

Supplementary Material to accompany the paper “Solvation of a Flexible Biomolecule in the Gas Phase: The Ultraviolet and Infrared Spectroscopy of Melatonin-Water Clusters” by Gina M. Florio and Timothy S. Zwier

Further Structural Analysis based on a comparison of the DFT calculations with experiment

This supplementary material contains further details of the conformational assignments of MEL-W₁ and MEL-W₂ using the harmonic vibrational frequencies and infrared intensities calculated at the DFT Becke3LYP/6-31+G*(5d) level of theory as a point of comparison with experiment. The calculated harmonic infrared frequencies have been scaled by 0.960, the factor required to align the indole NH to its experimental value. While this scale factor is appropriate for the indole NH stretch, it is slightly too large for the OH stretches.

1. MEL-W₁ A and B

We have assigned MEL-W₁ A to water bound to the carbonyl of the MEL A Gpy(*trans*-in)/anti conformer (structure III), and MEL-W₁ B to water bound to the carbonyl of the MEL B Anti(*trans*-out)/anti conformer (structure IV), based upon the nearly identical frequencies of the amide NH stretches in the water complex to that in the monomer. However, there are some aspects of the assignment of MEL-W₁ A to water bound to the carbonyl group of MEL A (structure III) that deserve comment. First, the assignment of the MEL monomer A as Gpy(*trans*-in)/anti was tentative due to the similarity in the IR spectra of Gpy and Gph conformers of this type, and was based upon the lower energy of the Gpy structure as determined by both the DFT and LMP2

calculations (Fig. 1). Second, the water cluster associated with the Gpy(*trans-in*)/anti conformer (structure III) is not predicted to be the lowest energy water cluster, although it is the lowest energy gauche oriented, single-donor, C=O bound cluster (Fig. 3). Third, the calculated infrared spectra (Fig. A) of the single-donor, C=O bound W_1 clusters, Gpy(*trans-out*)/anti (structure III), Gpy(*trans-in*)/anti (structure V), and Gph(*trans-out*)/anti (structure VII) are all very similar. However, based on the calculated energetics, the experimental infrared spectra, and the initial monomer population distribution, we concluded that MEL- W_1 A is the Gpy(*trans-out*)/anti- W_1 species.

It is also worth pointing out that while the calculations predict the Anti(*trans-in*)/anti- W_1 cluster (structure I) to be lower in energy than the Anti(*trans-out*)/anti- W_1 cluster (structure IV), we have chosen to assign MEL- W_1 B to the latter structure. This assignment is based on comparison of the RIDIR spectrum of MEL- W_1 B with the calculated infrared spectra for structures I and IV (Fig. A). The RIDIR spectrum of MEL- W_1 B shows the H-bonded OH stretch at 3490 cm^{-1} , which is between 30 and 50 cm^{-1} higher in frequency than the analogous OH stretches of the other MEL- W_1 clusters (Fig.6). The calculated infrared spectrum of the Anti(*trans-out*)/anti- W_1 cluster (structure IV) has the H-bonded OH stretch 33 cm^{-1} higher in frequency than the H-bonded OH stretch of Anti(*trans-in*)/anti- W_1 (structure I). The calculations also show that the Anti(*trans-out*)/anti- W_1 cluster has the highest frequency H-bonded OH stretch of all the single-donor C=O bound W_1 clusters (Fig. A), in keeping with the experimental assignment of MEL- W_1 B to structure IV.

2. MEL- $W_{1,2}$ Bridges: MEL- W_1 Y/Z and MEL- W_2 Y/Z

Based on the electronic and infrared spectroscopy, we have determined that there are two families of MEL-water bridges observed in the expansion: a) W_1 Y and W_2 Y and b) W_1 Z and W_2 Z. We have argued that these water bridges span the C=O and indole NH of MEL, based on the absence of a free indole NH at 3525 cm^{-1} and the presence of a strong H-bonded OH stretch in the W_1 Y and Z clusters where the $\text{OH}\cdots\text{O}=\text{C}$ stretch occurs. This section contains a discussion of the types of bridges that are minima on the MEL- W_1 potential energy surface, leading to an identification of possible structures for MEL- W_1 Y and Z. We will use the calculated infrared spectra of the MEL- W_1 bridges as our guide in this process (Fig. B). Of the twelve MEL- W_1 conformers determined to be minima on the DFT potential energy surface, five are W_1 bridges: structures II, VI, VII, IX, and XI.

These five structures include several that we anticipate to be inconsistent with experiment. Nevertheless, it is useful to view their predictions for the IR spectrum and energetics to compare with those of clusters that are more likely candidates for the MEL- W_1 Y and Z species. Based on the calculated frequencies and infrared intensities for structures XI, VI, and VII, we can rule these species out. Structure XI has the water molecule bridging across a *cis*-amide group, resulting in a set of strongly red-shifted H-bonded stretches below 3400 cm^{-1} and a free indole NH stretch. Structures VI and VII have the water accepting a H-bond from the amide NH and donating a H-bond to the π cloud (structure VI) or the indole N (structure VIII), resulting in a weakly H-bonded OH stretch around 3550 cm^{-1} .

The remaining two bridge structures (II and IX) are gauche pyrrole MEL structures with water bridging between the C=O and indole NH group. Structure II has a

trans-amide MEL conformation that is not observed in the expansion prior to introduction of water. Structure IX is essentially identical to structure II but has a *cis*-amide MEL conformation. This Gpy(*cis*-out)/anti conformer was tentatively assigned to the MEL E monomer observed in the expansion in the absence of water, with a population only about 1% of MEL A.

The calculated infrared spectra of structures II and IX are shown in Fig. B. Both spectra have aspects that correlate well with experiment. For instance, as expected, the formation of the H-bond to the C=O produces a red-shifted OH stretch with increased intensity, occurring at 3420 (II) and 3406 (IX) cm^{-1} . In these clusters the indole NH is a H-bond donor, and its frequency is also red-shifted from its free position at 3525 to 3468 (II) and 3454 (IX) cm^{-1} . The amide NH stretches are free and occur at 3468 in the *trans*-amide (II) and 3416 cm^{-1} in the *cis*-amide, very near their values in the absence of water binding. The side-chain in structure II is a ‘non-standard’ orientation of the monomer, and the amide NH stretch is shifted down in frequency by 6-17 cm^{-1} from the calculated *trans*-amide NH stretch of the MEL A-C conformers. It is not clear if this small red-shift is due to the unusual side-chain orientation or the formation of the H-bond at the C=O. The amide NH stretch of the optimized Gpy(*cis*-out)/anti monomer is at 3432 cm^{-1} , indicating that formation of the W_1 bridge across the C=O and indole NH of this monomer (structure IX) red-shifts the amide NH stretch by 16 cm^{-1} .

On the other hand, other aspects of the calculated spectra do not appear to properly reproduce the nature of the H-bond with the indole NH, which is a strained interaction. First, in both calculated spectra, the indole NH stretch is shifted to lower frequency due to H-bond formation, but the magnitude of the calculated shift is too small

to match up with any of the side bands around the $\text{C}=\text{O}\cdots\text{HO}$ transitions in the RIDIR spectra. This would indicate that the H-bond formed between the indole NH and water is calculated to be too weak by comparison to experiment. Second, according to the calculations, the intensity of the indole NH stretch fundamental should be increased significantly by H-bond formation. We would anticipate being able to observe this increase in intensity in the RIDIR spectra, yet the experimental spectra show only weak transitions flanking the $\text{OH}\cdots\text{O}=\text{C}$ band. If we were to assign one of the small transitions around the $\text{OH}\cdots\text{O}=\text{C}$ band in the RIDIRS of MEL- W_1 Y and Z to the indole NH, the calculated intensity is much too large, indicating the H-bond to the indole NH is too strong in the calculated structures. This is the opposite inference we drew from the frequency shift discussed above. It is possible that dispersion interactions, which are not accounted for properly in the DFT calculations, could be important due to the interaction of the water molecule with the π cloud in these bridge structures.

These puzzling aspects of the RIDIR spectra of MEL- W_1 Y and Z preclude firm assignments to specific structures based on the current set of MEL- W_1 minima. From the RIDIR spectra, we know that the water is behaving as a single-donor to the $\text{C}=\text{O}$ and that the indole NH stretch fundamental is perturbed from its 'free' position. The experimental frequencies of bands that are potential candidates for the amide NH stretch fundamental would seem to be too low to be consistent with a *trans*-amide MEL configuration. As a result, it seems more likely that MEL- W_1 Z and Y have melatonin structures that are *cis*-amides, based on the presence of the weak transitions to the red of the strong $\text{OH}\cdots\text{O}=\text{C}$ band, in the region where the free *cis*-amide NH stretches appear in *cis*-amide MEL monomers D and E. Many of the questions arising from the infrared spectra of MEL- W_1

Y and Z could be addressed via deuteration studies. Selective deuteration of the amide NH and indole NH sites would allow for determination of the position of these bands in the RIDIR spectra of MEL-W₁ Y and Z.

Finally, the only two MEL-W₂ clusters observed in the expansion are MEL-W₂ Y and Z that are extensions of the MEL-W₁ Y and Z bridges, respectively. The infrared spectra of these W₂ bridges is consistent with bridges that join the carbonyl oxygen with the indole NH group (C=O $\cdots$ HO $\cdots$ HO $\cdots$ HN indole), where each water molecule acts both as single donor and single acceptor of hydrogen bonds. Observation of only water bridges for the MEL-W₂ clusters is consistent with previous studies that have shown the preference for formation of water bridges between donor and acceptor sites on the same molecule.¹⁻⁴ It is striking that these are the only W₂ species observed in the expansion, despite the fact that the W₁ species that they share a likeness with (MEL-W₁ Y and Z) have such small population relative to the major MEL-W₁ A species.

References

- (1) Carney, J. R.; Dian, B. C.; Florio, G. M.; Zwier, T. S. *J. Am. Chem. Soc.* **2001**, *123*, 5596.
- (2) Bach, A.; Coussan, S.; Muller, A.; Leutwyler, S. *J. Chem. Phys.* **2000**, *113*, 9032.
- (3) Bach, A.; Coussan, S.; Muller, A.; Leutwyler, S. *J. Chem. Phys.* **2000**, *112*, 1192.
- (4) Robertson, E. G. *Chem. Phys. Lett.* **2000**, *325*, 299.

Table A: Summary of the UV-UV hole-burning results for MEL-W₁.

Species	Freq. (cm ⁻¹)	Relative Freq. (cm ⁻¹) ^a	Normalized Intensity ^b
MEL-W ₁ A	32442	0	100
	32461	19	7
	32478	36	3
	32484	42	16
	32487	45	9
	32497	55	4
	32501	59	8
	32505	63	2
	32514	72	2
	32517	75	2
	32520	78	4
	32529	87	3
	32802	360	1
	32831	389	2
	32892	450	2
	32920	478	4
	32942	500	12
	32963	521	2
	32984	542	2
	33142	700	4
	33159	717	12
	33179	737	2
	33239	797	2
	33304	862	2
MEL-W ₁ B	32956	0	8
MEL-W ₁ Y	32842	0	15
MEL-W ₁ Z	32673	0	9
	32725	52	10
	32753	80	2
	32776	103	3
	32777	104	2
	32778	105	2
	32924	251	3

^a Frequency shift relative to lowest energy transition of each species.^b Intensity normalized to MEL-W₁ A origin transition intensity, arbitrarily set to 100

Table B: Summary of the UV-UV hole-burning results for MEL-W₂.

Species	Freq. (cm ⁻¹)	Relative Freq. (cm ⁻¹) ^a	Normalized Intensity ^b
MEL-W ₂ Y	32814	0	90
MEL-W ₂ Z	32629	0	73
	32632	3	9
	32643	14	17
	32652	23	30
	32675	46	20
	32681	52	100
	32694	65	36
	32696	67	25
	32704	75	19
	32725	96	17
	32729	100	15
	32733	104	31
	32739	110	12
	32745	116	18
	32746	117	15
	32748	119	16

^a Frequency shift relative to lowest energy transition of each species.

^b Intensity normalized to most intense MEL-W₂ Z transition, arbitrarily set to 100.

Table C: Summary of the RIDIRS data for MEL-W₁ and MEL-W₂ clusters.

Species	Freq. (cm ⁻¹)	FWHM (cm ⁻¹)	Description
MEL-W ₁ A	3717	2	Free OH
	3525	2	Indole NH
	3749	2	Amide NH
	3441	8	C=O ^{··} HO
MEL-W ₁ B	3716	2	Free OH
	3527	2	Indole NH
	3494	^a	Amide NH
	3490	6 ^a	C=O ^{··} HO
MEL-W ₁ Y	3718	2	Free OH
	3457	6	C=O ^{··} HO ^{··} HN (Ind)
	3437	1	
	3415	2	
MEL-W ₁ Z	3718	2	Free OH
	3498	2	
	3464 / 3460	7 ^a	C=O ^{··} HO ^{··} HN (Ind)
	3444	2	
MEL-W ₂ Y	3724 / 3721	5 ^a	Free OH's
	3422	5 ^a	
	3408	18	OH/NH Bridge Fundamental
	3322	44	OH/NH Bridge Fundamental
	3226	13 ^a	OH/NH Bridge Fundamental
	3213	10 ^a	2ν _{water bend}
MEL-W ₂ Z	3725	2	Free OH
	3717	3	Free OH
	3462	2	
	3403	24	OH/NH Bridge Fundamental
	3337	39	OH/NH Bridge Fundamental
	3228	12 ^a	OH/NH Bridge Fundamental
	3212	10 ^a	2ν _{water bend}

^a Partially overlapped transition.

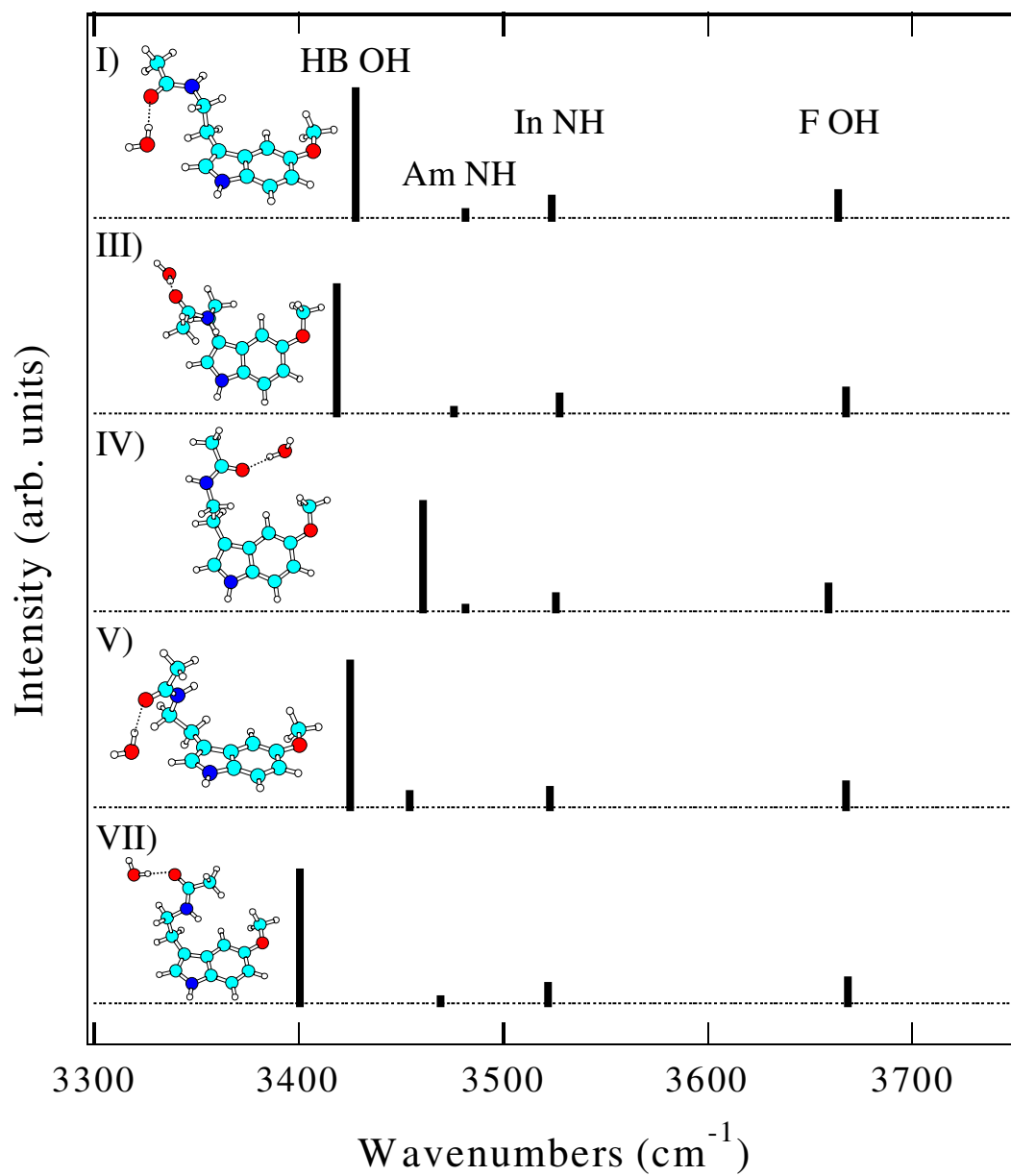
Fig. A: Calculated harmonic infrared spectra of carbonyl-bound MEL-W₁

Fig. B: Calculated harmonic infrared spectra of water bridges

