

Doktori disszertáció

Nacsa András Bence

Semleges és protonált gázfázisú aminosavak
ab initio konformeranalízise

Témavezető:

Dr. Czakó Gábor

Egyetemi docens



Szegedi Tudományegyetem

Természettudományi és Informatikai Kar

Fizikai Kémiai és Anyagtudományi Tanszék

Kémia Doktori Iskola

MTA-SZTE Lendület Elméleti Reakciódinamika Kutatócsoport

Szeged, 2024

TARTALOMJEGYZÉK

1. Rövidítések listája	1
2. Publikációs lista	2
3. Bevezetés	3
4. Irodalmi áttekintés	7
4.1. Glicin	7
4.2. Cisztein	9
4.3. Szerin	11
4.4. Protonaffinitás és gázfázisú bázicitás	12
5. Célkitűzés	15
6. Elméleti háttér	16
7. Eredmények és diszkusszió	23
7.1. Konformerek	23
7.1.1. Gly és GlyH ⁺ konformerek	24
7.1.2. Cys és CysH ⁺ konformerek	26
7.1.3. Ser és SerH ⁺ konformerek	36
7.2. Benchmark eredmények	40
7.2.1. Gly és GlyH ⁺ nagy pontosságú geometriák és energiák	41
7.2.2. Cys és CysH ⁺ nagy pontosságú geometriák és energiák	48
7.2.3. Ser és SerH ⁺ nagy pontosságú geometriák és energiák	57
7.3. Protonaffinitások és gázfázisú bázicitások	67
7.3.1. A glicin protonaffinitása és gázfázisú bázicitása	67
7.3.2. A cisztein protonaffinitása és gázfázisú bázicitása	71
7.3.3. A szerin protonaffinitása és gázfázisú bázicitása	73
8. Összefoglalás	76
9. Summary	80
10. Köszönetnyilvánítás	83
11. Hivatkozások	84

1. RÖVIDÍTÉSEK LISTÁJA

AE	minden elektron (all electron)
aug-cc-pVnZ	korreláció-konzisztens polarizált vegyérték n -zeta + diffúz függvények
cc-pVnZ-F12	explicit-korrelált módszerhez kifejlesztett korreláció-konzisztens polarizált vegyérték n -zeta
CC	csatolt-klaszter (coupled-cluster)
CI	konfigurációs kölcsönhatás (configuration interaction)
CCSD(T)	csatolt klaszter egyszeres, kétszeres és perturbatív háromszoros gerjesztésekkel (coupled-cluster with singles, doubles and perturbative triples)
CCSD(T)-F12	explicit-korrelált CCSD(T)
CCSDT	csatolt klaszter egyszeres, kétszeres és háromszoros gerjesztésekkel (coupled-cluster with singles, doubles and triples)
CCSDT(Q)	csatolt klaszter egyszeres, kétszeres, háromszoros és perturbatív négyszeres gerjesztésekkel (coupled-cluster with singles, doubles, triples and perturbative quadruples)
DK	Douglas–Kroll
FC	fagyasztott törzs (frozen core)
GB	gázfázisú bázicitás
GTO	Gauss-típusú pálya (Gauss-type orbital)
HF	Hartree–Fock
MP n	n -ed rendű Møller–Plesset
NIST	Nemzeti Szabványügyi és Technológiai Intézet (National Institute of Standards and Technology)
PA	protonaffinitás
PES	potenciálisenergia-felület
SCF	önkonzisztens tér (self consistent field)
STO	Slater-típusú pálya (Slater-type orbital)
STO- n G	n darab primitív GTO-ból álló bázisfüggvény

2. PUBLIKÁCIÓS LISTA

Doktori értekezés alapját képező közleményeim listája:

- [1] **A. B. Nacsa** & G. Czakó: Benchmark ab initio proton affinity of glycine, *Phys. Chem. Chem. Phys.*, 2021, **23**, 9663–9671.
- [2] **A. B. Nacsa** & G. Czakó: Benchmark ab initio determination of the conformers, proton affinities and gas-phase basicities of cysteine, *J. Phys. Chem. A*, 2022, **126**, 9667–9679.
- [3] **A. B. Nacsa**, M. Kígyósi, G. Czakó: Protonation of serine: Conformers, proton affinities and gas-phase basicities at the "gold standard" and beyond, *Phys. Chem. Chem. Phys.*, 2023, **25**, 8891–8902.

Doktori értekezéshez kapcsolódó közleményeim listája:

- [4] E. M. Orján, **A. B. Nacsa**, G. Czakó: Conformers of dehydrogenated glycine isomers, *J. Comput. Chem.*, 2020, **41**, 2001–2014.

Doktori értekezéshez nem kapcsolódó közleményeim listája:

- [5] **A. B. Nacsa**, V. Tajti, G. Czakó: Dynamics of the $\text{Cl}^- + \text{CH}_3\text{I}$ reaction on a high-level ab initio analytical potential energy surface, *J. Chem. Phys.*, 2023, **158**, 194306.
- [6] T. Gstir, T. Michaelsen, B. A. Long, **A. B. Nacsa**, A. Ayasli, D. Swaraj, F. Zappa, F. Trummer, S. G. Ard, N. S. Shuman, G. Czakó, A. A. Viggiano, R. Wester: The influence of fluorination on the dynamics of the $\text{F}^- + \text{CF}_3\text{CH}_2\text{I}$ reaction, *Phys. Chem. Chem. Phys.*, 2023, **25**, 18711–18719.
- [7] **A. B. Nacsa**, C. Tokaji, G. Czakó: High-level analytical potential-energy-surface-based dynamics of the $\text{OH}^- + \text{CH}_3\text{CH}_2\text{Cl}$ $\text{S}_{\text{N}}2$ and $\text{E}2$ reactions in full (24) dimensions, *Faraday Discuss.*, 2024, DOI: 10.1039/D3FD00161J.

3. BEVEZETÉS

Minimum konformereknek az olyan egyensúlyi szerkezeteket hívjuk, amelyek az adott molekula potenciálisenergia-felületletén (PES) globális vagy lokális minimumhoz tartoznak. Nagy, flexibilis molekulák esetén számos ilyen konformációs izomer létezik, melyeknek különböző fizikai, vagy akár kémiai tulajdonságai is lehetnek. A két, leggyakrabban alkalmazott kísérleti módszer, amellyel aminosavak konformereinek szerkezetét lehet vizsgálni, az a mikrohullámú és az infravörös spektroszkópia. Azonosításuk, megkülönböztetésük a különböző forgási és rezgési tulajdonságaikon alapszik (ezeknek a meghatározásában jelentős szerepet játszik az elmélet). A konformerek egyensúlyi eloszlását UV vagy infravörös-közeli (NIR) besugárzással változtathatjuk meg, azonban a méréseket többek között az nehezíti, hogy a konformerek között általában csekély az energiabeli különbség, így azokat nem lehet teljesen szelektíven előállítani, illetve az alagúteffektus következtében is könnyen átalakulhatnak egymásba. Ahhoz, hogy teljes képet kapjunk a molekuláris szintű tulajdonságokról, érdemes elméleti oldalról feltérképeznünk az összes lehetséges konformert. Az ilyen irányú kutatások az alapját képezik a számítógépes kémiának, a kémiai informatikának vagy akár a számítógépes gyógyszertervezésnek. A stabil egyensúlyi geometriák feltérképezéséről Hawkins 2017-ben közölt egy alapos összefoglaló cikket, amelyben bemutatja a két fő irányzatot.⁸ Az egyikhez – amely napjainkban egyre inkább feltörekvően van – tartoznak a gépi tanulás vagy mesterséges intelligencia alapú sztochasztikus eljárások, ezeket már aminosavak esetén is sikeresen alkalmazták.⁹⁻¹³ Az algoritmusok a molekulák konformációs terének valamilyen előzetes megfontolások alapján megtervezett leképezésén alapszanak. Megítélésem szerint ezek a módszerek a jövőben rendkívül hasznos módszerek lehetnek, hiszen a számítási igényeket drasztikusan le tudják csökkenteni, viszont egyelőre még nem adják vissza az összes eredményt, amit a „hagyományos” módszerekkel kaphatunk. Ez utóbbiakat képviseli a másik irányzat, amelyet egyébként mi is alkalmaztunk, a konformációs tér szisztematikus leképezése. Ekkor abból indulunk ki, hogy a molekulában lévő kötésszögek és kötэшosszak jóval „merevebbek”, mint a diéderes, vagy torziós szögek (a későbbiekben az utóbbi megnevezést fogom használni), ezért a molekula egyes részleteinek a megfelelő kötés mentén való szisztematikus elforgatásával megkereshetjük a lehetséges konformereket. Minél nagyobb szerkezetről van szó, általában annál több ilyen forgatási tengely jelenik meg, az előállítható lehetséges szerkezetek száma rohamosan nő, ezért ahhoz, hogy az elméleti számolások belátható időn belül eredményeket szolgáltatassanak, körültekintően kell

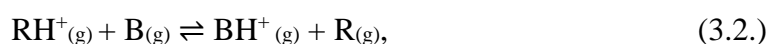
megválasztanunk a geometriai optimálásokhoz alkalmazott elméleti szintet illetve a forgatások során alkalmazott felbontást, vagyis azt, hogy hány fokként forgassunk. Az így kapott szerkezetek és az ezekre számolt rezgési és forgási paraméterek lehetővé teszik a konformerek spektroszkópiai azonosítását, kiváltképp a csillagközi térben, ahol az élet kialakulásához elengedhetetlen aminosavak keletkezésének és jelenlétének vizsgálata rendkívül izgalmas és népszerű téma.¹⁴⁻²² Mivel a semleges aminosav szerkezetek feltehetőleg nagyfokú hasonlóságot mutatnak a belőlük képződött vegyületek konformereivel, ezért ezek meghatározásánál is fel tudjuk őket használni. Erre példa lehet a dehidrogénezett glicin konformerek vizsgálata, amelyről egy korábbi publikációban értekeztünk,⁴ illetve az itt bemutatásra kerülő protonált aminosav szerkezetek meghatározása.

Milyen további lehetőségeink vannak, ha ismerjük egy vegyület, esetünkben aminosav (és annak protonált formájának) szerkezeit? Vizsgáljuk a következő gázfázisú reakciót:

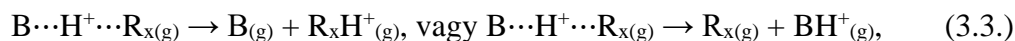


ahol BH^+ a protonált konjugált sav, B az analóg semleges bázis és H^+ pedig a szabad proton. A reakciót kísérő entalpiaváltozás a protonaffinitás (PA), míg a szabadentalpiaváltozás a gázfázisú bázicitás (GB). A meghatározott abszolút értékek felhasználhatóak a sav-bázis tulajdonságok számszerűsítésére, amelyek a pK_a -val ellentétben már magára a molekulára lesznek jellemzőek, oldószertől függetlenül. A gázfázisú protontranszfer reakciók természetesen a gyakorlatban is fellelhetőek, például kémiai ionizációt alkalmazó tömegspektrometriás módszerek esetén. Ide kapcsolódik még az úgynevezett „Mobile Proton Model”, amely arra is magyarázatot ad, hogy az aminosavak protonaffinitása hogyan befolyásolja a fehérjék fragmentálódását és hogy a C-, vagy az N-terminális lesz a töltéshordozó a tömegspektrometriás mérések során.²³⁻²⁸ Kutatók 1987-ben azt találták, hogy az aminosavak bázicitása összefüggésben van egyes enzimek aktivitásával,²⁹ ehhez kapcsolódóan a gáz-fázisú bázicitás értékek is relevánsak, hiszen az enzimek aktív helyein lejátszódó reakciók nagyban hasonlítanak a gáz-fázisú reakciókra.³⁰ Végül, de nem utolsósorban ismét megemlíteném a csillagközi térben lejátszódó ion-molekula reakciókat, mint a PA és GB értékek felhasználásának területét,¹⁷⁻¹⁹ hiszen a prekursoraiból képződő aminosavak nem csak gyökös, de ionos reakciókból is előállhatnak, ezek jellemzéséhez pedig szükségesek a fent említett termodinamikai értékek ismerete. Az abszolút mennyiségek meghatározása először is történhet elméleti úton, ahogyan az a disszertációhoz

vezető munka során is történt. Kísérletileg azonban körülményes, hiszen szükséges hozzá mindhárom, a reakcióban szereplő species képződéshőjének ismerete. Ez az eset csak kevés vegyület esetén áll fent, ilyen például az etilén, propén, vagy az izobutén, amelyek később relatív skálák felállítására is alkalmasak.³¹ Ha rendelkezünk néhány pontos és megbízható adattal, relatív értékeket már könnyebb meghatározni.³² A tömegspektrometriás méréseken alapuló kísérletek megtervezéséhez és kiértékeléséhez tartozó eljárásokat a következőkben ismertetem. Az első lehetséges módszer a gázfázisú egyensúlyi állandók vizsgálata a következő reakcióban:

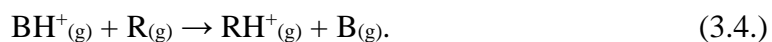


ahol R/RH^+ a referenciavegyület és annak protonált formája, míg B/BH^+ a vizsgálni kívánt vegyület (és protonált formája). Ha a termodinamikai egyensúly beállta után tömegspektrometriás mérésekkel meghatározzuk az ionok egymáshoz viszonyított mennyiségét, valamint ismerjük R és B mennyiségét, akkor az egyensúlyi állandó közvetlen kapcsolatot teremt a referencia és az ismeretlen gázfázisú bázicitás értékek között, adott hőmérsékleten. Azonban ha például B instabil molekula, reakcióba lép saját vagy a referenciavegyület protonált formájával, akkor protontranszferhez tartozó egyensúlyi állandó meghatározása e módon nehézkes, vagy akár lehetetlen. Ezekre az esetekre találták ki a „bracketing” módszert, amelynek alapja, hogy a BH^+ protonált formát több referenciavegyülettel reagáltatják (természetesen egyesével), majd azt vizsgálják, hogy történik-e protontranszfer. Amennyiben R_x referenciavegyülettel már nem, R_y referenciavegyülettel még igen, akkor a keresett gázfázisú bázicitás érték az R_x -hez és R_y -hoz tartozó értékek között lesz (innen ered a „bracketing” kifejezés), hiszen a protontranszfer reakció lejátszódásához negatív szabadentalpiaváltozás kötődik. Természetesen nem kívánt reakciók lejátszódása, izomerek vagy komplexek képződése ekkor is hátráltathatja erőfeszítéseinket. Egy másik elterjedt eljárás a következő, ütközés hatására lejátszódó disszociáció vizsgálata:



ahol $\text{B}\cdots\text{H}^+\cdots\text{R}$ egy proton által összetartott dimer ion. A két protonált forma képződésének aránya és a relatív protonaffinitások között közvetlen kapcsolat állítható fel, amennyiben a két termék képződéséhez kapcsolódó entrópiaváltozás hasonló. A két termék mennyiségének aránya azonos a két reakció sebességi együtthatójának arányával. Ez utóbbiakra Arrhenius-féle alakot felírva levezethető a relatív protonaffinitásokhoz vezető

formula, azonban ekkor is többféle referenciavegyületet kell alkalmaznunk. Mivel a sebességi együtthatókat használtuk fel, ezért ezt a módszert kinetikus módszernek nevezik. Egy újabb megközelítés a következő reakción alapszik:



A gázfázisú bázicitás illetve protonaffinitás értékek korrelálnak a (3.4.) reakcióhoz tartozó szabadentalpiaváltozással illetve a reakció hatásfokával, azaz a kísérletileg mért és az ütközési elmélet alapján számolt sebességi együtthatók hányadosával. A reakció hatásfokát mérve, különböző, ismert gázfázisú bázicitással rendelkező bázisok alkalmazása mellett az ismeretlen is meghatározható, ez a termokinetikus módszer.

Az aminosav konformerek egy másik, általam is érintett felhasználási területe a PES-ek illesztése. Ezeknek a felületeknek a globális és lokális minimum szerkezetek a legfontosabb pontjai (az átmeneti állapotokhoz tartozó szerkezetekkel egyetemben), nélkülük aligha tudnánk leírni egy molekulához tartozó PES-t. Természetesen egy szükséges, de nem elégséges feltétel áll fent. Ha ismerjük a konformerek relatív energiát, akkor Boltzmann-eloszlást feltételezve adott hőmérsékleten meghatározhatjuk, hogy milyen szerkezetek milyen valószínűséggel fordulnak elő a konformerelegyenben. A kísérleti eredmények azonban sokszor azt mutatják, hogy a helyzet ennél összetettebb, konformerek hiányoznak, vagy akár nem vártak jelennek meg. A konformációs dinamika mély, átfogó elméleti vizsgálata választ adhat ezekre a furcsaságokra, tanulmányozhatóak a lehetséges átmenetek, a konformerek közötti gátak illetve az alagúteffektusok. Ehhez pedig egy pontos potenciálisenergia-felületre van szükség, amelynek alapja a legfontosabb pontok ismerete. A konformációs dinamikán kívül kémiai reakciók is tanulmányozhatóak reaktív PES-ekkel. Ekkor azonban nem elég csupán a semleges aminosavak geometriájából kiindulnunk, hiszen a képződő termékeknek is megannyi izomere létezhet, ezért a termék szerkezetek is kulcsfontosságúak. Az egyik legegyszerűbb, aminosavakkal lejátszódó reakció a protonálódás, ekkor a termékoldalról a különböző (geometriájú és protonálódási helyű) protonált aminosav konformerek fogják szolgáltatni a minimum pontokat.

4. IRODALMI ÁTTEKINTÉS

4.1. Glicin

A glicin konformereikhez kapcsolódóan először Császárt említem meg,^{33–35} aki számításaival lefedte az alapokat a későbbi elméleti és kísérleti vizsgálatokhoz. 1992-ben³³ 13 semleges glicin szerkezetet vizsgált MP2, MP4, CCSD és CCSD(T) módszerek segítségével, 6-31++G** és 350 kontrahált Gauss-függvényből álló Partridge bázisokat alkalmazva. A geometriákat MP2/6-31++G** szinten határozta meg, a legfontosabb eredmény ezzel kapcsolatban az, hogy megmutatta, hogy a tizenhárom geometriából csak nyolc darab lesz minimum Hartree–Fock módszernél magasabb elméleti szintű számításokat használva. Az öt konformer, amelyek ekkor már nem minimum szerkezet, csupán a szimmetriájában tér el a hozzájuk tartozó valódi minimumoktól, az előbbieket síkszimmetrikusak, míg az utóbbiak aszimmetrikusak. Három évvel később a glicin (és alanin) alacsony energiájú szerkezeteit vizsgálta alaposabban és vetette össze a forgási állandókat a kísérletileg mértekkel.³⁴ 2007-ben pedig munkatársaival a glicin két legalacsonyabb energiájú konformerére (Ip és IIn) elvégezték az első csatolt-klaszter szintű geometriai optimalizációkat és vizsgálták a szerkezetek egymásba alakulásának gátmagasságát.³⁵ 2009-ben Balabin a DFT oldaláról közelítette meg a problémát és B3LYP/aug-cc-pVTZ elméleti szinten optimalizálta a glicin nyolc minimumát, meghatározta relatív energiáikat (különböző járulékokat figyelembe véve) és a harmonikus rezgési frekvenciáikat.³⁶ Számos elméleti (vagy elméleti és kísérleti) eredmény született még (és születik is napjainkkal bezárólag) a glicin konformerek szerkezetéről, azok átalakulásáról, azonban ezekben a csatolt-klaszter elméleti szint alkalmazása ritka, akkor is inkább csak single-point számolásra, vagy pedig valamilyen közelítő kompozit módszeren belül.^{4,11,12,37–43} A helyzetet talán Puzzarini 2013-as publikációjának a címe foglalja össze a legjobban: „Glycine Conformers: A Never-ending Story?”.³⁸ Kiemelném még egy korábbi munkánk 2020-ból,⁴ amelyben elsőként hajtottunk végre a kvantumkémiai „golden-standard” módszerével geometriai optimalizációkat glicin összes konformerére, vagyis az explicit-korrelált csatolt-klaszter módszerrel, aug-cc-pVDZ és aug-cc-pVTZ bázisokat alkalmazva, az előbbiekkal pedig a harmonikus rezgési analízist is elvégeztük. A protonált formához kapcsolódóan egy 1998-ban született cikkben megtalálható a három ismert N-protonált glicin MP2/6-31++G** elméleti szinten.⁴⁴ Noguera 2001-ben két N-protonált konformert azonosított DFT módszerekkel, amelyekhez a semleges glicin konformereiből jutott el.⁴⁵ Ugyan megemlíti, hogy a protonálódást lehetségesnek találták diol-forma kialakulásával is, HF szinten (rendre 6 és 2 O-protonált

szerkezetet találtak korábban az N-protonáltak mellett),^{46,47} azonban mivel az aminocsoport protonálódása a termodinamikailag kedvezőbb (becslésük szerint körülbelül 30 kcal/mol-lal), a karbonil-protonálódásnak nem jártak alaposabban utána. Két későbbi tanulmányban 2 amino- és 2 karbonil-protonált szerkezetet említenek.^{48,49}

A konformerek kísérleti vizsgálatához kapcsolódóan ismét előkerül Balabin neve, aki a számolásait immár mérésekkel is kiegészítette.⁵⁰ Mivel az irodalomban még nem volt közölve a glicin infravörös spektruma, így elsőként rögzítette a glicin Raman spektrumát, amelynek alacsony frekvenciájú régiója az elérhető felbontás mellett lehetővé tette három különböző konformer azonosítását (Ip, IIn, (irodalomban elsőként) IVn). Két évvel később ugyanezzel a berendezéssel, de más kísérleti körülmények között kimutatta a IIIIn szerkezet jelenlétét is.⁵¹ Tarczay igen aktív kutató ezen a területen, a glicin konformereinek vizsgálatával kapcsolatban itt két cikket emelném ki.^{52,53} 2012-ben a glicin közeli infravörös spektrumát rögzítették argon és kripton mátrixokban.⁵² A mátrix-izolált glicint ezután NIR lézerrel sugározták be, hogy megváltoztassák a konformerek eloszlását és tanulmányozhassák azok egymásba alakulását, amiket kvalitatívan le is tudtak írni. Új átmenetek felfedezése mellett először adódott kísérleti bizonyíték a VIp létezésére. Még ugyanebben az évben vizsgálták többek között a mátrixként használt gázok (argon, kripton, xenon és nitrogén) hatását a VIp konformer felezési idejére (alagút-effektussal való átalakulására).⁵³ Azt találták, hogy nitrogéngázban ez az idő jóval megnő, mivel a glicin molekulák komplexeket képeznek a nitrogénnel. A protonált glicin kísérleti vizsgálatáról összesen két publikáció található ugyanazoktól a szerzőktől, ugyanazokkal a konklúziókkal. Az első cikkben hat különböző aminosav protonálódási helyeit és protonált konformereit vizsgálták IRMPD (infrared multiple-photon dissociation)⁵⁴ spektroszkópiával.⁴⁸ A hagyományos IR spektroszkópiával szemben ez a kísérleti elrendezés könnyűvé teszi az ionok IR spektrumának felvételét. Megjegyezték, hogy a protonálódás a karbonilcsoporton is lejátszódhat, azonban nagy energiakülönbség nem teszi lehetővé a kialakulását (az adott kísérleti körülmények között). A számított és kísérleti spektrumok összevetésével csupán a legalacsonyabb energiájú, GH01 konformer jelenlétét tudták igazolni. A második cikkben a protonált glicin és annak oligomerjeit vizsgálták, a protonált glicinről ugyanazokat az eredményeket közölték,⁴⁹ mint korábban.

4.2. Cisztein

A konformációs tér leképezése a rendszer méretének (és flexibilis részeinek) növekedésével egyre nagyobb kihívássá válik. Ez nyilvánvalóan így van akkor is, ha a glicin aminosav után már a ciszteint szeretnénk vizsgálni, ezért az ide kapcsolódó források száma kevesebb, az alkalmazott elméleti szintek alacsonyabbak. Az első, szisztematikus feltérképezésre 1995-ből találunk leírást Gronert és O’Hair cikkében.⁵⁵ Az akkor elérhető számítógépes kapacitásokat maximálisan kihasználva 324 próbaszerkezeten végeztek el AM1⁵⁶ szintű optimalizációkat, ezeket a szerkezeteket pedig a korábban ismertetett módon, a megfelelő funkciós csoportok (bizonyos megfontolások melletti) elforgatásával állították elő. Az így kapott minimumokat HF/6-31G* szinten tovább optimalizálták, majd MP2/6-31+G* single-point számításokat végeztek rájuk, összesen 42 stabil cisztein konformert találtak. A következő hasonló publikációt 2007-ben tették közzé Dobrowolski és munkatársai.⁵⁷ Szintén ugyanebből a 324 próbaszerkezetből indultak ki, ezeket AM1 szinten optimalizálták, majd a minimumokat tovább optimalizálták az intramolekuláris kölcsönhatásokat pontosabban leíró B3LYP és MP2 módszerekkel aug-cc-pVDZ bázisokkal, végül 51 stabil szerkezetet határoztak meg. 2 évvel később megjelent az irodalomban fellelhető eddigi legrészletesebb vizsgálata a cisztein konformációs terének. Császár és Allen munkatársaikkal jóval nagyobb felbontás mellett vizsgálta a cisztein flexibilis funkciós csoportjait, összesen 11664 próbageometriát állítottak elő a megfelelő forgatásokkal.⁵⁸ Ezeket HF/3-21G szinten optimalizálták és 90 különböző minimum szerkezetet kaptak, azonban ezekből MP2/cc-pVTZ szinten már csak 71 darab maradt. A 11 legalacsonyabb relatív energiájú szerkezetet további analízisnek vetették alá és meghatározták a hozzájuk tartozó rezgési paramétereket. 2011-ben Bouchoux több aminosav sav-bázis termokémiájáról publikált.⁵⁹ A semleges cisztein próbageometriáit három tengely (C–C, C–N és C–O) forgatásával illetve egy második megközelítésben az AMOEBA erőter segítségével állította elő. Ennek pontos mikéntjét nem ismerteti, csak annyit, hogy néhány száz darab volt belőlük. Végezetül különböző DFT funkcionálokkal (B3LYP, B97-D és M06-2X) 87 geometriát azonosított. Noguera és munkatársai a protonálódott glicin geometriák mellett a cisztein protonálódását is vizsgálták.⁴⁵ A protonált cisztein szerkezetek meghatározásához a glicin és protonált glicin konformereit használták fel, helyettesítve az egyik hidrogént a –CH₂–SH oldallánccal. Összesen hat, amino-protonált szerkezet található meg a cikkben, illetve megemlítik, hogy a karbonil-protonált geometriák relatív energiája legalább 30 kcal/mol-lal magasabb az amino-protonáltaknál, ezért azokat nem vizsgálták tovább (a konformer(ek)e)t, amely(ek)ből

levonták ezt a következtetést, nem közölték). Bouchoux a semleges formánál közölt módon a protonált cisztein szerkezeteket is próbálta megkeresni, 21 darabot talált, a protonálódás helyét nem részletezte.⁵⁹ Ezek közül hat darab geometriája található meg a cikkhez tartozó ESI-ben, ezek mind N-protonáltak.

Dobrowolski és munkatársai 2007-ben egy hibrid, elméleti és kísérleti munkát is közöltek a cisztein konformereiről.⁶⁰ A nyolc legstabilabbnak talált cisztein konformert B3LYP/aug-cc-pVDZ és MP2/aug-cc-pVTZ szinten is geometriai optimalásnak vetették alá, illetve meghatározták a hozzájuk tartozó harmonikus és anharmonikus rezgési frekvenciákat. Ezen túl (noha ilyen méréseket nem végeztek), az elmélet segítségével a rezgési cirkuláris dikroizmus spektrumát is megadták az egyes konformereknek. Ez a módszer a királis molekulák azonosításának egy speciális válfaja.⁶¹ Később argon és nitrogén mátrixban is rögzítették a cisztein infravörös spektrumát alacsony hőmérsékleten. A kétféle mátrixban hasonlóan összetett spektrumot kaptak, annak alapos elemzésével, és a számolt értékek felhasználásával legalább három, valószínűleg hat, vagy több minimum szerkezet létezésére találtak bizonyítékot. 2008-ban Alonso lézerablációs, molekula nyalábos Fourier-transzformált mikrohullámú spektroszkópiával (LA-MB-FTMW) tanulmányozta a cisztein konformereit.⁶² A korábban publikált cisztein konformerek közül a 11 legalacsonyabb relatív energiájút MP2/6-311++G** szinten optimalta, majd meghatározta a mikrohullámú spektroszkópiához tartozó paramétereit. A kísérleti eredményekkel összevetve hat darab konformert tudott egyértelműen azonosítani, illetve kettő, várt szerkezet hiányát a relaxációval magyarázta. 2014-ben Tarczay és munkatársai mátrix-izolációs infravörös spektroszkópia segítségével próbálták feltérképezni a cisztein szerkezeteit.⁶³ A felvett spektrumok egyszerűsítésének érdekében NIR besugárzással változtatták a konformerek eloszlását, a glicinnél alkalmazottakhoz hasonlóan. A korábban publikált IR spektroszkópiás eredményeknél pontosabb eredményeket kaptak, a mikrohullámú spektroszkópiával meghatározott hat szerkezetből ötöt ők is ki tudtak mutatni, illetve azon felül még egy, eddig még nem kimutatott, nagyon rövid felezési idejű konformert. Sajnos a protonált cisztein monomerről nem állnak rendelkezésre publikált mérések, a témához legközelebb álló cikkben a tömegspektrumon egy protonált cisztein dimert tudtak azonosítani, ami vélhetően kénhidas kötéssel, és a két aminocsoport között egy protonnal van összekötve.⁶⁴

4.3. Szerin

1995-ben Gronert és O'Hair a cisztein mellett a szerin konformereit is keresték.⁵⁵ A 324 kiindulási szerkezetből az AM1 szintű optimalizálások után 73 egyedi geometriát kaptak, azonban ez a szám további optimalizálások után (HF/3-21G és HF/6-31G) 51-re redukálódott. 2005-ben Jarmelo⁶⁵ a megfelelő kötések mentén forgatva szintén 324 (HF/3-21G* szinten) optimalizálandó szerkezetet tudott felírni. A számolások elvégzése után 71 minimumot kapott, amelyekből B3LYP/6-311++G** szinten már csak 61 maradt. Még ugyanebben az évben Miao és munkatársai ugyancsak 324 próbageometriát állítottak elő és ezeket B3LYP/6-31G*, majd az így kapott konformereket B3LYP/6-311+G** szinten optimalizálták.⁶⁶ Végezetül 74 stabil minimumot tudtak azonosítani. Bouchoux 2011-es cikkében a szerinről is találunk eredményeket a cisztein mellett, azonban ahogy korábban is említettem, a próbageometriák előállításának módja nem teljesen egyértelmű.⁵⁹ Összességében 74 stabil minimumot talált, amelyből csupán a néhány legmélyebbet vizsgálta tovább. Az irodalomban eddig fellelhető legátfogóbb szisztematikus feltérképezése a szerin konformereinek 2016-ból származik, ekkor már jóval nagyobb adathalmaz feldolgozása vált lehetővé.⁶⁷ Különböző megfontolások mellett a flexibilis funkciós csoportok forgatásával összesen 15552 kezdeti szerkezetet állítottak elő (hat tengely körüli forgatással, azonban a cikk alapján elképzelhető, hogy ezek közül kettő azonos), amelyeket HF/6-31G* elméleti szinten optimalizáltak. Ez 89 egyedi geometriát eredményezett, amikből magasabb, MP2/cc-pVTZ szinten már csak 85 darab maradt. A tizenkét legalacsonyabb relatív energiájú szerkezetet további számításoknak vetették alá, a szerkezetüket pedig MP2/aug-cc-pVTZ szinten is meghatározták. Ismét Noguera publikációjára kell hivatkozzak, ha a protonált szerinről szeretnék beszélni.⁴⁵ A próbageometriák elkészítéséhez ezúttal a semleges illetve protonált glicin konformerekre értelemszerűen $-\text{CH}_2-\text{OH}$ oldalláncot helyettesített be. Hat darab (amino-)protonált minimumot talált, a karbonil-protonálódást a ciszteinhez hasonlóan megemlíti (a kapott relatív energiabeli különbség ekkor 33 kcal/mol), azonban a szerkezetet nem közli. Miao és társai a semleges forma mellett (többek között) a protonált szerin konformereit is vizsgálták.⁶⁶ A protonált formáknál említést tesznek az O-protonálódás lehetőségéről, de a hozzátartozó geometriákat Noguera cikke⁴⁵ alapján túl magas energiájuknak gondolták, így csak az N-protonálódással foglalkoztak. A PES feltérképezése után 14 amino-protonált minimumot közöltek, B3LYP/6-311+G** elméleti szinten optimalizálva. Ismét Bouchoux 2011-es cikkéhez visszakanyarodva, 13 különböző, N-protonált szerkezet beazonosításáról olvashatunk.⁵⁹

A semleges szerin szerkezeteivel kapcsolatban is jelentek meg kísérleti eredmények. 2004-ben Lambie és munkatársai⁶⁸ újraoptimálták az addig irodalomban megjelent szerin konformereket B3LYP/6-31++G** elméleti szinten, valamint elvégezték a rezgési analízisüket. Az így kapott spektrumok alapján négy alacsony energiájú szerin szerkezet jelenlétét tudták kimutatni mátrixizolációs Fourier-transzformációs infravörös spektroszkópiás módszerrel, 15 K-en, argongázban. Egy évvel később Jarmelo csoportja hasonlóan járt el,⁶⁵ a szerin PES-ét elemezve a talált minimum szerkezeteket szintén B3LYP/6-31++G** szinten optimalta és meghatározta a harmonikus rezgési frekvenciákat. Mátrixizolációs infravörös spektroszkópia méréssel (szintén argon mátrixban) kilenc különböző konformert azonosítottak. Az Alonso csoport a szilárd szerint lézerablációs technikával juttatta gázfázisba, majd felvette a Fourier-transzformált mikrohullámú spektrumát.⁶⁹ Az elméleti úton számolt forgási és kvadrupól csatolási állandók segítségével hét konformer megjelenésére mutattak bizonyítékot. Akárcsak a glicinre és a ciszteinre, úgy a szerinre is végzett méréseket Tarczay egy nemzetközi kollaboráció keretében.⁷⁰ Tizennégy alacsony relatív energiájú szerkezetnek kiszámolták a rezgési spektrumát és az infravörös intenzitásait. Ezután vákuumkamrába párologtattak szerint, ahol argonnal és nitrogénnél keverve mátrixizolációs infravörös spektroszkópiás méréseket végeztek, illetve NIR besugárással megváltoztatták a konformerek eloszlását. Összesítve a tizennégyből kilenc szerkezetet azonosítottak, illetve úgy gondolják, hogy egy másik, rövid felezési idejű konformer is jelen volt, azonban az ehhez tartozó geometria meghatározása nem volt lehetséges. A protonált szerin monomerről egy, a protonált glicinnél is hivatkozott kísérleti cikk található meg.⁴⁸ IRMPD spektroszkópiás mérések elvégzésével a két legalacsonyabb relatív energiájú N-protonált szerkezet volt kimutatható (SH01 és SH02). A legújabb tanulmányok már nem monomer, hanem dimer, vagy oktamer protonált szerinről szólnak.^{71,72}

4.4. Protonaffinitás és gázfázisú bázicitás

A protonaffinitás és gázfázisú bázicitás témájában számtalan elméleti és kísérleti cikk is született mind a glicin^{24,32,44,45,47,49,55,73-93}, mind a cisztein^{24,32,45,59,73,77,78,80,82,84,85,89,90,92-98}, mind pedig a szerinrel^{24,32,45,59,66,73-75,77-79,82,84,85,88-90,92,93} kapcsolatban. A mérési módszerekkel kapcsolatban a Bevezetés fejezetben írtam részletesebb ismertetőt. Az elméleti munkák hasonló elvet követnek, mint a mi munkánk: meghatározzák a semleges és a protonált forma geometriáját és relatív energiáját, majd ebből protonaffinitást és gázfázisú bázicitást számolnak. A számítástechnika rohamos fejlődése azonban lehetővé tette, hogy az

eddig publikációkban közzétett eredményeket jelentősen pontosítsuk, meghaladjuk. A rendszerek méretéből következően a konformerek kezelése általában DFT módszerekkel történik. A másik, hullámfüggvény alapú számolások oldaláról általában MP2 geometriájú, elvértve csatolt-klaszter (CC) single-point számolásokkal kapott relatív energiákkal találkozhatunk. Amennyiben magasabb szintű geometria is rendelkezésre áll, akkor csak a globális minimum geometriákból számolták ki a termodinamikai értékeket, azonban a gyakorlatban a konformerelegyek jelenléte a jellemző. Az utolsó pont, amiben még fejlődés érhető el, az a különböző protonálódási helyek vizsgálata: ugyan néhány publikáció említést tesz az aminocsoport mellett a karbonilcsoport protonálódásáról, azonban mivel az előbbi a termodinamikailag kedvezőbb, az utóbbira nem történtek kiterjedt vizsgálatok. Az energiaviszonyok felmérése mellett azért is érdekes az összes protonálódási hely elemzése (a karbonilcsoport mellett a nagyobb aminosavak oldalláncán is történhet protonálódás), mert sokszor az aminocsoport „nem áll rendelkezésre” a protonálódás szempontjából. Nem is kell bonyolult vegyületekre gondolnunk, elég, ha csupán aminosav dimerekre, oligomerekre gondolunk, amelyekben az O-terminális is egyre nagyobb hangsúlyt kap, illetve a szabad oldalláncok is. Ugyan a kémiai környezet ekkor megváltozik az adott funkciócsoport körül (illetve a kialakuló amidcsoportok is protonálhatóvá válnak), azonban a PA és GB értékekben ez vélhetően nem okoz drasztikus változást, de ennek vizsgálata további tudományos munkák feladata.

A kapott eredményeink későbbi, az irodalommal való összehasonlításához konkrét számadatakra van szükségünk. Szerencsére a bőséges mennyiségű elérhető protonaffinitás (és gázfázisú bázicitás) adatokból kiterjedt összefoglaló cikkek készültek. 1997-ben Harrison közölt egy cikket az aminosavak és peptidek PA és GB értékeiről,⁷⁸ ebben az addig elérhető kísérleti eredményeket foglalta össze. Azt találta, hogy a meghatározott GB értékek általánosságban jó egyezést mutatnak, azonban az ezekből számolt protonaffinitás már nem olyan jól definiált, az átszámításhoz szükséges entrópia járulék bizonytalansága miatt. Egy évvel később Hunter és Lias átfogó elemzést publikált különböző molekulák, gyökök és atomok PA és GB értékeiről és újragondolt kiértékeléséről.³² A csaknem 1700 kémiai specieszt tartalmazó gyűjteményben természetesen az aminosavakat is megtaláljuk. A publikáció jelentőségét jól mutatja, hogy rengeteg hivatkozás érkezik rá napjainkig is, illetve az amerikai egyesült államokbeli Nemzeti Szabványügyi és Technológiai Intézet (National Institute of Standards and Technology, NIST)⁹⁹ is számos vegyület esetén, köztük például a glicin és a cisztein, az általuk megadott adatokat tekinti meghatározónak. Bouchoux és

munkatársai számos munkával hozzájárultak a termokémia ezen területéhez, most azonban csak az összefoglalásukat tartalmazó publikációt emelném ki. Egy több részből álló sorozatban Hunter és Lias összefoglalásához hasonlóan összegyűjtötte és összevetette az irodalomban megtalálható kísérleti és elméleti protonálódáshoz kapcsolódó termodinamikai értékeket és elemezte azokat. A művek számunkra releváns, harmadik része éppen az aminosavakkal foglalkozott, az ott közölt értékek mind a kísérleti, mind az elméleti munkákkal összevethetővé teszik az eredményeinket.⁹⁰

5. CÉLKITŰZÉS

A gázfázisú aminosavak szerkezetének, és protonálódásának mind a kísérleti, mind az elméleti vizsgálata hosszú múltra tekint vissza. A számítástechnika, a számítógépes kapacitások robbanásszerű növekedése azonban lehetővé teszi, hogy az eddigi elméleti munkákat meghaladva, minden eddiginél pontosabban és részletesebben vizsgálhassuk az aminosavakat és a hozzájuk tartozó protonálódási folyamatokat. A doktori munkámhoz kapcsolódóan kitűzött feladataimat a következő pontokban tudom összefoglalni:

- I. A lehető legrészletesebb, szisztematikusan feltérképezése a glicin, a cisztein és a szerin aminosavak konformációs terének a HF módszeren túl. A lehetséges protonálódási helyek vizsgálata a kapott minimumokon, különös tekintettel az irodalomban még kevésbé elterjedt, termodinamikailag kedvezőtlenebb funkciós csoportokra. A protonált formákhoz tartozó stabil konformerek meghatározása.
- II. Az eredmények további korrigálására a legmélyebb energiájú struktúrákhoz kapcsolódóan, a kémiai pontosságot eredményező csatolt-klaszter módszer alkalmazása mind a geometriák, mind a harmonikus rezgési frekvenciák meghatározásához. A kapott értékek további pontosítására és a kvantumkémiai elhanyagolásokból eredő hibahatárainak becslésére különböző korrekciók kiszámítása: poszt-(T) korrekciók, törzselektron korrelációs korrekciók, másodrendű Douglas–Kroll relativisztikus korrekciók, zérusponyi energia-korrekciók.
- III. A glicin, a cisztein és a szerin protonaffinitásának és gázfázisú bázicitásának nagy pontosságú, elméleti meghatározása CC szinten, a II. pontban említett korrekciókat is figyelembe véve. A termodinamikai mennyiségek kiszámítása az összes lehetséges protonálódási helyre vonatkozóan, konformerelegyeket figyelembe véve.

6. ELMÉLETI HÁTTÉR

Az atomok, molekulák, vagy akár kémiai reakciók elméleti vizsgálatához a kvantumkémia eszköztárát kell alkalmaznunk. Az egyik alapegyenlet, amiből kiindulhatunk az az időfüggetlen, nem-relativisztikus Schrödinger-egyenlet:

$$\hat{H}\Psi = E\Psi, \quad (6.1.)$$

ahol Ψ a hullámfüggvény, E a teljes energia, míg \hat{H} a teljes energia, vagy más néven Hamilton-operátor. Ez egy sajátérték egyenlet, melynek analitikus megoldása néhány rendszertől eltekintve lehetetlen, numerikusan pedig a rendszer méretét (atomok és/vagy elektronok számát) növelve rohamosan nő a számítási igény is. A könnyebb kezelhetőség érdekében különböző közelítéseket vezettek be, amelyekkel jól leírható a vizsgált kvantumkémiai rendszer. Az első ilyen, a Born–Oppenheimer közelítés, mely azt mondja ki, hogy mivel az atommagok tömege 3–5 nagyságrenddel nagyobb, mint az elektronoké, így mozgásuk különválasztható. Ezáltal felírhatóvá válik az elektronokra vonatkozó Schrödinger-egyenlet egy adott magkonfigurációnál:

$$\hat{H}_e \Psi_{(x_1, x_2, \dots, x_N)} = E_e \Psi_{(x_1, x_2, \dots, x_N)}, \quad (6.2.)$$

ahol \hat{H}_e az elektron Hamilton-operátor, x_i az i -edik elektron koordinátája, míg N az elektronok száma. \hat{H}_e három részre bontható, attól függően, hogy az egyes tagok hány elektron koordinátájától függenek:

$$\hat{H}_e = \hat{H}_0 + \hat{H}_1 + \hat{H}_2, \quad (6.3.)$$

ahol \hat{H}_0 nulla-elektron operátort (mag-mag taszítás tag), \hat{H}_1 az egy-elektron operátorokat (elektron kinetikus energia operátor és a mag-elektron vonzás tag) illetve a \hat{H}_2 két-elektron operátort (elektron-elektron taszítás) tartalmazza. Ez utóbbi tag összezsacatolja az elektronok mozgását, amit szeretnénk feloldani. Ezt úgy tudjuk, ha behelyettesítünk egy átlagolt potenciáalteret, vagyis azt feltételezzük, hogy minden elektron a többi elektron átlagolt terében mozog. Ekkor a Hamilton-operátor felírható összeg alakban, a hullámfüggvényt kereshetjük szorzat alakban:

$$\Psi_{(x_1, x_2, \dots, x_N)} \approx \varphi_{1(x_1)} \cdot \varphi_{2(x_2)} \cdot \dots \cdot \varphi_{N(x_N)}, \quad (6.4.)$$

ahol $\varphi_{i(x_i)}$ az i -edik egyelektron pályafüggvény. Noha ez a szorzat az effektív egyelektron operátorok összegével felírható Hamilton-operátor sajátállapota, nem elégíti ki a

szimmetrizálási posztulátumot, így determináns alakban kell felírunk a hullámfüggvényt (Slater-determináns):

$$\Psi_{(x_1, x_2, \dots, x_N)} \approx \frac{1}{\sqrt{N!}} \begin{bmatrix} \varphi_1(x_1) & \cdots & \varphi_1(x_N) \\ \vdots & \ddots & \vdots \\ \varphi_N(x_1) & \cdots & \varphi_N(x_N) \end{bmatrix}. \quad (6.5.)$$

A soraiban az egyes egyrészesecske függvények változnak, míg az oszlopaiban a részesecskeindexek.

A Hartree–Fock (HF)¹⁰⁰ közelítés egy olyan független részesecske közelítés, amelyben egy Slater-determináns alakú hullámfüggvénnyel kiszámítva a teljes Hamilton-operátor várható értékét az a rendszer energiáját a lehető legjobban megközelíti, a legmélyebb értéket adja. A definícióból nyilvánvalóan következik, hogy egy variációs megoldásra lesz szükség. A variációs elv értelmében bármilyen próbafüggvénnyel kapott energia magasabb (vagy egyenlő) lesz, mint az egzakt alapállapot energiája. Az energia funkcionál, amelyet minimalizálnunk kell, a következőképpen néz ki:

$$E[\Psi] = \langle \Psi | \hat{H}_0 | \Psi \rangle + \langle \Psi | \hat{H}_1 | \Psi \rangle + \langle \Psi | \hat{H}_2 | \Psi \rangle. \quad (6.6.)$$

A Slater-Condon szabályokat alkalmazva a mátrixelemek kiszámításánál ezt az alakot kapjuk:

$$E[\Psi] = H_0 + \sum_{i=1}^N H_{ii} + \sum_{i>j}^N (J_{ij} - K_{ij}). \quad (6.7.)$$

A H_0 tag a mag-mag kölcsönhatási tag, egy konstans (mivel rögzített a magkonfiguráció), H_{ii} az egyelektronos tagok összege, J_{ij} a Coulomb-integrál és pedig a K_{ij} kicserélődési-integrál. Ezt a kifejezést kell variálnunk, úgy, hogy mellékfeltételben kikötjük (Lagrange-féle multiplikátoros módszerrel), hogy az egyrészesecske függvények ortonormáltak. A műveletet elvégezve, levezetések után kapjuk a kanonikus Hartree–Fock egyenleteket:

$$\hat{f}_{(\varphi_1, \varphi_2, \dots, \varphi_N)} |\varphi_i\rangle = \varepsilon_i |\varphi_i\rangle. \quad (6.8.)$$

ε_i az i -edik pályához tartozó pályaeenergia, φ_i a keresett i -edik egyelektron függvény, míg \hat{f} a Fock-operátor, melynek argumentumában megjelennek a keresett egyelektron pályák. Mivel az operátor maga is függ a meghatározandó pályáktól, ezért az egyenlet iteratív módon oldható meg, valamilyen kezdeti pályafüggvényekkel felépítjük a Fock-operátort, meghatározzuk az új pályákat, amelyekből ismét felépítünk egy új operátort és így tovább, egészen a konvergencia eléréséig. Ezt az eljárást nevezzük önkonzisztens tér módszernek

(angolul self-consistent field, SCF). Az egyenlet megoldása végtelen sok pályához vezet ($i > N$), ezek közül csak az N darab, legmélyebb energiát adó megoldás kerül bele a Slater-determinánsba (betöltött pályák), míg a kimaradók a virtuális pályáink. Hartree, Fock, Roothan és Hall azzal az ötlettel álltak elő, hogy ezt a problémát transzformáljuk egy algebrai sajátérték problémára, vezessünk be véges bázist az egyrészecske állapotér egy alterének kijelölésével. Ekkor az atom- és molekulapályák ezen bázisfüggvények lineáris kombinációjával állíthatók elő. A gondolatmenetet követve az alábbi egyenlethez jutunk:

$$\mathbf{f}\mathbf{c} = \mathbf{S}\mathbf{c}\boldsymbol{\varepsilon}, \quad (6.9.)$$

ahol \mathbf{f} a Fock-mátrix, \mathbf{c} az együtthatók mátrixa, $\boldsymbol{\varepsilon}$ a pályae energiákat tartalmazó diagonális mátrix és \mathbf{S} az átfedési-mátrix. A probléma megoldása itt is iteratíván történik, mivel a Fock-mátrix felépítéséhez szükségünk van az együtthatók értékeire. A HF módszerrel kapott energiák azonban gyakran igencsak távol állnak a valóságtól az elhanyagolások miatt. Elektronkorrelációnak hívjuk az egzakt és a legmélyebb Hartree–Fock megoldás (HF-limit) közötti különbséget, amely egyrészt a matematikai közelítésekből (statikus korreláció), másrészt pedig a független részecske közelítésből (dinamikus korreláció) ered. Az eredmények pontosítására további módszerek kerültek bevezetésre, ezek közül a legelterjedtebbeket — amelyeket én is alkalmaztam a munkám során — mutatom be.

A variációs eljárások mellett a következő fontos módszer a perturbációs számítás, több fajtája létezik, de az egyik legelterjedtebb a Rayleigh–Schrödinger perturbációs számítás. Az alapproblémánk az, hogy van egy Hamilton-operátorunk (\hat{H}), melynek a sajátállapotai (Ψ_i) és sajátértékei (E_i) nem ismertek:

$$\hat{H}|\Psi_i\rangle = E_i|\Psi_i\rangle. \quad (6.10.)$$

Viszont van egy hozzá nagyon hasonló, ismert megoldásokkal (φ_i sajátállapotok és ε_i sajátértékek) rendelkező másik Hamilton-operátorunk (\hat{H}_0), amivel közelítjük az ismeretlent:

$$\hat{H}_0|\varphi_i\rangle = \varepsilon_i|\varphi_i\rangle. \quad (6.11.)$$

Definiáljuk a két operátor különbségét, a perturbációt (\hat{W}):

$$\hat{W} = \hat{H} - \hat{H}_0. \quad (6.12.)$$

Sorfejtéses vagy algebrai levezetések után megkaphatjuk a különböző rendű korrekciókra vonatkozó képleteket mind az energiákra, mind a hullámfüggvényekre vonatkozóan.

Esetemben az alapállapotra vonatkozó formulák hasznosak, de természetesen gerjesztett állapotok is vizsgálhatóak. A kérdés az, hogy hogyan végezzük el a partíciónálást, vagyis hogyan válasszuk ki az ismert és az ismeretlen részeket. A Møller–Plesset (MP)¹⁰¹ perturbációs számítás értelmében a Hartree–Fock megoldásból indulunk ki, az ismert Hamilton-operátort a Fock-operátorokból állítjuk össze, míg a perturbáció ennek az egzakt Hamilton-operátortól való eltérése lesz. Mivel az egyelektron operátorok a két esetben megegyeznek, az eltérés a kételektron kölcsönhatás potenciáljának és az ezt közelítő potenciál különbségéből ered. Az alapállapot energiájának nulladrendű korrekciója ($E^{(0)}$) a HF pályae energiák (ε_i) összegével írható fel:

$$E^{(0)} = \sum_i \varepsilon_i. \quad (6.13.)$$

Az elsőrendű korrekcióra a következő képlet adódik:

$$E^{(1)} = \langle \varphi_0 | \widehat{W} | \varphi_0 \rangle. \quad (6.14.)$$

A perturbációs számítás első rendje még nem ad igazi korrekciót, hiszen a két értéket ($E^{(0)}$ és $E^{(1)}$) összeadva épp a HF energiát kapjuk vissza. Az első elektronkorrelációs energia járulékot a másodrendű korrekció fogja szolgáltatni, melyben már megjelennek a gerjesztett determinánsok is (ezeknek a Slater-determinánsoknak a felépítése során már virtuális pályákat is felhasználunk):

$$E^{(2)} = \sum_{k \neq 0} \frac{|\langle \varphi_0 | \widehat{W} | \varphi_k \rangle|^2}{\varepsilon_0 - \varepsilon_k}. \quad (6.15.)$$

A Brillouin tétel és a Slater–Condon szabályok értelmében csak azok a tagok nem lesznek nullák, amelyekben kétszeresen gerjesztett determinánsok szerepelnek, vagyis két virtuális pályát helyettesítettünk be. Fontos elismételni, hogy a perturbációs nem variációs módszer, azaz az eredményünk az egzakt megoldás alá is mehet. Sok rendszeren elvégzett számítások gyakorlati megfigyelése, hogy a másodrendben korrigált energia (MP2) még felette marad az egzakt értéknek, a harmadrendű forma alámegy (MP3), amit a negyedrendű korrigál (MP4). A gyakorlatban az MP2 módszer a legelterjedtebb perturbációs módszer, a kedvező számításigénye miatt nagy molekulákra is alkalmazható.

A csatolt-klaszter módszer (coupled-cluster, CC)^{102,103} bevezetése előtt néhány mondatban bemutatom a konfigurációs kölcsönhatás módszerét (configuration interaction, CI), amelyen a CC alapszik. Az alapötlet az, hogy válasszuk ki egy referencia Slater-determinánst (lehet a HF alapállapot, de nem csak és kizárólag, illetve némely rendszer esetén több referencia

determináns válik szükségessé), majd ebből állítsuk elő az 1-, 2-, 3-, ..., N -szeresen gerjesztett determinánsokat (φ_i^a , φ_{ij}^{ab} , ...) és ezek lineáris kombinációjával (együtthatóik: c_i^a , c_{ij}^{ab} , ...) állítsuk elő az n -edik állapothoz tartozó hullámfüggvényt:

$$\Psi_n = \varphi_0 + \sum_i \sum_a c_i^a \varphi_i^a + \sum_i \sum_{j>i} \sum_a \sum_{b>a} c_{ij}^{ab} \varphi_{ij}^{ab} + \dots \quad (6.17.)$$

További teendő az együtthatók meghatározása, melyekhez egy lineáris egyenletrendszeren keresztül jutunk el. Azt tapasztalták, hogy ugyan az első néhány tagnak a járuléka sokkal nagyobb, mint a későbbieknek, viszont az együtthatók csak együttesen határozhatóak meg. A rendszer méretének növekedésével a megoldás gyakorlatilag matematikailag kezelhetetlenné válik, pedig a teljes sorfejtés az egzakt megoldást adná. Ezt kiküszöbölve megszakíthatjuk a sorfejtést valahányszorosan gerjesztett determinánsoknál, viszont ez azt fogja eredményezni, hogy a módszer méret inkonzisztens lesz. Az eredeti elgondolást követve került bevezetésre a CC módszer: azoknak a gerjesztéseknek, amelyek előfordulnak a kifejtésben, szerepeljen az összes lehetséges szorzata is. Ez már orvosolni fogja a méret konzisztencia problémát, a hullámfüggvényt a következő alakban állítjuk elő:

$$\Psi_n = e^{\hat{T}} \varphi_0, \quad (6.18.)$$

ahol $e^{\hat{T}}$ a hullám operátor, míg \hat{T} a klaszter operátor, amelyben \hat{T}_N az N -szeres gerjesztéseket előállító operátor:

$$e^{\hat{T}} = \hat{1} + \hat{T} + \frac{1}{2} \hat{T}^2 + \frac{1}{3!} \hat{T}^3 + \dots, \quad (6.19.)$$

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots + \hat{T}_N, \quad (6.20.)$$

$$\hat{T}_1 = \sum_i \sum_a t_a^i \varphi_a^i, \quad (6.21.)$$

$$\hat{T}_2 = \sum_a \sum_{a>b} \sum_i \sum_{i>j} t_{ab}^{ij} \varphi_{ab}^{ij}. \quad (6.22.)$$

Ebben az esetben a klaszter operátorokban megjelenő t_i^a , t_{ij}^{ab} , ... együtthatókat kell meghatároznunk, amelyek neve amplitúdó. Mivel a kifejtésben ezen amplitúdók szorzata is szerepel, itt már a CI-vel ellentétben nem lineáris egyenletrendszert kapunk, ezeket projektált csatolt klaszter egyenleteknek hívjuk. Mivel itt is drasztikusan nő a számítási igény a rendszer méretének és a figyelembe vett gerjesztések számának növekedésével, ezért ebben az esetben is alkalmaznunk kell a megszakított sorfejtést, ám ezúttal megtehetjük ezt a méret konzisztencia probléma megjelenése nélkül. Ha a kétszeres gerjesztésekig hajtjuk végre a sorfejtést, akkor jutunk CCSD (coupled-cluster singles and doubles) módszerhez,

azonban a gyakorlatban inkább a CCSD(T) (coupled-cluster singles, doubles and perturbative triples) módszer az elterjedt, amely a háromszoros gerjesztéseket perturbációs módon veszi figyelembe, a pontosság elérheti a spektroszkópiai értékeket, azonban mégis jóval kisebb a számítási igénye, mint a CCSDT módszernek. Ennél még pontosabb a CCSDT(Q) (coupled-cluster singles, doubles, triples and perturbative quadruples) eljárás, amely a háromszoros gerjesztésekig nem tartalmaz elhanyagolást, a négyszeres gerjesztéseket pedig perturbatív eljárással vonja be. Továbbá bevezetésre kerültek az úgynevezett explicit-korrelált (F12)¹⁰⁴⁻¹⁰⁶ módszerek, amelyekben az $r_{1,2}$ elektron-elektron távolságot explicit módon beleépítjük a hullámfüggvénybe. Ezeket mind az MP2, mind a CCSD(T) módszerekkel szabadon kombinálhatjuk. Általánosságban elmondható, hogy ezek a kiegészített módszerek sokkal robusztusabbak, kevésbé érzékenyek arra, hogy milyen bázist választunk, illetve a báziskészlet-konvergencia is gyorsabbá válik.

A kvantumkémiai számolásokhoz a módszerek mellé megfelelő bázisokat is kell választanunk, amelyből felépítjük az pályáinkat, ezek általában két részből, egy radiális és egy szögfüggő részből állnak. A legriválisabb módszer, az LCAO-MO módszer (linear combination of atomic orbitals to molecular orbitals), amelyben a molekulákat leíró bázist a molekulát alkotó atomok atompályáiból állítjuk elő lineáris kombinációval. A hidrogén pályáit könnyen ki tudjuk számolni, de a többi atom esetében már nem tudunk ilyen pontosan eljárni, közelítő függvényeket kell alkalmaznunk. Az egyik ilyen nagy függvénycsalád a Slater típusú pályák (Slater-type orbitals, STO), melynek radiális része $e^{-\alpha r}$ tagot tartalmaz, a szögfüggő része hidrogénszerű gömbfüggvényekből épül fel. A kételektron integrálok kiszámításának egyszerűbbé tétele miatt bevezették a Gauss függvényeket (Gauss-type orbitals, GTO), amelyek szögfüggő részét egy Descartes-i koordinátáktól függő tényezővel helyettesítik, míg a radiális részben $e^{-\alpha r^2}$ jelenik meg. STO-k esetén az atommag helyén a derivált az exponens által meghatározott nemnulla érték, míg GTO-k esetén a függvény $r = 0$ -ban zérus. További különbség még, hogy az atomtól távoli tartományokban a GTO sokkal hamarabb lecseng, mint az STO. Azonban a számítások során a GTO-k hatalmas előnnyel rendelkeznek: bármely két Gauss-függvény szorzata egy Gauss-függvény, illetve minden integrálja kiszámítható analitikusan, vagy visszavezethető egy numerikusan könnyen kezelhető problémára. A Gauss-függvény alakjából eredő pontatlanságok javítására vezették be a kontrahált bázisfüggvények alkalmazását. Ezek a bázisfüggvények nem egyetlen Gauss-függvényből állnak, hanem többnek a lineáris kombinációjával

állíthatóak elő, a koefficiensek előre meghatározottak és rögzítettek. Ilyen például az STO- n G bázis, amelyben a Slater-típusú pályát n darab Gauss-függvényből alkotjuk meg.

A bázisokat úgy osztályozhatjuk, hogy egy-egy atompálya előállítására hány kontrahált vagy nem kontrahált bázisfüggvényt használunk. Ha egyet, akkor minimál vagy egyszeres zéta (single zeta), ha kettőt, akkor kétszeres zéta (double zeta), ha hármat, akkor háromszoros zéta (triple zeta) és így tovább bázisokról beszélhetünk. Pople és munkatársainak a nevéhez kötődnek a felhasított-vegyértékű (split valence) bázisok, amelyekben a belső héjakat alkotó pályákat egyetlen kontrahált bázisból állítjuk elő, míg a vegyértékűket alkotó pályákat többször.¹⁰⁷ Például a 3-21G bázis azt jelenti, hogy a törzselektronokhoz tartozó pályákat egyetlen, három Gauss-függvényből álló kontrahált függvény írja le, míg a vegyérték elektronok pályáit kettő darab, az egyik kettő, míg a másik csak egy primitív függvényből áll. Ezeket ki szokták még egészíteni polarizációs- illetve diffúziós-függvényekkel. A polarizációs-függvények alkalmazásakor olyan pályákat is figyelembe veszünk, amelyekhez nagyobb mellékkvantumszám tartozik, mint a betöltött pályákhoz. Ha csak nehéz atomokra alkalmazzuk, jelölése *, viszont ha hidrogénre is, akkor **. A diffúziós-függvények jelölése ezzel analóg, nehéz atomok esetén +, míg ha hidrogénre is alkalmazzuk, akkor ++. Ez utóbbiak lassan lecsengő függvények, alkalmazásuk hosszú távú kölcsönhatások leírásakor (H-kötés, van der Waals kölcsönhatások, molekula-komplexek) és anionok (kiterjedt elektronfelhő) vizsgálatokor célszerű.

Modernebb, korreláció-konzisztens (correlation-consistent, cc) bázisok Dunning nevéhez kötődnek, melyeket a korreláció-konzisztens módszerekhez fejlesztettek, de persze HF számolásokat is elvégezhetünk velük.¹⁰⁸ Ezek már alapértelmezetten tartalmazzak polarizációs függvényeket. Nevezéktanuk az általános cc-pV n Z alakkal írható fel, mely azt jelenti: korreláció-konzisztens, polarizált, n -szeres ζ vegyérték bázis. Természetesen ezek mellé is van lehetőség diffúziós-függvényeket alkalmaznunk, ennek jelölése az aug- előtag.¹⁰⁸ A korábban tárgyalt explicit-korrelált módszerek mellé továbbfejlesztették a bázisok ezen változatát, amelyekkel még pontosabb eredmények és gyorsabb konvergencia érhető el. Jelölésük cc-pV n Z-F12 és már tartalmazzak diffúziós-függvényeket is beépítve.¹⁰⁹

A kvantumkémiai számításokat a Czákó csoportban rendelkezésre álló számítógépeken futtattam, a MOLPRO¹¹⁰ és (a MOLPRO-hoz implementált) MRCC^{111,112} programok segítségével. A kapott eredmények elemzése általam írt Fortran95 illetve Python programokkal történt.

7. EREDMÉNYEK ÉS DISZKUSSZIÓ

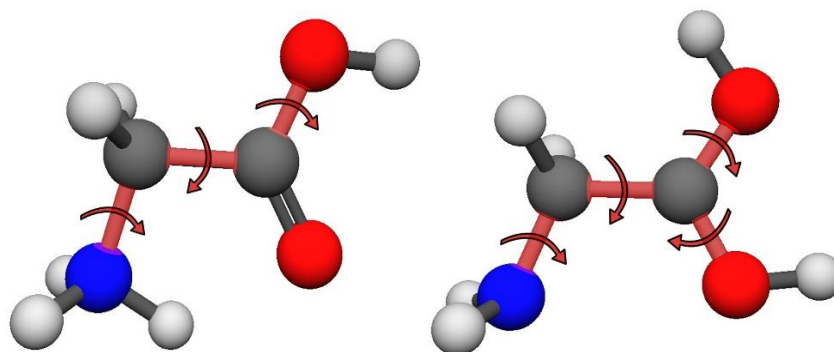
7.1. KONFORMEREK

A következő alfejezetekben a glicin, a cisztein és a szerin aminosavak semleges és protonált formáinak meghatározását ismertetem. Első lépésként a semleges formák konformációs térét térképeztük fel egy szisztematikus eljárás segítségével. Ehhez a molekulákat olyan kötéstengelyek mentén fogattuk periodikusan (60° -onként), amelyek a belső torziós mozgást megfelelően leírják. Az így kapott konformereket a keresés elméleti szintje után a pontosság szempontjából már megbízhatóbb elméleti szinten is optimáltuk, a harmonikus rezgési analízisükkel pedig megbizonyosodtunk, hogy valóban minimum szerkezetek. A protonált konformerek meghatározásához első lépésben a lehetséges protonálódási helyek validálására van szükség. Ehhez a semleges aminosavaknál talált geometriák felhasználásával próbaoptimalásokat végzünk úgy, hogy a különböző funkcióscsoportokhoz egy protont „kötünk”. Az ebből levont következtetések alapján pedig a semleges molekulákhoz hasonló szisztematikus eljárással meghatároztuk a különböző csoportokon protonálódott konformereket. A konformációs tér leírásához szükséges forgatási tengelyek száma a protonált esetben meg is nőhet, ezért ekkor különösen fontos a szimmetria tényezők figyelembevétele, amelyekkel a próbageometriák száma redukálható. A kapott geometriák szétválasztása három vizsgálaton alapszik. Az első, a relatív energiák összehasonlítása (a különbség legalább $0,01$ kcal/mol legyen), amely a glicinnél nagyobb aminosavak esetén nem volt önállóan módszer, hiszen a ciszteinnél és a szerinnél kapott nagyszámú konformer adathalmazában előfordulhat, hogy azonos, vagy nagyon közeli relatív energia tartozik különböző geometriákhoz. Az energia mellé ezért valamilyen geometriai összehasonlítást is be kell vezetnünk, ez történhet sokféleképpen, mi azt a feltételt szabtuk a különbözőségekre, hogy az A forgási állandó legalább $3 \times 10^{-4} \text{cm}^{-1}$ -ben legyen eltérő. Az utolsó dolog, amit még figyelembe kell vennünk, ha a különböző funkcióscsoportokon protonálódott próbageometriákat optimaljuk, az is előfordulhat, hogy a folyamat során a proton „átkerül” az egyik csoportról a másikra, hiszen energetikailag az kedvezőbb. Ennek elemzésére, és hogy a protonálódott konformereket megfelelő csoportba soroljuk, bevezettük a kötésrendmátrixot, amellyel megszabtuk, hogy mely atomok között feltételezünk kötést egy bizonyos csoportba való besoroláshoz (ezt alapértelmezetten a semleges aminosavak esetén is alkalmaztuk végül).

7.1.1. Gly és GlyH⁺: konformerek¹

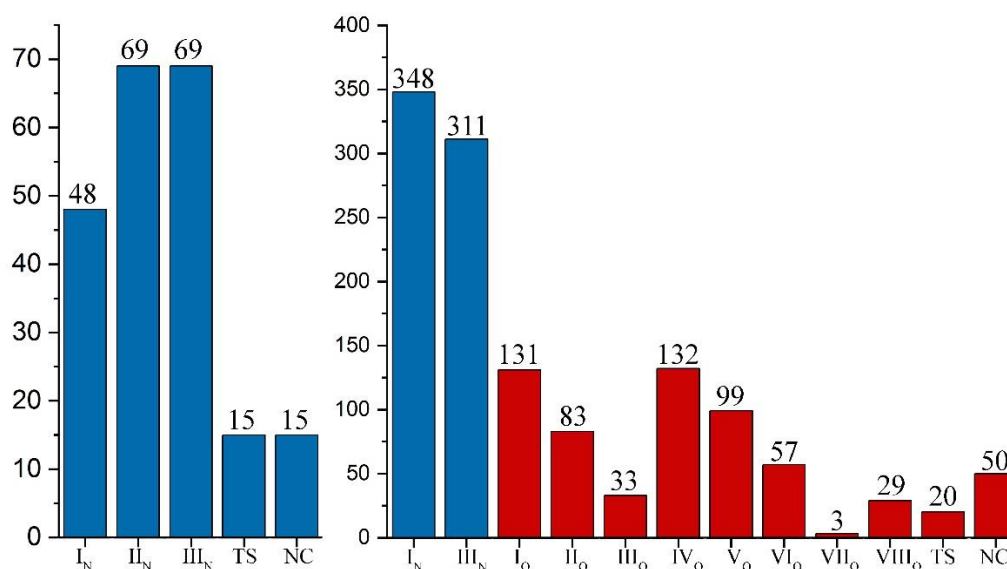
Mivel a glicin minimum konformerei az irodalomban jól ismertek,^{4,33,36} többek között mi is vizsgáltuk őket egy korábbi publikációnk kapcsán,⁴ ezért ezen aminosav esetén csak a protonált formák feltérképezése szükséges. Ennek első lépéseként a lehetséges protonálódási helyeket azonosítottuk be. A folyamat feltételezhetően lejátszódhat az amino-, a karbonil- és a hidroxilcsoportokon, a következő formákhoz vezetve: H₃N⁺-CH₂-COOH, H₂N-CH₂-C⁺(OH)₂ és H₂N-CH₂-CO(OH₂)⁺. Ennek megerősítésére vagy megcáfolására az irodalmi források elemzésén túl mi magunk is számításokat végeztünk különböző próbaszerkezeteken, amelyeket a semleges aminosav nyolc konformeréből állítottunk elő egy extra proton hozzáadásával a különböző funkciós csoportokon. Mindegyik csoport vizsgálatához kétszer nyolc próbaszerkezetet készítettünk, a két nyolc tagból álló sorozat közti különbség az újonnan kialakított funkciós csoport térbeli orientációjából adódik. Az aminocsoportoz tartozó próbaszerkezetekben az extra protont úgy helyeztük el, hogy a nitrogénatom körül a hidrogének (és a szén) közel tetraédres elrendeződésben legyenek, a második próbageometria sorozat esetén pedig a kialakított NH₃⁺ csoportot elforgattuk 60 fokkal. A karbonilcsoport protonálódásának vizsgálatánál az orientációs különbséget az adta, hogy az újonnan kialakított O-H az eredetihez képest *cisz* vagy *transz* állásban van, a hidroxilcsoport protonálódása pedig olyan szerkezetekhez vezet, amelyekben a két O-H kötés a molekula fősíkjában, vagyis az N-C-C síkban helyezkednek el, vagy arra merőlegesen. Az így kapott negyvennyolc (3 × 2 × 8) próbaszerkezetet MP2/aug-cc-pVDZ szintű geometriai optimalizációkat és harmonikus rezgési frekvencia számításokat futtattunk, hogy meggyőződhessünk, hogy tényleg minimum szerkezeteket kaptunk. Ugyan az amino- és a karbonilcsoport protonálódása stabil molekulaionokhoz vezetett, a hidroxilcsoporté már nem, az ehhez tartozó geometriák a C-OH₂ kötés néhány ångströmös megnyúlásával vízkation komplexekké alakultak.

A protonált glicin teljes konformációs terének feltérképezéséhez, vagyis az összes konformer megkereséséhez egy szisztematikus módszert alkalmaztunk, melynek vázlatát az 1. ábrán látható. A glicin I_p konformerén elhelyeztünk egy-egy protont az amino- (N-protonált szerkezet) vagy a karbonilcsoporton (O-protonált szerkezet) az ábra szerint, majd ebből a két szerkezetből elkészítettünk egy-egy geometriai készletet a belső torziós mozgások leírására. A megfelelő torziós szöveget 60 fokként változtattuk, ez torziós szögenként hat lépésből áll, mivel a 0° és a 360° ekvivalens szerkezetet eredményez. N-protonált esetben a hidroxil-, a karboxil- és a protonált aminocsoportot kell forgatni, így ösz-



1. ábra: Az N- (bal) és O-protonált (jobb) glicin konformációs terének szisztematikus feltérképezéséhez készített sematikus vázlat a forgatási tengelyek jelölésével.

szesen $6^3 = 216$ kiindulási geometria kapható, ez a szám a szimmetria figyelembevételével tovább redukálható. O-protonált esetben az amino-protonált esethez képest eggyel több funkciós csoportot kell forgatni, mivel megjelenik egy extra O–H, tehát $6^4 = 1296$ kiindulási szerkezetet kaptunk, ha a szimmetriát nem vettük figyelembe. A kiindulási geometriákat MP2/aug-cc-pVDZ szinten optimaltunk és kiszámítottuk a hozzájuk tartozó harmonikus rezgési frekvenciákat, hogy megbizonyosodjunk, valóban minimumokhoz tartoznak, nem pedig átmeneti állapotokhoz. A keresések eredményeit a 2. ábrán foglaltuk össze.



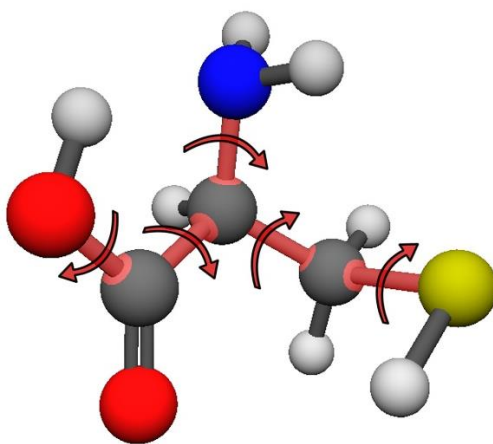
2. ábra: Az egyes próbageometriák MP2/aug-cc-pVDZ elméleti szintű geometriai optimalás utáni besorolása az eredetileg N- (bal) és O-protonált (jobb) próbageometriák esetén. A TS azt jelenti, hogy a számítás egy átmeneti állapothoz konvergált, míg a NC azt, hogy nem volt konvergencia 100 lépésen belül. Az ábra az [1]-es hivatkozás 3. ábrája.

A 216 N-protonált próbageometriából 15 darab olyan van, amely 100 lépésen belül nem felelt meg a programban¹¹⁰ definiált konvergencia kritériumoknak (NC), illetve szintén 15

darab (három különböző) átmeneti állapothoz konvergál (TS), ezek a kapott minimumokból származtathatóak. A többi esetben három különböző minimum konformert kapunk, amelyek előfordulási valószínűsége ~2:3:3. Az 1296 O-protonált próbaszerkezet csaknem fele N-protonált minimumokhoz (kivéve a Π_N -hez) konvergál, ez egyrészt geometriai megfontolásokkal (a geometriák részletesebben a 7.2.1. fejezetben kerülnek elemzésre) magyarázható, másrészt azzal, hogy a legmagasabb amino- és a legalacsonyabb karbonil-protonált szerkezetek között több, mint 20 kcal/mol-os relatív energia különbség van. A többi geometriából 50 nem optimálható (NC), 20 darab pedig négyféle átmeneti állapotba konvergál. A maradék szerkezetekből nyolc különböző minimum konformert tudunk azonosítani, ezek közül vannak gyakrabban (131-132) és ritkán (3!) előfordulóak is az adathalmazban.

7.1.2. Cys és CysH⁺: konformerek²

Míg a glicinnél a semleges forma konformerei jól meghatározottak az irodalomban, addig a ciszteinre már nem ilyen egyértelmű a helyzet, ezért a neutrális aminosav konformációs terét is fel kell térképeznünk. A kezdeti adatpontok felvételéhez a Wilke és munkatársai által publikált globális minimum szerkezet szolgált alapot,⁵⁸ ezen forgatjuk el 60 fokként a belső torziós mozgást leíró molekula részeket: az oldalláncot, az amino-, a karboxil-, a hidroxil- és a tiolcsoportot (lásd 3. ábra).



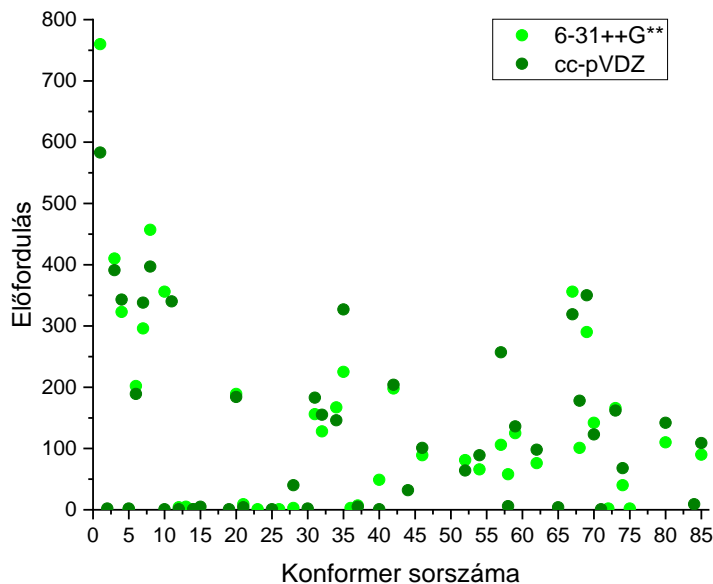
3. ábra: A semleges cisztein konformációs terének szisztematikus feltérképezéséhez készített sematikus vázlat a forgatási tengelyek jelölésével.

Minden forgatásnál 6 állás lehetséges (0° és a 360° ekvivalens), ezért összesen $6^5 = 7776$ adatpontunk lesz. Mivel a rendszer nem rendelkezik szimmetriával, így a pontok száma nem redukálható. A forgatásokkal elkészített próbageometriákat MP2 módszerrel optimáltuk a

következő bázisokkal kombinálva: 3-21G, 6-31G, 6-31++G, 6-31G**, 6-31++G** és cc-pVDZ. Az így kapott szerkezeteket konvergencia esetén konformerekhez rendeltük, amiket tovább optimaltunk MP2/aug-cc-pVDZ elméleti szinten. Attól függően, hogy melyik bázissal végeztük el a kezdeti forgatásokat, MP2/aug-cc-pVDZ szinten rendre 48, 44, 39, 51, 49 és 50 stacionárius pontot, azon belül 45, 41, 37, 47, 45 és 43 minimum konformert kaptunk. Megkülönböztethetünk még „egyedi” stacionárius pontokat és minimumokat, esetünkben az „egyedi” jelölés azt jelenti, hogy az adott szerkezetet nem találtuk meg semelyik, az adott bázisnál kisebb báziskészlettel elvégzett optimalások során sem. Összesítve 55 egyedi minimum konformert kaptunk a módszerünkkel. A 6-31++G bázis rosszabbul teljesít mind a kisebb, mind a nagyobb bázisoknál, de nem vonhatunk le általános következtetések a diffúz függvények hasznosságáról, mert ha azokat a 6-31G** báziskészlet mellé alkalmazzuk, öt egyedi konformert is találunk annak ellenére, hogy a talált minimumok száma 47-ről 43-ra csökken. Nagyobb rendszerek esetén, amikor nem cél az összes konformer megtalálása hasznosak lehetnek a 3-21G és 6-31G** bázisok, mert alacsony számítási költségek mellett nagyszámú minimum szerkezetet képesek megtalálni (az utóbbi ráadásul esetünkben az összes, 2,5 kcal/mol alatti konformerhez elvezetett). Kisebb molekulák esetében, ahol megengedhető, érdemes lehet még az általunk alkalmazottaknál nagyobb bázisokkal is számításokat végezni, de nem önállóan, hanem kisebbek kiegészítéseként.

Noha az irodalomban^{55,57-59} nincs teljes egyetértés a cisztein minimum konformereinek számáról, nem voltunk elégedettek az 55-tel. A legátfogóbb konformerkeresés Wilkéhez és munkatársaihoz kötődik, akik 71 minimum konformert találtak MP2/cc-pVTZ szinten.⁵⁸ Hogy közvetlen összehasonlítást tehesünk, újraoptimaltunk a geometriáinkat az imént említett elméleti szinten és harmonikus rezgési frekvencia számításokat is végeztünk, hogy megállapítsuk, a mi szerkezetünk is valóban minimum konformerek MP2/cc-pVTZ elméleti szinten. A szerkezeteink ezen az elméleti szinten is minimum konformerek, viszont kiderült, hogy a különbség nem csupán $71 - 55 = 16$ volt, hanem 31, ugyanis néhány általunk kapott minimum konformert ők nem találtak meg (kiemelendő ezek közül kettő, amelyek relatív energiája a globális minimumhoz képest kevesebb, mint 3 kcal/mol!). Ezután fordítva jártunk el: a tőlünk hiányzó minimumokra geometriai optimalásokat és harmonikus rezgési frekvencia számításokat végeztünk MP2/aug-cc-pVDZ szinten. A 31 hiányzó konformerből ezek után már „csak” 30 maradt, mivel két geometria ugyanahhoz a szerkezethez konvergált a magasabb elméleti szinten. Az egyik különbség a két munka között az volt, hogy Wilkéék

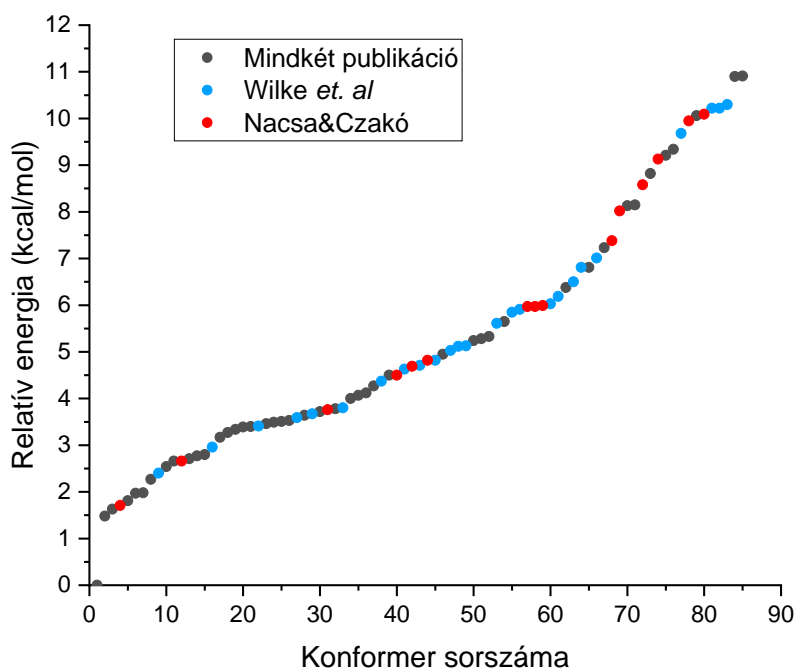
HF számításokat is végeztek 3-21G bázissal, ezért úgy döntöttünk, mi is megtegyük. Ezzel 2, eddig hiányzó szerkezetet találtunk meg, maradt még 28. Érdekes tapasztalat volt, hogy ezzel a módszerrel is megkaptunk néhány olyan konformert, amely a hivatkozott publikációban nincs jelen, azonban mi megkaptuk már őket az eredeti eljárásunk során. A következő különbség a forgatások szögeiben keresendő: míg mi mindent 60° -onként vizsgáltunk, addig ők az aminocsoportot és az oldalláncot 30° -onként, míg a többi részt csak 120° -onként. A két kiemelt molekuláris részleten elvégeztük mi is a 30° -os forgatásokat HF/3-21G szinten, a többi csoportnál eközben maradt a 60° -os felbontás. Ezen eljárás keretében mindössze egyetlen újabb minimum geometriát találtunk meg. Végül a maradék 27 minimum konformert hozzáadtuk a mi készletünkhöz, így összesen 85 minimum szerkezetről beszélhetünk MP2/aug-cc-pVDZ elméleti szinten. Nem vonható le egyértelmű következtetés, hogy mi okozza az eltérést a két eljárás között, a hiányzó szerkezetek között nem találtunk olyan strukturális összefüggéseket, amikből eljuthattunk volna a megoldáshoz. A probléma eredhet abból is, hogy különböző elektronszerkezet számító programokat használtunk, ezáltal különböző optimáló algoritmusokat, vagy a különböző kiindulási geometriából, amely az efféle konformer keresések alapja.



4. ábra: A különböző semleges cisztein konformerek előfordulásának gyakorisága a konformációs tér MP2 módszerrel, 6-31++G** és cc-pVDZ bázisokkal való feltérképezése során. Az ábra a [2]-es hivatkozás 1. ábrája.

Az egyes konformerek előfordulási gyakorisága az MP2/6-31++G** és MP2/cc-pVDZ számítások során a 4. ábrán látható. Általánosságban, ha egy szerkezet az egyik bázissal

gyakori, akkor a másikkal is (ez a „ritkaságra” is igaz), azonban van néhány kivétel. Például a 10-es konformert 356 esetben kaptuk meg az első bázist használva, míg a második alkalmazva már csupán egyetlen egyszer. Azok a szerkezetek, amelyek csak az egyik vagy csak a másik bázissal végzett számítások során adódtak általában kevesebb, mint tízszer, de inkább egy-két alkalommal fordultak elő a kapott struktúrák kiértékelése során. A két kivétel a 11-es és a 44-es szerkezet, amelyek rendre 340 és 32 esetben fordultak elő. A 6-31++G** bázis 6, a cc-pVDZ pedig 10 olyan minimum konformert talált meg, amelyet a másik bázis nem. Mivel nagyon sok konformer nem túl gyakori, ezért arra következtethetünk, hogy ez a szisztematikus módszer rendkívül érzékeny a kiindulási geometria megválasztására, még így is lehetnek elhanyagolt, eddig még nem publikált szerkezetek. A 85 végső konformer relatív energiaeloszlása az 5. ábrán látható, amelyen azt is jelöltük, hogy az adott szerkezet a saját szisztematikus feltérképezésünkből származik, vagy az irodalomból vettük.

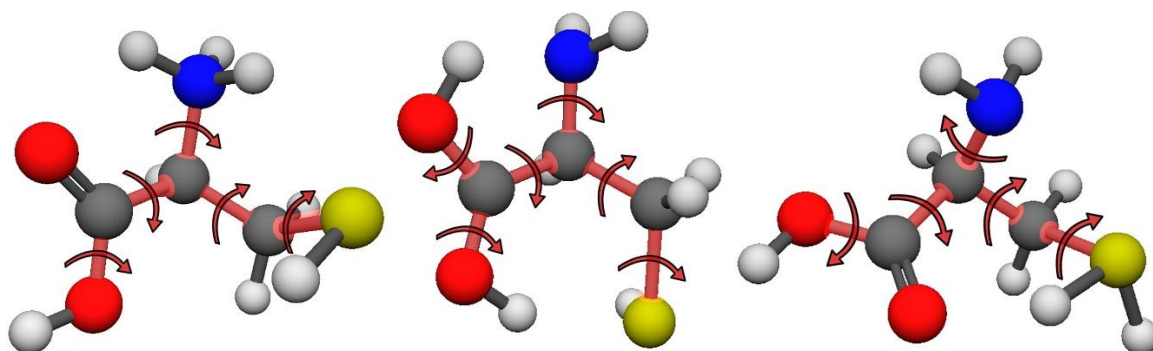


5. ábra: A 85 beazonosított semleges cisztein minimum konformer relatív energiájának eloszlása MP2/aug-cc-pVDZ elméleti szinten. Piros színnel jelöltük azokat, amelyeket csak a mi keresésünk,² kék színnel azokat, amelyeket csak Wilke és munkatársainak a keresése,⁵⁸ míg szürkével, amelyeket mindkét keresés megtalált. Az ábra a [2]-es hivatkozás 2. ábrájának átdolgozása. Az ábra a [2]-es hivatkozás 2. ábrája.

A pontokból látszik, hogy gyakorlatilag minden régióban van olyan geometria, amelyet csak az egyikünk módszere azonosított, ezért a két publikáció jó kiegészítése egymásnak. A legmagasabb relatív energia 10,91 kcal/mol, míg a legalacsonyabb 1,48 kcal/mol a globális minimumhoz képest, MP2/aug-cc-pVDZ elméleti szinten. Ezen két határoló érték közt a

többi viszonylag folyamatos három szakadással a 68-as (0,64 kcal/mol), 71-es (0,43 kcal/mol) és a 83-as (0,60 kcal/mol) konformereknél.

A protonált szerkezetek azonosításának érdekében megvizsgáltuk, hogy melyek lehetnek azok a részei a molekulának, amelyek képesek a proton felvételére. Kémiai intuíciónk alapján négy ilyen helyet feltételeztünk: a hidroxil-, a karbonil-, az amino- és a tiolcsoportot, amelyek protonálódása a következő kationokhoz vezet: $^+(H_2O)OC-CH(-NH_2)-CH_2-SH$, $(HO)_2C^+-CH(-NH_2)-CH_2-SH$, $HOOC-CH(-N^+H_3)-CH_2-SH$ és $HOOC-CH(-NH_2)-CH_2-S^+H_2$. A próbaszerkezetek elkészítéséhez fogtuk az általunk talált cisztein globális minimum konformert (MP2/aug-cc-pVDZ geometria) és azon a megfelelő helyekre elhelyeztünk egy extra protont kétfajta variációban: a két karbonil-protonált konformerben az OH csoportok *cisz* és *transz* állása volt a különbség, az aminocsoport esetén a glicinhez hasonlóan elvégeztünk egy 60° -os forgatást, a tiol- és hidroxil-protonációnál pedig páronként a H–O–H és a H–S–H síkok merőlegesek egymásra. Az MP2/aug-cc-pVDZ szintű számításaink azt mutatták, hogy a négyből három funkciós csoport protonálódása vezet stabil szerkezetekhez, a hidroxil-protonálódás nem, csakúgy, mint a glicin esetében. A protonálódás helye szerint tehát megkülönböztethetünk N-(amino-), O-(karbonil-) és S-(tiol-)protonált konformereket.

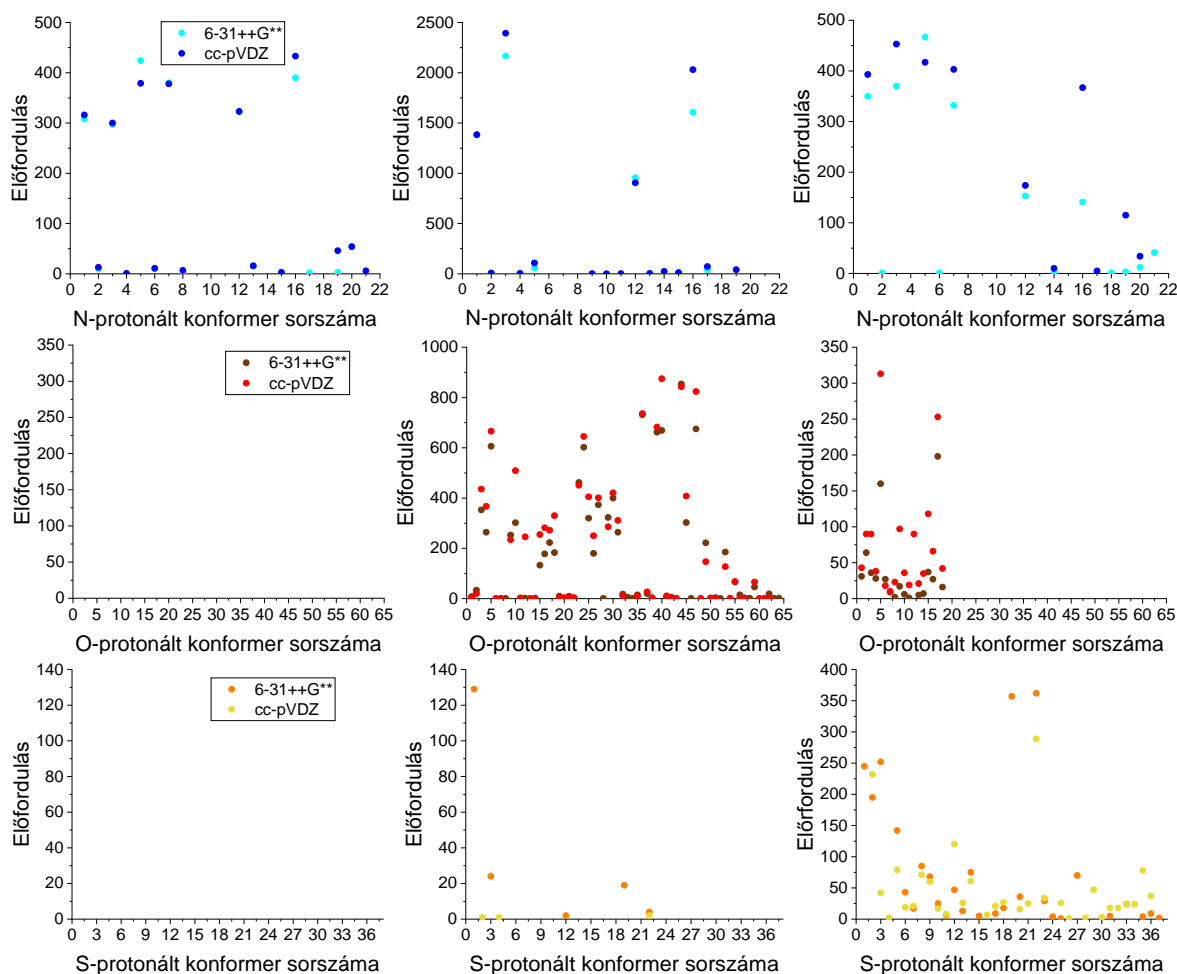


6. ábra: Az N- (bal), O- (közép) és S-protonált (jobb) cisztein konformációs terének szisztematikus feltérképezéséhez készített sematikus vázlat a forgatási tengelyek jelölésével.

A konformereket a korábban ismertetett szisztematikus eljárás keretében kerestük (6. ábra), azonban a semleges ciszteinnel ellentétben a protonált szerkezetben fellelhető szimmetria. Az N-protonált molekulában is öt darab szöget érdemes forgatni, mint a semleges aminosavban, azonban a protonált aminocsoport C_{3v} szimmetriáját figyelembe véve a kiindulási geometriák száma $6^4 \times 2 = 2592$ -re csökkenthető. Az O-protonált geometriából kiindulva már hat csoportot kellene, hogy forgassunk (az „új” OH részt is), ami $6^6 = 46656$ adatpont felvételét indokolná, azonban ez a szám megfelelezhető 23328-ra, ha kihasználjuk a

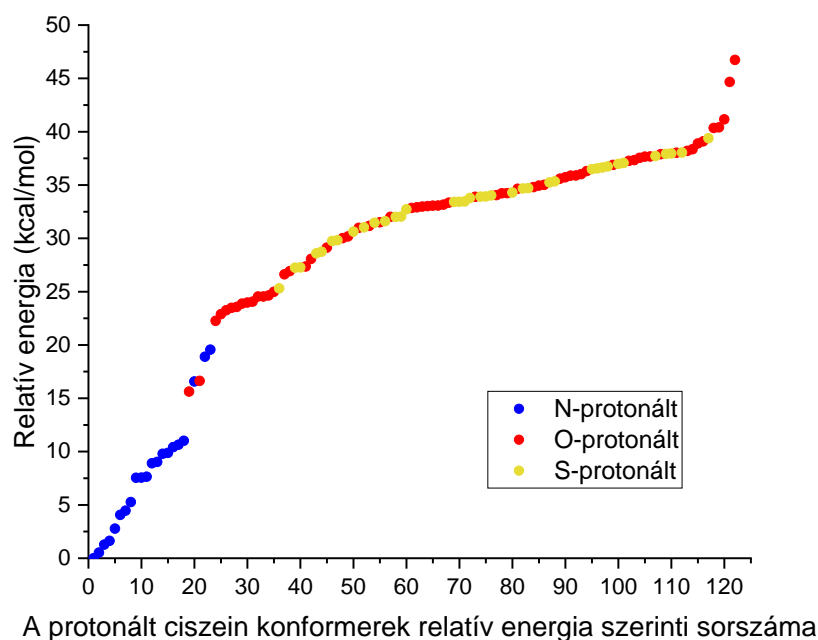
C^+O_2 molekularész C_{2v} szimmetriáját. A tiol-protonálódás nem jár új forgatandó molekularész kialakításával, viszont szimmetria sem lép fel, így $6^5 = 7776$ kezdeti geometria optimálása volt szükséges. A semleges ciszteinnél azt találtuk, hogy a cc-pVDZ és 6-31++G** bázisok kombinálva a leghatékonyabbak (az összes, általunk kapott konformert megtalálták és voltak olyan szerkezetek is, amiket csak az egyikük), ezért a forgatásokkal előállított próbageometriák optimálására ezeket alkalmazzuk az MP2 módszerrel. A végső, potenciális minimum szerkezeteket ismét MP2/aug-cc-pVDZ szinten optimáltuk valamint meghatároztuk a harmonikus rezgési frekvenciáikat. Ezek a számítások rendre 21, 64 és 37 N-, O- és S-protonált szerkezethez vezettek (illetve néhány „melléktermékhez”, de ezekről később fogok értekezni). Noha a globális minimum szerkezete megegyezik a korábbi tanulmányokban leírtakkal, a protonált minimum konformerek száma (122) jóval több a 21-hez⁵⁹ és a 6-hoz⁴⁵ képest. Az egyes konformerek előfordulási gyakorisága a három geometriai készletben (és a két különböző bázissal) a 7. ábrán van feltüntetve.

Az eredetileg N-protonált geometriákon (7. ábra, első oszlop) az optimálások során sem változik meg a protonálódás helye, viszont a 21-ből csak 16 amino-protonált konformert kapunk meg ebből az adathalmazból. A két báziskészlet jó egyezést mutat az egyes konformerek előfordulási gyakoriságában, viszont vannak olyan szerkezetek is, amelyek csak az egyik bázis alkalmazásával fordult elő: a 6-31++G** bázissal az N17-es míg a cc-pVDZ bázissal az N4-es és N21-es szerkezetek (előfordulásuk rendre 2, 1 és 6). Az eredetileg karbonil-protonált kiindulási szerkezetek (7. ábra, második oszlop) már vegyesebb geometriákhoz vezettek, voltak olyan optimálások, amelyek amino- vagy tiol-protonált konformerekhez konvergáltak. Ennek oka a konformerek relatív energiaeloszlásában keresendő, amiről később lesz szó. Az amino-protonált N9, N10 és N11-es geometriák csak ebből, az eredetileg O-protonált geometriai készletből kerültek beazonosításra, azon belül is csak a cc-pVDZ bázissal végzett optimálások esetén, rendre 2, 1 és 2 szerkezet esetén. Ezek nagyon ritka szerkezetek, főleg ha belegondolunk, hogy míg az eredetileg N-protonált próbageometriák száma több, mint kétezer, addig itt már több, mint húszezer az adatpontok száma. A tiol-protonált 37 minimum szerkezetből 7 darabot találunk meg ebben az adathalmazban, de egyiket sem csak és kizárólagosan itt. A karbonil-protonáltak közül pedig be tudtuk azonosítani mind a 64-et. Az egyes konformerek előfordulási valószínűsége megint csak hasonló a két bázisra nézve, a csak az egyik bázissal kapott konformereknél ez alacsony (9 különböző ilyen szerkezet van 6-31++G**-al, míg cc-pVDZ-vel 8). Az eredetileg tiol-protonált adatsorból kiindulva (7. ábra, harmadik oszlop) az



7. ábra: A különböző N-, O- és S-protonált (fentről lefelé) cisztein minimum geometriák előfordulása az MP2 módszerrel, 6-31++G** és cc-pVDZ bázisokkal elvégzett konformerkeresések során. Az oszlopok balról jobbra az eredetileg amino-, karbonil- és tiol-protonált próbaszerkezetekből kapott eredményeket foglalják össze. Az ábra a [2]-es hivatkozás 3. ábrája.

optimálások után megkapjuk az összes (37) tiol-protonált szerkezetet. Rajtuk kívül azonosítható még a legtöbb amino-protonált (14 a 21-ből, köztük az N18, amely az egész keresési folyamat során csupán egyszer fordult elő, ebben az adatsorban, a 6-31++G** bázissal) és néhány karbonil-protonált (az első 18) konformer, amelyek megjelenése ismét csak a relatív energiákkal magyarázható. Ebben az esetben a legnagyobb az eltérés a két bázissal kapott geometriák előfordulási gyakoriságában, hiszen például a globális minimum tiol-protonált S1 konformert csupán a 6-31++G**-hoz tartozó szerkezeteknél azonosítottuk, ott azonban 245 alkalommal. Ezen bázishoz 6, míg a cc-pVDZ-hez 8 egyedi konformer tartozik. Ennek a két báziskészletnek az alkalmazása a semleges formához hasonlóan itt is szükségesnek bizonyult, jól kiegészítik egymást, ha a számítási kapacitások lehetővé teszik, érdemes lehet akár még nagyobb bázisokat is kipróbálni.

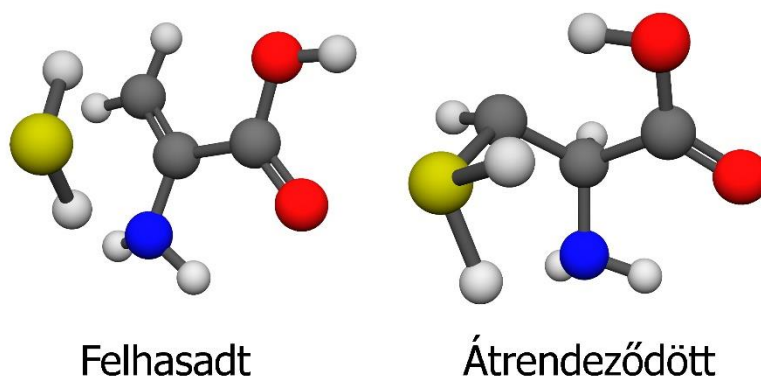


8. ábra: A beazonosított 122 protonált cisztein minimum konformer relatív energiájának eloszlása, MP2/aug-cc-pVDZ szinten. 21 N- (kék), 64 O- (piros) és 37 S-protonált (sárga) geometriát azonosítottunk. Az ábra a [2]-es hivatkozás 4. ábrája.

A protonált cisztein szerkezetek relatív energiaeloszlását (MP2/aug-cc-pVDZ szinten) a 8. ábrán foglaltuk össze. A legmélyebb energiák az amino-protonált szerkezetekhez tartoznak (ennek globális minimumát vesszük 0,0 kcal/mol relatív energiájúnak), a karbonil-protonáció globális minimuma 15,6 kcal/mol-lal, a tiol-protonációé pedig 25,3 kcal/mol-lal magasabb. A különböző csoportok protonálódása „átfed”, olyan értelemben, hogy a legmagasabb relatív energiájú N-protonált szerkezet előtt már kettő darab O-protonált szerkezet is megtalálható, majd később az S-protonált konformerek megjelenésével azok energiában vegyesen helyezkednek el a karbonil-protonált geometriákkal. Ezzel magyarázható az, hogy némelyik protonált kiindulási szerkezet megfelelő orientáció esetén egy másik helyen protonált geometriához konvergált az optimálások során.

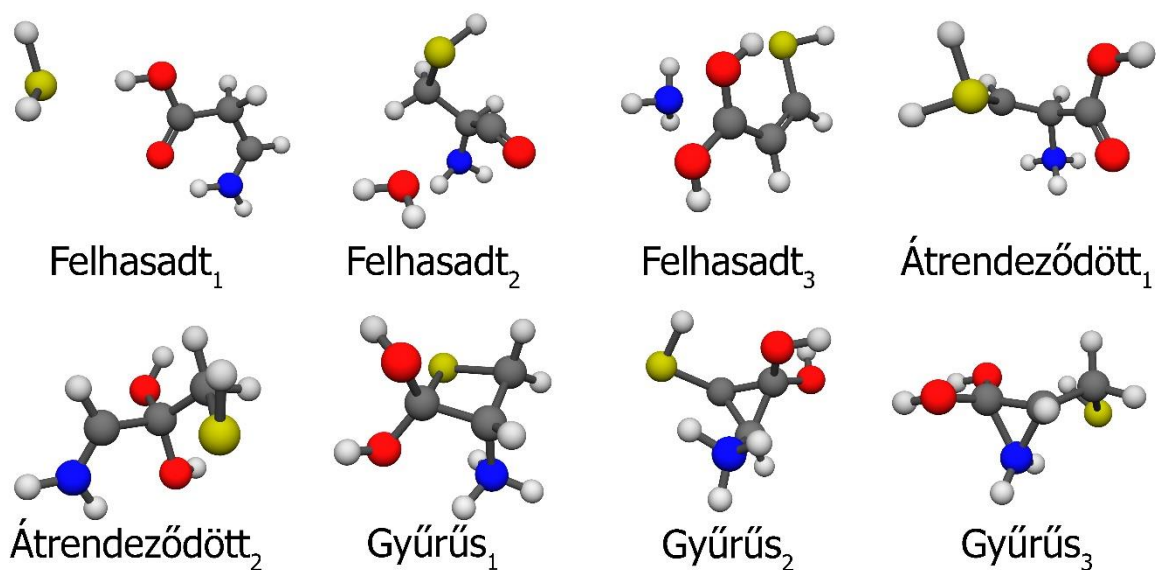
Miközben a semleges és protonált cisztein minimum szerkezeteit kerestük azonosítottunk néhány „mellékterméket” (hála a kötésrendek vizsgálatának), amelyek nem feltétlenül relevánsak a kutatásunk számára, azonban más munkákban érdekesek lehetnek. Ezekre is elvégeztük az MP2/aug-cc-pVDZ geometriai optimálásokat és a frekvencia számításokat és meggyőződünk arról, hogy valóban minimumokról beszélünk. A relatív energiákat a

semleges cisztein és az N-protonált cisztein globális minimumához képest adtuk meg (attól függően, hogy az adott „melléktermékek” mely konformerek keresése során fordultak elő).



9. ábra: A semleges cisztein konformációs terének feltérképezése során adódott két tipikus „melléktermék” MP2/aug-cc-pVDZ elméleti szintű geometriái. Összesen 23 „Felhasadt” és 20 „Átrendeződött” konformert találtunk. Az ábra a [2]-es hivatkozás 5. ábrája.

A semleges aminosavhoz kötődő ilyen struktúrákat két csoportba sorolhatjuk, amelyek a „Felhasadt” és az „Átrendeződött” neveket kapták, ezeknek egy-egy tagját a 9. ábrán mutatjuk be. Az előbbinél két kötés szűnik meg, egy S–C és egy C–H, majd a két levált fragmensből H₂S képződik, ami eltávolodik a hátramaradó nagyobb molekulától. A 23 ilyen szerkezet relatív energiája a 14,86–23,67 kcal/mol-os tartományban helyezkedik el. A legmagasabb relatív energiájú cisztein konformer 10,91 kcal/mol-os, ehhez képest van egy „ugrás” az eloszlásukban, viszont nagyobb molekuláknál ez eltűnhet, a szerkezetek szemrevételezése mellett érdemes lehet alkalmazni valamiféle megszorításokat a kötéshosszakra, vagy bevezetni kötésrend-mátrixokat, hogy megbizonyosodjunk arról, hogy valóban a keresett molekulát kapjuk a geometriai optimalás után. A másik „melléktermék” csoport az „Átrendeződött”. Ezek a geometriák első ránézésre tiol-protonáltak tűnhetnek, azonban a proton az oldallánc szénatomja mellől származik. 20 ilyen konformert különböztettünk meg, amelyeknek a relatív energiája 65,91 és 76,87 kcal/mol között van, vagyis a semleges ciszteintől távol. A protonált ciszteinek feltérképezése során talált „melléktermékeket” már három csoportba sorolhatjuk be: „Felhasadt”, „Átrendeződött” és „Gyűrűs”, ezeknek példái a 10. ábrán láthatóak. Egy „Felhasadt” szerkezet a semleges glicinnél ismertetett módon „képződik”, azonban itt további három osztályba sorolhatjuk őket attól függően, hogy az egyik újonnan képződött molekulánk H₂S („Felhasadt₁”), H₂O („Felhasadt₂”) vagy NH₃ („Felhasadt₃”). Rendre 40 (relatív energia: 8,00–49,34 kcal/mol), 5 (relatív energia: 31,85–41,16 kcal/mol) és 2 (26,76 és 29,97 kcal/mol) ilyen minimum geo-

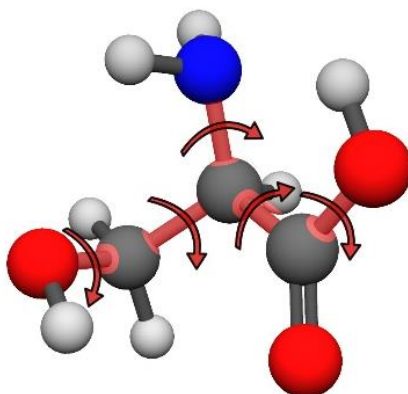


10. ábra: A protonált cisztein konformációs terének feltérképezése során adódott három tipikus „melléktermék” csoport MP2/aug-cc-pVDZ elméleti szintű geometriái. Összesen 40 + 5 + 2 (47) „Felhasadt”, 6 + 1 (7) „Átrendeződött” és 1 + 15 + 9 (25) „Gyűrűs” konformert találtunk. Az ábra a [2]-es hivatkozás 6. ábrája.

metriát találtunk. A célmolekulaionok (protonált cisztein) és ezen „melléktermékek” relatív energia tartománya átfed, úgyhogy a korábban említett geometriai megszorítások alkalmazása megkerülhetetlen. Megfontolandó, hogy a „Felhasadt₂” csoport inkább hidroxil-protonált szerkezet, de a megnyúlt C–O kötés (kb. 2,7 ångström) miatt inkább egy víz-kation komplexnek mondanánk (egyértésben a protonálódási helyek vizsgálata során kapott eredményeinkkel és az irodalommal). Hat darab „Átrendeződött₁” szerkezetünk van, amelyekben az oldallánc szénatomjához kötődő egyik hidrogén átkerül a lehetséges protonálódási helyek egyikére, ezek az 59,45–64,92 kcal/mol relatív energia régióban vannak. Azonosítottunk továbbá egy Átrendeződött₂ geometriát 25,37 kcal/mol relatív energiával, amelynél az N- és O-tartalmú ligandumok szénatomot „cseréltek”. Az új, korábban még elő nem fordult csoport a „Gyűrűs”, amiknél különböző három vagy négy atomból álló gyűrűk kialakulását figyeltük meg. A „Gyűrűs₁” egy ilyen konformert tartalmaz, amelyben egy C–C–C–S négytagú gyűrű van, relatív energiája 20,61 kcal/mol. A „Gyűrűs₂” és „Gyűrűs₃” osztályokban háromtagú, C–C–C vagy C–C–N gyűrű azonosítható. Az előbbi csoportban három ilyen szerkezet található 37,03; 37,25 és 37,57 kcal/mol relatív energiákkal. A maradék tizenkettő ilyen konformer a 62,18–73,61 kcal/mol-os, míg a kilenc C–C–N gyűrűs struktúra a 36,00–39,83 kcal/mol-os energiatartományban helyezkedik el.

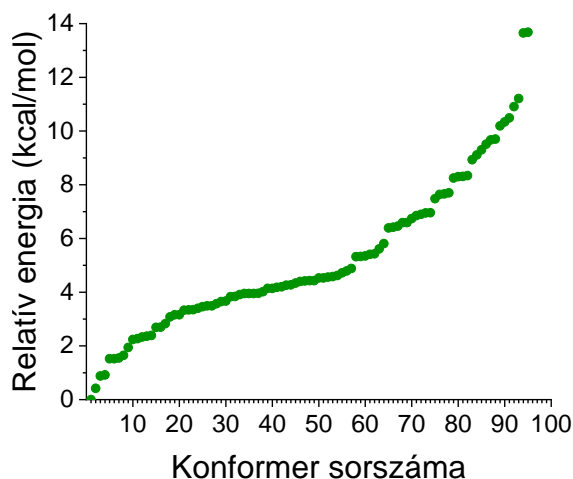
7.1.3. Ser és SerH⁺: konformerek³

A semleges szerin konformereinek meghatározásához egy korábbi, He és Allen által publikált globális minimum szerkezetet használtunk fel.⁶⁷ A konformációs tér feltérképezéséhez a belső torziós mozgást leíró szögeket változtattuk 60 fokként. Ahogyan az a 11. ábrán látható, ez a szerin amino-, karboxil-, két hidroxilcsoportjának és az oldalláncának a forgatásának felel meg.



11. ábra: A semleges szerin konformációs terének szisztematikus feltérképezéséhez készített sematikusan vázlat a forgatási tengelyek jelölésével. Az ábra a [3]-as hivatkozás 1. ábrájának átdolgozása.

Az így kapott $6^5 = 7776$ próbaszerkezetet az MP2 módszerrel optimalítottuk cc-pVDZ és 6-31++G** bázisokkal (a ciszteines tapasztalatainkból kiindulva ezen bázisok alkalmazása célravezető). Az első bázis 184, míg a második 168 potenciális minimum geometriát eredményezett a korábban ismertetett energia és geometriabeli megszorításainkat alkalmazva. Ha nagyobb, aug-cc-pVDZ bázist alkalmazva elvégezzük ezeknek a szerkezeteknek az optimalizálását és a harmonikus rezgési frekvenciáik meghatározását, összesen 95 minimum geometriát kapunk ezen az elméleti szinten. A nagy különbség azzal magyarázható, hogy számos szerkezet ugyanabba a konformerbe konvergál, de persze akadnak olyan geometriák is, amelyek átmeneti állapotokba optimalizálódnak, vagy egyszerűen csak nem konvergálnak. A végső, 95 minimum szerkezetnek a relatív energia eloszlása az 12. ábrán található. A pontok gyakorlatilag folyamatosan helyezkednek el 0,00-tól 11,21 kcal/mol-ig, majd itt egy körülbelül 2 kcal/mol-os ugrás látható, hiszen az utolsó két minimum konformer relatív energiája 13,65 kcal/mol és 13,68 kcal/mol. A legátfogóbb konformerkeresés He és Allen nevéhez köthető 2016-ból, ők összesen 85 konformert tudtak



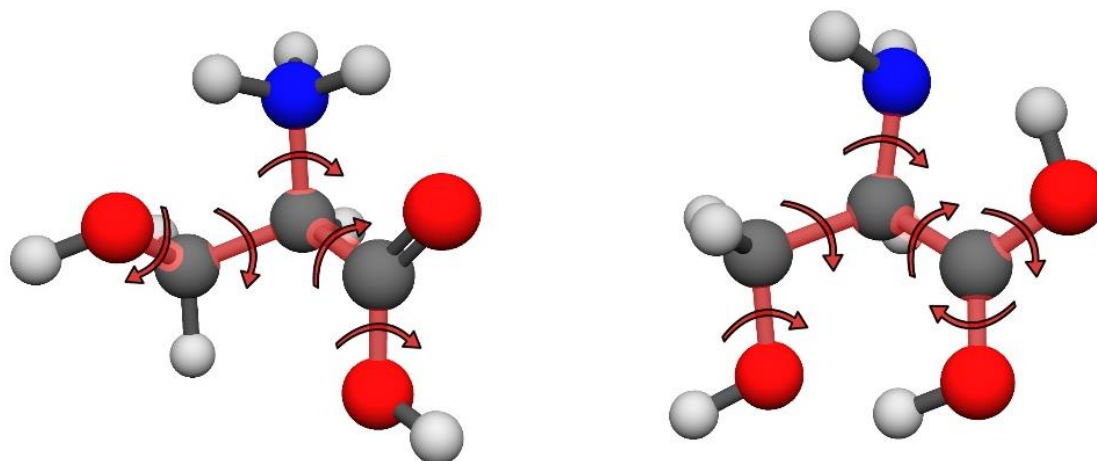
12. ábra: A beazonosított 95 semleges szerin minimum konformer relatív energiájának eloszlása, MP2/aug-cc-pVDZ szinten. Az ábra a [3]-as hivatkozás 2. ábrájának átdolgozása.

azonosítani.⁶⁷ A cc-pVDZ bázis tíz olyan szerkezethez vezetett el minket, amelyet a másik bázissal való keresés során nem találtunk volna meg (egyik ezek közül mindössze 1,52 kcal/mol relatív energiájú!), míg a 6-31++G** bázist alkalmazva csupán egy olyan konformer van (3,95 kcal/mol relatív energiával), ami nem azonosítható a másik adathalmazból.

A protonált szerin konformerek esetén szintén a korábban publikált szerin globális minimum konformert használtuk fel, a ciszteinhez hasonlóan négy lehetséges funkciós csoportra vonatkozóan végeztük el próbaszerkezetek optimalizálását: protont kapcsoltunk az amino-, a karbonil- és a két hidroxilcsoporthoz. Az így kapott pozitív töltésű molekulaionok a következők: $\text{HOOC-CH(-N}^+\text{H}_3\text{)-CH}_2\text{-OH}$, $(\text{HO})_2\text{C}^+\text{-CH(-NH}_2\text{)-CH}_2\text{-OH}$, $^+(\text{H}_2\text{O})\text{OC-CH(-NH}_2\text{)-CH}_2\text{-OH}$ és $\text{HOOC-CH(-NH}_2\text{)-CH}_2\text{-O}^+\text{H}_2$. Minden funkcióscsoport protonálódása esetén két próbageometriára végeztünk számításokat, a ciszteinnél ismertetett módon (természetesen az oldallánc tiolcsoportjának protonálódása esetünkben ekvivalens az oldallánc hidroxilcsoportjának protonálódásával). Az MP2/aug-cc-pVDZ geometriai optimalizálások és harmonikus rezgési frekvencia számításaink azt mutatták, hogy a glicinhez és ciszteinhez hasonlóan az α -szénatomhoz kapcsolódó hidroxilcsoport protonációja nem vezet stabil szerkezet kialakulásához, azonban a ciszteinnel ellentétben az oldallánc funkcióscsoportja sem protonálható, tehát ez esetben csak N- (amino-) és O- (karbonil-) protonálódásokat különböztethetünk meg.

A két protonált forma forgatásához készített vázlat a 13. ábrán látható. Az N-protonálódás esetén a kiindulási geometriák a semleges szerinhez hasonlóak, azonban a protonált

aminocsoport háromfogású szimmetriája miatt a szerkezetek száma harmadára, 2592-re redukálható. A karbonilcsoport protonálódásával egy újabb kötés mentén kell forgatásokat végrehajtanunk, azonban ekkor is felfedezhetjük a protonált karboxilcsoport síkszimmetriáját, így $6^6/2 = 23328$ elemből fog állni az O-protonált kiindulási szerkezeteket tartalmazó adathalmaz.

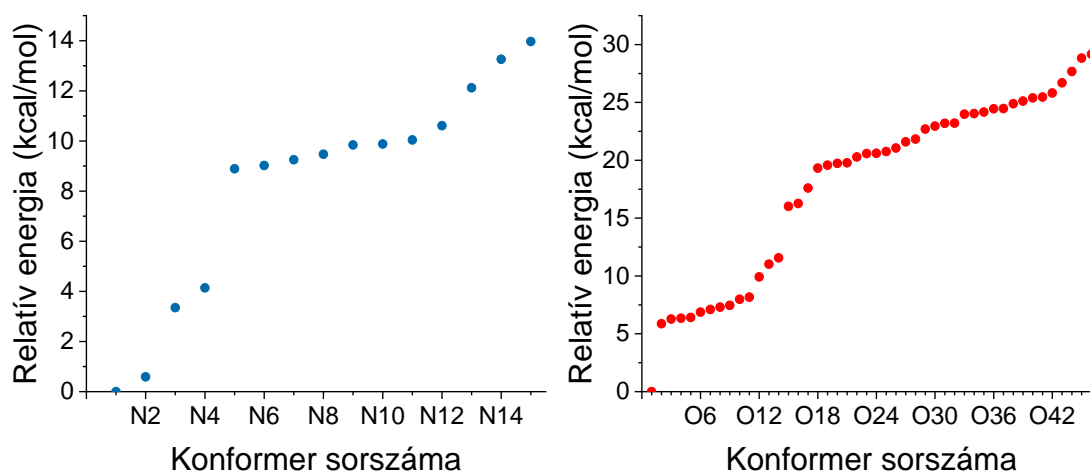


13. ábra: Az N- (bal) és O-protonált (jobb) szerin konformációs terének szisztematikus feltérképezéséhez készített sematikus vázlat a forgatási tengelyek jelölésével. Az ábra a [3]-as hivatkozás 1. ábrájának átdolgozása.

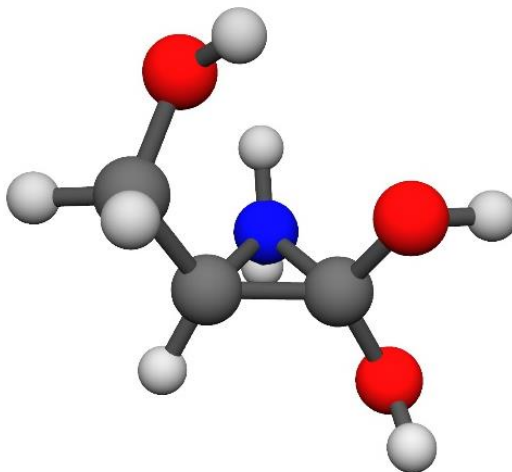
Az N-protonált molekulára a cc-pVDZ bázissal 32, míg 6-31++G** bázissal 27 különböző szerkezetet kaptunk az MP2/aug-cc-pVDZ elméleti szintű geometriai optimalizációkhoz és harmonikus rezgési frekvenciák meghatározásához, míg a karbonil-protonált szerin esetén ezek száma 136 és 129. Az eredményeket szűrve 15 amino-protonált szerint azonosíthatunk a magasabb elméleti szinten, ezek közül egyetlen van, ami csak az egyik bázissal való keresésből származik (N12, relatív energiája 10,61 kcal/mol).

Az amino-protonált konformerek a relatív energia eloszlása az 14. ábrán (bal oldal) látható. Négy szerkezet található az alacsony energiájú régióban ($\leq 4,14$ kcal/mol), míg a többi a 8,89–13,97 kcal/mol-os régióban. Karbonil-protonált minimum konformerekből 46 különbözőt találtunk ezen az elméleti szinten, ezek közül nyolcat csak a kisebb (cc-pVDZ), míg hatot csak a nagyobb (6-31++G**) bázissal elvégzett kezdeti optimalizációk után kaptunk meg. A minimum geometriák MP2/aug-cc-pVDZ relatív energia eloszlását az 14. ábrán (jobb oldal) tüntettük fel, az adatpontok változatosan helyezkednek el a 0,00–29,17 kcal/mol-os tartományban. Ez a görbe három további részre osztható fel, az első maga a globális minimum, a második 5,87 kcal/mol-tól tart 11,56 kcal/mol-ig, míg a többi adatpont

16,02–29,17 kcal/mol között helyezkedik el. A legmagasabb relatív energiájú N-protonált és az O-protonált globális minimum szerkezet között csupán 0,92 kcal/mol különbség van.



14. ábra: A beazonosított 15 N-protonált (bal oldal) és 46 O-protonált (jobb oldal) szerin minimum konformer relatív energiájának eloszlása, MP2/aug-cc-pVDZ szinten. Az ábra a [3]-as hivatkozás 2. ábrájának átdolgozása.



15. ábra: Az O-protonált szerin konformerkeresése során talált hat „melléktermék” közül egy struktúra, MP2/aug-cc-pVDZ szinten. Mind a hat szerkezet a „Gyűrűs” csoportba tartozik. Az ábra a [3]-as hivatkozás 3. ábrája.

Akárcsak a cisztein esetében, most is találtunk „melléktermék” szerkezeteket az O-protonált adatsor optimalálása során, hat darab „Gyűrűs”, háromtagú N–C–C gyűrűt tartalmazó geometriát, amelyek a három hidroxilcsoport térbeli orientációjában különböztek (15. ábra). Ezek relatív energiája az amino-protonált globális minimumhoz képest 37,10 kcal/mol és

41,44 kcal/mol között van, amely épp átfed a karbonil-protonált szerkezetek tartományával. Korábbi munkák során 14 N-protonált és 13 O-protonált szerkezetet tudtak azonosítani.

7. 2. Benchmark eredmények

A glicin esetén az összes, míg a cisztein és szerin esetén a tíz legmélyebb energiájú semleges és (különböző helyeken) protonált minimum konformerre meghatároztuk az irodalomban fellelhető eddigi legpontosabb, csatolt-klaszter szintű geometriákat és a hozzájuk tartozó relatív energiákat. Elvégeztük a struktúrák rezgési analízisét is harmonikus közelítésben és vizsgáltuk az elméleti szint hatását a zérusponti-energia korrekciókra. A semleges glicin konformerei valamelyest kivételet képeznek ez alól, némely eredmény egy, a disszertációhoz kapcsolódó, de annak alapját nem képező munkánkból származik, ezt természetesen jelölni fogom. Az elhanyagolt anharmonicitás jelentőségét már többen vizsgálták alaposan az aminosavak esetén,^{37,41,113} ez számunka feltehetőleg egy 0,01–0,1 kcal/mol-os bizonytalanságot okoz az imént említett korrekcióban. A kapott nagypontosságú energiák mellé meghatároztuk különböző kiegészítő korrekciók hatását, amelyek közül ugyan néhányat már a glicinre,³⁷ prolinra¹¹⁴ és a treoninra¹¹³ vizsgáltak, azonban átfogó tanulmány az összes aminosavra (és a hozzájuk tartozó PA/GB értékekre) nem található. Ilyen tagok a δT és $\delta(Q)$ (együttesen post-CCSD(T)), amelyek arra nyújtanak egy becslést, hogy a megszakított csatolt-klaszter sorfejtés esetén mekkora „hibát követünk el”, ha a háromszoros gerjesztéseket csak perturbatíván vesszük figyelembe, míg a négyszeres gerjesztéseket teljesen elhanyagoljuk. A kvantumkémiai számítások általában csak a vegyérték elektronok korrelációját veszik figyelembe, a törzselektronokét nem. Ezt az elhanyagolást számszerűsíti a $\Delta_{\text{törzs}}$, amely járulék meghatározásához a törzselektronok korreláltatásával és az anélkül kapott eredmények különbségét kell vennünk. Az utolsó kiegészítő korrekció a Δ_{rel} , amely a relativisztikus effektusokból ered, vagyis az elektronok gyors mozgásánál fellépő jelenségből. Ezt a másodrendű Douglas–Kroll relativisztikus korrekció¹¹⁵ kiszámításával határozzuk meg. Ahhoz, hogy figyelembe vegyünk azt is, hogy a molekuláknak 0 K-en is van rezgési energiája, meg kell határoznunk a zérusponti-energia korrekciót, amely értéke a harmonikus rezgési frekvenciák összegének a felével lesz egyenlő. Végezetül a meghatározott 0 K-es relatív entalpia értékeket a statisztikus termodinamika segítségével meghatározott termikus járulékok figyelembevételével továbbalakítottuk 298,15 K-es relatív entalpiákra és szabadentalpiákra. A kapott értékek hibabecslése kényes téma, hiszen bármennyire is szeretnénk megadni a bizonytalanságot számokban, ez kvantumkémiai számolásoknál nem egészen egyszerű. Mindenesetre a végső

mennyiségeknél a különböző elhanyagolásokat/közelítéseket figyelembe véve megpróbálunk egy becslést tenni a hibahatárra, ezek azonban nem tekinthetők „kőbe vésettnek”. Azt, hogy milyen hibaforrásokat veszünk figyelembe, a glicinnél fogom ismertetni, majd a cisztein és a szerin esetén is így fogunk eljárni.

7.2.1. Gly és GlyH⁺: nagypontosságú geometriák és energiák¹

A konformerkeresés során kapott MP2/aug-cc-pVDZ elméleti szintű protonált glicin minimumokat az explicit-korrelált csatolt-klaszter módszerrel (CCSD(T)-F12b),¹¹⁶ aug-cc-pVDZ bázissal optimáltuk és meghatároztuk harmonikus rezgési frekvenciáikat. A geometriák további optimálását elvégeztük aug-cc-pVTZ bázisok mellett is. Ezen a ponton vettük át egy korábbi munkánkból⁴ a semleges glicinre vonatkozó azonos elméleti szinten született adatokat és geometriákat. Végül mind a semleges, mind a protonált aminosavra single-point energia számításokat végeztünk aug-cc-pVQZ bázisok alkalmazásával. A kiegészítő korrekciók kiszámítását a legpontosabb, TZ szintű geometriákon végeztük.

A csatolt-klaszter háromszoros (δT)¹⁰² és perturbatív négyszeres ($\delta(Q)$)¹⁰³ korrekciókat meghatároztuk 3-21G, 6-31G és cc-pVDZ bázisokkal is, a legpontosabb értékeket a következőképpen kapjuk:

$$\delta T = \text{CCSDT/cc-pVDZ} - \text{CCSD(T)/cc-pVDZ}; \quad (7.2.1.1.)$$

$$\delta(Q) = \text{CCSDT(Q)/cc-pVDZ} - \text{CCSDT/cc-pVDZ}. \quad (7.2.1.2.)$$

A törzselektronok ($1s^2$ a szén-, nitrogén- és oxigénatomok esetén) korrelációját a következő „all electron” (AE) és „frozen-core” (FC) CCSD(T)-F12b/cc-pCVTZ-F12 elméleti szinten végzett számolások különbségeként vettük figyelembe:

$$\begin{aligned} \Delta_{\text{törzs}} = & \text{AE-CCSD(T)-F12b/cc-pCVTZ-F12} \\ & - \text{FC-CCSD(T)-F12b/cc-pCVTZ-F12}. \end{aligned} \quad (7.2.1.3.)$$

Az AE-CCSD(T)¹¹⁷ módszert aug-cc-pwCVTZ-DK¹¹⁸ bázissal használva meghatároztuk a másodrendű Douglas–Kroll relativisztikus korrekciót:

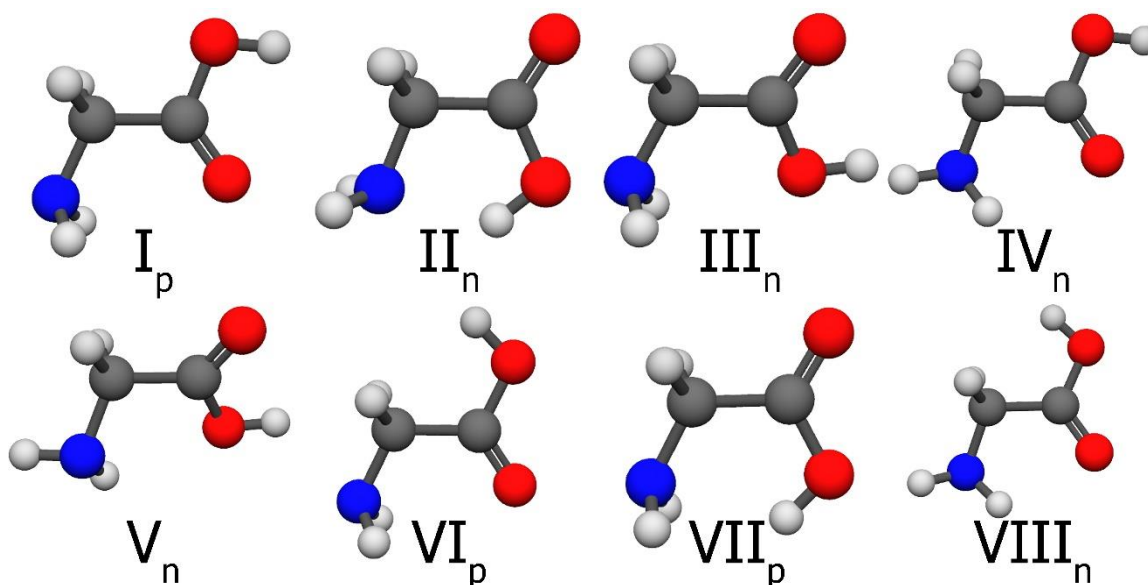
$$\begin{aligned} \Delta_{\text{rel}} = & \text{DK-AE-CCSD(T)/aug-cc-pwCVTZ-DK} \\ & - \text{AE-CCSD(T)/aug-cc-pwCVTZ}. \end{aligned} \quad (7.2.1.4.)$$

A kiegészítő korrekciókat a QZ szintű single-point relatív energiákhoz adva megkapjuk a legpontosabb egyensúlyi energiákat:

$$E_e = \text{CCSD(T)-F12b/aug-cc-pVQZ} + \delta T + \delta(Q) + \Delta_{\text{törzs}} + \Delta_{\text{rel}}. \quad (7.2.1.5.)$$

A CCSD(T)-F12b/aug-cc-pVDZ harmonikus rezgési frekvenciákat felhasználva kiszámolhatjuk a zérusponti-energia korrekciót (Δ_{ZPE}), ennek segítségével pedig a 0 K-es adiabatikus relatív energiák/relatív entalpiák is meghatározhatóak:

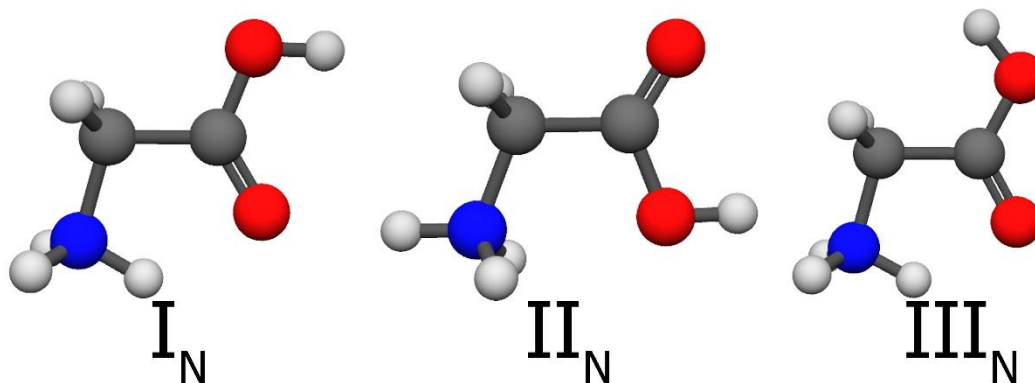
$$H_0 = E_e + \Delta_{\text{ZPE}}. \quad (7.2.1.6.)$$



16. ábra: A semleges glicin konformereinek CCSD(T)-F12b/aug-cc-pVTZ geometriái. A nevezéktan az irodalmat követi,³³ a római számok a relatív energiák sorrendje szerint alakul (egyetlen kivétel: a IV_n mélyebb energiájú, mint a III_n), a kis „p” a „planar”, vagyis sík szimmetrikus (C_s), míg a kis „n” a „non-planar”, vagyis nem sík szimmetrikus (C_1) szerkezetre utal.

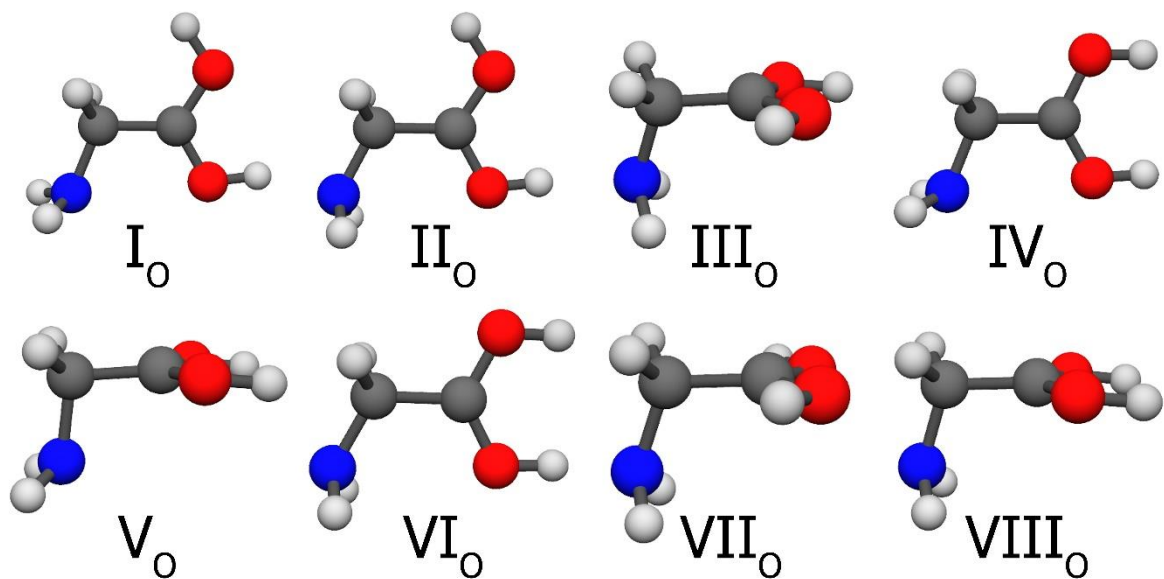
A semleges glicin nyolc minimum konformere jól ismert az irodalomból, ezek a 16. ábrán láthatóak. Elnevezésükben a római szám a relatív energiák sorrendjét jelöli nagyon alacsony elméleti szinteken, azonban az már 1992-ben kiderült, hogy a IV_n mélyebb energiával rendelkezik, mint a III_n.³³ A legújabb eredmények szerint (sokáig ez is vita tárgyát képezte) közülük három, az Ip, VI_p és VII_p síkszimmetrikus, míg a többinek nincs szimmetriája. A C_1 pontcsoportú szerkezeteknél vélhetően a hidrogénhidás kölcsönhatás kedvezőbben tud kialakulni a szimmetriából való kimozdulással. Az egyes konformerek a belső torziós mozgással, vagyis az amino-, karboxil- és hidroxilcsoportok elfordulásával könnyen egymásba alakíthatóak. Kísérletileg a nyolcból ötöt sikerült azonosítani, az IR spektrumuk alapján, Balabinnak^{50,51} az Ip, II_n, III_n és (elsőként) IV_n-t sikerült, míg a Tarczay csoportnak^{52,53} ezeken felül a VI_p-t is. Az általunk azonosított három N-protonált konformer

a 17. ábrán látható, elnevezésükben a római számok a CCSD(T)-F12b/aug-cc-pVQZ single-point relatív energiák sorrendjét követik, az „N” az alsó indexben pedig a protonálódás hely-



17. ábra: A három N-protonált glicin geometriája CCSD(T)-F12b/aug-cc-pVTZ elméleti szinten. Elnevezésük a CCSD(T)-F12b/aug-cc-pVQZ single-point relatív energiájuk sorrendjét követi (római számok), az alsó indexben az „N” pedig a protonálódás helyére (aminocsoport) utal.

ét jelöli. Mindhárom amino-protonált glicin C_s szimmetriájú, nagyban hasonlítanak a semleges glicin I_p formájára. Az I_N -t úgy származtathatjuk, hogy az I_p aminocsoportján lévő nemkötő elektronpár megköt egy protont és a protonált funkciócsoportot elfordul 60° -kal (ellenkező esetben egy átmeneti állapothoz tartozó geometriát kapunk). Ekkor intramolekuláris hidrogénkötés tud kialakulni a protonált aminocsoport egy protonja és a karbonilcsoport között. A III_N -t hasonlóan állítható elő a VI_p konformerből. A II_N párja pedig a III_N , ez esetben hidrogénkötés nem alakulhat ki a protonált aminocsoporttal, azonban a protonált konformer már síkszimmetrikus lesz. Ahogyan a 2. ábra kapcsán említettem, az O-protonált konformerek feltérképezése során a II_N geometriába egy számolás sem konvergált. Ennek nagyon egyszerű oka van: a protonált aminocsoport és a karbonilcsoport egymáshoz képest *transz* állásban helyezkednek el, tehát a proton nem tud „átvándorolni” a protonált karbonilcsoportból az aminocsoportra a geometriai optimalás során. A háromból egy, az I_N létezését tudták eddig még csak kimutatni, IRMPD-s mérés során (GH01 néven).⁴⁸ Az O-protonált konformereket a 18. ábrán tüntettük fel, megnevezésük ugyanazt a logikát követi, mint az N-protonáltaknál (ekkor a protonálódás helyét „O” jelöli alsóindexben). Négy hasonlít közülük a semleges aminosav konformereire (I_o , II_o , IV_o , VI_o), a többinél egy új motívum alakul ki, a karboxilcsoport (közel) merőleges helyzetbe fordul a molekula fősíkjához (N–C–C) képest. Megfigyelhető még az is, hogy ezekből a geometriákból hiányoznak a stabilizáló intramolekuláris hidrogénkötések. Ha az egyik hidroxilcsoport (akár az „eredeti”, akár a protonált karbonilcsoportból kialakuló) protonja kölcsönhatást ala-



18. ábra: A nyolc O-protonált glicin geometriája CCSD(T)-F12b/aug-cc-pVTZ elméleti szinten. Elnevezésük a CCSD(T)-F12b/aug-cc-pVQZ single-point relatív energiájuk sorrendjét követi (római számok), az alsó indexben az „O” pedig a protonálás helyére (karbonilcsoport) utal.

kítana ki az aminocsoport nemkötő elektronpárjával, a proton „átkötődne” oda, így mélyebb energiájú állapotba kerülne (ahogyan az a 2. ábrán, a próbageometriák eredményeinek elemzésénél látszik).

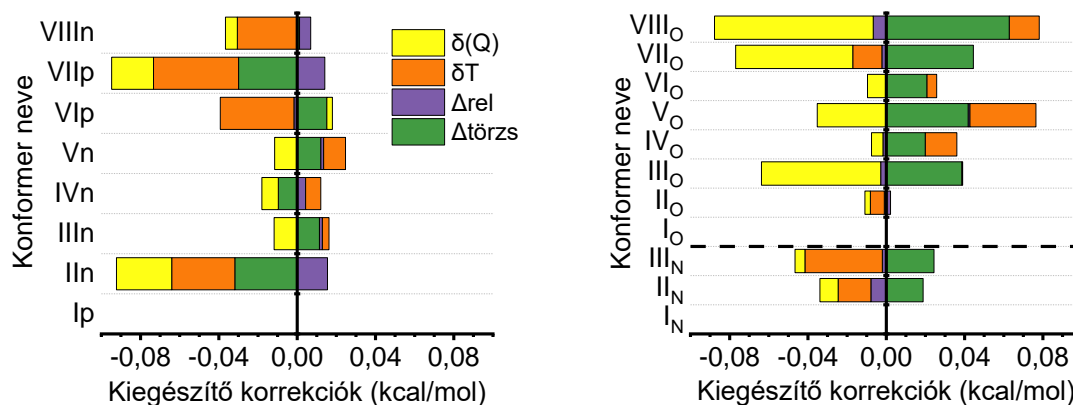
A glicinre és protonált glicinre vonatkozó relatív energiák és korrekciók értékei az 1. táblázatban láthatóak. A korreláció-konzisztens aug-cc-pVDZ báziskészletet alkalmazva jó egyezést kaptunk az MP2 és a CCSD(T)-F12b módszerek között, az eltérés átlagosan csupán 0,14 kcal/mol. Az egyetlen kiugró adat az amino-protonált II_N szerkezethez tartozik, amelyre az MP2 módszer körülbelül 0,6 kcal/mol-lal mélyebb energiát ad. MP2 módszerről CCSD(T)-F12b-re váltva változás figyelhető meg az $\text{IV}_\text{O}/\text{V}_\text{O}$ és $\text{VI}_\text{O}/\text{VII}_\text{O}$ O-protonált konformerek energiasorrendjében, a különbség a második esetben még tovább nő a báziskészlet növelésével. A CCSD(T)-F12b módszert alkalmazva látható a gyors báziskészlet konvergencia, hiszen míg az aug-cc-pVDZ és aug-cc-pVTZ bázisok által adott eredmények között átlagosan 0,05 kcal/mol eltérés van, a legmagasabb érték 0,09 kcal/mol (VI_O), addig tovább növelve a bázist aug-cc-pVQZ-re, az átlagos eltérés már csupán 0,01-0,02 kcal/mol, tehát a báziskészlet konvergencia kielégítő.

1. táblázat: A semleges és protonált glicin konformerek relatív egyensúlyi energiái, a kiegészítő korrekciói, 0 és 298,15 K-es relatív entalpiái és 298,15 K-es relatív szabadentalpiái. A mértékegység minden esetben kcal/mol. A táblázat az [1]-es hivatkozás 1. táblázatának átdolgozása.

Konformer neve	MP2	CCSD(T)-F12b			Σ_{kor}^e	ΔE_e^f	$\Delta_{\text{ZPE}}^{g,k}$	ΔH_0^h	$\Delta H_{298,15}^i$	$\Delta G_{298,15}^j$
	DZ ^{a,k}	DZ ^{b,k}	TZ ^{c,k}	QZ ^d						
I_p	0,00	0,00	0,00	0,00	+0,00	0,00	+0,00	0,00	0,00	0,00
II_n	0,54	0,66	0,68	0,72	-0,08	0,64	+0,24	0,87	0,22	1,82
III_n	1,59	1,73	1,73	1,73	+0,00	1,73	+0,04	1,77	1,77	1,20
IV_n	1,25	1,23	1,23	1,24	-0,01	1,24	-0,02	1,22	1,17	1,43
V_n	2,43	2,59	2,62	2,65	+0,01	2,65	+0,08	2,74	2,66	2,90
VI_p	4,86	4,79	4,80	4,81	-0,02	4,80	-0,17	4,63	4,68	4,65
VII_p	6,06	5,92	5,89	5,91	-0,08	5,84	-0,10	5,74	5,64	6,02
VIII_n	6,25	6,05	6,06	6,08	-0,03	6,06	-0,14	5,93	5,90	6,16
I_N	0,00	0,00	0,00	0,00	+0,00	0,00	+0,00	0,00	0,00	0,00
II_N	4,36	4,94	4,97	4,99	-0,02	4,97	-0,01	4,96	4,50	6,17
III_N	7,55	7,49	7,51	7,52	-0,02	7,51	-0,13	7,37	7,31	8,08
I_O	0,00	0,00	0,00	0,00	+0,00	0,00	+0,00	0,00	0,00	0,00
II_O	0,62	0,70	0,67	0,67	-0,01	0,67	+0,02	0,69	0,58	0,88
III_O	1,86	2,00	2,00	2,00	-0,02	1,96	+0,34	2,31	2,15	2,66
IV_O	3,98	3,86	3,80	3,79	+0,03	3,82	-0,31	3,51	3,61	3,12
V_O	3,94	3,95	3,90	3,91	+0,04	3,94	+0,12	4,06	4,01	4,39
VI_O	5,20	5,15	5,06	5,06	+0,02	5,08	-0,25	4,83	4,75	5,03
VII_O	4,95	5,18	5,18	5,17	-0,03	5,13	+0,32	5,45	5,33	5,66
VIII_O	5,67	5,78	5,72	5,72	-0,01	5,70	+0,13	5,82	5,68	6,28

^a MP2/aug-cc-pVDZ relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^b CCSD(T)-F12b/aug-cc-pVDZ relatív energiák a CCSD(T)-F12b/aug-cc-pVDZ geometriákkal számítva. ^c CCSD(T)-F12b/aug-cc-pVTZ relatív energiák a CCSD(T)-F12b/aug-cc-pVTZ geometriákkal számítva. ^d CCSD(T)-F12b/aug-cc-pVQZ relatív energiák a CCSD(T)-F12b/aug-cc-pVTZ geometriákkal számítva. ^e A kiegészítő korrekciók összege: $\delta T + \delta(Q) + \Delta_{\text{törzs}} + \Delta_{\text{rel}}$, definíciók: 7.2.1.1–7.2.1.4. egyenletek. ^f Egyensúlyi relatív energiák, definíció: 7.2.1.5. egyenlet. ^g Zérusponyi energia-korrekció CCSD(T)-F12b/aug-cc-pVDZ szinten. ^h Adiabatus relatív energiák, definíció: 7.2.1.6. egyenlet. ⁱ 298,15 K-es relatív entalpiák. ^j 298,15 K-es relatív szabadentalpiák. ^k Az oszlopok semleges glicinre vonatkozó értékei egy, a disszertációhoz kapcsolódó, de alapját nem képező publikációból származnak.⁴

A kiegészítő korrekciókat grafikusán ábrázoltuk, a 19. ábrán láthatóak. A csatolt-klaszter háromszoros (δT) és perturbatíván négyszeres ($\delta(Q)$) gerjesztéseket figyelembe vevő post-CCSD(T) korrekciókról elmondható, hogy cc-pVDZ bázis mellett a járulékuk 0,00 és 0,08 kcal/mol közé esik. Míg a δT tagokról nem vonhatunk le általános következtetéseket, viszont a ($\delta(Q)$) tagok mindig negatív vagy zérus értékek. A kettő tag összege mindig lecsökkenti a relatív energiákat, kivéve az IV_n, II_O, VI_O és IV_O konformereket, az első három szerkezetre nem változik egy századon belül, míg az utolsó esetben megnő egy századdal.



19. ábra: A semleges (bal), és az N- és O-protonált (jobb) glicin konformerekre kapott kiegészítő korrekciók. Definíciók: 7.2.1.1–7.2.1.4. egyenletek.

2. táblázat: A post-CCSD(T) korrelációs korrekciók (kcal/mol) báziskészlet konvergenciája a semleges és protonált glicin konformerekre vonatkozóan. A korrekciókat a 7.2.1.1. és 7.2.1.2. egyenletek definiálják, az alkalmazott bázisok a 3-21G, 6-31G és a cc-pVDZ (VDZ) voltak. A táblázat az [1]-es hivatkozás 2. táblázatának átdolgozása.

Konformer neve	δT			$\delta(Q)$			$\delta T + \delta(Q)$		
	3-21G	6-31G	VDZ	3-21G	6-31G	VDZ	3-21G	6-31G	VDZ
Ip	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00
IIn	-0,03	-0,03	-0,02	-0,03	-0,03	-0,04	-0,06	-0,05	-0,06
IIIn	+0,00	+0,00	+0,00	-0,01	-0,01	-0,01	-0,01	-0,01	-0,01
IVn	+0,01	+0,00	+0,00	-0,01	+0,00	+0,00	+0,00	+0,00	+0,00
Vn	+0,01	+0,01	+0,00	-0,01	-0,01	-0,01	+0,00	+0,00	-0,01
VIp	-0,04	-0,04	-0,03	+0,00	+0,00	+0,00	-0,03	-0,04	-0,03
VIIp	-0,04	-0,04	-0,03	-0,02	-0,02	-0,03	-0,06	-0,06	-0,06
VIIIIn	-0,03	-0,04	-0,03	-0,01	+0,00	+0,00	-0,04	-0,04	-0,03
IN	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00
II_N	-0,02	-0,02	-0,01	-0,01	-0,03	-0,02	-0,03	-0,04	-0,03
III_N	-0,04	-0,04	-0,03	-0,01	+0,00	-0,01	-0,04	-0,05	-0,04
Io	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00	+0,00
IIo	-0,01	+0,00	+0,00	+0,00	+0,00	+0,00	-0,01	+0,00	+0,00
IIIo	+0,00	+0,01	-0,01	-0,06	-0,07	-0,07	-0,06	-0,06	-0,08
IVo	+0,02	+0,02	+0,01	-0,01	-0,01	-0,01	+0,01	+0,02	+0,01
Vo	+0,03	+0,04	+0,02	-0,04	-0,03	-0,04	+0,00	+0,01	-0,02
VIo	+0,00	+0,01	+0,01	-0,01	-0,01	-0,01	+0,00	+0,01	+0,00
VIIo	-0,01	-0,01	-0,02	-0,06	-0,07	-0,06	-0,07	-0,07	-0,08
VIIIo	+0,02	+0,03	+0,01	-0,08	-0,09	-0,09	-0,07	-0,06	-0,08

A post-CCSD(T) korrekciót megvizsgáltuk még kisebb, 6-31G és 3-21G bázisokkal is, melynek eredményei összefoglalva a 2. táblázatban vannak. Jól látszik, hogy kisebb bázisokat alkalmazva is 0,01 kcal/mol-on belül ugyanazokat az értékeket kapjuk, viszont a számítások idején rengeteget spórolhatunk. A következő vizsgált kiegészítő járuléka a törzselektronok korreláltatásából ered ($\Delta_{\text{törzs}}$). Ennek nagysága átlagosan 0,02 kcal/mol, előjele vegyes a semleges glicinre, azonban a protonált szerkezetekre mindig pozitív. A

másodrendű DK relativisztikus korrekció (Δ_{rel}) értéke elhanyagolható, a legtöbb esetben az egy század kcal/mol-t sem éri el.

Az 1. táblázathoz visszatérve megállapíthatjuk, hogy a korrekciók összege nem haladja meg az 0,1 kcal/mol-t sem a semleges, sem pedig a protonált glicinen, a legnagyobb értéket a II_n és VII_p konformerek esetén éri el, ez pedig -0,08 kcal/mol. Látható, hogy a konformerek végső, ezeket a korrekciókat is tartalmazó relatív energiasorrendje (E_e) azonos lesz a korrekciók nélkülivel, ennek hibáját a báziskészlet effektus (QZ és TZ energiák közötti különbség) átlagával és a kiegészítő korrekciók átlagával becsülhetjük. Ezek az értékek a semleges glicinnél $\sqrt{0,03^2+0,02^2} = 0,04$ kcal/mol, az N-protonált glicinnél 0,03 kcal/mol és az O-protonált glicinnél 0,02 kcal/mol. Jelentősebb hozzájárulást képvisel a zérusponyi energia-korrekció (Δ_{ZPE}), aminek értéke átlagosan 0,15 kcal/mol, változó előjellel, míg néhány karbonil-protonált szerkezetre a 0,3 kcal/mol-t is meghaladhatja, az adiabatikus energiasorrendet (ΔH_0) azonban ez sem változtatja meg. Mivel a rezgési frekvenciákat harmonikus közelítésben számítjuk az elhanyagolt anharmonicitás hibát okoz. Ennek kiküszöbölésére alkalmazható a másodrendű rezgési perturbációs elmélet és/vagy a merev rotor analízis,⁴¹ ahogyan azt el is végezték a glicinre³⁷ és a treoninra¹¹³ korábban. Az effektus nagysága az 0,01-0,10 kcal/mol tartományban van, ezért ebből (a legrosszabbat feltételezve) és az E_e bizonytalanságára kapott mennyiségekből kiindulva a 0 K-es relatív entalpia értékeinkre $\sqrt{0,04^2+0,10^2} \approx 0,10$ kcal/mol-os pontosságot javasolunk a semleges és két különböző helyen protonálódott glicin esetén is. Ha a statisztikus termodinamika segítségével kiszámítjuk a termikus járulékokat, akkor megkaphatjuk a 298,15 K-re vonatkozó relatív entalpia értékeket ($\Delta H_{298,15}$). Mivel ezek a mennyiségek nagyon érzékenyek az alacsony frekvenciájú rezgések pontosságára, ezért a 298,15 K-es relatív entalpia (és később szabadentalpia) értékek hibabecslése alaposabb rezgési analízist igényelne, ami azonban túlmutat a disszertáció kitűzött céljain. A 298,15 K-es relatív entalpia értékek valamivel kisebbek, mint a 0 K-esek, kivéve két konformer esetén, ahol növekedés tapasztalható: VI_p (0,04 kcal/mol) és IV_o (0,1 kcal/mol). A II_n és II_N konformerek kiugró értékeket adnak, az eltérés 0,5-0,6 kcal/mol között van, azonban ha az MP2/aug-cc-pVDZ frekvenciákat vesszük figyelembe a termikus korrekciók kiszámításánál, akkor nem lesznek kiugró értékek. Ez arra utalhat, hogy a MOLPRO-val számolt (alacsony) harmonikus rezgési frekvenciák csatolt-klaszter módszerrel még nagyobb bizonytalanságúak lehetnek, mint az MP2 módszer esetén. A relatív szabadentalpia értékek 298,15 K-en történő meghatározásához a TS tagot kell még figyelembe vennünk, ahol S az

entrópiát jelöli. A konformerek különböző entrópia járuléka természetesen a különböző rezgési (különböző rezgési módusok és frekvenciák) és forgási (különböző forgási állandók) járulékokból ered. Az azonos hőmérsékleten vett relatív entalpia értékektől $\pm(0,2-0,8)$ kcal/mol-lal térnek el a szabadentalpia értékek és a II_N és a II_N konformer itt is kiugró értéket mutat, az entrópia korrekció rájuk vonatkozóan 1,59 és 1,68 kcal/mol. Ha ismét az MP2/aug-cc-pVDZ frekvencia értékeket használjuk, akkor ezek a konformerek is átlagos járulékkal rendelkeznek, vagyis a nagy eltérés az alacsony rezgési frekvenciák bizonytalanságának tudható be.

7.2.2. Cys és CysH⁺: nagy pontosságú geometriák és energiák²

A semleges ciszteinek egy nagyságrenddel több konformerjét azonosítottuk, mint a glicinnek, sőt, a protonálódása három funkciós csoporton is lejátszódhat, szintén nagymennyiségű geometriát eredményezve. Ezeket MP2/aug-cc-pVDZ szinten meghatároztuk, azonban a magasabb szintű számolások előtt a számítógépes kapacitások és a számítási idők figyelembe vételével kompromisszumokat kell kötnünk, ezért úgy döntöttünk, hogy minden esetben csak a tíz legalacsonyabb relatív energiájú konformert vetjük alá további vizsgálatoknak. Az általunk vizsgált körülmények között (0 és 298,15 K) ezeknek a populációja számottevő, így a későbbi protonaffinitás és gázfázisú bázicitás értékek meghatározására nézve ez nem fog hibát jelenteni. A tíz legmélyebb relatív energiájú szerkezet kiválasztása minden esetben az MP2/aug-cc-pVDZ relatív energiák alapján történt, ennek kiemelése azért fontos, mert magasabb szinten könnyen elképzelhető, hogy változik az energiasorrend.

A négyszer tíz MP2/aug-cc-pVDZ szintű geometriát CCSD(T)-F12a/cc-pVDZ-F12 szinten optimáltuk (a semleges formára a harmonikus rezgési analízist is elvégeztük), majd ezeken a geometriákon cc-pVTZ-F12 és cc-pVQZ-F12 bázisokkal single-point energiákat számoltunk. Az utóbbi bázissal a CCSD(T)-F12b módszert is kombináltuk, az irodalomban fellelhető források szerint QZ bázisokkal ez ad pontosabb értéket.^{104,106,110} A továbbiakban meghatározott korrekciók és energiák fajtája analóg a glicinnél felsoroltakkal, azonban számos változás van az alkalmazott módszerekben és/vagy bázisokban, ezért jobbnak látom újradefiniálni őket.

A megszakított csatolt-klaszter sorfejtéshez kapcsolódó korrekcióinkat (δT és $\delta(Q)$) a következő egyenletekkel számoltuk, a glicinnél elégségesnek bizonyult 6-31G bázissal:

$$\delta T = \text{CCSDT}/6-31G - \text{CCSD(T)}/6-31G, \quad (7.2.2.1.)$$

$$\delta(Q) = \text{CCSDT}(Q)/6-31G - \text{CCSDT}/6-31G. \quad (7.2.2.2.)$$

A törzselektronok korreláltatása a szén, nitrogén és oxigénatomok esetén az $1s^2$ elektronokat foglalja magába, azonban a kén esetén csupán a közvetlenül vegyértékhej alatti, $2s^2$ és $2p^6$ elektronokat korreláltatjuk a vegyértékelektronok mellett:

$$\begin{aligned} \Delta_{\text{törzs}} = & \text{AE-CCSD(T)-F12a/cc-pCVTZ-F12} \\ & - \text{FC-CCSD(T)-F12a/cc-pCVTZ-F12}. \end{aligned} \quad (7.2.2.3.)$$

A relativisztikus effektust ismét a másodrendű Douglas–Kroll korrekcióval vettük figyelembe:

$$\begin{aligned} \Delta_{\text{rel}} = & \text{DK-AE-CCSD(T)/aug-cc-pwCVTZ-DK} \\ & - \text{AE-CCSD(T)/aug-cc-pwCVTZ}. \end{aligned} \quad (7.2.2.4.)$$

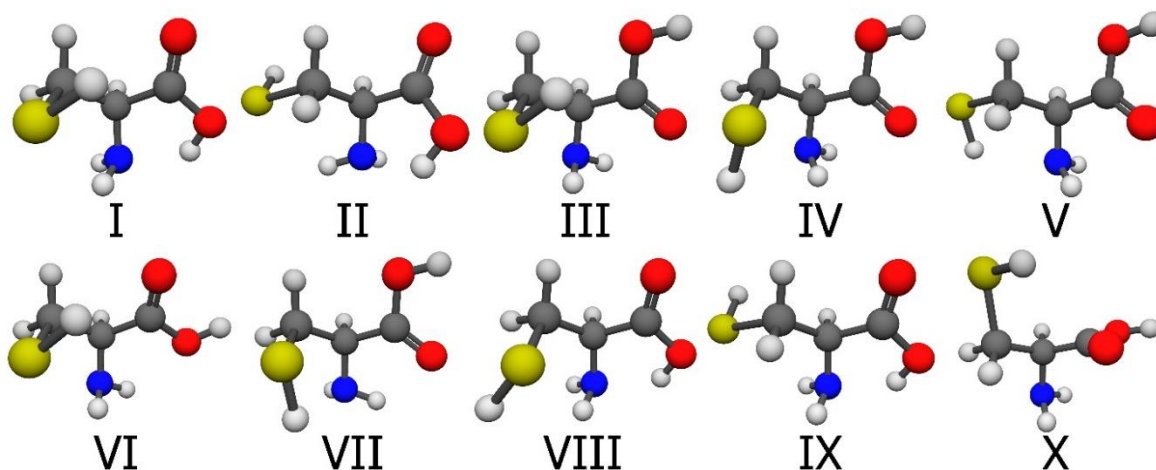
Az egyensúlyi energiákat a kiegészítő korrekciók illetve a QZ single-point energiák összegeként kapjuk:

$$E_e = \text{CCSD(T)-F12b/cc-pVQZ-F12} + \delta T + \delta(Q) + \Delta_{\text{törzs}} + \Delta_{\text{rel}} \quad (7.2.2.5.)$$

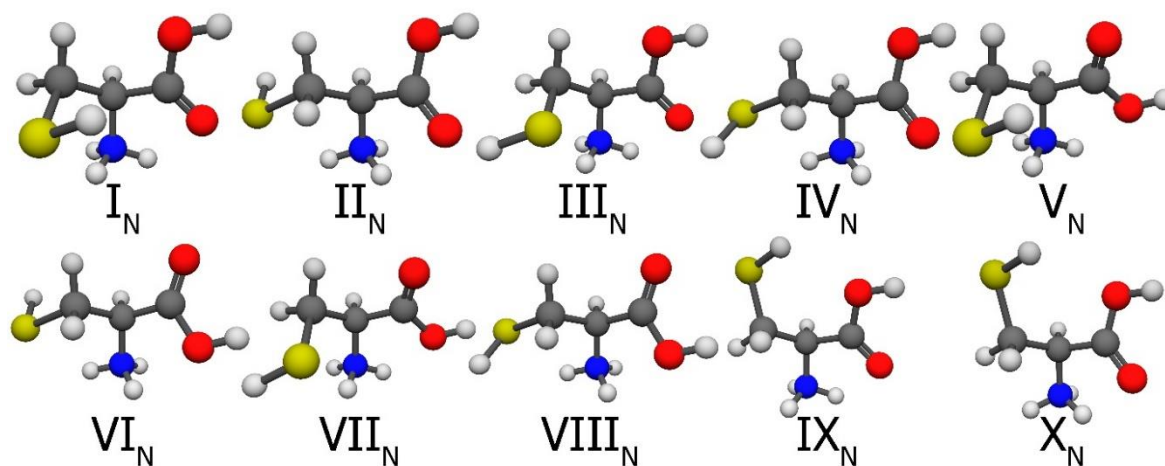
A zéruspointi-energia korrekciót (Δ_{ZPE}) az MP2/aug-cc-pVDZ harmonikus rezgési frekvenciákból kapjuk, majd ennek segítségével meghatározzuk a 0 K-es adiabatikus relatív energiákat/relatív entalpiákat:

$$H_0 = E_e + \Delta_{\text{ZPE}} \quad (7.2.2.5.)$$

A kiválasztott, CCSD(T)-F12a/cc-pVDZ-F12 szinten optimált semleges cisztein konformerek a 20. ábrán, az N-protonáltak a 21. ábrán, az O-protonáltak a 22. ábrán és az S-protonáltak pedig a 23. ábrán láthatóak. A római számok a CCSD(T)-F12b/cc-pVQZ-F12 szinten a CCSD(T)-F12a/cc-pVDZ-F12 geometriákon számított relatív energiák sorrendjét követik, az alsó index pedig a protonálódás helyét jelöli. Mint várható volt, a szerkezeteket egy vagy több intramolekuláris hidrogénkötés stabilizálja. Az amino-protonálódás gátolja a nitrogénatom akceptor szerepét, míg az O-protonálódás ugyanezen karaktert gátolja a karbonilcsoport oxigénatomján. A cisztein globális minimum szerkezete (20. ábra) a glicin II_n és az alanin¹¹⁹ II_a szerkezetére hasonlít, ha az alfa és béta szénatomhoz tartozó torziós szögeket nézzük, illetve szintén egy, a nitrogénatom nemkötő elektronpárja és a hidroxilcsoport hidrogénje között kialakuló hidrogénhid stabilizálja a molekulát. Ezen kívül

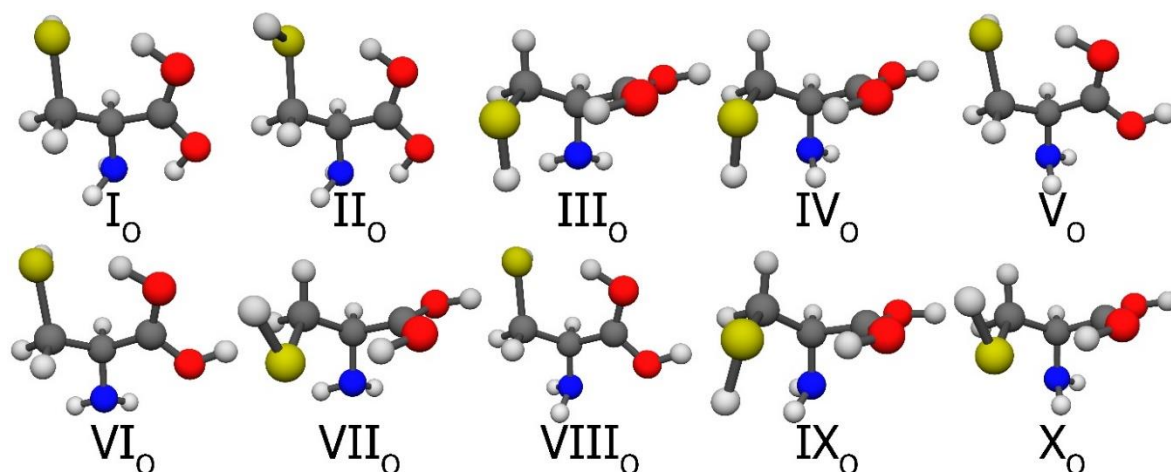


20. ábra: A semleges cisztein tíz konformerének CCSD(T)-F12a/cc-pVDZ-F12 geometriája. Kiválasztásuk az MP2/aug-cc-pVDZ relatív energiák alapján történt. A római számok a CCSD(T)-F12b/cc-pVQZ-F12 single-point energiák sorrendjét követik. Az ábra a [2]-es hivatkozás 7. ábrája.

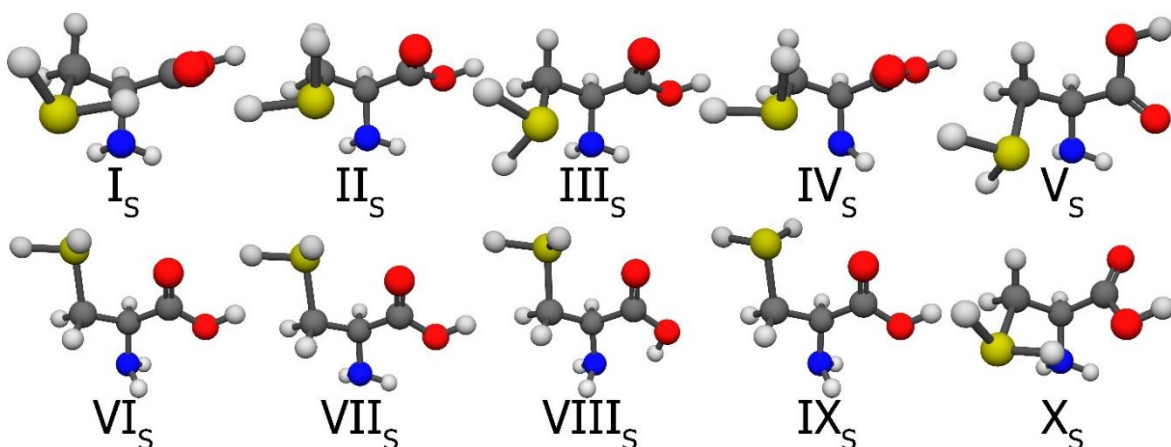


21. ábra: Az N-protonált cisztein tíz konformerének CCSD(T)-F12a/cc-pVDZ-F12 geometriája. Kiválasztásuk az MP2/aug-cc-pVDZ relatív energiák alapján történt. A római számok a CCSD(T)-F12b/cc-pVQZ-F12 single-point energiák sorrendjét követik, míg az „N” az alsóindexben a protonálódás helyét jelöli. Az ábra a [2]-es hivatkozás 8. ábrája.

még egy jóval gyengébb kölcsönhatás figyelhető meg a tiolcsoport hidrogénje és a karbonilcsoport oxigénjének nemkötő elektronpárja között (távolságuk $\sim 2,7$ ångström). Az amino-protonált minimumot szintén a glicinhez és alaninhoz hasonlóan egy erős intramolekuláris hidrogénkötés stabilizálja, amely a protonált aminocsoport és a karbonilcsoport oxigénjének nemkötő elektronpárja között jön létre. A karbonil-protonált szerkezeteknél a glicinnel ellentétben azonosíthatjuk stabilizáló kölcsönhatások létrejöttét, a globális minimumban rögtön egy erősebb O–H \cdots N és egy gyengébb O–H \cdots S hidrogénkötés kialakulását, de ezek a motívumok visszatérőek a többi O-protonált konformer esetén is. A tiol-protonált globális minimum szerkezet esetében a protonált tiolcsoport egyik hidrogénje



22. ábra: Az O-protonált cisztein tíz konformerének CCSD(T)-F12a/cc-pVDZ-F12 geometriája. Kiválasztásuk az MP2/aug-cc-pVDZ relatív energiák alapján történt. A római számok a CCSD(T)-F12b/cc-pVQZ-F12 single-point energiák sorrendjét követik, míg az „O” az alsóindexben a protonálódás helyét jelöli. Az ábra a [2]-es hivatkozás 9. ábrája.



23. ábra: Az S-protonált cisztein tíz konformerének CCSD(T)-F12a/cc-pVDZ-F12 geometriája. Kiválasztásuk az MP2/aug-cc-pVDZ relatív energiák alapján történt. A római számok a CCSD(T)-F12b/cc-pVQZ-F12 single-point energiák sorrendjét követik, míg az „S” az alsóindexben a protonálódás helyét jelöli. Az ábra a [2]-es hivatkozás 10. ábrája.

kölcsönhatásba kerül a nitrogénatom és az egyik oxigénatom nemkötő elektronpárjával. Ezeknél a szerkezeteknél nem figyelhető meg erős stabilizáló kölcsönhatás kialakulása. A már kísérletileg is azonosított konformerekkel való összehasonlításunk a 3. táblázatban látható. Mint az látható, az összes, rezgési vagy forgási spektrum alapján kísérletileg meghatározott konformert sikerült elméleti úton is felderítenünk. A nagyszámú szerkezetből adódóan gyakran vannak átfedések a spektrumokban, amelyek az egyes konformerek egyértelmű azonosítását megnehezítik, ezért jelenlétük nem igazolható minden kétséget kizáróan. Az energetikai viszonyokat is figyelembe véve láthatunk egyezéseket a sorrendek-

3. táblázat: A kísérletileg kimutatott semleges cisztein konformerek összevetése az általunk találtakkal.

Jelen munka ^{a,2}	Dobrowolski ^{b,60}	Alonso ^{c,62}	Tarczay ^{b,63}
1/I/(1)	1	IIb	1
2/II/(8)	5 ^e	Ib	5
3/III/(2)	2	IIa	2
4/IV/(3)	7 ^e		
5/V/(4)	3	Ia	3
6/VII/(5)			
7/VI/(6)		III _β b	9
8/VIII/(9)			
9/X/(7)			
10/IX/(10)		III _β c	13 ^e
65/-/-			14 ^f

^a A(z) római számok (arab számok) [zárójeles arab számok] a CCSD(T)-F12b/cc-pVQZ-F12 relatív energiák (MP2/aug-cc-pVDZ relatív energiák) [298.15 K-es relatív szabadentalpiák] sorrendjét követik. ^b Az elnevezések a B3LYP/aug-cc-pVTZ elméleti szintű relatív energiáit követik. ^c A konformerek elnevezése hidrogénkötéssel kialakuló motívumokon alapszik. ^e Ezen konformerek jelenlétét sem kizárni, sem megerősíteni nem tudták a hozzájuk tartozó publikáció szerint. ^f Elnevezése nem követi a B3LYP/aug-cc-pVTZ elméleti szintű relatív energiák sorrendjét, ez a szerkezet egy elméletileg jósolt, magasabb energiájú és rövid felezési idejű konformer.

ben is, azonban vélhetően a mi eredményink a magasabb elméleti szint használatának következtében pontosabbak. A protonált cisztein kísérleti vizsgálatáról legjobb tudomásunk szerint nem lelhető fel irodalmi forrás.

A semleges ciszteinhez tartozó relatív energiák, entalpiák, szabadentalpiák és a megfelelő korrekciók a 4. táblázatban vannak összefoglalva. A római számok a legmagasabb elméleti szinten (CCSD(T)-F12b/cc-pVQZ-F12) meghatározott relatív energiák sorrendjét követik, míg az arab számok az MP2/aug-cc-pVDZ relatív energiáikét. A különbség az MP2/aug-cc-pVDZ és a CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák között két részből tevődik össze: egyrészt természetesen a magasabb elméleti szint pontosabb eredményeket szolgáltat, másrészt pedig a geometriában is változás történik az újabb optimalás során. Ezt, a geometriából adódó komponens megvizsgálhatjuk, ha kiszámoljuk a különbségét az MP2 és a CC módszerrel az MP2 geometrián kapott relatív energiaértékeknek. A semleges cisztein esetén ez a járulék néhány század kcal/mol, ami azt jelenti, hogy az MP2 módszer már elég pontos geometriához konvergál. A két módszerrel a megfelelő optimalások után kapott relatív energiák jó egyezést mutatnak néhány század kcal/mol-on belül, kivéve az utolsó két konformer esetében, ahol az eltérés nagyobb, $-0,37$ és $+0,15$ kcal/mol. Háromszoros és négyszeres ζ bázisokkal single-point számolásokat elvégezve az értékek $0,01$ – $0,04$ kcal/mol-lal változnak, ami a jó báziskészlet konvergenciát mutatja. A cc-pVQZ-F12 báziskészletet mind a CCSD(T)-F12a, mind a CCSD(T)-F12b módszerrel kipróbáltuk,

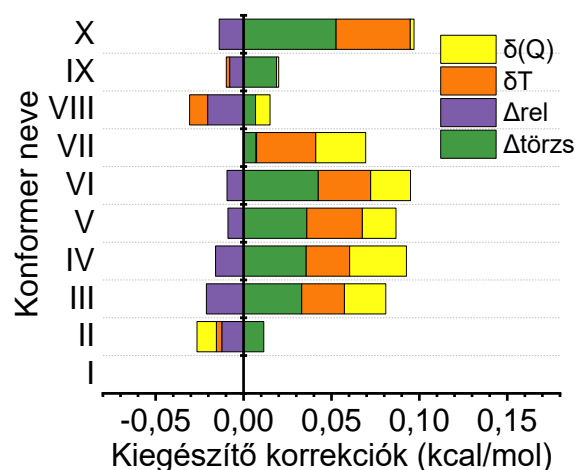
4. táblázat: A továbbvizsgált semleges cisztein konformerek relatív egyensúlyi energiái, a kiegészítő korrekciói, 0 és 298,15 K-es relatív entalpiái és 298,15 K-es relatív szabadentalpiái. A mértékegység minden esetben kcal/mol. A táblázat a [2]-es hivatkozás 2. táblázatának átdolgozása.

Konformer neve ^a	MP2	CC//MP2	CCSD(T)-F12a			CCSD(T)-F12b	Σ_{korrr}^h	ΔE_e^i	Δ_{ZPE}^j	ΔH_0^k	$\Delta H_{298,15}^l$	$\Delta G_{298,15}^m$
	DZ ^b	DZ ^c	DZ ^d	TZ ^e	QZ ^f	QZ ^g						
I/1	0,00	0,00	0,00	0,00	0,00	0,00	+0,00	0,00	+0,00	0,00	0,00	0,00
II/2	1,48	1,48	1,49	1,53	1,53	1,53	-0,01	1,52	-0,06	1,46	1,58	1,32
III/3	1,63	1,63	1,61	1,62	1,61	1,60	+0,06	1,66	-0,50	1,16	1,42	0,43
IV/4	1,71	1,75	1,74	1,75	1,74	1,74	+0,08	1,81	-0,44	1,37	1,62	0,72
V/5	1,81	1,78	1,78	1,81	1,80	1,80	+0,08	1,87	-0,44	1,43	1,68	0,81
VI/7	1,98	1,95	1,95	1,95	1,94	1,94	+0,09	2,03	-0,28	1,75	1,86	1,26
VII/6	1,97	1,95	1,95	1,96	1,95	1,95	+0,07	2,02	-0,55	1,46	1,65	1,11
VIII/8	2,27	2,10	2,10	2,11	2,11	2,11	-0,02	2,09	-0,15	1,94	2,08	1,73
IX/10	2,54	2,15	2,17	2,19	2,19	2,19	+0,01	2,20	-0,27	1,93	2,54	1,90
X/9	2,39	2,54	2,55	2,58	2,58	2,58	+0,08	2,66	-0,36	2,30	2,13	1,31

^a A(z) római számok (arab számok) a CCSD(T)-F12b/cc-pVQZ-F12 (MP2/aug-cc-pVDZ) relatív energiák sorrendjét követik. ^b MP2/aug-cc-pVDZ relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^c CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^d CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^e CCSD(T)-F12a/cc-pVTZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^f CCSD(T)-F12a/cc-pVQZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^g CCSD(T)-F12b/cc-pVQZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^h A kiegészítő korrekciók összege: $\delta T + \delta(Q) + \Delta_{\text{törzs}} + \Delta_{\text{rel}}$, definíciók: 7.2.2.1–7.2.2.4. egyenletek. ⁱ Egyensúlyi relatív energiák, definíció: 7.2.2.5. egyenlet. ^j Zérus ponti energia-korrekció MP2/aug-cc-pVDZ szinten. ^k Adiabátikus relatív energiák, definíció: 7.2.2.6. egyenlet. ^l 298,15 K-es relatív entalpiák. ^m 298,15 K-es relatív szabadentalpiák.

hiszen az irodalomban QZ bázisokhoz már az F12b változatot javasolják, azonban a semleges ciszteinnél a különbség elhanyagolhatóan adódott, hiszen az egy század kcal/mol-t csupán a III-as konformer esetén érte el.

A semleges ciszteinre vonatkozó kiegészítő korrekciókat a 24. ábrán láthatóak. Közülük a belső elektronok korreláltatása eléggé hasznosnak bizonyul, mert a $\Delta_{\text{törzs}}$ értéke a 0,05 kcal/mol-t is elérheti (X-es konformer), átlagosan pedig 0,03 kcal/mol. A Douglas–Kroll relativisztikus korrekció (Δ_{rel}) kevésbé meghatározó, mivel a 0,02 kcal/mol-t nem haladja meg, de általában inkább egy század kcal/mol körüli a járuléka (minden esetben negatív az előjele). A post-CCSD(T) korrekciók is fontos szerepet játszanak a kémiai pontosság elérésében. A δT és a $\delta(Q)$ járulékok összege 0,05–0,06 kcal/mol a legtöbb konformerre, kivétel ez alól a II-es konformer, amely esetében -0,01 kcal/mol (az egyetlen negatív érték), a X-es konformer, ahol 0,04 kcal/mol értéket vesz fel, a IX-es konformernél pedig elhanyagolható. Összefoglalva az általunk vizsgált kiegészítő korrekciók a relatív energiákat körülbelül 0,05 kcal/mol-lal változtatják meg, de némely szerkezetre akár a 0,08–0,09 kcal/mol-t is elérhetik.



24. ábra: A továbbvizsgált semleges cisztein konformerekre kapott kiegészítő korrekciók. Definíciók: 7.2.2.1–7.2.2.4. egyenletek.

Ahogy az a 4. táblázatban is látszik, a VI-os és a VII-es konformer nagyon közel van egymáshoz energiában, magasabb elméleti szinten a VI-osnak mélyebb az energiája, de a korrekciók figyelembevételére után már a VII-es kerül előrébb a sorban. A relatív egyensúlyi energiák hibáját a glicinnél ismertetett módon becsülhetjük, ez a semleges ciszteinnél $\sim 0,06$ kcal/mol. A harmonikus rezgési frekvenciákat az MP2/aug-cc-pVDZ elméleti szint mellett magasabb, CCSD(T)-F12a/cc-pVDZ-F12 szinten is meghatároztuk. Figyelembe véve a jóval nagyobb számítási igényt és azt, hogy az eredmények csupán csekély mértékben (átlagosan 0,08 kcal/mol) változnak, későbbi vizsgálataink során eltekintünk a frekvenciák meghatározásától magasabb szinten (viszont figyelembe vesszük a pontosság megbecsülésénél). A 4. táblázatban az MP2/aug-cc-pVDZ Δ_{ZPE} korrekciókat tüntettük fel. Mivel ezek a számolások harmonikus közelítésben történtek, várhatóan egy 0,1 kcal/mol-os bizonytalanság is megjelenik a zérus ponti energia-korrekcióban, amely az elhanyagolt anharmonicitásból ered.^{37,41,113} A 0 K-es relatív entalpiaértékeknél sorrendváltás figyelhető meg a II–III és VIII–IX konformerek esetén, míg a különbség a VI–VII szerkezetek között megnő. A hibabecslésnél a glicinnél említett tényezőknél felül az MP2–CC harmonikus rezgési frekvenciák átlagos eltérését is figyelembe vesszük, így egy $\sim 0,15$ kcal/mol-os közelítő értéket kapunk. A termikus hozzájárulások figyelembevételével és az entrópiás tagok kivonása után megkapjuk a 298,15 K-es relatív entalpia és relatív szabadentalpia értékeket, azonban ezek pontossága bizonytalan, a glicinnél említett, alacsony frekvenciákra való érzékenységből kifolyólag.

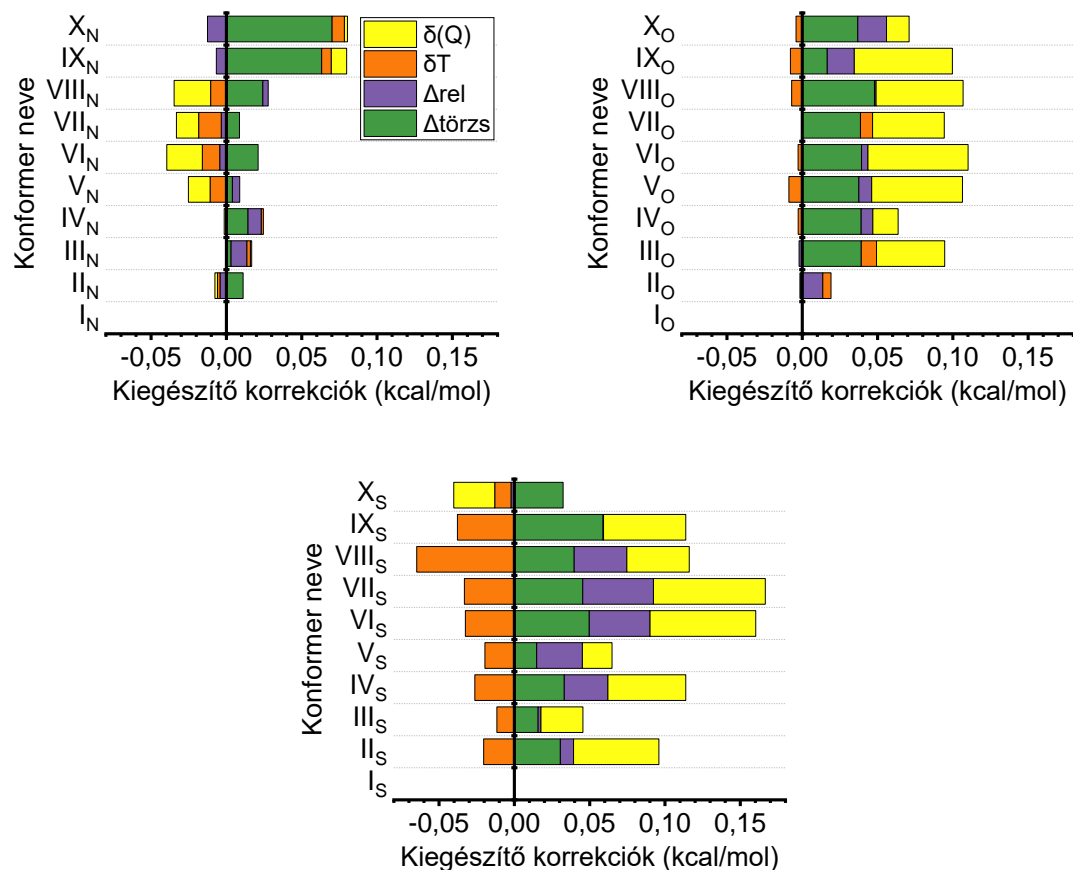
5. táblázat: A továbbvizsgált protonált cisztein konformerek relatív egyensúlyi energiái, a kiegészítő korrekciói, 0 és 298,15 K-es relatív entalpiái és 298,15 K-es relatív szabadentalpiái. A mértékegység minden esetben kcal/mol. A táblázat a [2]-es hivatkozás 2. táblázatának átdolgozása.

Konformer neve ^a	MP2	CC//MP2	CCSD(T)-F12a			CCSD(T)-F12b	Σ_{korrt}^h	ΔE_e^i	Δ_{ZPE}^j	ΔH_0^k	$\Delta H_{298,15}^l$	$\Delta G_{298,15}^m$
	DZ ^b	DZ ^c	DZ ^d	TZ ^e	QZ ^f	QZ ^g						
I _N /N1	0,00	0,00	0,00	0,00	0,00	0,00	+0,00	0,00	+0,00	0,00	0,00	0,00
II _N /N2	0,52	0,52	0,54	0,56	0,55	0,55	+0,00	0,56	+0,06	0,62	0,69	0,54
III _N /N3	1,27	1,45	1,44	1,39	1,38	1,39	+0,02	1,40	-0,14	1,27	1,37	1,04
IV _N /N4	1,62	1,73	1,75	1,71	1,70	1,71	+0,02	1,73	-0,06	1,67	1,83	1,36
V _N /N5	2,77	3,16	3,19	3,20	3,20	3,20	-0,02	3,18	+0,00	3,18	3,19	3,20
VI _N /N6	4,07	4,40	4,38	4,43	4,42	4,43	-0,02	4,41	-0,06	4,35	4,48	4,01
VII _N /N7	4,46	4,94	4,96	4,93	4,93	4,93	-0,02	4,90	-0,15	4,76	4,88	4,55
VIII _N /N8	5,27	5,69	5,68	5,67	5,67	5,67	-0,01	5,67	-0,18	5,49	5,72	4,96
IX _N /N10	7,56	7,15	7,17	7,18	7,19	7,20	+0,07	7,27	-0,14	7,14	7,32	6,73
X _N /N9	7,54	7,31	7,31	7,31	7,32	7,33	+0,07	7,39	-0,18	7,21	7,44	6,33
I _O /O1	0,00	0,00	0,00	0,00	0,00	0,00	+0,00	0,00	+0,00	0,00	0,00	0,00
II _O /O2	1,00	1,12	1,12	1,07	1,07	1,07	+0,02	1,09	-0,08	1,01	1,07	0,83
III _O /O3	6,63	6,71	6,76	6,81	6,83	6,82	+0,09	6,91	-0,70	6,21	6,24	5,91
IV _O /O4	7,25	7,19	7,28	7,27	7,26	7,26	+0,06	7,32	-0,35	6,97	7,05	6,80
V _O /O6	7,84	7,46	7,56	7,57	7,56	7,55	+0,10	7,65	-0,61	7,04	7,34	7,01
VI _O /O7	7,93	7,55	7,65	7,66	7,66	7,65	+0,11	7,76	-0,55	7,21	7,15	6,83
VII _O /O5	7,62	7,70	7,78	7,78	7,79	7,79	+0,09	7,88	-0,54	7,34	7,29	6,72
VIII _O /O8	8,24	7,84	7,91	7,91	7,90	7,89	+0,10	7,99	-0,51	7,48	7,58	7,13
IX _O /O9	8,34	7,89	7,98	7,97	7,96	7,95	+0,09	8,05	-0,41	7,63	7,81	7,31
X _O /O10	8,43	8,51	8,60	8,54	8,53	8,53	+0,07	8,60	-0,51	8,09	8,25	7,78
I _S /S1	0,00	0,00	0,00	0,00	0,00	0,00	+0,00	0,00	+0,00	0,00	0,00	0,00
II _S /S3	1,97	0,66	0,81	0,83	0,79	0,79	+0,08	0,86	+0,84	1,70	2,14	2,03
III _S /S2	1,94	0,40	1,11	1,15	1,11	1,11	+0,03	1,14	+0,87	2,01	1,81	1,76
IV _S /S4	3,29	1,72	2,14	2,13	2,08	2,08	+0,09	2,17	+0,61	2,78	3,11	1,77
V _S /S5	3,44	2,17	2,49	2,48	2,44	2,43	+0,05	2,48	+0,66	3,13	3,31	2,62
VI _S /S6	4,44	2,66	3,06	3,03	2,98	2,98	+0,13	3,11	+0,78	3,88	4,00	3,64
VII _S /S7	4,52	2,71	3,12	3,07	3,03	3,02	+0,13	3,16	+0,70	3,86	3,94	3,75
VIII _S /S9	5,71	4,69	4,01	3,98	3,95	3,94	+0,05	4,00	+1,18	5,18	6,19	6,69
IX _S /S10	6,15	3,63	4,68	4,66	4,62	4,61	+0,08	4,69	+0,78	5,47	4,99	4,51
X _S /S8	5,32	4,26	5,03	5,07	5,04	5,04	-0,01	5,03	+0,69	5,72	5,40	5,22

^a A(z) római számok (arab számok) a CCSD(T)-F12b/cc-pVQZ-F12 (MP2/aug-cc-pVDZ) relatív energiák sorrendjét követik, a betűk pedig a protonálódás helyére utalnak. ^b MP2/aug-cc-pVDZ relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^c CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^d CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^e CCSD(T)-F12a/cc-pVTZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^f CCSD(T)-F12a/cc-pVQZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^g CCSD(T)-F12b/cc-pVQZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^h A kiegészítő korrekciók összege: $\delta T + \delta(Q) + \Delta_{\text{törzs}} + \Delta_{\text{rel}}$, definíciók: 7.2.2.1–7.2.2.4. egyenletek. ⁱ Egyensúlyi relatív energiák, definíció: 7.2.2.5. egyenlet. ^j Zérusponyi energia-korrekció MP2/aug-cc-pVDZ szinten. ^k Adiabatus relatív energiák, definíció: 7.2.2.6. egyenlet. ^l 298,15 K-es relatív entalpiák. ^m 298,15 K-es relatív szabadentalpiák.

A protonált ciszteinre kapott eredményeinket az 5. táblázatban foglaltuk össze. A jelölés hasonló elvet követ, mint a semleges aminosav esetén, az alsó index a sorrendet (CCSD(T)-F12b/cc-pVQZ-F12 sorrendben), az első karakter (MP2/aug-cc-pVDZ sorrendben) a protonálódás helyét jelöli. Az MP2/aug-cc-pVDZ szerkezetek elég pontosak az aminosav protonáció esetén, a geometria effektus 0,02 kcal/mol átlagosan, a karbonil-protonált konformerek esetén ez az effektus valamivel nagyobbá válik, átlagosan 0,06 kcal/mol-lá. A tiol-protonálódásnál azonban azt találtuk, hogy az MP2 geometriák nem kielégítőek, hiszen itt az eltérés 0,55 kcal/mol átlagban, de az 1,00 kcal/mol-t is meghaladja az IXs konformernél. Az átlagos különbség az MP2/aug-cc-pVDZ és CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák között (természetesen az adott elméleti szinten elvégzett geometriai optimalizálás után) 0,29 kcal/mol az N-, 0,21 kcal/mol az O- és jóval nagyobb, 1,15 kcal/mol az S-protonált minimumok esetében. Magában a relatív energiasorrendben is több változás történik az elméleti szint növelésével, a csatolt-klaszter módszert alkalmazva viszont már mindegyik bázis esetén ugyanazt a sorrendet kapjuk, a jó báziskészlet konvergencia itt is megfigyelhető. Ez alól talán a tiol-protonált szerkezetek a kivételek, ahol a TZ és QZ bázisokkal kapott eredmények között még mindig 0,04 kcal/mol az átlagos eltérés. Jelentős változást a semleges aminosavhoz hasonlóan itt sem tapasztalunk attól függően, hogy az explicit-korrelált CC módszerek közül az F12a vagy F12b változatot használjuk.

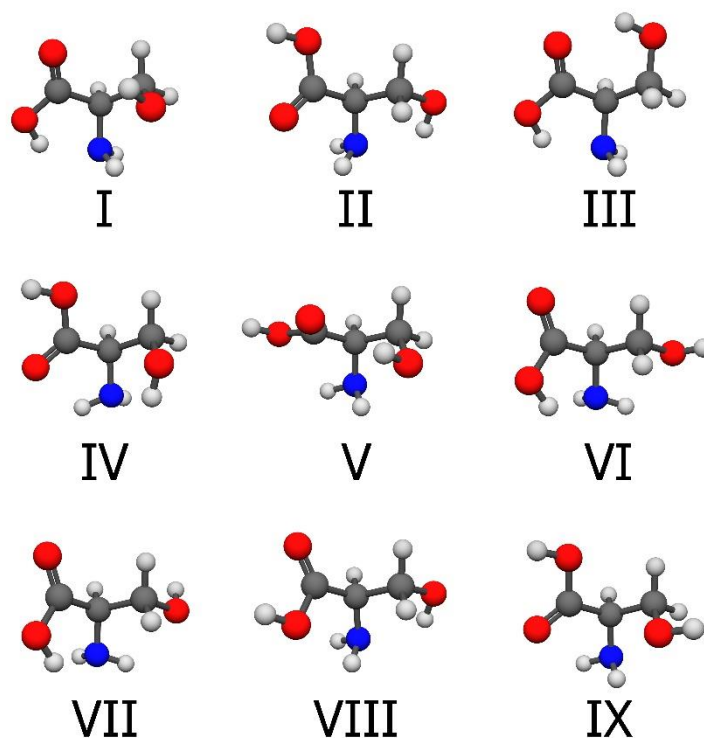
A protonált ciszteinre kapott kiegészítő korrekciókat a 25. ábrán tüntettük fel. Ezeket összegezve az 5. táblázatban eljutunk a végső, relatív egyensúlyi energiákhoz, amelyek sorrendjében a semleges ciszteinnel ellentétben nem történt változás, a hibahatárookra pedig rendre 0,03; 0,08 és 0,08 kcal/mol-os becsléseket tehetünk az N-, O- és S-protonált konformerek esetén. A zérusponthoz viszonyított energia-korrekciók az N-protonált geometriák esetén ~0,11 kcal/mol-osak, előjelük vegyes, az O-protonáltaknál ezek a tagok negatívak minden esetben, átlaguk 0,47 kcal/mol, végül pedig az S-protonáltaknál pozitívak, átlagosan 0,79 kcal/mol-osak. Az adiabatikus relatív energiáknál az energiasorrendben egy helyen történik csere, mégpedig a VI és VII között. Ezen mennyiségek pontosságát a semleges formához hasonlóan, körülbelül $\pm 0,15$ kcal/mol-lal jellemezhetjük. A hőmérsékletet 298,15 K-re „emelve” a relatív entalpiák néhány tizedet változnak, kivétel ez alól például a VIII konformer, ahol több, mint 1 kcal/mol-t. Az entrópiák figyelembevételével meghatározott relatív szabadentalpiák pedig 0,01 és 1,34 kcal/mol közötti értékkel alacsonyabbak (kivéve a VIII esetén, ahol növekedést tapasztaltunk) a megfelelő entalpia értékeknél.



25. ábra: A továbbvizsgált protonált cisztein konformerekre kapott kiegészítő korrekciók. Definíciók: 7.2.2.1–7.2.2.4. egyenletek.

7.2.3. Ser és SerH⁺: nagypontosságú geometriák és energiák³

A semleges és protonált minimum konformerek nagy száma akárcsak a cisztein esetében, a szerinnél sem teszi lehetővé, hogy mindegyik szerkezetre további csatolt-klaszter számításokat futtassunk le. Mivel az általunk meghatározni kívánt termodinamikai mennyiségek 298,15 K-en csekély mértékben függenek a magasabb relatív energiájú konformerektől, ezért a magasabb elméleti szintű számításokat és a korrekciók meghatározását ismét csak a tíz, MP2/aug-cc-pVDZ elméleti szinten legalacsonyabb relatív energiával rendelkező szerkezetekre végeztük el. A korrekciók definíciói és a számítások menete megegyeznek a 7.2.2.1–7.2.2.5. egyenleteknél felvázoltakkal. A CC geometriai optimalások után kapott kilenc semleges szerin minimum konformert a 26. ábrán tüntettük fel. A minimum geometriák nevezéktanában a római számok a CCSD(T)-F12b/cc-pVQZ-F12 relatív energiák, az arab számok pedig az MP2/aug-cc-pVDZ relatív energiák sorrendjét követik. Az, hogy a tíz szerkezet helyett végül miért csak kilencet említünk azzal magyaráz-



26. ábra: A semleges szerin kilenc konformerének CCSD(T)-F12a/cc-pVDZ-F12 geometriája. Kiválasztásuk az MP2/aug-cc-pVDZ relatív energiák alapján történt. A római számok a CCSD(T)-F12b/cc-pVQZ-F12 single-point energiák sorrendjét követik. Az ábra a [3]-as hivatkozás 4. ábrája.

ható, hogy az 5-ös konformer magasabb szinten a globális minimum szerkezetéhez, vagyis az I-es konformerhez konvergál. Annak eldöntésére, hogy az 1-es vagy az 5-ös konformer tartozik a CC szintű globális minimumhoz, összehasonlítottuk a forgási állandóikat az I-es konformerével és azt kaptuk, hogy az 1-es. Ez azért lesz fontos, mert későbbi számításinknál fel fogjuk használni ezen a szinten meghatározott harmonikus rezgési frekvenciákat, amelyek viszont eléggé különböznek. Az összes semleges szerin konformert, kivéve az V-öt és a IX-est, egy erős O–H···N kölcsönhatás stabilizálja, amely az aminocsoport és a két hidroxilcsoport egyike között jön létre. Gyengébb, O–H···O és N–H···O intramolekuláris kölcsönhatásokat is megfigyelhetünk az összes konformer esetén. Kísérleti oldalról már több kutató vizsgálta a semleges szerin konformereket. Az irodalomban fellelhető eredményeket a 6. táblázatban gyűjtöttük össze és hasonlítottuk az általunk találtakhoz.

A saját szerkezeteinket háromféle nevezéktan szerint (MP2/aug-cc-pVDZ relatív energiák, CCSD(T)-F12b/cc-pVQZ-F12 relatív energiák és a relatív szabadentalpiák) tüntettük fel és

6. táblázat: A kísérletileg kimutatott semleges szerin konformerek összevetése az általunk találtakkal. A táblázat a [3]-as hivatkozás 2. táblázatának átdolgozása.

Jelen munka ^{a,3}	Blanco ^{b,69}	Jarmelo ^{c,65}	Lambie ^{d,68}	Najbauer ^{e,f,70}
1/I/(2)	IIb	2	SER2	2
2/II/(1)	Ia	1	SER1	1
3/IV/(4)	IIb	4	SER4	4
4/III/(3)	IIc	3	SER3	3
5/I/(2) ^g				
6/V/(5)	III _β b	7		7 ^h
7/VI/(8)	IIa	5		6
8/VII/(9)		9		
9/VIII/(6)		6		5
10/IX/(7)				
11		8		
12	III _β c			10 ^h

^a A(z) római számok (arab számok) [zárójeles arab számok] a CCSD(T)-F12b/cc-pVQZ-F12 relatív energiák (MP2/aug-cc-pVDZ relatív energiák) [298.15 K-es relatív szabadentalpiák] sorrendjét követik. ^b A konformerek elnevezése hidrogénkötéssel kialakuló motívumokon alapszik. ^c A konformerek elnevezése a B3LYP/6-311++G(d,p) relatív energiák zéruspontri energia-korrekció utáni sorrendjén alapszik. ^d A konformerek elnevezése a B3LYP/6-311++G(d,p) relatív energiák zéruspontri energia-korrekció utáni sorrendjén alapszik. ^e A konformerek elnevezése a 0 K-es relatív szabadentalpiaértékeket követi (B3LYP/6-31++G** szinten, harmonikus oszcillátor közelítéssel). ^f A szerzők egy további, rövid felezési idejű konformert is detektáltak, azonban ehhez nem tudtak geometriát rendelni. ^g A szerzők nem tudták a két konformert egyértelműen megkülönböztetni, de arra a következtetésre jutottak, hogy mindkettő jelen volt a mérések során.

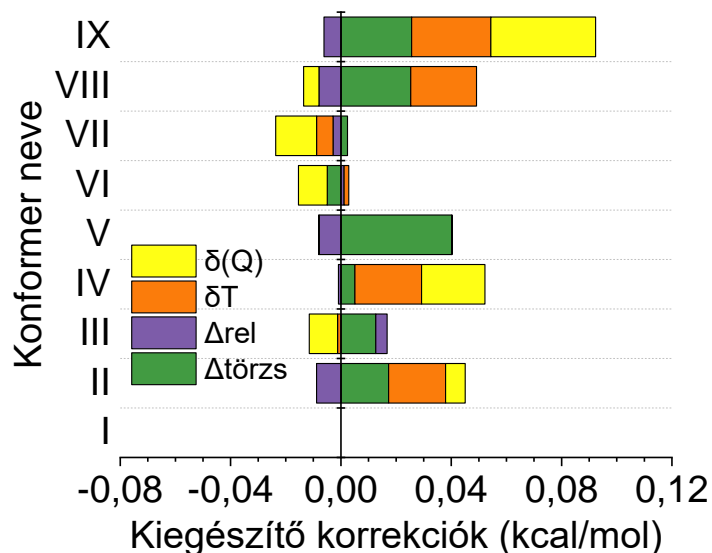
„párosítottuk” őket a kísérletileg kimutatottakkal. A konformerek relatív energia és/vagy szabadentalpia sorrendje nagyrészt megegyezik az irodalomban találhatóakkal. További kiemelendő tény, hogy a kísérletileg detektált tíz szerin konformert (amelyek közül néhányat több csoport is kimutatott) mind megtaláltuk a szerin potenciálisenergia-felületletén és mellettük még számos, akár alacsony relatív energiájú szerkezetet.

A semleges szerinre vonatkozó mennyiségeink a 7. táblázatban találhatóak. A ciszteinnél tapasztalt jelenségek miatt itt is megvizsgáltuk a CC módszerrel az MP2 geometriákra kapott relatív energiák értékét is. A CC módszerrel az MP2 és CC geometriákra meghatározott relatív energiák között az átlagos különbség 0,01 kcal/mol, tehát elmondható, hogy a semleges szerin esetében már az alacsonyabb elméleti szint is pontos geometriát jósol. A minimumok sorrendjében több változás is megfigyelhető a csatolt-klaszter módszer és az egyre nagyobb bázisok alkalmazásával. A CCSD(T)-F12a single-point számítások cc-pVTZ-F12 és cc-pVQZ-F12 bázisokkal kiváló báziskészlet konvergenciát mutatnak, a különbségek tized kcal/mol körüliek. A cc-pVQZ-F12 bázis használata mellett megvizsgáltuk a kétféle explicit-korrelált módszerrel (F12a és F12b) kapható relatív energiákat, azonban számottevő különbség ennél az aminosavnál sem látható. A kiszámolt kiegészítő korrekciókat pedig grafikon formájában az 27. ábrán foglaltuk össze.

7. táblázat: A továbbvizsgált semleges szerin konformerek relatív egyensúlyi energiái, a kiegészítő korrekciói, 0 és 298,15 K-es relatív entalpiái és 298,15 K-es relatív szabadentalpiái. A mértékegység minden esetben kcal/mol. A táblázat a [3]-as hivatkozás 1. táblázatának átdolgozása.

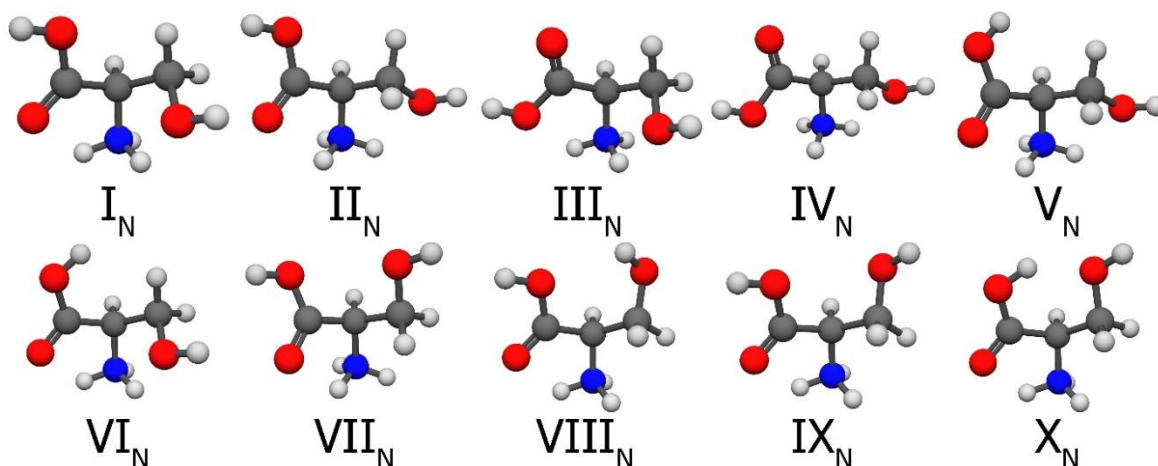
Konformer neve ^a	MP2	CC//MP2	CCSD(T)-F12a			CCSD(T)-F12b	Σ_{korrr}^h	ΔE_e^i	Δ_{ZPE}^j	ΔH_0^k	$\Delta H_{298.15}^l$	$\Delta G_{298.15}^m$
	DZ ^b	DZ ^c	DZ ^d	TZ ^e	QZ ^f	QZ ^g						
I/1	0,00	0,00	0,00	0,00	0,00	0,00	+0,00	0,00	+0,00	0,00	0,00	0,00
II/2	0,42	0,52	0,50	0,50	0,49	0,48	+0,04	0,52	-0,27	0,25	0,44	-0,09
III/4	0,92	0,62	0,63	0,63	0,64	0,64	+0,01	0,64	+0,10	0,74	0,68	0,84
IV/3	0,88	0,88	0,87	0,86	0,85	0,85	+0,05	0,90	-0,34	0,56	0,69	0,44
V/6	1,52	1,48	1,47	1,47	1,47	1,47	+0,03	1,50	-0,16	1,35	1,48	1,17
VI/7	1,55	1,82	1,82	1,82	1,82	1,82	-0,01	1,81	-0,16	1,65	1,78	1,58
VII/8	1,65	1,88	1,87	1,87	1,87	1,87	-0,02	1,85	-0,09	1,76	1,84	1,74
VIII/9	1,94	2,20	2,20	2,20	2,19	2,19	+0,04	2,22	-0,27	1,95	2,15	1,48
IX/10	2,24	2,47	2,45	2,46	2,45	2,45	+0,09	2,53	-0,61	1,92	2,28	1,50

^a A(z) római számok (arab számok) a CCSD(T)-F12b/cc-pVQZ-F12 (MP2/aug-cc-pVDZ) relatív energiák sorrendjét követik. ^b MP2/aug-cc-pVDZ relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^c CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^d CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^e CCSD(T)-F12a/cc-pVTZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^f CCSD(T)-F12a/cc-pVQZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^g CCSD(T)-F12b/cc-pVQZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^h A kiegészítő korrekciók összege: $\delta T + \delta(Q) + \Delta_{\text{törzs}} + \Delta_{\text{rel}}$, definíciók: 7.2.2.1–7.2.2.4. egyenletek. ⁱ Egyensúlyi relatív energiák, definíció: 7.2.2.5. egyenlet. ^j Zérusponyi energia-korrekció MP2/aug-cc-pVDZ szinten. ^k Adiabatikus relatív energiák, definíció: 7.2.2.6. egyenlet. ^l 298,15 K-es relatív entalpiák. ^m 298,15 K-es relatív szabadentalpiák.



27. ábra: A továbbvizsgált semleges szerin konformerekre kapott kiegészítő korrekciók. Definíciók: 7.2.2.1–7.2.2.4. egyenletek. Az ábra a [3]-as hivatkozás 5. ábrájának átdolgozása.

Az $1s^2$ elektronok korreláltatásából eredő $\Delta_{\text{örzs}}$ tag átlagosan 0,02 kcal/mol nagyságú, pozitív előjellel (kivéve VI-os konformer), de az V-ös konformer esetén a 0,04 kcal/mol-t is eléri. A másodrendű Douglas–Kroll relativisztikus korrekciók az eredményeinket nem befolyásolják, még kerekítéssel is éppen csak elérik a század kcal/mol-t néhány esetben. A δT és $\delta(Q)$ korrekciók vegyes nagyságúak (–0,02 és 0,04 kcal/mol között) és előjelűek, a teljes post-CCSD(T) járulékok fontosnak bizonyulnak, hiszen értékük 0,05 kcal/mol a IV-es és 0,07 kcal/mol a IX-es konformer esetében. A kiegészítő korrekciók összege átlagosan 0,04 kcal/mol, a legnagyobb értéket a IX-es minimumnál éri el (0,09 kcal/mol). Ezeket hozzáadva a CCSD(T)-F12b/cc-pVQZ-F12 single-point relatív energiákhoz megkaphatjuk a relatív egyensúlyi energiákat, a konformerek sorrendjében ekkor nem történik változás. Az egyensúlyi energiák pontosságát hasonlóképpen becsültük, ahogyan eddig: figyelembe vettük a TZ és QZ bázisokkal kapott eredmények közötti különbséget (0,01 kcal/mol) és az összegzett kiegészítő korrekciók átlagát (0,04 kcal/mol), így eljutva a $\sqrt{0,01^2 + 0,04^2} = 0,04$ kcal/mol-os bizonytalansághoz. A semleges szerin esetében is kiszámoltuk a harmonikus rezgési frekvenciákat MP2/aug-cc-pVDZ elméleti szint mellett CCSD(T)-F12a/cc-pVDZ-F12 szinten is. Itt is azt tapasztaltuk, hogy a ZPE korrekciók változása elhanyagolható (a különbség átlagosan 0,06 kcal/mol, amely még a ciszteinnél látottaknál is kisebb) a számítási idők drasztikus megnövekedéséhez képest. Ebből kifolyólag a további termodinamikai paraméterek meghatározásához az MP2 módszerrel kapott harmonikus rezgési frekvenciákat használjuk fel (a két érték közötti eltérés pedig egy hibabecslési alap lesz), valamint a protonált formákra nem végeztünk CC rezgési analízist. A Δ_{ZPE} tagok hozzáadásával megkapjuk a 0 K-re vonatkozó relatív entalpia értékeket, ekkor két helyen is változás történik a konformerek sorrendjében (III–IV és VIII–XI). Az egyensúlyi energiáknál említett hibaforrások mellett ekkor megjelenik az anharmonicitás is, illetve az MP2-CC frekvenciák közötti átlagos eltérés, ezek figyelembe vételével javasoljuk a $\sim 0,15$ kcal/mol bizonytalanságot. 298,15 K-re átszámítva a relatív entalpiákat a konformerek eredeti sorrendje visszaáll, azonban ha a relatív szabadentalpiákat is meghatározzuk, számtalan csere történik a szerkezetek között, még a globális minimum is a II-es lesz. A 298,15 K-re való átszámításnál különösen nagy szerepe van az alacsony rezgési frekvenciáknak, amelyek esetünkben a harmonikus közelítés miatt könnyen pontatlanok lehetnek, ezért az utolsó két termodinamikai paraméternek biztosan még nagyobb a hibája, mint a 0 K-re vonatkozó relatív entalpiáknak.



28. ábra: Az N-protonált szerin tíz konformerének CCSD(T)-F12a/cc-pVDZ-F12 geometriája. Kiválasztásuk az MP2/aug-cc-pVDZ relatív energiák alapján történt. A római számok a CCSD(T)-F12b/cc-pVQZ-F12 single-point energiák sorrendjét követik, míg az „N” az alsóindexben a protonálódás helyét jelöli. Az ábra a [3]-as hivatkozás 6. ábrája.

A tíz N-protonált minimum konformert a 28. ábrán mutatjuk be. A római számok a CCSD(T)-F12b/cc-pVQZ-F12 relatív energiák sorrendjét követik, az alsó index pedig a protonálódás helyére utal. Nyilvánvalóan, az O–H···N hidrogénkötések kialakulása nem lehetséges a protonálódás miatt, ezért csak gyengébb, N–H···O és O–H···O–H kölcsönhatások alakulnak ki. Kísérletileg kettő amino-protonált konformert tudtak azonosítani, az I_N/N1 (mint SH01) és a VIII_N/N8 (mint SH02) szerkezeteket, IRMPD spektroszkópiával.⁴⁸

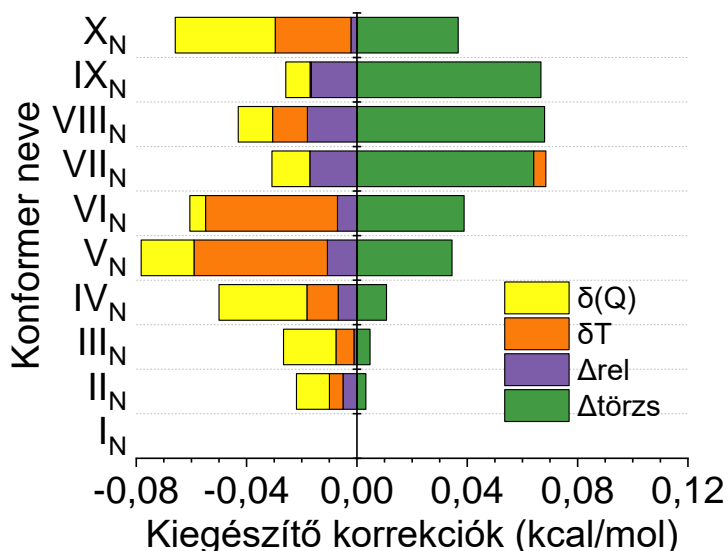
Az amino-protonált szerkezetekre kiszámolt relatív energiák, korrekciók, illetve relatív entalpiák/szabadentalpiák a 8. táblázatban láthatóak. Az MP2 geometriákon végzett CC számolások eredményei néhány tized kcal/mol-on belül megegyeznek a CC geometriai optimalás után meghatározottakkal, ami azt jelenti, hogy nincs jelentős geometriai effektus ebben az esetben. Ha azonban a pontosabb, CC módszert használjuk, két konformer, az N6-os és N7-es helyet cserél a sorrendben. A háromszoros és négyszeres ζ báziskészletek ismét jó konvergenciát mutatnak, a legnagyobb eltérés a TZ és QZ értékek között az X_N szerkezet esetén van, értéke körülbelül három század kcal/mol. Az F12a és F12b módszerek között az amino-protonált konformereknél sincs jelentős különbség.

Ahogy az a 29. ábrán látható, a törzselektron korrekció elhanyagolható az első három konformer esetén, azonban a VII_N, VIII_N és a IX_N szerkezetekre a 0,06–0,07 kcal/mol-t is eléri, átlagértéke 0,04 kcal/mol. A relativisztikus korrekciók ennél csekélyebbek, abszolút értékben a 0,02 kcal/mol-os maximális értéket veszik fel (VII_N, VIII_N és IX_N esetén), de átlá-

8. táblázat: A továbbvizsgált N-protonált szerin konformerek relatív egyensúlyi energiái, a kiegészítő korrekciói, 0 és 298,15 K-es relatív entalpiái és 298,15 K-es relatív szabadentalpiái. A mértékegység minden esetben kcal/mol. A táblázat a [3]-as hivatkozás 3. táblázatának átdolgozása.

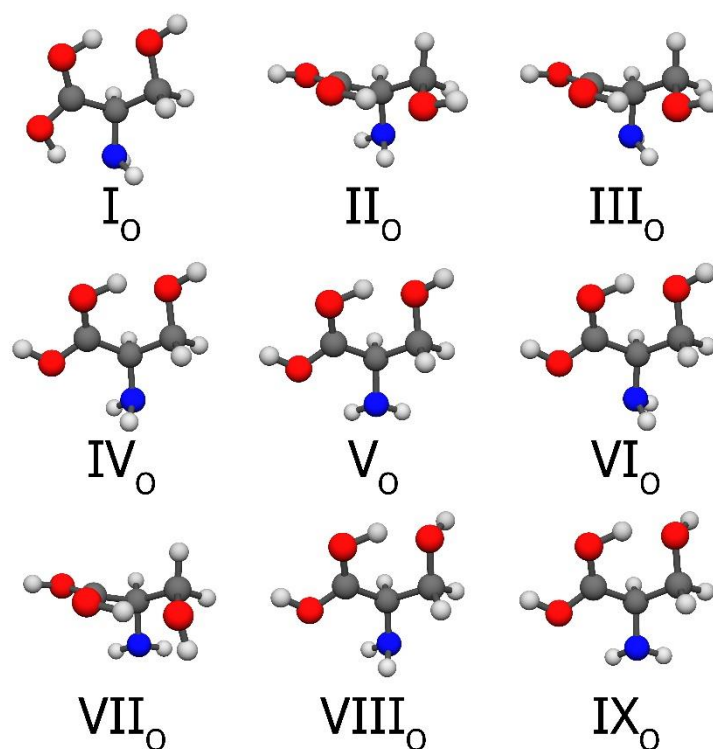
Konformer neve ^a	MP2	CC//MP2	CCSD(T)-F12a			CCSD(T)-F12b	Σ_{korrt}^h	ΔE_e^i	Δ_{ZPE}^j	ΔH_0^k	$\Delta H_{298,15}^l$	$\Delta G_{298,15}^m$
	DZ ^b	DZ ^c	DZ ^d	TZ ^e	QZ ^f	QZ ^g						
I _N /N1	0,00	0,00	0,00	0,00	0,00	0,00	+0,00	0,00	+0,00	0,00	0,00	0,00
II _N /N2	0,59	0,67	0,67	0,65	0,65	0,65	-0,02	0,63	-0,01	0,62	0,72	0,44
III _N /N3	3,35	3,69	3,67	3,69	3,68	3,69	-0,02	3,67	-0,11	3,55	3,65	3,30
IV _N /N4	4,14	4,50	4,49	4,50	4,49	4,50	-0,04	4,46	-0,15	4,31	4,50	3,57
V _N /N5	8,89	8,88	8,89	8,88	8,90	8,90	-0,04	8,85	-0,24	8,61	8,71	8,53
VI _N /N7	9,25	9,07	9,08	9,09	9,11	9,11	-0,02	9,09	-0,40	8,68	8,76	8,58
VII _N /N6	9,02	9,10	9,08	9,11	9,11	9,11	+0,04	9,15	-0,22	8,93	9,10	8,60
VIII _N /N8	9,47	9,10	9,11	9,13	9,13	9,12	+0,02	9,15	-0,27	8,88	9,08	8,41
IX _N /N9	9,84	9,50	9,51	9,53	9,53	9,53	+0,04	9,57	-0,42	9,15	9,43	8,32
X _N /N10	9,88	9,54	9,55	9,56	9,59	9,59	-0,03	9,56	-0,08	9,48	9,38	9,67

^a A(z) római számok (arab számok) a CCSD(T)-F12b/cc-pVQZ-F12 (MP2/aug-cc-pVDZ) relatív energiák sorrendjét követik, az "N" a protonálódás helyére utal. ^b MP2/aug-cc-pVDZ relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^c CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^d CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^e CCSD(T)-F12a/cc-pVTZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^f CCSD(T)-F12a/cc-pVQZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^g CCSD(T)-F12b/cc-pVQZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^h A kiegészítő korrekciók összege: $\delta T + \delta(Q) + \Delta_{\text{törzs}} + \Delta_{\text{rel}}$, definíciók: 7.2.2.1–7.2.2.4. egyenletek. ⁱ Egyensúlyi relatív energiák, definíció: 7.2.2.5. egyenlet. ^j Zérusponyi energia-korrekció MP2/aug-cc-pVDZ szinten. ^k Adiabatus relatív energiák, definíció: 7.2.2.6. egyenlet. ^l 298,15 K-es relatív entalpiák. ^m 298,15 K-es relatív szabadentalpiák.



29. ábra: A továbbvizsgált N-protonált szerin konformerekre kapott kiegészítő korrekciók. Definíciók: 7.2.2.1–7.2.2.4. egyenletek. Az ábra a [3]-as hivatkozás 5. ábrájának átdolgozása.

gosan inkább a 0,01 kcal/mol-osak, előjelük mindig negatív. A háromszoros és négyszeres gerjesztések figyelembevételéből származó korrekciók értékei szintén negatív előjelűek, mindkét komponens fontos, nagyságuk átlagosan 0,03 kcal/mol, a legnagyobb érték 0,07 kcal/mol (V_N). A 8. táblázatban láthatóan a QZ energiákhoz az összes kiegészítő korrekciót hozzáadva megkapjuk az egyensúlyi relatív energiákat, amelyek bizonytalanságára 0,03 kcal/mol-t jóslunk, a minimum konformerek sorrendje nem változik. A ZPE tagok meghatározásával (amelyek minden esetben negatív előjelűek) eljutunk az adiabatikus relatív energiákhoz, amiknek a bizonytalanságát 0,15 kcal/mol-nak becsüljük. A termikus járulékokkal (-0,10 – 0,28 kcal/mol) a 298,15 K-es relatív entalpiákat kapjuk meg, majd ezekből az entrópiikus tagok (-1,11 – 0,29 kcal/mol) kivonásával pedig a relatív szabadentalpiákat.



30. ábra: Az O-protonált szerin kilenc konformerének CCSD(T)-F12a/cc-pVDZ-F12 geometriája. Kiválasztásuk az MP2/aug-cc-pVDZ relatív energiák alapján történt. A római számok a CCSD(T)-F12b/cc-pVQZ-F12 single-point energiák sorrendjét követik, míg az „O” az alsóindexben a protonálódás helyét jelöli. Az ábra a [3]-as hivatkozás 7. ábrája.

A tíz, MP2/aug-cc-pVDZ elméleti szinten legalacsonyabb relatív energiájú O-protonált konformerből kilencet kapunk CCSD(T)-F12a/cc-pVDZ-F12 szinten, ezek a 30. ábrán láthatóak, a római számok a CCSD(T)-F12a/cc-pVQZ-F12 single-point relatív energiák sorrendjét követik, míg alsó indexben a protonálódás helye van jelölve. Akárcsak a semleges

szerinnél, úgy itt is, kettő MP2 szerkezet, az O2 és az O9 ugyanahhoz a csatolt-klaszter szerkezethez (IIo) konvergál. Csupán a globális minimum esetén figyelhető meg a nagyon erős O–H···N hidrogénkötés kialakulása, ami magyarázza azt a tényt, hogy a második legmélyebb energiájú minimumnak a relatív energiája már ~6 kcal/mol-lal nagyobb. A másik, intramolekuláris stabilizáló kölcsönhatás, amely még megfigyelhető a szerkezetekben az az O–H···O–H kölcsönhatás, azonban ez jóval gyengébb, mint az imént említett hidrogénkötés. Kísérleti eredmények az O-protonált szerkezettel kapcsolatban nem találhatók az irodalomban.

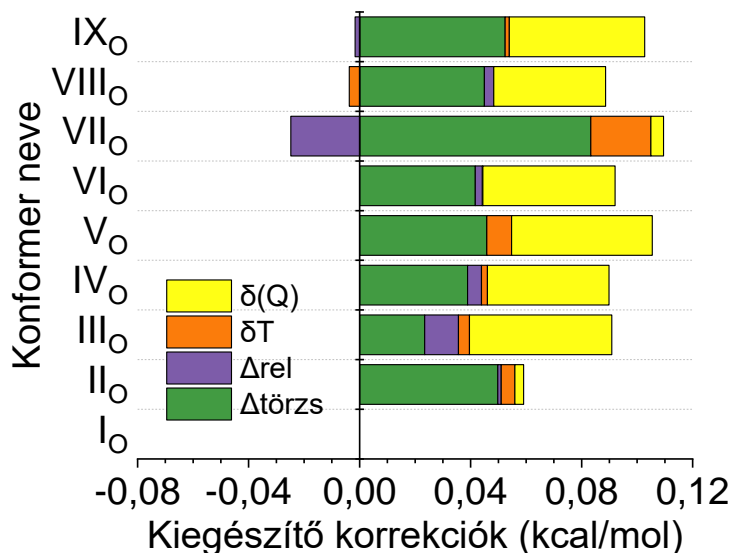
9. táblázat: A továbbvizsgált O-protonált szerin konformerek relatív egyensúlyi energiái, a kiegészítő korrekciói, 0 és 298,15 K-es relatív entalpiái és 298,15 K-es relatív szabadentalpiái. A mértékegység minden esetben kcal/mol. A táblázat a [3]-as hivatkozás 4. táblázatának átdolgozása.

Konformer neve ^a	MP2	CC//MP2	CCSD(T)-F12a			CCSD(T)-F12b	Σ_{kor}^h	ΔE_e^i	Δ_{ZPE}^j	ΔH_0^k	$\Delta H_{298,15}^l$	$\Delta G_{298,15}^m$
	DZ ^b	DZ ^c	DZ ^d	TZ ^e	QZ ^f	QZ ^g						
I _o /O1	0,00	0,00	0,00	0,00	0,00	0,00	+0,00	0,00	+0,00	0,00	0,00	0,00
II _o /O2	5,87	6,12	6,14	6,12	6,11	6,11	+0,06	6,17	-0,38	5,80	5,87	5,76
III _o /O3	6,27	6,20	6,23	6,22	6,20	6,20	+0,09	6,29	-0,37	5,93	6,12	5,77
IV _o /O4	6,35	6,32	6,34	6,30	6,29	6,29	+0,09	6,38	-0,67	5,71	5,83	5,54
V _o /O6	6,87	6,87	6,89	6,86	6,86	6,85	+0,11	6,96	-0,59	6,37	6,49	6,20
VI _o /O7	7,10	6,94	6,96	6,93	6,91	6,91	+0,09	7,00	-0,57	6,43	6,59	6,23
VII _o /O5	6,41	7,00	6,94	6,98	6,99	6,98	+0,08	7,07	-0,98	6,09	6,13	5,98
VIII _o /O8	7,31	7,24	7,21	7,18	7,17	7,17	+0,08	7,25	-0,78	6,47	6,63	6,26
IX _o /O10	7,98	7,94	7,93	7,90	7,89	7,89	+0,10	7,99	-0,72	7,27	7,43	7,07

^a A(z) római számok (arab számok) a CCSD(T)-F12b/cc-pVQZ-F12 (MP2/aug-cc-pVDZ) relatív energiák sorrendjét követik, az "O" a protonálódás helyére utal. ^b MP2/aug-cc-pVDZ relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^c CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák az MP2/aug-cc-pVDZ geometriákkal számítva. ^d CCSD(T)-F12a/cc-pVDZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^e CCSD(T)-F12a/cc-pVTZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^f CCSD(T)-F12a/cc-pVQZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^g CCSD(T)-F12b/cc-pVQZ-F12 relatív energiák a CCSD(T)-F12a/cc-pVDZ-F12 geometriákkal számítva. ^h A kiegészítő korrekciók összege: $\delta T + \delta(Q) + \Delta_{\text{törzs}} + \Delta_{\text{rel}}$, definíciók: 7.2.2.1–7.2.2.4. egyenletek. ⁱ Egyensúlyi relatív energiák, definíció: 7.2.2.5. egyenlet. ^j Zérusponthi energia-korrekció MP2/aug-cc-pVDZ szinten. ^k Adiabatus relatív energiák, definíció: 7.2.2.6. egyenlet. ^l 298,15 K-es relatív entalpiák. ^m 298,15 K-es relatív szabadentalpiák.

Az O-protonált szerinre kapott további eredményeinket a 9. táblázatban foglaltuk össze. Az MP2 és CC geometriák nagyon jó egyezést mutatnak, a CC optimalás során a relatív energiák csupán néhány század kcal/mol-lal térnek el egymástól. Az O5, O6 és O7 konformerek sorrendje változik az elméleti szint növelésével, míg végül TZ és QZ bázisokat alkalmazva már ugyanaz marad. A báziskészlet konvergenciát tekintve ennél az aminosavnál itt a legnagyobb a különbség a háromszoros és négyszeres ζ bázisok között, de még így is

0,02 kcal/mol alatt van. A QZ bázis alkalmazva gyakorlatilag azonos eredményeket kapunk az F12a és F12b módszerekkel, akárcsak eddig.



31. ábra: A továbbvizsgált O-protonált szerin konformerekre kapott kiegészítő korrekciók. Definíciók: 7.2.2.1–7.2.2.4. egyenletek. Az ábra a [3]-as hivatkozás 5. ábrájának átdolgozása.

A kiegészítő korrekciók, amelyek a 31. ábrán láthatóak, az O-protonált formákra a legjelentősebbek. A törzselektronok korreláltatásából származó korrekciók egészen nagyok, átlagosan 0,05 kcal/mol-osak, de a VII_o konformer esetén a 0,08 kcal/mol-t is meghaladja, ezzel szemben a másodrendű relativisztikus korrekciók század kcal/mol alattiak, kivéve a III_o, IV_o és VII_o szerkezeteknél (rendre 0,01; 0,01 és –0,02 kcal/mol). A post-CCSD(T) tagok az O-protonált forma esetén a legnagyobbak, és minden esetben (kivéve a VII_o) a perturbatív négyszeres gerjesztések járuléka lesz a nagyobb. Összegezve, a $\delta T + \delta(Q)$ korrekciók pozitív előjelűek, átlagosan 0,04 kcal/mol-osak, míg a maximális érték, amit felvehetnek, az 0,06 kcal/mol. A kiegészítő korrekciók együttvéve ezekre a konformerekre a legnagyobbak, átlagosan 0,09 kcal/mol, (a semleges formáknál ez az átlag 0,04 kcal/mol és az N-protonált formáknál pedig 0,03 kcal/mol). A relatív egyensúlyi energiák felírásával nem változik meg a konformerek sorrendje, ekkor a bizonytalanságot 0,09 kcal/mol körülnek jósoljuk. Ahogyan a 9. táblázatban is látszik, a zérusponthoz energiakorrekciók minden esetben negatív előjelűek, a szerin és protonált formáit tekintve ez is az O-protonált szerkezetekre éri el a legnagyobb értékeket (például a VII_o konformernél –0,98 kcal/mol). A 0 K-es relatív entalpiánál már azonban változás történik a sorrendben, a IV_o kerül a

második, a VIIo pedig a negyedik helyre, a hibahatárra $\pm 0,15$ kcal/mol-os becslést tehetünk. A 298,15 K-re való átváltáshoz egy 0,04 és 0,19 kcal/mol közötti tagot kell még figyelembe vennünk, míg a relatív szabadentalpiák kiszámításához egy $-0,11$ és $-0,37$ kcal/mol közötti entrópikus járulékot.

7.3. Protonaffinitások és gázfázisú bázicitások

A korábban meghatározott semleges és protonált aminosav formák és a statisztikus termodinamika segítségével ki tudjuk számolni a glicinhez, ciszteinhez és szerinhez tartozó abszolút protonaffinitásokat és gázfázisú bázicitásokat. He és munkatársai azt találták,¹²⁰ hogy az MP2 módszerrel meghatározott termodinamikai mennyiségek nagyszámú tesztmolekula esetén jelentős hibával terheltek a csatolt-klaszter eredményekhez képest. Ebből kiindulva célul tűztük ki, hogy minél pontosabban meghatározzuk a PA és GB értékeket. Ezt több megközelítésből érhetjük el: akárcsak a különálló konformereknél, úgy itt is figyelembe vesszük azt, hogy bármennyire is magas elméleti szinten számoltunk, még mindig alkalmazunk közelítéseket. Az így kapott értékek jó alapot szolgáltatnak arra is, hogy megbecsülhessük az eredményeink pontosságát. A következő szempont, hogy nem csupán a globális minimumok protonálódását vizsgáltuk az amino-protonált globális minimummá. A hagyományos Boltzmann-eloszlást feltételezve meghatároztuk azt is, hogy mekkora különbséget kapunk, ha a legmélyebb minimumok helyett konformerelegyeket vesszük figyelembe. Végezetül az összes (konformerkeresésnél stabilnak bizonyult) protonálódott funkcióscsoport kialakulásának energetikáját vizsgáltuk, amelynek ugyan gyakorlati jelentősége is van, az irodalomban még is elhanyagolt volt a termodinamikailag legkedvezőbb amino-protonálódás javára.

7.3.1. A glicin protonaffinitása és gázfázisú bázicitása

A glicinre meghatározott protonaffinitás és gázfázisú bázicitás értékeink valamint a hozzájuk tartozó analóg korrekciók a 10. táblázatban találhatóak. A gyakorlattel való összehasonlíthatóság érdekében a 0 K-es értékek mellett itt is meghatároztuk a 298,15 K-re vonatkozó mennyiségeket is. A két termodinamikai mennyiséget megadtuk különböző variációkra, tekintettel arra, hogy csak a globális minimum szerkezeteket vesszük figyelembe (I_p-I_N és I_p-I_O), vagy konformer elegyeket, illetve hogy a protonálódási folyamat melyik funkcióscsoporton játszódik le. Ez utóbbi azért fontos, mert az N-protonálódás termodinamikailag jóval kedvezőbb, így konformer elegyekre végzett számolásaink során a Boltzmann-eloszlás miatt az O-protonálódás elhanyagolódna, ha nem

10. táblázat: A glicinhez tartozó protonaffinitás, gázfázisú bázicitás értékek, valamint a hozzájuk tartozó korrekciók, a mértékegység minden esetben kcal/mol. A táblázat az [1]-es hivatkozás 3. táblázatának átdolgozása.

	ΔE_{QZ}^a	δT^b	$\delta(Q)^b$	$\Delta_{\text{törzs}}^b$	Δ_{rel}^b	ΔE_c^c	Δ_{ZPE}^d	ΔH_0^e	$\Delta H_{298,15}^f$	$\Delta G_{298,15}^g$
Ip – In	219,73	+0,01	-0,01	+0,11	-0,02	219,82	-9,14	210,68	212,14	204,90
Ip – Io	194,04	+0,06	-0,16	+0,04	-0,06	193,91	-7,69	186,22	187,49	180,22
Átlag N^h	220,04	+0,01	-0,01	+0,11	-0,02	220,13	-9,13	211,00	212,43	204,75
Átlag O^h	194,18	+0,06	-0,16	+0,04	-0,06	194,06	-7,68	186,38	187,64	180,21

^a CCSD(T)-F12b/aug-cc-pVQZ egyensúlyi protonaffinitások a CCSD(T)-F12b/aug-cc-pVTZ geometriákon számítva. ^b A kiegészítő korrekciók, definíciók: 7.2.1.1–7.2.1.4. egyenletek. ^c A kiegészítő korrekciók figyelembevételével számolt egyensúlyi protonaffinitások. ^d Zérusponyi energia-korrekció CCSD(T)-F12b/aug-cc-pVDZ szinten. ^e 0 K-es protonaffinitások ($\Delta E_c + \Delta_{\text{ZPE}}$). ^f 298.15 K-es protonaffinitások. ^g 298.15 K-es gázfázisú bázicitások. ^h A konformerek populációit a Boltzmann-eloszlás segítségével határoztuk meg.

különítenék el. A post-CCSD(T) korrekciók páronként ellentétes előjelűek, az aminoszén-protonálódás esetén a két tag kiejti egymást (értékeik 0,01 kcal/mol), míg O-protonálódás esetén összesítve egy körülbelül 0,11 kcal/mol-os járulékot adnak.

11. táblázat: A post-CCSD(T) korrelációs korrekciók (kcal/mol) báziskészlet konvergenciája a glicin PA/GB értékekre vonatkozóan. A korrekciókat a 7.2.1.1. és 7.2.1.2. egyenletek definiálják, az alkalmazott bázisok a 3-21G, 6-31G és a cc-pVDZ (VDZ) voltak. A táblázat az [1]-es hivatkozás 2. táblázatának átdolgozása.

	δT^a			$\delta(Q)^b$			$\delta T + \delta(Q)$		
	3-21G	6-31G	VDZ	3-21G	6-31G	VDZ	3-21G	6-31G	VDZ
Ip – Ip_N	-0,04	-0,02	+0,01	+0,05	+0,00	-0,01	+0,01	-0,02	+0,00
Ip – Ip_O	+0,04	+0,06	+0,06	-0,07	-0,15	-0,16	-0,03	-0,09	-0,11
Átlag N^a	-0,04	-0,02	+0,01	+0,05	+0,01	-0,01	+0,01	-0,02	+0,00
Átlag O^a	+0,04	+0,06	+0,06	-0,07	-0,15	-0,16	-0,03	-0,09	-0,11

^a A konformerek populációit a Boltzmann-eloszlás segítségével határoztuk meg.

A 11. táblázatban a PA és GB értékekre vonatkozóan is megvizsgáltuk a δT és $\delta(Q)$ korrekciók báziskészlet függését. N-protonálódás esetén a 3-21G bázissal kapott δT és $\delta(Q)$ korrekciók 0,05-0,06 kcal/mol-lal eltérnek a nagyobb, 6-31G vagy cc-pVDZ bázisokkal számolt értékektől, míg ha a teljes post-CCSD(T) korrekciót nézzük, a különbség csupán 0,03 kcal/mol körüli. A báziskészlet-effektus a karbonil-protonálódásra már jóval szembetűnőbb. Noha a δT korrekció jó konvergenciát mutat (0,04; 0,04 és 0,06 kcal/mol a 3-21G, 6-31G és cc-pVDZ bázisokkal), addig a $\delta(Q)$ korrekciónál a nagyobb bázisok használata válik szükségessé, a kapott értékek rendre -0,07; -0,15 és -0,16 kcal/mol. Az összesített értékek -0,03; -0,09 és -0,11 kcal/mol a három vizsgált bázissal, amiből az látszik, hogy a 3-21G bázis használata még nem kielégítő, azonban a 6-31G már jó választás lehet a cc-pVDZ bázissal szemben, hiszen nagyon jó közelítő eredményhez juthatunk vele jelentősen rövidebb számítási költség árán. A törzselektron korrekciók pozitív járulékkal

bírnak, míg a DK relativisztikus korrekcióból eredőek előjele negatív. Az aminocsoporton lejátszódó protonálódásnál a $\Delta_{\text{törzs}}$ okoz nagyobb eltérést, míg a karbonil-protonálódásnál a Δ_{rel} . Összegezve az eddigi korrekcióinkat, az aminocsoportoz tartozó PA-ban 0,09 kcal/mol, míg a karbonilcsoportoz tartozó PA-ban 0,12 kcal/mol változást tapasztaltunk.

Az egyensúlyi protonaffinitás értékek (ΔE_e) úgy kaphatóak, ha vesszük a semleges és protonált aminosav egyensúlyi energiájának különbségét. Ebből a 0 K-re vonatkozó protonaffinitás a ZPE korrekciók hozzáadásával kapható, ez utóbbi értéke mindkét protonálódási hely esetén kb. $-8/-9$ kcal/mol. A termodinamikai paramétereket a relatív entalpiákhoz és szabadentalpiákhoz hasonlóan meghatároztuk 298,15 K-en is. Véges hőmérsékleten a rezgési és forgási termikus járulékok mellett a proton translációs entalpiájával is számolnunk kell, amelynek értéke ezen a hőmérsékleten 1,48 kcal/mol. A 0 K-es és a 298,15 K-es protonaffinitás értékek között 1,5 vagy 1,3 kcal/mol a különbség (amino- és karbonil-protonálódás), ami azt mutatja, hogy a rezgési és forgási tagok jóval kisebbek, mint a protonhoz tartozó entalpia érték. Az entrópiákat is figyelembe véve eljutunk a gázfázisú bázicitásokhoz, amik átlagosan 7,4 kcal/mol-lal alacsonyabbak a megfelelő protonaffinitásoknál. Ebből az értékből ismét az látszik, hogy a szabad proton járuléka a legfontosabb, hiszen az entalpiájának és szabadentalpiájának a különbsége $1,48 - (-6,27) = 7,75$ kcal/mol. A meghatározott termodinamikai mennyiségek nagyban függenek attól, hogy melyik protonálódási helyet vesszük figyelembe, de attól is, hogy csak a globális minimum konformereket vesszük figyelembe, vagy konformer elegyeket. A két funkciós csoportoz tartozó értékek közti különbség a ΔE_e esetén a legnagyobb, kb. 26 kcal/mol, míg a többi paraméternél valamivel kisebb, $\sim 24,7$ kcal/mol. Ha konformer elegyekkel számolunk, akkor az egyensúlyi, a 0 K-es és a 298,15 K-se protonaffinitás értékek 0,3 kcal/mol-lal növekednek N-protonálódásra nézve, míg O-protonálódásra 0,15 kcal/mol-lal. GB esetén az elegy-effektus $-0,15$ kcal/mol és $-0,01$ kcal/mol az N- és O-protonálódási helyekre. Ezek a számok azt mutatják, hogy a magasabb energiájú konformerek jelenléte sem elhanyagolható. A végső, 298,15 K-es protonaffinitás értékek [globális minimumokat/konformer elegyeket] figyelembe véve [212,14/212,43] kcal/mol az aminocsoport esetében, míg a karbonil-protonálódásra [187,49/187,64] kcal/mol. Az azonos hőmérsékletű gázfázisú bázicitások pedig [204,90/204,75] kcal/mol (N-protonálódás) és [180,22/180,21] kcal/mol (O-protonálódás). Ugyan ~ 5 kcal/mol-lal a javasolt érték alatt vagyunk, de Noguera jó becslést tett arra, hogy mekkora az energetikai különbség a két protonálódási hely között.⁴⁵ Mivel a relatív entalpia és szabadentalpia értékek meghatározásánál azt tapasztaltuk, hogy azok

nagyban függenek attól, hogy az MP2/aug-cc-pVDZ vagy a CCSD(T)-F12b/aug-cc-pVDZ rezgési frekvenciákat használjuk fel, ezért a PA-okat és GB-okat is kiszámoltuk mindkét frekvencia felhasználásával. A PA értékek (mind 0, mind 298,15 K-en) 0,10 kcal/mol-on belül azonosak mindkét funkciós csoportra nézve, a GB pedig 0,50-0,55 kcal/mol-lal tér el amino-protonálódásnál, 0,20-0,25 kcal/mol-lal pedig a karbonil-protonálódásnál. Akárcsak a relatív entalpiáknál és szabadentalpiáknál, így itt is megpróbálunk javaslatot tenni a meghatározott termodinamikai értékeket hibájára. Az egyensúlyi protonaffinitások (ΔE_e) bizonytalanságát a korrekciók összegével közelítjük, ez az N-protonálódás esetén 0,10 kcal/mol, míg az O-protonálódásnál 0,04 kcal/mol. A 0 K-es PA értékeknél már egy újabb hibaforrást is figyelembe kell vennünk: He és munkatársai azt találták,¹²⁰ hogy az elhanyagolt anharmonicitás akár 0,15 kcal/mol-os hibát is okozhat különböző molekulák GB értékeiben, így mi ugyanezt feltételezzük a PA (és később GB) értékekre is, így kapjuk rendre a $\sqrt{0,10^2 + 0,15^2} \approx 0,20$ kcal/mol és 0,15 kcal/mol-os hibahatárokat az amino- és karbonil-protonációra. A 298,15 K-re való átváltásnál az MP2 és CC frekvenciákkal kapott értékek különbségéből következtetünk az alacsony rezgési frekvenciák bizonytalanságára, a legrosszabb eshetőségeket figyelembe véve, az eredményeket inkább felfelé kerekítve. Ekkor a 298,15 K-es protonaffinitásokra $\sqrt{0,20^2 + 0,10^2} \approx 0,25$ kcal/mol és $\sim 0,20$ kcal/mol, míg a gázfázisú bázicitásokra $\sqrt{0,55^2 + 0,25^2} \approx 0,60$ kcal/mol és $\sim 0,35$ kcal/mol értékeket javasolunk az amino- és a karbonilcsoportok esetén.

12. táblázat: A glicinhez tartozó protonaffinitás és gázfázisú bázicitás értékek összehasonlítása az irodalmi értékekkel. A mértékegység minden esetben kcal/mol.

	Protonaffinitás		Gázfázisú bázicitás	
	Érték	Bizonytalanság	Érték	Bizonytalanság
Harrison ⁷⁸	210,5	2-3	202,7	1-2
Hunter&Lias ³²	211,7	<1,9	203,5	<1,9
Bouchoux ⁹⁰	211,9	-	204,0	-
Jelen munka	212,4	0,3	204,8	0,6

A glicinre javasolt PA és GB értékeinket a 12. táblázatban hasonlítottuk össze az irodalmi résznel kiemelt három összefoglaló publikációval. Jól látható, hogy hibahatáron belüli egyezés tapasztalható, az eredményeink valamivel magasabbak, viszont az általunk alkalmazott magas elméleti szint miatt jövőbeli munkáknál a javasolt érték az „irányunkba” tolódhat el.

7.3.2. A cisztein protonaffinitása és gázfázisú bázicitása

A ciszteinre kiszámolt protonaffinitások és gázfázisú bázicitások (0 és 298,15 K-re vonatkozóan is, illetve a hozzájuk tartozó korrekciók értékei) a 13. táblázatban láthatóak.

13. táblázat: A ciszteinhez tartozó protonaffinitás, gázfázisú bázicitás értékek, valamint a hozzájuk tartozó korrekciók, a mértékegység minden esetben kcal/mol. A táblázat a [2]-es hivatkozás 6. táblázatának átdolgozása.

	ΔE_{QZ}^a	δT^b	$\delta(Q)^b$	$\Delta_{\text{törzs}}^b$	Δ_{rel}^b	ΔE_c^c	Δ_{ZPE}^d	ΔH_0^e	$\Delta H_{298,15}^f$	$\Delta G_{298,15}^g$
I-I _N	222,89	-0,03	-0,01	+0,10	+0,03	222,98	-8,49	214,49	215,79	208,44
I-I _O	208,57	+0,02	-0,04	+0,05	+0,02	208,62	-7,45	201,17	202,74	194,56
I-I _S	198,22	-0,03	+0,01	-0,06	+0,05	198,18	-5,46	192,72	194,02	186,86
Átlag N ^k	223,60	-0,03	-0,01	+0,10	+0,04	223,69	-8,73	214,96	216,39	208,21
Átlag O ^k	209,47	+0,02	-0,04	+0,04	+0,02	209,51	-7,67	201,83	203,55	194,16
Átlag S ^k	199,13	-0,04	+0,02	-0,06	+0,06	199,10	-5,79	193,31	194,74	186,40

^a CCSD(T)-F12b/cc-pVQZ-F12 egyensúlyi protonaffinitások a CCSD(T)-F12a/cc-pVDZ-F12 geometriákon számítva. ^b A kiegészítő korrekciók, definíciók: 7.2.1.1–7.2.1.4. egyenletek. ^c A kiegészítő korrekciók figyelembevételével számolt egyensúlyi protonaffinitások. ^d Zérusponyi energia-korrekció MP2/aug-cc-pVDZ szinten. ^e 0 K-es protonaffinitások ($\Delta E_c + \Delta_{\text{ZPE}}$). ^f 298.15 K-es protonaffinitások. ^g 298.15 K-es gázfázisú bázicitások. ^h A konformerek populációit a Boltzmann-eloszlás segítségével határoztuk meg.

Ismét szétválasztottuk a különböző funkciócsoportok protonálódását, hiszen a glicinhez hasonlóan itt is csak az N-protonálódás dominálna. Megvizsgáltuk azt is, hogy mekkora hatása van annak, ha csak a globális minimum szerkezeteket vesszük figyelembe, vagy konformer elegyekkel számolunk (Boltzmann-eloszlást feltételezve). A δT korrekciók negatív előjelűek, kivéve az O-protonálódás esetében és 0,02–0,04 kcal/mol közti értékeket vesznek fel. A $\delta(Q)$ tagok is negatívak, ez alól kivétel most a S-protonálódás, értékük 0,01 és 0,04 kcal/mol közötti. Összegezve a post-CCSD(T) korrekciókat az mondható el, hogy az az amino-protonálódásnál a legjelentősebb, [–0,04/–0,05] kcal/mol a [globális minimumokat/konformer elegyeket] tekintve, míg a többi funkciócsoport esetén egyöntetűen –0,02 kcal/mol. A törzselektronok korreláltatása az előzőeknél jobban befolyásolja a végső eredményeinket, hiszen a hozzá tartozó korrekció értéke [0,10/0,10] kcal/mol az amino-, [0,05/0,04] kcal/mol a karbonil- és [–0,06/–0,06] kcal/mol a tiol-protonálódásra. Az utolsó, általunk vizsgált kiegészítő járulék a relativisztikus korrekcióból származó pozitív mennyiség a 0,02–0,06 kcal/mol-os tartományban, az S-protonálódás esetén a legrelevánsabb ([0,05/0,06 kcal/mol]). A négy eddig ismertett korrekció összege az amino-protonálódás esetén a legnagyobb, 0,09 kcal/mol (mindkét esetre nézve), ebből kifolyólag az általunk meghatározott egyensúlyi protonaffinitás értékek bizonytalanságára egy 0,10 kcal/mol-os becslést tettünk. Ha ehhez az értékhez hozzáadjuk a ZPE korrekciót, akkor megkapjuk a 0 K-re vonatkozó PA-t. Ez a tag [–8,5/–8,7], [–7,5/–7,7] és [–5,5/5,8]

kcal/mol az N-, O- és S-protonálódások esetében. A termikus korrekciókat kiszámolva és figyelembe véve a proton translációs entalpiáját (1,48 kcal/mol) meghatároztuk a 298,15 K-es protonaffinitásokat. A változás a hőmérséklet „megemelésével” ebben az esetben is közel van a protonhoz tartozó entalpiaértékkel, vagyis a termikus tagok ehhez képest kevésbé jelentősek (akárcsak a glicinnél). Végül az entrópikus tagok hozzáadásával megkapjuk a 298,15 K-re vonatkozó GB-okat, amelyek rendre [7,35/8,18], [8,17/9,39] és [7,16/8,34] kcal/mol-lal alacsonyabbak a megfelelő amino-, karbonil- és tiol-protonált protonaffinitás értékeknél. Ismételten a proton járuléka (7,75 kcal/mol) lesz a legjelentősebb, illetve a konformer elegyek esetén a keveredési entrópia tag is számottevő. A PA és GB mennyiségek között 10–20 kcal/mol-os különbségek vannak attól függően, hogy melyik funkciócsoporton történik a protonálódás, termodinamikailag az amino-protonálódás a legkedvezőbb, ezt követi a karbonil-, majd a tiol-protonálódás. A különböző funkciócsoportok protonálódásáról említést tartalmazó publikáció több, mint kétszeresen túlbecsüli az amino- és a karbonilcsoport protonálódása közötti energiabeli különbséget a cisztein esetén.⁴⁵ Ahhoz, hogy egy javaslatot tehesünk a meghatározott termodinamika mennyiségek pontatlanságára, a következő dolgokat kell figyelembe vennünk: a kiegészítő korrekciók összegének nagysága, az elhanyagolt anharmonicitás, az alacsony harmonikus rezgési frekvenciák bizonytalansága illetve az, hogy MP2 frekvenciákkal számolunk, nem pedig CC frekvenciákkal. Az egyensúlyi protonaffinitások bizonytalanságára 0,09/0,05/0,03 kcal/mol-t javaslunk az N-, O- és S-protonálódásra. A 0 K-es protonaffinitásnál a glicinnél említett elhanyagolt anharmonicitásból eredő 0,15 kcal/mol-os tag mellett még az MP2-CC frekvenciák különbségét is figyelembe vesszük (0,08 kcal/mol-nak adódott a semleges ciszteinre), így $\sqrt{0,09^2 + 0,15^2 + 0,08^2} \approx 0,20$ kcal/mol-os bizonytalanság becslést tehetünk mindhárom protonálódás esetére, felfelé kerekítve. A PA és GB értékek esetén a tiol-protonálódásra nem áll rendelkezésre arról információ, hogy mekkora a különbség abból, ha MP2 vagy CC frekvenciákkal számoljuk ki (az alacsony rezgési frekvenciák bizonytalanságát is ezzel közelítjük), ezért a glicinnél kapott rosszabb, amino-protonálódási értéket vesszük figyelembe a kiszámításához. Végezetül a 298,15 K-es protonaffinitásokra $\pm 0,25$ kcal/mol-os (mindhárom esetben), míg a gázfázisú protonaffinitásokra $\pm 0,60/\pm 0,40/\pm 0,60$ kcal/mol-os hibahatárokat mondhatunk.

A ciszteinre kapott PA és GB értékeinket a 14. táblázatban vetettük össze az irodalomban legpontosabbnak tekinthető mennyiségekkel. Akár a glicinnél, úgy most is a mérések/kiértékelések bizonytalanságán belüli értékeket kaptunk, kitűnő egyezés figyelhető

meg. Már a glicinnél is látszott, de itt talán jobban, hogy Harrison 1997-es cikke⁷⁸ után szükség volt a javasolt értékek módosítására, hiszen az látszólag 1-2 kcal/mol-lal alábecsüli a feltehetőleg pontos értékeket/értékeinket.

14. táblázat: A ciszteinhez tartozó protonaffinitás és gázfázisú bázicitás értékek összehasonlítása az irodalmi értékekkel. A mértékegység minden esetben kcal/mol.

	Protonaffinitás		Gázfázisú bázicitás	
	Érték	Bizonytalanság	Érték	Bizonytalanság
Harrison ⁷⁸	214,0	2-3	206,2	1-2
Hunter&Lias ³²	215,7	<1,9	207,6	<1,9
Bouchoux ⁹⁰	215,7	-	207,8	-
Jelen munka	216,4	0,4	208,2	0,6

7.3.3. A szerin protonaffinitása és gázfázisú bázicitása

A szerinhez kapcsolódó termodinamikai mennyiségek meghatározásához is a kémiai pontosságot elérő módszerként hivatkozott csatolt-klaszter módszert alkalmaztuk, valamint analóg módon a konformerek esetén is vizsgált korrekciókat.

15. táblázat: A szerinhez tartozó protonaffinitás, gázfázisú bázicitás értékek, valamint a hozzájuk tartozó korrekciók, a mértékegység minden esetben kcal/mol. A táblázat a [3]-as hivatkozás 5. táblázatának átdolgozása.

	ΔE_{QZ}^a	δT^b	$\delta(Q)^b$	$\Delta_{\text{törzs}}^b$	Δ_{rel}^b	ΔE_e^c	Δ_{ZPE}^d	ΔH_0^e	$\Delta H_{298,15}^f$	$\Delta G_{298,15}^g$
I-I _N ^h	225,05	-0,04	-0,04	+0,12	-0,02	225,07	-8,57	216,50	217,76	209,72
I-I _O ^h	211,63	+0,01	-0,08	+0,09	-0,05	211,60	-7,83	203,77	205,33	196,46
Átlag N ⁱ	225,43	-0,03	-0,04	+0,13	-0,03	225,46	-8,74	216,72	218,05	209,86
Átlag O ⁱ	212,22	+0,02	-0,07	+0,10	-0,05	212,21	-8,00	204,20	205,87	196,36

^a CCSD(T)-F12b/cc-pVQZ-F12 egyensúlyi protonaffinitások a CCSD(T)-F12a/cc-pVDZ-F12 geometriákon számítva. ^b A kiegészítő korrekciók, definíciók: 7.2.1.1–7.2.1.4. egyenletek. ^c A kiegészítő korrekciók figyelembevételével számolt egyensúlyi protonaffinitások. ^d Zérusponti energia-korrekció MP2/aug-cc-pVDZ szinten. ^e 0 K-es protonaffinitások ($\Delta E_e + \Delta_{ZPE}$). ^f 298,15 K-es protonaffinitások. ^g 298,15 K-es gázfázisú bázicitások. ^h A relatív szabadentalpiák esetén a II-es konformer lesz az új globális minimum, ezt figyelembe vettük a gázfázisú bázicitások meghatározásakor. ⁱ A konformerek populációit a Boltzmann-eloszlás segítségével határoztuk meg.

A 15. táblázatban láthatóak a szerin aminosavra kiszámolt mennyiségeink, a korrekciók értékei, a 0 K és 298,15 K-es protonaffinitások valamint a 298,15 K-es gázfázisú bázicitás. Ezeket megadtuk a különböző lehetséges protonálódási helyekre (amino- és karbonilcsoport, amelyek közül az előbbi a domináns) vonatkozóan is, illetve úgy is, hogy csak a globális minimum szerkezeteket vesszük figyelembe, vagy az összes, jelentős mértékben betöltött állapotot. Visszautalva a 7. táblázatra, amelyben a semleges szerin termodinamikai paramétereit is ismertettük, újból kiemelendő, hogy a relatív szabadentalpiákat tekintve a semleges szerinnél már a II-es konformer lesz a globális minimum, így a csupán a

legmélyebb energiájú konformereket figyelembevevő számolásoknál erre mindenképpen ügyelni kell. A δT korrekcióknak van az egyik legkisebb jelentősége, hiszen az amino-protonálódásra $[-0,04/-0,03]$ kcal/mol, attól függően, hogy [globális minimumokat/konformerelegyeket] veszünk figyelembe, a karbonil-protonálódásra még csekélyebb, $[0,01/0,02]$ kcal/mol-os értékeket kaptunk. A $\delta(Q)$ tagok valamivel nagyobbak, $-0,04$ kcal/mol az N-protonálódásra, míg az O-protonálódásnál $[-0,08/-0,07]$ kcal/mol-osak. A törzselektronok korreláltatása okozza a legnagyobb változást a kiegészítő korrekciók közül, hiszen ez az effektus mindkét protonálódási helynél $0,1$ kcal/mol fölött van (kivéve, ha a karbonil-protonálódást nézzük csak a legmélyebb energiájú semleges és protonált formát figyelembe véve, akkor $0,09$ kcal/mol). Az utolsó ilyen tagunk a Δ_{rel} , N-protonálódást tekintve ez $[-0,02/-0,03]$ kcal/mol, a másik funkcióscsoport esetén pedig $[-0,05/-0,05]$ kcal/mol. Az egyensúlyi protonaffinitás értékekhez hozzáadva a $8-9$ kcal/mol körüli zérus ponti energia-korrekciókat megkapjuk a 0 K-es protonaffinitás értékeket. Ez utóbbiaknál az, hogy konformer elegyekkel számolunk egy $-0,22$ kcal/mol-os, vagy egy $-0,44$ kcal/mol-os különbséget jelent az amino- és karbonilcsoportnál. A végső, 0 K-es protonaffinitás értékeink $216,72$ kcal/mol az N- és $204,20$ kcal/mol az O-protonálódásra. $298,15$ K hőmérsékletre átszámolva ezek az értékek $218,05$ kcal/mol-ra (N) és $205,87$ kcal/mol-ra (O) változnak, a 0 K-es értéktől való különbségeik ismét jól közelítik a proton transzlációs entalpiáját ($1,48$ kcal/mol), a termikus járulékok kb. egy nagyságrenddel kisebbek ehhez képest. Az entrópikus tagok kiszámolása után megkapjuk a $298,15$ K-re vonatkozó gázfázisú bázicitásokat, amelyek $209,86$ kcal/mol (N) és $193,36$ kcal/mol (O). Akárcsak a ciszteinnél, úgy a szerinnél is termodinamikailag sokkal kedvezőtlenebbnek tartják a karbonilcsoport protonálódását, mint az az eredményeink alapján elmondható.⁴⁵

A kapott termodinamikai mennyiségek hibájára ismét próbáltunk becsléseket tenni. Az egyensúlyi protonaffinitás értékénél a hibahatár meghatározásához a korrekciók átlagos értékét ($0,02$ kcal/mol mindkét funkcióscsoport esetén) vettük alapul. A 0 K-es értékeknél már több „hibaforrás” is megjelenik a zérus ponti energia-korrekciókból származóan. Az MP2 és CC harmonikus rezgési frekvenciák közötti különbséget a semleges szerinnél $0,06$ kcal/mol-nak találtuk, így ezt vettük figyelembe a $0,15$ kcal/mol mellé (anharmonicitás), így kapjuk a körülbelül $0,2$ kcal/mol-t (szintén mindkét esetre). A $298,15$ K-es PA értékre $0,25$ kcal/mol-os lehet a pontatlanság. Az utolsó bizonytalanság, amit meg kell fontolnunk, az a gázfázisú bázicitásokhoz tartozik. N-protonálódásra $0,60$ kcal/mol, O-protonálódásra pedig $0,40$ kcal/mol-os közelítő értékeket adhatunk meg. A 16. táblázatban „összemértük” a

termodinamikai mennyiségeinket a referencia értékekkel. Akárcsak a másik két aminosav esetén, az egyezés itt is kimondottan jó, bizonytalanságon belüli. A Harrison⁷⁸ által javasolt értékek még jobban (mint a glicinnél és ciszteinnél) eltérnek akár a később javasoltaktól, akár az általunk meghatározottaktól.

16. táblázat: A ciszteinhez tartozó protonaffinitás és gázfázisú bázicitás értékek összehasonlítása az irodalmi értékekkel. A mértékegység minden esetben kcal/mol. A táblázat a [3]-as hivatkozás 6. táblázatának átdolgozása.

	Protonaffinitás		Gázfázisú bázicitás	
	Érték	Bizonytalanság	Érték	Bizonytalanság
Harrison ⁷⁸	215,2	2-3	207,6	1-2
Hunter&Lias ³²	218,4	<1,9	210,4	<1,9
Bouchoux ⁹⁰	217,8	-	209,7	-
Jelen munka	218,1	0,4	209,9	0,6

8. ÖSSZEFOGLALÁS

A gázfázisú aminosavak (és a hozzájuk tartozó protonált formák) minimum konformereinek vizsgálata hosszú múltra tekint vissza, mind kísérleti, mind elméleti oldalról. Mikrohullámú és infravörös spektroszkópiás mérésekkel már számos aminosav konformert tudtak azonosítani forgási vagy rezgési attribútumaik alapján, azonban a meghatározásaikat számos dolog nehezíti. Az aminosavak könnyen bomló vegyületek, kiváltképp, ha sikerül őket gázfázisba juttatnunk, valamint egy-egy konformer szelektív előállításuk nehézkes, ugyanis az egyes szerkezetek könnyen átalakulhatnak egymásba (csekély az energiagát), ami egy összetett spektrumhoz vezet. Számos kutató (köztük mi is) az elméleti oldalról közelítette meg a problémát, és kvantumkémiai számítások segítségével próbálta feltérképezni az aminosavak konformációs terét. Ez történhet szisztematikus módon, vagy algoritmusok/mesterséges intelligencia alapú eljárások felhasználásával is, azonban ahhoz, hogy egy átfogó tanulmányt készíthessünk, a számítástechnika fejlődésére van/volt szükség. A tudomány mai állása hasonló a vegyületek protonálódását jellemző protonaffinitások és gázfázisú bázicitások esetén is. A kísérleti módszerek csak nagyon kevés esetben teszik lehetővé ezen termodinamikai mennyiségek abszolút meghatározását, általában valamilyen referenciaértékhez tudunk csak viszonyítani, ami azt jelenti, hogyha pontatlan a referenciánk, a kapott értékünk is az lesz. A problémán ismét az elméleti kémia ezköztára segíthet, azonban az irodalomban fellelhető források jelentős elhanyagolásokat és alacsony elméleti szinteket tartalmaznak.

Doktori munkám célja a glicin, cisztein és szerin aminosavak minden eddiginél széleskörűbb szisztematikus, *ab initio* szerkezeti vizsgálata, beleértve a protonált formáikat is, valamint a protonálódást jellemző PA és GB értékek meghatározása, különböző szempontok figyelembevételével. A glicin nyolc minimum konformere az irodalomban már jól meghatározott, így ebben az esetben nem volt szükség a konformációs tér leképezésére, azonban a másik két semleges aminosav esetén már nem ilyen egyértelmű a helyzet. A konformációs tér leképezéséhez a megfelelő torziós szögek forgatásával állítottunk elő próbageometriákat, majd azokat különböző bázisokkal, MP2 szinten optimáltuk. Az így kapott szerkezeteket osztályoztuk a relatív energiáik, a forgási állandóik és a kötésrendmátrixaik szerint, majd MP2/aug-cc-pVDZ elméleti szinten optimáltuk, valamint hogy meggyőződjünk arról, hogy a geometriák minimumokhoz tartoznak, harmonikus rezgési analízisüket is elvégeztük. Ezzel a módszerrel a ciszteinre (némi utómunkálattal) 85, míg a szerinre 95 minimum geometriát találtunk, köztük olyanokat, amelyek az alacsony relatív

energiájú régióban helyezkednek el. A protonált formák meghatározásának első lépése a lehetséges protonálódási helyek vizsgálata volt. Ez glicin esetén az amino-, karbonil- és hidroxilcsoportokat jelentette, míg a cisztein és szerin esetén már az oldallánc funkcióscsoportja (tiol- és hidroxilcsoport) is számításba jöhetett. A kérdés eldöntésére előzetes számításokat végeztünk, amelyek eredménye azt mutatta, hogy az α -szénatomon lévő hidroxilcsoport protonálódása egyik esetben sem vezet stabil szerkezethez. Érdekes még, hogy míg a cisztein esetén az oldallánc funkcióscsoportja protonálható, addig az analóg szeriné már nem. A protonált konformerek keresését már a glicinhez kapcsolódóan is el kellett végezzük, ekkor a forgatásokkal előállított próbaszerkezeteket alapértelmezetten MP2/aug-cc-pVDZ szinten optimáltuk harmonikus rezgési frekvenciáik meghatározása mellett (az adatpontok mennyisége ezt lehetővé tette), míg a másik két protonált aminosavnál a semleges formákhoz hasonlóan, először kisebb bázisokkal optimáltunk, majd az így kapott struktúrákat vetettük alá aug-cc-pVDZ bázissal való számításoknak. Ezen az elméleti szinten 3/8 N-/O-protonált glicin, 21/64/37 N-/O-/S-protonált cisztein és 15/46 N-/O-protonált szerin minimum konformert tudtunk azonosítani. A szakirodalomban fellelhető összes, kísérletileg meghatározott konformert be tudtuk azonosítani a semleges és protonált aminosavak esetén is.

Noha a konformerekkel kapcsolatos eredményeink már önmagukban helytállóak lennének az irodalomban, ennél is magasabb szintű számításokat végeztünk, a glicin és protonált glicin esetén az összes, míg a cisztein és szerin, valamint protonált formáik esetén a tíz legalacsonyabb (MP2/aug-cc-pVDZ szinten) relatív energiájú szerkezetre. Megjegyzendő, hogy egy korábbi, a disszertációhoz kötődő, de annak alapját nem képező publikációnkból⁴ a semleges glicinre rendelkezésre álló magasabb, CCSD(T)-F12b/aug-cc-pVTZ szintű geometriák, és DZ bázissal meghatározott rezgési frekvenciák, ezért azokat csak a protonált formákra kellett meghatározzuk. Ezután QZ bázissal meghatároztuk az összes szerkezetre a single-point energiákat. A cisztein és szerin (valamint protonált változataik) esetén az MP2/aug-cc-pVDZ geometriákat CCSD(T)-F12a/cc-pVDZ-F12 szinten optimáltuk, semleges esetben harmonikus rezgési analízist is végeztünk, majd TZ és QZ bázisokkal (ez utóbbi esetben kétféle (F12a és F12b) explicit korrelált módszerrel) single-point energiákat határoztunk meg. A tapasztalatunk az volt, hogy az MP2 módszerrel is már pontos geometriákat kapunk, azonban a kéntartalmú aminosav esetén ez az állítás megdőlni látszik, akár 1 kcal/mol-os geometriai effektusokkal is találkozhatunk. A meghatározott relatív energiák mellé az összes semleges és protonált aminosavra különböző kiegészítő

korrekciókat határoztunk meg (definícióik: 7.2.1.1–7.2.1.4., valamint 7.2.2.1–7.2.2.4. egyenletek): a megszakított csatolt-klaszter sorfejtésből származó δT és $\delta(Q)$ korrekciókat, a törzselektronok korreláltatásából származó $\Delta_{\text{törzs}}$ tagot, valamint a másodrendű Douglas–Kroll relativisztikus korrekciót. Ezek nagyjából a tized és század kcal/mol között vannak, a legelenyésőbb talán a Δ_{rel} , azonban néhány cisztein konformernél a többi korrekcióval összemérhetővé növekszik. Végezetül pedig a gyakorlati felhasználhatóság érdekében vizsgáltuk a 298,15 K-es relatív entalpiákat és relatív szabadentalpiákat.

A disszertáció utolsó kitűzött célja a protonálódáshoz tartozó PA és GB értékek meghatározása és vizsgálata volt. Ez az elérhető szakirodalomhoz képest több szempontból is új eredményekhez vezetett: minden eddiginél pontosabb, csatolt-klaszter szintű relatív energiákból indultunk ki, valamint itt is vizsgáltuk a konformereknél említett korrekciók hatását. Ez utóbbiak nagyságrendileg szintén a tized és század kcal/mol között vannak, azonban a másodrendű DK relativisztikus korrekció is hasonló járulékot ad, mint a többi tag. A protonálódásnál a legtöbb publikációval ellentétben nem csak a semleges aminosav globális minimumának protonálódását vizsgáltuk (a globális minimum aminosavvá), hanem konformerelegyeket feltételeztünk, amelyek populációit Boltzmann-eloszlással közelítettük. Ez a végső értékekben jónéhány tized kcal/mol-lal módosította a javasolt értékeinket. Végezetül pedig az összes lehetséges protonálódási helyhez tartozó termodinamikai mennyiségeket meghatároztuk, a korábbi publikációk csupán az energetikailag legkedvezőbb amino-protonálódásra koncentrálnak. A becsléseik szerint a karbonil-protonálódás legalább 30 kcal/mol-lal kedvezőtlenebb, a glicinnél ez még egy pontos becslés, azonban a ciszteinnél és szerinnél kevesebb, mint fele a különbség a két csoport protonálódása között. A meghatározott mennyiségeinket összevetettük három, az addigi kísérleti és elméleti eredményeket összefoglaló publikációval, azokkal jó egyezést kaptunk.

Természetesen mindig van mit pontosítani, továbbgondolni, ez ebben az esetben is így van. A konformerekkel kapcsolatosan érdemes lenne megvizsgálni, hogy konkrétan ezekre a molekulákra mekkora hibát hordoz az anharmonicitás elhanyagolása a relatív entalpiákban/szabadentalpiákban, de ez a végső protonaffinitás és gázfázisú bázicitás értékeknél is érdekes lehet. A következő lehetőség a szisztematikus konformkeresés robusztusságának vizsgálata lehet, ahogyan a ciszteinnél is látható, a módszer valószínűleg érzékeny a kiindulási geometriára és a megfelelő forgatások felbontására. A gépi tanulás vagy algoritmus alapú módszerek validálásához elengedhetetlen a hagyományos

módszerekkel történő alapos vizsgálat. Végezetül, a következő lépés a konformerek felhasználásával PES-ek fejlesztése lehet, amelyekkel a konformációs dinamika válik tanulmányozhatóvá, de akár egyszerű, protonálódási reakciókat is érdemes lehet vizsgálni.

9. SUMMARY

The studies of gas-phase amino acid minimum conformers (and their protonated counterparts) date back a very long time both experimentally and theoretically. With microwave and infrared spectroscopy measurements many amino acid conformers have been identified, based on the rotational and vibrational properties, although there are many difficulties in handling them. They decompose easily, especially if we manage to get them into gas-phase, and to selectively produce one conformer is challenging, since they can easily transform into other forms (the energy barriers are not so high), which leads to complex spectra. Many scientists (including us) approached this problem from the theoretical side and tried to map the conformational space of amino acids with quantum chemical computations. This can be done with systematic mapping or with the help of methods developed with algorithms and artificial intelligence, nevertheless, to do an comprehensive study, we need(ed) the rapid development of computers. The situation is similar in the case of the proton affinities and gas-phase basicities. Experimental methods rarely allow determining these absolute thermodynamical values, in general, we can only compare to references, if the reference values are inaccurate, the results will be too. We can overcome this problem with the tools of theoretical chemistry, although publications in the literature utilize many approximations and low level of theory.

My doctoral thesis aims at the most comprehensive, systematical, *ab initio* structural study of the glycine, cysteine and serine amino acids (and their protonated counterparts), also to determine the PA and GB values, with taking many perspectives into accounts. The eight minima of the glycine is well known in the literature, thus in this case, we do not need to map the conformational space, while in the case of the two other amino acids, the situation is not so straightforward. To map the conformational space, we generated test geometries by rotating the corresponding torsion angles and then optimized with different basis sets with the MP2 method. The resulting structures were classified according to their relative energies, rotational constants and bond order matrices, then optimized at the MP2/aug-cc-pVDZ theoretical level, and to make sure that the geometries belong to minima, their harmonic vibrational analysis was also performed. Using this method, 85 minimum geometries were found for cysteine (with some post-processing) and 95 for serine, including some in the low relative energy region. The first step in determining the protonated forms was to investigate the possible protonation sites. In the case of glycine, this meant the amino, carbonyl and hydroxyl groups, while in the case of cysteine and serine, the functional group of the side

chain (thiol and hydroxyl groups) should be considered. To address this issue, we carried out preliminary calculations, the results of which showed that protonation of the hydroxyl group on the α -carbon atom does not lead to a stable structure in either case. It is also interesting to note that while in the case of cysteine the functional group of the side chain can be protonated, that of the analogous serine cannot. The search for protonated conformers had to be carried out for the glycine too, in this case the test structures generated by rotations were optimized by default at the MP2/aug-cc-pVDZ level with determination of their harmonic vibrational frequencies (the amount of data points allowed this). For the other two protonated amino acids, and also for the neutral forms, we first performed the optimizations with smaller basis sets and then subjected the newly found structures to computations with the aug-cc-pVDZ basis set. At this theoretical level, we were able to identify 3/8 N-/O-protonated glycine, 21/64/37 N-/O-/S-protonated cysteine and 15/46 N-/O-protonated serine minimum conformers. All experimentally determined conformers in the literature could be identified for both neutral and protonated amino acids.

Although our results on the conformers themselves would be novelties in the literature on their own, we have performed even higher level calculations for all structures of the glycine and protonated glycine, and for the ten lowest (MP2/aug-cc-pVDZ level) relative energy structures for the cysteine and the serine (and their protonated forms). It should be noted that from an earlier publication,⁴ which is related to the dissertation, but this work is not based on that, higher-level geometries at CCSD(T)-F12b/aug-cc-pVTZ and vibrational frequencies with DZ basis set are available for neutral glycine, and therefore these quantities had to be determined only for the protonated forms. We then determined the single-point energies for all structures with QZ basis. In the case of cysteine and serine (and their protonated counterparts), the MP2/aug-cc-pVDZ geometries were optimized at the CCSD(T)-F12a/cc-pVDZ-F12 level, harmonic vibrational analysis was also performed in the neutral case, and single-point energies were determined with TZ and QZ bases (in the latter case using two (F12a and F12b) explicitly-correlated methods). Our experience was that we already obtain accurate geometries with the MP2 method, but this claim seems to be refuted in the case of the sulphur-containing amino acid forms, with geometric effects as large as 1 kcal/mol. In addition to the relative energies determined, various additional corrections have been defined for all neutral and protonated amino acids (defined in Equations 7.2.1.1 to 7.2.1.4 and 7.2.2.1 to 7.2.2.4): the δT and $\delta(Q)$ are corrections (from the incomplete coupled-cluster series), the Δ_{core} term (comes from the correlation of the core electrons), and the second-order

Douglas–Kroll relativistic correction. These are roughly in the range of tenths to hundredths of kcal/mol, perhaps the most negligible is Δ_{rel} , but for some cysteine conformers they increase to a value comparable with the other corrections. Finally, for practical use, we investigated the relative enthalpies and relative Gibbs free energies at 298.15 K.

The final goal of this doctoral thesis was to determine and investigate the PA and GB values associated with protonation. This led to new results in several aspects compared to the available literature: we started from more accurate relative energies (at the coupled-cluster level) than ever before and we also investigated the effect of the corrections mentioned for the separate conformers. These corrections also range in magnitude from tenths to hundredths of kcal/mol, but the second-order DK relativistic correction gives a similar contribution as the other terms. For protonation, in contrast to most publications, we did not only consider the protonation of the global minimum of the neutral amino acid (to the global minimum to amino-protonated amino acid), but we assumed conformer mixtures, the populations were calculated by the Boltzmann-distribution. This changed our proposed values by several tenths of kcal/mol in the final values. Finally, we determined the thermodynamic quantities associated with all possible protonation sites, previous publications only focused on the most energetically favourable amino-protonation. They estimate that carbonyl protonation is at least 30 kcal/mol less favourable, for glycine this is still an accurate estimate, but for cysteine and serine the difference between the protonation of the two functional group is less than half of that. Our determined quantities were compared with three publications summarising experimental and theoretical results to date, and good agreement was obtained.

Of course, there is always room for clarification and further reflection, and this is also the case here. With respect to the conformers, it would be worthwhile to investigate how much error is introduced by neglecting anharmonicity in relative enthalpies/Gibbs free energies for these molecules in particular, but this may also be of interest for the final proton affinity and gas-phase basicity values. The next option could be to investigate the robustness of systematic conformational analysis, as shown for cysteine, the method is likely to be sensitive to the initial geometry and the resolution of the corresponding rotations. To validate machine learning or algorithm-based methods, it is essential to thoroughly determine the conformers using traditional methods. Finally, the next step could be the development of PESs using conformers to study conformational dynamics, but it may also be worthwhile to investigate simple reactions, like the protonation.

10. KÖSZÖNETNYÍLVÁNÍTÁS

Szeretnék köszönetet mondani Dr. Czakó Gábornak a PhD tanulmányaim alatt nyújtott támogatásáért, témavezetésért, sokszori rugalmasságáért. Kiemelném azt, hogy a kutatómunkában „hagyott gondolkozni”, önálló ötleteket megvalósítani, saját megoldásokat keresni, ugyanakkor, ha olyan problémákba ütköztem, amelyekkel nem boldogultam egyedül, akkor a helyes irányba terelt.

Köszönöm Dr. Tóth Ágotának, hogy a Szegedi Tudományegyetem Kémia Doktori Iskolájában végezhettem el a képzésem. Bármilyen kérdéssel, személyes ügyel bátran fordulhattam hozzá, mint a Doktori Iskola vezetőjéhez.

Hálával tartozom a Czakó csoport valamennyi tagjának, de kiemelném Szűcs Tímeát, Gruber Balázst, Tajti Viktort és Cangtao „Frank” Yint, akikkel számtalan szakmai és nem szakmai programon, beszélgetésben vehettem részt, ezzel is előrelendítve a tanulmányaimat, barátságainkat.

Végül, de nem utolsó sorban köszönet illeti a Családom: a Feleségem, a Szüleim és a Bátyám. Ezalatt a négy év alatt bármikor számíthattam rájuk, munkásságom lelkes figyelemmel követték, ahogyan csak tudták, segítették a maguk módján.

11. HIVATKOZÁSOK

- 1 A. B. Nacsa & G. Czakó, *Phys. Chem. Chem. Phys.*, 2021, **23**, 9663–9671.
- 2 A. B. Nacsa & G. Czakó, *J. Phys. Chem. A*, 2022, **126**, 9667–9679.
- 3 A. B. Nacsa, M. Kigyósi & G. Czakó, *Phys. Chem. Chem. Phys.*, 2023, **25**, 8891–8902.
- 4 E. M. Orján, A. B. Nacsa & G. Czakó, *J. Comput. Chem.*, 2020, **41**, 2001–2014.
- 5 A. B. Nacsa, V. Tajti & G. Czakó, *J. Chem. Phys.*, 2023, **158**, 194306.
- 6 T. Gstir, T. Michaelson, B. A. Long, A. B. Nacsa, A. Ayasli, D. Swaraj, F. Zappa, F. Trummer, S. G. Ard, N. S. Shuman, G. Czakó, A. A. Viggiano & R. Wester, *Phys. Chem. Chem. Phys.*, 2023, **25**, 18711–18719.
- 7 A. B. Nacsa, C. Tokaji & G. Czako, *Faraday Discuss.*, 2024, DOI:10.1039/D3FD00161J.
- 8 P. C. D. Hawkins, *J. Chem. Inf. Model.*, 2017, **57**, 1747–1756.
- 9 D. Ferro-Costas, I. Mosquera-Lois & A. Fernández-Ramos, *J. Cheminform.*, 2021, **13**, 100.
- 10 L. Fang, E. Makkonen, M. Todorović, P. Rinke & X. Chen, *J. Chem. Theory Comput.*, 2021, **17**, 1955–1966.
- 11 G. Mancini, M. Fusè, F. Lazzari, B. Chandramouli & V. Barone, *J. Chem. Phys.*, 2020, **153**, 124110.
- 12 M. Ropo, M. Schneider, C. Baldauf & V. Blum, *Sci. Data*, 2016, **3**, 160009.
- 13 N. Kishimoto & H. Waizumi, *Chem. Phys. Lett.*, 2017, **685**, 69–76.
- 14 S. Ioppolo, G. Fedoseev, K.-J. Chuang, H. M. Cuppen, A. R. Clements, M. Jin, R. T. Garrod, D. Qasim, V. Kofman, E. F. van Dishoeck & H. Linnartz, *Nat. Astron.*, 2021, **5**, 197–205.
- 15 Y. Kuan, S. B. Charnley, H. Huang, W. Tseng & Z. Kisiel, *Astrophys. J.*, 2003, **593**, 848–867.
- 16 J. L. Bada, D. P. Glavin, G. D. McDonald & L. Becker, *Science*, 1998, **279**, 362–365.
- 17 E. Herbst, *Chem. Soc. Rev.*, 2001, **30**, 168–176.
- 18 V. Blagojevic, S. Petrie & D. K. Bohme, *Mon. Not. R. Astron. Soc.*, 2003, **339**, L7–L11.
- 19 J. L. Snow, G. Orlova, V. Blagojevic & D. K. Bohme, *J. Am. Chem. Soc.*, 2007, **129**, 9910–9917.
- 20 J. E. Elsila, J. P. Dworkin, M. P. Bernstein, M. P. Martin & S. A. Sandford, *Astrophys. J.*, 2007, **660**, 911–918.
- 21 V. Barone & C. Puzzarini, *EPJ Web Conf.*, 2020, **246**, 00021.
- 22 K. Kobayashi, H. Mita, Y. Kebukawa, K. Nakagawa, T. Kaneko, Y. Obayashi, T.

- Sato, T. Yokoo, S. Minematsu, H. Fukuda, Y. Oguri, I. Yoda, S. Yoshida, K. Kanda, E. Imai, H. Yano, H. Hashimoto, S. Yokobori & A. Yamagishi, *Astrobiology*, 2021, **21**, 1479–1493.
- 23 A. R. Dongré, J. L. Jones, Á. Somogyi & V. H. Wysocki, *J. Am. Chem. Soc.*, 1996, **118**, 8365–8374.
- 24 C. Bleiholder, S. Suhai & B. Paizs, *J. Am. Soc. Mass Spectrom.*, 2006, **17**, 1275–1281.
- 25 B. Paizs & S. Suhai, *Mass Spectrom. Rev.*, 2005, **24**, 508–548.
- 26 R. Boyd & Á. Somogyi, *J. Am. Soc. Mass Spectrom.*, 2010, **21**, 1275–1278.
- 27 V. H. Wysocki, G. Tsaprailis, L. L. Smith & L. A. Breci, *J. Mass Spectrom.*, 2000, **35**, 1399–1406.
- 28 P. B. Armentrout, A. L. Heaton & S. J. Ye, *J. Phys. Chem. A*, 2011, **115**, 11144–11155.
- 29 A. J. Russell, P. G. Thomas & A. R. Fersht, *J. Mol. Biol.*, 1987, **193**, 803–813.
- 30 M. Meot-Ner, *J. Am. Chem. Soc.*, 1984, **106**, 278–283.
- 31 S. G. Lias, J. F. Liebman & R. D. Levin, *J. Phys. Chem. Ref. Data*, 1984, **13**, 695–808.
- 32 E. P. L. Hunter & S. G. Lias, *J. Phys. Chem. Ref. Data*, 1998, **27**, 413–656.
- 33 A. G. Császár, *J. Am. Chem. Soc.*, 1992, **114**, 9568–9575.
- 34 A. G. Császár, *J. Mol. Struct.*, 1995, **346**, 141–152.
- 35 V. Kasalová, W. D. Allen, H. F. Schaefer, E. Czinki & A. G. Császár, *J. Comput. Chem.*, 2007, **28**, 1373–1383.
- 36 R. M. Balabin, *Chem. Phys. Lett.*, 2009, **479**, 195–200.
- 37 V. Barone, M. Biczysko, J. Bloino & C. Puzzarini, *Phys. Chem. Chem. Phys.*, 2013, **15**, 10094–10111.
- 38 V. Barone, M. Biczysko, J. Bloino & C. Puzzarini, *Phys. Chem. Chem. Phys.*, 2013, **15**, 1358–1363.
- 39 V. Barone, M. Fusè, F. Lazzari & G. Mancini, *J. Chem. Theory Comput.*, 2023, **19**, 1243–1260.
- 40 F. Gabas, R. Conte & M. Ceotto, *J. Chem. Theory Comput.*, 2017, **13**, 2378–2388.
- 41 A. Karton, L.-J. Yu, M. K. Kesharwani & J. M. L. Martin, *Theor. Chem. Acc.*, 2014, **133**, 1483.
- 42 H.-W. KE, L. RAO, X. XU & Y.-J. YAN, *J. Theor. Comput. Chem.*, 2008, **07**, 889–909.
- 43 M. Kieninger, S. Suhai & O. N. Ventura, *J. Mol. Struct. THEOCHEM*, 1998, **433**, 193–201.
- 44 K. Zhang & A. Chung-Phillips, *J. Phys. Chem. A*, 1998, **102**, 3625–3634.

- 45 M. Noguera, L. Rodríguez-Santiago, M. Sodupe & J. Bertran, *J. Mol. Struct. THEOCHEM*, 2001, **537**, 307–318.
- 46 F. Jensen, *J. Am. Chem. Soc.*, 1992, **114**, 9533–9537.
- 47 I. A. Topol, S. K. Burt, M. Toscano & N. Russo, *J. Mol. Struct. THEOCHEM*, 1998, **430**, 41–49.
- 48 R. Wu & T. B. McMahon, *ChemPhysChem*, 2008, **9**, 2826–2835.
- 49 R. Wu & T. B. McMahon, *J. Phys. Chem. B*, 2009, **113**, 8767–8775.
- 50 R. M. Balabin, *J. Phys. Chem. Lett.*, 2010, **1**, 20–23.
- 51 R. M. Balabin, *Phys. Chem. Chem. Phys.*, 2012, **14**, 99–103.
- 52 G. Bazsó, G. Magyarfalvi & G. Tarczay, *J. Mol. Struct.*, 2012, **1025**, 33–42.
- 53 G. Bazsó, G. Magyarfalvi & G. Tarczay, *J. Phys. Chem. A*, 2012, **116**, 10539–10547.
- 54 N. C. Polfer, *Chem. Soc. Rev.*, 2011, **40**, 2211.
- 55 S. Gronert & R. A. J. O’Hair, *J. Am. Chem. Soc.*, 1995, **117**, 2071–2081.
- 56 M. J. S. Dewar, E. G. Zoebisch, E. F. Healy & J. J. P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902–3909.
- 57 J. C. Dobrowolski, J. E. Rode & J. Sadlej, *J. Mol. Struct. THEOCHEM*, 2007, **810**, 129–134.
- 58 J. J. Wilke, M. C. Lind, H. F. Schaefer, A. G. Császár & W. D. Allen, *J. Chem. Theory Comput.*, 2009, **5**, 1511–1523.
- 59 V. Riffet, G. Frison & G. Bouchoux, *Phys. Chem. Chem. Phys.*, 2011, **13**, 18561.
- 60 J. C. Dobrowolski, M. H. Jamróz, R. Kołos, J. E. Rode & J. Sadlej, *ChemPhysChem*, 2007, **8**, 1085–1094.
- 61 G. Magyarfalvi, G. Tarczay & E. Vass, *WIREs Comput. Mol. Sci.*, 2011, **1**, 403–425.
- 62 M. E. Sanz, S. Blanco, J. C. López & J. L. Alonso, *Angew. Chemie Int. Ed.*, 2008, **47**, 6216–6220.
- 63 E. E. Najbauer, G. Bazsó, S. Góbi, G. Magyarfalvi & G. Tarczay, *J. Phys. Chem. B*, 2014, **118**, 2093–2103.
- 64 J. D. Steill, J. Szczepanski, J. Oomens, J. R. Eyler & A. Brajter-Toth, *Anal. Bioanal. Chem.*, 2011, **399**, 2463–2473.
- 65 S. Jarmelo, L. Lapinski, M. J. Nowak, P. R. Carey & R. Fausto, *J. Phys. Chem. A*, 2005, **109**, 5689–5707.
- 66 R. Miao, C. Jin, G. Yang, J. Hong, C. Zhao & L. Zhu, *J. Phys. Chem. A*, 2005, **109**, 2340–2349.
- 67 K. He & W. D. Allen, *J. Chem. Theory Comput.*, 2016, **12**, 3571–3582.
- 68 B. Lambie, R. Ramaekers & G. Maes, *J. Phys. Chem. A*, 2004, **108**, 10426–10433.

- 69 S. Blanco, M. E. Sanz, J. C. López & J. L. Alonso, *Proc. Natl. Acad. Sci.*, 2007, **104**, 20183–20188.
- 70 E. E. Najbauer, G. Bazsó, R. Apóstolo, R. Fausto, M. Biczysko, V. Barone & G. Tarczay, *J. Phys. Chem. B*, 2015, **119**, 10496–10510.
- 71 F. X. Sunahori, G. Yang, E. N. Kitova, J. S. Klassen & Y. Xu, *Phys. Chem. Chem. Phys.*, 2013, **15**, 1873–1886.
- 72 V. Scutelnic, M. A. S. Perez, M. Marianski, S. Warnke, A. Gregor, U. Rothlisberger, M. T. Bowers, C. Baldauf, G. von Helden, T. R. Rizzo & J. Seo, *J. Am. Chem. Soc.*, 2018, **140**, 7554–7560.
- 73 Z. B. Maksić & B. Kovačević, *Chem. Phys. Lett.*, 1999, **307**, 497–504.
- 74 Z. B. Maksić, B. Kovačević & R. Vianello, *Chem. Rev.*, 2012, **112**, 5240–5270.
- 75 S. Gronert, D. C. Simpson & K. M. Conner, *J. Am. Soc. Mass Spectrom.*, 2009, **20**, 2116–2123.
- 76 E. F. Strittmatter & E. R. Williams, *Int. J. Mass Spectrom.*, 1999, **185**, 935–948.
- 77 T. C. Dinadayalane, G. N. Sastry & J. Leszczynski, *Int. J. Quantum Chem.*, 2006, **106**, 2920–2933.
- 78 A. G. Harrison, *Mass Spectrom. Rev.*, 1997, **16**, 201–217.
- 79 A. Baeten, F. De Proft & P. Geerlings, *Int. J. Quantum Chem.*, 1996, **60**, 931–939.
- 80 C. Pepe, S. Rochut, J. P. Paumard & J. C. Tabet, *Rapid Commun. Mass Spectrom.*, 2004, **18**, 307–312.
- 81 S. Bouchonnet & Y. Hoppilliard, *Org. Mass Spectrom.*, 1992, **27**, 71–76.
- 82 G. Bouchoux, S. Huang & B. S. Inda, *Phys. Chem. Chem. Phys.*, 2011, **13**, 651–668.
- 83 C. J. Cassady, S. R. Carr, K. Zhang & A. Chung-Phillips, *J. Org. Chem.*, 1995, **60**, 1704–1712.
- 84 G. Bojesen, *J. Am. Chem. Soc.*, 1987, **109**, 5557–5558.
- 85 G. S. Gorman, J. P. Speir, C. A. Turner & I. J. Amster, *J. Am. Chem. Soc.*, 1992, **114**, 3986–3988.
- 86 F. M. Siu & C. M. Che, *J. Phys. Chem. A*, 2006, **110**, 12348–12354.
- 87 A. Chung-Phillips, *J. Phys. Chem. A*, 2005, **109**, 5917–5932.
- 88 A. Moser, K. Range & D. M. York, *J. Phys. Chem. B*, 2010, **114**, 13911–13921.
- 89 G. Bouchoux & J.-Y. Salpin, *Eur. J. Mass Spectrom.*, 2003, **9**, 391–402.
- 90 G. Bouchoux, *Mass Spectrom. Rev.*, 2012, **31**, 391–435.
- 91 E. Uggerud, *Theor. Chem. Accounts Theory, Comput. Model. (Theoretica Chim. Acta)*, 1997, **97**, 313–316.
- 92 G. Bouchoux, D.-A. Buisson, C. Colas & M. Sablier, *Eur. J. Mass Spectrom.*, 2004, **10**, 977–992.

- 93 C. Afonso, F. Modeste, P. Breton, F. Fournier & J.-C. Tabet, *Eur. J. Mass Spectrom.*, 2000, **6**, 443–449.
- 94 S. Gronert, D. C. Simpson & K. M. Conner, *J. Am. Soc. Mass Spectrom.*, 2009, **20**, 2116–2123.
- 95 N. F. Brás, M. A. S. Perez, P. A. Fernandes, P. J. Silva & M. J. Ramos, *J. Chem. Theory Comput.*, 2011, **7**, 3898–3908.
- 96 B. Hernández, F. Pflüger, A. Adenier, S. G. Kruglik & M. Ghomi, *Phys. Chem. Chem. Phys.*, 2011, **13**, 17284.
- 97 F. Freeman, I. T. Adesina, J. Le La, J. Y. Lee & A. A. Poplawski, *J. Phys. Chem. B*, 2013, **117**, 16000–16012.
- 98 M. V. Roux, C. Foces-Foces, R. Notario, M. A. V. Ribeiro da Silva, M. das D. M. C. Ribeiro da Silva, A. F. L. O. M. Santos & E. Juaristi, *J. Phys. Chem. B*, 2010, **114**, 10530–10540.
- 99 National Institute of Standards & Technology, <https://www.nist.gov/>, Utolsó látogatás dátuma: 2024.03.23.
- 100 J. C. Slater, *Phys. Rev.*, 1951, **81**, 385–390.
- 101 C. Møller & M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618–622.
- 102 J. Noga & R. J. Bartlett, *J. Chem. Phys.*, 1987, **86**, 7041–7050.
- 103 M. Kállay & J. Gauss, *J. Chem. Phys.*, 2005, **123**, 214105.
- 104 M. K. Kesharwani, N. Sylvetsky, A. Köhn, D. P. Tew & J. M. L. Martin, *J. Chem. Phys.*, 2018, **149**, 154109.
- 105 T. B. Adler, G. Knizia & H.-J. Werner, *J. Chem. Phys.*, 2007, **127**, 221106.
- 106 D. Manna, M. K. Kesharwani, N. Sylvetsky & J. M. L. Martin, *J. Chem. Theory Comput.*, 2017, **13**, 3136–3152.
- 107 W. J. Hehre, R. Ditchfield & J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257–2261.
- 108 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007–1023.
- 109 J. G. Hill, S. Mazumder & K. A. Peterson, *J. Chem. Phys.*, 2010, **132**, 054108.
- 110 M. S. & others. H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, Molpro, version 2015.1, a package of ab initio programs., <http://www.molpro.net>, Utolsó látogatás dátuma: 2024.03.23.
- 111 M. Kállay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csontos, J. Csóka, P. B. Szabó, L. Gyevi-Nagy, B. Hégyel, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei & Á. Ganyecz, *J. Chem. Phys.*, 2020, **152**, 074107.
- 112 P. D. M. & Á. G. M. Kállay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csontos, J. Csóka, P. B. Szabó, L. Gyevi-Nagy, B. Hégyel, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, MRCC, a quantum chemical program suite, www.mrcc.hu, Utolsó látogatás dátuma: 2024.03.23.
- 113 T. Szidarovszky, G. Czako & A. G. Császár, *Mol. Phys.*, 2009, **107**, 761–775.

- 114 E. Czinki & A. G. Császár, *Chem. – A Eur. J.*, 2003, **9**, 1008–1019.
- 115 M. Douglas & N. M. Kroll, *Ann. Phys. (N. Y.)*, 1974, **82**, 89–155.
- 116 G. Knizia, T. B. Adler & H.-J. Werner, *J. Chem. Phys.*, 2009, **130**, 054104.
- 117 K. Raghavachari, G. W. Trucks, J. A. Pople & M. Head-Gordon, *Chem. Phys. Lett.*, 1989, **157**, 479–483.
- 118 W. A. de Jong, R. J. Harrison & D. A. Dixon, *J. Chem. Phys.*, 2001, **114**, 48.
- 119 A. Á. Dékány & G. Czakó, *J. Comput. Chem.*, 2022, **43**, 19–28.
- 120 X. He, L. Fusti-Molnar & K. M. Merz, *J. Phys. Chem. A*, 2009, **113**, 10096–10103.