

Vízmodellek összehasonlítása szerkezeti információtartalom szempontjából

Soczó Bálint

BME Fizika BSc, II. évfolyam

Témavezetők:

Némethné Pethes Ildikó

tudományos munkatárs, Wigner FK

Dr. Asbóth János

Egyetemi docens

BME TTK, Elméleti Fizika Tanszék

és tudományos főmunkatárs, Wigner FK

Budapest

2023

Tartalomjegyzék

1. Bevezetés	2
1.1. Célkitűzés	2
1.2. A molekulák vizsgálatának történelme	3
1.3. A molekuladinamika elméleti alapjai	4
1.4. A molekuladinamika mai állása, mik a megoldatlan problémák a témakörben?	7
1.5. A vízmodellek fejlődése a molekuladinamikában	8
2. Módszertan	8
2.1. GROMACS	8
2.2. Egy MD szimuláció szakaszai	9
2.3. Algoritmus optimalizálások	10
2.4. A szimulált eredmények szerkezetvizsgálati feldolgozása	11
2.4.1. A modellek szerkezeti információtartalma	11
2.4.2. Szerkezetvizsgálati alapfogalmak	11
2.5. Vizsgált vízmodellek	13
2.5.1. 3 site-os modellek	14
2.5.2. 4 site-os modellek	14
2.5.3. 5 és több site-os modellek	15
3. Munkafolyamat	18
3.1. Induló konfigurációk létrehozása	18
3.2. A szimuláció szakaszai	19
3.3. A szimulációk további paraméterei	20
3.4. A szimulációk vezérlése, az eredmények kiértékelésének folyamata	20
4. Eredmények	21
4.1. A modellek illeszkedése a kísérleti adathoz	21
4.2. A modellek számításigényének összehasonlítása	23
4.3. A vízmodellek szerkezetének elemzése	24
4.3.1. SPC, SPC/E, SPC-FW	25
4.3.2. TIP3P, TIP3-FB	27
4.3.3. OPC, OPC3	29

4.3.4. TIP4P alapú modellek	30
4.3.5. Polarizálható modellek	31
4.3.6. TIP5P, TIP5P-EW, TIP5P-2018, TIP7P	32
5. Összefoglalás	33
5.1. Kitekintés	35

Absztrakt

A víz a minket körülvevő legkézenfekvőbb anyag, számos meglepő tulajdonsággal. A víz elméleti kutatása során még nem sikerült olyan modellt alkotni, mely a víz kísérletileg ismert jellemzőit pontosan képes lenne reprodukálni. Molekuladinamikai szimulációkhoz több mint 50 kölcsönhatási potenciálmodell jelent meg az elmúlt évtizedekben, melyek egy vagy több specifikus célt figyelembe véve készültek. Kutatásom célja ezen modellek vizsgálata, összehasonlítása volt, szerkezeti információ szempontjából. Munkám során klasszikus molekuladinamikai szimulációkat készítettem 2-4 ezer vízmolekulából álló rendszereken. Számos különböző Lennard-Jones típusú effektív párpotenciált vizsgáltam meg. A kapott részecskekonfigurációkból meghatároztam a parciális radiális eloszlásfüggvényeket. Ezekből neutron és röntgendiffrakciós szerkezeti függvényeket számítottam ki, melyeket az irodalomban található [1] [2] kísérleti görbékkel vetettem össze. A vizsgált modellek között mind friss, mind régebben megjelent, de máig népszerű potenciálok szerepeltek. Megállapítottam, hogy a szerkezet leírását tekintve nem fedezhető fel egyértelmű javuló tendencia. Az újabb modellek közül a TIP5P-2018 [3] szerepelt a legjobban, amely körülbelül hasonló pontossággal adja meg a szerkezetet, mint a már 18 éves TIP4P-2005-ös [4] modell.

1. Bevezetés

1.1. Célkitűzés

Dolgozatom célja különböző vízmodellek összehasonlítása. Megvizsgáltam mind újabb, mind a leggyakrabban használt korábbi vízmodelleket, hogy mennyire egyeznek szerkezeti szempontból a kísérleti elvárásainkkal. A régi modellek, mint viszonyítási pontok is hasznosak. Az újabb modellek többnyire bonyolultabbak, hiszen az alkotóik kihasználják a megnövekedett számítási kapacitást. A kérdés az, hogy a nagyobb komplexitás szerkezeti szempontból is pontosabb képet ad-e. Erre keresem a választ.

A víz, bár az egyik legegyszerűbb molekula, még sincs arról pontos modellünk, hogy viselkedik mikroszinten, más vízmolekulák környezetében. Arról még kevesebbet tudunk, amikor a vízmolekula más molekulákkal is interakcióba lép. A víz esetében a szerkezet vizsgálata különösen fontos, hiszen az egyedi viselkedése nagyrészt az intermolekuláris hidrogénkötéseknek tudható be. Így tehát feltételezhető, hogy egy olyan modell, ami pontosabban leírja a szerkezetet, azaz a molekulák közötti kötések távolságát, gyakoriságát, valamint a kötések szögeloszlását, az más szempontból is pontosabb lesz, mint egy olyan modell, ami a szerkezet figyelembevétele nélkül készült. A vizet abból a szempontból is kritikus megérteni, mert az élet elengedhetetlen építőeleme, mindenütt megtalálható. Így, ha a fehérjéket, nukleinsavakat, vagy akár csak a cukrokat jobban meg akarjuk érteni, elengedhetetlen, hogy az őket feloldó vagy körbevevő vízről is pontos képünk legyen, főleg szobahőmérsékleten, ahol a biológiai folyamatok túlnyomó többsége lezajlik.

1.2. A molekulák vizsgálatának történelme

A molekulákat nincs száz éve, hogy mai formájukban ismerjük, nem meglepő tehát, hogy még ma is sok hozzájuk kapcsolódó jelenség van, amit nem tudunk megmagyarázni. A molekulákkal foglalkozó tudományág gyökerei az 1800-as évek elejére nyúlnak vissza, amikor Dalton atomelmélete megalapozta az anyag részecskékből való felépítésének gondolatát. Azonban a molekuláris szintű vizsgálatok korlátozottak voltak, mivel nem voltak elérhetőek kellő pontosságú kísérleti eszközök és technikák.

Az első modellek ezért mérések helyett többnyire elméleti alapokra támaszkodtak. Boltzmann és Maxwell a 19. század végén kidolgozták a gázok viselkedését leíró statisztikus mechanikát, amely az egyes részecskék mozgását a gáz hőmérsékletével és a nyomásával kapcsolta össze. Azonban a molekulák mozgásának részletes megértéséhez és modellezéséhez új eszközökre volt szükség, főleg, ha nem gáz halmazállapotú anyagokat is vizsgálni akartak. Hétköznapi körülmények között, gázok esetében a molekulák egymással ritkán és röviden lépnek interakcióba. Ezért ezen kölcsönhatás figyelmen kívül hagyásával is viszonylag pontos modellt kapunk.

Ez azonban folyadékoknál és szilárd anyagoknál már nem így van, mert az anyagok tulajdonságai, viselkedése nagyon jelentős részben az intermolekuláris kölcsönhatásoknak tudható be. A van der Waals (vdW) erők olyan gyenge, rövid hatótávolságú kölcsönhatások, amelyek két molekula között jönnek létre az elektronfelhők ideiglenes polarizációjából. Ezek az erők hozzájárulnak az anyagok tulajdonságaihoz, mint például a kompresszibilitáshoz vagy viszkozitáshoz.

Attól, hogy jobban megértettük az egyes molekulák felépítését és a közöttük fellépő kölcsönhatásokat, még hosszú út áll előttünk az anyagok viselkedésének feltárásához. Hiszen sok anyag felfogható

egy hatalmas molekuláris rendszerként, amely molekulái folyamatosan kölcsönhatnak egymással és a környezetükkel. Sok test közötti kölcsönhatásokat és az ebből eredő mozgást numerikus módszerekkel lehet kiszámítani, de ez már kisebb rendszerek és csekély időtartamok mellett is nagyon számításigényes.

Nem meglepő, hogy a molekulák ilyen, kvantitatív vizsgálata a számítógépekkel párhuzamosan fejlődött. Amikor még nem volt lehetőség annyi számítás elvégzésére, akkor még nem mozgásegyenleteket oldottak meg. Ehelyett véletlenszerűen mozgatták az atomokat, és minimalizálni igyekeztek egy függvényt, ami többnyire valamilyen egyszerűsített potenciálfüggvény volt. Ezek a Monte-Carlo szimulációk.

A számítási kapacitás növekedésével lehetővé vált a molekuladinamikai (továbbiakban MD) szimulációk futtatása. A szimulációk során a részecskék az atomok közötti kölcsönhatásokat kiszámítva a Newtoni mozgásegyenleteknek eleget téve mozognak. Így már nem kell véletlen mozgatásokra hagyatkozva vizsgálni egy rendszert, hanem az magától "mozog be" az egyensúlyi helyzetbe, ahol a potenciális energia minimális. Mivel a Monte-Carlo szimulációkkal szemben az MD szimulációk nem statikusak, lehetővé válik dinamikus jelenségek, mint például a fázisátmenetek vizsgálata is.

1.3. A molekuladinamika elméleti alapjai

Az egyes atomokra ható szabaderők a potenciális energia r szerinti deriválásával határozhatók meg: $F_i = -\frac{\partial V}{\partial r_i}$. A szimuláció során a részecskék mozgását ezek az erők és a kényszerek határozzák meg. Ezért az, hogy az MD szimulációk milyen pontos képet adnak a valós jelenségekről, nagyban múlik azon, hogy milyen pontos potenciálmodellt használunk. A legegyszerűbb potenciálmodellek az effektív párpotenciálok, amelyek csak két részecske közötti kölcsönhatásból származó potenciált vesznek figyelembe. Bár kémiai folyamatok, mint egyes oldódások részleteit nem olyan jól írják le, sokkal könnyebb kiszámítani őket. De, amíg egy rendszer nyugalmi viselkedését szeretnénk vizsgálni, a párpotenciálok kellően pontosak lehetnek.

A 6-N típusú Lennard-Jones (LJ) potenciál (1. ábra) a részecskék között fellépő van der Waals kölcsönhatást modellezi. A pozitív tagja egy "puha-taszító" kölcsönhatás, a negatív, hatodfokú tagja pedig egy vonzó kölcsönhatás, ami a London-diszperziós kölcsönhatást írja le. A vonzó tagnak van fizikai alapja, a másik tag kitevője viszont azért gyakran 12, hogy az a diszperziós tag négyzete legyen, így könnyebb legyen kiszámítani. A 6-12-es változatot (1) írja le:

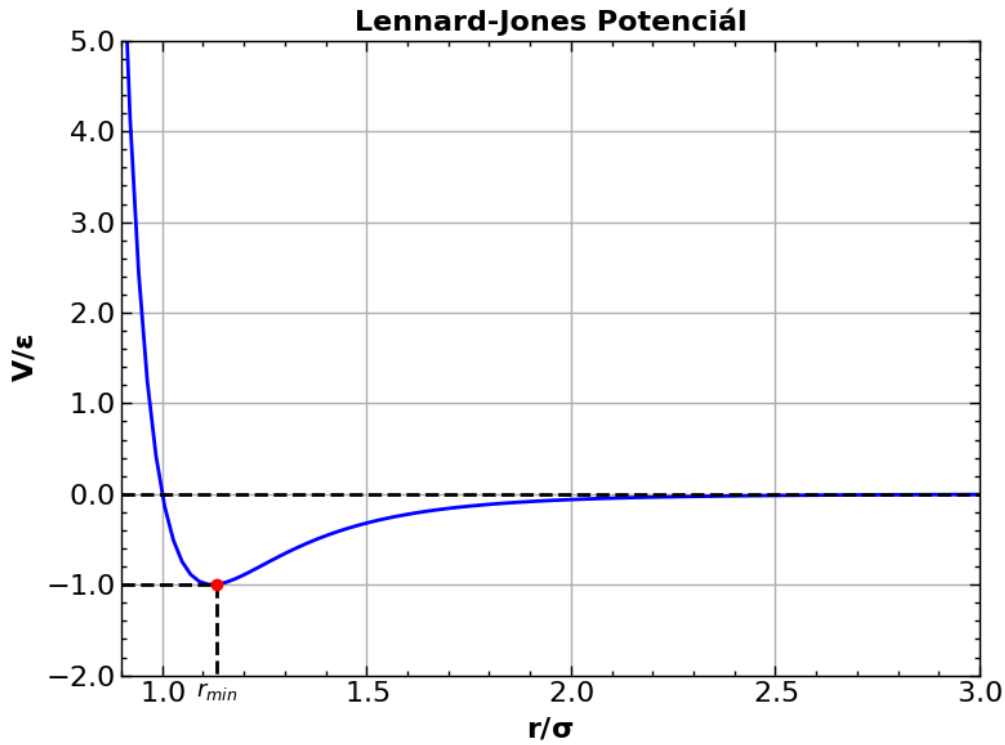
$$V_{LJ}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (1)$$

Ahol r a központi atomtól mért távolság, σ az energiaminimum helyét befolyásoló paraméter, ε pedig a potenciál nagyságát skálázó paraméter. A potenciál értéke $r_{\min} = 2^{\frac{1}{6}}\sigma$ helyen $-\varepsilon$, ez a potenciálgödör mélysége.

A legtöbb vízmodellnél csak az oxigénatomnak vannak nem nulla értékű LJ paraméterei, de, ha két különböző atom között számítjuk ki a kölcsönhatást, azt valamilyen kombinációs szabály alapján tehetjük meg. Én a munkám során a Lorenz-Berthelot szabályt ((2) és (3)) követtem.

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \quad (2)$$

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \quad (3)$$



1. ábra. A Lennard-Jones potenciált szemléltető ábra.

Fontos megjegyezni, hogy kisebb molekulák esetében a molekula atomjai közötti van der Waals kölcsönhatást nem vesszük figyelembe, de nagyobb molekuláknál fontos szerepet játszik a molekula atomjai közötti van der Waals kölcsönhatás is.

A modellek sokszínűsége szemléltetendő érdemes még megemlíteni egy, a dolgozatomban nem tárgyalt másik effektív párpotenciált, a Buckingham potenciált, ami nem polinomiális, hanem exponenciális:

$$V_{\text{Buckingham}}(r) = A \exp\left(\frac{-r}{\rho}\right) - \frac{C}{r^6} \quad (4)$$

A diszperziós kölcsönhatáson kívül még nagy szerepet játszik a rendszer potenciális energiájában a töltött részecskék között fellépő elektrosztatikus kölcsönhatás is:

$$V_{\text{Coulomb}}(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r} \quad (5)$$

Ahol q_i és q_j a két részecske töltése, r pedig az őket összekötő szakasz távolsága.

Mind az elektrosztatikus, mind a diszperziós kölcsönhatás viszonylag gyorsan lecseng, ahogy nő a távolság, így nem indokolt a távoli molekulák közötti kölcsönhatások kiszámítása. Ez a szimuláció pontosságán nem változtat számottevően, de nagyobb rendszerek esetén, számítási költségben több nagyságrendnyi különbség is lehet. Ez a két energia alkotja a nem-kötő (non-bonded) energiát, amivel az MD szimulációk során a rendszert jellemezzük:

$$V_{\text{non-bonded}}(r) = V_{\text{LJ}}(r) + V_{\text{Coulomb}}(r) \quad (6)$$

Vannak olyan modellek, amelyek nem csak a molekulák közötti kölcsönhatásokat, hanem a molekulákat egyben tartó kölcsönhatási energiákat is figyelembe veszik. Ekkor a molekulákat nem merev kényszerek határozzák meg, hanem potenciálok. Kötő (bonded) energiából hármat különböztetünk meg (7) (8) (9):

$$U_{\text{bonded}} = \frac{1}{2} k_{\text{bond}} (r - r_{\text{eq}})^2 \quad (7)$$

Ahol k_{bond} a kötéshosszra jellemző rugóállandó, r_{eq} pedig az egyensúlyi helyzet távolsága.

$$U_{\text{bend}} = \frac{1}{2} k_{\text{angle}} (\theta - \theta_{\text{eq}})^2 \quad (8)$$

Ahol k_{angle} a kötésszögre jellemző rugóállandó, θ_{eq} pedig az egyensúlyi helyzet szöge. A harmadik kötő kölcsönhatás a diéderes (dihedrális) szögekre bevezetett kényszer, ez azonban vízmodelleknél nem használatos, így itt csak említés szintjén szerepel.

$$U_{\text{dihedral}} = \frac{1}{2} k_{\text{dihedral}} (\phi_{ijkl} - \phi_{\text{eq}})^2 \quad (9)$$

A kötő kölcsönhatások több különböző függvénnyel is megvalósíthatóak, de mivel az általam vizsgált vízmodellek a harmonikus potenciált használják, a többi bemutatása kívül esik a dolgozatom hatókörén. Az alapján, hogy a modell figyelembe veszi-e a molekula kötő atomjai közti kölcsönhatási energiát vagy merev kényszerekkel tarja egyben a molekulát, megkülönböztetünk merev és rugalmas modelleket. A munkám során elsősorban merev modellekkel dolgoztam, ez alól kivételt jelent a rugalmas SPC-FW [5] modell, amelynek, mint azt későbbiekben látni fogjuk, vannak jellegzetes szerkezeti tulajdonságai.

1.4. A molekuladinamika mai állása, mik a megoldatlan problémák a témakörben?

A molekulák mozgását tehát teljes egészében a potenciálfüggvény határozza meg. A pontosabb potenciálfüggvények kvantummechanikai alapúak, tehát a szimulálásukkor a Schrödinger-egyenletet kell megoldani. Ezek jelentősen számításigényesebbek, mint a klasszikus modellek. A szimulációkat legjobban korlátozó tényező a számítási kapacitás. Aki ilyen szimulációkat futtat, egyensúlyt kell találnia az időlépés, a szimulált részecskék száma, a potenciálfüggvény komplexitása, valamint a molekulák pontos modellezése között. Egy egyszerűbb szimuláció képes kis időlépésekkel, sok molekulát szimulálni belátható időn belül, de nem feltétlenül tükrözi a valóságot. Egy kvantummechanikai alapú, elektronhéjakat szimuláló modell viszont túl nagy számítási igényű ahhoz, hogy néhány száz molekulánál többet lehessen egyszerre szimulálni, és még a kisebb részecskeszámú futások is jelentős időt vesznek igénybe.

A klasszikus modellek egyik előnye, hogy a potenciálfüggvények lehetnek kvantummechanikai alapúak, még ha a szimuláció maga nem is az. A potenciálparaméterek értékeit ezért szinte mindig kvantummechanikai számításokkal határozzák meg, így a számítási igény nem változik, de a szimuláció pontosabb lesz. A klasszikus modellek emellett rugalmasabbak és könnyebben paraméterezhetők, ezért sokkal produktívabban lehet velük dolgozni. Persze, ha például bonyolultabb kémiai reakciókat szeretnénk pontosan modellezni, amelyek elektronátmenettel járnak, akkor kénytelenek vagyunk kvantumos szimulációkat futtatni.

Egyes molekulák vizsgálatához sok különböző modellt készítettek az évek során. A modelleket sok esetben biológiai és gyógyszerkutatásokhoz fejlesztették ki, ahol nem a szobahőmérsékleti, atomi szintű szerkezet pontos reprodukálása volt a fő szempont. Az, hogy a modell egy-egy tulajdonság szempontjából minél pontosabban tükrözze a valóságot, sokszor csak a többi tulajdonság rovására érhető el. Tehát a modellek egyike sem általános modell, hanem egy vagy több tulajdonság pontos reprodukálásához lett paraméterezve. Vannak olyan modellek, amelyek jól visszaadnak termodinamikai tulajdonságokat, mint például a hőtágulási együtthatót vagy az olvadáspontot. Vannak más modellek, amelyek elektrosztatikai szempontból modellezik jól a valós molekulákat. Ezen tulajdonságokra hangolás közben viszont a szerkezet nem feltétlen szempont, így bár olyan modelleket kapunk, amelyek jól megjósolják, hogyan viselkedik a molekula bizonyos szempontokból, a modell mégsem tükrözi szerkezeti szempontból a valóságot.

1.5. A vízmodellek fejlődése a molekuladinamikában

A víz az egyik első folyadék volt, amit elkezdtek MD szimulációkkal vizsgálni. Az egyik első vízmodell a Mercedes Benz [6] volt. Arról kapta a nevét, hogy a modellben éppúgy, mint az autómárka logójában, a vízmolekula három atomja 120° -os szöget zár be egymással. Ez azonban csak 2 dimenzióban írja le valamelyest jól a valóságot, 3D-ben már nem működik.

Ezzel egyidejűleg létrejöttek pontosabb modellek is, amelyek a víz 3 atomján kívül, még két nemkötő elektronpár (lone-pair) részecskét is tartalmaztak. Bár ezen modellek közül az ST2 [7] a víz néhány tulajdonságát egész pontosan visszaadta, az 1970-es években még nem volt akkora számítási kapacitás, hogy 100-200 molekulánál nagyobb rendszert szimuláljanak.

Több részecskéből álló rendszerek vizsgálatához szükség volt egyszerűbb modellekre. Ezt a célt szolgálta Berendsen SPC [8] (simple point charge) modellje, ami már csak a két hidrogénatomból és az oxigénatomból áll. A modellt a mai napig gyakran használják egyszerűsége miatt, egyik újraparaméterezett változata, a TIP3P [9] is nagyon elterjedt. Ahogy fejlődtek a számítógépek, lehetővé vált több részecskéből álló modellek, mint a TIP5P [10], flexibilis modellek, mint a SPC/FW [5], és polarizálható modellek, mint a SWM4-NDP [11] szimulálása is.

2. Módszertan

2.1. GROMACS

A GROMACS [12] az egyik legnépszerűbb molekuladinamikai szimulációs szoftver. Jól karbantartott, és számos hasznos eszközzel rendelkezik. A program talán legnagyobb előnye, hogy nagyon hatékonyan használja ki a CPU, sőt még a GPU erőforrásokat is. Emiatt komplexebb vagy több molekulából álló rendszerek szimulálásához is kiváló. Kihhasználja a szimulációs számítások párhuzamosíthatóságát is, ami tovább növeli a teljesítményét.

A GROMACS sok potenciálmodellt (force field) előre beépítve tartalmaz, így a régebbi modellek egy része eleve rendelkezésemre állt. A program további nagy előnye, hogy sok beépített segédprogramot is tartalmaz, amelyek a futás előkészítéséhez, és az eredmények kiértékeléséhez is nagyon hasznosak, és mivel a GROMACS fájlformátumokat figyelembe véve készültek, nagyon hatékonyak is. Emellett mind a főprogram, mind a segédprogramok használata jól dokumentált, ami sokat segített a program használatának elsajátításában.

2.2. Egy MD szimuláció szakaszai

Egy MD szimulációnak 3 fő szakasza van: energiaminimalizálás (em), ekvilibráls (eq1) és maga a szimuláció, a produkciós futás. Ezeket a szakaszokat tovább szokták finomítani, hogy a rendszer biztosan stabil állapotban kezdje meg a produkciós futást, azaz a dinamikai jellemzők eloszlása konstans legyen. Ha minden szimulált kényszert egyszerre vezetnénk be, akkor a tapasztalat szerint, nem lehetne garantálni, hogy a rendszer képes lenne egyidejűleg megfelelni nekik.

Az energiaminimalizálás általában kevesebb időt vesz igénybe, mint a többi szakasz. A molekuláknak ilyenkor még nincs sebessége, hanem valamilyen szélsőértékkereső algoritmus alapján vannak mozgatva. Amint a rendszer elérte a maximum lépések számát, vagy a részecskék között ható erő abszolútértékének maximuma a felhasználó által megadott értéknél kisebbé válik, az em véget ér.

Az em szakaszban a Steepest Descent [13] algoritmus mozgatja a molekulákat. A szakasz célja eltüntetni a molekulák között fellépő legerősebb feszültségeket, amelyek akkor vannak jelen, ha a molekulák kezdetben túl közel vannak egymáshoz. Kisebb r értéknél a LJ-potenciál taszító tagja nagyon gyorsan növekszik, így hatalmas erő lép fel két közeli molekula között. Az algoritmus egyszerűségének köszönhető, hogy amennyiben a potenciálmodell fizikai szempontból értelmes, képes szinte bármilyen rendszerben a molekulák között fellépő erőt jelentősen lecsökkenteni. Ezért nem kell különböző algoritmusokkal kísérletezni, ami nagyban növeli az automatizálhatóságot.

A részecskék véletlenszerű mozgatása megfelelő geometriai kényszer nélkül valótlán konfigurációhoz vezethet. Ezért már az első szakasztól kezdve szükség van egy ilyen kényszerre. Víz esetében a Settle [14] algoritmust használjuk, ami további kényszerekkel együtt növeli a molekula stabilitását.

A következő szakasz az ekvilibráls, ekkor a részecskékhez sebességértékeket rendelünk, áttérünk egy statikus rendszerből egy dinamikusba. Mivel itt már mozgásegyenletet oldunk meg, az itt használt differenciálegyenlet-megoldó leapfrog algoritmus (integrátor) idő alapú. Az algoritmus diszkrét időpontokban végzi el az objektumok pozícióinak és sebességeinek frissítését, hogy megoldja a Newton-féle mozgásegyenleteket. Magukat az időlépéseket tovább osztja két részre, és a következő pozíciót a sebesség egy fél időlépésnyivel előrehaladottabb értékével számítja ki (10). Ez pontosabb, mintha egyszerűen a pillanatnyi sebességet extrapolálnánk.

$$r_{i+1} = r_i + v_{i+\frac{1}{2}} * \Delta t \quad (10)$$

Az ekvilibráls célja a rendszert a szimuláció során használni kívánt kényszerek fokozatos bevezetésével egyensúlyi helyzetbe hozni, hogy a produkciós futásról már tetszőleges ideig lehessen adatot gyűjteni.

A részecskék sebességének generálása egy megadott hőmérsékleti értékhez igazodik. Miközben a részecskék mozognak a hőmérséklet egy ún. termosztáttal állandó értéken tartjuk. A termosztát működését nagyban befolyásolja a hőmérsékletcsatolási állandó értéke, amely azt határozza meg, hogy milyen gyorsan képes reagálni a rendszer a hőmérséklet-fluktuációra. A csatolási állandó idő mértékegységű, ha az értéke túl nagy, a termosztát lassabban képes reagálni a rendszer hőmérsékletváltozásaira. Cserébe viszont kevesebb számítást kell végezni, és a rendszer is szabadabban tud mozogni, ezért csökken a hőmérséklet-korrekciókból származó zaj. Az optimális csatolási állandó értéke minden rendszernél eltérő.

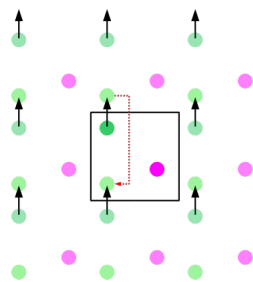
A bevezetett kényszerektől függően több különböző típusú MD szimulációt különböztetünk meg. Én NVT és az NPT szimulációkat futattam. Az NVT szimulációkban egy merev falú dobozt szimulálunk, ami képtelen összemenni vagy tágulni. Ez egy egyszerűen implementálható kényszer, így ilyen rendszereknél csak a termosztát az, ami a kényszerek közül számításigényes.

Az állandó nyomású NPT szimulációk előnye, hogy közelebb áll sok valós rendszer körülményeihez, hátránya viszont, hogy komplexebb, mint az NVT szimulációk, hiszen a nyomást egy barosztáttal állandón kell tartani, miközben a dobozméret folyamatosan változik.

A végső szakasz a produkciós futás. Miután előkészítettük a rendszert, készen áll a hosszabb megfigyelésre. A szakasz célja a molekulák helyzetéről és a rendszer többi jellemzőjéről adatgyűjtés, az itt keletkező trajektóriákból számíthatók a vizsgált anyagok tulajdonságai.

2.3. Algoritmus optimalizálások

Az eddig megemlített algoritmusokon kívül léteznek még olyan megfontolások, amelyek a szimulációk teljesítményét és a pontosságát nagymértékben növelik. Ezekből most kettőt emelek ki. Az első megoldás a periodikus határfeltétel (PBC). A szimulációk során véges dobozokkal dolgozunk, de a rendszer végtelenné tehető a dobozok egymás mellé másolásával. Ha a részecske eljut a doboz szélére, akkor nem visszapattan, hanem a szemközti oldalon újra belép. Az erőket úgy lehet ezzel a módszerrel számolni, hogy egy részecske körül minden másik részecske távolságának meghatározásakor megnézzük, hogy melyik periodikus képe van a legközelebb a központi részecskéhez. Ha nem így tennénk, akkor a doboz szélén levő atomok sokkal kevesebb másik atommal lépnének kölcsönhatásba, mint a doboz közepén lévők, ami értelmetlen eredményhez vezetne.



2. ábra. A PBC-t szemléltető ábra. A lila színű központi atomhoz képest a zöld atom távolsága a PBC-nek köszönhetően nem változik meg egy doboz élhossznyival, amikor a doboz egyik végéről hirtelen a másikba kerül [15].

A másik nagyon hasznos módszer a Particle Mesh Ewald (PME) [16], amely az Ewald-féle összegzési módszert alkalmazza az elektrosztatikus kölcsönhatások effektívebb kiszámításához. Az eredeti, $\mathcal{O}(n^2)$ -es összegzés helyett, amiben minden részecskét minden másikkal összepárosítunk, a PME egy sokkal jobban skálázható $\mathcal{O}(n \log n)$ -es algoritmus.

2.4. A szimulált eredmények szerkezetvizsgálati feldolgozása

2.4.1. A modellek szerkezeti információtartalma

A szimulációk során használt kényszerek miatt, amelyek egyben tartják a molekulákat, a szimuláció nem fog olyan intramolekuláris információt tartalmazni, amelyet nem a kényszereken keresztül implementáltunk. Ez alól valamelyest kivételt képeznek a flexibilis modellek, de ott is van jelen kényszer, csak az valamivel megengedőbb. Viszont végsősoron az intermolekuláris szerkezet tartalmaz olyan információt, ami nem paramétere a szimulációnak.

Az új információ abban rejlik, hogy a szimuláció során egymáshoz képest milyen távolságra és milyen szögben állnak be a molekulák, valamint, hogy milyen gyakran alakul ki közöttük hidrogénkötés. Emellett a különböző elméleti modellek kísérleti adatokkal való összevetése rávilágíthat az elkerülhetetlen kompromisszumok mértékére, amelyek a specifikus, nem szerkezeti paraméterek optimalizálása érdekében történtek. Erre példa az OPC, ami sok egyéb tulajdonságot jól visszaad [17], de a kísérleti szerkezettel kevésbé egyezik, mint sok másik modell 3.

2.4.2. Szerkezetvizsgálati alapfogalmak

Az atomi szintű szerkezet kvantitatív vizsgálatához elengedhetetlen a párkorrelációs függvények (pf) használata. A függvény egy központi atomtól r távolságra a relatív sűrűséget adja meg. A függvény

segítségével le lehet írni egy atom környezetét, az átlagos sűrűségtől való eltérést vizsgálva pedig rendezettséget kereshetünk a rendszerünkben.

$$g(r) = \frac{\rho_r}{\rho_{\text{átlag}}} \quad (11)$$

Ha például a $g(r)$ érték egy tartományon jól elkülöníthetően nagyobb értéket vesz fel, mint a környezete, akkor az lehet az első, vagy a többedik szomszéd távolsága az atomtól. Többkomponensű rendszerekben a parciális párkorrelációs függvény egy adott típusú központi atom körüli szerkezetet írja le, de a környezetében levő atomok közül csak egy bizonyos fajta referencia atomot veszünk figyelembe a $g(r)$ értékek kiszámításakor. Például, az OH parciális párkorrelációs függvény a hidrogénatomok relatív sűrűségét adja meg az oxigénatomoktól vett távolság függvényében. Mivel itt sok atom körüli sűrűség-fluktuációt vizsgálunk, az egyedi rendellenességek kiátlagolódnak, ami átlagtól való eltérés marad, az - kellően nagy minta esetén - rendszerszintű jelenségnek tekinthető.

A szimulált rendszerből számolt párkorrelációs függvényeket valahogyan össze kell tudni vetni valós adatokkal. Párokrelációs információt azonban közvetlenül nem tudunk mérni. Ezért szükség van közvetett leképezési módszerekre, amelyekből valamilyen módon előállítható a párkorrelációs függvény, és a parciális párkorrelációs függvények is.

Erre a célra gyakran neutron- és röntgen-diffrakciós technikákat szoktak használni. A kapcsolatot a kísérleti adat és a szimulált párkorrelációs függvény között a mért intenzitásokból számolható szerkezeti függvény adja meg (12). A szerkezeti függvény ($S(Q)$) az intenzitás eloszlását az inverz tér ($1/\text{\AA}$) függvényében írja le. Egykomponensű rendszer esetében a Q -térből inverz Fourier transzformálással áttérve az $r(\text{\AA})$ -térbe megkapjuk a kísérleti párkorrelációs függvényt:

$$S(Q) = 1 + 4\pi\rho \int r[g(r) - 1] \frac{\sin(Qr)}{Q} dr \quad (12)$$

A többkomponensű rendszerekben azonban a parciális $g(r)$ függvények és az azokból számított parciális szerkezeti függvények (13) különböző súlyokkal (14) járulnak hozzá a teljes párkorrelációs függvényhez.

$$S_{ij}(Q) = 1 + \frac{4\pi\rho}{Q} \int r (g_{ij}(r) - 1) \sin(Qr) dr \quad (13)$$

$$S(Q) = \sum_{i \leq j} \left[w_{ij}^{(X,N)}(Q) \left(1 + \frac{4\pi\rho}{Q} \int_0^\infty r (g_{ij}(r) - 1) \sin(Qr) dr \right) \right] \quad (14)$$

Egy többkomponensű rendszer teljes szerkezeti függvényéből nem lehet Fourier transzformálással egyértelműen előállítani a parciális párkorrelációs függvényeket.

A szerkezeti függvényből nem lehet előállítani a parciális párkorrelációs függvényeket, de fordítva ennek semmi akadálya. Tehát van egy, a valóságot leíró szerkezeti függvényünk. Amit még nem tudunk, hogy milyen atomi konfiguráció eredményez ilyen alakú függvényt, a célunk ezt a konfigurációt

megtalálni. Ezért van szükség -többek között- a szimulációra. A szimulált szerkezetről nem tudjuk, hogy jól írja-e le a valóságot, de az atomok helyzetét tetszőlegesen pontosan kinyerhetjük a szimuláció kimenetéből. A cél tehát a szimulációval egy olyan konfigurációt találni, amelynek a parciális párkorrelációs függvényeiből előállított szerkezeti függvény jól illeszkedik a kísérleti szerkezeti függvényhez. Ezt az illeszkedést a R_w (R-faktor) értékkel számszerűsítjük. Minél kisebb az R-faktor, annál jobb az illeszkedés a két függvény között.

$$R_w = \frac{\sqrt{\sum_i (S_{\text{exp}}(Q_i) - S_{\text{sim}}(Q_i))^2}}{\sqrt{\sum_i (S_{\text{exp}}(Q_i))^2}} \quad (15)$$

Ahol $S_{\text{exp}}(Q_i)$ a kísérleti szerkezeti függvény értéke, $S_{\text{sim}}(Q_i)$ a szimulált szerkezeti függvény értéke, az i index pedig az i -edik adatpontot jelöli.

2.5. Vizsgált vízmodellek

A kutatás során összesen 18 vízmodellt hasonlítottam össze, ezek vagy részecskeszámukban vagy töltésmegoszlásukban különböznek. A víz természetesen továbbra is csak 3 atommal van modellezve, a több részecskéből álló modellekben a többi részecske vagy merev virtuális, vagy pedig egy, szintén virtuális héj (shell) részecske. GROMACS-ban a merev virtuális részecskéket egyszerűen virtuális részecskéknek, a szintén virtuális héj részecskéket, pedig csak shell részecskéknek hívják, a továbbiakban én is ezt a konvenciót követem.

A virtuális részecskék segítségével a molekulák polaritását és az elektronfelhő alakját lehet pontosabban modellezni. Ezeknek a részecskéknek a tömegük mindig 0, töltésük viszont van, és az atomoktól egy előre megadott távolságra helyezkednek el. A shell részecskék pedig a többi molekula által indukált polarizációt képesek modellezni. A shell részecske egy központi atomhoz, az általam vizsgált vízmodellek esetében az oxigénhez van rugalmasan rögzítve. A shell részecske és az atom közötti távolságtól függ egy rugalmas potenciálfüggvény, így a teljes energiába beleszámít a rugalmas potenciális energia is. A shell részecske pontos modellezéséhez gyakrabban kell a rendszer energiáját mintavételezni, ezért ezek a modellek jelentősen nagyobb számítási igényűek.

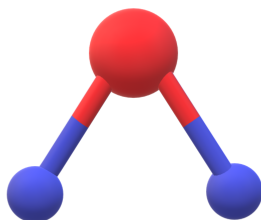
Érdemes még részletesebben kitérni a vízmodellek különböző típusaira, mert a későbbiekben tárgyalt eltérések a modellek szerkezetében, azok felépítésében keresendők. Jó tudni, hogy MD-ben a részecskéket site-oknak hívják, utalva arra, hogy a szimulációban egy részük virtuális. Ezen site-ok között számítunk ki kölcsönhatásokat a szimulációk során.

2.5.1. 3 site-os modellek

Az általam vizsgált 3 site-os modellek az SPC [8], SPC/E [18], SPC-FW [5], TIP3P [9], TIP3P-FB [19], OPC3 [20]. Ezek a modellek állnak a legközelebb a naiv vízmodellhez, mind a hat esetben egy egyszerű H_2O molekuláról van szó (lásd 3. ábra). A modellek látszólagos azonossága ellenére lényegi különbségek vannak közöttük. Különbözhetnek a töltésmegoszlásukban, valamint az atomok közti távolságban, esetleg az általuk bezárt szögben. Az itt említett modellek paraméterei a 1. táblázatban vannak feltüntetve. Egyre pontosabb kvantummechanikai számításokkal egyre jobban lehetett finomítani az egyes atomokra eső töltést, ami az elektronfelhő alakját tükrözi. Az eltérésnek egyéb oka lehet, például egyes mennyiségre való optimalizálás, hogy az adott mennyiség kísérleti adatokkal jobban egyezzen.

A modellek, az SPC-FW kivételével merevek, tehát nem egy függvény, hanem egy kényszer tartja egyben őket. A kötéstáv és a HOH kötésszög előre meg van adva, és az egyes molekulák alakjában nincs különbség. A SPC-FW esetében, ami rugalmas, a kötésszög és kötéstáv egy-egy kényszer helyett egy harmonikus potenciáltól függ (7. és 8. egyenletek). Ennek a megközelítésnek egyértelmű hátránya, hogy a teljes energia két további taggal bővül, ami számításigényesebb. Persze a modell ettől pontosabb lesz, hiszen ténylegesen merev kényszerek a természetben nem léteznek, a merev test csak közelítés.

A különböző modelltípusokat az alábbiakban -kizárólag illusztrációs szándékkal készült- ábrákon szemléltetem, amelyeket Paint3D-ben [21] készítettem.



3 site

3. ábra. A vízmolekula sematikus rajza. A piros gömb az oxigénatomot, a kék gömbök a hidrogénatomokat jelképezik.

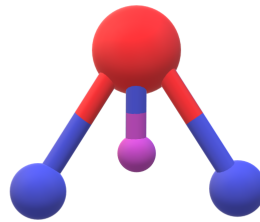
2.5.2. 4 site-os modellek

A 4 site-os modelleknél a negyedik részecske egy virtuális részecske, ami a két hidrogén és az oxigénatom síkjában található (lásd 4. ábra). A részecske a HOH szög szögfelezőjénél helyezkedik el, az oxigénatomtól modellenként más-más távolságra. A virtuális részecske segítségével pontosabban lehet modellezni

Modell	Publikálás éve	Szerző	Site	$d_{OH}(\text{nm})$	$d_{HH}(\text{nm})$	Θ_{HOH}	σ_{OO}	ε_{OO}	$q_O(\text{e})$	$q_H(\text{e})$	$d_{OM}(\text{nm})$	$q_M(\text{e})$
SPC	1981	Berendsen et al. [8]	3	0,1	0,1633	109,47°	0,316556	0,65017	-0,82	0,41	-	-
SPC/E	1987	Berendsen et al. [18]	3	0,1	0,1633	109,47°	0,316556	0,65017	-0,8476	0,4238	-	-
SPC/Fw	2006	Wu et al. [5]	3	0,1012	0,1690	113,24°	0,3165492	0,65029946	-0,82	0,41	-	-
TIP3P	1983	Jorgensen et al. [9]	3	0,09572	0,1514	104,52°	0,315066	0,636272	-0,834	0,417	-	-
TIP3P-FB	2014	Wang et al. [19]	3	0,10118	0,1639	108,15°	0,3178	0,65214	-0,84844	0,42422	-	-
OPC3	2016	Izadi, Onufriev [20]	3	0,09789	0,1599	109,47°	0,317427	0,68369	-0,8952	0,4476	-	-
TIP4P	1983	Jorgensen et al. [22]	4	0,09572	0,1514	104,52°	0,315066	0,636272	0	0,52	0,015	-1,04
TIP4P/Ew	2004	Horn et al. [23]	4	0,09572	0,1514	104,52°	0,316435	0,680946	0	0,52422	0,0125	-1,04844
TIP4P/2005	2005	Abascal, Vega [4]	4	0,09572	0,1514	104,52°	0,31589	0,774916	0	0,5564	0,01546	-1,1128
TIP4P-FB	2014	Wang et al. [19]	4	0,09572	0,1514	104,52°	0,31655	0,74928	0	0,52587	0,010527	-1,05174
OPC	2014	Izadi et al. [17]	4	0,08724	0,1371	103,6°	0,316655	0,89036	0	0,6791	0,01594	-1,3582

1. táblázat. Az általam vizsgált 3 és 4 részecskéből álló vízmodellek paramétereit. A virtuális részecskét a konvenciónak megfelelően M-mel jelöltem. Ezekben a modellekben csak az oxigénatomoknak vannak nem 0 értékű LJ paramétereit, így az OH és HH interakcióknál nincs LJ-féle potenciális energia.

a vízmolekulák dipólmomentumát. Az ilyen modelleknél a negatív töltés nem az oxigénatomhoz, hanem a virtuális részecskéhez tartozik. Így a negatív töltés nem az oxigénatom középpontjában lesz, hanem attól tetszőlegesen eltolható. Az általam vizsgált ilyen modellek a TIP4P [22], TIP4P-EW [23], TIP4P-FB [19], TIP4P-2005 [4], OPC [17]. Ezek mind merev modellek, amelyek töltéseloszlásait különböző technikákkal számították ki, a szimuláció kimenetéből kiszámolt különböző mennyiségek pontos modellezése érdekében.



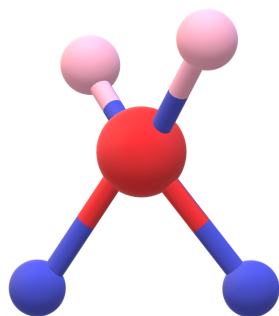
4 site

4. ábra. A 4 site-os vízmolekula sematikus rajza. A lila gömb a H_2O molekula által meghatározott síkban található virtuális részecskét jelöli.

2.5.3. 5 és több site-os modellek

Az öt site-os modellek közül két különböző típust vizsgáltam. Az egyik lehetőség, hogy modellezzük a hidrogénkötésben gyakran résztvevő nemkötő elektronpárokat, két ún. lone-pair részecskével (5. ábra).

Ilyen modellek a TIP5P [10], TIP5P-EW [24], TIP5P-2018 [3] modellek. Ez a három modell egyedül a töltéseloszlásában különbözik. A modellek balról jobbra egymás leszármazottjainak tekinthetők.

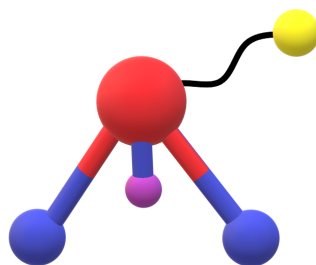


5 site, 2 lone-pair

5. ábra. Az 5 site-os molekula sematikus rajza, a két lone-pair részecskével, amelyeket a rózsaszín gömbök jelölnek.

A másik lehetőség, hogy a dipólmomentumot reprezentáló, HOH-szögfelezőnél található virtuális részecskén kívül a molekulát egy, az oxigénatomhoz kötött héj részecskével polarizálhatóvá tesszük (4. ábra). Ilyen modellek az SWM4-DP [25], és az SWM4-NDP [11]. A két modell között annyi a különbség, hogy az SWM4-DP esetében a héj részecske (drude particle) pozitív töltésű, az SWM4-NDP-nél negatív töltésű, ahogy az a valóságban is elvárható a polarizált atomtól.

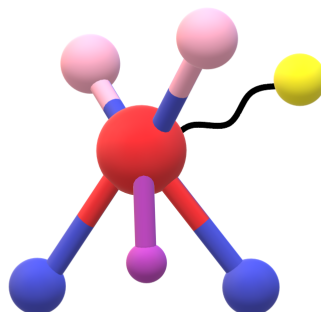
Félreértésre adhat okot, hogy az MD-ben a shell részecskéket nem mindig kezelik site-nak, ezért is SWM4 a modellek neve. Tehát ezek a modellek felfoghatóak 4 site-os vagy 5 site-osnak is. Én, mivel komplexitásban közelebb állnak az 5 site-os modellekhez, oda soroltam őket, de a paramétereket összefoglaló táblázatban (2) és az ábrámon a gyakoribb konvenciót használtam.



4 site, polarizálható

6. ábra. A sárga részecske a shell részecskét jelöli, a fekete hullámos vonallal a polarizálható kötést, mint "rugót" illusztráltam.

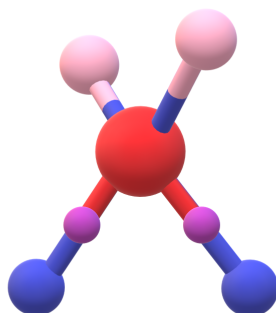
Az általam vizsgált legmodernebb és egyben legbonyolultabb klasszikus vízmodellek a 7 site-os molekulák. Ezek közül két merev modellt vizsgáltam meg: az SWM6-ot [26] és a TIP7P-t [27]. Az SWM6 lényegében a két 5 site-os konfiguráció uniója, tehát tartalmazza a két lone-pair molekulát, de polarizálható és a dipólmomentumot pontosabban reprodukáló virtuális site is jelen van a modellben.



SWM6

7. ábra. Az SWM6-os vízmodell sematikus rajza.

A TIP7P viszont más oldalról közelíti meg a víz modellezését. A 2 lone-pair részecskén kívül két másik virtuális részecskén helyezi el az oxigénatom negatív töltését. Ezek az oxigén és a hidrogénatomok között találhatóak meg félúton (7. ábra). Maguk az atomok pozitív töltésűek, és az atomtörzset reprezentálják, míg a negatív töltés a virtuális részecskéken oszlik el (lásd 2. táblázat).



TIP7P

8. ábra. A TIP7P modell sematikus rajza.

Modell	Publikálás éve	Szerző(k)	Site	$d_{OH}(\text{nm})$	Θ_{HOH}	$d_{HH}(\text{nm})$	$\sigma_{OO}(\text{nm})$	$\varepsilon_{OO}(\text{kJ/mol})$	$q_O(\text{e})$	$q_H(\text{e})$
SWM4-DP	2003	Lamoureux et al. [25]	4+1	0,09572	104,52°	0,1514	0,31803	0,86056512	-1,77185	0,5537
SWM4-NDP	2006	Lamoureux et al. [11]	4+1	0,09572	104,52°	0,1514	0,318395	0,88257296	1,71636	0,55733
TIP5P	2000	Mahoney, Jorgensen [10]	5	0,09572	104,52°	0,1514	0,312	0,66944	0	0,241
TIP5P/Ew	2004	Rick [24]	5	0,09572	104,52°	0,1514	0,3097	0,744752	0	0,241
TIP5P/2018	2018	Khalak et al. [3]	5	0,09572	104,52°	0,1514	0,3145	0,79	-0,64111	0,394137
SWM6	2013	Yu et al. [26]	6+1	0,09572	104,52°	0,1514	0,31983264	0,677808	1,91589	0,5307
TIP7P	2019	Zhao et al. [27]	7	0,09572	104,52°	0,1514	0,3156	0,6284368	0,11094	0,58014

Modell	$d_{OM}(\text{nm})$	$q_M(\text{e})$	$d_{OL}(\text{nm})$	$q_L(\text{e})$	Θ_{LOL}	$d_{OM_2}(\text{nm})$	$q_{M_2}(\text{e})$	$\sigma_{HH}(\text{nm})$	$\varepsilon_{HH}(\text{kJ/mol})$	Shell atomja	$q_S(\text{e})$	$k_D(\text{kJ}/(\text{molnm}^2))$
SWM4-DP	0,023808	-1,1074	-	-	-	-	-	-	-	O	1,77185	418000
SWM4-NDP	0,024034	-1,11466	-	-	-	-	-	-	-	O	-1,71636	418000
TIP5P	-	-	0,07	-0,241	109,47°	-	-	-	-	-	-	-
TIP5P/Ew	-	-	0,07	-0,241	109,47°	-	-	-	-	-	-	-
TIP5P/2018	-	-	0,07	-0,07358	109,47°	-	-	-	-	-	-	-
SWM6	0,0247	-1,1334	0,0315	-0,108	101,098°	-	-	-	-	O	-1,62789	418000
TIP7P	-	-	0,041	-0,17724	109,47°	0,04786	-0,45837	0,14476	0,01046	-	-	-

2. táblázat. Az 5 és 7 site-os, valamint polarizálható modellek paramétereit. Különböző atomok közötti kölcsönhatást a Lennard jones paraméterek mértani közepével számoljuk ki $\varepsilon_{OH} = \sqrt{\varepsilon_{HH}\varepsilon_{OO}}$

3. Munkafolyamat

3.1. Induló konfigurációk létrehozása

A szimulációk futtatásához szükséges egy kiinduló konfigurációs fájl, ami a részecskék kezdeti koordinátáit tartalmazza. Ehhez először készíteni kell egy néhány molekulát tartalmazó dobozt, majd ezekből felépíthető egy véletlenszerűen elrendezett kezdeti konfiguráció. Az összes szimulációm kocka alakú dobozban futtatam. A kezdetben még merev dobozoknak 5nm-es élhosszt állítottam be. 1000kg/m³-es sűrűséggel számolva 4000 vízmolekula ~4,93nm-es élhosszú kockának megfelelő térfogatot tölt ki. A becsült 4,93nm-es élhossznál kicsit nagyobbra szabtam a doboz méretét, hogy a szimuláció energia-minimalizálási szakaszában szabadabban mozogjanak az atomok, így gyorsabban konvergál a rendszer energiája a minimumértékhez.

A különböző szerzők által készített vízmodellek [3, 4, 5], [8, 9, 10, 11], [18, 19, 20, 22], [23, 17, 24, 25, 26, 27] alapján elkészítettem a GROMACS szimulációkhoz szükséges forcefield és topológia fájlokat. A modellek többsége nincs benne alapértelmezetten a GROMACS-ban. Ezért nekem kellett a megfelelő virtuális részecskéket tartalmazó molekulát is elkészíteni. A publikációk megadják a részecskék geometriáját, ebből kiszámoltam a részecskék relatív helyzetét. Ezután az egyik részecskét (általában oxigént) origónak választva megadtam egy molekula koordinátáit, ami mintájára a többi molekulát felépíti a GROMACS. Az immár több molekulából álló koordináta fájl segítségével a kiinduló dobozt felépítettem, majd a dobozzal a szimulációt futtattam.

3.2. A szimuláció szakaszai

Ahogy az megszokott az MD szimulációk futtatásakor, én mind az energia-minimalizálást, mind az egyensúlyi állapotot egy-egy további szakaszra osztottam, így összesen öt különböző részre bontottam a szimulációt.

A két em szakasz nem sokban tér el egymástól, ugyanúgy a Steepest Descent algoritmus alapján történik a részecskék mozgatása. A két szakasz között azonban van két lényegi különbség. Míg az első szakaszban hagyjuk, hogy a molekulák geometriáját változtassa a program, a második szakaszban, ha a modell maga nem rugalmas, ez már nem megengedett, a kötéstávot és a kötőszöveget merev kényszerekkel állandón tartjuk. Mivel nem célunk megváltoztatni a molekulák geometriáját, az első szakasz csak 1000 lépésből áll, míg a második 50000-ből. Viszont a tapasztalat szerint az első szakasz nélkül nem konvergál a részecskék között ható erő egy kellően alacsony értékhez (10^2 N nagyságrend).

Az egyensúlyi állapot kettébontásával lehetővé válik a szimulációkba fokozatosan bevezetni a kényszereket. Az összes modellt NPT szimulációként futtattam, ahol előnyös csak a hőmérsékletkényszer után bevezetni a nyomáskényszert. Az egyensúlyi állapot első szakaszában bevezetjük a hőmérsékletkényszert, és hagyjuk elég sokáig - esetemben 100 ps elégnek bizonyult - futni a szimulációt, ahhoz, hogy egyensúlyi helyzetbe kerüljön.

Én a szimulációk során a Nosé-Hoover [28] termosztátot használtam. Előnye, hogy pontos, de emiatt komplexebb is, mint a nagyobb molekulákra optimalizált termosztátok. Azonban a víz egy egyszerűbb molekula, így a nagyobb számításigény nem okoz problémát. További aggály az egyszerűbb algoritmusokkal kapcsolatban, hogy mivel egyszerűsítésekkel élnek, tovább absztrahálják a szimulációt, eltávolítják a valós viselkedéstől. Egy egyszerű lehetséges algoritmus az lenne például, hogy a részecskék mozgási energiáját önkényesen úgy változtatjuk, hogy az átlagos mozgási energia megfeleljen a kívánt hőmérsékletnek. Ezzel viszont alapvetően belenyúlunk a rendszerbe, nem a valóságban megfigyelt módon viselkedne a rendszer és így a dinamikai jellemzők eloszlásai nem adnák vissza a kísérletileg mért eloszlásokat.

Az egyensúlyi állapot második szakaszában a rendszerbe bevezetjük a nyomáskényszert, amit barosztát valósít meg. A barosztát egy olyan algoritmus, amely a doboz dimenzióinak mozgatásával éri el, hogy a rendszerben uralkodó nyomás állandó maradjon. Az általam használt algoritmus a Parrinello-Rahman [29] [30], amely a Nosé-Hoover algoritmushoz hasonló elven működik. A barosztát a becsült dobozméretet beállítja a rendszer állapotára jellemző mennyiségre.

A szimulációkat 10 ns-ig futtattam, ami alatt 1000 konfigurációt mentettem le, ezekből számoltam ki a párkorrelációs függvényeket. A lépésköz általában fs nagyságrendbe esik bele, így egy futáshoz

több millió lépés kiszámítása szükséges, mivel én 2 fs-os lépésközzel dolgoztam, nekem ez produkciós futásonként 5 millió lépést jelentett.

3.3. A szimulációk további paraméterei

A munkám során a megszokott 3, 4, és 5 részecskés modellek mellett futtattam két komplexebb, 7 részecskéből álló modellt is, amelyek nagyobb számítási igényűek. Ezért ezeket a modelleket 2000 molekulával szimuláltam, míg a többi modell esetén 4000 molekulát használtam. Az általam referenciaként használt két adat közül, a neutrondiffrakció [2] 283 K-en, a röntgendiffrakció [1] 295 K-en készült. Ezzel összhangban modellenként két szimulációt futtattam, egyet 283 K-en, egyet pedig 295 K-en.

A szimulációk a GROMACS 2020.5-ös [31] verzióval készültek. Egy szimulációs lépésköz 2 fs volt, az egyensúlyozás 2,1 ns-ig, a produkciós futás 10 ns-ig tartott. A hőmérséklet csatolási tényező 2 ps, a barosztát esetében a csatolási tényező 4 ps volt. A szimulációk állását 5000 időlépésenként (10 ps-onként) mentettem le, a párkorrelációs függvényeket ezek átlagából számoltam. A 7 részecskés modelleknél a stabilitás érdekében a hőmérséklet csatolási tényezője 0,5 ps-ra, a barosztát csatolási tényezője 2 ps-ra csökkentettem.

3.4. A szimulációk vezérlése, az eredmények kiértékelésének folyamata

Egy szimuláció 5 szakaszból áll, minden modellt 2 hőmérsékleten futtatam, ami a 18 modell mellett 180 külön futásnak felel meg. Ezek rendszerezett vezérléséhez és az adatok feldolgozásához elengedhetetlen volt minden részt automatizálni. Tehát egészen a futások kezdetétől, a kiértékelésen át az ábrák ábrázolásáig minden automatizált programok/szkriptek segítségével történt. Ezeket a szkripteket Bash-ben és Pythonban írtam. A különböző mennyiségek kiértékeléséhez felhasználtam a GROMACS által tartalmazott függvényeket. A Fourier-transzformáció számításához pedig a kutatócsoportunk belső szkriptjét használtam, melyet Dr. Temleitner László készített. Az ábrákat Pythonban készítettem a Matplotlib [32] csomag segítségével.

A szimulációs futásokból lementett részecskekonfigurációkból kiszámoltam a parciális párkorrelációs függvényeket. A parciális pf-ekből előállítottam a szerkezeti függvényeket, majd ezeket összevetettem a kísérleti adatokkal. A számolt adatok két kísérleti görbével lettek összehasonlítva, egy neutrondiffrakcióval [2] (továbbiakban ND) és egy röntgendiffrakcióval [1] (továbbiakban XRD). Az XRD adat szerzői az $S_{OO}(Q)$ parciális szerkezeti függvényt adták meg. Így az csak az oxigénatomok járulékát tartalmazza.

A kísérleti adatokat renormáltam a szimulált eredményhez, annak érdekében, hogy a mérésben fennmaradó bizonytalanságot csökkentsem. A renormálásnál a szerkezeti függvény értékeit megszorozhatjuk egy számmal (újraskálázás), valamint hozzáadhatunk egy Q -ban n -ed rendű, polinomiális tagot. Én a renormáláskor skáláztam a szerkezeti függvény értékét, amelyhez hozzáadtam egy legfeljebb másodfokú polinomot. Az így kiszámolt R-faktorok (illeszkedés) értékét nagyban befolyásolja, hogy az adat melyik tartományához renormáljuk azt. Ha kis Q értékekre rossz az illeszkedés, az arra utal, hogy az atomoktól nagy távolságokra nem egyezik a modell a kísérleti értékekkel. Ha nagy Q értékeknél van gond az illeszkedéssel, akkor az intramolekuláris szerkezet vagy az intermolekuláris kötések vannak a kísérlet szerint máshogy elrendezve, mint ahogy azt a modell alapján jósoltuk.

4. Eredmények

4.1. A modellek illeszkedése a kísérleti adathoz

A szimulációkból, a kísérleti adatokhoz számolt illeszkedés a 3. és 4. táblázatokban van összefoglalva. Az első táblázatban a $0-20 \text{ \AA}^{-1}$ tartományra, a másodikban a $0-10 \text{ \AA}^{-1}$ és $10-20 \text{ \AA}^{-1}$ tartományokra kiszámított értékek vannak. Látható, hogy a modellek között lényegi eltérések vannak. Néhány modell esetében már az intramolekuláris kötésben is alapvető eltérés figyelhető meg a többihez képest. Ezekben az esetekben az újraparaméterezés az illeszkedést negatívan befolyásolja. Ez megfigyelhető az OPC-nél, ahol a kötésértékek eltérnek a többi modellettől, és az ND adathoz való illeszkedés az összes közt a legrosszabb. A modellt a szerzői az elektromos tulajdonságok pontos reprodukálására készítették. Ezzel cikkük alapján sikerrel is jártak [17], de szemmel láthatólag a szerkezet leírásának rovására.

A vízmolekula belső geometriája mára már viszonylag pontosan meghatározott. Mivel célunk a hosszabbtávú szerkezet vizsgálata, azért felmerül a kérdés, hogy az intramolekuláris szerkezet "feláldozásával" milyen paramétereket lehet pontosabban leírni. Nem kizárható, hogy egy OPC-hez hasonló, de a hosszútávú szerkezet leírására optimalizált modell illeszkedése hosszútávon jó lehet, azzal egyidejűleg, hogy a molekula geometriája nem pontos.

Érdekes még megfigyelni, hogy az újabb modellek, amelyeket más paraméterekre optimalizáltak, rosszabbul teljesítenek, mint a régebbi modellek, amelyekből a készítőik kiindultak. Erre példa az TIP4P és OPC modellek közötti jelentős eltérés (19% és 40%). Ha egy modellt néhány paraméterre optimalizálunk, akkor látszik, hogy szerkezeti szempontból sok esetben rosszabb modellt kapunk. A víznek számos olyan tulajdonsága van, amelyre lehetne illeszteni a modell paramétereit, így csak szerkezeti tulajdonságokat vizsgálva nem lehet megállapítani, hogy az újabb modellek általánosságban

pontosabbak-e, mint az elődeik. Szerencsénkre a modellek készítői tudják, hogy mire alkották meg a modelleket, így azok a tulajdonságok, amelyeket a legpontosabban leír a modelljük, a modellekhez tartozó publikációban általában fel vannak tüntetve. Amire viszont fontos figyelni az újabb modelleket illetően, hogy nem a szerkezet volt a készítésükkor az elsődleges cél, így sokszor még a régebbi, egyszerűbb modellekkel dolgozva is pontosabb szerkezeti függvényeket kapunk.

MODEL	$R_{wND}(\%), 0-20\text{\AA}^{-1}$	$R_{wXRD}(\%), 0-20\text{\AA}^{-1}$
SPC	26	7
SPCE	26	8
SPC-FW	27	7
TIP3P	30	12
TIP3P-FB	28	9
OPC	40	5
OPC3	22	7
TIP4P	19	6
TIP4P-2005	18	6
TIP4P-EW	19	7
TIP4P-FB	19	7
SWM4-DP	19	4
SWM4-NDP	21	5
SWM6	21	10
TIP5P	23	10
TIP5P-EW	23	7
TIP5P-2018	22	5
TIP7P	22	7

3. táblázat. A szimulációs trajektóriákból kiszámított szerkezeti függvények illeszkedése a kísérleti, renormált Neutrontdiffrakciós és Röntgendiffrakciós adatokból származó görbékhez. A második és a harmadik oszlop értékeinek illeszkedése a $0-20\text{\AA}^{-1}$ tartományon lett kiszámítva. A kísérleti adatok ezen tartományon történő legjobb illeszkedéshez lettek renormálva mindkét esetben. A mérési adatok nagyobb Q értéken is elérhetőek voltak, de akkora értékekre már nagyon kicsi az intenzitás, így nagy lenne a zaj (bizonytalanság).

MODEL	$R_{w_{\text{ND}}}(\%), 0-10\text{\AA}^{-1}$	$R_{w_{\text{XRD}}}(\%), 0-10\text{\AA}^{-1}$	$R_{w_{\text{ND}}}(\%), 10-20\text{\AA}^{-1}$	$R_{w_{\text{XRD}}}(\%), 10-20\text{\AA}^{-1}$
SPC	13	7	49	25
SPCE	12	7	50	46
SPC-FW	10	6	57	39
TIP3P	27	11	99	33
TIP3P-FB	13	8	29	33
OPC	33	4	60	42
OPC3	13	7	24	27
TIP4P	14	6	24	34
TIP4P-2005	12	5	25	17
TIP4P-EW	13	7	24	30
TIP4P-FB	13	6	25	21
SWM4-DP	13	4	26	17
SWM4-NDP	16	5	25	16
SWM6	16	9	23	83
TIP5P	19	9	22	63
TIP5P-EW	18	7	22	82
TIP5P-2018	17	5	25	20
TIP7P	17	6	27	18

4. táblázat. R-faktorok, a $0-10\text{\AA}^{-1}$ és $10-20\text{\AA}^{-1}$ tartományokra kiszámítva. A számításokkor a kísérleti adat mindig csak az adott tartományhoz lett normálva.

4.2. A modellek számításigényének összehasonlítása

Még mielőtt a konkrét modellek elemzésébe belekezdenénk, érdemes egy további, nagyon fontos szempontot figyelembe venni. A modellek fizikai tartalma mellett a komplexitásuk is nagyban befolyásolja a felhasználhatóságukat. Például a TIP5P és az SWM6 modellek R-faktorai összemérhetőek, mégis az utóbbi ugyanakkora időtartamon történő szimulálásához, saját méréseim alapján 4-szer annyi idő kell, fele annyi szimulált molekula mellett (5. táblázat).

Egy modell szimulálásánál nem mindig lehet előre tudni, hogy milyen paraméterek mellett lesz stabil a szimuláció. Egy modell lefuttatásának ideje valójában a tiszta futásidő többszöröse, mert sok esetben többször is le kell futtatni, mire megtaláljuk a megfelelő futási paramétereket. Összetettebb

rendszereknél, ahol a víz, mint oldószer van jelen, már eleve komplex molekulákat szimulálunk. Ha a másik molekula mellé még a vízmolekula is bonyolult, abban az esetben a futásidő akár hetekig is eltarthat. Emellett az ilyen rendszereknél sokszor hosszabb időtartományt kell szimulálni, hogy a lassabb lefolyású folyamatokat is meg lehessen figyelni.

Modell	Futásidő (ns/nap)
SPC	1001
SPC/E	996
SPC/FW	844
OPC3	1018
TIP3P	997
TIP3P-FB	993
OPC	680
TIP4P	681
TIP4P-2005	679
TIP4P-EW	688
TIP4P-FB	690
SWM4-DP	108
SWM4-NDP	100
TIP5P	469
TIP5P-2018	478
TIP5P-EW	468
SWM6	62
TIP7P	293

5. táblázat. ns/napra normált számítási igény a különböző modellekre, 4000 molekula esetén. A szimulációk az alábbi hardveren futottak:

CPU: Intel(R) Core(TM) i7-8700K CPU @ 3,70GHz

GPU: NVIDIA GeForce RTX 3080 10GB

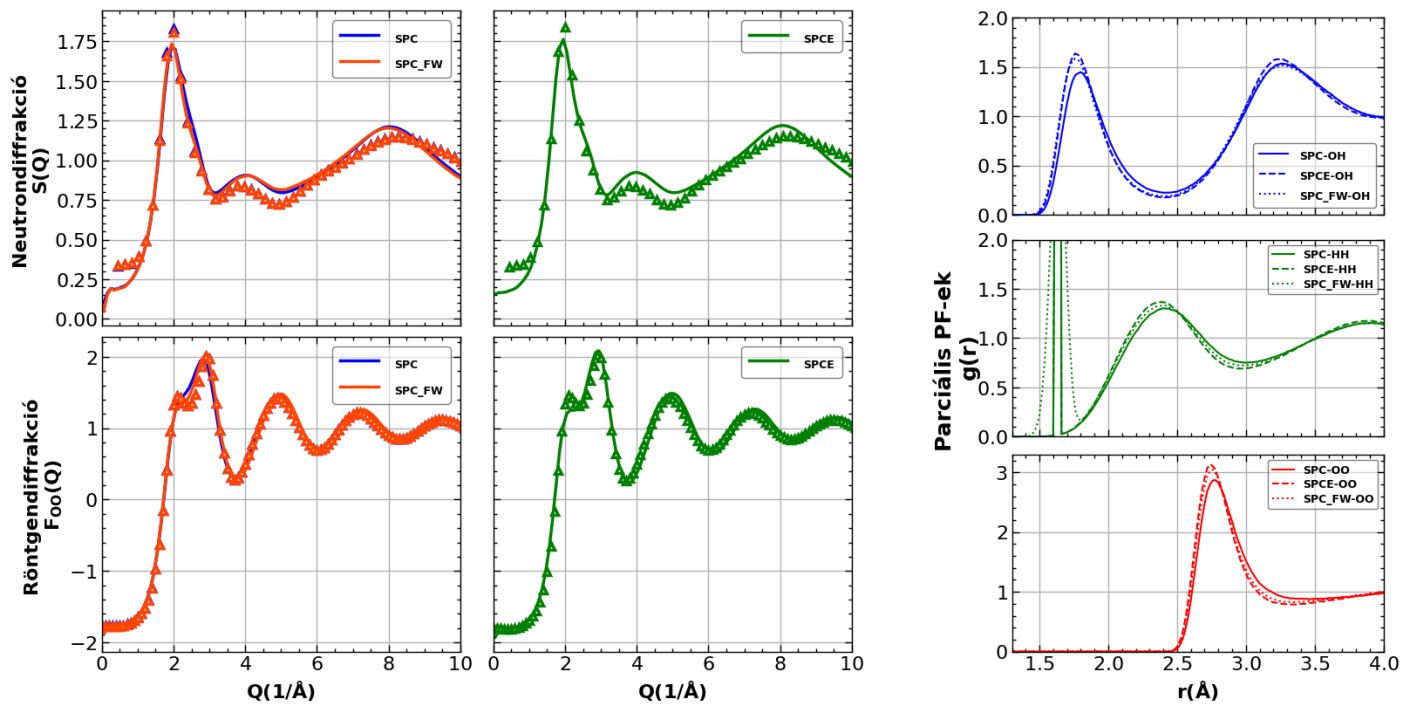
4.3. A vízmodellek szerkezetének elemzése

Az alábbi ábrák segítségével vizuálisan is kereshetjük az eltérések okát a hasonló modellek között. Az első két oszlopban a modellek illeszkedése a renormált kísérleti görbékkel látható, itt a teli vonal a

szimulált adat, a háromszög jelölők pedig a kísérleti adatok értékeit reprezentálják. Az ND és az XRD illeszkedésekről is két-két ábrát készítettem, hogy egy ábrára legfeljebb 2 kísérlet-szimuláció adatpár kerüljön. Az egy oszlopban levő ábrák x-tengelye közös.

Az átláthatóság kedvéért nem minden kísérleti adatpontot ábrázoltam, ez a görbék alakján nem változtat. Az ND és XRD görbék illeszkedését csak mindig olyan tartományon mutatom be, ahol szemre is jól látszik az eltérés vagy egyezés a mérés és a szimuláció között. A harmadik oszlopban a különböző parciális pf-ek láthatóak. A parciális pf-eket legtöbbször csak kisebb r tartományon ábrázoltam, nagyobb r értékekre a $g(r)$ -ek tartanak az 1-hez, mert ellentétben a kristályokkal, a folyékony víznél nincs hosszútávú rendezettség. Ezt szemlélteti a 11. ábra, ahol 8 Å után az oszcillációk lecsengnek.

4.3.1. SPC, SPC/E, SPC-FW



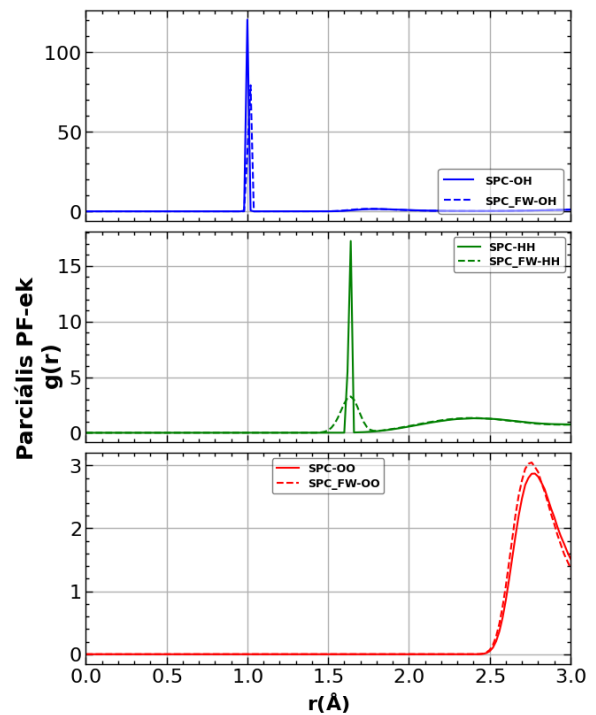
9. ábra. Az SPC, SPC/E, és a rugalmas SPC-FW modellek összehasonlítása

A három SPC modell nem sokban különbözik egymástól, sem szerkezeti, sem komplexitási szempontból, a számítási igényük is hasonló. A 9. ábra grafikonjain látszik, hogy az XRD (Röntgendiffrakciós, bal alsó két ábra) adattal mindhárom modellnél jobb az illeszkedés, mint az ND (Neutrontdiffrakciós, bal felső két ábra) adattal. Jól látható, hogy az alsó ábrákon a háromszög jelölökkel ábrázolt kísérleti adat kis hibával illeszkedik a folytonos görbékkel ábrázolt szimulált adatokhoz.

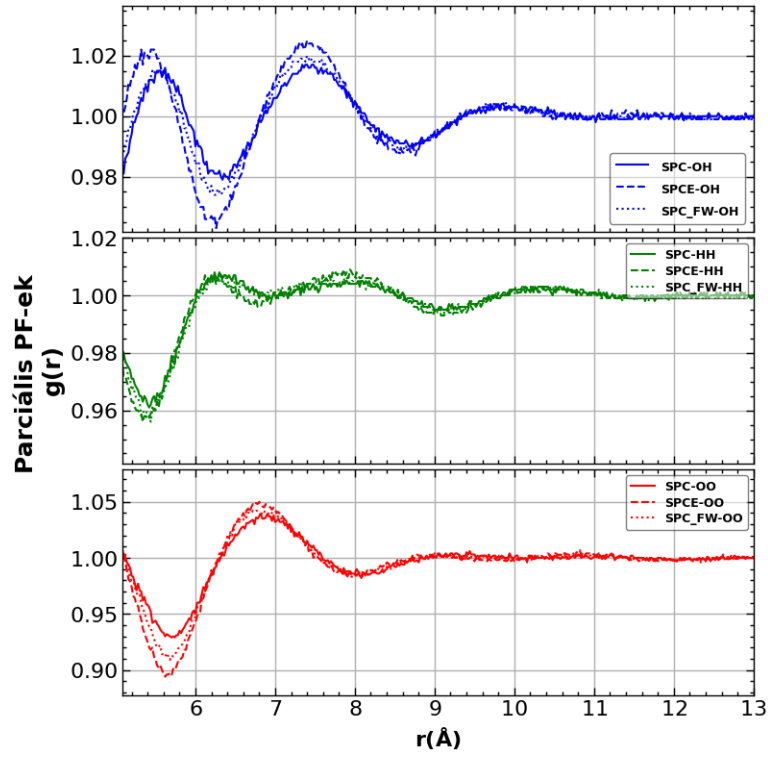
A bal alsó ábrán (9) a két modell (SPC és SPC-FW) szerkezeti függvénye annyira hasonló, hogy a kék színnel jelölt SPC adatpontjai alig látszanak. Így a modellbéli, első O-O atomtávolságok jól egyeznek a kísérlettel. A többi kötéstávról ez kevésbé mondható el.

Ha összevetjük az illeszkedéseket, a $0-10 \text{ \AA}^{-1}$ tartományon (4. táblázat), amely r -térben a központi atomoktól nagyobb távolságra levő kötésekről ad információt, akkor kvantitatívan sem találunk jelentős eltérést a két modell illeszkedése között. Ezen a tartományon SPC-FW-nek van a legjobb illeszkedése, mind az ND, mind az XRD adatahoz. A flexibilis modell tehát valamelyest pontosabban írja le a hosszabb távú szerkezetet, a számítási igénye sem sokkal nagyobb, a három modell közül ez tűnik a legalkalmasabbnak szerkezetvizsgálatra.

Ezen kívül érdemes még kitérni a rugalmas és a merev modellek közötti különbségre, amelyet 10. ábrán jól nyomon követhetünk. A modellek paraméterezésekor a vízmolekula geometriáját explicit megadjuk, így a kialakuló csúcsok a kristályos anyagok párkorrelációs függvényeinek hegyes csúcsaira emlékeztetnek, hiszen a vízmolekulákon belül a modelljeink tökéletes rendezettséget követelnek meg. Ha hagyunk valamennyi szabadságot a rendszernek a kötésszöget és a kötéstávolságot változtatni, kiderül, hogy ezek a kényszerek nem adódnak a modell többi paraméteréből. Az OH kötéstáv nagyjából olyan jól meghatározott rugalmas esetben, mint merev kényszerekkel. Itt a párkorrelációs függvények mintavételezési felbontása (esetemben $0,02 \text{ \AA}$) szab határt a pontosságnak. Ezzel szemben az elsődleges HH távolság bizonytalan-sága nagyjából $\pm 0,15 \text{ \AA}$. Így a kötéstáv a rugalmas kényszer mellett jól teljesül, de a kötésszög értéke már bizonytalanabb, hozzávetőleg $\pm 15^\circ$ -ot jelent állandó kötéstávot feltételezve. Ennyi alapján még elhamarkodott lenne kijelenteni, hogy a valóságban is ilyen sokféle értéket vehet fel a HOH kötésszög, de az eredmények alapján nem is kizárható. A vízmolekulák HOH kötésszögét viszont már régebben kimérték [33], ennél jóval nagyobb pontossággal.

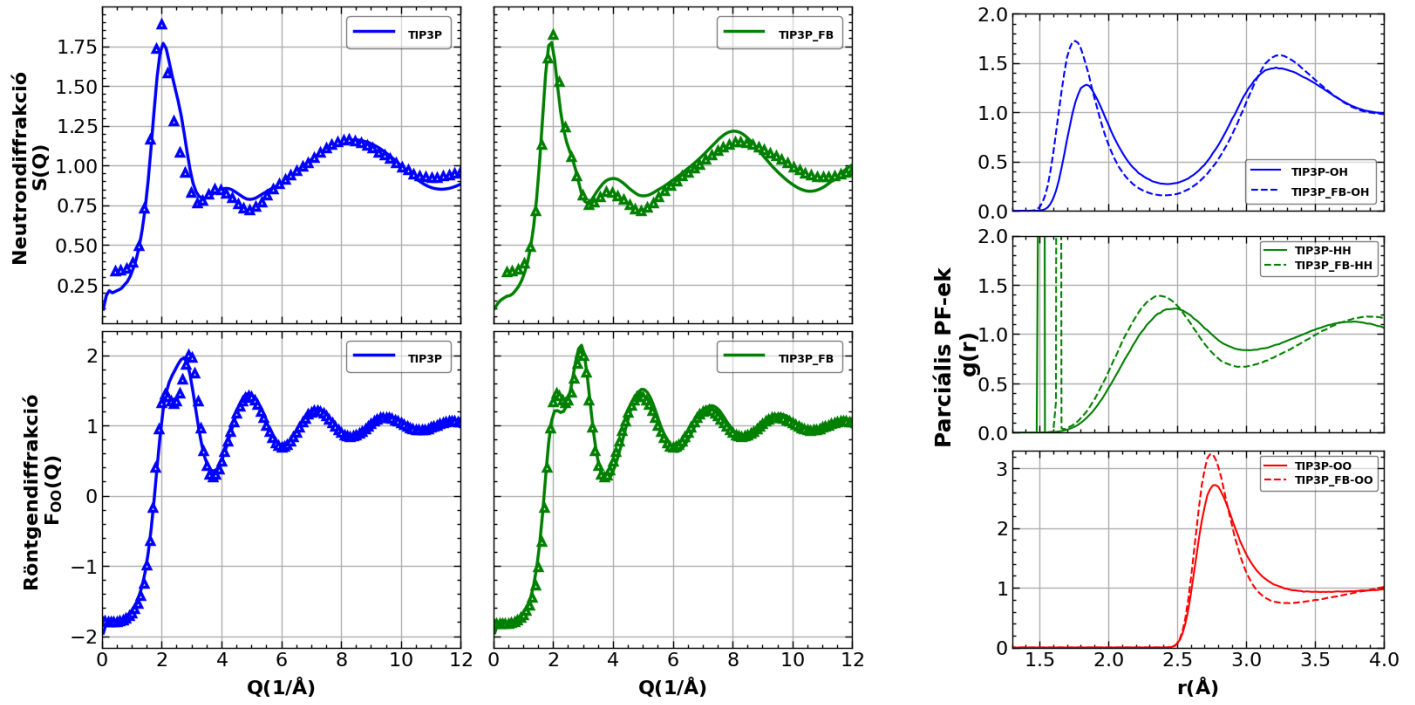


10. ábra. A merev SPC, és a rugalmas verziója, az SPC-FW összehasonlítása intramolekuláris szerkezet szempontjából.



11. ábra. Az SPC, SPC/E, és a rugalmas SPC-FW modellek pf-jeinek összehasonlítása nagyobb r tartományon is.

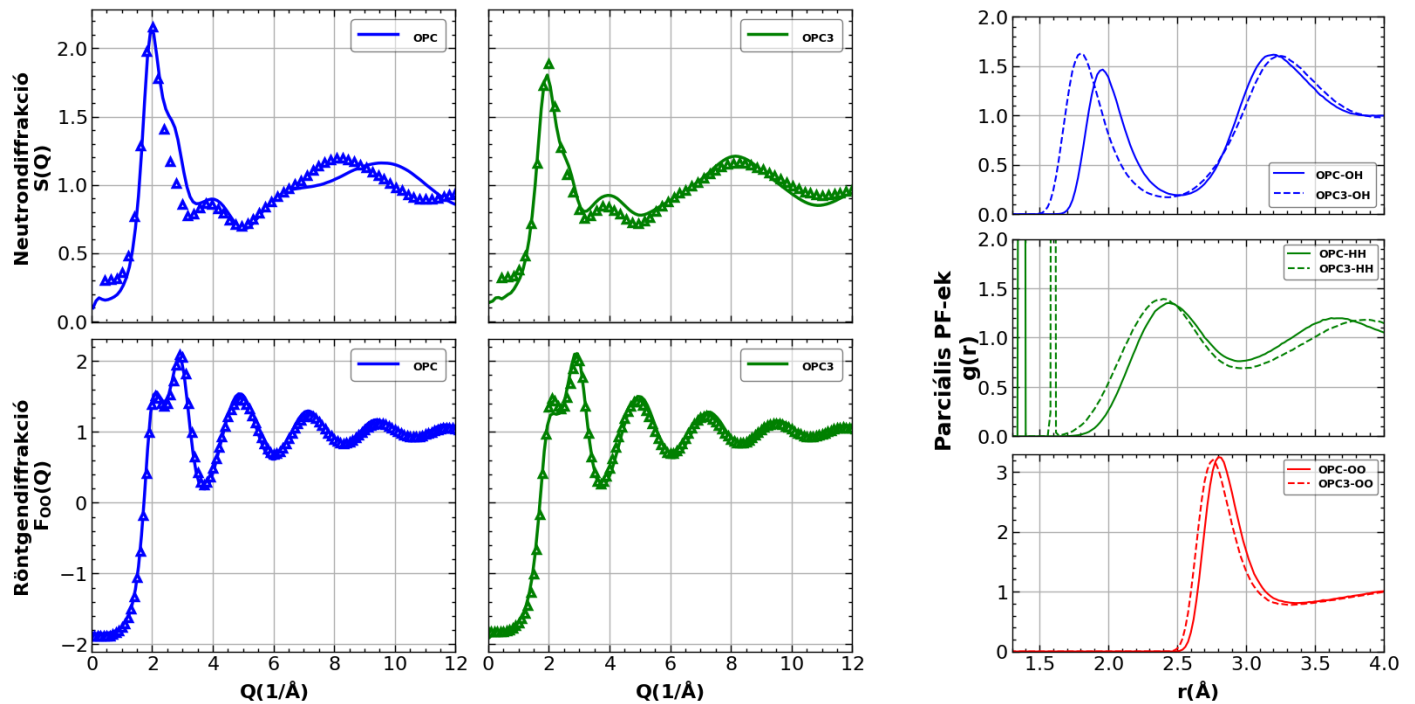
4.3.2. TIP3P, TIP3-FB



12. ábra. A TIP3P és a több, mint 30 évvel újabb TIP3-FB összehasonlítása

A két modell futásidőben lényegében megegyezik, az újraparaméterezéstől nem is várnánk, hogy számottevően változzon a számítási igénye a modellnek. Érdekes, hogy a két modell parciális pf-ei között is jelentős eltérések figyelhetők meg. A TIP3P-FB-nél az első HH csúcs távolsága nagyobb, mint a TIP3P-nél, amit a paraméterek alapján meg is követeltünk. Ami viszont a szimuláció eredménye, hogy a TIP3P-FB esetében mind az első intermolekuláris OH, mind az első OO csúcsok amplitúdója nagyobb, valamint a csúcsok valamelyest jobban meghatározottak. Emellett a második OH csúcs (az első csúcs az ábrákon nincs feltüntetve, 1 Å-nél van) a TIP3P-FB-nél kisebb távolságra van. Mivel a második OH csúcs feleltethető meg a hidrogénkötéseknek, ebben a hidrogénkötések távolsága valamennyivel kisebb, mint a TIP3P-nél. Ami még érdekes, hogy a TIP3P-FB mindhárom parciális pf-énél a csúcsok között a minimumhely jobban közelíti a nullát. A jobban meghatározott és a környezetüktől jobban elkülönülő csúcsok nagyobb rendezettségére utalnak. A kísérlethez a TIP3P-FB jobban illeszkedik, így ez modell közelebb áll a valósághoz. Ha azonban megnézzük nagyobb Q értékekre az ND adatot, mind fázisban, mind amplitúdóban jelentős eltérések figyelhetők meg a kísérletből származó és a szimulációs szerkezeti függvények között, mindkét modell esetében. Az R-faktorok is ezt tükrözik, 10-20 Å⁻¹-on, a modellek sok más modellnél nagyobb R-faktorral rendelkeznek. Ha viszont megnézzük a 0-10 Å⁻¹ tartományt, akkor a TIP3P-FB az ND adathoz jól illeszkedik, míg a TIP3P ND-re számított R-faktora ezen a tartományon több, mint kétszer akkora. Összességében TIP3P-FB bármely tartományra nézve jobb illeszkedést mutat mind ND-vel, mind XRD-vel, mint az elődje. A modell készítői által végzett számítások alapján [19] sok más tulajdonságot is pontosabban ad vissza, mint a TIP3P. Így a cikk alapján a leggyakoribb felhasználásokat tekintve a TIP3P elavult, érdemes helyette a TIP3P-FB-t használni. Ezt az állítást az én eredményeim is megerősítik.

4.3.3. OPC, OPC3

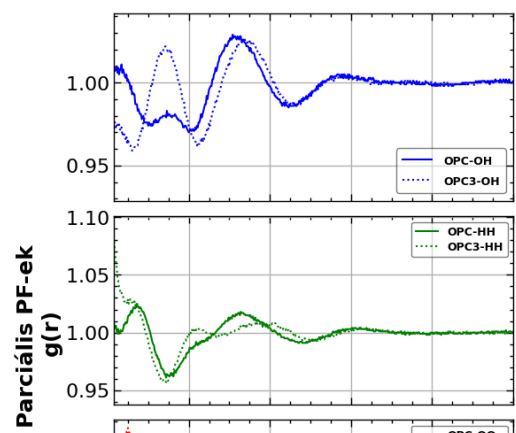


13. ábra. Az OPC és OPC3 modellek összehasonlítása

Amikor az OPC-t és az OPC3-at hasonlítjuk össze, akkor egy 4 és egy 3 site-os modellt vizsgálunk, így természetes, hogy jelentős eltérések lesznek. A két modellt felhasználási szempontból nem is lehet direktben összehasonlítani. Viszont a 3 site-os és 4 site-os modellek összehasonlítására tökéletesen alkalmasak. A modellek ugyanazon elv mentén készültek, és a szerzőik is megegyeznek. Így tehát a legfőbb tényező, amiben különböznek, az a virtuális részecske, amit az OPC tartalmaz, az OPC3 nem.

Az első OO kötéstávból látható (13. ábra), hogy a molekulák egymástól nagyjából ugyanolyan távolságra találhatók. Ez várható volt, hiszen a LJ paraméterek távolságot szabályzó σ értékei a két modellnél szinte teljesen megegyeznek. Ami érdekes, hogy ennek ellenére az OPC3 esetében az első intermolekuláris OH csúcs távolsága kisebb, mint az OPC-nél. Ez a csúcs az első intermolekuláris kötéseknek felel meg, azaz az OPC3 hidrogénkötés hossza kisebb, mint az OPC-nek. Ez a különbség, bár kisebb mértékben, de megfigyelhető a második HH-csúcsoknál is.

Az OO kötéstávok hasonlósága miatt nem meglepő, hogy az XRD-vel összehasonlítva hasonló R-faktor adódik mindkét modell esetében. Viszont, ha összevetjük az OPC intermolekuláris OH kötéstávolságát a többi modellel, akkor lát-



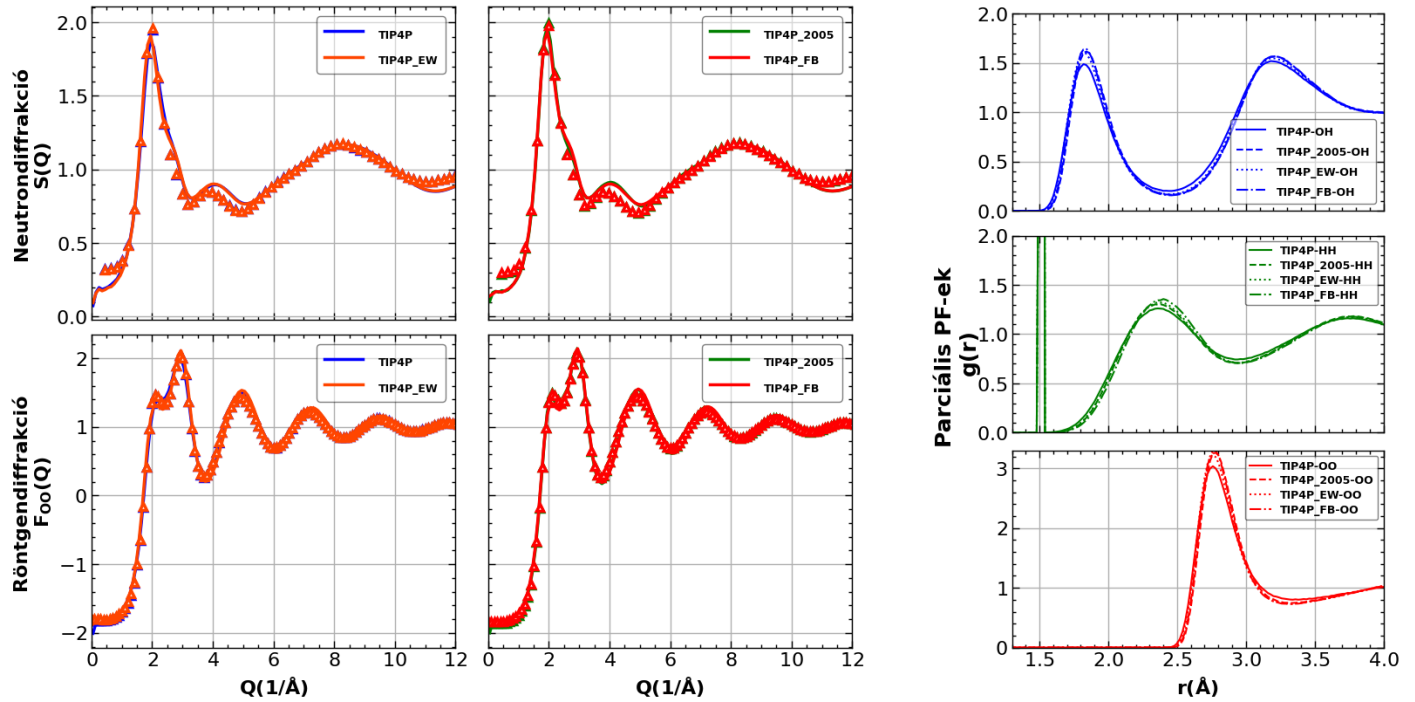
hatjuk, hogy az OPC esetében ez a távolság nagyjából 0,2 Å-mel nagyobb, mint a többi modell esetében.

Ez az eltérés az 4-14 Å tartományon még látványosabb. Az OO csúcsok távolsága nagyjából megegyezik, de az OH és HH csúcsok között fázis és amplitúdóbeli eltérés van. Ez az eltérés a többi modelltől meglátszik az OPC illeszkedésében: ha az ND adattal hasonlítjuk össze, akkor az összes modell közül a legrosszabb az illeszkedése.

Ez persze nem jelenti, hogy ez lenne a legrosszabb modell az összes közül, mert fontos nem elfelejteni, hogy a szerkezet csak egy szempont.

Az OPC készítésekor jól látszik, hogy nem a szerkezet volt az elsődleges szempont [17], így nem meglepő, ha ilyen téren a modell rosszabbul teljesít. Az OPC3 viszont a 3 site-os modellek közül a legjobb illeszkedést mutatja az ND adattal. Ha a szimuláció komplexitása fontos szempont, a modell jól használható szerkezetvizsgálatra.

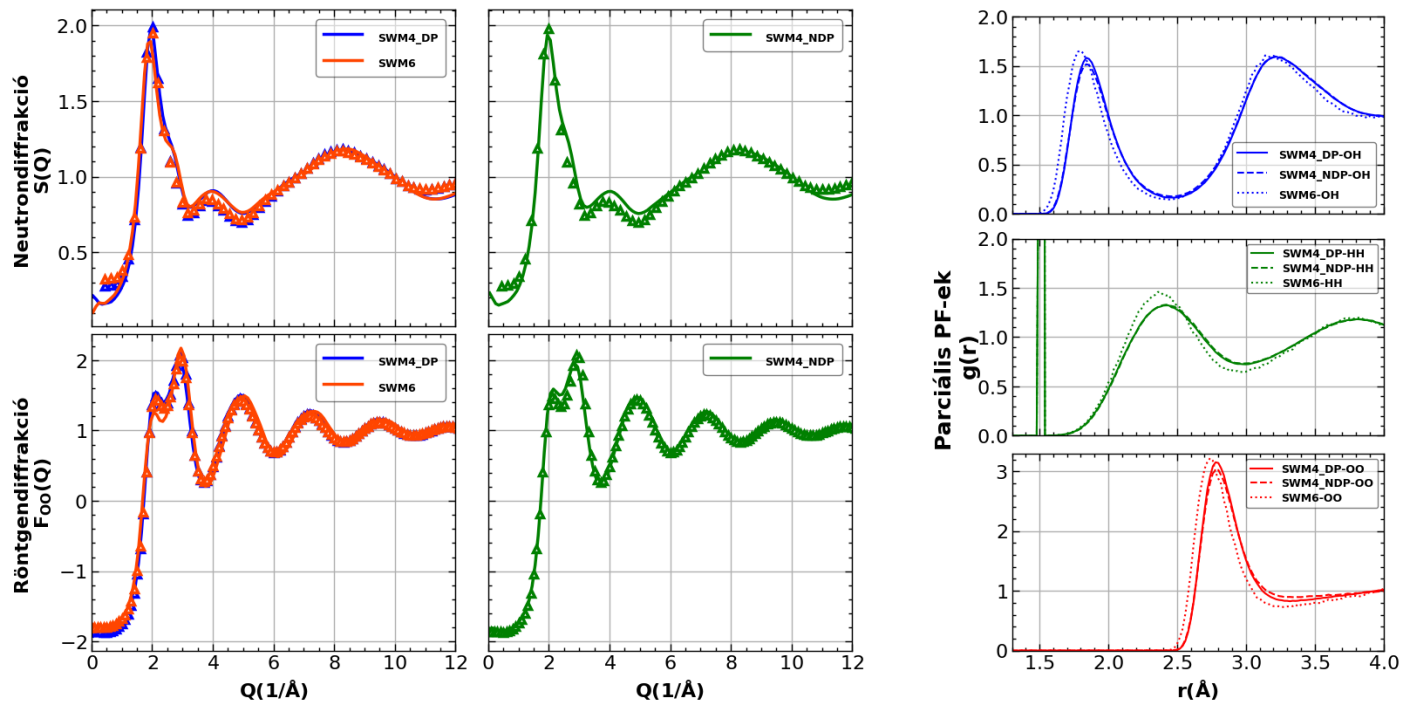
4.3.4. TIP4P alapú modellek



15. ábra. Négy TIP4P-típusú modell összehasonlítása

A TIP4P modell mintájára több modell is készült. Az általam vizsgált TIP4P, TIP4P-EW, TIP4P-2005 és TIP4P-FB modellekben a HOH kötőszög és az OH intramolekuláris távolság azonos. Az eltérés a virtuális részecske oxigénatomtól mért távolságában és a 4 részecske töltéseiben van. Nem meglepő tehát, hogy a pf-ek között csak nagyon minimális különbség van. Ez az illeszkedésre is elmondható, ami nem meglepő eredmény, tekintve, hogy hasonló pf-ekből számolt szerkezeti függvények is hasonlóak. Összességében a 4 modell közül a TIP4P-2005 illeszkedése a legjobb. Viszont, annak függvényében, hogy szerkezet mellett más tulajdonságok pontos leírása is fontos-e számunkra, az optimális választás eltérő lehet.

4.3.5. Polarizálható modellek



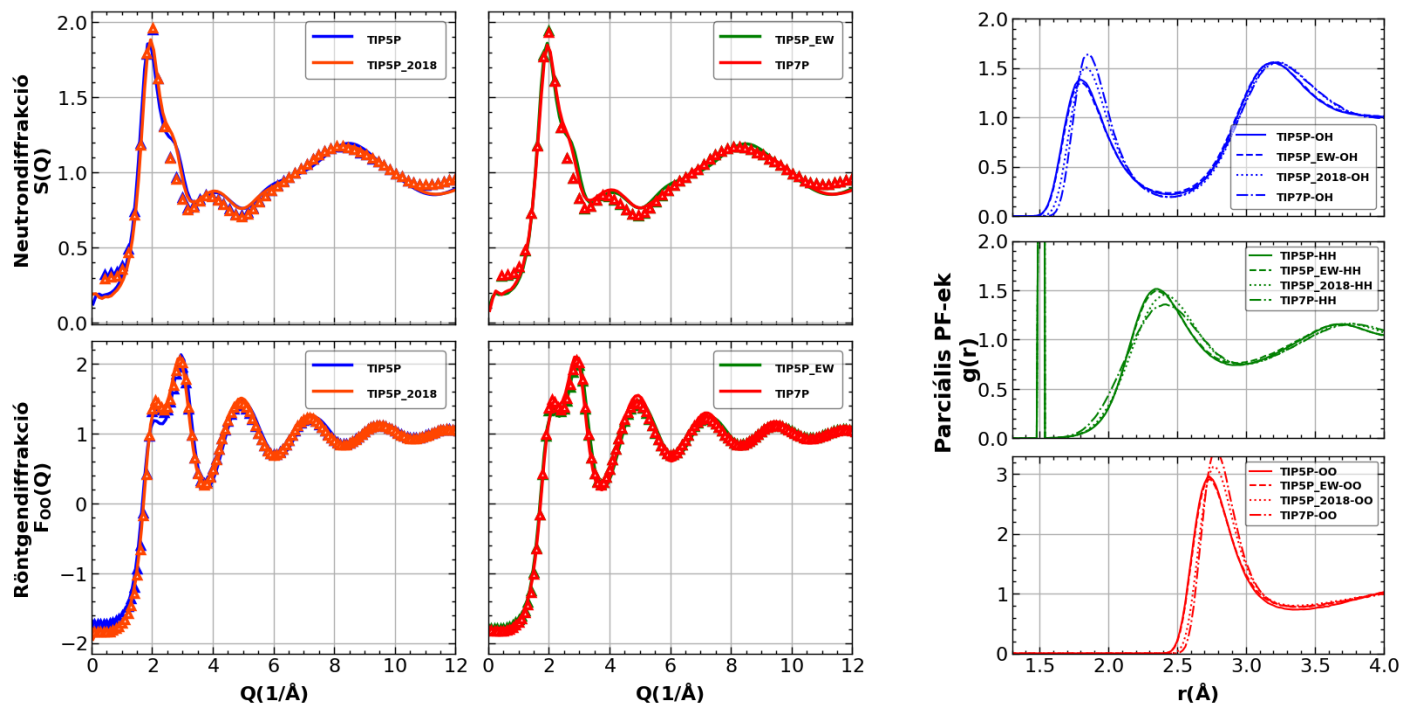
16. ábra. Az általam vizsgált polarizálható modellek összehasonlítása

Bár az SWM6-os modell, mind újabb, mind komplexebb, mint az SWM4-es modellek, mégis az ND adathoz hasonlóan, de az XRD adathoz kevésbé illeszkedik, mint elődei. Ez nem várt eredmény, tekintve, hogy a modell kifejezetten a víz folyékony fázisának a pontosabb leírásával készült, a SWM4-NDP-ből kiindulva [26].

Az SWM6 a hidrogén kötések pontosabb modellezése érdekében két LP-site-ban tér el az SWM4-es modellektől. Mégis, ez a fejlesztés nem jelenik meg az R-faktorokban, nem sikerült javulást elérni az ND adathoz való illeszkedésben. Az SWM4-DP R-faktora például kevesebb, mint fele az SWM6-énak, az XRD adat mindhárom vizsgált tartományán.

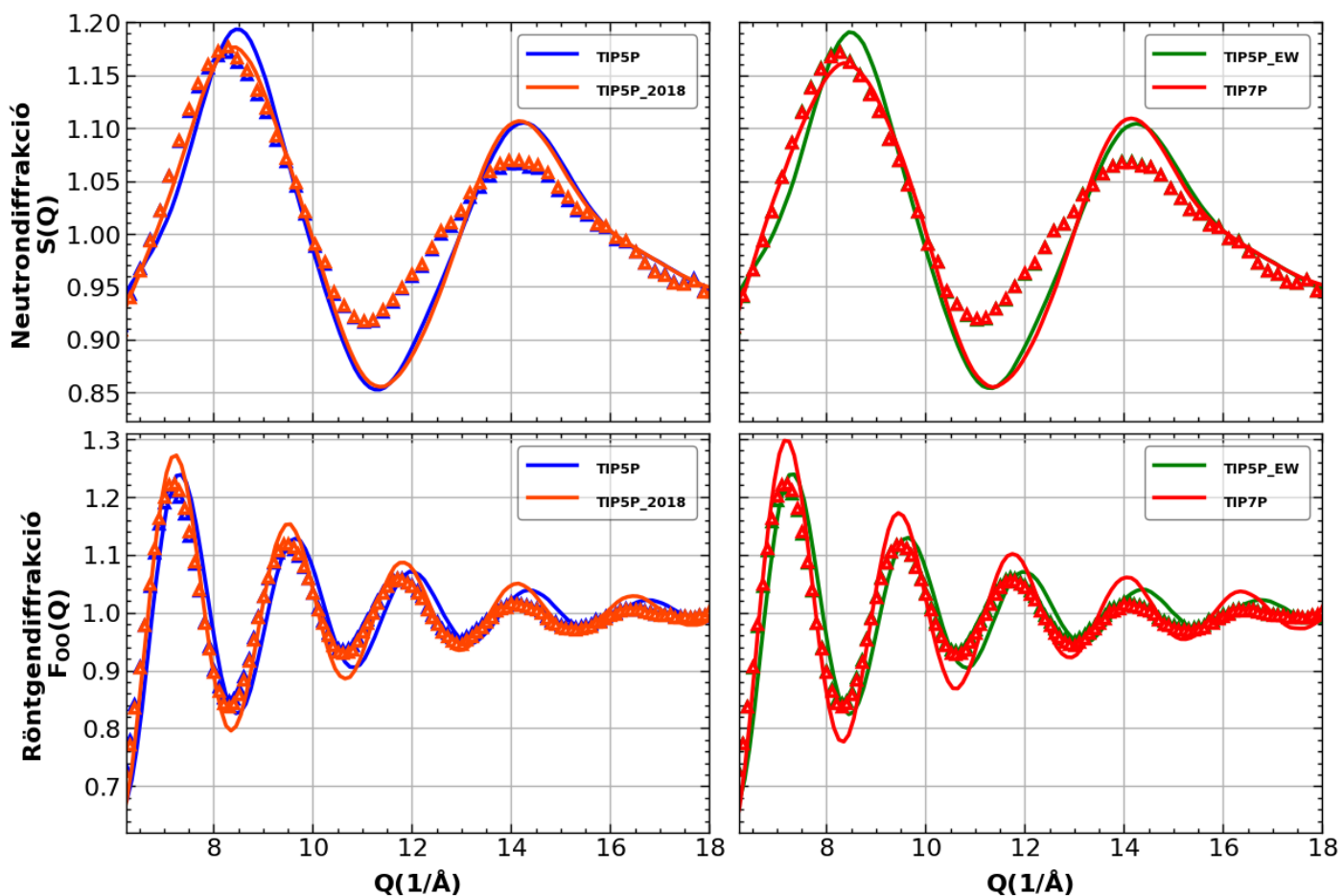
Ami hasonlóképp meglepő eredmény, hogy az SWM4-DP modell illeszkedése az SWM4-NDP-nél is nem elhanyagolhatóan jobb. Az SWM4-NDP lényegében ugyanaz, mint az eredeti SWM4-DP, csak a paraméterek újra lettek optimalizálva negatív shell részecske töltéshez. Az illeszkedésbeli eltérés a 16. ábrán is nyomon követhető. Az SWM6 esetében sem az amplitúdót, sem a fázist nem sikerült reprodukálnia a modellnek. Az SWM4-NDP fázisa nagyjából egyezik az XRD adattal, de az amplitúdó az első csúcsnál nem ugyanaz, mint a kísérleti görbéken, sem XRD-nél, sem ND-nél. Az SWM4-DP viszont a fázist és az amplitúdót is jól visszaadja.

4.3.6. TIP5P, TIP5P-EW, TIP5P-2018, TIP7P



17. ábra. Az általam vizsgált 5 és 7 részecskéből álló vízmodellek összehasonlítása

Az 5 és 7 részecskéből álló modellek párkorrelációs függvényeit megnézve látszik, hogy a paraméterekben nem a TIP5P-típusú és a TIP7P között van különbség, hanem az újabb TIP7P és TIP5P-2018 kötéstávjai térnek el a TIP5P és a TIP5P-EW-től. Az illeszkedések alapján (3. táblázat) az újabb modellek szerkezeti szempontból pontosabbak. A 18. ábrán látható, hogy a régebbi modelleknél nem csak az amplitúdó, hanem a fázis sem egyezik a kísérleti görbékkel, az újabb adatok fázisa viszont nagyrészt azonos a kísérletével. Amplitúdóbeli eltérés itt is megfigyelhető, mind az ND, mind az XRD görbékhez képest.



18. ábra. Szerkezeti függvények nagyobb Q értékeken ábrázolva

A TIP5P-2018 újraparaméterezése a TIP5P modellnek, így mivel jobban visszaadja a szerkezetet, mint az elődje és a TIP5PEW, így szerkezetvizsgálatra a TIP5P-2018-at érdemes használni. A 18. táblázatból kiolvasható, hogy a TIP7P számítási igénye még fele annyi szimulált molekula esetén is nagyobb, mint a TIP5P modelleké, elvégre több részecske közötti kölcsönhatást kell kiszámítani molekulánként. Mindemellett a TIP7P neutrondiffrakciós R-faktora nem sokkal jobb, mint a TIP5P-2018 modellé, a röntgendiffrakciós adathoz az illeszkedése pedig 1,5%-kal rosszabb. Amennyiben nem prioritás számunkra a TIP7P egyik olyan egyéb tulajdonsága, amelyben közelebb áll a valósághoz, mint a TIP5P-2018, az extra komplexitás nem éri meg, szerkezetvizsgálatra a TIP5P-2018 alkalmasabb.

5. Összefoglalás

A munkám során klasszikus molekuladinamikai szimulációkat futtatam, 18 vízmodell elemzéséhez. A szimulációkból kinyert trajektóriákból meghatároztam a modellekre jellemző szerkezeti függvényeket, amelyeket kísérleti adattal összevetve, számszerűsítettem a modellek illeszkedését a kísérlethez.

A 18 vizsgált vízmodell közül szerkezetvizsgálati szempontból több modell is jól használható, nincs egyértelmű optimális választás. Több modell közül azt érdemes kiválasztani, amely éppen a céljainknak legjobban megfelel. A 3 részecskéből álló modellek nagyon egyszerűek, így a számítási igényük az összes modell közül a legkisebb. Olyan rendszerekben, ahol kénytelenek vagyunk nagyobb kiterjedésű és több részecskéből álló szimulációs dobozzal dolgozni, az egyszerűbb modellek használata nagyon előnyös. Például, a több száz atomból álló fehérjemolekulák esetében nem tudjuk elkerülni, hogy sok vízmolekulát használjunk, hiszen, ha a fehérjét nem veszi teljesen körül az oldószere, akkor az megmászhatja az eredményt. Ilyen rendszereknél az OPC3 modell használata a javasolt, amely a 21,88%-os neutrodiffrakciós és 6,89%-os röntgendiffrakciós R-faktoraival a 3 részecskés modellek közül a legpontosabb.

Kisebb molekulák, esetleg tiszta víz rendszerek szerkezetvizsgálatakor, ahol nem kell olyan nagy dobozt szimulálni és a számítási igény kisebb szempont, rendelkezésre állnak pontosabb modellek is, mint a legegyszerűbb modellek. A legnagyobb változást az illeszkedésben a dipólmomentumot reprezentáló, csak töltéssel rendelkező virtuális részecske jelentette. Ezeknek a 4 részecskéből álló modelleknek mindkét típusú kísérleti adathoz való illeszkedése jobb, mint a 3 részecskés modelleknek. Az OPC-n kívül bármelyik ilyen modell jól használható szerkezetvizsgálathoz, annak függvényében, hogy milyen egyéb tulajdonságok pontos szimulálása fontos számunkra. Összességében a TIP4P-2005 illeszkedése a legjobb.

Ha megnézzük az 5 részecskéből álló modelleket, megállapítható, hogy a két nemkötő elektronpár virtuális részecskével való modellezésével jelentős javulás figyelhető meg a 3 részecskés modellekhez képest. De az ilyen modellek illeszkedése nem olyan jó, mint a TIP4P-típusú modelleké. Ezért a 4 részecskés modellek egyszerűségüknél fogva jobban használhatók, mint az 5 részecskések.

Ha a 4 és 5 részecskés modelleknél is javult az illeszkedés, akkor felmerül a kérdés, hogy vajon a két modell ötvözte, egy 6 részecskés modell pontosabb lehet-e, mint az egyszerűbb modellek. Az SWM6 illeszkedése, amely ráadásul még polarizálható is, ezt nem támasztja alá, mert az R-faktorai nagyobbak, mint a TIP4P-típusú modelleké, de legalább az 5 részecskés modellekhez képest a neutrodiffrakciós adattal való illeszkedés valamelyest javul.

A TIP7P esetében hasonló a helyzet, bár az illeszkedése jó, de ha figyelembe vesszük a nagyobb komplexitását, nem biztos, hogy szerkezet szempontjából megéri a többlet számítási igény.

A polarizálható modelleket messze a legszámításigényesebb szimulálni, az oszcilláló rendszerek szimulációjához kellően finom időlépésre van szükség, különben a szimuláció a energetikailag instabillá válhat. Mindemellett a polarizáció egy nagyon fontos jelenség, amelyet nagy előny, hogy képesek va-

gyunk szimulálni. Érdekesmód a valósághűbb ábrázolása a molekuláknak, az általam vizsgált modellek esetében nem jelent pontosabb szerkezeti modellt. Ha a polarizálható mellett a négy nem polarizálható részecskéből álló SWM4-DP és SWM4-NDP modelleket nézzük, az R-faktoraik összemérhetőek a nem polarizálható 4 részecskés modellekével. A polarizálhatóság láthatóan ezeknél a modelleknél nem változtat érdemileg az illeszkedés minőségén. Ugyanez a helyzet a korábban már említett SWM6-tal is. Itt nem lehet egyértelműen megmondani, hogy csak az adott modellek paraméterezése volt szerkezeti pontatlan, vagy a polarizáció tényleg nem vezet pontosabb szerkezethez. Ehhez további polarizálható modelleket kellene megvizsgálni.

5.1. Kitekintés

A dolgozatom célja az egyes modellek összehasonlítása mellett az is volt, hogy megállapítsam, van-e javuló tendencia a modellek szerkezeti információtartalmát illetően. Bár összességében ilyen javulás nem figyelhető meg, fontos tudni, hogy a modernebb modellek többsége nem szerkezetvizsgálathoz lett paraméterezve. Ha ezt figyelembe vesszük, az a tény, hogy az újabb modellek R-faktora nem romlott a régiekhez képest, de emellett sok más fizikai mennyiséget pontosabban visszaad, véleményem szerint jelentős előrelépés.

Ezek ellenére felmerül a kérdés, hogy a klasszikus molekuladinamika eszköztárával képesek vagyunk-e még pontosabb modelleket előállítani szerkezeti szempontból is.

Bár elsőre úgy tűnhet, hogy nem, de figyelembe kell venni, hogy még sok más típusú modell létezik, amit lehet vizsgálni. Például léteznek olyan polarizálható vízmodellek, amelynek minden atomjához tartozik egy polarizálható részecske, egy híres példa erre a BK3 [34]. Ezenkívül létezik három részecskés modell is, amelyhez tartozik egy extra polarizálható részecske. Az általam eddig vizsgált összes modell a Lennard-Jones effektív párpotenciálon alapul, lehetséges, hogy ez a limitáló tényező, lehet használni bonyolultabb potenciálokat is. A BK3 például már nem LJ potenciálon alapul. Ezeket a bonyolultabb modelleket egy logikus következő lépés lenne megvizsgálni. Emellett az eddig megszerzett tudás a különböző vízmodellek alkalmazhatóságáról szerkezetvizgálatra lehetővé teszi számomra többkomponensű rendszerek vizsgálatát is, ahol a víz valami bonyolultabb molekulát vesz körül.

Jelenleg a tudományágat limitáló legfőbb tényező a számítási kapacitás, ami viszont évről-évre rohamosan növekszik, ezt párosítva a mesterséges intelligencián alapuló modellekkel, valószínű, hogy a jövőben még pontosabb, akár klasszikus vízmodelleket is képesek leszünk alkotni.

Hivatkozások

- [1] Lawrie B. Skinner. “Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range”. *The Journal of Chemical Physics* 138 (2013), 74506. old. DOI: 10.1063/1.4790861.
- [2] Alan K. Soper. “Disordered Atom Molecular Potential for Water Parameterized against Neutron Diffraction Data. Application to the Structure of Ice Ih”. *The Journal of Physical Chemistry B* 119 (2015), 9244–9253. old. DOI: 10.1021/jp509909w.
- [3] Yuriy Khalak, Björn Baumeier és Mikko Karttunen. “Improved general-purpose five-point model for water: TIP5P/2018”. *The Journal of Chemical Physics* 149 (2018), 224507. old. DOI: 10.1063/1.5070137.
- [4] J. L. F. Abascal és C. Vega. “A general purpose model for the condensed phases of water: TIP4P/2005”. *The Journal of Chemical Physics* 123 (2005), 234505. old. DOI: 10.1063/1.2121687.
- [5] Yong Wu, Harald L. Tepper és Gregory A. Voth. “Flexible simple point-charge water model with improved liquid state properties”. *Journal of Chemical Physics* 124 (2006), 24503. old.
- [6] Kevin A. T. Silverstein, Anthony D. J. Haymet és Ken A. Dill. “A simple model of water and the hydrophobic effect”. *Journal of the American Chemical Society* 120 (1998), 3166–3175. old.
- [7] Frank H. Stillinger és Aneesur Rahman. “Improved simulation of liquid water by molecular dynamics”. *The Journal of Chemical Physics* 60 (1974), 1545–1557. old.
- [8] H. J. C. Berendsen és tsai. “Intermolecular Forces”. *Intermolecular Forces*. Szerk. B. Pullman. Dordrecht: Reidel, 1981, 331. old.
- [9] William L. Jorgensen és tsai. “Comparison of simple potential functions for simulating liquid water”. *The Journal of Chemical Physics* 79 (1983), 926. old.
- [10] Michael W. Mahoney és William L. Jorgensen. “A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions”. *The Journal of Chemical Physics* 112 (2000), 8910–8922. old.
- [11] Guillaume Lamoureux és tsai. “A polarizable model of water for molecular dynamics simulations of biomolecules”. *Chemical Physics Letters* 418 (2005), 241–245. old.

- [12] Mark James Abraham és tsai. “GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers”. *SoftwareX* 1 (2015), 19–25. old. DOI: 10.1016/j.softx.2015.06.001.
- [13] Juan C. Meza (if available). *Osti.gov*. Accessed on 2023. 10. 22. URL: <https://www.osti.gov/servlets/purl/983240>.
- [14] Shuichi Miyamoto és Peter A. Kollman. “Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models”. *Journal of Computational Chemistry* 13 (1992), 952–962. old.
- [15] *Periodic boundary conditions illustration*. Accessed on: 2023. 10. 25. URL: https://en.wikipedia.org/wiki/File:Periodic_boundary_conditions.gif.
- [16] Tom Darden, Darrin York és Lee Pedersen. “Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems”. *The Journal of Chemical Physics* 98 (1993), 10089–10092. old.
- [17] Saeed Izadi, Ramu Anandakrishnan és Alexey V. Onufriev. “Building Water Models: A Different Approach”. *The Journal of Physical Chemistry Letters* 5 (2014), 3863–3871. old.
- [18] H. J. C. Berendsen, J. R. Grigera és T. P. Straatsma. “The missing term in effective pair potentials”. *The Journal of Physical Chemistry* 91 (1987), 6269–6271. old.
- [19] Lee-Ping Wang, Todd J. Martinez és Vijay S. Pande. “Building Force Fields: An Automatic, Systematic, and Reproducible Approach”. *The Journal of Physical Chemistry Letters* 5 (2014), 1885–1891. old. DOI: 10.1021/jz500737m.
- [20] Saeed Izadi és Alexey V. Onufriev. “Accuracy limit of rigid 3-point water models”. *The Journal of Chemical Physics* 145 (2016), 74501. old.
- [21] *Paint3D*. Computer software. Accessed on September 27, 2023. Microsoft Corporation, 2023. URL: <https://apps.microsoft.com/detail/paint-3d/9NBLGGH5FV99?hl=en-us&gl=US>.
- [22] William L. Jorgensen és Jeffry D. Madura. “Temperature and size dependence for Monte Carlo simulations of TIP4P water”. *Molecular Physics* 56 (1985), 1381–1392. old.
- [23] Hans W. Horn és tsai. “Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew”. *The Journal of Chemical Physics* 120 (2004), 9665–9678. old.
- [24] Steven W. Rick. “A reoptimization of the five-site water potential (TIP5P) for use with Ewald sums”. *The Journal of Chemical Physics* 120 (2004), 6085–6093. old. DOI: 10.1063/1.1652434.

- [25] Guillaume Lamoureux, Alexander D. MacKerell Jr és Benoît Roux. “A simple polarizable model of water based on classical Drude oscillators”. *Journal of Chemical Physics* 119 (2003), 5185–. old.
- [26] Wenbo Yu és tsai. “Six-site polarizable model of water based on the classical Drude oscillator”. *The Journal of Chemical Physics* 138 (2013), 34508. old. DOI: 10.1063/1.4774577.
- [27] Chong-Li Zhao és tsai. “Seven-Site Effective Pair Potential for Simulating Liquid Water”. *The Journal of Physical Chemistry B* 123 (2019), 4594–4603. old. DOI: 10.1021/acs.jpcb.9b03149.
- [28] Shuichi Nosé. “A molecular dynamics method for simulations in the canonical ensemble”. *Molecular Physics* 52 (1984), 255–268. old.
- [29] Shuichi Nosé és Michael L. Klein. “Constant pressure molecular dynamics for molecular systems”. *Molecular Physics* 50 (1983), 1055–1076. old.
- [30] Michele Parrinello és Abderrezak Rahman. “Polymorphic transitions in single crystals: A new molecular dynamics method”. *Journal of Applied Physics* 52 (1981), 7182–7190. old.
- [31] Erik Lindahl és tsai. *GROMACS 2020.5 Source code*. Open Software. 2020.5. verzió. Published on January 6, 2021. 2021. URL: [URL%20of%20the%20source%20code%20archive,%20e.g.,%20https://example.com/gromacs2020.5](https://example.com/gromacs2020.5).
- [32] Matplotlib Development Team. *Matplotlib*. <https://matplotlib.org/>. Accessed on October 28, 2023. 2023.
- [33] Shuzo Shibata és L. S. Bartell. “Electron-Diffraction Study of Water and Heavy Water”. *The Journal of Chemical Physics* 42 (1965), 1147–1151. old. DOI: 10.1063/1.1696094.
- [34] Péter T. Kiss és András Baranyai. “A systematic development of a polarizable potential of water”. *The Journal of Chemical Physics* 138 (2013), 204507. old. DOI: 10.1063/1.4807600.

Függelék – Saját eredmények

Felvetett kérdés a dolgozat során

A különböző vizsgált vízmodellek közül melyek milyen mértékben alkalmazhatók a szerkezeti jellemzők vizsgálatára, szobahőmérsékleten. Továbbá annak a megvizsgálása, hogy a komplexebb, nagyobb számításigényű vagy modernebb modellek jobbak-e szerkezetvizsgálatra, mint a régebben készültek.

Módszertan

A vízmodelleket a GROMACS klasszikus molekuladinamikai szimuláció szoftver segítségével vizsgáltam. Az általam vizsgált modellekből számolt neutron-és röntgendiffrakciós szerkezeti függvényeket számoltam, amelyeket összevetettem kísérleti eredményekkel [1][2], szobahőmérsékleten.

Kapott eredmények

A modellek illeszkedését a 3 ; 4 táblázatokban foglaltam össze, a futásidőről pedig a 5 táblázat tartalmaz információt. Összességében 3 modellt emelnék ki, amelyeket jól alkalmazhatónak tartok szerkezetvizsgálatra, OPC3-at, SWM4-DP-t és a TIP4P-2005-öt. Az OPC3, egy egyszerű modell, amiatt a három modell közül messze a leghatékonyabb szimulálni, nagy rendszerek esetén jó választás lehet szerkezetvizsgálatra. Ha a számítási kapacitás nem szempont, akkor az SWM4-DP is jó választás, ez adja vissza a legjobban a kísérleti röntgendiffrakciós adatahoz. Összességében viszont a TIP4P-2005 illeszkedik a legjobban a kísérleti adatokhoz és a számításigénye még nem túl nagy, nagyjából az OPC3 kétszerese.