



optibrium



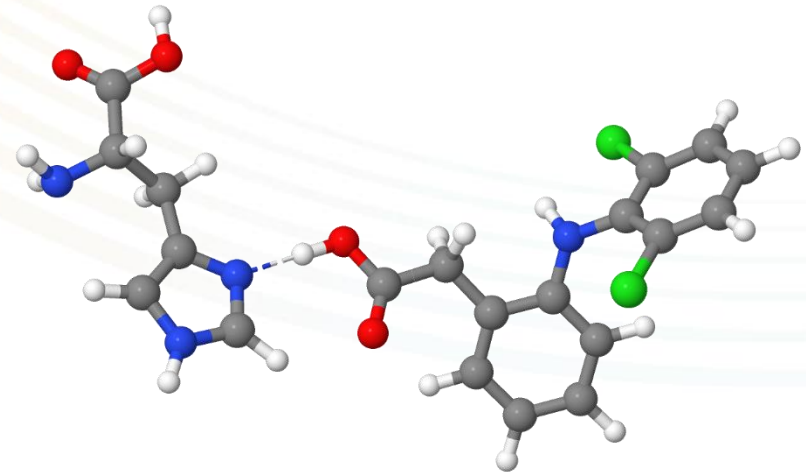
Vesiniksidemete tugevus: *ab initio* täpsus,
jõuväljade kiirus

18. aprill 2019

Mario Öeren – mario@optibrium.com

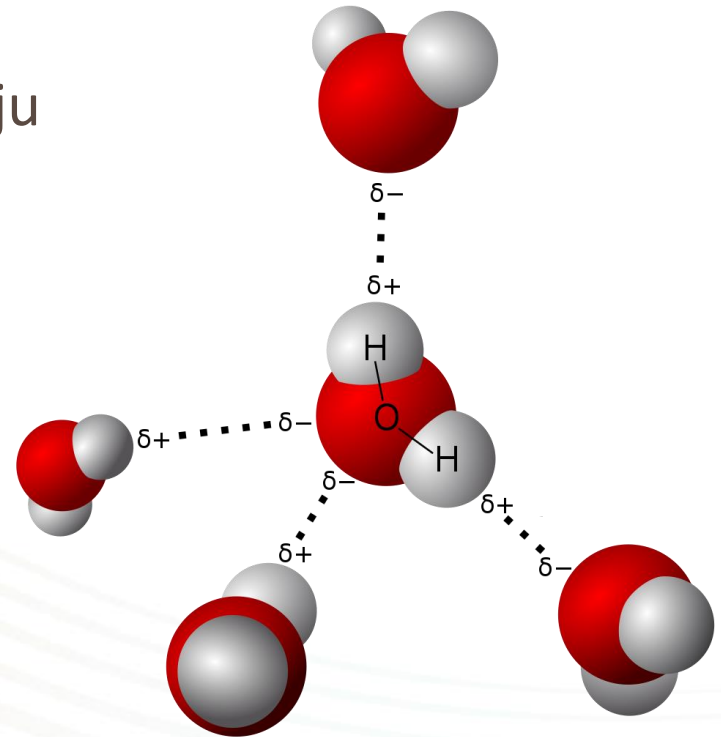
Sisukord

- Vesiniksidemed
 - Olulisus
- Gaussian Approximation Potential (GAP)
 - Praktiline näide
- Vesiniksidemete mudel
 - Kuidas rakendada masinõpet
- Tulemused
 - Mudeli täpsus
 - Ajavõit



Vesiniksidemed

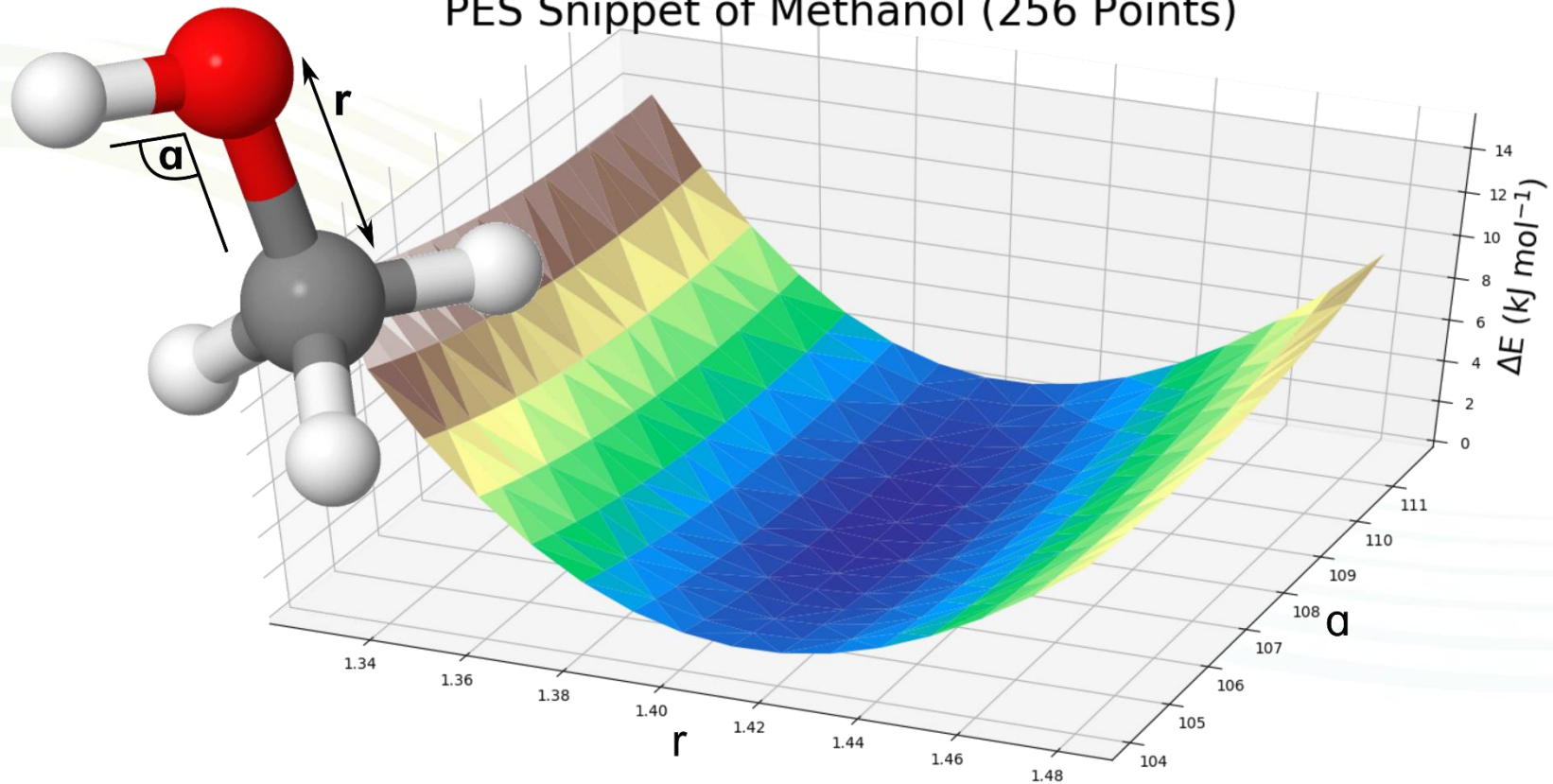
- Peamiselt elektrostaatiline vastasmõju
 - Prootoni doonor
 - Prootoni aktseptor
- Tähtsus
 - Lahustite füüsilised omadused
 - Biomolekulide struktuur
- Ravimitööstus
 - Virtuaalne skriinimine (virtual screening)
 - Ravimite optimeerimine (lead optimisation)



Praktiline näide

B3LYP/def2-SVP

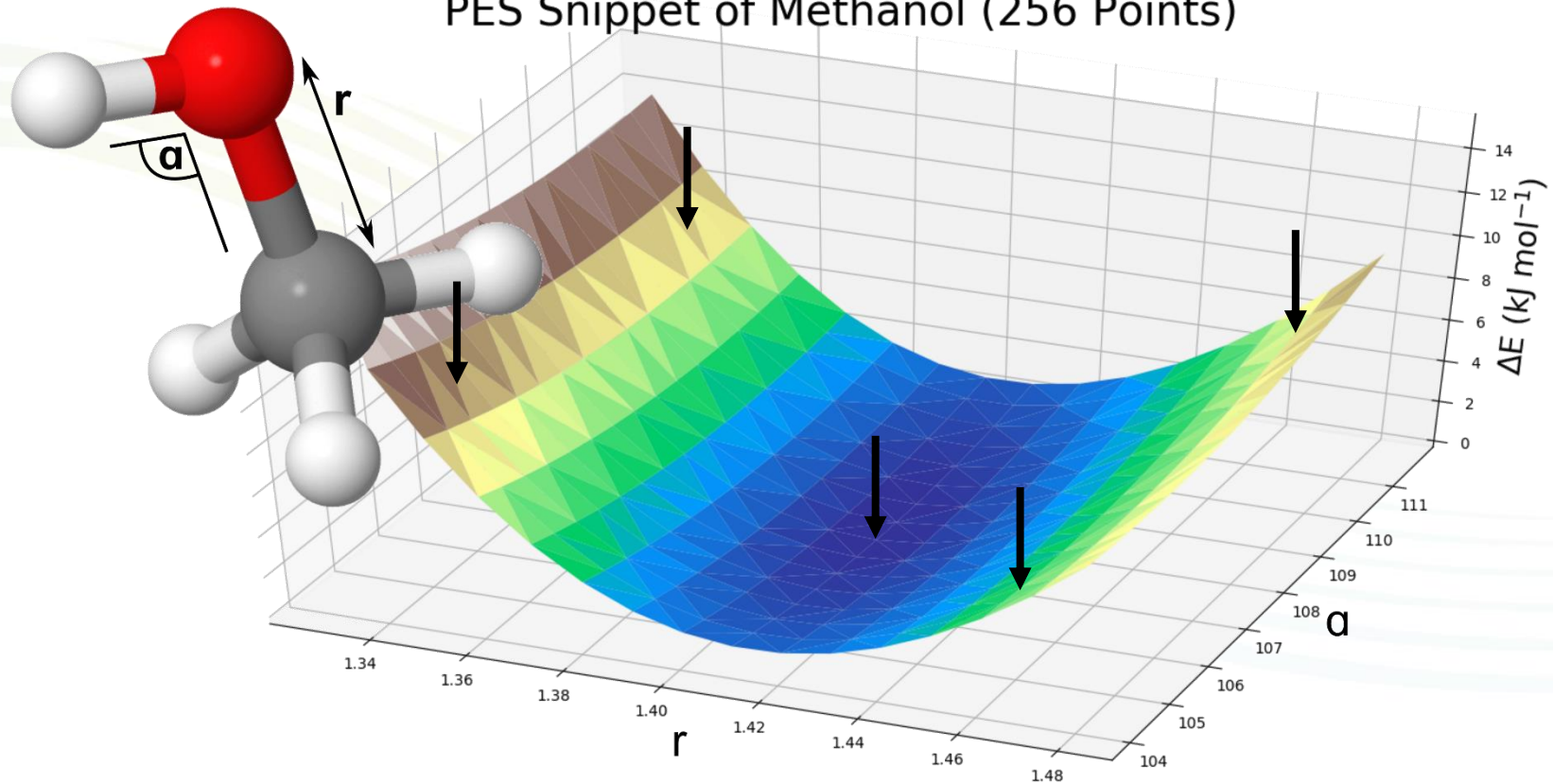
PES Snippet of Methanol (256 Points)



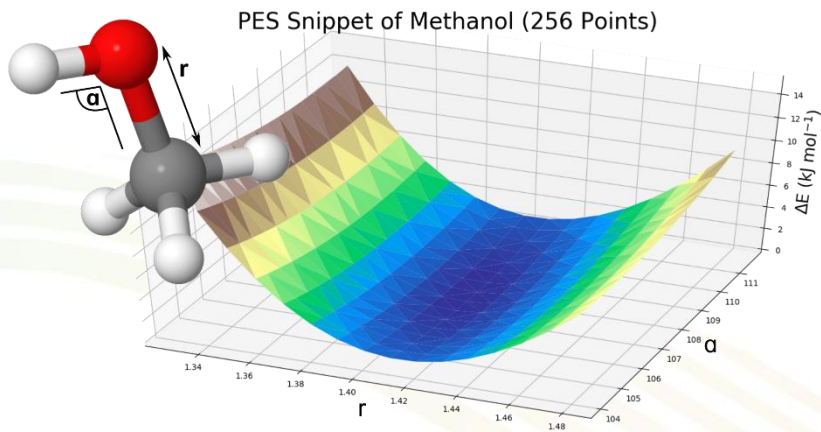
Praktiline näide

B3LYP/def2-SVP

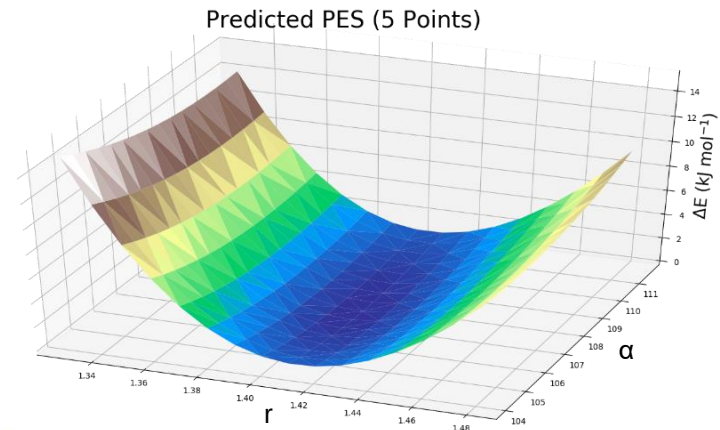
PES Snippet of Methanol (256 Points)



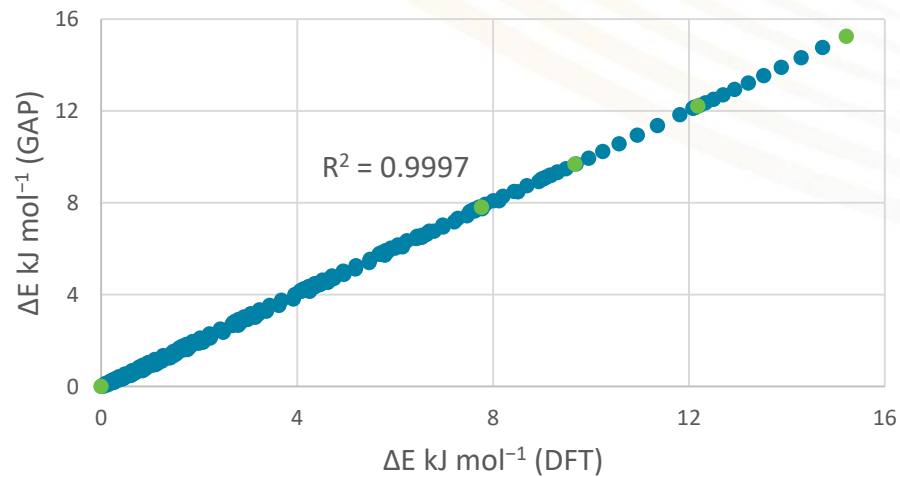
Praktiline näide



B3LYP/def2-SVP



GAP (energia, gradient)

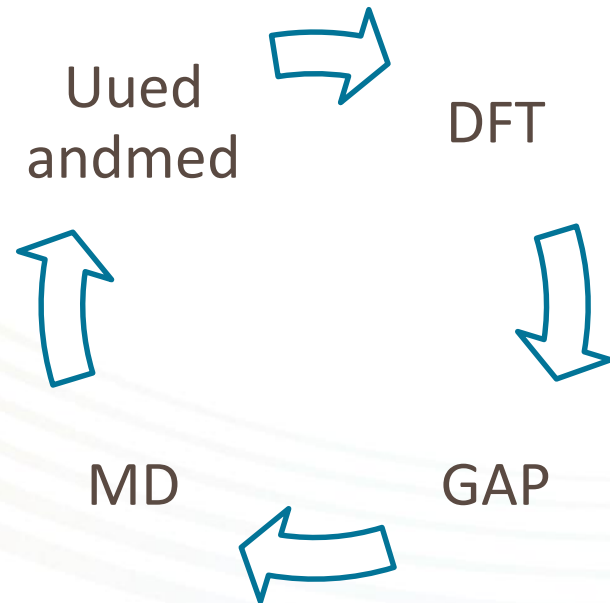


- Test Set (251 Points)
- Training Set (5 Points)

ΔE kirjeldab uuritava struktuuri ning energeetiliselt soodsaima struktuuri energia vahet

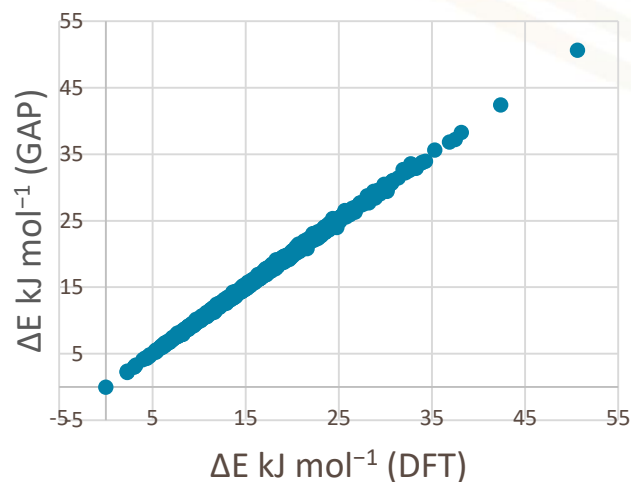
Metanool

- Potentsiaalse energia pinna valimi koostamine
 - Süstemaatiline
 - Molekulaardünaamika
- Andmestik
 - 100 punkti treenimiseks
 - 1000 punkti testimiseks

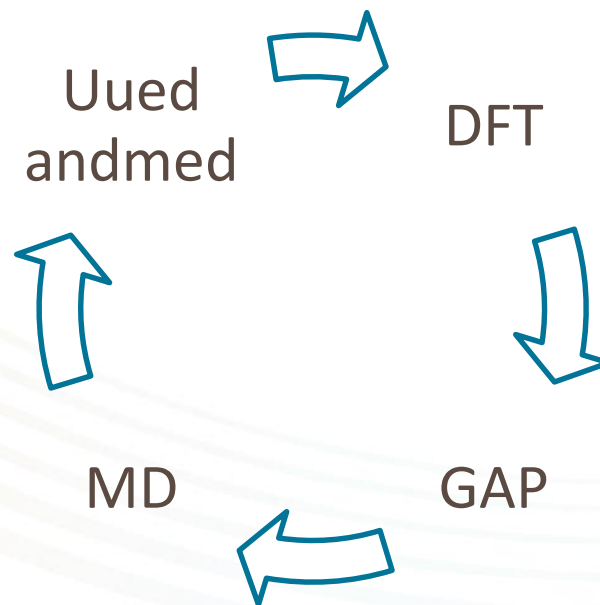


Metanool

- Potentsiaalse energia pinna valimi koostamine
 - Süstemaatiline
 - Molekulaardünaamika
- Andmestik
 - 100 punkti treenimiseks
 - 1000 punkti testimiseks



• 1000 Points



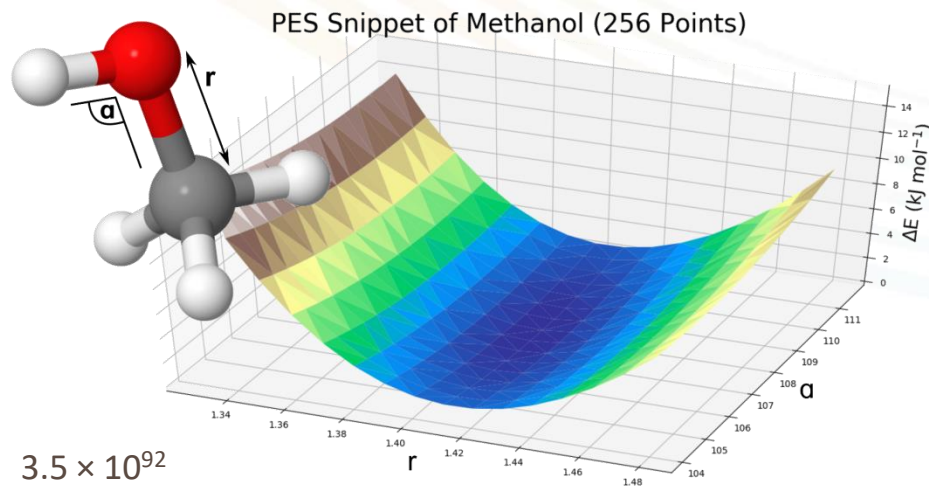
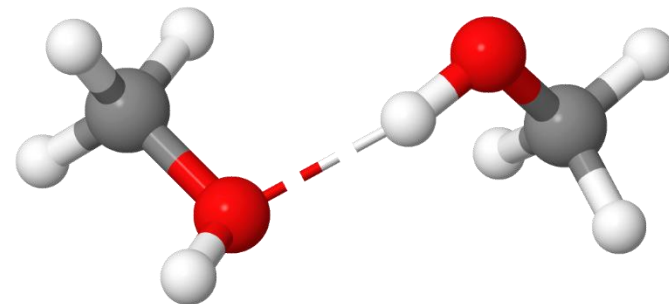
$$R^2 = 0.987$$

RMSE: 0.2 kJ mol^{-1}

Suurim viga: 0.88 kJ mol^{-1}

Dimeer

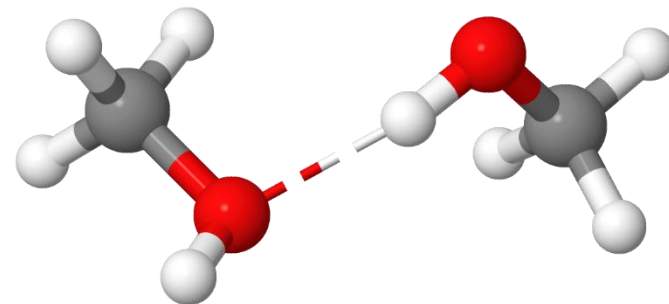
- Keeruline potentsiaalse energia pind
 - Üle 20 muutuja
 - Vesinikside
 - Suur valik konformeere



Dimeer

- Keeruline potentsiaalse energia pind

- Üle 20 muutuja
- Vesinikside
- Suur valik konformeere

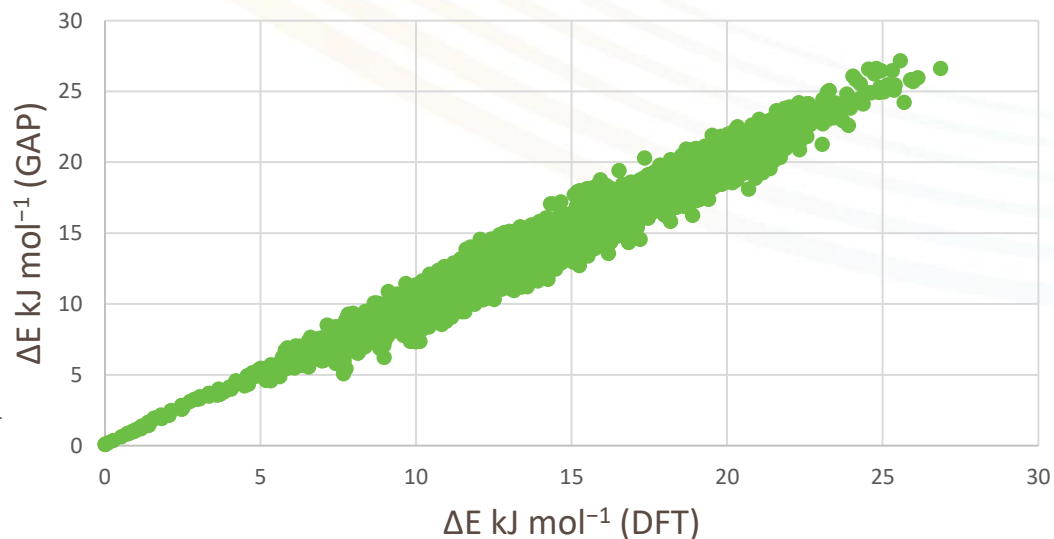


1500 andmepunkti treenimiseks ja
10 000 andmepunkti testimiseks

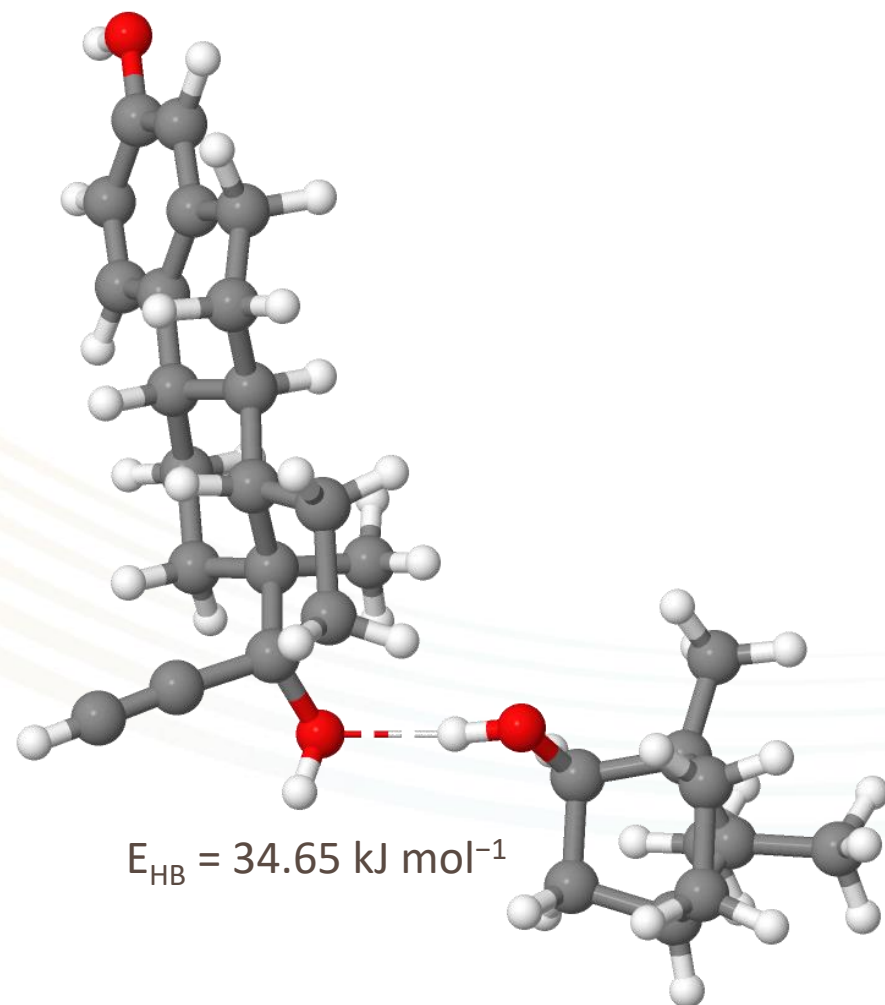
$R^2 = 0.9682$

RMSE: 0.7 kJ mol^{-1}

Suurim viga: 3.0 kJ mol^{-1}

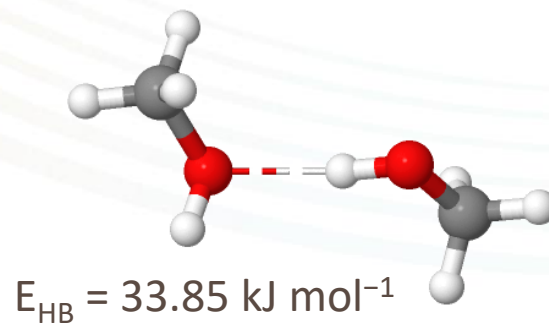


Vesiniksidemete tugevuse hindamine



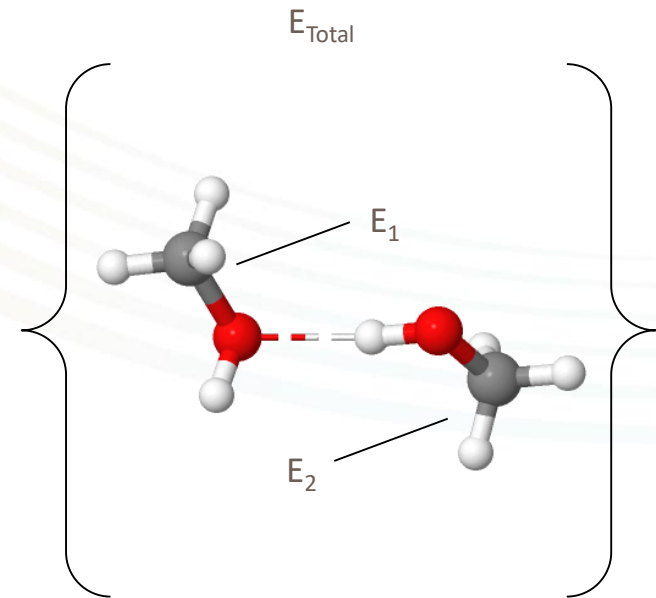
Vesiniksidemete tugevuse hindamine

- Fragmendipõhine lahendus



Vesiniksidemete tugevuse hindamine

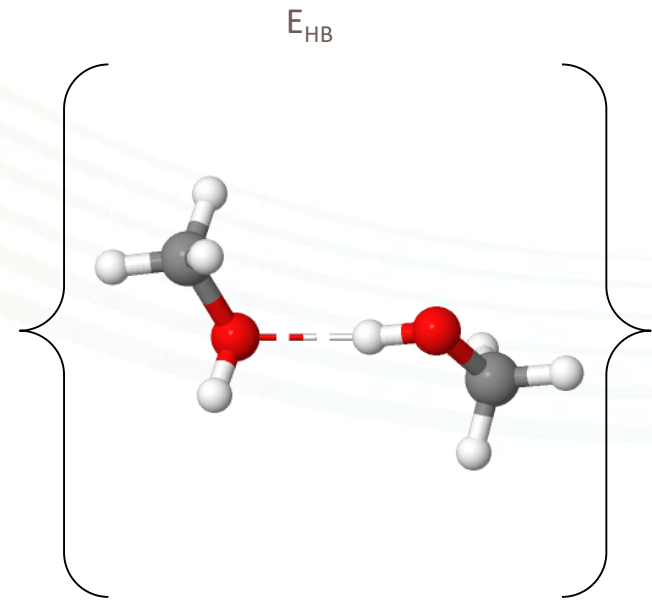
- Fragmendipõhine lahendus
- Vesiniksideme energia
 - Metanooli molekulid
 - Metanooli dimeer



$$E_{\text{HB}} = E_{\text{Total}} - E_1 - E_2$$

Vesiniksidemete tugevuse hindamine

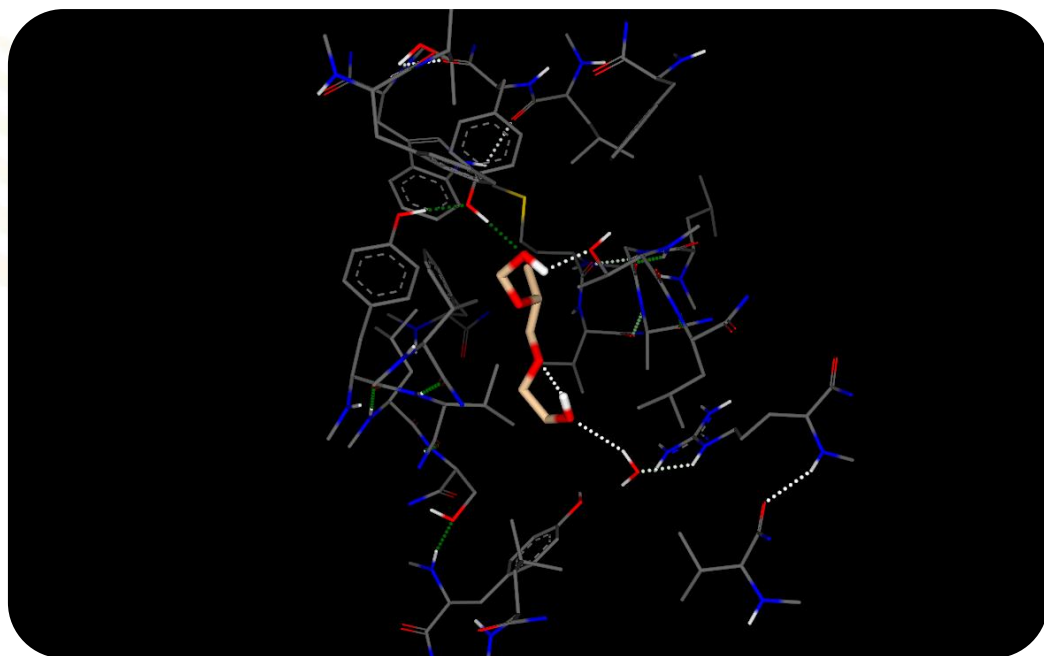
- Fragmendipõhine lahendus
- Vesiniksideme energia
 - Metanooli molekulid
 - Metanooli dimeer
- Vastasmõju potentsiaal
 - Kasutame E_{HB}
- Eelised
 - Täpsem
 - Kiirem



Vesiniksidemete tugevuse hindamine

- Fragmendipõhine lahendus
- Vesiniksideme energia
 - Metanooli molekulid
 - Metanooli dimeer
- Vastasmõju potentsiaal
 - Kasutame E_{HB}
- Eelised
 - Täpsem
 - Kiirem

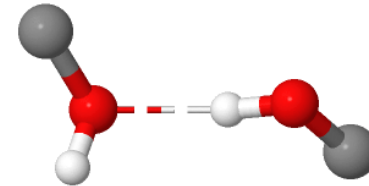
PDB ID: 5FDQ



Pildid on loodud SeeSAR™ mooduliga StarDrop™ programmis.

Vesiniksidemete tugevuse hindamine

- Fragmendipõhine lahendus
- Vesiniksideme energia
 - Metanooli molekulid
 - Metanooli dimeer
- Vastasmõju potentsiaal
 - Kasutame E_{HB}
- Eelised
 - Täpsem
 - Kiirem



Vesiniksidemete tugevuse hindamine

- Fragmendipõhine lahendus

- Vesiniksideme energia

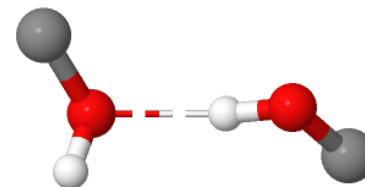
- Metanooli molekulid
- Metanooli dimeer

- Vastasmõju potentsiaal

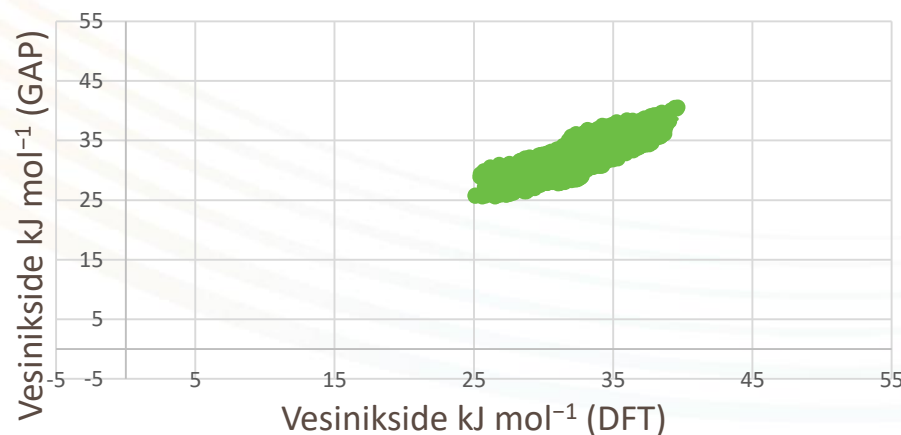
- Kasutame E_{HB}

- Eelised

- Täpsem
- Kiirem



10 000 andmepunkti testimiseks



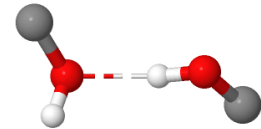
$$R^2 = 0.7646$$

RMSE: 1.2 kJ mol^{-1}

Suurim viga: 4.4 kJ mol^{-1}

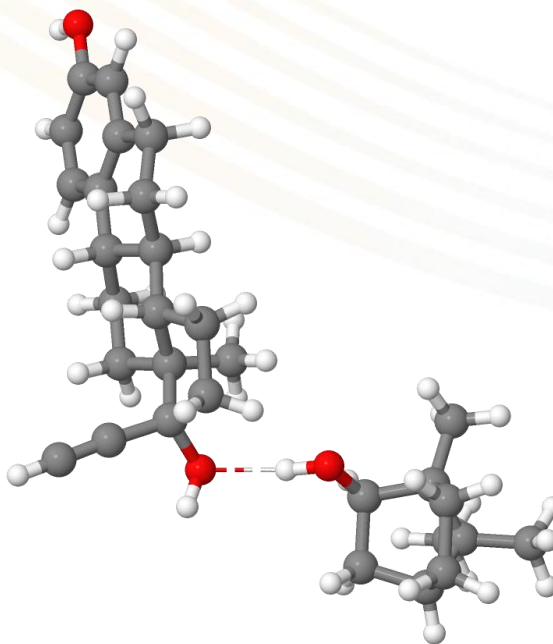
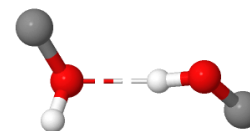
Kokkuvõte

- Vesiniksidemete mudel – *ab initio* täpsus kasutades GAPe
 - RMSE: 1.2 kJ mol⁻¹
 - Ajavõit: 120 sekundit (DFT), 0.02 sekundit (GAP)



Kokkuvõte

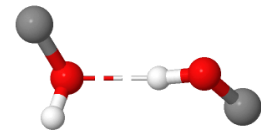
- Vesiniksidemete mudel – *ab initio* täpsus kasutades GAPe
 - RMSE: 1.2 kJ mol⁻¹
 - Ajavõit: 120 sekundit (DFT), 0.02 sekundit (GAP)
 - Suured molekulid (12)
RMSE: 0.66 kJ mol⁻¹, suurim viga: 1.28 kJ mol⁻¹



Kokkuvõte

- Vesiniksidemete mudel – *ab initio* täpsus kasutades GAPe

- RMSE: 1.2 kJ mol⁻¹
- Ajavõit: 120 sekundit (DFT), 0.02 sekundit (GAP)
- Suured molekulid (12)
RMSE: 0.66 kJ mol⁻¹, suurim viga: 1.28 kJ mol⁻¹



- Rakendatavus

- Teised dimeerid

metanool : formaldehüüd

metanool : metaanhappe anioon

metanool : metaanhape

primaarne amiin : primaarne amiin

- Konjugatsioon

- Ei piirdu ainult vesiniksidemete ja DFT-ga

Tänan tähelepanu eest!

- Peter Hunt
- Matthew Segall
- Gábor Csányi
- David Ponting



UNIVERSITY OF
CAMBRIDGE



NWChem

HIGH-PERFORMANCE COMPUTATIONAL
CHEMISTRY SOFTWARE

**GAP
SOAP**

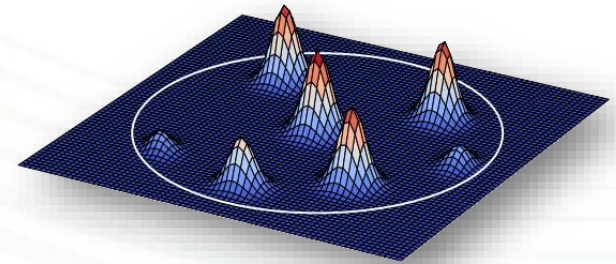


matplotlib



Sissejuhatus

- Sisendandmed *ab initio* arvutustest:
 - Cartesiuse koordinaadid
 - Energiaväärtused
 - Gradiendiväärtused
- Struktuur masinõppeks
 - Muutumatu permutatsioonide ning pöörde-, peegeldus- ja nihkesümmeetria suhtes
- Smooth Overlap of Atomic Positions (SOAP)
 - Aatomite ümbruskonna tihedusfunktsioon
 - Kirjeldab keerulisi ning suuri struktuure
 - Bartók *et al* Physical Review B 87, 184115 (2013)
- Jõuvälja treenimine kasutades Gaussi protsessi
 - Gaussian Approximation Potential (GAP)
 - Bartók *et al* International Journal of Quantum Chemistry 2015, 115, 1051 – 1057



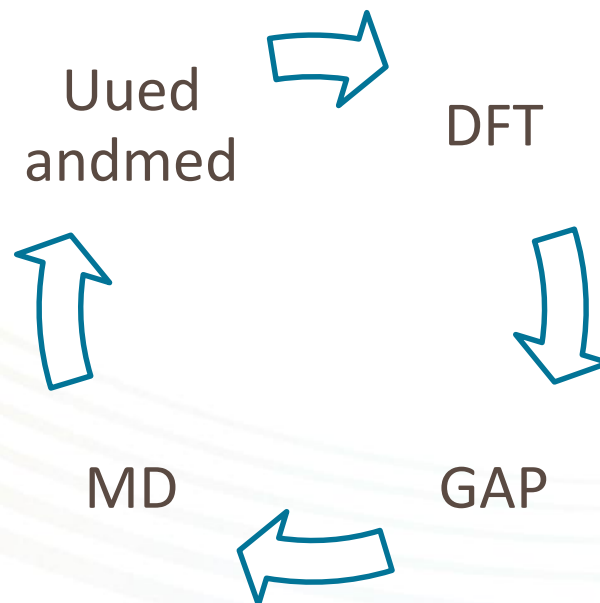
Metanool

- Potentsiaalse energia pinna valimi koostamine

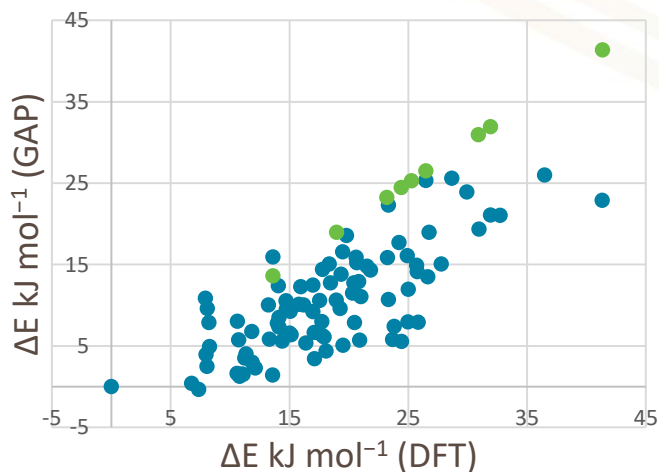
- Süstemaatiline
- Molekulaardünaamika

- Andmestik

- 100 punkti treenimiseks
- 1000 punkti testimiseks



- 91 Points
- 9 Points



$$R^2 = 0.5895$$

RMSE: 9.1 kJ mol^{-1}

Suurim viga: 18.9 kJ mol^{-1}

Metanool

- Potentsiaalse energia pinna valimi koostamine

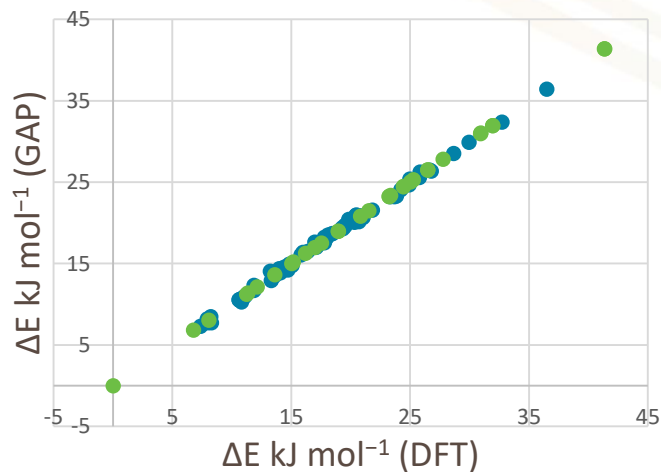
- Süstemaatiline
- Molekulaardünaamika

- Andmestik

- 100 punkti treenimiseks
- 1000 punkti testimiseks

• 68 Points

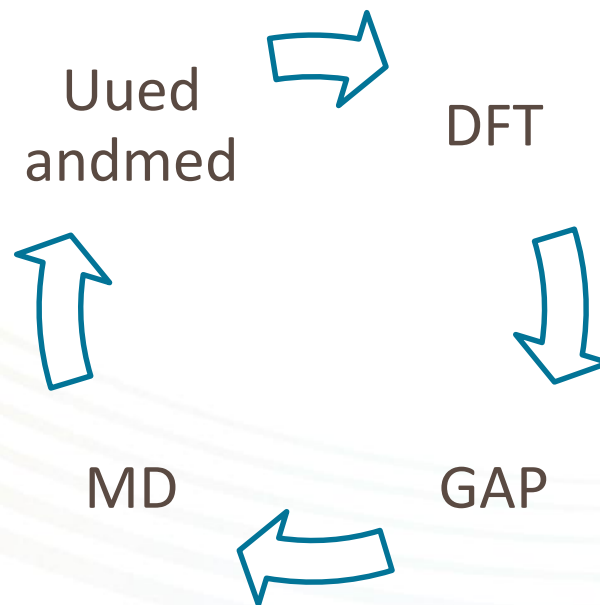
• 32 Points



$$R^2 = 0.9975$$

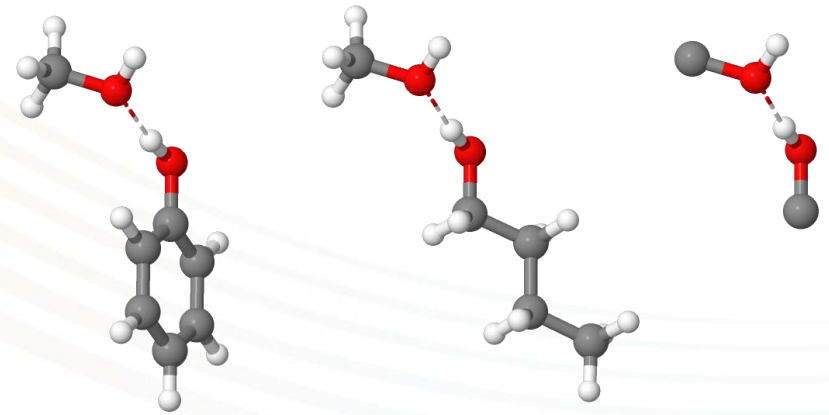
RMSE: 0.3 kJ mol^{-1}

Suurim viga: 0.79 kJ mol^{-1}



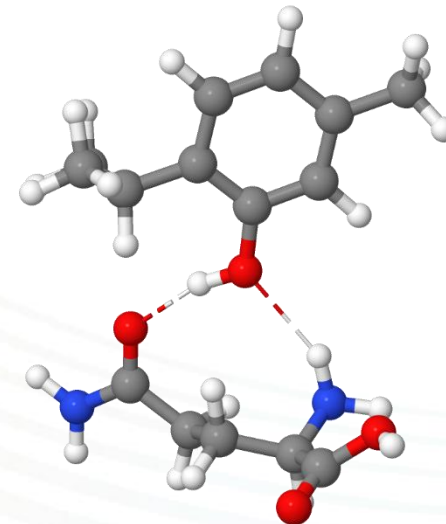
Tuleviku perspektiiv

- Über potentsiaal
 - Täiendavad dimeeride tüübid
 - Laenguga dimeerid
 - Konjugeerunud süsteemid
- Keerulisemad interaktsioonid
 - Keerulised dimeerid
 - Kolmas monomeer



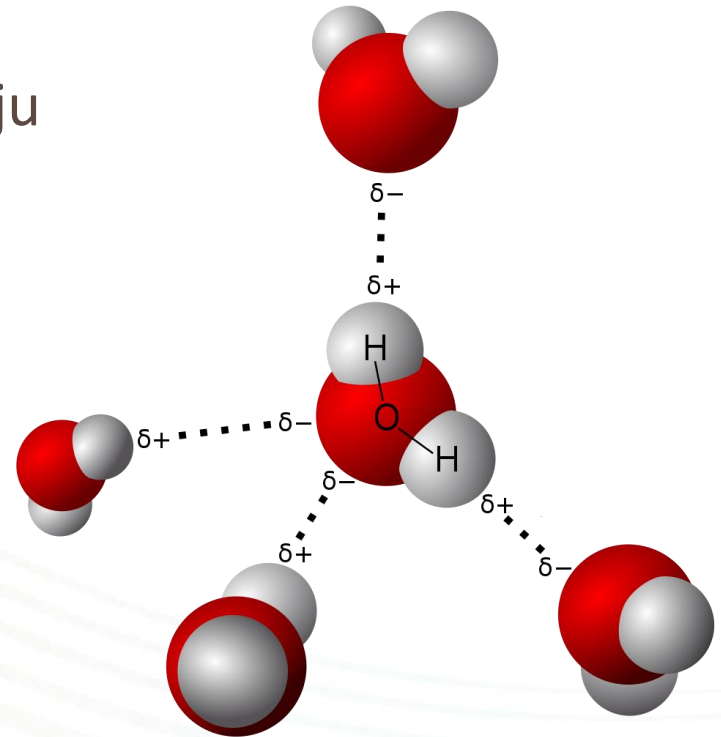
Tuleviku perspektiiv

- Über potentsiaal
 - Täiendavad dimeeride tüübid
 - Laenguga dimeerid
 - Konjugeerunud süsteemid
- Keerulisemad interaktsioonid
 - Keerulised dimeerid
 - Kolmas monomeer



Vesiniksidemed

- Peamiselt elektrostaatiline vastasmõju
 - Prootoni doonor
 - Prootoni aktseptor
- Tähtsus
 - Lahustite füüsilised omadused
 - Biomolekulide struktuur
- Ravimitööstus
 - Virtuaalne skriinimine (virtual screening)
 - Ravimite optimeerimine (lead optimisation)



https://en.wikipedia.org/wiki/Hydrogen_bond#/media/File:3D_model_hydrogen_bonds_in_water.svg