Complex. *Biochemistry*, 1993;32:12290-12298.
- 38 Love RA, Stroud RM. The crystal structure of alpha-bungarotoxin at 2.5 Å resolution: relation to solution structure and binding to acetylcholine receptor. *Protein Eng* 1986;1:37-46.
- 39 Samson AO, Scherf T, Eisenstein M, Chill JH, Anglister J. The Mechanism for Acetylcholine Receptor Inhibition by α -Neurotoxins and Species-specific Resistance to α -Bungarotoxin Revealed by NMR. *Neuron* 2002;35:319-332.
- 40 Le Du MH, Marchot P, Bougis PE, Fontecilla-Camps JC. 1.9-Å Resolution Structure of Fasciculin-1, an Anti-acetylcholinesterase Toxin from Green Mamba Snake Venom. *J Biol Chem* 1992;267:22122-22130.
- 41 Bilwes A, Rees B, Moras D, Ménez R, Ménez A. X-ray Structure at 1.55 Å of Toxin γ , a Cardiotoxin from *Naja nigricolis* Venom. *J Mol Biol* 1994;239:122-136.
- 42 O'Connell JF, Bougis PE, Wuthrich K. Determination of the NMR Solution Structure of Cardiotoxin CTX-IIb from *Naja mossambica mossambica*. *Eur J Biochem* 1993;213:891-900.
- 43 Zinn-Justin S, Roumestand C, Gilquin B, Bontems F, Menez A, Toma F. 3-Dimensional Solution Structure of a Curare-mimetic Toxin from *Naja nigricollis* Venom: A Proton NMR and Molecular Modeling Study. *Biochemistry* 1992;31:11335-11347.

-
- 44 Mitchell SS, Shon KJ, Foster MP, Davis DR, Olivera BM, Ireland CM. Tree-Dimensional Solution Structure of Conotoxin ψ -PIIIE, an Acetylcholine Gated Ion Channel Antagonist. *Biochemistry* 1998;37:1215-1220.
- 45 Van Wagoner RM, JacobsenRB, Olivera BM, Ireland CM. Characterization and Tree-Dimensional Solution Structure Determination of ψ -Conotoxin PIIIF, a Novel Antagonist of Acetylcholine Receptors. *Biochemistry* 2003;42:6353-6362.
- 46 Hill JM, Alewood PF, Craik DJ. Solution structure of the sodium channel antagonist conotoxin GS: a new molecular caliper for probing sodium channel geometry. *Structure*, 1997;5:571-583.
- 47 Hill JM, Alewood PF, Craik DJ. Conotoxin TVIIA, a novel peptide from the venom of *Conus tulipa* 2. Three-dimensional solution structure. *Eur J Biochem*, 2000;267:4649-4657.
- 48 Wakamatsu K, Kohda D, Hatanaka H, Lancelin J-M, Ishida Y, Oya M, Nakamura H, Inagaki F, Sato K. Structure-Activity Relationships of μ -Conotoxin GIIIA: Structure Determination of Active and Inactive Sodium Channel Blocker Peptides by NMR and Simulated Annealing Calculations. *Biochemistry* 1992;31:12577-12584.
- 49 Hill JM, Alewood PF, Craik DJ. Three-Dimensional Solution structure of μ -Conotoxin GIIIB, a Specific Blocker of Skeletal Muscle Sodium Channels. *Biochemistry* 1996;35:8824-8835.
- 50 Le Bonniec BF, Guinto ER, MacGillivray RTA, Stone ST, Esmon CT. The Role of Thrombin's Tyr-Pro-Trp Motif in the Interaction with Fibrinogen, Thrombomodulin, Protein-C, Antithrombin-III, and the Kunitz Inhibitors. *J Biol Chem* 1993;268:19055-19061.
- 51 Banner DW, Hadvary P. Crystallographic Analysis At 3.0-Å Resolution of the Binding to Human Thrombin of Four Active Site-directed Inhibitors. *J Biol Chem* 1991;266:20085-20093.
- 52 Xiao G, Todorova M, Jagadeesh J, Drohat AC, Stivers JT, Gilliland GL. Crystal structure of *Escherichia coli* uracil DNA glycosylase and its complexes with uracil and glycerol: Structure and glycosylase mechanism revisited. *Proteins: Structure, Function and Genetics*, 1999;35:13-24.
- 53 Mol CD, Arvai AS, Sanderson RJ, Slupphaug G, Kavli B, Krokan HE, Mosbaugh DW, Tainer JA. Crystal-Structure of Human Uracil-DNA Glycosylase in Complex with a Protein Inhibitor - Protein Mimicry of DNA. *Cell*, 1995;80:869-878.
- 54 Savva R, McAuley-Hecht K, Brown T, Pearl L. The structural basis of specific base-excision repair by uracil-DNA Glycosylase. *Nature*, 1995;373:487-493.
- 55 Kay BK, Williamson MP, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J*, 2000;14:231-241.
- 56 Ogura K, Nagata K, Horiuchi M, Ebisui E, Hasuda T, Yuzawa S, Nishida M, Hatanaka H, Inagaki F. Solution structure of N-terminal SH3 domain of Vav and the recognition site for Grb2 C-terminal SH3 domain. *J Biomol NMR* 2002;22:37.

-
- 57 Vidal M, Goudreau N, Cornille F, Cussac D, Gincel E, Garbay C. Molecular and cellular analysis of Grb2 SH3 domain mutants: interaction with Sos and dynamin. *J.Mol.Biol.* 1999;290:717-730.
- 58 Pisabarro MT, Serrano L, Wilmanns M. Crystal Structure of the Abl-SH3 Domain Complexed with a Designed High-affinity Peptide Ligand: Implications For SH3-Ligand Interactions. *J Mol Biol* 1998;281:513-521.
- 59 Zou P, Pinotsis N, Lange S, Song Y-H, Popov A, Mavridis I, Mayans OM, Gautel M, Wilmanns M. Palindromic assembly of the giant muscle protein titin into the sarcomeric Z-disk. *Nature* 2006;439:229-233.
- 60 Atkinson RA, Joseph C, Kelly G, Muskett FW, Frenkiel TA, Nietlisbach D, Pastore A. Ca²⁺-independent binding of an EF-hand domain to a novel motif in the α -actinin-titin complex. *Nat Struct Biol* 2001;8:853-857.
- 61 Buevich AV, Dai Q-H, Liu X, Brodsky B, Baum J. Site-Specific NMR Monitoring of cis-trans Isomerization in the Folding of the Proline-Rich Collagen Triple Helix. *Biochemistry*, 2000;39:4299-4308 és további hivatkozások a publikációban.
- 62 Kramer RZ, Vitagliano L, Bella J, Berisio R, Mazzarella L, Brodsky B, Zagari A, Berman HM. X-ray crystallographic determination of a collagen-like peptide with the repeating sequence (Pro-Pro-Gly). *J Mol Biol* 1998;280:623-638.
- 63 Kramer RZ, Bella J, Mayville P, Brodsky B, Berman HM. Sequence dependent conformational variations of collagen triple-helical structure. *Nat Struct Biol* 1999;6:454-457.
- 64 Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. Protein Disorder Prediction: Implication for Structural Proteomics. *Structure* 2003;11:1453-1459.
- 65 Tompa P. The functional benefits of protein disorder. *J Mol Struct (THEOCHEM)* 2003;666-667:361-371 és további hivatkozások a publikációban.
- 66 Paterson D, Nordberg A. Neuronal nicotinic receptors in the human brain. *Prog Neurobiol* 2000;61:75-111.
- 67 Bourne PC, Isupov MN, Littlechild JA. The atomic-resolution structure of a novel bacterial esterase. *Structure* 2000;8:143-151.
- 68 Ghosh M, Anthony C, Harlos K, Goodwin M, Blake C. The refined structure of the quinoprotein methanol dehydrogenase from *Methylobacterium-extorquens* at 1.94 angstrom. *Structure* 1995;3:177-187.
- 69 Afolabi PR, Mohammed F, Amaratunga K, Majekodunmi O, Dales SL, Gill R, Thompson D, Cooper JB, Wood SP, Goodwin PM, Anthony C. Site-directed mutagenesis and X-ray crystallography of the PPQ-containing quinoprotein methanol dehydrogenase and its electron acceptor, cytochrome *c_L*. *Biochemistry* 2001;40:9799-9809.
- 70 Zheng YJ, Xia Z, Chen Z, Mathews FS, Bruice TC. Catalytic mechanism of quinoprotein methanol dehydrogenase: a theoretical and X-ray crystallographic investigation. *Proc Natl Acad Sci USA* 2001;98:432-434.
- 71 Xia Z, Dai W, Zhang Y, White SA, Boyd GD, Mathews FS. Determination of the gene sequence and the three-dimensional structure at 2.4 Å resolution of methanol dehydrogenase from *Methylophilus W3A1*. *J Mol Biol* 1996, 259, 480.

-
- 72 Oubrie A, Rozeboom HJ, Kalk KH, Huizinga EG, Dijkstra BW. Crystal structure of quinoxinohemoprotein alcohol dehydrogenase from *Comamonas testosteroni*. *J Biol Chem* 2002;277:3727-3732.
- 73 Chen Z, Matsushita K, Yamashita T, Fujii T, Toyama H, Adachi O, Bellamy HD, Matthews FS. Structure at 1.9 Å resolution of a quinoxinohemoprotein alcohol dehydrogenase from *Pseudomonas putida* HK5. *Structure* 2002;10:837-849.
- 74 Keitel T, Diehl A, Knaute T, Stezowski JJ, Höhne W, Görisch H. X-ray structure of the quinoprotein ethanol dehydrogenase from *pseudomonas aeruginosa*: basis of substrate specificity. *J Mol Biol* 2000;297:961-974.
- 75 Blake CCF, Ghosh M, Harlos K, Avezoux A, Anthony C. The active site of methanol dehydrogenase contains a disulphide bridge between adjacent cysteine residues. *Nat Struct Biol* 1994;1:102-103.
- 76 Avezoux A, Goodwin MG, Anthony C. The role of the novel disulphide ring in the active site of the quinoprotein methanol dehydrogenase from *Methylobacterium extorquens*. *Biochem J* 1995;307:735-741.
- 77 Wang X, Connor M, Smith R, Maciejewski MW, Howden MEH, Nicholson GM, Christie MJ, King GF. Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge. *Nat Struct Biol* 2000;7:505-513.
- 78 Yoon HJ, Hashimoto W, Miyake O, Murata K, Mikami B. Crystal structure of alginate lyase A1-III complexed with trisaccharide product at 2.0 Å resolution. *J Mol Biol* 2001;307: 9-16.
- 79 Yoon HJ, Mikami B, Hashimoto W, Murata K. Crystal structure of alginate lyase A1-III from *Sphingomonas* species A1 at 1.78 Å resolution. *J Mol Biol* 1999;290:505-514.
- 80 Tepljakov A, Polyakov K, Obmolova G, Strokopytov B, Kuranova I, Osterman A, Grishin N, Smulevitch S, Zagnitko O, Galperina O, Matz M, Stepanov V. Crystal structure of carboxypeptidase T from *Thermoactinomyces vulgaris*. *Eur J Biochem* 1992;208:281-288.
- 81 Bianchet MA, Odom EW, Vasta GR, Amzel M. A novel fucose recognition fold involved in innate immunity. *Nat Struct Biol* 2002;9:628-634.
- 82 Hunter HN, Fulton DB, Ganz T, Vogel HJ. The solution structure of human hepcidin, a peptide hormone with antimicrobial activity that is involved in iron uptake and hereditary hemochromatosis. *J Biol Chem* 2002;277:37597-37603.
- 83 Gehrman J, Alewood PF, Craik DJ. Structure determination of the three disulfide bond isomers of α -conotoxin GI: a model for the role of disulfide bonds in structural stability. *J Mol Biol* 1998;278:401-415.
- 84 Bellizzi JJ, Widom J, Kemp C, Lu JY, Das AK, Hofmann DJ, Clardy J. The crystal structure of palmitoyl protein thioesterase 1 and the molecular basis of infantile neuronal ceroid lipofuscinosis. *Proc Natl Acad Sci USA* 2000;97:4573-4578.
- 85 Das AK, Bellizzi JJ, Tandel S, Biehl E, Clardy J, Hofmann SL. Structural basis for the insensitivity of a serine enzyme (palmitoyl-protein thioesterase) to phenylmethylsulfonyl fluoride. *J Biol Chem* 2000;275:23847-23851.

-
- 86 Nurizzo D, Turkenburg JP, Charnock SJ, Roberts SM, Dodson EJ, McKie VA, Taylor EJ, Gilbert HJ, Davies GJ. Cellvibrio japonicus α -L-arabinase 43A has a novel five-blade β -propeller fold. *Nat Struct Biol* 2002;9:665-668.
- 87 Baker J, Here WJ. Geometry optimization in Cartesian coordinates: the end of the Z-matrix? *J Comput Chem* 1991;12:606-610.
- 88 Schlegel HB. A comparison of geometry optimization with internal, Cartesian, and mixed coordinates. *Int J Quantum Chem, Quantum Chem Symposium* 1992;26:243-252.
- 89 Viviani W, Rivail J-L, Perczel A, Csizmadia IG. Peptide models. 3. Conformational potential energy hypersurface of formyl-L-valinamide. *J Am Chem Soc* 1993;115:8321-8329.
- 90 Hudáky P, Perczel A, Csizmadia IG. Deciphering factors which determine the Ramachandran surface of peptides. The application of isodesmic surfaces, $\Delta E_{ID}(\varphi, \psi)$, to analyze the contribution of rotating moieties to the shape of potential energy surfaces. *J Mol Struct (THEOCHEM)* 2000;500:59-96.
- 91 Zimmerman SS, Pottle MS, Némethy G, Scheraga HA. Conformational analysis of the 20 naturally occurring amino acid residues using ECEPP. *Macromolecules*, 1977;10:1-9.
- 92 Scarsdale JN, Van Alsenoy C, Klimkowski VJ, Schäfer L, Momany FA. Ab initio studies of molecular geometries 27. Optimized molecular structures and conformational analysis of $N\alpha$ -acetyl-N-methylalaninamide and comparison with peptide crystal data and empirical calculations. *J Am Chem Soc* 1983;105:3438-45.
- 93 Perczel A, Ángyán JG, Kajtár M, Viviani W, Rivail J-L, Marcoccia JF, Csizmadia IG. Peptide Models 1. Topology of Selected Peptide Conformational Potential-Energy Surfaces (Glycine and Alanine Derivatives). *J Am Chem Soc* 1991;113:6256-6265.
- 94 Jalkanen KJ, Suhai S. *N*-Acetyl-L-alanine-*N'*-methylamide: a density functional analysis of the vibrational absorption and vibrational circular dichroism spectra. *Chem Phys* 1996;208:81-116.
- 95 Hudáky P, Perczel A, Csizmadia IG. Amino acid conformational analyses of proteins (ACAP program). *J. Mol. Struct. (THEOCHEM)* 362 (1996) 263-273.
- 96 Székely Z, Kónya Z, Becskei A, Goldring WPD, Perczel A, Penke B, Molnár J, Michejda CF, Aszalós A, Csizmadia IG. Suggested binding mechanism of the HIV-gp120 to its CD4 receptor. *J Mol Struct (THEOCHEM)* 1996;367:159-186.
- 97 Karplus PA. Experimentally observed conformation-dependent geometry and hidden strain in proteins. *Protein Science*, 1996;5:1406-1420.
- 98 Efimov AV. Standard structures in prroteins. *Prog Biophys Mol Biol* 1993;60:201-239.
- 99 Richardson JS, Richardson DC, in *Prediction of Protein Structure and the Principles of Protein Conformation* (ed. Fasman, G.) New York, Plenum Press (1989) p. 1-98.
- 100 Rooman MJ, Kocher JPA, Widak SJ. Extracting information on folding from amino-acid-sequence – accurate predictions for protein regions with preferred conformation in the absence of tertiary interactions. *Biochemistry* 1992;31:10226-10238.
- 101 Thornton J, Jones DT, MacArthur MW, Orengo CM, Swindells MB. Protein Folds: Towards Understanding Folding from Inspection of Native Structures *Phil Trans Roy Soc London* 1995;B348:71-79.

-
- ¹⁰² Magyar nyelvű szakirodalonként pl. Veszprémi T, Fehér M. A kvantumkémia alapjai és alkalmazása. Műszaki könyvkiadó, 2002.
- ¹⁰³ Gaussian 98 (Revision A.1), Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ES, Pople JA. Gaussian, Inc., Pittsburgh PA, 1998.
- ¹⁰⁴ Gaussian 03 (Revision B.01), Frisch M J, Trucks G W, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson G A, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain M C, Farkas O, Malick D K, A Rabuck D, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA, Gaussian, Inc., Pittsburgh PA, 2003.
- ¹⁰⁵ (a) Pulay P. Convergence acceleration of iterative sequences - the case of SCF iteration. *Chem Phys Lett* 1980;73:393-398, (b) Pulay P. Improved SCF convergence acceleration. *J Comput Chem* 1982;3:556-560, (c) Császár P, Pulay P. Geometry optimization by direct inversion in the iterative subspace. *J Mol Struct (THEOCHEM)* 1984 (d)114:31-34, (d) Farkas Ö, Schlegel HB. Methods for optimizing large molecules - Part III. An improved algorithm for geometry optimization using direct inversion in the iterative subspace (GDIIS). *Phys Chem Chem Phys Lett* 2002;4:11-15.
- ¹⁰⁶ Allen FH, Kennard O. 3D Search and Research Using the Cambridge Structural Database. *Chemical Design Automation News*, 1993;8(1):1 & 3-37, www.ccdc.cam.ac.uk/products/csd.
- ¹⁰⁷ Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Research* 2000;28:235-242, www.rcsb.org.
- ¹⁰⁸ (a) Hobohm U, Scharf M, Schneider R, Sander C. Selection of representative protein data sets. *Protein Sci* 1992;1:409-417., (b) Hobohm U, Sander C. Enlarged representative set of protein structures. *Protein Sci* 1992;1:409-417.
- ¹⁰⁹ Perczel A, McAllister MA, Csaszar P, Csizmadia IG. Peptide models 6. New beta-turn conformations from ab initio calculations confirmed by X-ray data of proteins. *J Am Chem Soc* 1993;115:4849-4858.

-
- 110 Venkatachalam, CM. Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. *Biopolymers*, 1968;6:1425-1436.
- 111 Wilmot CM, Thornton JM. β -turns and their distortions: a proposed new nomenclature. *Protein Eng* 1990;3:479-493.
- 112 Lee HJ, Song JW, Choi YS, Park HM, Lee KB. A theoretical study of conformational properties of N-methyl azapeptide derivatives. *J Am Chem Soc* 2002;124:11881-11893.
- 113 Ashish, Grover A, Kishore R. Characterization of a novel type VII beta-turn conformation for a bio-active tetrapeptide rigidin - A synergy between theoretical and experimental results. *Eur J Biochem* 2000;267:1455-1463.
- 114 Han W, Jalkanen KJ, Elstner M, Suhai S. Theoretical Study of Aqueous *N*-Acetyl-L-alanine-*N'*-Methylamide: Structures and Raman, VCD, and ROA Spectra. *J Phys Chem B* 1998;102:2587-2602.
- 115 Baldoni HA, Rodriguez AM, Zamora MA, Zamarbide GN, Enriz RD, Farkas Ö, Császár P, Torday LL, Sosa CP, Jákli I, Perczel A, Papp JG, Hollosi M, Csizmadia IG. Peptide models XXIV: An ab initio study on N-formyl-L-prolinamide with *trans* peptide bond. The existence or non-existence of α_L and ϵ_L conformations. *J Mol Struct (THEOCHEM)* 1999;465:79-91.
- 116 Kang YK. Ab initio molecular orbital calculations on low-energy conformers of N-acetyl-*N'*-methylprolineamide. *J Phys Chem*. 1996;100:11589-11595.
- 117 Fischer S, Dunbrack RL, Karplus M Jr. Cis-trans imide isomerization in the proline dipeptide. *J Am Chem Soc* 1994;116:11931.
- 118 Czugler M, Sasvári K, Hollósi M. Crystal Structure of cyclo-(Gly-L-Pro-L-Pro-Gly-L-Pro-L-Pro) Trihydrate. Unusual Conformational Characteristics of a Cyclic Hexapeptide. Crystal Structure of cyclo-(Gly-L-Pro-L-Pro-Gly-L-Pro-L-Pro) Tihydrate. Unusual Conformational Characteristics of a Cyclic Hexapeptide. *J Am Chem Soc* 1982;104:4465-4469.
- 119 Fuess H, Bats J.W. *Am Cryst Assoc, Abstr Papers (Summer)*, 1983;29.
- 120 Nakashima T, Yamane T, Tanaka I, Ashida T. Structure of cyclo(-L-Pro-L-Pro-Gly-L-Pro-L-Leu-Gly-) methanol molvate monohydrate, $C_{25}H_{38}N_6O_6 \cdot CH_3OH \cdot H_2O$. *Acta Crystallogr, Sect C*, 1984;40:171-174.
- 121 MacArthur MW, Thornton JM. Influence of Proline Residues on Protein Conformation. *J Mol Biol* 1991;218:397-412.
- 122 Ho BK, Bresseur R. The Ramachandran plots of glycine and pre-proline. *BMC Struct Biol* 2005;5:14, és további hivatkozások a publikációban.
- 123 Perczel A, Jákli I, McAllister MA, Csizmadia I.G. Relative stability of major types of beta-turns as a function of amino acid composition: A study based on ab initio energetic and natural abundance data. *Chem-Eur J* 2003;9:2551-2566.
- 124 Adzhubei AA, Sternberg MJE. Left-handed Polyproline II Helices Commonly Occur in Globular Proteins. *J Mol Biol* 1993;229:472-493.
- 125 Kakinoki S, Hirano Y, Oka M. On the Stability of Polyproline-I and II Structures of Proline Oligopeptides. *Polymer Bulletin* 2005;53:109-115.

-
- 126 Vitagliano L, Berisio R, Mazzarella L, Zagari A. Structural Bases of Collagen Stabilization Induced by Proline Hydroxylation. *Biopolymers* 2001;58:459-464.
- 127 Perczel A, McAlister MA, Császár P, Csizmadia IG. Peptide models 9. A complete conformational set of For-Ala-Ala-NH₂ from ab initio computations. *Can J Chem* 1994;72:2050-2070.
- 128 Ramek M, Yu CH, Schäfer L. Ab initio conformational analysis of the model tripeptide N-formyl-L-alanyl-L-alanine amide. *Can J Chem* 1998;76:566-575.
- 129 Brejc K, van Dijk WJ, Klaassen RV, Schuurmans M, van der Oost J, Smit AB, Sixma TK. Crystal structure of an Ach-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* 2001;411:269-276.
- 130 Zeng H, Moise L, Grant MA, Hawrot E. The solution structure of the complex formed between α -bungarotoxin and an 18-mer cognate peptide derived from the $\alpha 1$ subunit of the nicotinic acetylcholine receptor from the Torpedo californica. *J Biol Chem* 2001;276:22930-22940.
- 131 Zeng H, Hawrot E. NMR-based binding screen and structural analysis of the complex formed between α -cobratoxin and an 18-mer cognate peptide derived from the $\alpha 1$ subunit of the nicotinic acetylcholine receptor from Torpedo californica. *J Biol Chem* 2002;277:37439-37455.
- 132 Moise L, Piserchio A, Basus VJ, Hawrot E. NMR structural analysis of α -bungarotoxin and its complex with the principal α -neurotoxin-binding sequence on the $\alpha 7$ subunit of a neuronal nicotinic acetylcholine receptor. *J Biol Chem* 2002;277:12406-12417.
- 133 Samson AO, Scherf T, Eisenstein M, Chill JH, Anglister J. The mechanism for acetylcholine receptor inhibition by α -neurotoxins and species-species resistance to α -bungarotoxin revealed by NMR. *Neuron* 2002;35:319-332.
- 134 Keitel T, Diehl A, Knaute T, Stezowski JJ, Höhne W, Görisch H. X-ray structure of the quinoprotein ethanol dehydrogenase from pseudomonas aeruginosa: basis of substrate specificity. *J Mol Biol* 2000;297:961-974.
- 135 Ghosh M, Anthony C, Harlos K, Goodwin M, Blake C. The refined structure of the quinoprotein methanol dehydrogenase from methylobacterium-extorquens at 1.94 angstrom. *Structure* 1995;3:177-187.
- 136 Zheng YJ, Xia Z, Chen Z, Mathews FS, Bruice TC. Catalytic mechanism of quinoprotein methanol dehydrogenase: a theoretical and X-ray crystallographic investigation. *Proc Natl Acad Sci USA* 2001;98:432-434.
- 137 Xia Z, Dai W, Zhang Y, White SA, Boyd GD, Mathews FS. Determination of the gene sequence and the three-dimensional structure at 2.4 Å resolution of methanol dehydrogenase from Methylophilus W3A1. *J Mol Biol* 1996, 259, 480.
- 138 Oubrie A, Rozeboom HJ, Kalk KH, Huizinga EG, Dijkstra BW. Crystal structure of quinohemoprotein alcohol dehydrogenase from Comamonas testosteroni. *J Biol Chem* 2002;277:3727-3732.
- 139 Chen Z, Matsushita K, Yamashita T, Fujii T, Toyama H, Adachi O, Bellamy HD, Matthews FS. Structure at 1.9 Å resolution of a quinohemoprotein alcohol dehydrogenase from Pseudomonas putida HK5. *Structure* 2002;10:837-849.

-
- ¹⁴⁰ Hata Y, Matsuura Y, Tanaka N, Ashida T, Kakudo M. Tert-butyloxycarbonyl-L-cysteinyl-L-cysteine disulfide methyl-ester. *Acta Crystal B* 1977;33:3561-3564.
- ¹⁴¹ Weiss MS, Jabs A, Hilgenfeld R. Peptides bonds revisited. *Nat Struct Biol* 1998;5:676.
- ¹⁴² Jabs A, Weiss M, Hilgenfeld R. Non-proline cis peptide bonds in proteins. *J Mol Biol* 1999;286:291-304.
- ¹⁴³ Creighton CJ, Reynolds CH, Lee DHS, Leo GC, Reitz AB. Conformational analysis of the eight-membered ring of the oxidized cysteinyl-cysteine unit implicated in nicotinic acetylcholine receptor ligand recognition. *J Am Chem Soc* 2001;123:12664-12669.
- ¹⁴⁴ Perczel A, Farkas Ö, Jákl I, Topol IA, Csizmadia IG. Peptide models XXXIII. Extrapolation of low-level Hartree-Fock data of peptide conformation to large basis set SCF, MP2, DFT, and CCSD(T) results. The Ramachandran surface of alanine dipeptide computed at various levels of theory. *J Comput Chem*, 2003;24:10261042.