

# Infrared spectra of small biomolecules from first-principle molecular dynamics simulations and effective normal mode analysis

R. Vuilleumier, M.-P. Gaigeot and D. Borgis

*Département de chimie, Ecole Normale Supérieure (Paris)*

*LAMBE, Université d'Evry*

# Introduction

- Normal mode analysis at the optimized geometry

Determination of normal modes from diagonalisation of the Hessian

Modes (phonons in crystals) and frequencies at T=0

No anharmonic effects

- Vibrational spectra from MD

IR spectra including temperature and environmental effects

Obtained from the current-current autocorrelation function (linear response theory)

$$\omega \varepsilon''(\omega) = \frac{2\pi}{3V} \beta \int_{-\infty}^{+\infty} \langle \mathbf{j}(0) \cdot \mathbf{j}(t) \rangle e^{i\omega t} dt$$

# ab initio Molecular Dynamics

forces on the atoms evaluated *on the fly* by an electronic structure calculation

Density Functional Theory is used for the electronic structure calculation

Plane waves, pseudopotentials

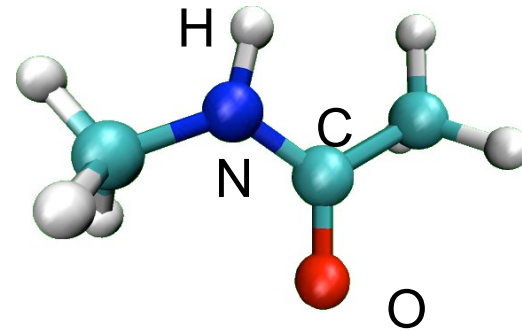
Electrons and nuclei propagated simultaneously by using an extended Lagrangian: Car-Parrinello Molecular Dynamics

Dipole moment also evaluated *on the fly*

Good agreement with experimental spectra in a large range of situations

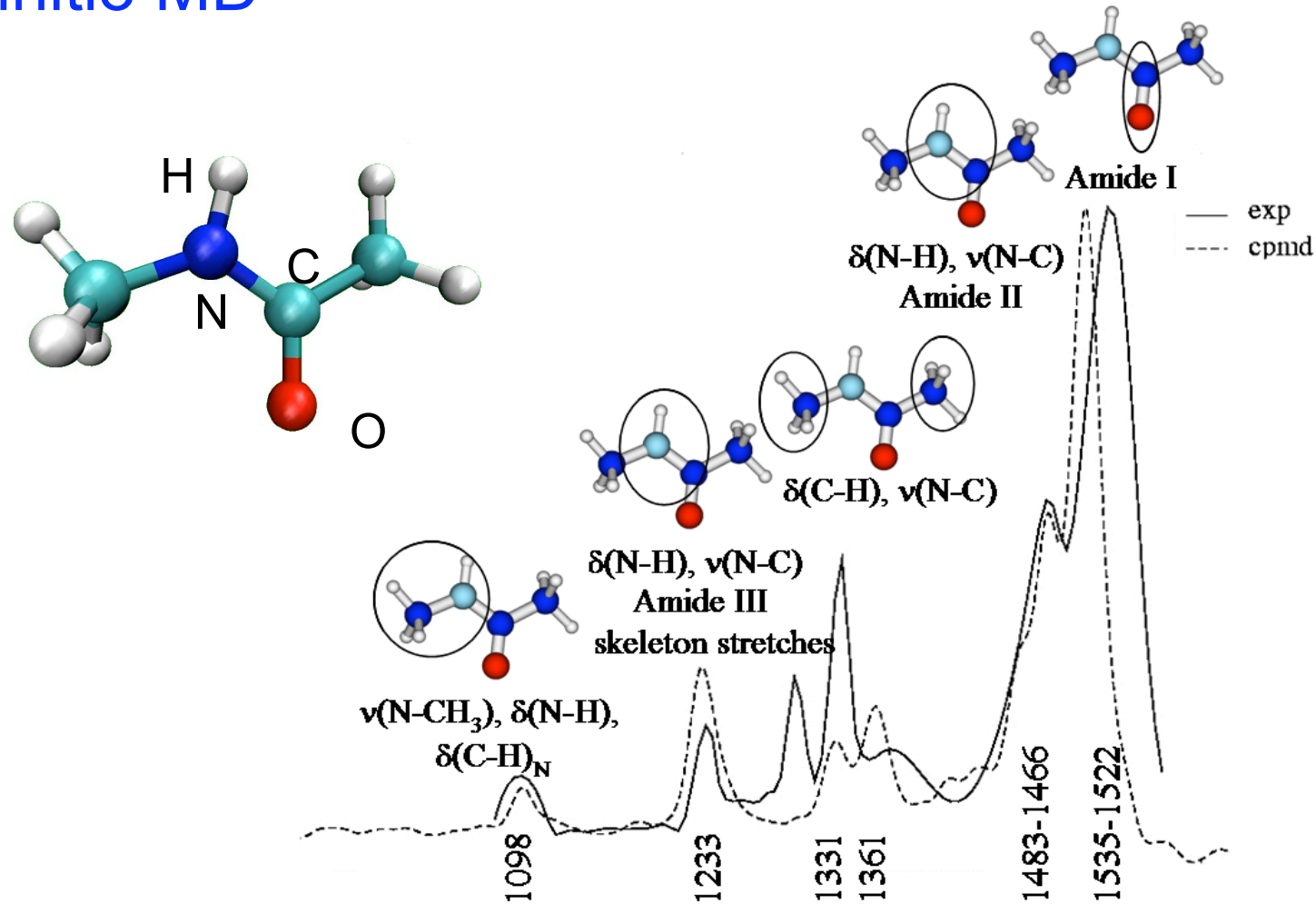
# IR spectrum of N-methyl acetamide in solution by ab initio MD

- NMA in 50 water molecules
- BLYP functional
- cutoff=70 Ry, TM pp
- 7 ps of NVE dynamics at 300 K (average)



- IR spectrum from TF of auto-correlation function of NMA dipole
- Dipole moment of NMA obtained from assigning Maximally localized Wannier orbitals

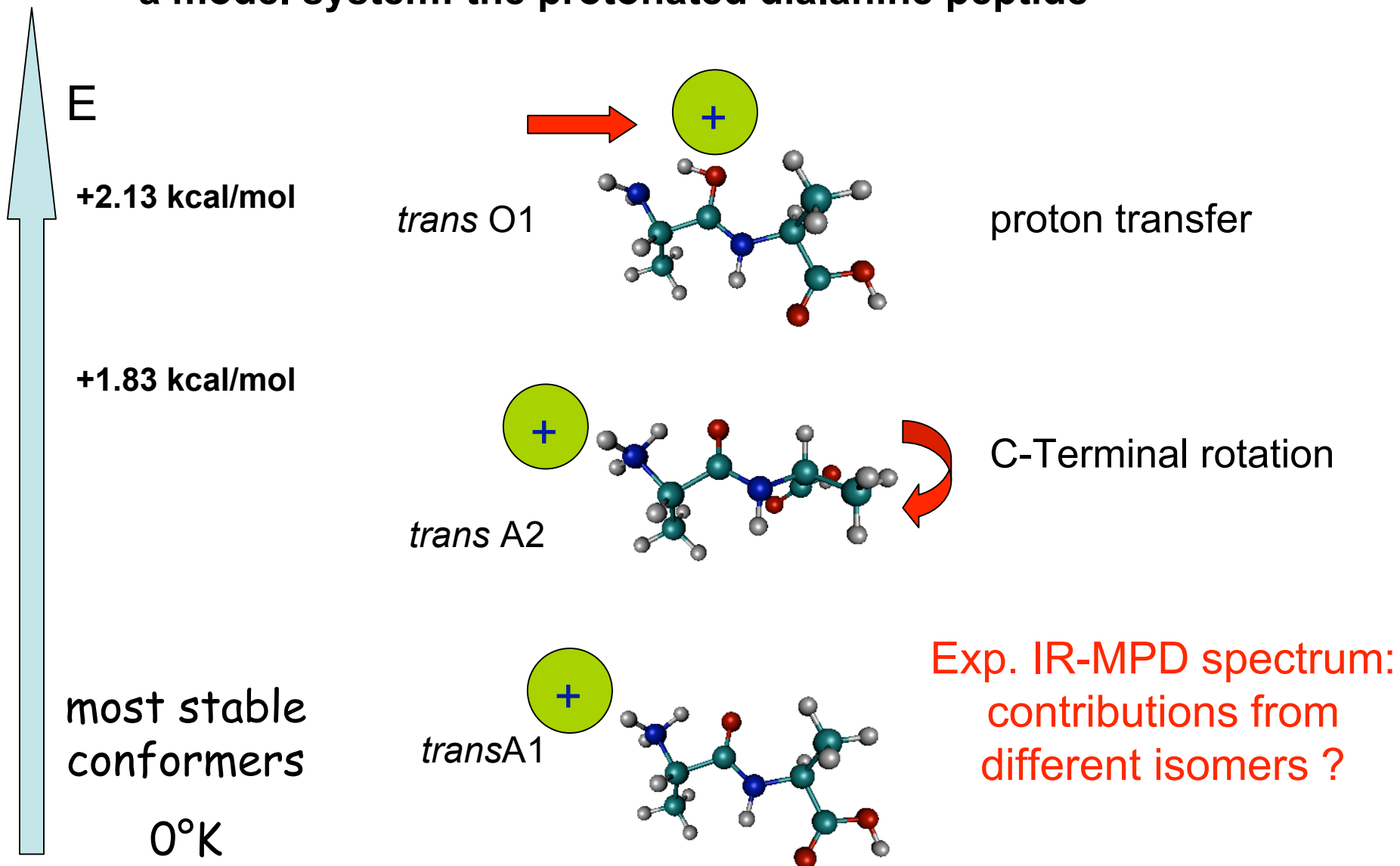
# IR spectrum of N-methyl acetamide in solution by ab initio MD



Interpretation of these spectra ?

# IRMPD 300K gas phase spectroscopy of the $\text{Ala}_2\text{H}^+$ peptide

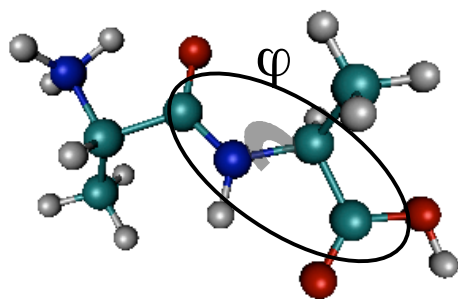
a model system: the protonated dialanine peptide



# IRMPD 300K gas phase spectroscopy of the Ala<sub>2</sub>H<sup>+</sup> peptide

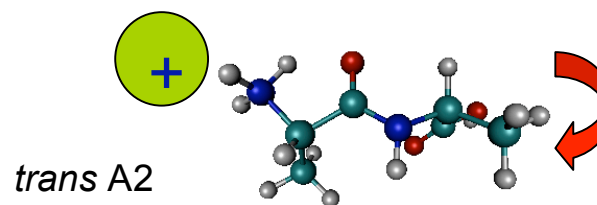
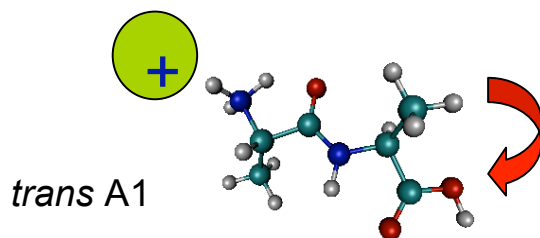
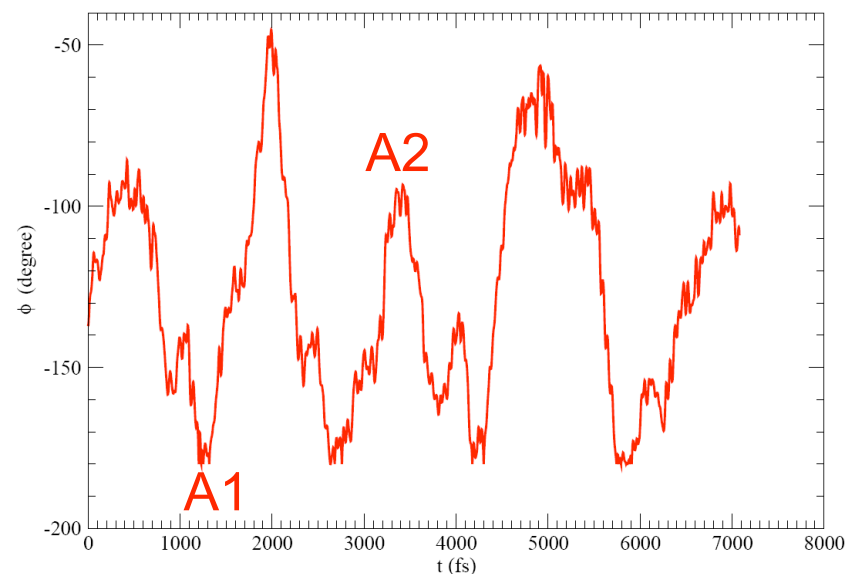
## CPMD at 300 K

Continuous conformational dynamics  
between 2 conformers transA1 et transA2



$$\phi_{\text{transA1}} = -162^\circ$$

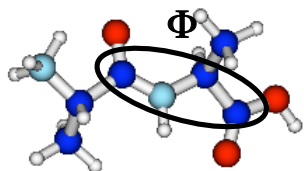
$$\phi_{\text{transA2}} = -74^\circ$$



The relative weight of each conformation probed during the dynamics  
will be naturally taken into account  
in the calculation of the infrared spectrum

# IRMPD 300K gas phase spectroscopy of the Ala<sub>2</sub>H<sup>+</sup> peptide

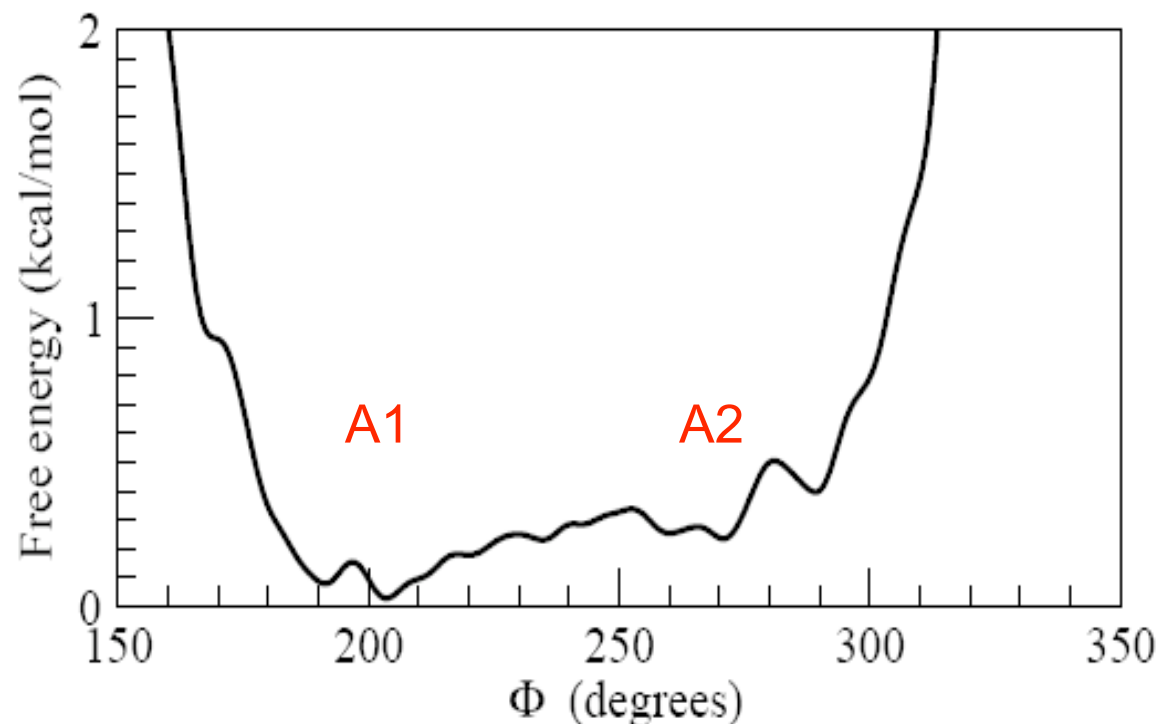
Free energy profile measured along the 300K dynamics



$$G = -kT \ln P(\Phi)$$

$$\Phi_{\text{transA1}} = 198^\circ$$

$$\Phi_{\text{transA2}} = 284^\circ$$



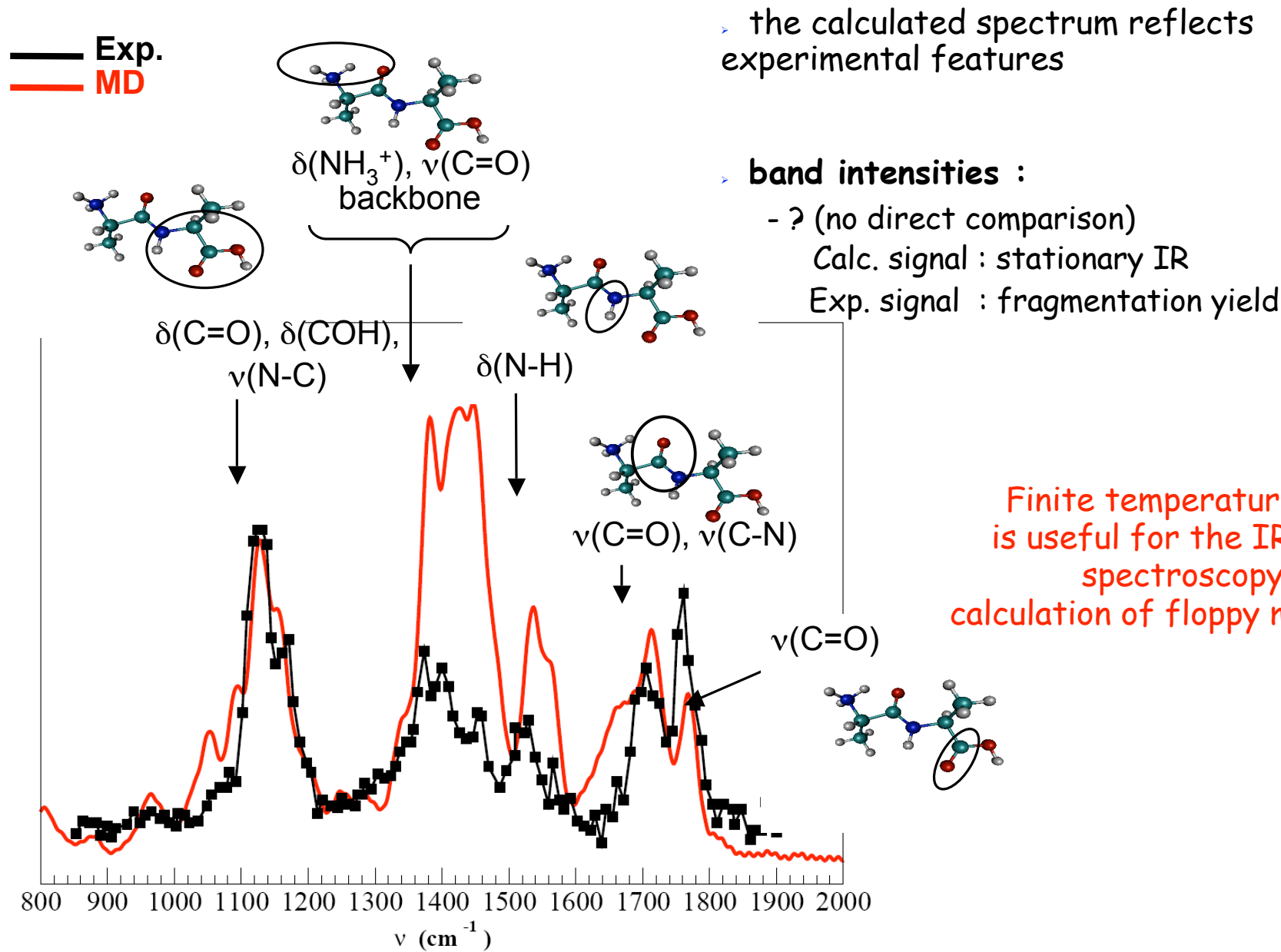
No energy barrier for going from transA1 to transA2 :  $\Delta G \approx 0.3$  kcal/mol

A reminder :  $\Delta G \approx 2$  kcal/mol with static 0K calculations in the harmonic approximation

Entropic effects are important in structural equilibria



# IRMPD 300K gas phase spectroscopy of the Ala<sub>2</sub>H<sup>+</sup> peptide



> the calculated spectrum reflects experimental features

> **band intensities :**

- ? (no direct comparison)

Calc. signal : stationary IR

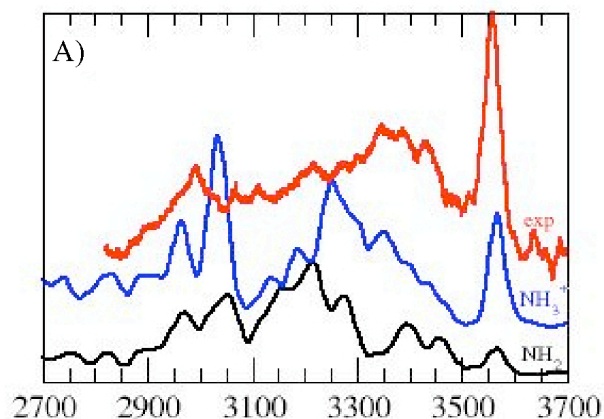
Exp. signal : fragmentation yield

Finite temperature MD is useful for the IR-MPD spectroscopy calculation of floppy molecules

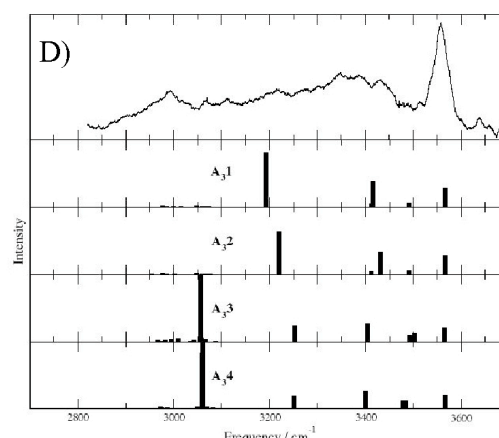
Average over 10 different dynamics

# IR-MPD gas phase spectroscopy of the Ala<sub>3</sub>H<sup>+</sup> peptide

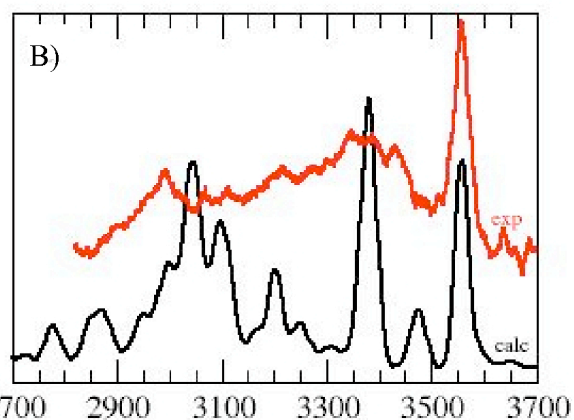
Exps. Oxford Snoek, Vaden, de Boer [J.Chem. Theor. Comput. 5, 1068-78 \(2009\)](#)



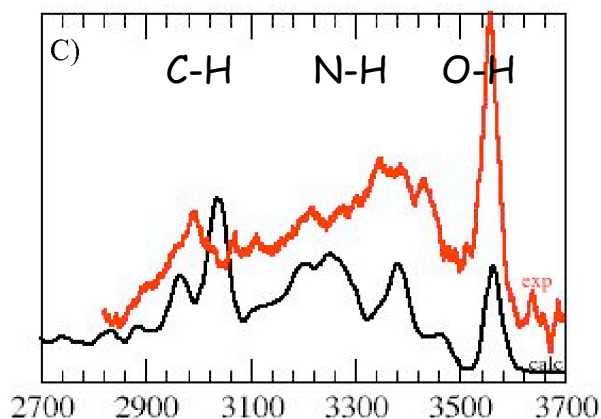
Dynamics : NH<sub>2</sub> & NH<sub>3</sub><sup>+</sup> families



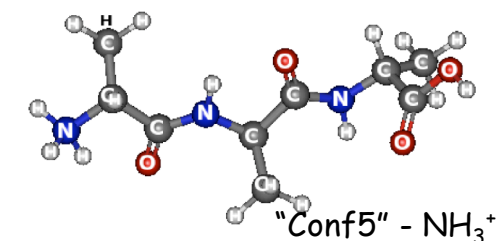
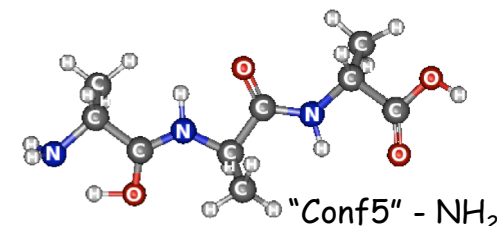
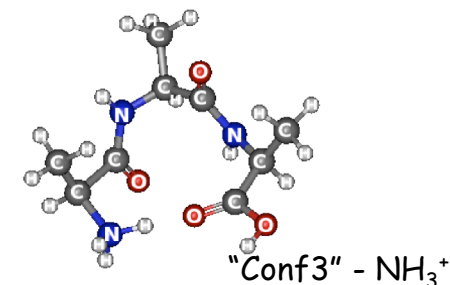
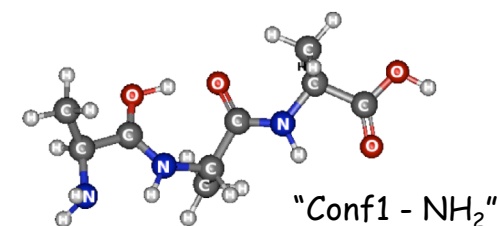
Static calculations



Dynamics : folded NH<sub>3</sub><sup>+</sup> family



Dynamics : average over all families



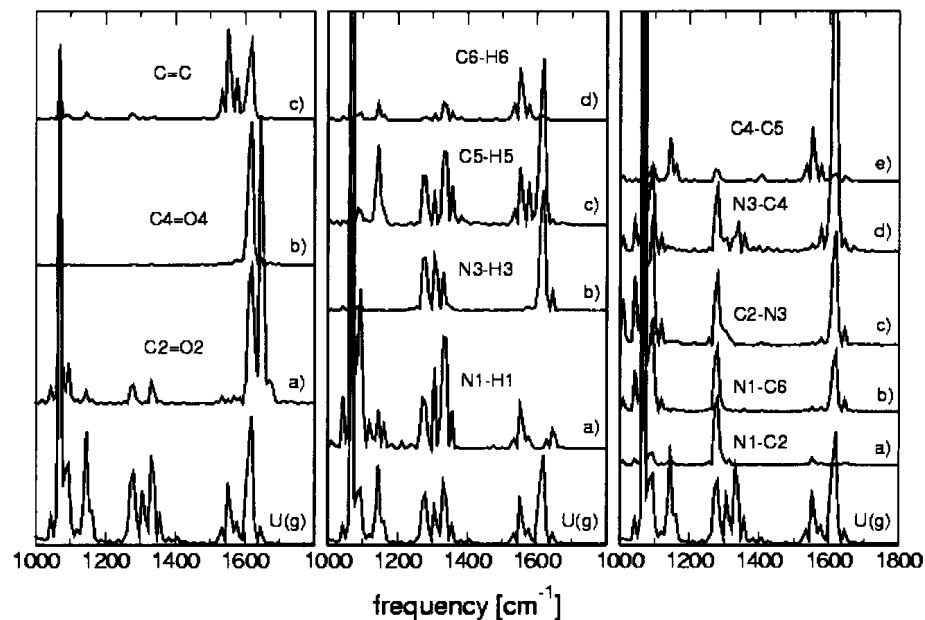
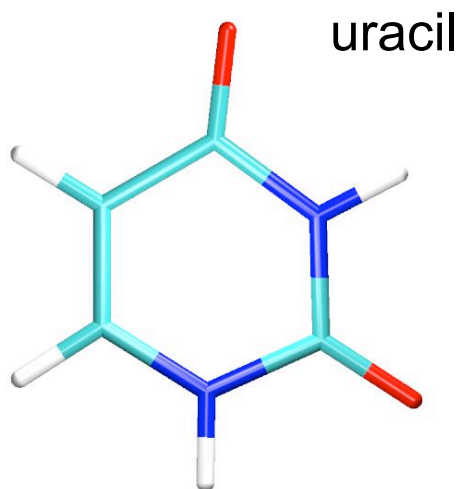
Exp. Spectrum features obtained as a result of all conformations  
The dynamics of the N-H groups give rise to the complex 3100-3500 cm<sup>-1</sup> band  
This complexity can not be obtained without dynamics

# Interpretation of the spectra

- What is the nature of the molecular motion underlying the IR bands?
- What is the origin of the band intensity?

# I – band by band decomposition of IR spectra

- no easy **interpretation of bands** in terms of internal motions (complex molecules with number of atoms larger than 4-5)
- interpretation usually based on comparison with **velocity autocorrelation functions**



vibrational density of states of uracil in aqueous solution (CPMD)  
from M.-P. Gaigeot and M. Sprik, *JPC B* **107**, 10344 (2003)

# Maximally localized modes

- velocities in cartesian coordinates written as **linear combination of modes**

$$\dot{x}_i(t) = Z_{ik}\dot{q}_k(t)$$

- that are decoupled at equilibrium

$$\langle \dot{q}_k(0)\dot{q}_l(0) \rangle = \delta_{kl}$$

- and are **maximally localized in frequency** by minimisation of

$$\Omega^{(n)} = \sum_k \left( \frac{1}{\pi} \int_0^{+\infty} d\omega \omega^{2n} P_k^q(\omega) - \left( \frac{1}{\pi} \int_0^{+\infty} d\omega \omega^n P_k^q(\omega) \right)^2 \right)$$

where  $P_k^q(\omega) = \int_{-\infty}^{+\infty} \langle \dot{q}_k(0)\dot{q}_k(t) \rangle \cos \omega t dt$  is the power spectrum of  $q$

# Generalized Eigenvalue problem

Minimization of  $\Omega^{(n)}$  leads to a generalized eigenvalue problem:

$$\left[ \int d\omega \omega^n P_{kj}^x(\omega) \right] Z_{ji} = \lambda_i \left[ \int d\omega P_{kj}^x(\omega) \right] Z_{ji}$$

with the matrices

$$P_{kj}^x(\omega) = \int \langle \dot{x}_k(0) \dot{x}_j(t) \rangle \cos \omega t dt$$

satisfying

$$\int d\omega P_{kj}^x(\omega) = \langle \dot{x}_k(0) \dot{x}_j(0) \rangle \left( = \frac{k_B T}{m_k} \delta_{kj} \right)$$

## Effective normal modes

Minimizing  $\Omega^{(n)}$  is equivalent, for  $n = 2$ , to solve the generalized eigenvalue problem

$$\langle F_i F_j \rangle z_j = \omega^2 \langle p_i p_j \rangle z_j$$

By integration by part we have

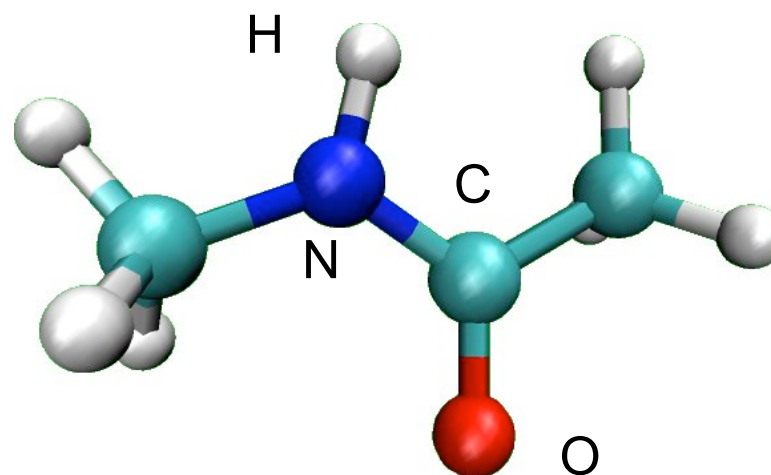
$$\langle F_i F_j \rangle = k_B T \left\langle \frac{\partial^2 E}{\partial x_i \partial x_j} \right\rangle$$

(conformational temperature)

We are thus lead to solve  $\left\langle \frac{\partial^2 E}{\partial x_i \partial x_j} \right\rangle z_j = \omega^2 m_i z_i$

## Use of internal coordinates instead of cartesian coordinates

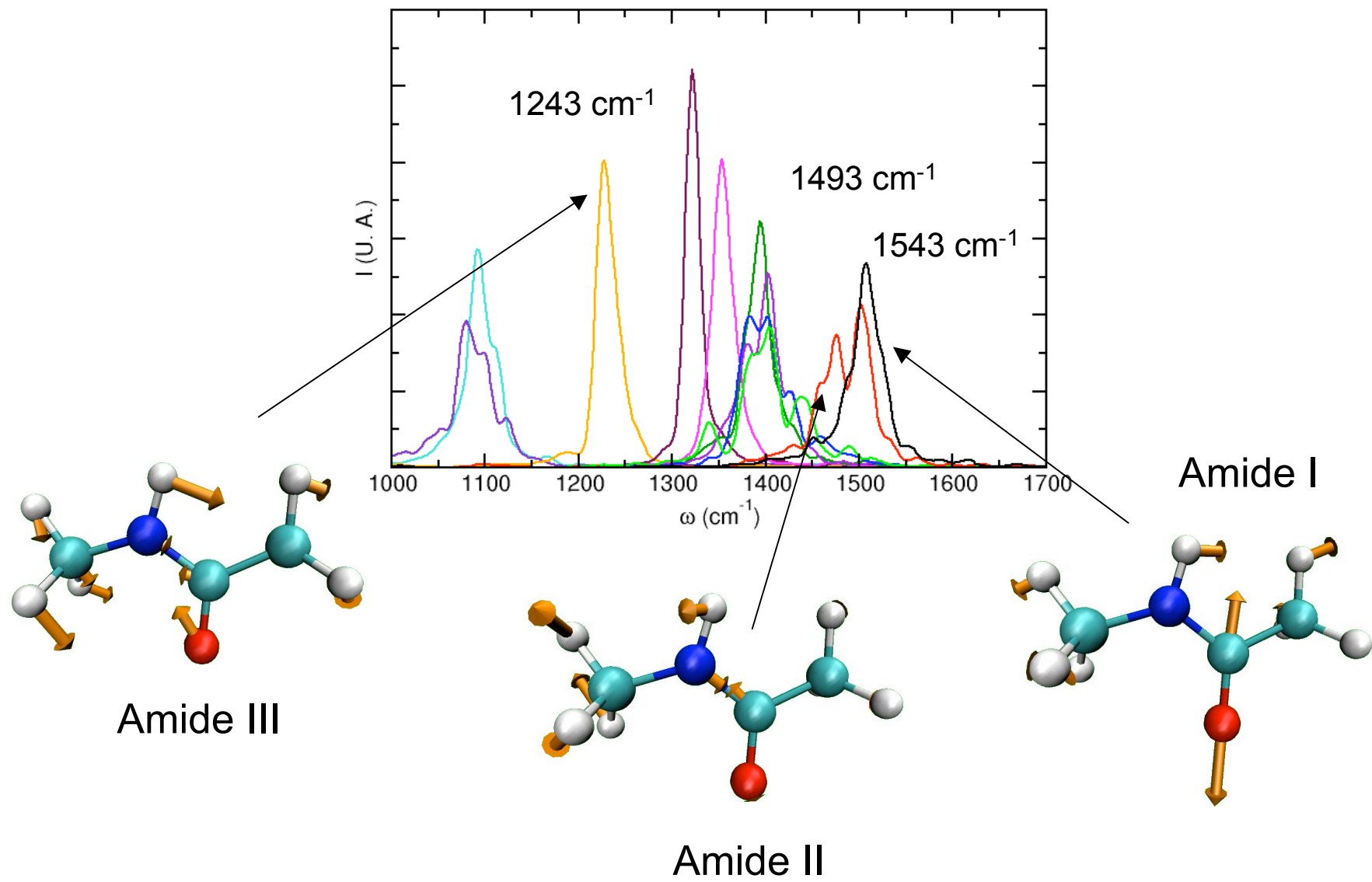
NMA is highly flexible ( $\text{CH}_3$ )



- cartesian coordinates not appropriate
- internal coordinates: 11 stretches, 3 torsions, 14 bends, 2 out-of-plane
- linear combination of internal coordinates to maximally localize the resulting power spectra
- no need for Eckart frame



# Amide modes of NMA in water at 300 K



# Comparison between gas phase and aqueous phase

Potential Energy Distribution (PED) used to describe the effective normal modes

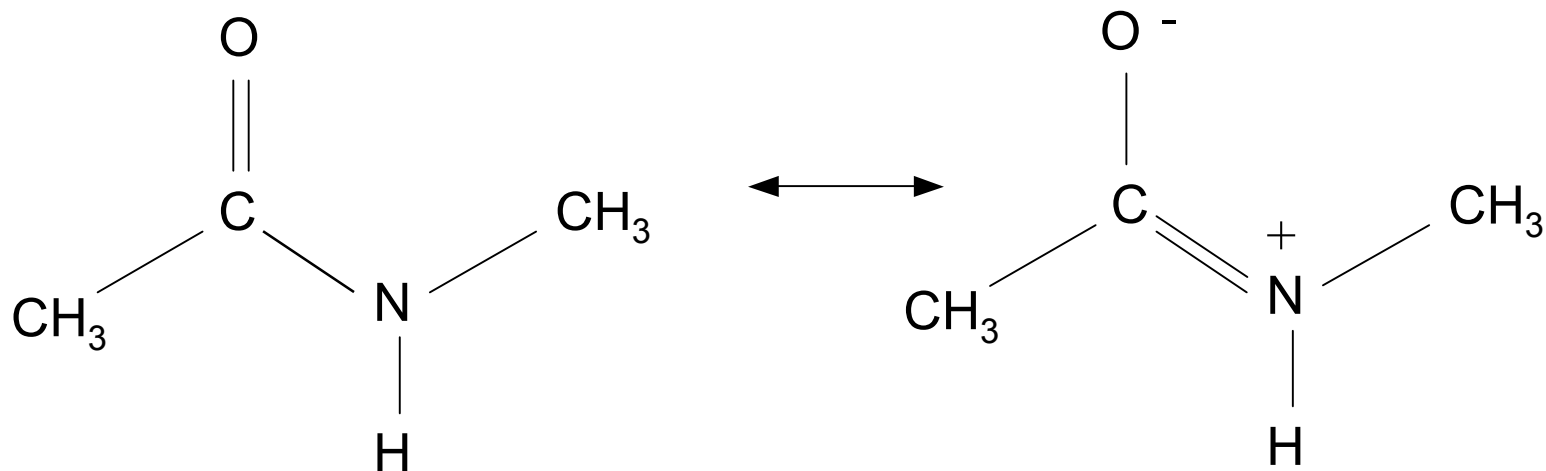
- Aqueous phase:

- Amide I (1516 cm<sup>-1</sup>): st C=O 54%, b NH 37%
- Amide II (1493 cm<sup>-1</sup>): b NH 48%, st CN 24%, C3 sym 12%
- Amide III (1243 cm<sup>-1</sup>): st CN 39%, b NH 35%, st C=O 12%

- Gas phase:

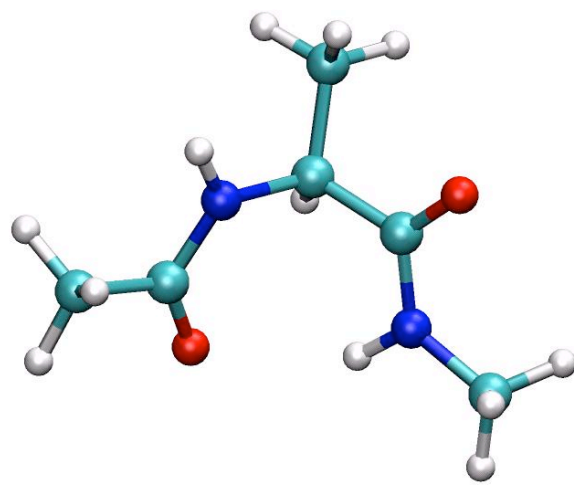
- Amide I (1598 cm<sup>-1</sup>): st C=O 96%, b NH 1,3%
- Amide II (1426 cm<sup>-1</sup>): b NH 47%, st CN 10%, C3 sym 14%
- Amide III (1179 cm<sup>-1</sup>): st CN 26%, b NH 25%, st C=O 21%

## Stabilization of the zwitterionic form in water

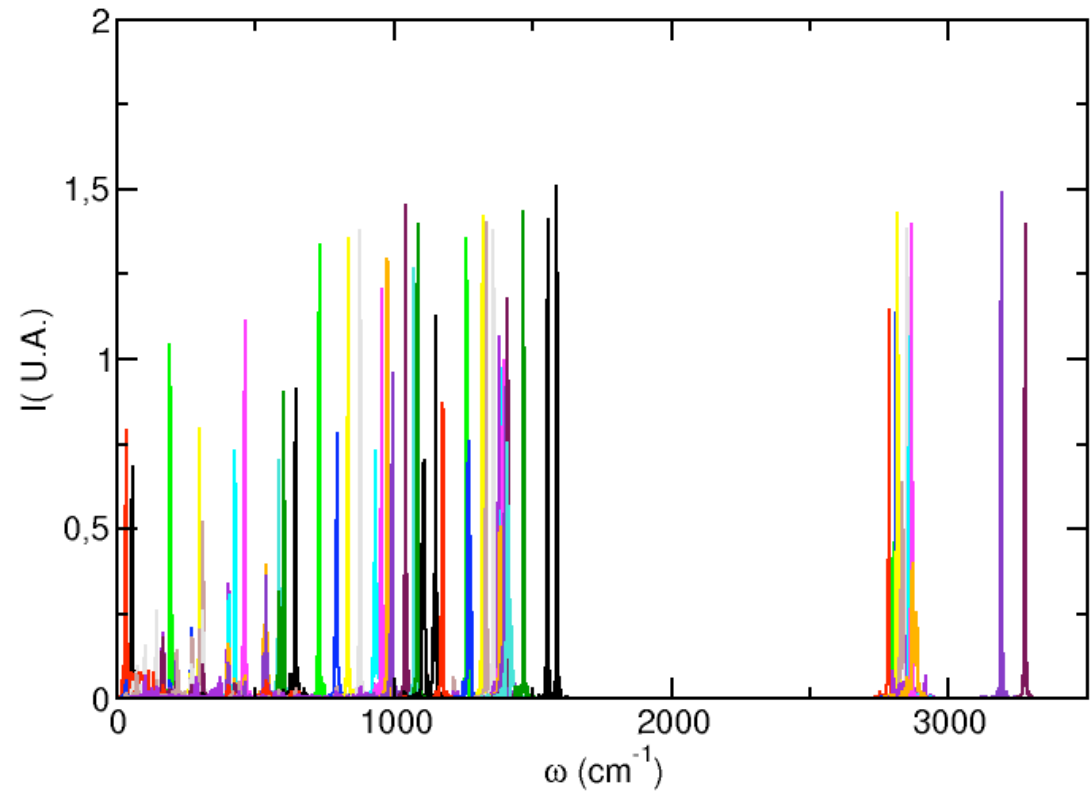


- lower C=O stretch frequency
- stiffer C-N stretch
- in agreement with dipole moment change between gas phase and solution
- partly due to H-bonding to water

# Alanine dipeptide in gas phase (configuration C7 equa)

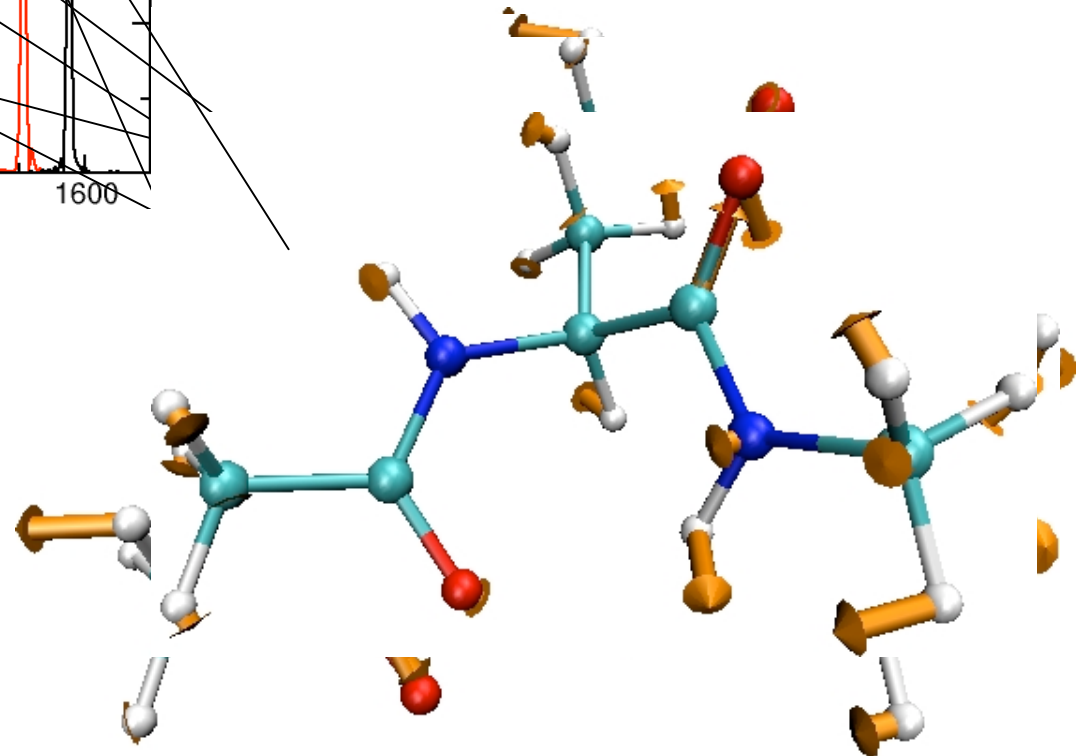
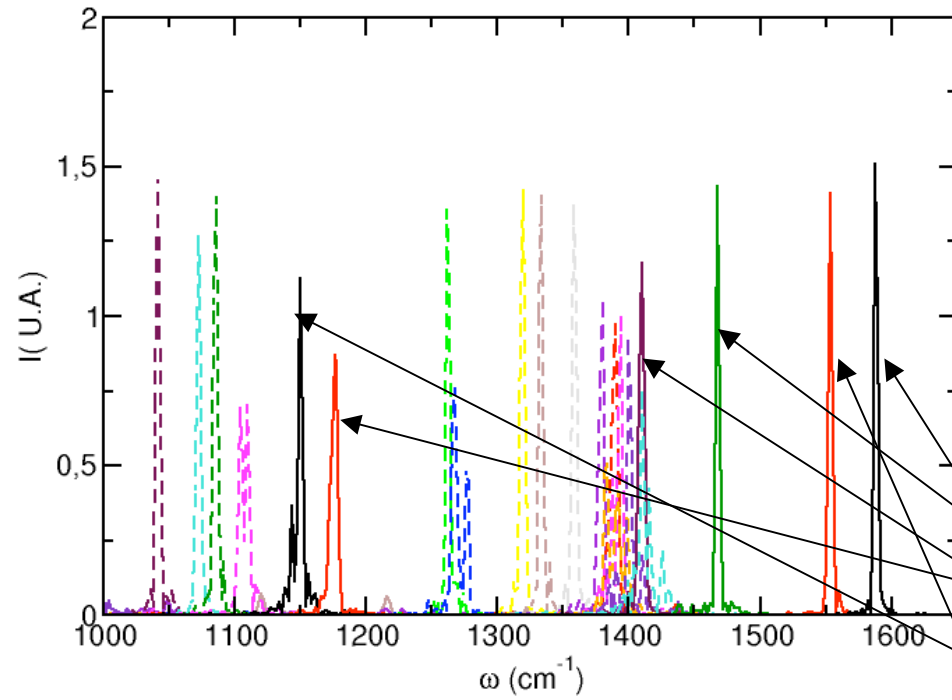


T=20 K



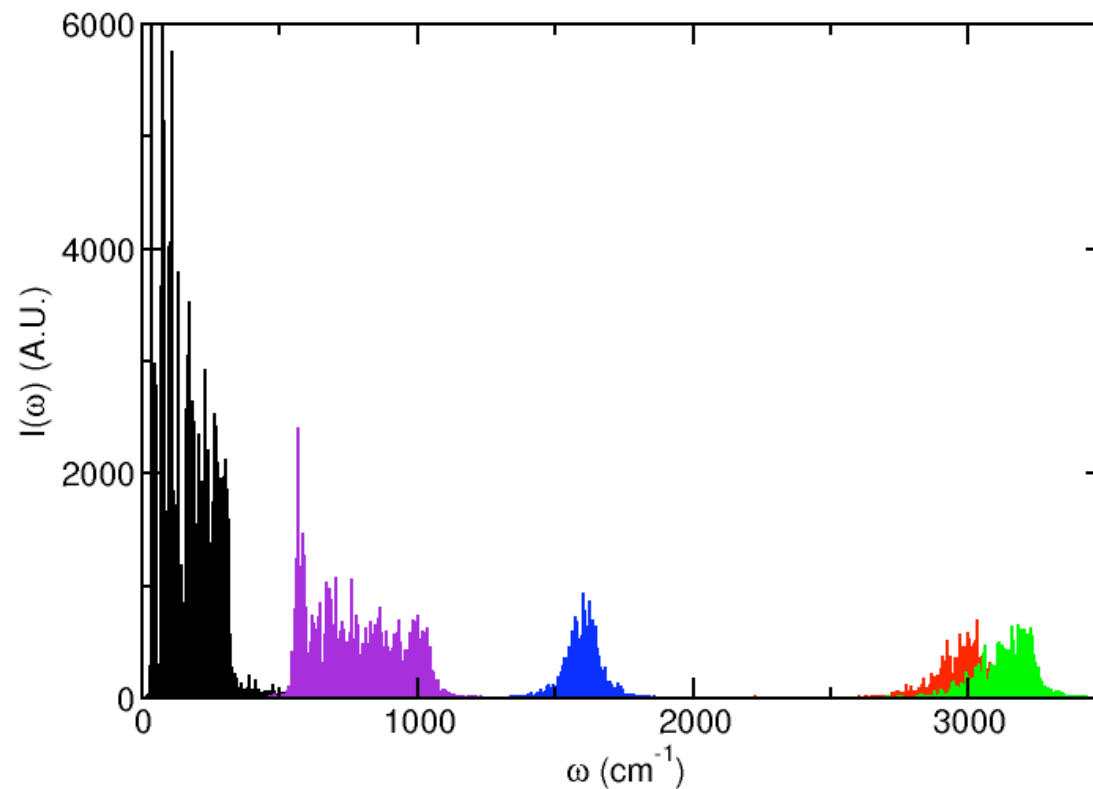
60 modes

# Alanine dipeptide in gas phase: amide modes



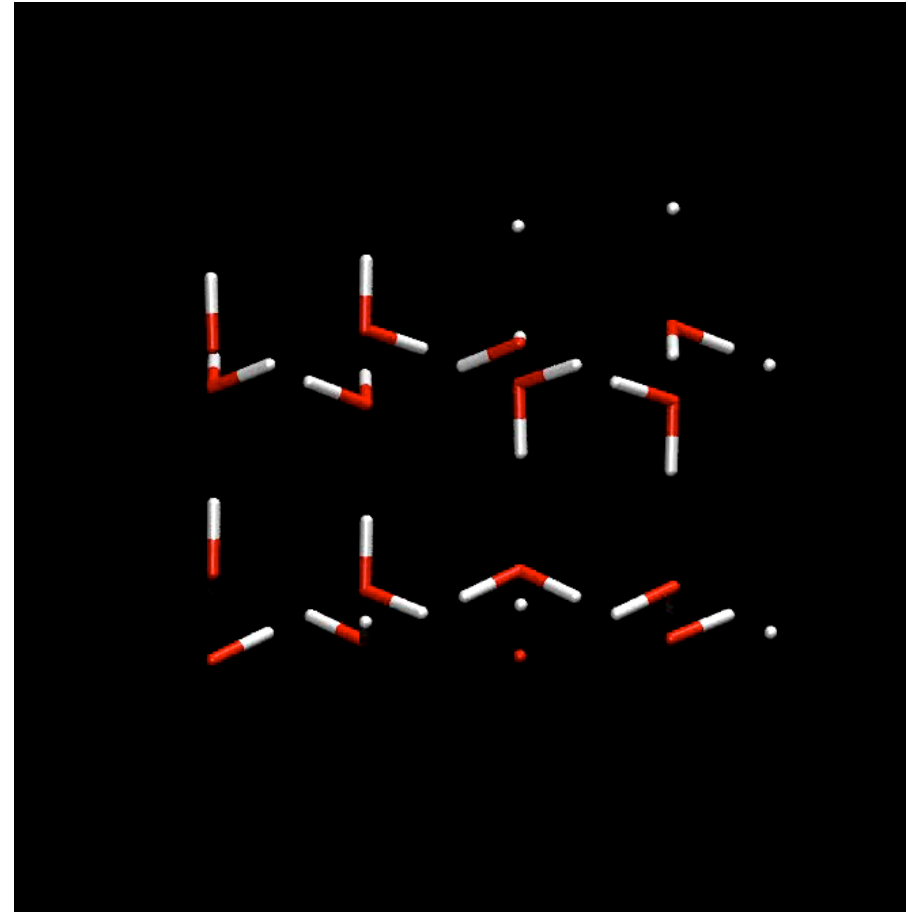
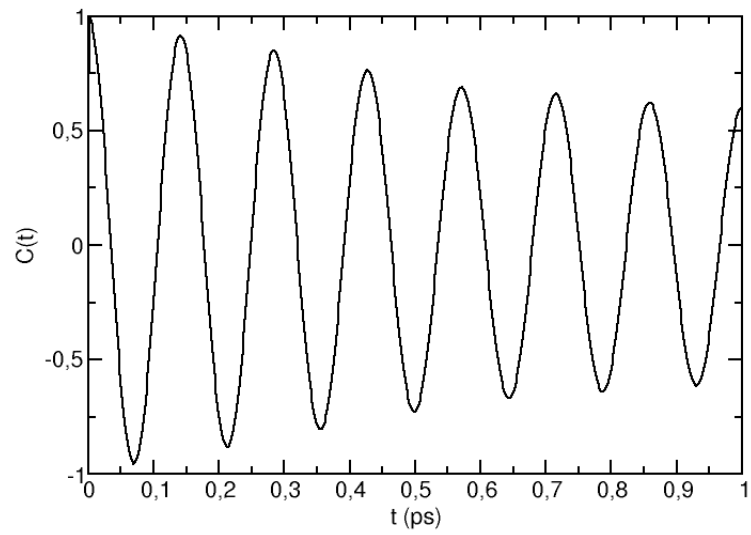
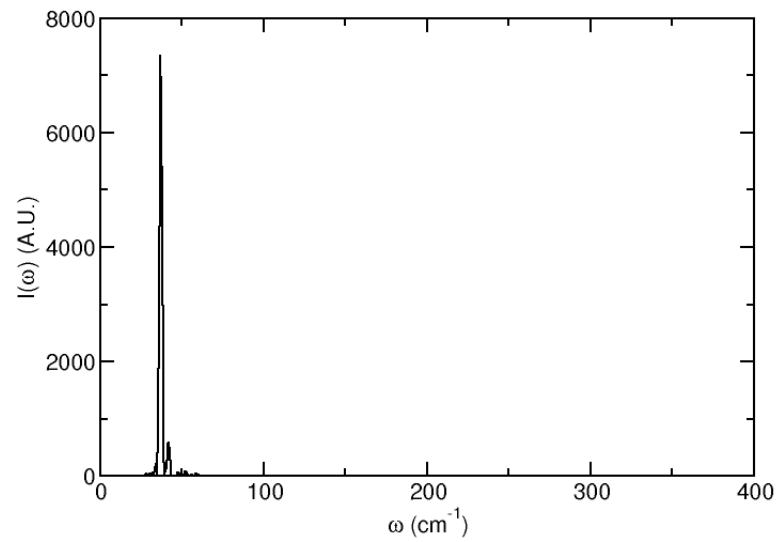
# Effective normal modes (phonons) in ice Ih

- AIMD simulation of 16 water molecules at 240 K (16 ps)
- 141 effective normal modes
- no frozen phonons approximation

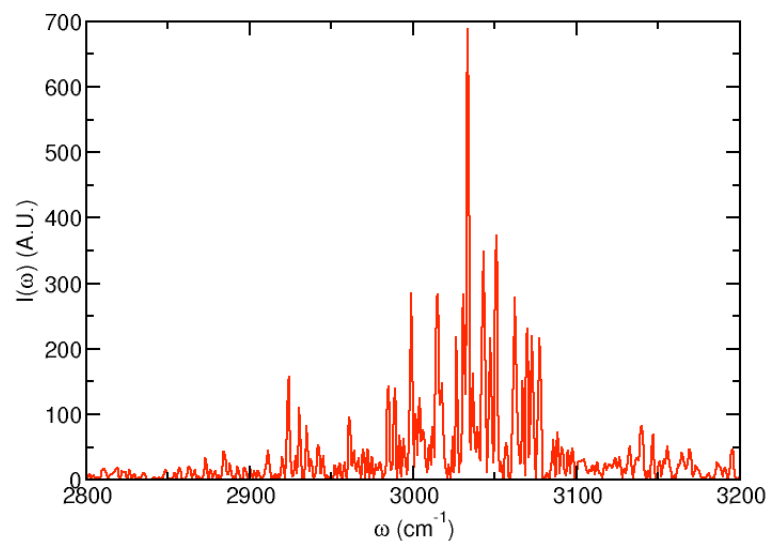


- bands made of 45, 48, 16, 16 and 16 modes respectively

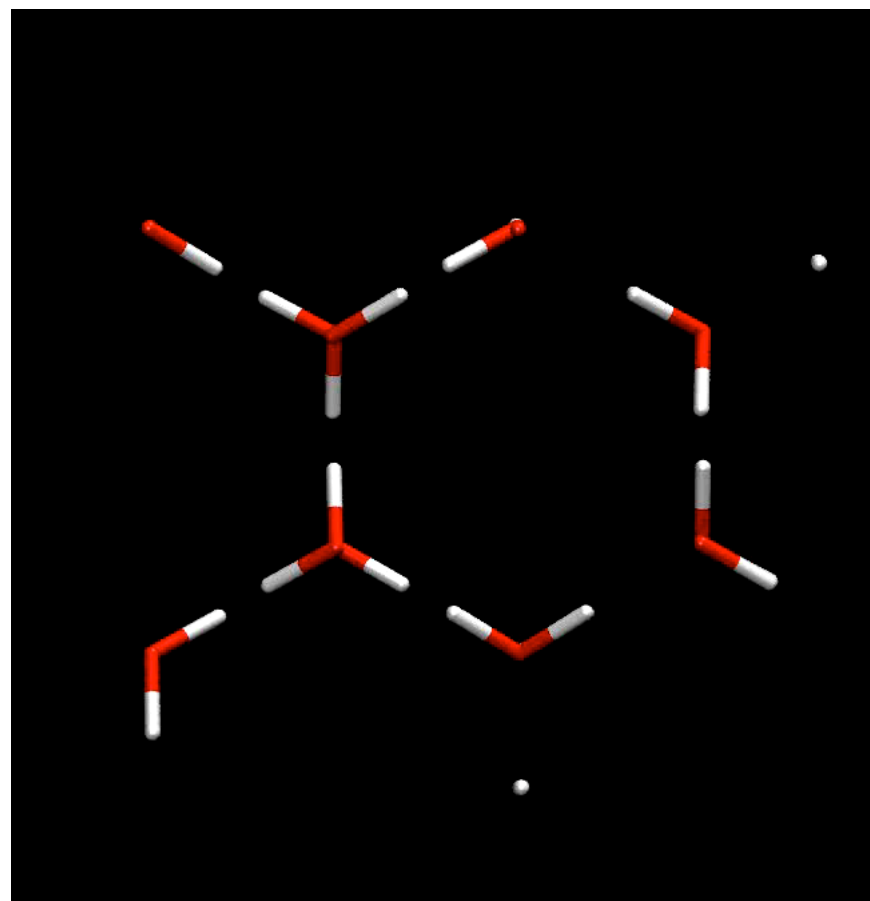
# Lowest modes are very harmonic



## One mode in the OH stretch region



- mainly one peak character but still delocalized
- mixing of symmetric and asymmetric OH stretch





## II - Infrared intensities: atomic polar tensor

The infrared spectrum is obtained from the current-current autocorrelation function:

$$\omega \varepsilon''(\omega) = \frac{2\pi}{3V} \beta \int_{-\infty}^{+\infty} \langle \mathbf{j}(0) \cdot \mathbf{j}(t) \rangle e^{i\omega t} dt$$

with the current:  $\mathbf{j}(t) = \left( \sum_i \left( \frac{\partial \mathbf{M}}{\partial \mathbf{r}} \right) \cdot \mathbf{v}_i \right) (t)$

IR intensity related to the atomic polar tensors

# APT from ab initio Molecular Dynamics

- *Density Functional Theory based* ab initio Molecular Dynamics
- APT calculated in the framework of Density Functional Perturbation Theory

- Maxwell Relation: 
$$\frac{\partial M^\beta}{\partial r_i^\alpha} = \frac{\partial^2 E_{tot}}{\partial r_i^\alpha \partial E^\beta} = \frac{\partial F_i^\alpha}{\partial E^\beta}$$

- Linear response to a uniform electric field (also polarizability)

see also: A. Pasquarello et R. Resta, *Phys. Rev. B* **68**, 174302 (2003)

# Infrared and Raman intensities of bands

APT in localized modes

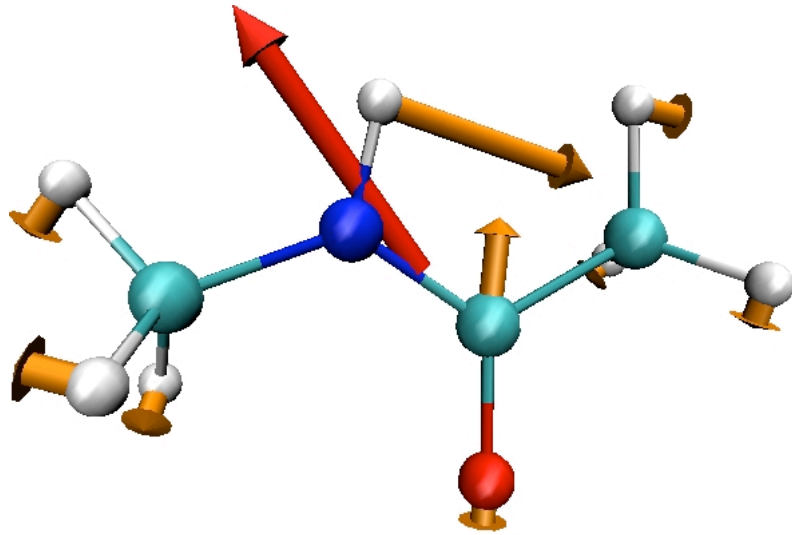
$$\frac{\partial \vec{M}}{\partial q_i} = Z_{ij} \frac{\partial \vec{M}}{\partial x_j} \quad \text{transition moment (vector)}$$

Oscillator strength of mode  $i$

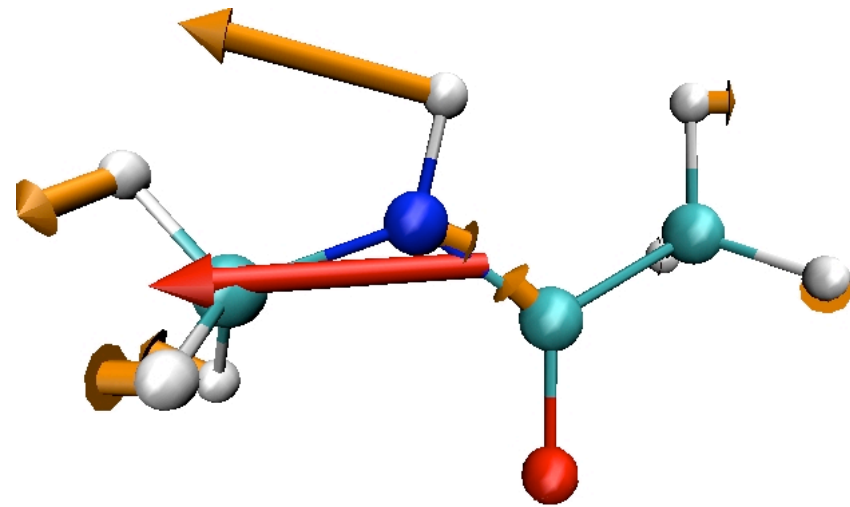
$$f_i \propto \left\langle \left| \frac{\partial \vec{M}}{\partial q_i} \right|^2 \right\rangle \approx \left\langle \left| \frac{\partial \vec{M}}{\partial q_i} \right|^2 \right\rangle \quad \text{(infrared)}$$

$$f_i \propto \left\langle \left| \frac{\partial \bar{\alpha}}{\partial q_i} \right|^2 \right\rangle \quad \text{(Raman)}$$

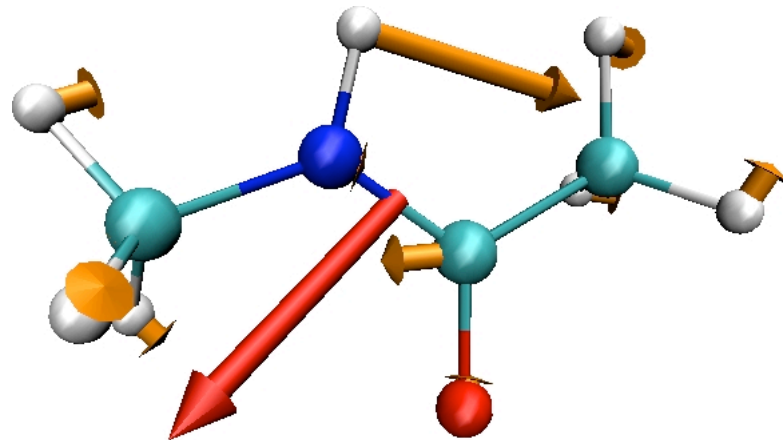
# NMA transition moments and intensities



Amide I:  $I=1$

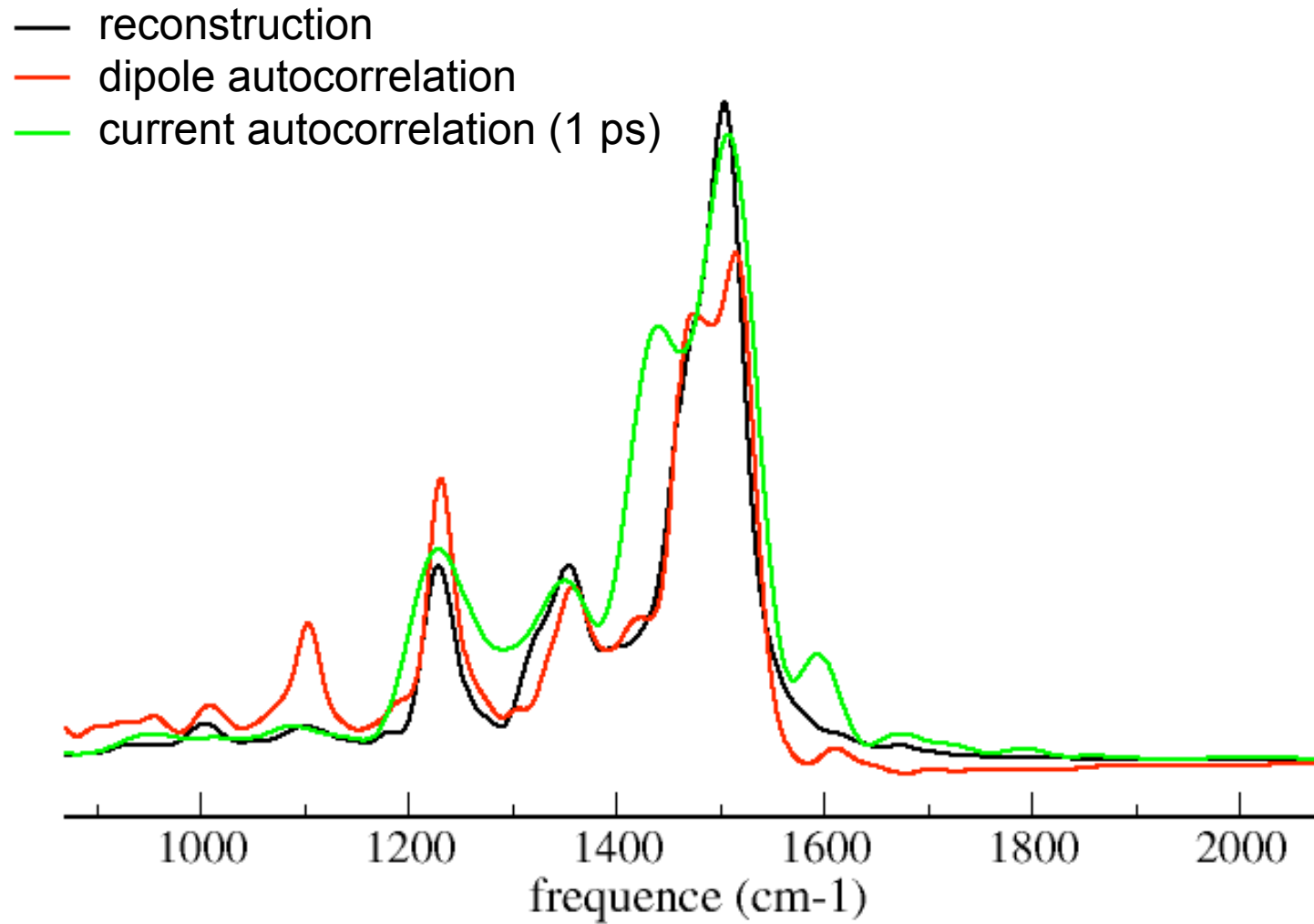


Amide II:  $I=2.06$

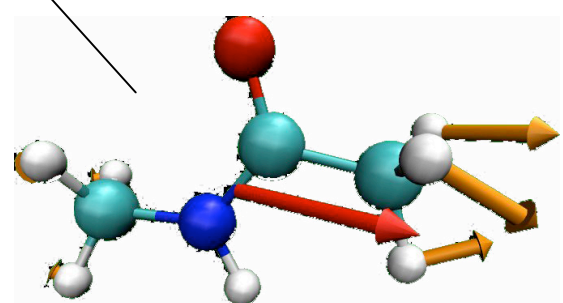
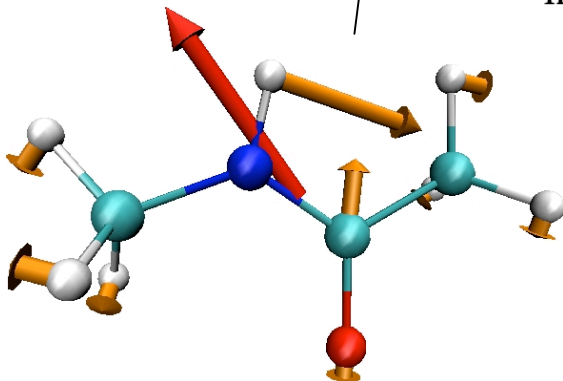
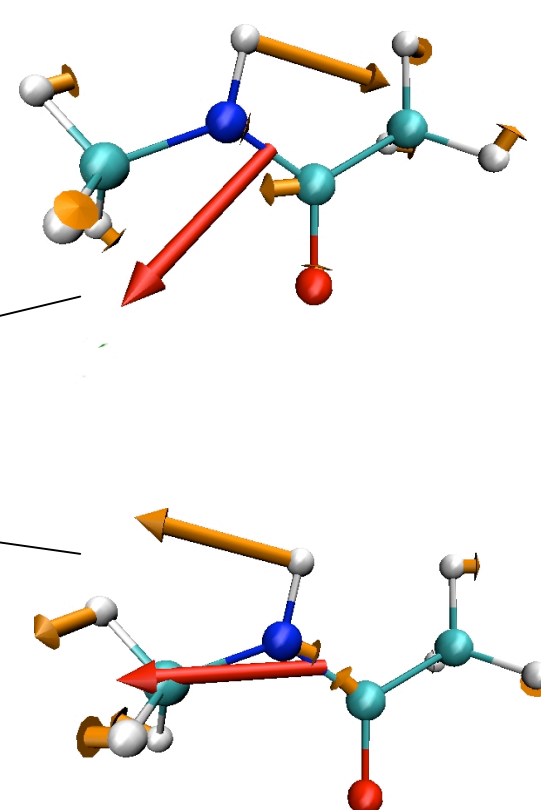
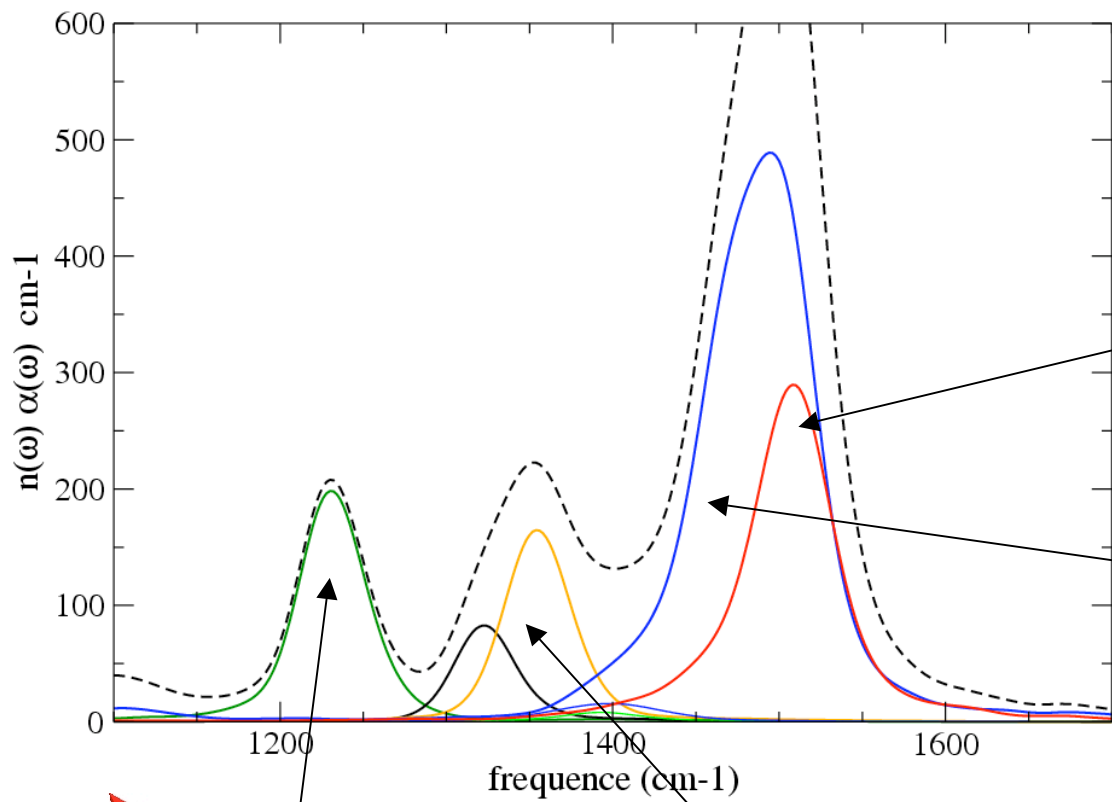


Amide III:  $I=0.53$

# Reconstruction of the Infrared Spectrum of NMA



# Main peaks of NMA spectrum



# Cheap approximate way for calculating IR and Raman intensities

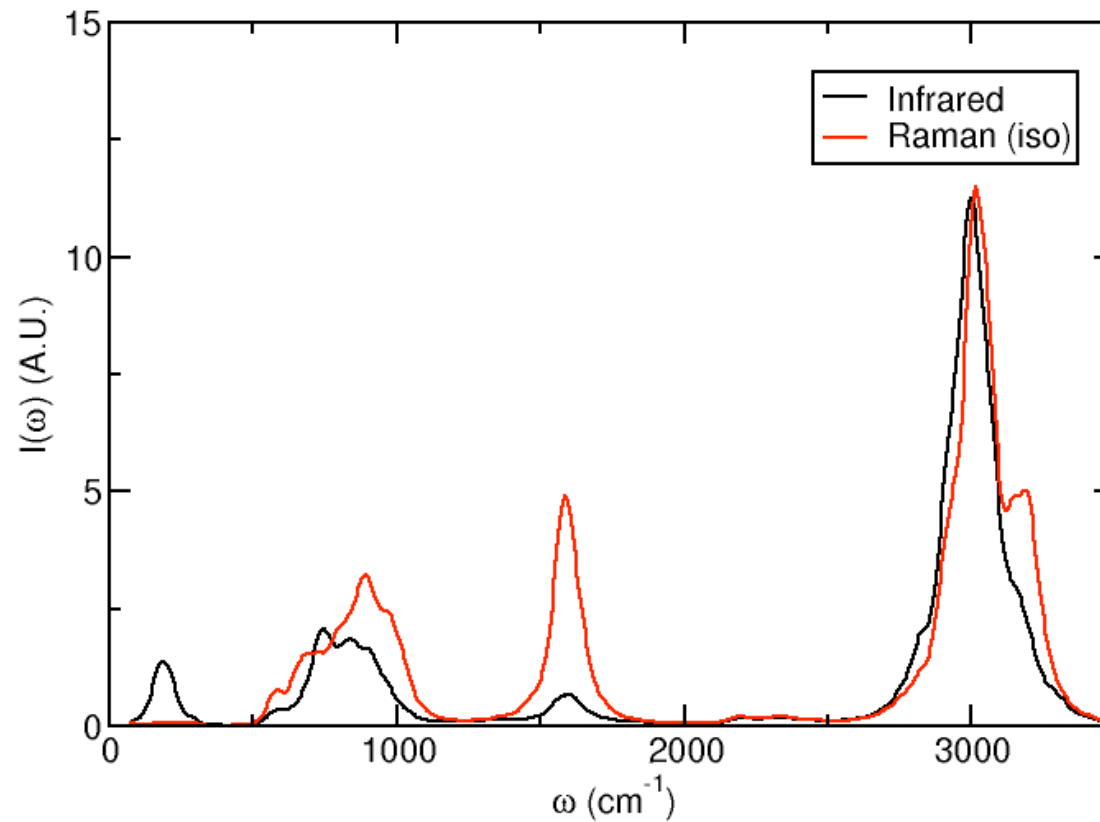
approximate  $\left\langle \frac{\partial \vec{M}}{\partial x_i} \frac{\partial \vec{M}}{\partial x_j} \right\rangle$  by  $\left\langle \frac{\partial \vec{M}}{\partial x_i} \right\rangle \left\langle \frac{\partial \vec{M}}{\partial x_j} \right\rangle$

two useful properties (integration by parts):

$$k_B T \left\langle \frac{\partial \vec{M}}{\partial x_i} \right\rangle = \langle \vec{M} F_i \rangle \quad k_B T \left\langle \frac{\partial \bar{\alpha}}{\partial x_i} \right\rangle = \langle \bar{\alpha} F_i \rangle$$

best linear function of coordinates to represent the total dipole or polarizability (electric hamonicity)

# Polarized Raman spectrum of ice Ih from bands decomposition

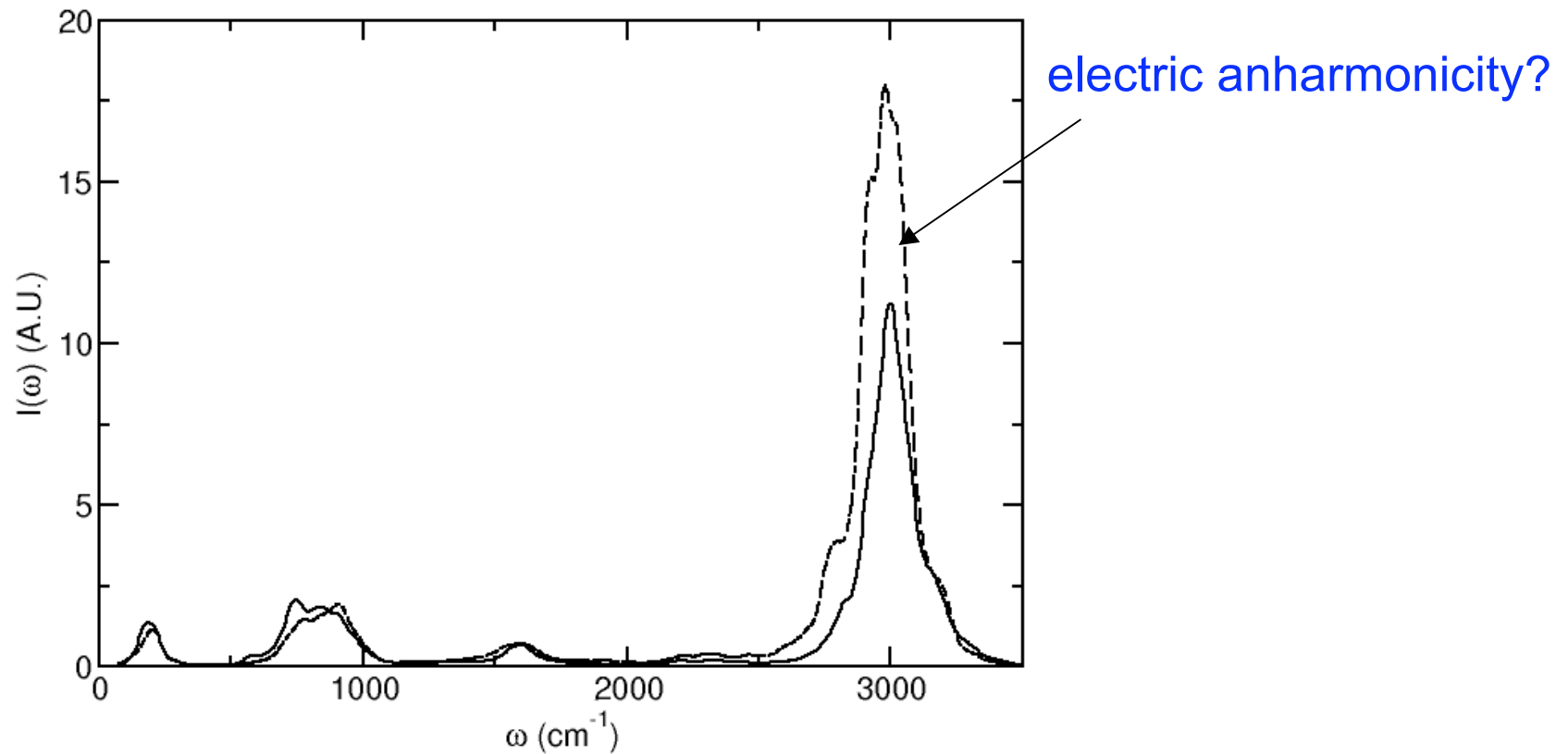


polarizability tensor calculated from linear response for 1000 configurations



# Infrared spectrum of ice Ih

- IR spectrum as a sum of bands
- - - IR spectrum directly from the autocorrelation function of the total dipole



# Conclusion

- IR spectra from autocorrelation function of the current (classical dynamics)
- use of ab initio MD to describe 1) finite temperature equilibrium, 2) bond anharmonicity and 3) electronic polarization
- applications for systems in gas phase, liquid phase, solids...
- Interpretation localization principle in frequency
- generalization of the normal mode analysis at finite temperature and for anharmonic systems
- suggests a way to study overtone and combination bands
- we now want to move towards non-linear spectroscopies which involve multiple time correlation functions

# Acknowledgements

Codruta Marinica (Post-Doc, now at LCAM-Orsay France)

Alvaro Cimas-Samaniego (Post-Doc)

Michaël Martinez (PhD, now Post-Doc in Heidelberg Germany)

Frederic Bellencontre (PhD, now Engineer)

Groups of P Maître & G Ohanessian - Orsay France

Group of JP Schermann, C Desfrançois, G Grégoire - Villetaneuse France

Group of JP Simons/L. Snoek - Oxford UK

Michiel Sprik - Cambridge UK

Riccardo Spezia & J.Y. Salpin (Evry)