

Bírálat

Tóth Gergely MTA doktori értekezéséről

Tóth Gergely „Adatok értelmezése szimulációkkal és a többváltozós statisztika eszközeivel a kémiában” című MTA doktori értekezése olvasásához nagy várakozással fogtam hozzá. Várakozásom legfőbb oka az volt, hogy a jelölt szakmai pályafutását a ”kezdetektől figyelemmel kísértem”, ha lehet így fogalmazni akkor, amikor azok a bizonyos kezdetek mindkettőnk számára azok voltak, első tétova szakmai lépéseinket egymás mellett tettük meg. És tudományos pályafutásunk hosszú időn át párhuzamosan haladt.

Elöljáróban le kell szögeznem, hogy a doktori munkával kapcsolatos véleményem kettős, ennek megfelelően a továbbiakban külön értékelem a jelölt eddigi, az értekezés alapját jelentő szakmai munkásságát illetve magát a disszertációt mint önálló tudományos művet.

A disszertáció mögött álló publikációs tevékenység jelentős: a dolgozat 65, zömében rangos nemzetközi folyóiratokban megjelent cikken alapul. E nagyszámú publikáció három évtizednyi kutatómunka gyümölcse. A jelölt tudományos munkássága igen szerteágazó, azonban jól sikerült megfognia az ezt végigkísérő vezérfonalat, nevezetesen hogy munkássága végig a numerikus adatelemzés körül forgott, legyen szó eközben statisztikus mechanikáról vagy éppen kemometriáról. A szerző egy igazi öntörvényű kutató, akit pályafutása során végig az őszinte kíváncsiság hajtott, érdeklődésében soha nem befolyásolták az aktuális tudományos divatok, hivatkozások gyűjtésére alkalmas ”hot” témák, munkáit eleinte egy-egy idősebb, tapasztaltabb kollégával való laza, alkalmi együttműködésekben csinálta, később alapvetően nagyobb csoporthoz, iskolához nem tartozó, néhány diákot irányító magányos kutatóként végezte. Ennek a hozzáállásnak a tiszteletet parancsoló mivolta és előnyei mellett megvannak a maga kétségtelen hátrányai is. Nem véletlen hogy a színvonalas, kiterjedt és sok publikációt termelő életmű hivatkozottsága semmiképpen nem mondható kimagaslónak (1000 alatti független hivatkozás). A hivatkozások számomra meglepően alacsony száma azt is mutatja, hogy a jelölt életművével nem gyakorolt akkora hatást a nemzetközi tudományos életre, és nem végzett olyan mértékű iskolateremtő tevékenységet amelyet

tudása, képességei és a sok évtizedes kemény szakmai munkája alapján akár végezhetett is volna. Cserébe megmaradt mindig új utakat kereső, a külsőségektől többé-kevésbé független, szabad kutató szellemnek. Félreértések elkerülése végett fontos itt leszögezni, hogy a fenti kritikám semmiképpen sem jelenti azt, hogy a jelöltnek az eddigi tudományos és iskolateremtő tevékenysége ne haladná meg feltétlenül az MTA Doktora fokozatnál elvártakat, mindössze annyit jelent, hogy mindez akár ennél még sokkal jelentősebb is lehetett volna.

Maga a dolgozat MTA doktori értekezéshez képest meglepően rövid, mindössze 110 oldal terjedelmű, melyet köszönetnyilvánítás, tartalom- és hivatkozásjegyzék illetve egy rövid függelék egészít ki. A dolgozat 136 oldalas összterjedelme még mindig alatta marad az akadémiai doktori értekezések szakmai részénél megszokottnak. A dolgozat szerkesztését illetően a legmeghökkenőbb rész a módszertani összefoglalás teljes hiánya. Az együtt 8 oldal összterjedelmű rövid bevezetőt és a célkitűzéseket tárgyaló fejezetek után rögtön következnek az eredményeket taglaló fejezetek, amit egy 3 oldalas összefoglalás követ. A jelölt magyarázatával, miszerint a használt módszerek túl szerteágazóak volnának egy értekezésben való áttekintéshez, nem értek egyet. Maga a dolgozat az alanyi költészet jegyeit mutatja – a módszerek rigorózus ismertetése helyett személyes sztorizgatást kapunk, ami átível az egész dolgozaton. Olvashatunk arról, hogy hallgató korában egy adott gyakorlaton milyen mérést kellett csinálnia, munkahelyén melyik kolléga mikor tért haza külföldről, konferenciákon kik hogyan nem értették meg a munkáit, mikor voltak belső költözések a munkahelyén és kivel osztozott az irodáján, melyik eredményének a publikálása milyen nehézségekbe ütközött, vagy éppen posztdoktori éve alatt az intézetében más csoportokban milyen kutatómunka folyt. Ezek a történetek nyilvánvalóan a jelöltnek az adott kérdés vizsgálatával kapcsolatos eredeti személyes motivációt hivatottak illusztrálni, azonban véleményem szerint egy ilyen disszertációban ezek irreleváns kérdések, különösen akkor, ha a használt módszerek ismertetése lényegében elmaradt. Így például a 3. fő fejezet, azaz a három nagy kutatási irány közül az első alapját képező fordított (Reverse) Monte Carlo (RMC) módszer teljes ismertetése a függelékben egy alig fél oldalas szócikkre korlátozódik, míg a jelölt által munkatársaival együtt kifejlesztett fordított (Reverse) molekuláris dinamika (RMD) módszerről semmit nem tudunk meg a dolgozatból. Mindez természetesen megnehezíti a

bíráló dolgát is, különösen azoknál a részeknél – nálam ez elsősorban az 5. illetve részben a 4. fő fejezetet jelentette – ahol a disszertációból kimaradt módszertani illetve elméleti háttérismeretek nincsenek kellő mélységben meg.

Hasonló kifogásaim vannak a fő fejezetek szerkesztésével kapcsolatosan is. Az adott területen végzett munka és kapott eredmények sokszor mindenféle tudományos ívet nélkülözve vannak egymás után téve, azt a benyomást keltve a bírálóban, mintha a jelölt egyszerűen összefűzte volna a fejezet alapjául szolgáló publikációk összefoglalásait.

A disszertációnak mint tudományos műnek az értékelésénél az alapvető kérdés az, hogy mit várunk el egy MTA doktori értekezéstől. Ha az a cél hogy a disszertáció megmutassa, hogy a jelölt mögött álló szakmai életút megfelel az MTA Doktora címnél elvártaknak, akkor ezt a disszertáció – a mögöttes publikációkkal együtt – teljesíti. Részemről ez az elvárás a rövidített disszertációk esetén, ahol a mögöttes publikációk a disszertáció függelékét képezik. Ha a személyes sztorikat elhagyjuk, jelen disszertáció leginkább egy rövidített értekezésre emlékeztet. Saját véleményem szerint egy MTA Doktori értekezésnek a fentiekén túl önálló tudományos műnek is kellene lennie, amely tankönyvszerű rigorózsággal foglalja össze a bemutatott munka elméleti háttérét, és amit a tanszék leendő fiatal kutatói a következő évtizedekben tankönyv jelleggel is forgathatnak. Ennek az elvárásnak a jelen dolgozat meg sem próbál eleget tenni.

A dolgozat érdemi, eredményeket bemutató része három fő fejezetre tagolódik. A 3. fejezetben az RMC módszerrel kapcsolatos korai kutatásait taglalja a jelölt. Ezek a munkák egymással csak lazán függenek össze. Szó van itt a módszer egyértelműségének kérdéséről, a fluktuációk szerepéről, a szerkezeti vagy párkorrelációs függvényt érdemes-e illeszteni kérdéséről, részecskék belső szerkezetének meghatározásáról. E sokféleség azt a tényt tükrözi, hogy e vizsgálatok a jelölt pályájának korai szakaszában, közvetlenül a doktori fokozat megszerzését követően születtek – a jelöltnek ekkor még nem volt önálló kutatói profilja, a munkák inkább az akkoriban a jelölt körül, vele laza szakmai együttműködésben dolgozó idősebb kollégák akkori érdeklődését mutatják. A jelölthöz hasonlóan abban az időben az RMC módszer is a helyét kereste. Ezt mára megtalálta, a diffrakciós kísérleteket végző kutatók adatfeldolgozó, kísérletkiértékelő arzenáljába került több más hasonló célt szolgáló módszerrel (Empirical Potential Monte Carlo) együtt, így mára e munkák jelentősége csekély.

A 4. fejezet a statisztikus mechanikai inverz probléma, azaz a párpotenciálok párkorrelációs függvényekből való származtatásának kérdése körül végzett saját kutatásokat foglalja össze. (Az említett függvények fordított kapcsolata, azaz párpotenciálokból párkorrelációs függvények származtatása triviális, ezt bármely klasszikus szimuláció megteszi.) A probléma jelentőségét emeli, hogy a párpotenciálokkal ellentétben a párkorrelációs függvények kísérletileg hozzáférhető mennyiségek. A fejezetből némiképp kilóg az első alfejezet, mely a jelölt korai éveiben végzett, párpotenciál fejlesztésével kapcsolatos munkáit összegzi. Ezek a munkák, bár nyilvánvalóan a jelölt tudományos munkásságának részei, még mindig nem mutatják a későbbi, önálló kutatói karakterét. E saját kutatói profil első markáns megnyilvánulása a fentebb említett inverz probléma megoldására kidolgozott négyféle eljárás. Az ezen munkákat ismertető fejezetek véleményem szerint a dolgozat legfontosabb részei közé tartoznak. Érdekes rész a kinetikai Monte Carlo módszerrel kapcsolatos fejezet is, kár hogy magáról a módszerről itt sem ír lényegében semmit a szerző. Szép és elegáns munka itt a karbamid kristályosodásával kapcsolatos vizsgálat.

Az 5. fejezet a jelölt kemometriai jellegű kutatásait foglalja össze. Itt teljesedik ki a jelölt önálló kutatói arcéle, iskolateremtő tevékenysége is lényegében ehhez a területhez köthető. A fejezetben bemutatja a szeriálással, azaz adatmátrixokban sorok és/vagy oszlopok célzott felcserélésével, az egyes adatsorokban esetlegesen fellelhető trendek meglétének vizsgálatával, adott modellek tesztkészletének kiválasztási tartományával (allokáció) illetve validációs stratégiák vizsgálatával kapcsolatos kutatásait. Különösen elegánsnak találtam az allokáció fontosságát illusztráló, egy tripeptid Ramachandran diagramjával kapcsolatos munkáját, még ha nyilvánvalóan nem is ez az itt bemutatott eredmény legfontosabbika.

Összefoglalva, a Tóth Gergely disszertációja mögötti tudományos teljesítmény eleget tesz az MTA doktora fokozat megszerzésével szemben támasztott követelményeknek, míg maga a disszertáció egy személyes hangvételű szakmai fejlődéstörténet és nem egy tankönyvszerű szakmai összefoglalása a használt módszereknek, elméleteknek és eredményeknek. Mivel az MTA Doktori disszertációval szembeni elvárásokkal kapcsolatos vitát nem az én tisztem itt és most eldöntenem, így az

értekezéssel kapcsolatos fenntartásaim ellenére a dolgozat téziseit új tudományos eredményeknek elfogadom, és az alább ismertetendő kérdéseimre adott válaszoktól függetlenül javaslom az értekezés nyilvános vitára bocsátását, sikeres védelem esetén pedig az MTA Doktora fokozat odaítélését.

Az értekezéshez az alábbi kérdéseket és megjegyzéseket fűzöm.

- A jelöltnek az RMC módszerrel kapcsolatos egyik fontos megállapítása, hogy az általa generált konfigurációk nem képezik a kanonikus sokaság reprezentatív mintáját. Az ezzel kapcsolatos eredményeket és érveket teljes mértékben elfogadva meg kell jegyezmem, hogy nem is szánták annak (szándéka szerint az RMC módszer egy, a korábbiaknál teljesebb és alaposabb kiértékelési módszere bizonyos, elsősorban diffrakciós kísérleteknek), illetve tudomásom szerint senki nem is állította, hogy az volna. (Más kérdés hogy abban az időben egyes szerzők implicit módon, vélhetőleg a kérdést maguknak fel sem téve úgy kezelték a módszert mintha az általa generált konfigurációk valamely, a kanonikussal legalábbis egyenértékű sokaságot alkotnának).

- A 25. oldalon, a micellák RMC vizsgálatánál írja a jelölt: "...egy...*hidrofób belső mag* és az azt körülvevő *hidrofil lelógó molekula részekből* áll a micella.", majd pár sorral lejjebb: "... a *hidrofil magot* egy nagy gömb, míg...*hidrofób láncot* egy-egy kisebb gömb tartalmazta." A két egymásnak ellentmondó mondat közül melyik a helyes? Az első leírás felel meg a micellák tipikus szerkezetének (hidrofób mag, hidrofil felszín), míg a későbbi leírás arra utal mintha itt a második változatról lenne szó. Vagy ezek inverz micellák?

- A 4.2 fejezetben (Potenciál meghatározás „on fly” a 90-es években) részletesen ír a korabeli, szemiempirikus kvantumkémiai módszerekkel (MNDO, és változatai) végzett QM6MM számításairól. Milyen pluszt adtak abban az időben ezek az akkor kb. 50 részecskét kezelni tudó módszerek a klasszikus (akkoriban nagyságrendileg 1000 molekulával elbíró) szimulációs módszerekhez képest?

- A 4.10. ábra (44. oldal) jobb oldalán az ábrafelirat szerint pátkorrelációs függvények összehasonlítása szerepel, míg a valóságban itt is pátpotenciálokat láthatunk. Ez nyilván

egy egyszerű hiba, kérem azonban a jelöltet, hogy a teljesség kedvéért válaszában mutassa meg az ide eredetileg szánt, párkorrelációs függvényeket összehasonlító ábrát.

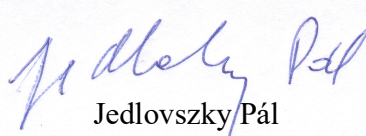
- A "potential matching" módszert a BJH (flexibilis) vízmodellre alkalmazva a módszer lényegében kudarcot vallott (51. oldal, "...a potenciál nem volt összehasonlítható az eredeti... intermolekuláris kölcsönhatással"). Korábban arról írt a jelölt, hogy Lennard-Jones részecskékre viszont működött a módszer. Mi a jelölt véleménye, víz esetén a sikertelenség oka a Lennard-Jones és Coulomb kölcsönhatások együttes jelenléte, vagy a molekulák nem gömbszerű alakja, esetleg a több parciális párpotenciál jelenléte volt? Mire számítana alkáli halogenid olvadékok esetében?

- A fenti kérdés folytatásaként kérem, hogy ismét csak a teljesség kedvéért válaszában, ha lehet, mutassa be, hogyan egyezik a BJH vízmodellre a PM módszerrel kapott három parciális párpotenciál (4.16. ábra felső panelje) az eredeti modellével (az ábra az eredeti potenciálokat nem, csak azok Lennard-Jones és Coulomb járulékát külön-külön mutatja).

- Az 5.1 fejezetben (69. oldal) azt írja a jelölt hogy a lokális hasonlóságot az 5.1 egyenlettel definiálta, majd három sorral lejjebb az áll hogy "a lokális hasonlóságot a következő egyenletekkel definiáltuk", mely állítás most az 5.2 és 5.3 egyenletekre utal. Melyik egyenlet definiálja valójában a lokális hasonlóságot, és mit ír(nak) le a további egyenlet(ek)?

- A Neumann-féle trend teszt leírásánál (77. oldal) a H_0 hipotézist a leírás szerint akkor vetjük el, ha $d_N < d_{Ncrit}$, ez utóbbi értékre viszont csak egy példát kapunk ($N=30$, $\alpha=0.05$ esetén $d_{Ncrit}=1.418$). Honnan tudható (hogyan adható meg) ez a kritikus érték általános esetben?

Eger, 2026 január 29.



Jedlovsky Pál
az MTA Doktora