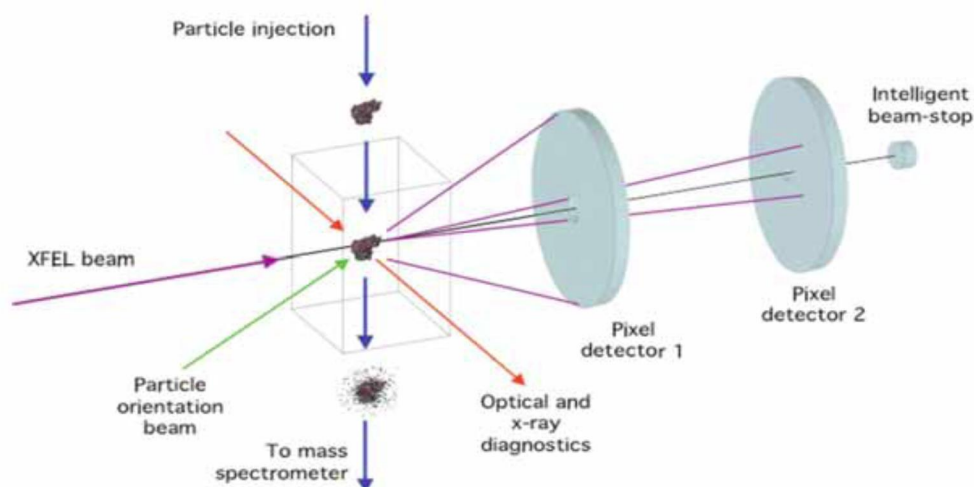


Ultragyors diffrakciós egyrészecske leképezés (2010-2015)

Bevezetés

A szilárd anyagok, molekulák és élő szervezetek részletes szerkezetének megértése elengedhetetlen a tudomány és az ipar számára az anyagtudománytól a mérnöki alkalmazásokig, a biológiától és az orvostudománytól a biotechnológiáig és a gyógyszeriparig. Atomi szintű szerkezeti információt leggyakrabban kristályokon végzett (röntgen- elektron- és neutron-) diffrakciós kísérletekből nyerhetünk. Sajnos nem minden anyag vagy részecske kristályosítható. Például sok fehérje ilyen, beleértve majdnem az összes sejtmembrán-fehérjét.

Diffrakciós kísérletek egyetlen részecskén elvben szolgáltatathatnának elegendő információt a részecske atomi szerkezetéről. Sajnos általában a sugárkárosodás (amely még a kristályoknál is komoly problémát okoz) tönkreteszi a mintát, mielőtt lemérhetnénk a diffrakciós képet. Először Solem [1986], majd Neutze és munkatársai [2000] javasoltak egy megoldást a problémára. Javaslatuk lényege, hogy egy nagyon rövid (10-100 fs) és nagyon erős ($>10^{12}$ foton/impulzus) röntgenimpulzust használva még azelőtt lehet használható információt szerezni a minta szerkezetéről, mielőtt a károsodás bekövetkezne. A méréshez szükséges impulzusokat a közelmúltban kifejlesztett röntgen szabadelektron lézerek (XFEL) szolgáltatják. A mérés vázlatja az 1. ábrán látható. A részecskéket egy injektor juttatja a röntgennyaládba, a diffrakciós képeket 2D detektorok mérik. Azok a diffrakciós képek használhatók, ahol a röntgenimpulzus csak egy részecskét talál el. Az eltalált részecske a röntgenimpulzus hatására ionizálódik és felrobban, de ha a röntgenimpulzus elég rövid [Jurek, Faigel & Tegze 2004], akkor a diffrakciós kép még a részecske eredeti szerkezetéről ad információt. Egy diffrakciós kép általában nem tartalmaz elég fotont a részecske szerkezetének meghatározásához, ezért többnyire sok ezer, esetleg millió képet kell azonos részecskékről készíteni. A nyaládba belőtt részecskék orientációja tetszőleges lehet, a diffrakciós képek feldolgozásának első és talán legnehezebb feladata az ismeretlen orientációk meghatározása. Pályázatunk ennek a nehéz problémának a megoldására fókuszált.



1. ábra. Egyrészecske diffrakciós kísérlet vázlatja [European XFEL Technical Design Report]

Elméleti munka és eredmények

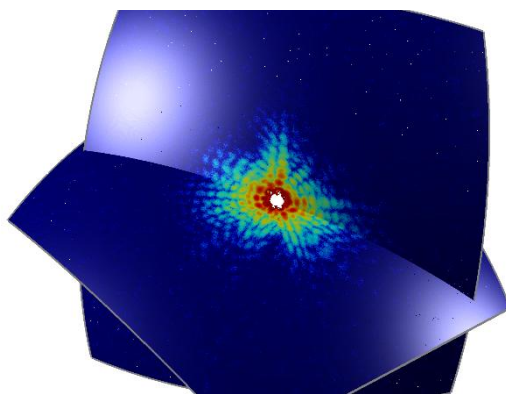
Az első fő feladatunk az volt, hogy áttekintsük a különböző orientációs módszereket, kiválasszuk a célnak legmegfelelőbbeket és kipróbáljuk azokat egyszerűbb adatokon és alacsonyabb dimenzióban. Az orientációs módszerek a diffrakciós képek két tulajdonságán alapulhatnak: 1. A képek a háromdimenziós térben metszik egymást, ezért van közös részük. 2. Az egymáshoz közeli orientációjú képek hasonlóak. A munkatervben felsorolt módszerek közül viszonylag hamar elvetettük az önszervező térképen (és a hasonló sokdimenziós térbeágyazási technikákon) alapuló eljárásokat. Ezek csak a 2. tulajdonságot használják ki, ezen kívül viszonylag bonyolultak és nagy a számításigényük. Így két irány maradt: a közös vonalon alapuló és az iteratív módszerek. Mivel mindkét irány jó eredményeket ígért, ezért párhuzamosan dolgoztunk egy-egy ilyen módszer kifejlesztésén.

Közös ív módszer

A kétdimenziós diffrakciós képek a háromdimenziós reciprok térben gömbsüvegeken helyezkednek el úgy, hogy a képek középpontja (a nulla szórási szöghöz tartozó pont) egybeesik a reciprok tér origójával. A képek metszésvonalai a reciprok térben körívek. A feladat az, hogy megtaláljuk a képpárok közötti közös ívet, és ez alapján meghatározzuk a két kép relatív orientációját. A relatív orientációk alapján felépíthető egy orientációs térkép, ami megadja, hogy egy diffrakciós kép hogyan helyezkedik el a reciprok térben. Shneerson és munkatársai [2008] már kidolgoztak egy eljárást arra az esetre, amikor a közös ív egy egyenessel közelíthető, azonban ilyenkor a relatív orientáció három Euler-szögéből csak kettő határozható meg.

Mi egy olyan módszert fejlesztettünk ki, amely nem igényli az említett közelítést és a közös ív görbületét használja fel a képpárok relatív szögének meghatározására. Az algoritmus a relatív szögek alapján meghatározza a képek egymás közötti orientációját, majd azt képhármasok közötti összefüggések segítségével tovább finomítja. Az orientációk meghatározását jelentősen bonyolítja, ha a részecske valamilyen szimmetriával rendelkezik. Ezért az algoritmust továbbfejlesztettük úgy, hogy minden előzetes ismeret nélkül felismerje és meghatározza a részecske szimmetriáját és azt figyelembe vegye az orientációk meghatározásánál.

A módszer működését két részecskén teszteltük. Az egyik egy vírusmodell volt, amely belső szerkezettel és pszeudoszimmetriával is rendelkezett. Ez utóbbi jelentősen nehezíti az orientációk meghatározását. A másik egy nagyméretű szimmetrikus fehérjemolekula (RuBisCo) volt. Minkét esetben sikerült a diffrakciós képek orientációját nagy (néhány tized fok) pontossággal meghatározni (2. ábra). Eredményeink az „Acta Crystallographica Section A: Foundations of Crystallography” folyóiratban jelentek meg [Bortel & Tegze, 2011].



2. ábra. A vírusmodell két diffrakciós képe, ahogy a közös ív módszer meghatározta a relatív orientációjukat. Jól látható a tökéletes egyezés a metszésvonal mentén.

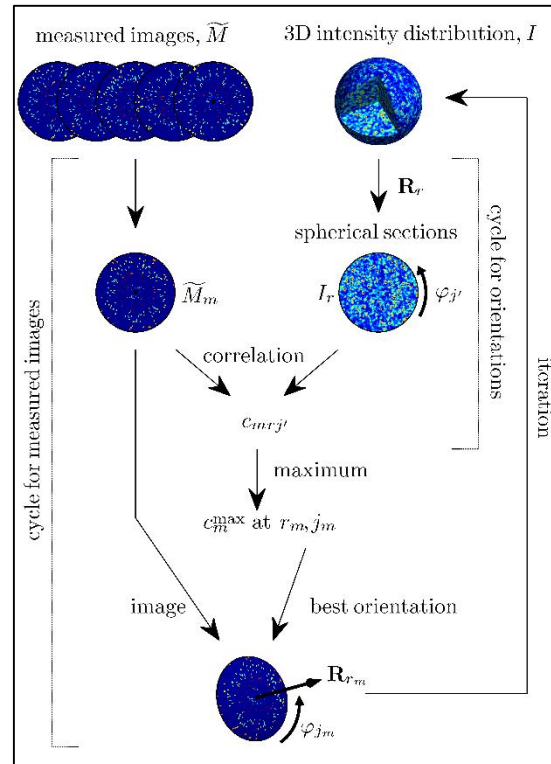
A közös ív módszer megbízhatóan működött azokban az esetekben, amikor az zaj nem volt nagyon nagy. Nagyon nagy zaj esetén viszont a közös ívben található viszonylag kevés képpont már nem tartalmaz elegendő információt a relatív orientáció meghatározására. Ezért szükség volt egy olyan algoritmusra is, amely egyszerre veszi figyelembe az összes diffrakció képen található teljes információt.

Korreláció maximalizáló (CM) módszer

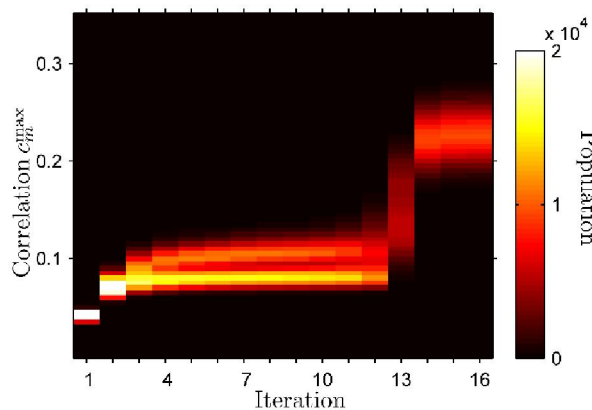
Az iterációs orientációs módszerek lényege, hogy egy (kezdetben tetszőleges) háromdimenziós eloszláshoz hasonlítjuk az egyes diffrakciós képeket, és azokban az orientációkban ahol jó az egyezés, hozzákeverjük. Az így feljavított eloszlással pedig megismételjük az eljárást a következő iterációban. Jó esetben a lépésenként javított eloszlás a keresett eloszláshoz konvergál. Az irodalomban eddig egy ilyen jellegű módszer található [Loh & Elser, 2009], az azonban nagyon bonyolult, lassú és ezért alkalmatlan a röntgen szabadelektron lézerekkel előállított nagy mennyiségű (millió nagyságrendű) diffrakciós kép feldolgozásához.

A célunk egy viszonylag egyszerű és így gyors, de mégis megbízhatóan működő algoritmus kidolgozása volt. Az alapötlet az volt, hogy minden iterációban minden képet minden lehetséges orientációban hozzápróbálunk az előző iterációban kapott háromdimenziós eloszláshoz, és a képeket a legjobb illeszkedés orientációjában összerakva alkotunk egy teljesen új eloszlást. Ezt az elgondolást először kétdimenziós képek és 8 különböző orientáció (90 fokos forgatások és tükrözések) esetére próbáltuk ki. A módszert tovább teszteltük kétdimenziós képekből vett egydimenziós ívek orientációjának meghatározására. Ez a feladat az eredeti háromdimenziós probléma modellje eggyel alacsonyabb dimenzióban. Az eredmények azt mutatták, hogy elegendően nagyszámú „mérés” esetén a módszer gyorsan konvergál még erősen zajos adatok esetén is.

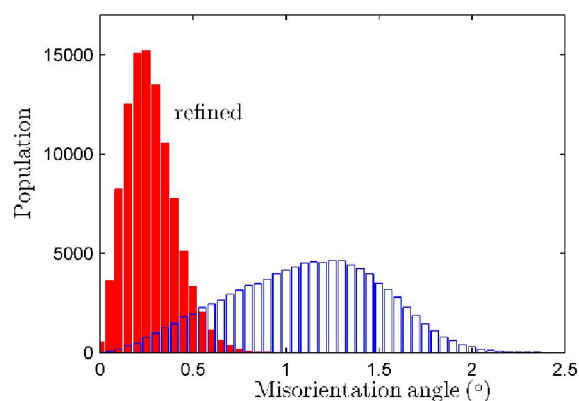
A következő lépés az iterációs módszerünk alkalmazása a háromdimenziós feladatra. Ahhoz, hogy az algoritmus nagyméretű biológiai molekulák többszázezer vagy millió nagyfelbontású diffrakciós képeinek orientálására is alkalmas legyen, az algoritmus legidőigényesebb részeit nagyfokú párhuzamosítással grafikus processzorokra, CUDA nyelven írtuk meg. Az algoritmus elvi vázlatát a 3. ábrán látható. A módszert sikeresen teszteltük egy nagy fehérjemolekula (NapAB) szimulált diffrakciós képein. Az algoritmus képes volt a 100000 erősen zajos diffrakciós kép nagy pontosságú orientációjára kizárólag a képekben található információk alapján (4. és 5. ábra). Úgy találtuk, hogy módszerünk nagyságrendekkel gyorsabb, mint a korábbi orientációs eljárások [Loh & Elser, 2009; Fung *et al.*, 2009] és a számítási idő is lassabban növekszik a molekula méretével. Megmutattuk, hogy az orientált képekből összeállított



3. ábra. A CM algoritmus elvi vázlatát

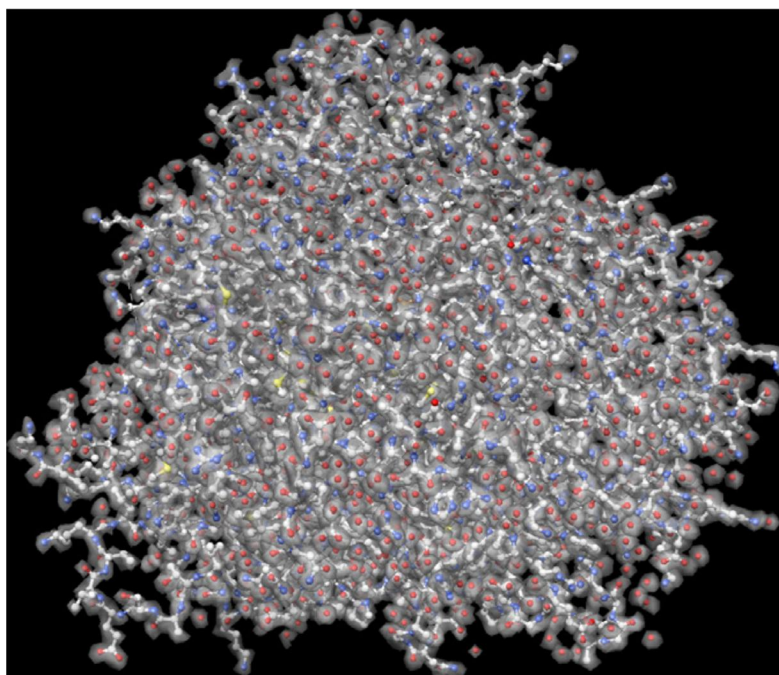


4. ábra. A korrelációk maximumainak fejlődése. A hirtelen ugrás jelzi a konvergenciát.



5. ábra. Az orientáció hibája a konvergencia elérése (kék) és további finomítás után (piros).

háromdimenziós intenzitás-eloszlásból meghatározható a molekula szerkezete atomi felbontással. Az ismeretlen fázisok meghatározásához a hibrid input-output algoritmust használtuk. (6. ábra)



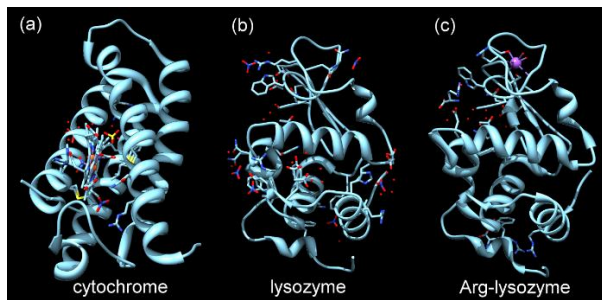
6. ábra. A NapAB fehérjemolekula orientált diffrakciós képeiből meghatározott elektronsűrűség-eloszlása. Jól látható az egyezés a színes gömbökkel jelölt atomi pozíciókkal.

Az eredményekből írt cikkünk a Journal of Structural Biology folyóiratban jelent meg [Tegze & Bortel, 2012]. A témából Tegze Miklós meghívott előadást tartott a Celsius-Linné Symposium-on (Uppsala, Svédország, 2012. február 16-17.) „Single molecule structure from noisy diffraction images” címmel.

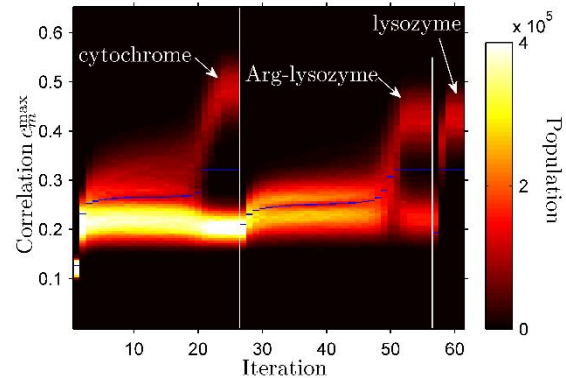
Az orientációs módszer továbbfejlesztése keverékek és szimmetrikus objektumok esetére

Megvizsgáltuk, hogy a fenti módszer hogyan tehető alkalmassá különböző részecskék diffrakciós képeinek a szétválogatására. Erre több okból lehet szükség: a biológiai objektumokat nem mindig könnyű szétválasztani, a fehérjemolekulák különféle konformációkat vehetnek föl, valamint szennyező anyagok is kerülhetnek a nyaládba.

Háromféle molekula (lizozim, citokróm és egy arginin-lizozim komplex) különböző arányú keverékeit vizsgáltuk. A három molekula hasonló méretű. A lizozim és az arginin-lizozim komplex szerkezete csak kissé tér el egymástól, a citokróm molekula szerkezete viszont lényegesen különböző. Megmutattuk, hogy a továbbfejlesztett algoritmus a korrelációs küszöb megfelelő megválasztásával nem csak a lizozim és citokróm, hanem az egymáshoz nagyon hasonló lizozim és arginin-lizozim molekulák (7. ábra) diffrakciós képei is szétválaszthatók a képek orientálásával egyidejűleg (8. ábra).



7. ábra. A citokróm, a lizozim és az arginin-lizozim molekulák szerkezete

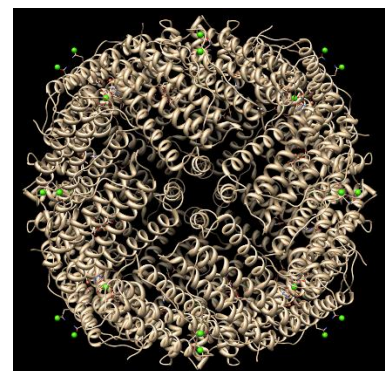


8. ábra. Az algoritmus működése a három molekula 1:1:1 arányú keverékére. A korreláció hirtelen megnövekedése az egyes komponenseknek megfelelő konzisztens 3D intenzitás-eloszlás megtalálását jelenti

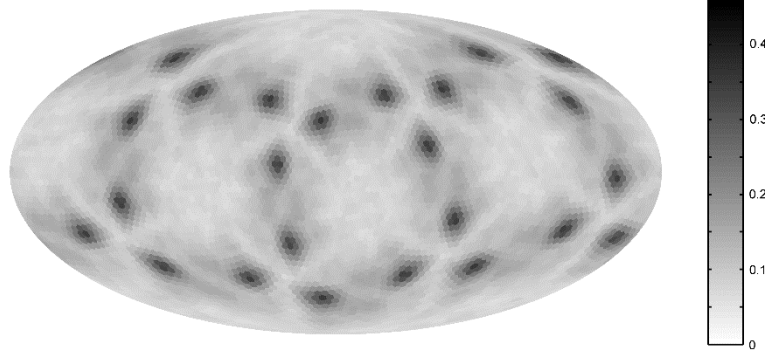
Az eredményekből írt cikkünk a Journal of Structural Biology folyóiratban jelent meg [Tegze & Bortel, 2013].

Az orientációs algoritmus részletes tesztelése során azt találtuk, hogy szimmetrikus vagy közel szimmetrikus objektumok esetében lassan, sőt egyes esetekben egyáltalán nem konvergál. Ennek oka az, hogy nem állapítható meg, hogy egy diffrakciós kép a szimmetria-ekvivalens helyzetek közül melyikhez tartozik. Így előfordulhat, hogy az algoritmus az orientációs térnek csak egy részébe helyez képeket, a többi része pedig üres marad.

Kidolgoztunk egy olyan algoritmust, amely szimmetrikus részecskék esetén is lehetővé teszi az orientációk gyors meghatározását. A módszer lényege, hogy a diffrakciós képeket nem csak egy orientációban, hanem – az egymáshoz közeli orientációk kizárásával – a szimmetriának megfelelő számban, több orientációban is hozzáadjuk a háromdimenziós intenzitás-eloszláshoz. Ehhez nincs szükség semmilyen *a priori* információra a szimmetriáról, a szimmetria az eredményül kapott eloszlásból utólag meghatározható. A módszert az ikozaédres szimmetriájú ferritin molekula (9. ábra) szimulált diffrakciós képein teszteltük. Azt találtuk, hogy míg az eredeti módszerünk nem volt képes az orientációk megtalálására, az új, módosított algoritmus nagy biztonsággal meghatározta az orientációkat még erősen zajos diffrakciós képek esetén is (10. ábra).



9. ábra. A ferritin molekula szerkezete.

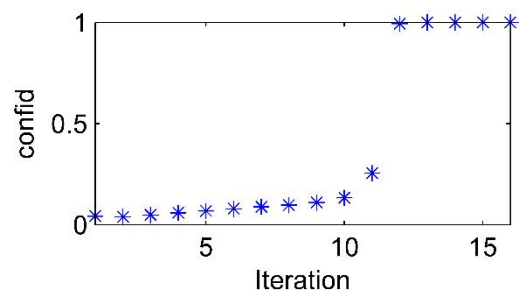


10. ábra. A korrelációtérkép a konvergencia elérése után. Jól látszik a szimmetriának megfelelő 24 egyforma maximum.

A közel szimmetrikus (pszeudo-szimmetrikus) objektumok esetében a fenténél is összetettebb eljárásra van szükség. Az iterációk során először úgy járunk el, mintha az objektum szimmetrikus lenne, vagyis a diffrakciós képeket sok, közelítőleg illeszkedő orientációban helyezük el. Ekkor egy majdnem szimmetrikus 3D intenzitás-eloszlást kapunk, ami viszont még nem felel meg a részecske szerkezetének, viszont jól kitölti a rendelkezésre álló teret. Ezután úgy folytatjuk az eljárást, mintha az objektumnak nem lenne szimmetriája, vagyis minden képet csak egy, (a legjobban illeszkedő) helyzetben használunk fel. Ilyenkor az eljárás már gyorsan konvergál a helyes megoldáshoz. Az eredmények publikálására a közeljövőben kerül sor.

Konfidencia paraméter

Az orientálás sikerességét szimulált diffrakciós képek esetén könnyű megállapítani. Valódi mérési adatok esetén ez már sokkal nehezebb feladat. Általában az iterációs eljárás során a korrelációk értékében bekövetkező ugrás jelzi a konvergenciát. Előfordul néha, hogy a konvergencia fokozatos, vagy olyan gyors, hogy beleolvad a kezdeti nagyobb változásba. Ilyenkor nem lehet egyértelműen megállapítani, hogy az orientálás sikeres volt. Ezért szükségesnek találtuk egy olyan paraméter meghatározását, amely egyértelműen jelzi a megtalált orientációk helyességét. Sikerült egy ilyen konfidencia paramétert találnunk, amely jelzi, ha a 3D intenzitás-eloszlás konzisztens a mérési adatokkal, azok statisztikus hibáin belül. A konfidencia paraméter megbízhatóságát szimulált adatokon ellenőriztük (11. ábra). Az eredmények publikálása folyamatban.



11. ábra. A konfidencia paraméter változása az iteráció függvényében

Orientálás a töredékek irányeloszlása alapján

Megvizsgáltunk egy teljesen új lehetőséget is a diffrakciós képek orientálására. Amikor a szabadelektron lézer nagyon intenzív röntgennyalábja eltalálja a molekulát, akkor az a nagyfokú ionizáció következtében fellépő taszító erők hatására felrobban. A szétrepülő molekulatöredékek, atomok detektálásával is kapható információ a részecske orientációjáról. Részletes, a szétrobbanó molekula ionjait és elektronjait is egyenként nyomon követő molekuladinamikai számításokkal megmutattuk, hogy a detektált részecskék szögeloszlásából meghatározható a molekula robbanás előtti orientációja. Az új, a diffrakciós képektől független

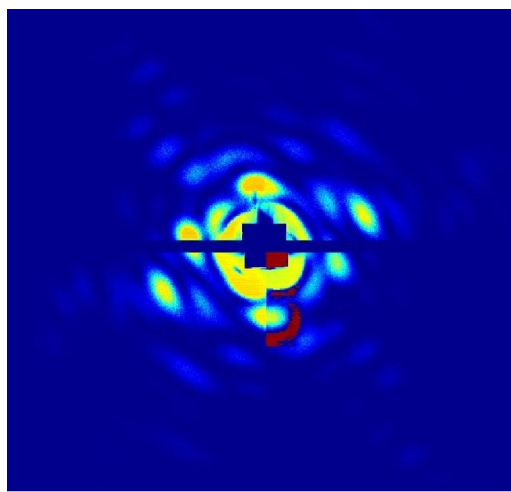
orientálási módszer lehetővé teheti az ultragyors egyrészecske diffrakciós módszer alkalmazását olyan esetekre is, amikor a diffrakciós képekből az alacsony jel/zaj miatt a részecskék orientációja nem határozható meg. Az eredményekből készült cikk az EPL folyóiratban jelent meg [Jurek & Faigel, 2013].

Kísérleti munka és előzetes eredmények

A pályázat utolsó évében két mérésorozatban is személyesen vettünk részt a stanfordi LCLS szabadelektron lézernél. Mindkét mérésorozat széles nemzetközi együttműködéssel szerveződött.

Vírusok szerkezetének meghatározása

Az első mérés kisméretű vírusok szerkezetének meghatározására irányult. A vírusokat az Uppsalai Egyetemen tervezett és épített injektor juttatta a röntgennyaládba. A röntgennyalábot



12. ábra. A trinity-vírusról készült egyrészecske diffrakciós kép

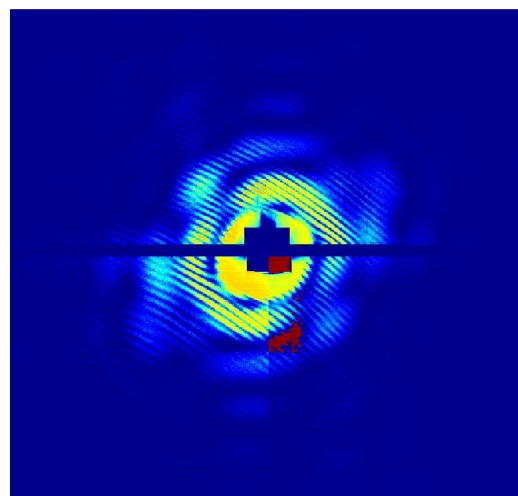
az eredeti tervek szerint 100 nm átmérőre kellett volna fókuszálni, de technikai problémák miatt ez csak kb. 500 nm-re sikerült. Ez 25-szörös intenzitáscsökkenést jelentett a tervekhez képest. A találati arány időnként elérte a néhány százalékot, ami (figyelembe véve az LCLS 120 Hz-es ismétlési frekvenciáját), óránként tízezres nagyságrendű diffrakciós képet eredményezett. A sok TB adat feldolgozása még folyamatban van, de az már most is látszik, hogy a kisebb röntgen-intenzitás miatt az adatok sokkal zajosabbak, mint azt eredetileg vártuk.

A második kísérlet fő célja a kétrészecske holográfia megvalósítása volt (ld. alább), de ennél a mérésnél is keletkezett nagyszámú (több százezer) diffrakciós felvétel egyetlen vírusról is. Mivel itt a fókuszolt mérete jobban megközelítette a kívánt értéket, a

diffrakciós képek is jobb minőségűek lettek. Az egyik vizsgált vírusról (trinity-vírus) készült diffrakciós képek (12. ábra) orientációját sikerült az általunk kifejlesztett CM módszerrel meghatározni és a kétdimenziós képekből létrehozni egy konzisztens háromdimenziós intenzitás-eloszlást. A fázisprobléma megoldásához és a vírus szerkezetének meghatározásához még további munka szükséges.

Kétrészecske holográfia kísérlet

Az LCLS-nél végzett második kísérlet egy teljesen új elven alapul. Ebben a kísérletben nem csak a vizsgálni kívánt részecskét (vírust) juttattuk a röntgennyaládba, hanem egyidejűleg egy ismert szerkezetű referenciárészecskét is. Esetünkben a referenciárészecske egy xenon klaszter volt, amit egy második, az előzőre és a röntgennyalábra merőlegesen elhelyezett injektor juttatott a találkozási pontba. A két részecskéről szórt



13. ábra. Trinity virus és xenon klaszter kétrészecske hologramja

röntgensugárzás interferenciát hoz létre, ami hologramként értelmezhető (13. ábra). Ez elvben lehetőséget ad arra, hogy egyetlen mérésből kapjunk 3D szerkezeti információt az ismeretlen részecske belsejéről. Mivel ilyen típusú méréseket ezelőtt még senki sem végzett, ezért itt az elvégzendő feladat nem csak a képek kiértékelése, hanem új kiértékelő algoritmusok és számítógépprogramok létrehozása. A szokásos holográfia esetén a tárgyról szórt nyaláb intenzitása kicsi a referencianyaláb intenzitásához képest. Itt nem ez a helyzet, ezért a szokásos holografikus kiértékelési módszerek nem alkalmazhatók. Ezen a problémán jelenleg is dolgozunk.

Számítástechnikai fejlesztések

Az orientációs probléma megoldása és a molekuladinamikai modellezés is nagyon nagy számítástechnikai kapacitást igényel. A számítások idejét párhuzamos programozással és nagyteljesítményű grafikus processzorok alkalmazásával csökkentettük elfogadható mértékűre. Ehhez először beszereztünk egy nagyteljesítményű grafikus processzorokkal felszerelt személyi számítógépet. A számítógép négy GTX 480 grafikus kártyát tartalmazott összesen 1920 processzormaggal. A rendszert később egy nyolc grafikus processzorkártya befogadására alkalmas GPU-expander egység beszerzésével fejlesztettük tovább. A régi grafikus kártyákat hat, a jelenlegi csúcstechnológiát képviselő GTX Titan Z kártyával váltottuk ki. Az így használható összesen 34560 processzormag jelentős teljesítménynövekedést jelentett.

Hivatkozások

Bortel G. & Tegze M. 2011 Common arc method for diffraction pattern orientation, *Acta Cryst.* **A67**, 533–543.

Fung R., Shneerson V., Saldin D.K. and Ourmazd A. 2008 Structure from fleeting illumination of faint spinning objects in flight, *Nature Phys.* **5**, 64–67.

Jurek Z., Faigel G. & Tegze M. 2004 Dynamics in a cluster under the influence of intense femtosecond hard X-ray pulses, *Eur. Phys. J. D* **29**, 217–229.

Jurek Z. & Faigel G. 2013 Orienting single-molecule diffraction patterns from XFELs using heavy-metal explosion fragments, *EPL* **101**, 16007.

Loh N.-T.D. & Elser V. 2009 Reconstruction algorithm for single-particle diffraction imaging experiments, *Phys. Rev. E* **80**, 026705.

Neutze R., Wouts R., van der Spoel D., Weckert E. & Hajdu J. 2000 Potential for biomolecular imaging with femtosecond X-ray pulses, *Nature* **406**, 752–757.

Shneerson V.L., Ourmazd A., Saldin D.K. 2008 Crystallography without crystals. I. The commonline method for assembling a three-dimensional diffraction volume from single-particle scattering, *Acta Cryst.* **A64**, 303–315.

Solem J.C. 1986 Imaging biological specimens with high-intensity soft x rays, *J. Opt. Soc. Am. B* **3**, 1551–1565.

Tegze M. & Bortel G. 2012 Atomic structure of a single large biomolecule from diffraction patterns of random orientations, *J. Struct. Biol.* **179**, 41–45.

Tegze M. & Bortel G. 2013 Selection and orientation of different particles in single particle imaging, *J. Struct. Biol.* **183**, 389–393.