Modeling two-dimensional infrared spectroscopy of hydrogen bonded systems

De Carvalho Vicente Da Cunha, Ana

SUMMARY

The hydrogen bond is a non-covalent bond that can be found in various systems, such as proteins, liquids, and even some crystals. Being non-covalent this unique interaction is extraordinarily labile, and the breakage and formation occur on the pico second timescale. One example where hydrogen bonds are particularly important is proteins, in which the different levels of organization depend on the formation of hydrogen bonds between the different aminoacids of the backbone. Hence, this interaction acts like a spring that keeps these systems together, conferring enough flexible, for structural rearrangements, and the necessary strength to keep them assembled in organized motifs. Since, these rearrangements are typically fast (picoseconds), powerful experimental techniques such as X-ray diffraction and NMR are not sensitive to the time scale of such fluctuations. Thus, to probe the dynamical properties of hydrogen bonded systems, a femto second time resolved technique, which enable to follow the changes of the system in real time, is required. Two-dimensional infrared spectroscopy (2D IR) is a correlation technique that allows to follow the evolution of a system as a function of time. Here, the molecular system is perturbed by a laser pulse to a non equilibrium state, and its relaxation to equilibrium is probed by a weak laser pulse. Because of the time delay between the excitation and the detection pulses, this technique allows to determine dynamically the spectral changes, and consequently enabling the analysis of the structural fluctuations as a function of time. Furthermore, 2D IR spectroscopy is also sensitive to environmental fluctuations carrying time-resolved information regarding the molecular couplings. A 2D IR spectrum portrays, not only, dynamical, but also, structural information, which can be seen by the peak shapes, and positions, as well as by the formation and intensity of cross peaks. The first contain information regarding changes in the environment of the molecule, whereas the cross peaks portray structural information concerning the angle between and the proximity of oscillators. Due to the crowded nature of these spectra the interpretation based on experiments alone is very challenging. Thus, computational spectroscopy methods appear as a solution. We use molecular dynamics combined with response function calculations, which enables the prediction of linear and 2D IR spectra. Hence, the time-dependent vibrational

Hamiltonian for one target mode is constructed with the atomic coordinates generated by the molecular dynamics simulations. The molecular frequencies and transitions dipoles are obtained from electrostatic maps, which relate the electrostatic environment generated by the point charges of the force field with the frequencies of the transition dipoles. The couplings between the different molecules are calculated using coupling models. The spectra are then modeled using the time-dependent Hamiltonian as an input for the response function calculations. This method enables the spectral modeling of large systems with a low computational cost. In this thesis, we have investigated the dynamics of alcohols, with different alkyl chains, either in bulk or in solution. We have also assessed the best combination of electrostatic maps used to construct the vibrational Hamiltonian, and the molecular dynamics force field. This information was later used to simulate protein like liquids, and the temperature collapse of an elastine like peptide.

Computational spectroscopy methods have proven useful to interpret experimental two-dimensional spectra, since it is possible to isolate signals or parts of the studied system, as well as to analyze microscopic properties. As a consequence these methods provide information regarding molecular structure, and details of their fluctuations. However, the results are dependent on the chosen force field or electrostatic map. Thus, in Chapter 4 we assessed possible combinations of force fields, electrostatic maps, and coupling models, for the amide I mode of proteins. The modeled spectra obtained with different models of three well-known proteins were compared with the experimental counterpart, and the quality of the calculated spectra of the considered combinations was quantified. Here, we found that the OPLS-AA force field gives the best modeled spectra when compared to the experimental ones. This force field performed best with all mappings and coupling models, independently of the electrostatic map and coupling model. Making use of this benchmark, users of these computational spectroscopy methods can obtain the optimal spectra using the best combination. Furthermore, this also calls the attention that the electrostatic map development should also take into account transferability between force fields and coupling models, in order to reduce spectral artifact that may lead to an incorrect interpretation. This benchmark was considered when modelling the spectra of a protein-like liquid (Chapter 5) and an elastine-like pepide (Chapter 6).

In Chapter 5, we modeled a protein like liquid with the purpose of mimicking the properties of intrinsically disordered protein (IDPs) domains, as they play important roles in many biological functions, such as signalling and DNA

translation/transcription. However, these protein motifs are difficult to model due to their ill defined and labile structure. Small molecules present a solution for this problem, since they portray in the liquid phase the same lability as IDPs. One example of such a building block is the N-methylacetamide (NMA) molecule, which contains a peptide bond, and in bulk it forms long hydrogen bonded chains. Thus, it can be used to mimic the properties of intrinsically disordered proteins. Here, we investigated the structure and dynamics of bulk NMA using linear and 2D IR. We have found that in the bulk liquid NMA aggregates in four different structures, in which the chain one is dominating. Furthermore, we have found that vibrational dynamics is quite pronounced in the chain like structures, and consequently, is the main contributor for the sub band on the blue side of the linear spectra, as well as to the cross peaks seen in the 2DIR spectra. The delocalization of vibrational modes is corroborated by the fast population transfer decay when compared to the orientational correlation function. This is further confirmed by the fast anisotropy decay. Thus, this study is important to disentangle spectral signatures of highly disordered systems.

The elasticity of animal tissues is given by a protein called elastin, which is present in many different organs such as muscles, skin, and the lungs. The elastic and self-assembly properties of this protein are due to the molecular structure, which contains alternating hydrophobic and hydrophilic domains. The latter plays an important role in self-assembly processes. The properties of the hydrophobic domains can be mimicked using a peptide consisting of repeats of aminoacid sequence Val-Pro-Gly-Val-Gly, named elastine-like peptides (ELPs). However, due to its short length, this 5-residue reference ELP does not show a coacervation (electrostatically-driven phase separation) transition. For the 90-repeat ELP, a coacervation transition was found, induced by increased temperatures. Concomitant with the coacervation transition, desolvation of the peptide backbone occurs. In Chapter 6, we have found that the coacervation transition of these peptides is not related to conformational changes, since the amide group of valine is shielded from the solvent. Thus, the slow dynamics are most probably due to the fact that the Val(1) residue forms an intrapeptide hydrogen bond with the Val(4) residue, which is present in the aggregated form of ELP90.

In Chapter 7, we studied the dynamical properties of bulk alcohols with a different alkyl chain. Here, we have found that on the sub-picosecond time scale the dynamics of bulk alcohols is dominated by the librational motions of the O-H stretch, whereas the slower motion arises from the hydrogen bonding exchange dynamics. Also, the larger alkyl alcohols have a slower hydrogen

bonding exchange, and consequently a slower anisotropic decay. From this study, it was still not evident to what extent the spectral diffusion and vibrational lifetimes were dominated by inter or intramolecular degrees of freedom.

In Chapter 8, we have investigated the dynamics of diluted alcohols with different alkyl chains, with the purpose of understanding whether the stretching mode dynamics depends on the intra- or inter-molecular interactions. This is achieved by removing the solvent effects and diluting the alcohol in a more inert hydrogen bond accepting solvent. The results demonstrate that the vibrational lifetimes arise from the intermolecular interactions, as changing the alkyl groups does not significantly change the observed dynamics when studied in an the same solvent.

Through this work we have found that 2D IR spectroscopy is a powerful tool to distinguish molecular properties of various systems, enabling to infer the influence of hydrogen bonding on the structural and dynamical properties. We have managed to answer crucial questions regarding methodological issues, such as the optimal simulation methodology and the interplay between delocalization of vibrations and hydrogen bond strength and exchange. One useful example is the amide I band spectral benchmark that can be used by other researchers to improve the quality of their calculated spectra or to infer about possible spectral artifacts arising from the model/force field combination. The findings regarding hydrogen bonded chains in protein like liquids enable to understand more about the assembly of small molecules in bulk, and from that to draw conclusions regarding the dynamics of more complex systems. However, while the presented research provided answers to many questions new ones emerged. Thus, the development of coupling models between different molecular vibrations, like the water bend and the amide I band, becomes crucial for the simulation of more realistic spectra. Such, would allow a more detailed interpretation of the spectral signatures seen in numerous experimental spectra. Furthermore, these new methods would enable spectral calculations of proteins in different conformations, likewise experiments. Thus, the urgent need for developing computational broad band spectra implies the construction of new coupling models, as well as an improvement of the existing ones.