# Ab initio molecular dynamics for biomolecular spectroscopy

#### Mariana Rossi

#### Fritz-Haber-Institut der MPG, Faradayweg 4-6, 14195 Berlin



DFT and beyond - FHI-aims Developers' and Users' Meeting 2012, Berlin

<u>Ion mobility spectroscopy</u>

Evidence of helical secondary structure

for Ac-Ala<sub>n</sub>-LysH<sup>+</sup>, n=5-20 in the gas-phase

Hudgins, Ratner, Jarrold, J.Am. Chem. Soc **120**, 12974 (1998)



lon mobility spectroscopy

Evidence of helical secondary structure

for Ac-Ala<sub>n</sub>-LysH<sup>+</sup>, n=5-20 in the gas-phase

Hudgins, Ratner, Jarrold, J.Am. Chem. Soc **I 20**, I 2974 (1998)





#### IR spectroscopy

Detailed structure sensitivity

[1] Stearns, Seaiby, Boyarkin, Rizzo, Phys. Chem. Chem. Phys. 11, 125 (2009)
[2] Kupser, Pagel, Oomens, Polfer, Koksch, Meijer, von Helden, J.Am. Chem. Soc. 132, 2085 (2010) Others!



## Outline

- Conformational search and IR spectra
- Benchmark against "beyond DFT" methods
- Including nuclear quantum effects new features of FHI-aims

# Methodology

- All-electron / localized basis code (DFT and beyond) developed in the Fritz-Haber-Institut: FHI-aims<sup>[1]</sup>
- Accurate NAO basis sets for DFT
- PBE exchange correlation functional
- Inclusion of van der Waals (vdW) effects via a C<sub>6</sub>[n]/R<sup>6</sup> term derived from self consistent electronic density<sup>[2]</sup>
  - Benchmarked against CCSD(T) for alanine based polypeptides<sup>[3]</sup>



[1] V. Blum et al. Comp. Phys. Comm. **180**, 2175 (2009)
 [2] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009)
 [3] A. Tkatchenko, M. Rossi, V. Blum, J. Ireta, and M. Scheffler, Phys. Rev. Lett. **106**, 118102, (2011)





 Huge conformational freedom:



- Huge conformational freedom:
  - 10 torsional angles



- Huge conformational freedom:
  - 10 torsional angles
  - 5 peptide bonds



- Huge conformational freedom:
  - 10 torsional angles
  - 5 peptide bonds
  - Long lysine side chain



- Huge conformational freedom:
  - 10 torsional angles
  - 5 peptide bonds
  - Long lysine side chain
- E.g., assume 5 values per torsional mode  $\Rightarrow 5^{10}!$



- Huge conformational freedom:
  - 10 torsional angles
  - 5 peptide bonds
  - Long lysine side chain
- E.g., assume 5 values per torsional mode  $\Rightarrow 5^{10}!$



## Strategy to find conformers: Ac-Ala<sub>n</sub>-LysH<sup>+</sup>, n=4-8

Nconformers

Input structures generation: Basin-hopping search (TINKER) with opls-aa force field

*O*(10<sup>5</sup>)

## Strategy to find conformers: Ac-Ala<sub>n</sub>-LysH<sup>+</sup>, n=4-8





#### Number of PBE+vdW relaxations for Ac-Ala<sub>n</sub>-LysH<sup>+</sup>, n=4-8

size	# of relax.
4	1068
5	1000
6	800
7	800
8	820



#### Energy hierarchies: PES and harmonic free energies $\alpha$ -helix 3<sub>10</sub>-helix 0.2 $E-E_{\alpha-helix}$ (eV) Ib, 3 la 0 -0. I Families PES Families PES Families PES Families PES Families PES 5 8 4 6

Number of Ala residues





-0.1







- 3<sub>10</sub>-helix



— α-helix

**—** 3<sub>10</sub>-helix











$$F_{vib}(T) = E_{PES} + \Delta U_{vib}(T) - TS_{vib}(T) = E_{PES} + \sum_{i} \left[ \frac{\hbar \omega_i}{2} + k_B T \ln \left( 1 - \exp^{-\frac{\hbar \omega_i}{k_B T}} \right) \right]$$







Example: Ac-Ala<sub>8</sub>-LysH<sup>+</sup> - first vibrational mode, anharmonic

• 21 ps of ab initio molecular dynamics (PBE+vdW), microcanonical ensemble





Example: Ac-Ala<sub>8</sub>-LysH<sup>+</sup> - first vibrational mode, anharmonic

• 21ps of ab initio molecular dynamics (PBE+vdW), microcanonical ensemble



## What stabilizes helices in Polyalanine?



Conspiracy of:

### What stabilizes helices in Polyalanine?



Conspiracy of:

Hydrogen bonds & cooperativity (need sufficient chain length)

van der Waals (favors compact structures)

Entropy: Low-frequency vibrations (disfavors compact structures with increasing T)

## Vibrational spectroscopy: Ac-Ala15-LysH<sup>+</sup>

#### $\alpha$ -helical Ac-Ala<sub>15</sub>-LysH<sup>+</sup>



<u>Experiment</u>: von Helden, Kupser, Bierau, Meijer, Molecular Physics, FHI Berlin

Infrared multiphoton dissociation spectroscopy, FELIX free electron laser

Room temperature



## Vibrational spectroscopy: Ac-Ala15-LysH<sup>+</sup>

#### $\alpha$ -helical Ac-Ala<sub>15</sub>-LysH<sup>+</sup>



<u>Experiment</u>: von Helden, Kupser, Bierau, Meijer, Molecular Physics, FHI Berlin

Infrared multiphoton dissociation spectroscopy, FELIX free electron laser

Room temperature



## Vibrational spectroscopy: Ac-Ala<sub>15</sub>-LysH<sup>+</sup>

 $\alpha$ -helical Ac-Ala<sub>15</sub>-LysH<sup>+</sup>




## Vibrational spectroscopy: Ac-Ala15-LysH<sup>+</sup>

25 ps Born-Oppenheimer molecular dynamics, DFT-PBE+vdW M. Rossi, V. Blum, P. Kupser, G. von Helden, F. Bierau, K. Pagel, G. Meijer, and M. Scheffler, J. Phys. Chem. Lett. 1, 3465 (2010)

Theory: PBE+vdW, shifted, not scaled



## Vibrational spectroscopy: Ac-Ala15-LysH<sup>+</sup>

25 ps Born-Oppenheimer molecular dynamics, DFT-PBE+vdW M. Rossi, V. Blum, P. Kupser, G. von Helden, F. Bierau, K. Pagel, G. Meijer, and M. Scheffler, J. Phys. Chem. Lett. 1, 3465 (2010)



Theory: PBE+vdW, shifted, not scaled

Experiment (IRMPD, 300K) Theory (harmonic)  $I(\omega) \propto \omega^2 \int_{-\infty}^{\infty} dt \, \underline{\langle \vec{M}(t) \cdot \vec{M}(0) \rangle} \, e^{iwt}$ 

dipole-dipole time correlation function

## Vibrational spectroscopy: Ac-Ala15-LysH<sup>+</sup>

25 ps Born-Oppenheimer molecular dynamics, DFT-PBE+vdW M. Rossi, V. Blum, P. Kupser, G. von Helden, F. Bierau, K. Pagel, G. Meijer, and M. Scheffler, J. Phys. Chem. Lett. 1, 3465 (2010)



Theory: PBE+vdW, shifted, not scaled







# Quantifying continuous spectra: The Pendry R-factor



#### Experiment-theory comparison:

- continuous spectra (300 K)
- peak positions matter, not so much absolute intensities

# Quantifying continuous spectra: The Pendry R-factor



# Quantifying continuous spectra: The Pendry R-factor









### AIMD derived IR spectra: importance of vdW

#### <u>Ac-Ala<sub>10</sub>-LysH<sup>+</sup>, < T > = 300K</u>





### AIMD derived IR spectra: importance of vdW

#### <u>Ac-Ala<sub>10</sub>-LysH<sup>+</sup>, < T > = 300K</u>



### AIMD derived IR spectra: importance of vdW

#### <u>Ac-Ala<sub>10</sub>-LysH<sup>+</sup>, < T > = 300K</u>



## Is DFT(+vdW) accurate enough?



# Is DFT(+vdW) accurate enough?

• Even with optimized basis functions (capturing most of the core correlation), feasible "higher level" methods are perturbative



# Is DFT(+vdW) accurate enough?

• Even with optimized basis functions (capturing most of the core correlation), feasible "higher level" methods are perturbative



### Is the classical treatment of the nuclei accurate enough?

- In many cases, it seems so...
- ... but quantum nuclear effects do play a role in many systems
- Example:



• No good match to experiment in this case - quantum effects?

Time dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\Psi\rangle = \hat{H} |\Psi\rangle$$

$$\Psi(t)\rangle = \exp(-i\hat{H}t/\hbar) |\Psi(0)\rangle = \hat{U}(t) |\Psi(0)\rangle$$

$$Iime \text{ evolution propagator}$$
Probability
$$P(x_1, 0; x_2, t) = |\langle x_1 | \hat{U}(t) | x_2 \rangle|^2$$

(x<sub>2</sub>, t) (x<sub>1</sub>, 0) time

All paths contribute:

$$\langle x_1 | \hat{U}(t) | x_2 \rangle \propto \sum_{paths} \exp\left(-\frac{i}{\hbar} \int_{path} \mathcal{L}dt\right) = \sum_{paths} \exp\left(-\frac{i}{\hbar} S_{path}\right)$$

• There is a relation between the quantum canonical density matrix and the quantum time evolution propagator:

 $\hat{O}$ 

$$\beta = 1/(k_B T) \qquad |\Psi(t)\rangle = \exp(-i\hat{H}t/\hbar)|\Psi(0)\rangle = \hat{U}(t)|\Psi(0)\rangle$$
$$= \exp(-\beta\hat{H}t) \qquad \hat{U}(t) = \exp(-i\hat{H}t/\hbar)$$

• There is a relation between the quantum canonical density matrix and the quantum time evolution propagator:

$$\beta = 1/(k_B T) \qquad |\Psi(t)\rangle = \exp(-i\hat{H}t/\hbar)|\Psi(0)\rangle = \hat{U}(t)|\Psi(0)\rangle$$
$$\hat{O} = \exp(-\beta\hat{H}t) \qquad \hat{U}(t) = \exp(-i\hat{H}t/\hbar)$$

$$\hat{\rho}(\beta) = \hat{U}(-i\beta\hbar)$$

 $P(x_1, 0; x_2, -i\beta\hbar) = |\langle x_1 | \hat{U}(-i\beta\hbar) | x_2 \rangle|^2 = |\langle x_1 | \hat{\rho}(\beta) | x_2 \rangle|^2$ 

• There is a relation between the quantum canonical density matrix and the quantum time evolution propagator:

$$\beta = 1/(k_B T) \qquad |\Psi(t)\rangle = \exp(-i\hat{H}t/\hbar)|\Psi(0)\rangle = \hat{U}(t)|\Psi(0)\rangle$$
$$\hat{O} = \exp(-\beta\hat{H}) \qquad \hat{U}(t) = \exp(-i\hat{H}t/\hbar)$$

$$\hat{\rho}(\beta) = \hat{U}(-i\beta\hbar)$$

 $P(x_1, 0; x_2, -i\beta\hbar) = |\langle x_1 | \hat{U}(-i\beta\hbar) | x_2 \rangle|^2 = |\langle x_1 | \hat{\rho}(\beta) | x_2 \rangle|^2$ 

• One can solve everything in imaginary time (canonical density matrix!) and then go back to the real time

For evaluating observables (quantum canonical averages), one needs the partition function

 $Z = \operatorname{Tr}(\hat{\rho}) = \langle \mathbf{x} | \exp(\beta \hat{H}) | \mathbf{x} \rangle$ 

For evaluating observables (quantum canonical averages), one needs the partition function

$$Z = \operatorname{Tr}(\hat{\rho}) = \langle \mathbf{x} | \exp(\beta \hat{H}) | \mathbf{x} \rangle$$

 Performing a Trotter factorization of the trace - "dividing in P slices" taking care of the non-commuting nature of x and p operators:

$$Z \approx \frac{1}{(2\pi\hbar)^{NP}} \int d\mathbf{p}^{NP} \int d\mathbf{x}^{NP} \exp\left[\frac{-\mathcal{H}_P(\mathbf{x}, \mathbf{p})}{Pk_BT}\right] \leftarrow \text{Exact for } P \rightarrow \infty$$
  
Hamiltonian of  
a classical ring  
polymer  $\rightarrow \mathcal{H}_P(\mathbf{x}, \mathbf{p}) = \sum_{I}^{N} \sum_{k}^{P} \left[\frac{[\mathbf{p}_I^{(k)}]^2}{2M_I} + \frac{M_I \omega_P^2}{2} (\mathbf{x}_I^{(k)} - \mathbf{x}_I^{(k+1)})^2\right] + \sum_{k}^{P} V(\mathbf{x}^{(k)})$ 

For evaluating observables (quantum canonical averages), one needs the partition function

 $Z = \operatorname{Tr}(\hat{\rho}) = \langle \mathbf{x} | \exp(\beta \hat{H}) | \mathbf{x} \rangle$ 

 Performing a Trotter factorization of the trace - "dividing in P slices" taking care of the non-commuting nature of x and p operators:

$$Z \approx \frac{1}{(2\pi\hbar)^{NP}} \int d\mathbf{p}^{NP} \int d\mathbf{x}^{NP} \exp\left[\frac{-\mathcal{H}_P(\mathbf{x}, \mathbf{p})}{Pk_BT}\right] \leftarrow \text{Exact for } P \rightarrow \infty$$
Hamiltonian of  
a classical ring  
polymer  $\mathcal{H}_P(\mathbf{x}, \mathbf{p}) = \sum_{I}^{N} \sum_{k}^{P} \left[\frac{[\mathbf{p}_I^{(k)}]^2}{2M_I} + \frac{M_I \omega_P^2}{2} (\mathbf{x}_I^{(k)} - \mathbf{x}_I^{(k+1)})^2\right] + \sum_{k}^{P} V(\mathbf{x}^{(k)})$ 
  
Classical  
 $q_i$ 
 $V(\mathbf{q}, \mathbf{q})$ 
 $q_i$ 
 $Quantum$ 
 $V(\mathbf{q}^{(i)}, \mathbf{q}^{(i)})$ 
 $q_i$ 
 $Quantum$ 
 $V(\mathbf{q}^{(i)}, \mathbf{q}^{(i)})$ 
 $q_i$ 
 $Quantum$ 
 $V(\mathbf{q}^{(i)}, \mathbf{q}^{(i)})$ 
 $Quantum$ 
 $Quantum$ 
 $V(\mathbf{q}^{(i)}, \mathbf{q}^{(i)})$ 
 $Quantum$ 

For evaluating observables (quantum canonical averages), one need partition function

$$Z = \operatorname{Tr}(\hat{\rho}) = \langle \mathbf{x} | \exp(\beta \hat{H}) | \mathbf{x} \rangle$$

 Performing a Trotter factorization of the trace - "dividing in P slices In FHI-aims we now have:

Hai

ac

 "In-code" implementation of PIMD, Centroid MD (CMD), and Ring-Polymer MD (RPMD) - Xinzheng Li, P 8—

 $\hbar$ 

 Interface to external python wrapper "WRAPPI" that works through internet sockets and allows to parallelize even over several machines



M. Ceriotti, G. Bussi, M. Parrinello, JCTC 2010, 6, 1170-1180 (http://gle4md.berlios.de/)

$$\begin{aligned} \dot{x} &= p/m & \mathbf{A}_{p}\mathbf{C}_{p} + \mathbf{C}_{p}\mathbf{A}_{p}^{T} = \mathbf{B}_{p}\mathbf{B}_{p}^{T} \\ \begin{pmatrix} \dot{p} \\ \dot{\mathbf{s}} \end{pmatrix} &= \begin{pmatrix} -V'(x) \\ 0 \end{pmatrix} - \begin{pmatrix} a_{pp} & \mathbf{a}_{p}^{T} \\ \mathbf{a}_{p} & \mathbf{A} \end{pmatrix} \begin{pmatrix} p \\ \mathbf{s} \end{pmatrix} + \begin{pmatrix} b_{pp} & \mathbf{b}_{p}^{T} \\ \mathbf{b}_{p} & \mathbf{B} \end{pmatrix} \begin{pmatrix} \xi \end{pmatrix} & \begin{array}{c} \text{matrix} \\ \text{(fluctuations)} \\ \end{array} \end{aligned}$$

M. Ceriotti, G. Bussi, M. Parrinello, JCTC 2010, 6, 1170-1180 (http://gle4md.berlios.de/)

$$\begin{aligned} \dot{x} &= p/m & \mathbf{A}_{p}\mathbf{C}_{p} + \mathbf{C}_{p}\mathbf{A}_{p}^{T} = \mathbf{B}_{p}\mathbf{B}_{p}^{T} \\ \begin{pmatrix} \dot{p} \\ \dot{\mathbf{s}} \end{pmatrix} &= \begin{pmatrix} -V'(x) \\ 0 \end{pmatrix} - \begin{pmatrix} a_{pp} & \mathbf{a}_{p}^{T} \\ \mathbf{a}_{p} & \mathbf{A} \end{pmatrix} \begin{pmatrix} p \\ \mathbf{s} \end{pmatrix} + \begin{pmatrix} b_{pp} & \mathbf{b}_{p}^{T} \\ \mathbf{b}_{p} & \mathbf{B} \end{pmatrix} \begin{pmatrix} \xi \end{pmatrix} & \begin{array}{c} \text{matrix} \\ \text{(fluctuations)} \\ \end{array} \end{aligned}$$

Integrating out the s degrees of freedom

$$\dot{p} = -\frac{\partial V}{\partial x} - \int_{\infty}^{t} \frac{K(t-\tau)p(\tau) + \zeta(t)}{\int_{\text{memory kernel}}} \int_{\infty}^{t} \frac{K(t-\tau)p(\tau) + \zeta(t)}{\zeta(t)} dt = 0$$

M. Ceriotti, G. Bussi, M. Parrinello, JCTC 2010, 6, 1170-1180 (http://gle4md.berlios.de/)

$$\dot{x} = p/m \qquad \mathbf{A}_{p}\mathbf{C}_{p} + \mathbf{C}_{p}\mathbf{A}_{p}^{T} = \mathbf{B}_{p}\mathbf{B}_{p}^{T}$$

$$\begin{pmatrix} \dot{p} \\ \dot{\mathbf{s}} \end{pmatrix} = \begin{pmatrix} -V'(x) \\ 0 \end{pmatrix} - \begin{pmatrix} a_{pp} & \mathbf{a}_{p}^{T} \\ \mathbf{a}_{p} & \mathbf{A} \end{pmatrix} \begin{pmatrix} p \\ \mathbf{s} \end{pmatrix} + \begin{pmatrix} b_{pp} & \mathbf{b}_{p}^{T} \\ \mathbf{b}_{p} & \mathbf{B} \end{pmatrix} (\xi) \qquad \begin{array}{c} \text{matrix} \\ \text{(fluctuations)} \end{array}$$

Integrating out the s degrees of freedom

$$\dot{p} = -\frac{\partial V}{\partial x} - \int_{\infty}^{\iota} \frac{K(t-\tau)p(\tau) + \zeta(t)}{\int_{\text{memory kernel}} \zeta(t)} \int_{\infty}^{\iota} \frac{K(t-\tau)p(\tau) + \zeta(t)}{\zeta(t)} dt = 0$$

- Freedom to choose A and C
  - If one wants canonical sampling fluctuation dissipation theorem:  $H(t) = \langle \zeta(t)\zeta(0) \rangle = k_B T K(t)$   $\mathbf{C}_p = k_B T$
  - But FDP can be broken to simulate, e.g. quantum distributions/fluctuations or to converge PIMD simulations using less beads

M. Ceriotti, G. Bussi, M. Parrinello, JCTC 2010, 6, 1170-1180 (http://gle4md.berlios.de/)

$$\dot{x} = p/m \qquad \mathbf{A}_{p}\mathbf{C}_{p} + \mathbf{C}_{p}\mathbf{A}_{p}^{T} = \mathbf{B}_{p}\mathbf{B}_{p}^{T}$$

$$\begin{pmatrix} \dot{p} \\ \dot{\mathbf{s}} \end{pmatrix} = \begin{pmatrix} -V'(x) \\ 0 \end{pmatrix} - \begin{pmatrix} a_{pp} & \mathbf{a}_{p}^{T} \\ \mathbf{a}_{p} & \mathbf{A} \end{pmatrix} \begin{pmatrix} p \\ \mathbf{s} \end{pmatrix} + \begin{pmatrix} b_{pp} & \mathbf{b}_{p}^{T} \\ \mathbf{b}_{p} & \mathbf{B} \end{pmatrix} (\xi) \qquad \begin{array}{c} \text{matrix} \\ \text{(fluctuations)} \end{array}$$

Integrating out the s degrees of freedom

$$\dot{p} = -\frac{\partial V}{\partial x} - \int_{\infty}^{\iota} \frac{K(t-\tau)p(\tau) + \zeta(t)}{\int_{\text{memory kernel}} \zeta(t)} \int_{\infty}^{\iota} \frac{K(t-\tau)p(\tau) + \zeta(t)}{\zeta(t)} dt = 0$$

- Freedom to choose A and C
  - If one wants canonical sampling fluctuation dissipation theorem:  $H(t) = \langle \zeta(t)\zeta(0) \rangle = k_B T K(t)$   $\mathbf{C}_p = k_B T$
  - But FDP can be broken to simulate, e.g. quantum distributions/fluctuations or to converge PIMD simulations using less beads

Download input matrices (in FHI-aims input format) from: <u>http://gle4md.berlios.de</u>

### What can we do now?

T = 300K

Classical

Quantum thermostat (qt)

WRAPPI: PIMD + qt





## Conclusions

## Conclusions

- Helix stability in isolated polyalanine emerges as a conspiracy of:
  - Favorable hydrogen bond network
  - van der Waals interactions
  - Vibrational entropy at finite T

## Conclusions

- Helix stability in isolated polyalanine emerges as a conspiracy of:
  - Favorable hydrogen bond network
  - van der Waals interactions
  - Vibrational entropy at finite T
- Density functional theory (PBE+vdW) for peptide secondary structure:
- Helix stability in isolated polyalanine emerges as a conspiracy of:
  - Favorable hydrogen bond network
  - van der Waals interactions
  - Vibrational entropy at finite T
- Density functional theory (PBE+**vdW**) for peptide secondary structure:
  - Captures relevant interactions responsible for the stability of secondary structure motifs

- Helix stability in isolated polyalanine emerges as a conspiracy of:
  - Favorable hydrogen bond network
  - van der Waals interactions
  - Vibrational entropy at finite T
- Density functional theory (PBE+vdW) for peptide secondary structure:
  - Captures relevant interactions responsible for the stability of secondary structure motifs
  - Verification of structure predictions against experiment (vibrational spectroscopy), if anharmonic effects are taken into account

- Helix stability in isolated polyalanine emerges as a conspiracy of:
  - Favorable hydrogen bond network
  - van der Waals interactions
  - Vibrational entropy at finite T
- Density functional theory (PBE+vdW) for peptide secondary structure:
  - Captures relevant interactions responsible for the stability of secondary structure motifs
  - Verification of structure predictions against experiment (vibrational spectroscopy), if anharmonic effects are taken into account
  - Transferable accuracy across wide configurational landscape

- Helix stability in isolated polyalanine emerges as a conspiracy of:
  - Favorable hydrogen bond network
  - van der Waals interactions
  - Vibrational entropy at finite T
- Density functional theory (PBE+vdW) for peptide secondary structure:
  - Captures relevant interactions responsible for the stability of secondary structure motifs
  - Verification of structure predictions against experiment (vibrational spectroscopy), if anharmonic effects are taken into account
  - Transferable accuracy across wide configurational landscape
- Quantum nuclear effects must be included in several situations now possible in FHI-aims!

- Helix stability in isolated polyalanine emerges as a conspiracy of:
  - Favorable hydrogen bond network
  - van der Waals interactions
  - Vibrational entropy at finite T
- Density functional theory (PBE+vdW) for peptide secondary structure:
  - Captures relevant interactions responsible for the stability of secondary structure motifs
  - Verification of structure predictions against experiment (vibrational spectroscopy), if anharmonic effects are taken into account
  - Transferable accuracy across wide configurational landscape
- Quantum nuclear effects must be included in several situations now possible in FHI-aims!



