Potential for the detection of molecular complexes and determination of interaction geometry by 2DIR: application to protein sciences

Rui Guo, Margherita Miele, Elizabeth M. Gardner, Frederic Fournier, Kathryn M. Kornau and David R. Klug*

Institute of Chemical Biology, Imperial College London, London SW7 2AZ, United Kingdom
Email: d.klug@imperial.ac.uk

The ability to detect molecular complexes and determine their geometries is crucial to our understanding of all biological phenomena, including protein structures and functions. We recently demonstrated that a novel 2DIR technique, EVV 2DIR spectroscopy, can be used for this purpose. In this paper, we evaluate the potential utility of the method for the analysis of protein composition, structure and function. In order to do this we apply computational tools to a group of selected biological systems, for which our calculated spectra all showed features that can in principle be detected with existing sensitivities. We also investigate the possibility of using our technique to detect and analyse hydrogen-bonded systems through a tyrosine-water model.

I. Introduction

Reversible molecular interactions mediated by electrostatics and hydrogen bonding dominate biological plasticity at the molecular level. For example, the 20,400 human genes result in a protein interactome estimated to be ~600,000 distinct pairwise protein-protein interactions. Similarly nearly all of the 3-5000 important metabolites interact with at least one protein and most drugs function by binding to recognition motifs of large biological macromolecules. Finally gene expression is largely mediated by the interaction of nucleic acids with proteins but also by nucleic acids with themselves. Above and beyond the interaction of discrete molecular entities with each other, the self-assembly of biological membranes and the assembly of proteins are themselves driven by a combination of intramolecular electrostatics and hydrogen-bonding.

Detecting, measuring the geometries of and understanding biomolecular interactions present a significant challenge to both theory and experiment. In particular there is a pressing need for new experimental tools able to detect such interactions, ideally in a label-free way with high sensitivity and high throughput whilst retaining an ability to resolve molecular details when and where required. The existing set of electronic and vibrational spectroscopies is certainly sensitive enough for these purposes and could provide the throughput required, however they are generally unable to resolve the interactions of interest due to spectral congestion. This leads to the introduction of labelling methods, and although Fluorescence Resonance Energy Transfer (FRET) spectroscopy has been an extremely successful tool in detecting the co-location of proteins within a cell, the requirement for heterolabelling of two proteins is a major disadvantage and there is no known way to generalise it to multiple proteins let alone clinical samples. By way of contrast, Nuclear Magnetic Resonance (NMR) spectroscopy can provide exquisite structural and chemical details but it requires large amounts of material, is very low-throughput and many biomolecules and biomolecular complexes are too large to be analysed by NMR methods due to the need for rotational averaging to produce narrow spectral lines.

We have been developing a particular form of coherent two-dimensional infrared spectroscopy (2DIR) with a view to use it as a biomolecular analysis tool. This particular 2DIR pulse sequence discussed in this paper is actually a triple resonance method which combines infrared absorption with Raman scattering. We have previously shown how this form of 2DIR, Electron-Vibration-Vibration (EVV) 2DIR spectroscopy, can be used to identify proteins, observe reactive intermediates in enzymes and image biological tissue. We have also recently demonstrated the ability of EVV 2DIR to both detect molecular complexes and to directly determine the geometry of interaction of the molecular groups involved. The crucial step in this study was that we showed how it is possible in a generic way, to invert 2DIR spectroscopic data to obtain the geometry of the molecular interaction because there is a simple and straightforward correspondence between spectra and geometry for the case of through-space vibrational coupling in the absence of a chemical bond. In this previous work we detected and measured the interactions between aromatic groups but there is no theoretical reason why it can not be applied to the interactions of aliphatic groups too.

In this paper we calculate the effects of through-space vibrational coupling to components of three large
biological systems which also contain interacting pairs of aromatic groups, a fairly common motif in biological macromolecules. In doing so we predict that it should be possible in principle to detect whether such interactions are present using 2DIR spectroscopy, and also in some cases to determine the geometry of the interaction.

A more challenging problem is the case of hydrogen bonding. Although hydrogen bonding is dominated by electrostatic interactions, it does also include an important contribution from exchange energy. As the contribution of the exchange interaction becomes higher, interactions between the bonding partners start to become dominated by through-bond interactions which involve substantial contributions from mechanical anharmonicities. In light of this it is an open question as to whether there is a systematic variation of such interactions are present using 2DIR spectroscopy, and also in some cases to determine the geometry of hydrogen-bonded chemical groups. In this paper we therefore explore a simple hydrogen-bonding model system, that of a tyrosine sidechain hydrogen-bonded to a single water molecule, to examine the potential of EVV 2DIR for detecting the presence of hydrogen bonds and determining the geometry of the hydrogen-bond interaction.

The purpose of the calculations presented in this paper is to gain a feeling for whether intermolecular cross-peaks are likely to be rare or common in biological systems, whether they tend to be strong or weak and whether there are many or few such features associated with inter-group interactions. This then leads to a view as to whether it will be possible in general to use 2DIR methods for: a) the detection of such intermolecular interactions and b) whether geometry determination will be possible in principle.

II. Theory and Methods

Theoretical and experimental development of EVV 2DIR spectroscopy can be followed in details in other papers.²⁻⁴ Here we gave but a brief summary of its main features most relevant to our current discussions. EVV 2DIR spectroscopy uses three pico- to femto-second laser pulses centred at frequencies \( \omega_a \), \( \omega_b \) and \( \omega_c \) to excite and detect the nonlinear vibrational response of the system, with \( \omega_a \) being infrared and \( \omega_b \) in the visible region. EVV 2DIR signals, seen as crosspeak, can only appear between coupled vibrations.

There are three coherent pathways involved in generating EVV 2DIR signals as shown by double-sided Feynman diagrams in Fig. 1, and the signal \( \omega_d \) is generated at frequency \( \omega_a - \omega_b + \omega_c \). The central frequency and delay of each of the three laser pulses can be separately controlled, leading to dé-congestion of 2DIR spectra and suppression of non-resonant background.

\[
\begin{align*}
\frac{\partial^2 \mu}{\partial Q_1 \partial Q_2} &= \frac{\partial^2 \mu}{\partial Q_1 \partial Q_2} \left[ T_{11}^{\gamma} \frac{\partial \mu}{\partial Q_1} + T_{22}^{\gamma} \frac{\partial \mu}{\partial Q_2} + T_{12}^{\gamma} \frac{\partial \mu}{\partial Q_1} \right]. 
\end{align*}
\]

by application of a dipole-dipole interaction model. Just like electrical anharmonicities and mechanic anharmonicities intrinsic to a molecule, this induced electrical anharmonicity can also lead to nonlinear EVV
2DIR signals. In many cases, contributions from these induced electrical anharmonicities become the dominating or even the only contributors to the signals. Combined with measurements using different permutations of laser beam polarisations, these signals have the potential to elucidate the geometrical arrangements of the interacting pair of chemical groups, as have been demonstrated in the benzonitrile-phenylacetylene system previously\(^5\). For the polarisation cases PPP and PPS, the nonlinear susceptibilities from the interaction induced electrical anharmonicities can be expressed as:

\[
\chi_{\text{PPP}}^{(3)} \propto \text{Tr} \left( \frac{\partial \alpha}{\partial Q_1} \left( \frac{\partial^2 \mu}{\partial Q_2 \partial Q_1} - \frac{\partial \mu}{\partial Q_1} \right) \frac{\partial \mu}{\partial Q_2} + \frac{1}{2} \frac{\partial \alpha}{\partial Q_1} \left( \frac{\partial^2 \mu}{\partial Q_1 \partial Q_2} - \frac{\partial \mu}{\partial Q_1} \right) \frac{\partial \mu}{\partial Q_2} \right) \right. \\
\chi_{\text{PPS}}^{(3)} \propto 2 \text{Tr} \left( \frac{\partial \alpha}{\partial Q_1} \left( \frac{\partial^2 \mu}{\partial Q_2 \partial Q_1} - \frac{\partial \mu}{\partial Q_1} \right) \frac{\partial \mu}{\partial Q_2} + \frac{1}{2} \frac{\partial \alpha}{\partial Q_1} \left( \frac{\partial^2 \mu}{\partial Q_1 \partial Q_2} - \frac{\partial \mu}{\partial Q_1} \right) \frac{\partial \mu}{\partial Q_2} \right) 
\]

The above two equations also establish the foundation for determination of interaction geometries by EVV 2DIR spectroscopy, i.e. the dependency of interaction-induced electrical anharmonicity on interaction tensor \(\chi^{(3)}\), and the dot-product dependency of \(\chi^{(3)}\) on these molecular properties. This can be translated into direct relationships between experimentally measurable signal intensities and the distance and relative angles between two interacting chemical groups. For example, in the special case of two interacting linear vibrations, the signal ratio of an interaction-induced crosspeak under PPP and PPS setups can be written as:

\[
\left| \frac{\chi_{\text{PPP}}^{(3)}}{\chi_{\text{PPS}}^{(3)}} \right| = \left| \frac{2 + k \cos \phi - \cos^2 \phi}{1 + 3k \cos \phi + 2 \cos^2 \phi} \right|^2 
\]

where \(\phi\) is defined as the vector angle formed between the two bond vectors, and \(k\) is a constant involving their individual vibrational spectral properties. Combined with the fact that interaction-inducing 2DIR signals are inversely proportional to their distances to the power of six, both angles and the distance between the two interacting partners can be readily determined through direct experimental measurements.

The basic procedure for calculating conventional EVV 2DIR spectra has been described in details elsewhere\(^6\)\(^\text{-}^9\). For interaction EVV 2DIR spectra, interaction induced electric anharmonicities were calculated according to Equ. 1 and then induced EVV 2DIR signals were calculated similarly to the procedure for the normal intra-molecular coupling cases according to Equ. 2. The spectral properties for each vibrational mode of all the molecules reported here were calculated at B3LYP/6-311++G(d,p) level with Gaussian 03\(^1\). Using these anharmonic parameters EVV 2DIR spectra over any interested spectral area involving all vibrational modes of the molecules can be calculated. Only one coherent pathway of EVV 2DIR spectroscopy, pathway I as in Fig. 1, was included in our calculations since this is the pathway used most extensively in experiments. All vibrational frequencies mentioned here were harmonic frequencies scaled by a constant factor (0.96) and in the cases of combination bands, sum frequencies of their component modes were used. This does not in any case affect the results reported here since with only one pathway included in our calculations, anharmonic shiftings of combination bands become trivial\(^7\). Moreover, for our automated calculations with large numbers of vibrational modes involved, Gaussian lineshapes with uniform widths of 10 cm\(^{-1}\) were used as typical of the linewidths seen in these systems. This approximate approach is useful in guiding experimentation to the best regions of the spectrum to look for possible candidates for complex detection and geometry determination. Note that in the following discussion locations of crosspeaks are denoted in the form of \(\omega_x / \omega_y (\omega_\text{KAMAN})\). \(\omega_x\) and \(\omega_y\) are the central frequencies of the two infrared pulses and \(\omega_\text{KAMAN}\) is the other component of the combination band at \(\omega_x\). Also in the calculated spectra reported below, sometimes stronger features have to be truncated to allow weaker features to be seen, thus will look more like rings instead of peaks.

III. Results and discussions

A. p53-MDM2

p53 is a major tumour suppressor protein that regulates genes inducing either cell cycle arrest or apoptosis. It is also a key regulator of cell function in general and is known to interact with over 200 other proteins. Several X-ray crystal structure studies showed that MDM2, a negative regulator of p53, binds to a specific 15-residue \(\alpha\) helix of p53 which are known to be involved in p53 transactivation, indicating that MDM2 inhibits the tumour suppressing functions of p53 by concealing its transactivation domain. The ability to monitor this binding between p53 and MDM2 could potentially be a useful tool particularly as many drugs under development target this same interaction. The interactions between the two parts are largely hydrophobic, in the form of nonspecific van der Waals contacts. In particular, on MDM2 there is a deep
hydrophobic cleft to which the amphipathic α helix of p53 binds. Two X-ray crystal structures were selected for this study, *Xenopus Laevis* MDM2 bound to human p53 (PDB: 1YCQ) and human MDM2 (HDM2) bound to human p53 (PDB: 1T4F). In both structures, there is a close contact between the sidechain of a phenylalanine on p53 (F19 in both structures), and the sidechain of a tyrosine on MDM2/HDM2 (Y63 in 1YCQ and Y67 in 1T4F) as shown in Fig. 2 (a). In 1T4F the distance between the two aromatic rings is about 5.4 Å while in 1YCQ the distance is slightly closer at ~ 5.0 Å. Calculations were carried out for the couplings between the PHE and TYR sidechains and on both of the two structures. Using p-methylphenol and toluene as sidechain models for tyrosine and phenylalanine, the interaction-induced electrical anharmonicities were calculated as well as the corresponding EVV 2DIR spectra. Fig. 2 (b) and (c) showed the calculated spectra for the p53-MDM2 as described in 1YCQ, overlapped with calculated EVV 2DIR spectra of the two sidechain monomers. The results for 1T4F are very similar. As shown in the spectra, there are but a few crosspeaks from interactions between PHE and TYR sidechains which stand out around one spectral region at ~ 1140/4180 cm\(^{-1}\). There are a few more at 1225/4265 cm\(^{-1}\), but these are shown buried under monomer crosspeaks which are more than one order of magnitude stronger, making their detection challenging. The vibrations involved in the interaction-induced crosspeak at 1139/4178(3039) cm\(^{-1}\) are the strongly IR-active COH bending mode of TYR at 1139 cm\(^{-1}\) coupled with one strong Raman band of the C-H stretching modes of PHE at 3039 cm\(^{-1}\). A slightly weaker interaction crosspeak can be seen at 1139/4198(3059) cm\(^{-1}\), which is the coupling between the same COH bending mode and another C-H stretching mode of PHE at 3059 cm\(^{-1}\).

Interestingly, TYR itself also has a C-H stretching vibration at 3039 cm\(^{-1}\), almost identical to the mode of PHE at 3039 cm\(^{-1}\). However, calculated intensity of the nonlinear signal from the intra-ring couplings between this C-H stretching mode of TYR and its own COH bending mode was found to be more than 200 times weaker than the signal intensity from TYR-PHE through-space coupling. Further analysis showed that this is because in the intra-TYR case, contributions from mechanical anharmonicity and electrical anharmonicity to the nonlinear signal are of similar values but of opposite sign, thus largely cancelling out each other, while the signal in the PHE-TYR coupling case is necessarily solely from electrical anharmonicity, thus turns out to be much stronger. This is an interesting case in which the crosspeak from an intermolecular vibrational coupling turns out to be much stronger than a similar crosspeak from an intramolecular vibrational coupling, even the modes involved are almost the same. Nevertheless, the interaction-induced crosspeaks at 1139/4198(3059) cm\(^{-1}\) and 1139/4178(3039) cm\(^{-1}\) are still 6 and 10 times weaker compared to another close-lying strong intra-TYR crosspeak at 1139/4158(3019) cm\(^{-1}\), making it a challenging task for the experimentalists to observe them.

Upon changes of laser polarization setups from the PPP to PPS case, the PHE-TYR crosspeak at 1139/4198(3059) cm\(^{-1}\) becomes only 0.7 times weaker, while the crosspeak at 1139/4178(3039) cm\(^{-1}\) reduces to one sixteenth of its intensity under PPP setup. This makes the former crosspeak the relatively stronger one. However, in PPS setup another TYR crosspeak from intra-TYR coupling at 1139/4202(3063) cm\(^{-1}\) becomes 8 times stronger and thus overlapping with the crosspeak at 1139/4198(3059) cm\(^{-1}\), making it difficult to detect.

**B. SPE7-DNP system**

Antibody SPE7 is a monoclonal immunoglobulin E which was recently reported to show multispecificity by adopting different binding-site conformations when binding to several unrelated antigens\(^\text{14}\). This conformational diversity of one antibody was used to explain why a limited repertoire of antibodies can bind and protect us against “an almost infinite diversity of invading antigens”\(^\text{14}\). It has been shown that there are at least four different binding-site conformations that SPE7 can take up, one of which, the Ab\(^3\) conformation (PDB: 1OAU), is what SPE7 takes when it binds to 2,4-dinitrophenyl-serine (DNP-Ser). In this conformation, the sidechain of TRP93 stacked very close to the dinitrophenyl ring. The distance between the centres of the two aromatic ring centres is only about 3.5 Å as shown in Fig. 3 (a).
To apply our theory of coupled vibrations to this system, 1,3-dinitrobenzene (DNB) was used as the model for DNP-Ser, and 4-methylindole was used as the sidechain model for tryptophan. Fig. 3 (b) was the calculated 2DIR spectra showing crosspeaks induced through interactions between TRP and DNB. Compared with the spectra shown in Fig. 2 (b), in this system there are much more interaction-induced crosspeaks. As an example, there are a series of strong interaction induced crosspeaks along $\omega \approx 1530 \text{ cm}^{-1}$ and some along $\omega \approx 1380 \text{ cm}^{-1}$ spectral region, providing experimentalists many candidates to choose from. Since the interaction-induced signals are very sensitive to the distance between this interacting pair, the observation and measurement of multiple interaction-induced crosspeaks will provide mutually complementary information about the binding between the antibody and the hapten. One of the strongest is the interaction crosspeak at 1529/2915(1387) cm$^{-1}$, which involves the coupling between the DNB vibrational mode at 1529 cm$^{-1}$ and a TRP ring mode at 1387 cm$^{-1}$. The DNB mode is largely the antisymmetric NO$_2$ stretching vibration but also mixed with some phenyl ring vibrations. Since the intensity is strong and there is no any crosspeaks around, this crosspeak could be a good candidate for experimental exploration. Upon change of laser polarization from PPP to PPS, its intensity is reduced by more than 50 times. Just as the case in p53-MDM2, TRP itself can also form a crosspeak at almost the same position (at 1530/2916(1387) cm$^{-1}$). However this crosspeak is only strong enough to show up in the PPS setup, while in PPP configuration it is almost 2000 times weaker. Therefore although a signal can be detected at this region under both PPP and PPS polarization setups, the nature of the crosspeak changes from a DNB-TRP-interaction crosspeak in PPP setup to an intra-TRP crosspeak in the PPS case. The crosspeak at 1529/2720(1192) cm$^{-1}$ is another promising candidate which may be used as an indicator of the interaction between DNB and TRP. The vibrational modes involved in this crosspeak are the same DNB mode for $\omega_5$, but the TRP mode is a different ring mode involving more CN stretching motion. Upon change of laser polarization setup from PPP to PPS, the signal becomes only 3 times weaker but with no any other crosspeaks around, detection of this crosspeak in both setups should be easier.
C. Myristoyltransferase and inhibitor

*Candida albicans* N-Myristoyltransferase (Nmt) is a monomeric enzyme that catalyzes the transfer of the fatty acid myristate from myristol-CoA to the N-terminal glycine residue of a variety of eukaryotic and viral proteins. Researches have established Nmt as an attractive target for antifungal drugs. Two types of inhibitors have been discovered, one of them a series of small molecules with a common benzofuran core. These inhibitors bind to a groove generated by some structural rearrangement of the enzyme. When this happens there will be a close contact between the benzofuran ring of the inhibitor and a TYR sidechain of Nmt. The benzofuran core is located in the centre of a deep pocket, surrounded by hydrophobic residues, largely composed of aromatic amino acids. In particular the benzofuran ring is stacked parallel to Y225 (PDB: 1IYL). In the current research, we focused on this stacked interaction between Y225 and the benzofuran core. A simple model was built mimicking the inhibitor core as shown in Fig. 4 (a), and interaction-induced EVV 2DIR spectra were calculated between the model molecule and TYR sidechain, shown in Fig. 4 (b). Particularly interesting among the many interaction-induced crosspeaks, are those lying along the $\omega_1 \sim 1140 \text{ cm}^{-1}$ spectral region, as shown in more detail in Fig. 4 (c). As an example, the crosspeak at $\sim 1139/4235 \text{ cm}^{-1}$ actually consists of two crosspeaks: one at $1139/4236(3097) \text{ cm}^{-1}$, from a coupling between the TYR COH bending mode and one of the C-H stretching modes on benzofuran ring; and another at $1139/4232(3093) \text{ cm}^{-1}$ from a coupling between the same TYR COH bending mode and another C-H stretching mode on benzofuran ring. The signal ratio between the two crosspeaks is about 7.8:1 in PPP polarization setup. However when the laser polarization setup changes from PPP to PPS case, the signal ratio between the two crosspeaks becomes about 226:1. This is due to the different nature of the two benzofuran vibrational modes involved. Thus upon laser polarization change, the first crosspeak only becomes slightly weaker, while the other one becomes almost 50 times weaker. So the contribution in the crosspeak at $\sim 1139/4235 \text{ cm}^{-1}$ in the PPS setup becomes almost purely from the crosspeak at $1139/4236(3097) \text{ cm}^{-1}$. Another potential candidate for experimentalist is the crosspeak at $1139/4203(3064) \text{ cm}^{-1}$ which comes from the coupling between TYR COH bending and yet another benzofuran C-H stretching mode. Upon laser polarization change, this crosspeak becomes almost 60 times weaker, thus disappeared in the calculated spectra in Fig. 4 (c).
tyrosine is the modification of the original vibrational coupling scenario within each component molecule. It has been found that upon formation of a hydrogen bond between R-A-H and B, A-H stretching frequency shifts lower, while R-A-H bending frequency shift higher, which are usually also accompanied by significant band broadening.

Conventional IR and Raman spectroscopy have been shown to be extremely useful in detecting the existence of hydrogen bonding. Hydrogen bonding is known to play crucial structural and functional roles in biological systems and is possibly the most important intermolecular interaction in these systems.

To explore the possibility of using EVV 2DIR spectroscopy to detect the formation of hydrogen bonds and to investigate the geometrical arrangements of hydrogen bonds, a simple system composed of a tyrosine sidechain and a water molecule was used, as shown in Fig. 5 (a). This system was first optimized to its equilibrium minimum and then anharmonic parameters were calculated at this geometry. When compared to the calculated 2DIR spectra of tyrosine and water monomers, substantial changes in the spectra become obvious as shown in Fig. 5 (b). One effect of the formation of a hydrogen bond between water molecule and tyrosine is the modification of the original vibrational coupling scenario within each component molecule. For example as shown in more details in Fig. 5 (c), the crosspeak at 1591(1591) cm\(^{-1}\) in tyrosine-water system is almost identical to the tyrosine crosspeak at 1590(1590) cm\(^{-1}\), both involving the self-coupling of the same phenyl ring mode, yet in the complex the nonlinear signal is almost 8 times stronger. This is particularly interesting as the vibrational frequency of the mode involved actually changes so little. Further analysis shows that this is at least partly due to the increase of electrical anharmonicity upon the formation of hydrogen bonding between water and tyrosine. The prominent water crosspeak at 1539(1539) cm\(^{-1}\) due to its bending vibrations shifted to higher frequencies at 1558(1558) cm\(^{-1}\) and the signal intensity became 3.5 times weaker. More dramatic for this crosspeak is the intensity change when laser polarisation changes: with water alone, a polarisation change from PPP to PPS leads to a weaker signal by a factor of 14, yet when hydrogen bond was formed, the same polarisation change leads to the PPS signal weaker by a factor of 211 than the PPP signal. In both cases, contributions to the nonlinear signals came from electrical anharmonicities. Yet in the complex, the contribution becomes smaller. The increase in sensitivity largely comes from changes of transition polarizability of the water bending vibration, which has one of its three principal tensor components reduced upon hydrogen-bond formation.

Another effect of the hydrogen bond, which might be expected, is the appearance of new crosspeaks in EVV 2DIR spectra. This can be either a) between two vibrational modes, one localized on the H-donor, another localized on the H-acceptor, or b) between one vibrational mode localized on one of the two molecules and one of the newly created vibrational modes of the complex. However in tyrosine-water system, there is no new crosspeaks of both types in the spectral region shown in Fig. 5 (b).

On the other hand, it would be interesting to understand how nonlinear signal changes with respect to small changes in geometrical relationship between the proton donor and proton acceptor. To do so, we kept the geometries of the monomers frozen, and made manual geometry changes of the complex along the water bending direction, i.e. changing the \(\theta\) as shown in Fig. 6 (a), with \(\theta = 0\) defined for the equilibrium structure. For each \(\theta\), numerical differencing was used to calculate anharmonic parameters at such a slightly changed geometry, followed by calculations of its corresponding EVV 2DIR spectra. The intensity changes of two selected crosspeaks, one at \(\sim 1315/3045(3045)\) cm\(^{-1}\) and another at \(\sim 1560/2730(1170)\) cm\(^{-1}\) was drawn with respect to the change of \(\theta\) as shown in Fig. 6 (a). Both crosspeaks come entirely from couplings between vibrational modes of tyrosine, yet their intensities nevertheless showed changes along with the change of the bending angle \(\theta\). In particular, the crosspeak at \(\sim 1560/2730(1170)\) cm\(^{-1}\) experience changes of more than one order of magnitude, yet the other one at \(\sim 1315/3045\) cm\(^{-1}\) changes comparably less. This difference in behaviour may be related to the fact that the former is largely contributed by electrical anharmonicity while the latter has almost equal contributions from both mechanical and electrical anharmonicities.
Naturally, although each geometry calculated here was treated as if it were static, in room temperature and surrounded by solvents geometry of the tyrosine-water system will be statistically distributed within the configuration space spanned by the bending angle $\theta$ and all the other molecular coordinates. Shown in Fig. 6 (b) is the calculated frequency of the vibrational modes involved in the two crosspeaks with respect to the change of bending angle $\theta$. These frequency shifts, combined with the intensity changes discussed above could therefore show up as complicated lineshapes depending on just how much of the configuration space is sampled and the timescale of that sampling. As has been demonstrated for the liquid water system$^{16}$, such lineshapes could carry rich information about the dynamical behaviour and environmental effect of the hydrogen-bonded system. There are however thought to be many cases for proteins where hydrogen bonds form well-defined geometries with narrow configurational distributions. It is therefore instructive to explore the sensitivity of EVV 2DIR signals to small geometrical changes in a hydrogen bond, as shown in Fig. 6 (a) to determine whether in such cases structural information might be recovered from EVV 2DIR measurements.
Fig. 5 a) Optimized geometry of tyrosine sidechain + water, also shown and b) calculated EVV 2DIR spectra of the complex (blue) compared to monomeric spectra of tyrosine (red) and water (black); c) the same spectra in detail as discussed in the text.
Fig. 6 a) Change of calculated signal intensities for two crosspeaks of tyrosine-water complex with respect to the change of bending angle $\theta$; b) Frequency changes with respect to the change of bending angle $\theta$.

As can be seen from Fig. 5 b), there are numerous EVV 2DIR crosspeaks that occur exclusively due to hydrogen bonding of the tyrosine. This suggests that it will be possible to determine whether amino acids are hydrogen bonded or not and therefore to detect the formation of a molecular complex. As a bonus it seems that variations in the hydrogen-bonding angle $\theta$ lead to well-defined, and in many cases monotonic, changes to both frequency and intensity of EVV 2DIR crosspeaks. Although we have as yet no method for inverting such data to obtain a structure, the existence of many such features suggests that it should be possible to extract angular information from such data.

IV. Summary

In this paper we predict the result of intermolecular couplings between chemical groups in three non-bonded systems, the p53-MDM2, SPE7-DNB and Nmt-Inhibitor systems. Our calculated EVV 2DIR spectra predict that for each system there are detectable features which come from the electrostatic interactions between two molecular groups, and we highlight these as potential candidates for further experimental exploration. In all cases, spectra were calculated for two different laser polarizations as this allows geometric information to be recovered above and beyond the detection of the complexation. As a first step to explore the potential of EVV 2DIR spectroscopy for the study of hydrogen-bonded systems, we calculated EVV 2DIR spectra of a simple hydrogen-bonded system, the tyrosine-water complex, to demonstrate the sensitivity of its nonlinear signals to both the presence of the hydrogen-bond and to small changes in the hydrogen bonding geometry.

Acknowledgements

This research was supported by the Engineering and Physical Sciences Research Council (EPSRC) through the Single Cell Proteomics Project, the Biotechnology and Biological Science Research Council (BBSRC) through the Chemical Biology Centre Doctoral Training Centre.
References


