

Reaktive kraftfelter

– en genvej til realistiske simuleringer af kemiske reaktioner?

I artiklen præsenteres det reaktive kraftfelt, ReaxFF, som en mulig vej mod mere realistiske simuleringer af kemiske reaktioner ved brug af begrænsede computerressourcer.

Af Peter Bjerre Jensen, Risø DTU og
Peter Fristrup, Institut for Kemi, DTU

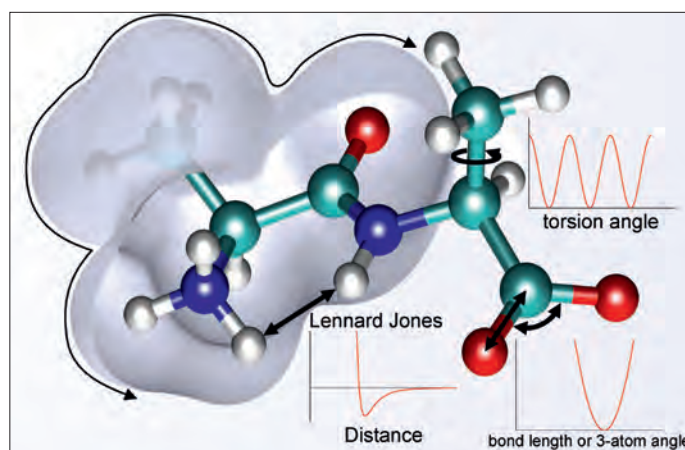
Modellering af kemiske systemer ved anvendelse af matematiske modeller er et vigtigt forskningsområde, som har mange fordele ift. "traditionel" eksperimentel kemi. Computere kan anvendes til at få detaljeret indsigt i, hvad der sker på atomart niveau, hvilket kan give ideer til optimeringer af kemiske reaktioner. Allerede nu anvendes computermodellering til analyse af f.eks. spektroskopiske data. I fremtiden vil det formodentligt også være hurtigere at lave computerekspirerter end "rigtige" eksperimenter – dette afhænger dog selvfølgelig af erfaring samt tilgængelig computerkraft. Eksperimenter vil dog altid være nødvendige, da man formodentligt ikke har en helt nøjagtig model. Dertil kommer, at syntesen af et givent stof som eksempelvis et medikament jo indlysende nok ikke kan erstattes af en computermodel.

Modellerne kan dog ofte bruges som et screeningsværktøj, der ved at hjælpe med udvælgelsen af de mest relevante eksperimenter medfører, at omkostningerne i forbindelse med eksempelvis indkøb og bortskaffelse af kemikalier nedsættes.

Generelle kraftfelter

Molekylmekanik (engelsk: molecular mechanics) er baseret på, at molekyler, eller samlinger af interagerende molekyler, kan beskrives udelukkende ud fra atomkernernes positioner. Da elektronerne ikke inkluderes i beskrivelsen, har metoderne den umiddelbare ulempe, at man ikke kan forudsige elektroniske egenskaber. Imidlertid viser det sig, at man ved at opstille meget simple matematiske udtryk er i stand til at beskrive forskellige systemer ret præcist – specielt hvis man har trænet sit kraftfelt (engelsk: force field) på en lignende samling af molekyler.

Ud over atompositionerne er man også nødt til at specificere bindingerne i systemet. Ofte indføres flere forskellige atomtyper, som f.eks. forskellige typer af kulstof, afhængigt af hybridisering (sp, sp² og sp³). Derudover findes der flere forskellige kraftfelter udviklet til at beskrive forskellige systemer. AMBER anvendes primært til proteiner og DNA, hvorimod MM2, MM3 og MM4 er udviklet til at beskrive små organiske forbindelser. Alt dette besværliggør opsætningen en smule, men indtil for nylig har det været de eneste tilgængelige beregningsmetoder for større systemer (>1000 atomer), og det er derfor bredt anvendt. Der findes også mere avancerede og præcise kvantemekaniske metoder baseret på Schrödinger-ligningen, men de kan kun behandle et par hundrede atomer med den nuværende tilgængelige computerkraft.



Figur 1. Illustration af de enkelte energibidrag (ligning 1) i molekylmekanik [1].

Opbygning af et typisk kraftfelt

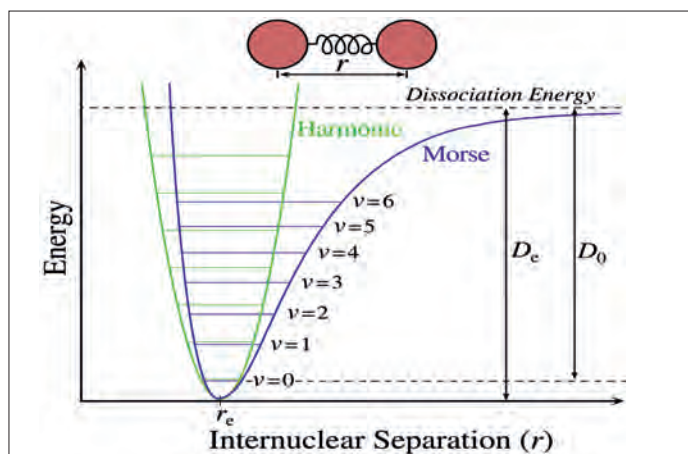
Alle kraftfelter er bygget op på næsten samme måde, hvor energien i systemet opdeles i flere led, som vist i ligning 1. I dette eksempel er kun stræk (str), bøjning (bend), torsioner (tors), samt "non bonded"-interaktioner i form af van der Waals (vdW) og elektrostatiske (el) interaktioner inkluderet. Dette er det mindste antal led, der normalt er nødvendige. Mange kraftfelter indfører ekstra special-led, til f.eks. at behandle hydrogen-bindinger, hvis det er relevant for de studerede systemer.

$$E_{FF} = E_{str} + E_{ben} + E_{tors} + E_{vdW} + E_{el} \quad 1$$

Som illustreret i figur 1 anvendes der relativt simple udtryk for de forskellige led.

F.eks. approksimeres energien, når en binding strækkes eller komprimeres som et simpelt harmonisk potentiale (2), på trods af at det er kendt, at Morse-potentialet er en mere korrekt beskrivelse generelt (figur 2).

Dette er en meget forsimplet beskrivelse af en binding, men det viser sig, at beskrivelsen er relativt præcis, når bindingerne deformeres +/- 0,1 Å. Det er i mange tilfælde tilstrækkeligt, da man kun ønsker at beskrive relakserede systemer ved relativt lave temperaturer. I ligning 1 er de første 3 led intramolekulære og beregnes for alle forbundne atomer, mens de sidste to intermolekulære led beregnes for alle atomer, der er separeret med mindst tre bindinger. Ofte defineres en "cut-off" radius på f.eks. ►



Figur 2. Illustration af forskellen på et harmonisk potential og Morse-potentialet for kemisk binding [2].

$$E_{str} = \sum_{bonds} k_{AB}(R - R_0)^2 \quad 2$$

10 Å, som er den afstand, hvor interaktionen er så lille, at det er spild af (computer)tid at regne den ud.

Som vist i ligning 2 indføres der konstanter for stivheden i bindingerne (dvs. str). På samme måde er der konstanter for, hvor fleksibelt (dvs. bend og tors) molekylet er samt konstanter i forbindelse med "non bonded"-interaktioner. Disse konstanter er ikke umiddelbart direkte målbare eller mulige at beregne, og man må derfor komme med et startgæt, for herefter at optimere dem. Ideen med molekylmekanik er at beskrive en gruppe af ensartede molekyler ud fra kendskab til en lille delmængde – f.eks. vil det forventes, at alle systemer af lineære alkaner kan beskrives ud fra et kraftfelt trænet med data for de ti mindste. Inputdataene til den træning kan være enten eksperimentelle eller højkvalitets beregnede værdier. Typisk benyttes en kombination af bindingslængder, vinkler og torsionspotentialer i denne proces. Træningen eller optimeringen af kraftfeltet er meget simpel at formulere matematisk, men svær at løse i praksis. Typisk løses det ved at minimere fejlfunktionen, ErrF (3), der er et mål for, hvor stor fejl kraftfeltet i gennemsnit har ift. inputdataene.

$$\sum_i^{data} (\chi_{i,FF} - \chi_{i,exp})^2 \cdot \omega^2 \quad 3$$

Forskellene i summen kan være på alle målbare værdier og kan have forskellige vægte, ω , afhængigt af hvor vigtige datapunkterne anses for at være, og hvor store statistiske fejl der observeres. Problemet med at optimere ErrF er, at der i et standard kraftfelt er mindst 50 parametre, som kan ændres for at få den optimale beskrivelse og dermed finde minimum for ErrF. Ofte er disse parametre stærkt korrelerede, hvilket gør, at det er fordelagtigt at anvende mere avancerede multiparameter-optimeringer. Dette kan opnås ved en række forskellige matematiske metoder, som f.eks. Newton-Raphson-algoritmen, der dog har den store ulempe, at den kræver beregning af differentialkvotienterne for ErrF. Dette kan oftest ikke lade sig gøre analytisk, og derfor må løsningen findes numerisk, hvilket er meget ressourcekrævende.

Det er vigtigt at huske, at et kraftfelt (formodentligt) kun kan beskrive samme type molekyler, som det er trænet med. Som minimum skal man forvente, at den gennemsnitlige fejl ved beregningerne er lig den gennemsnitlige fejl i datasættet

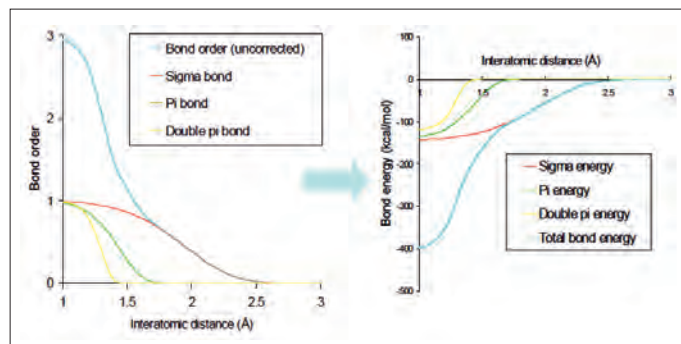
anvendt til træningen. Alligevel anvendes kraftfelter ofte, da det er den eneste brugbare metode til at studere større systemer. Da bindingerne specificeres på forhånd og normalt ikke kan ændres undervejs i simuleringen, er et kraftfelt ikke i stand til at beskrive kemiske reaktioner.

Et reaktivt kraftfelt

I 2001 publicerede van Duin *et al.* en artikel om forbrænding af kulbrinter ved anvendelse af et specielt reaktivt kraftfelt (engelsk: Reactive Force Field), ReaxFF [3]. ReaxFF adskiller sig fra et almindeligt kraftfelt, som beskrevet tidligere, ved at alle led i det generelle udtryk (1), er gjort afhængige af bindingsordenen, BO, som vi kender f.eks. fra kulstofs sigma og pi-bindinger (figur 3).

BO beregnes ud fra afstanden mellem to atomer og beskrives ved en eksponentialfunktion, der går mod nul, når atomerne bevæger sig væk fra hinanden og derfor ikke forventes at interagere mere.

ReaxFF er meget let at sætte op modsat andre kraftfelter, da det eneste der kræves for at lave en beregning er de initiale xyz-koordinater for alle atomer samt angivelse af, hvilke grundstoffer der er tale om. Der findes kun en "atomtype" pr. grundstof. Programmet finder selv ud af, hvilke atomer der er tæt på hinanden og dermed danner bindinger eller interagerer generelt. At der ikke findes flere forskellige atomtyper har også den betydning, at kraftfeltet er generelt og kan anvendes til at

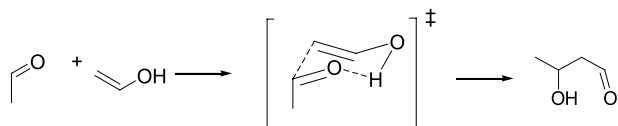


Figur 3. Illustration af bindingsorden (venstre) og bindingsenergi (højre) som funktion af afstanden mellem to kulstofatomer i ReaxFF-beskrivelsen [4].

beskrive mange forskellige kemiske systemer. En klar fordel er, at kraftfeltet er i stand til at behandle flere faser samtidig. Det at atomerne indledningsvist befinder sig tættere pakket i en del af systemet, betyder at kraftfeltet gennem simuleringer vil holde fast (hvis det er energimæssigt fordelagtigt), mens en fase uden om, f.eks. en gas vil forblive gas (dog vil noget evt. adsorberes til overfladen af den faste del).

ReaxFF er langsommere end ordinære kraftfelter, da der beregnes på interaktioner mellem alle atomer, men dog betydeligt hurtigere end kvantemekaniske metoder, der normalt anvendes til beskrivelser af kemiske reaktioner. ReaxFF har været anvendt til beskrivelse af meget forskelligartede systemer som f.eks. den nævnte kulbrinteoxidering, overgangsmetalkatalyseret kulstofrør-dannelse, sprængstoffer, vandadsorption på ZnO-overflader og til beskrivelse af biologiske systemer. For at kunne beskrive disse forskellige systemer har det været nødvendigt at parameterisere kraftfeltet med nye træningsdata, men generelt har dette ikke ødelagt den grundlæggende struktur. Ofte er det faktisk ganske enkelt at tilføje nye parametre for nye grundstoffer.

På Institut for Kemi, DTU arbejder vi med udviklingen af ReaxFF til studier af kemisk og enzymatisk katalyse [5]. Her tillader ReaxFF simuleringer af reaktioner med mange tusinde

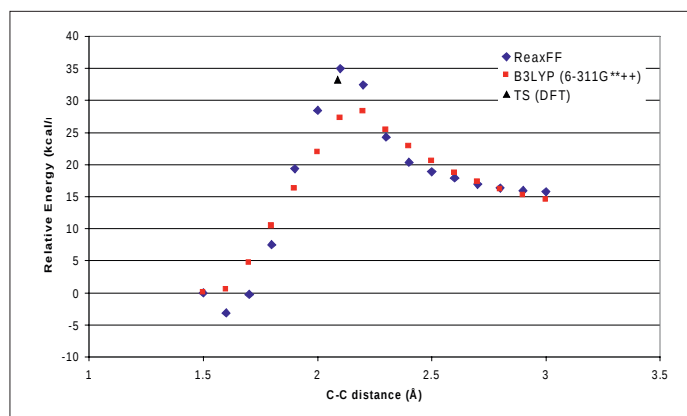


Figur 4. Aldol-reaktionen mellem to acetaldehyd-molekyler hvoraf det ene er på enol-form.

atomer, hvilket udføres på en enkelt CPU, noget som ikke forventes at være muligt med kvantemekaniske metoder de kommende mange år.

Som eksempel på nøjagtigheden af den nuværende version af ReaxFF har vi undersøgt aldol-reaktionen (figur 4). Her dimeriseres to acetaldehyd-molekyler, hvoraf det ene (det nukleofile) er på enol-formen. Reaktionen danner en C-C-binding og der sker samtidig en proton-overførsel. På figur 5 ses energien som funktion af C-C-afstanden, og det er tydeligt, at der er meget fin overensstemmelse mellem ReaxFF og energier bestemt med tæthedsfunktionalteori (DFT).

Generelt kan der ikke forventes fuldstændig overensstemmelse med molekyllmodellering og eksperimenter, men det er vores håb, at ReaxFF vil vise sig at være hurtigt og præcist nok til at kunne bruges i forudsigelser af kemisk reaktivitet og selektivitet, specielt i forbindelse med indledende katalysatorscreening.



Figur 5. Energidiagram for aldol-reaktionen vist i figur 4. X-aksen er afstanden mellem de to kulstofatomer, hvorimellem bindingen dannes.

Projektet "Atomistic Modelling of Chemical Reactions – A New Approach in Drug Design" er støttet af Det Frie Forskningsråd, Teknologi og Produktion i perioden 2009-2012. Desuden ønsker vi at takke Prof. William A. Goddard III (California Institute of Technology) og Prof. Adri C. T. van Duin (Penn State University) for samarbejde om udviklingen af ReaxFF. Projektet blev startet under et PostDoc-ophold (PF) finansieret af Carlsbergfondet og Lundbeckfonden (2008-9). Arbejdet er videreført af Peter B. Jensen i et eksamensprojekt støttet af Novozymes A/S. Vi vil desuden takke Allan Svendsen (Novozyms) og Hans Jørgen Aagaard Jensen (SDU) for faglig sparring.

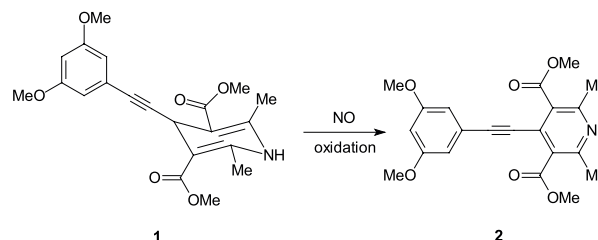
E-mail-adresse
Peter Fristrup: pf@kemi.dtu.dk

Referencer

1. Boas F. E.; Harbury, P. B. *Curr. Opin. Struct. Biol.* **2007**, *17*, 199.
2. Illustration lavet af Mark Somoza, 26. marts 2006.
3. van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A., III *J. Phys. Chem. A.* **2001**, *105*, 9396.
4. Gengivet med personlig tilladelse fra Adri van Duin.
5. Mere information om igangværende projekter kan findes her: <http://www.organic.kemi.dtu.dk/Research/PeterFrstrup.aspx>

... Diagnose af lungesygdomme ved hjælp af NO

Forøget koncentration af NO i udåndingsluften er en biomarkør for forskellige lungesygdomme som tuberkulose og lungecancer. Man er derfor interesseret i at udvikle metoder, der kan detektere en sådan forøgelse. Forskere ved University of Michigan har iagttaget at dihydropyridinen **1**, som ikke er plan og er letopløselig, ved oxidation omdannes til den plane pyridinforbindelsen **2**, der danner en gel. Man har udført denne oxidation



ved hjælp af NO som oxidationsmiddel og fundet, at graden af geldannelse er afhængig af mængden af NO. Man håber at kunne udnytte denne iagttagelse til udvikling af en sensor, der kan måle NO i udåndingsluften.

Carl Th.

Analyte-Triggered Gelation, *Journal of the American Chemical Society*, 130, 2008, side 16496.

C. I-Analyse
Gasanalyse – Gasalarmer – CEM systemer – Rådgivning – Service – Nanopartikler – Aerosoler – IAQ

En Målbar forskel i gasanalyse



Delta-F O₂ Trace analyser
ppt / ppb / ppm / %



Servomex
O₂ / CO / CO₂ mm. Process analyser



IQM 60 Indeks klimamålestation/logger
CO₂/O₃/VOC/rH/Temp/PM 10/2,5 mm.

C. I-Analyse · Nattergalevej 7 · DK-3650 Ølstykke · Tlf.: +45 4717 9299 · Mobil: +45 2961 1783
www.ihle-andersen.dk · claus@ihle-andersen.dk