

## A bírálóbizottság értékelése

Tóth Gergely *Adatok értelmezése szimulációkkal és a többváltozós statisztika eszközeivel a kémiában* című akadémiai doktori értekezésében kémiai (különösen is fizikai-kémiai és analitikai kémiai) módszerek számítógépes adatelemzésével foglalkozott. Az értekezésben bemutatott kutatómunka három fő tématerületre osztható, melyet a közös eszköztár, a számítógépes adatfeldolgozás köt össze. A szerteágazó tématerület okán a jelölt az elért eredményeit 20 tézispontban foglalta össze.

Az értekezés első része a fluidumok diffrakciós méréseiből nyert adatok feldolgozásához kapcsolódik, melyben a jelölt az itt alkalmazott adatelemző módszert, a fordított Monte-Carlo (RMC) szimulációt vizsgálta több aspektusból. Célja az RMC módszerrel kapott konfigurációk, ill. az ezeket leíró szerkezettel kapcsolatos függvények egyértelműségének, valamint a valós viszonyok reprezentálásának mértékének vizsgálata. Az elért eredményeket az 1-3 tézispontok tartalmazzák, melyekben a jelölt megállapította, hogy az RMC módszer nem megfelelően reprodukálja a különböző szerkezeti, termodinamikai és dinamikai tulajdonságok átlagát és fluktuációját, valamint, hogy a módszer a kanonikus sokaságtól eltérően mintavételezi a konfigurációs teret. Kiemelten fontos eredmény az a megállapítás, hogy a párkorrelációs függvény meghatározza saját fluktuációját.

A második rész középpontjában a statisztikus mechanika inverz problémája áll, azaz, hogy milyen módszerekkel lehet makroszkopikus jellemzők (pl. termodinamikai és szerkezeti adatok) ismeretében kölcsönhatási potenciálokat meghatározni. Ezen a tématerületen belül a jelölt úttörő munkát végzett, négy, különböző megoldásokon (kvantumkémiai számításokon, szemiempirikus módszereken, fordított Monte Carlo szimuláción és a mesterséges ideghálókön) alapuló adatelemző módszert fejlesztett ki. Ehhez a tématerülethez a 4-11 tézispontok tartoznak.

Az értekezés harmadik része a kemometriához tartozik, ahol a jelölt kutatásai az adatmátrixok átrendezési lehetőségei és a modellvalidálás köré csoportosulnak. Az itt elért eredményeket a 12-20 tézispontok tartalmazzák. Kiemelt jelentőségű eredmény a szeriális módszere, azaz az adatmátrixok sorainak és oszlopainak célzott átrendezése úgy, hogy abból elsőre nem látható, rejtett információk táruljanak fel. A modellvalidálás témakörében végzett kutatási jelentős mértékben hozzájárulnak a validálás során használt tesztkészlet, valamint a validációs paraméterek megfelelő kiválasztásához.

A doktori értekezés több évtizedes, szerteágazó kutatói pálya összegzése, mely egyértelműen bemutatja a jelölt adatelemző módszerek világában való jártasságát és kritikai gondolkodásra való képességét. Noha az értekezés szerkesztése, hangvétele eltér a megszokottól, és ez megnehezítette a bírálók munkáját, az elért eredmények alapján Tóth Gergely valamennyi tézispontja új tudományos eredmény, melyek alapján a pályázó megfelel az MTA doktori elvárásoknak.