

Doktori disszertáció tézisei

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

BEVEZETÉS

Az aminosavak konformereinek, vagyis stabil minimum geometriáinak vizsgálata felkapott téma mind a kísérleti, mind az elméleti kutatók körében, hiszen flexibilis kötéseiknek köszönhetően számos stabil formával rendelkeznek. Az egyes konformereket a különböző rezgési és forgási tulajdonságaik („ujjlenyomataik”) alapján lehet beazonosítani infravörös vagy mikrohullámú spektroszkópiás mérések segítségével. Ehhez elméleti oldalról a molekulák konformációs terét kell feltérképeznünk és beazonosítani a potenciálisenergia-felületükön (PES) található minimumokhoz tartozó elrendeződéseket. Ez történhet algoritmusok és/vagy mesterséges intelligencia segítségével, vagy pedig „hagyományos”, szisztematikus kereséssel, ez utóbbit került alkalmazásra a disszertációhoz kötődő munkában is. A lényege, hogy nagyszámú kiindulási szerkezetet állítunk elő geometriai optimálásokhoz a megfelelő kötések mentén való szisztematikus elforgatással. A legkisebb aminosavnak, a glicinnek, már évtizedekkel ezelőtt elméleti úton meghatározták a nyolc stabil minimum konformerét, azokból ötöt ki is tudtak mutatni kísérletileg, azonban a nagyobb aminosavak – mint a cisztein és szerin – esetén már nem ilyen egyértelmű az irodalom felelhető források konklúziója. Az elméleti úton meghatározott geometriák, relatív energiák és rezgési frekvenciák lehetővé teszik a konformerek azonosítását, például az asztrokémikusok számára, akik az élet keletkezéséhez elengedhetetlen molekulák képződését, jelenlétét és reakcióit vizsgálják a csillagközi térben. Az egyik legegyszerűbb reakció, amit jellemezhetünk, az a protonálódás:



ahol B az analóg semleges bázis, H^+ a szabad proton, BH^+ pedig a protonált konjugált sav. A reakciót kísérő változások számszerűsítésére a protonaffinitás (PA, entalpia) és gázfázisú bázicitás (GB, szabadentalpia) mennyiségek kerültek bevezetésre. Abszolút értékeket leginkább csak elméleti úton lehet meghatározni, kísérletileg általában csak relatív értékekhez juthatunk különböző protonálódási reakciók kinetikájának, arányának és termodinamikájának tanulmányozásával. Az aminosavak PA és GB mennyisége számos folyamatot befolyásol: a fehérjék fragmentálódását, enzimek aktivitását, de természetesen az űrben lejátszódó ion-molekula reakciókat is.

A rendelkezésre álló szakirodalom elemzése után három főbb célt tűztünk ki a disszertációval kapcsolatban. Az első, hogy minden eddiginél részletesebb kereséssel, a Hartree–Fock módszerrel továbblépve beazonosítsuk a semleges és protonált glicin, cisztein és szerin aminosavak konformereit. A protonált formák lehetséges protonálódási helyek vizsgálata is

szükséges, hiszen eddig jellemzően csak a termodinamikailag legkedvezőbb, aminosav protonálódást elemezték, azonban a nagy pontosságú kvantumkémiai számítások még precízebben megadhatják a különböző funkciócsoportok közötti különbségeket. A következő lépés ugyanezen semleges és protonált molekulákra vonatkozóan az első csatolt-klaszter szintű geometriák és/vagy harmonikus rezgési frekvenciák meghatározása, valamint különböző elméleti korrekciók hatásának vizsgálata. Az utolsó feladat pedig a három aminosav protonaffinitásának és gázfázisú bázicitásának meghatározása csatolt-klaszter szinten, a különböző korrekciókat is figyelembe véve a különböző protonálódási helyekre vonatkozóan mind globális minimumok, mind konformerelegyeket figyelembe véve. Noha a konformerkeresésnek és PA/GB értékek meghatározásának még mindig vannak olyan aspektusai, amiket nem vizsgáltunk, ez a munkánk értékét nem csorbítja. A disszertáció alapját képező publikációink megannyi új, eddig még nem vizsgált eredményt tartalmaznak, azokat számos kutatási formában fel lehet használni: reaktív vagy konformációs PES-ek fejlesztése, gépi tanulás alapú algoritmusok validálása, anharmonikus rezgési analízis elvégzése, stb.

MÓDSZEREK

A semleges és protonált aminosavak esetén is a konformációs tér leképezéséhez az adatpontokat egy szisztematikus módszerrel, a megfelelő torziós szögek 60° -onkénti (ciszteinnél bizonyos tengelyek mentén 30° -onkénti) változtatásával állítottam elő, semleges formáknál az irodalomban fellelhető globális minimumokból kiindulva, míg protonált esetben a protonálódási helyek vizsgálatakor kapott legstabilabb szerkezeteket forgatva. A kiindulási geometriák optimalizálásához és a harmonikus rezgési analízis elvégzéséhez alkalmazott elméleti szint a protonált glicin esetén az MP2/aug-cc-pVDZ volt. A semleges cisztein esetében az optimalizálásokhoz az MP2 módszert alkalmaztunk 3-21G, 6-31G, 6-31++G, 6-31G**, 6-31++G** és cc-pVDZ bázisokkal (később HF/3-21G is), míg a protonált cisztein, valamint a szerin két formájánál már csupán a 6-31++G** és cc-pVDZ bázisokkal. A két nagyobb aminosav esetén az így kapott stacionárius pontokat újraoptimaltunk MP2/aug-cc-pVDZ szinten és meghatároztuk harmonikus rezgési frekvenciáikat. A protonálódási helyek vizsgálatához ugyanezen az elméleti szinten végeztünk optimalizálásokat és rezgési analízist különböző próbaszerkezeteken, amelyeket a semleges aminosavakból állítottunk elő, a funkciócsoportjaikhoz egy protont kötve különböző variációkban. A semleges glicinnél egy korábbi munkánkból átvett CCSD(T)-F12b/aug-cc-pVTZ geometriákon végeztünk single-point energiaszámításokat aug-cc-pVQZ bázissal. A cisztein és a szerin esetén (mind semleges mind

protonált esetben) a konformerek nagy száma miatt csupán a 10 legalacsonyabb MP2/aug-cc-pVDZ szintű relatív energiával rendelkező szerkezetet vizsgáltuk tovább. CCSD(T)-F12a/cc-pVDZ-F12 szintű geometriai optimálásokat (és semleges esetben harmonikus rezgési analízist) végeztünk, majd meghatározzuk a single-point energiákat TZ és QZ bázis mellett is (ez utóbbi esetben a CCSD(T)-F12b módszerrel is). Az elérhető legpontosabb geometriákon megvizsgáltuk a poszt-(T) korrekciók, a törzselektronok korreláltatásából származó korrekciók, valamint a másodrendű Douglas–Kroll relativisztikus korrekciók értékeit is. A kvantumkémiai számításokhoz a MOLPRO 2015-ös verzióját használtam, illetve a poszt-CCSD(T) korrekciók meghatározásánál szükség volt az MRCC programra. A kutatás során kapott nagymennyiségű adat feldolgozása, kiértékelése és osztályozása az általam írt Fortran és Python kódokkal történt.

EREDMÉNYEK

A következő tézispontok mindegyike a három, a disszertáció alapját képező közleményen alapszik, így ezeket külön nem jelölöm.

T1. *A semleges cisztein 85, míg a semleges szerin esetén 95 stabil minimum konformer szerkezetét és harmonikus rezgési frekvenciáját határoztuk meg MP2/aug-cc-pVDZ elméleti szinten a konformációs terük szisztematikus feltérképezése (és néhány kiegészítés) után.*

Mivel a semleges glicin konformerei az irodalomban már jól meghatározottak, többek között mi is foglalkoztunk már velük korábban, ezért a disszertációhoz kapcsolódó munka keretében csak a cisztein és a szerin konformereinek megtalálása volt a feladat. Ezt egy szisztematikus módszerrel, a megfelelő torziós szögek 60° -os változtatásával végeztük el, amely nagyszámú adatponthoz vezetett. A próbageometriákat MP2 módszerrel különböző bázisok mellett optimáltuk, a cisztein esetén a 6-31++G** és cc-pVDZ bizonyultak hasznosnak, így a szerinnél már csak ezeket használtuk a keresésre. A ciszteinre kapott szerkezeteket az irodalmi források kritikus elemzésével kiegészítettük, majd a két aminosavnál kapott geometriákat tovább optimáltuk és meghatároztuk harmonikus rezgési frekvenciáikat MP2/aug-cc-pVDZ elméleti szinten. Mindkét aminosavra vonatkozóan számos új geometriát azonosítottunk, köztük olyanokat is, amelyek az alacsony relatív energiájú régióban helyezkednek el.

T2. *Beazonosítottuk a glicin, cisztein és szerin aminosavak lehetséges protonálódási helyeit (amino-, karbonil-, valamint cisztein esetén a tiolcsoport), majd az így kapott formák minimum geometriáit is megkerestük a megfelelő potenciálisenergia-felületeken. MP2/aug-cc-pVDZ szinten 3 N- és 8 O-protonált glicin, 21 N-, 64 O-, és 37 S-protonált cisztein, valamint 15 N- és 46 O-protonált szerin minimumot találtunk, illetve meghatároztuk a harmonikus rezgési frekvenciáikat.*

A glicin esetén az amino-, karbonil- és hidroxilcsoportok protonálódása volt valószínűsíthető. Az irodalommal egybehangzóan a mi vizsgálataink is azt mutatták, hogy a hidroxilcsoport protonálódásával nem tud stabil forma kialakulni, csak a másik kettő esetben. A két nagyobb aminosavnál kémiai megfontolások alapján az oldallánc tiol- és hidroxilcsoportja is protonálódhat, ám ezek közül csak a tiol-protonálódás vezet stabil molekulához. Ugyan termodinamikailag az aminocsoportok protonálódása a legkedvezőbb, mi a többi, az irodalomban eddig elhanyagolt formát is kerestük. A módszer a semleges aminosavaknál alkalmazotthoz hasonló volt, a protonált glicinek esetén a keresést MP2/aug-cc-pVDZ elméleti szinten végeztük, a protonált cisztein és szerin esetében azonban a kiindulási geometriák

nagyságrendekkel nagyobb száma ezt nem tette lehetővé, így ezekre a 6-31++G** és cc-pVDZ bázisokat alkalmaztuk, majd a potenciális minimum geometriákat újraoptimáltuk (és meghatároztuk harmonikus rezgési frekvenciáikat) aug-cc-pVDZ bázis mellett. Ezen formákra is igaz, hogy számos új, korábbi publikációkban nem szereplő geometriát találtunk.

T3. *A legmélyebb energiájú semleges és protonált glicin, cisztein és szerin konformerekre elsőként explicit-korrelált csatolt-klaszter szintű relatív energiákat és/vagy geometriákat és/vagy harmonikus rezgési frekvenciákat határoztunk meg, valamint azt találtuk, hogy a megszakított csatolt-klaszter sorfejtés, a törzselektronok korreláltatása és a másodrendű Douglas–Kroll relativisztikus korrekció a tized-század kcal/mol-os nagyságrendben befolyásolják az eredményeink pontosságát.*

Semleges glicin esetén egy korábbi munkánkból rendelkezésre álltak CCSD(T)-F12b/aug-cc-pVTZ szintű geometriák és relatív energiák, ezért a korrekciók meghatározása mellett csupán a QZ energiák meghatározására volt szükség ebben az esetben. A másik két aminosavnál a tíz legmélyebb MP2/aug-cc-pVDZ relatív energiájú konformert vizsgáltuk tovább (mind semleges, mind protonált esetben, a protonálódási helyekre nézve külön-külön) csatolt-klaszter módszerekkel. A geometriai optimalásokat CCSD(T)-F12a/cc-pVDZ-F12 szinten végeztük el, semleges esetben a harmonikus rezgési frekvenciákat is meghatároztuk. TZ és QZ bázisokkal single-point energiákat határoztunk meg ezeken a szerkezeteken. A korrekciók a tized és század kcal/mol-os nagyságrendbe esnek, közülük néhány (semleges/protonált) cisztein konformertől eltekintve a relativisztikus korrekció értéke század kcal/mol alatt van. Különböző megfontolások téve a kapott relatív energiák/entalpiák pontosságára megadhatunk egy becsült hibahatárt.

T4. *Meghatároztuk a glicinhez, a ciszteinhez és a szerinhez tartozó csatolt-klaszter szintű protonaffinitás és gázfázisú bázicitás értékeket az összes lehetséges protonálódási helyre vonatkozóan, illetve a T3. pontban említett korrekciókon kívül elemeztük annak a hatását is, ha konformer elegyeket veszünk figyelembe a globális minimumok helyett. Az amino-protonálódásra vonatkozó értékeink jó egyezést mutatnak az irodalmi, leginkább kísérleti munkákon alapuló javaslatokkal, a karbonilcsoport protonálódását viszont termodinamikailag jóval kedvezőbbnek találtuk a korábbi feltevésekhez képest.*

Az irodalomban felelhető publikációk a protonálódást jellemző két mennyiséget csupán a termodinamikailag legkedvezőbb, amino-protonálódásra határozták meg. Az eddigi legpontosabb, explicit-korrelált csatolt-klaszter szintű relatív energiáink segítségével

meghatároztuk ezeket az értékeket az összes funkcióscsoportra vonatkozóan (csupán a globális minimumokat vagy konformerelegyeket figyelembe véve), ezzel megmutattuk, hogy a korábban becsülteknél jóval alacsonyabb (különösen a nagyobb aminosavak esetén) az energiabeli különbség az egyes helyek protonálódása között. A T3. pontban felsorolt korrekciók nagyságát itt is vizsgáltuk és segítségükkel (valamint rezgési frekvenciák meghatározása során tett közelítéseinket figyelembe véve) becslést tettünk az eredményeink pontosságára. A glicinre javasolt PA/GB értékeink $212,4 \pm 0,3 / 204,8 \pm 0,6$ kcal/mol, ciszteinre $216,4 \pm 0,4 / 208,2 \pm 0,6$ kcal/mol, míg a szerinre $218,1 \pm 0,4 / 209,9 \pm 0,6$ kcal/mol.

PUBLIKÁCIÓS LISTA (MTMT azonosító: 10073207)

Doktori értekezés alapját képező közlemények listája (\sum IF = 10,145)

A. B. Nacsa & G. Czakó: Benchmark ab initio proton affinity of glycine, *Phys. Chem. Chem. Phys.* **23**, 9663 (2021), IF = 3,945

A. B. Nacsa & G. Czakó: Benchmark ab initio determination of the conformers, proton affinities and gas-phase basicities of cysteine, *J. Phys. Chem. A* **126**, 9667 (2022), IF = 2,9

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.* **25**, 8891 (2023), IF = 3,3

Doktori értekezéshez kapcsolódó, de alapját nem képező közlemény (\sum IF = 3,376)

E. M. Orján, **A. B. Nacsa**, G. Czakó: Conformers of dehydrogenated glycine isomers, *J. Comput. Chem.* **41**, 2001 (2020), IF = 3,376

Doktori értekezéshez nem kapcsolódó közleményeim listája (\sum IF = 11,1)

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.* **158**, 194306 (2023), IF = 4,4

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.* **25**, 18711 (2023), IF = 3,3

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.* DOI: 10.1039/D3FD00161J (2024), IF = 3,4

Összes IF = 24,621

ELŐADÁSOK

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 E2 reactions in full (24) dimensions

A. B. Nacsa, C. Tokaji, G. Czakó

New directions in molecular scattering Faraday Discussion, Edinburgh, Egyesült Királyság, 2024

A szerin kvantumkémiai konformációs analízise és ab initio protonaffinitása

M. Kígyósi, **A. B. Nacsa**, G. Czakó

XXXVI. Országos Tudományos Diákköri Konferencia, Kémia és Vegyipari szekció, Szeged, Magyarország, 2023

Dehidrogénezett és protonált aminosavak ab initio konformeranalízise

A. B. Nacsa, E. M. Orján, G. Czakó

Magyar Tudományos Akadémia Reakciókinetikai és Fotokémiai Munkabizottsága, Mátrafüred, Magyarország, 2022

Ab initio conformational analysis of dehydrogenated and protonated amino acids

A. B. Nacsa, E. M. Orján, G. Czakó

Magyar Tudományos Akadémia Anyag- és Molekulaszerkezeti Munkabizottsága, Mátrafüred, Magyarország, 2022

A glicin protonaffinitásának nagy pontosságú ab initio meghatározása

A. B. Nacsa, G. Czakó

KeMoMo-QSAR Szimpózium, Szeged, Magyarország, 2021

POSZTEREK

Theoretical insights into the $\text{Cl}^- + \text{CH}_3\text{I}$ and $\text{F}^- + \text{CF}_3\text{CH}_2\text{I}$ reactions

A. B. Nacsa, V. Tajti, G. Czakó

Dynamics of Molecular Collisions, Snowbird, UT, USA, 2023

Towards the first-principles dynamics of the $\text{Cl}^- + \text{CH}_3\text{I}$ and $\text{F}^- + \text{CF}_3\text{CH}_2\text{I}$ reactions

A. B. Nacsa, V. Tajti, G. Czakó

Molecular Interactions and Dynamics Gordon Research Conference, Easton, MA, USA, 2022