

## **Supporting Information**

### **Calculation of Vibrational Shifts of Nitrile Probes in the Active Site of Ketosteroid Isomerase upon Ligand Binding**

Joshua P. Layfield and Sharon Hammes-Schiffer\*

*Department of Chemistry, 600 South Mathews Avenue,  
University of Illinois at Urbana-Champaign, Urbana, IL 61801; shs3@illinois.edu*

## PM3 Reparameterization Procedure

We derived a set of PM3 parameters by minimizing the difference in the frequencies calculated with DFT and PM3 for a series of configurations obtained from a molecular dynamics (MD) simulation of methyl thiocyanate (MeSCN) in water. After equilibration, we propagated a 500 ps classical MD trajectory for MeSCN solvated in a box of 2900 TIP3P water molecules and obtained configurations every 1 ps. We calculated the nitrile stretching frequency at the B3LYP/6-311++G(d,p) level of theory using the Morse potential analysis described above. The MeSCN molecule and the twelve water molecules nearest to its nitrogen atom, as determined by the oxygen-nitrogen distances, were included in the QM region. The additional water molecules were treated as point charges with the TIP3P partial charges ( $q_{\text{O}}=-0.834$  and  $q_{\text{H}}=0.417$ ). This level of theory was chosen based on previous work using a similar QM/MM method to calculate the vibrational frequencies of the chemically similar azide<sup>29,30</sup> and acetonitrile.<sup>32,33</sup> Analysis of the time-dependence of the frequencies calculated for the 500 configurations generated from the MD trajectory illustrated that these configurations are not correlated.

We implemented a non-linear least-squares algorithm to vary the 18 free nitrogen PM3 parameters, as well as two parameters that govern the interactions between the QM and the MM region.<sup>56</sup> In this procedure, we minimized the root-mean-squared-deviation (RMSD) between the nitrile frequencies calculated with DFT and PM3 for the 500 configurations sampled in the MD trajectory. We performed this optimization including the MeSCN molecule and either zero, one, or two water molecules in the QM region. As will be discussed below, the calculation is converged with only one water molecule in the QM region.

The optimized set of PM3 parameters is provided in Table S1 of Supporting Information. Note that all of the nitrogen PM3 parameters are within 20% of the original PM3 parameters and that the average deviation is <7%. The QM/MM interaction parameter  $\alpha$ , which governs the range of interaction between QM and MM atoms, deviates by nearly 50% from the original parameterization, but this value is nearly identical to the value used by Skinner and co-workers for the calculation of the

asymmetric stretching mode of azide in water.<sup>29</sup> The other QM/MM parameter,  $\rho$ , deviates only slightly from the original value of zero. Using these optimized parameters, the average PM3 frequency differs from the average DFT frequency by less than  $0.1 \text{ cm}^{-1}$ , and the RMSD between the DFT and PM3 frequencies for the 500 configurations is  $4.9 \text{ cm}^{-1}$ . As a point of comparison, calculations with the original PM3 parameters result in an RMSD of  $117 \text{ cm}^{-1}$  compared to the DFT calculations. The correlation between the DFT and PM3 frequencies for the 500 configurations is depicted in Figure S1 of Supporting Information. The scatter in the frequencies for the new parameter set is smaller than the measured full width at half maximum (FWHM) for all of the KSI systems studied, indicating that this reparameterized PM3 potential should be adequate to describe the shifts in the vibrational frequencies upon ligand binding.

**Table S1: Original and Optimized PM3 Nitrogen Parameters and QM/MM Interaction Parameters ( $\alpha$  and  $\rho$ )**

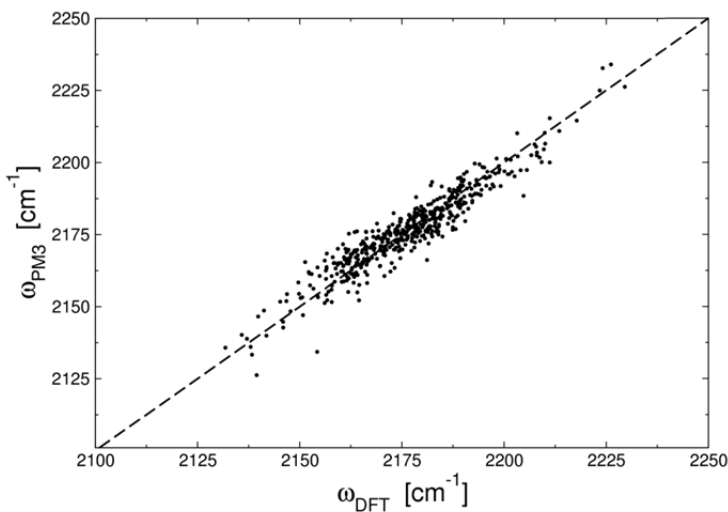
parameter	Original	optimized
$U_{ss}$ (eV)	-49.3357	-48.5707
$U_{pp}$ (eV)	-47.5097	-44.9956
$\beta_s$ (eV)	-14.0625	-14.5429
$\beta_p$ (eV)	-20.0438	-16.0230
$\zeta_s$ (eV)	2.0281	2.0196
$\zeta_p$ (eV)	2.3137	2.3520
$\alpha$ ( $\text{\AA}^{-1}$ )	2.8305	2.7917
$G_{ss}$ (eV)	11.905	11.5942
$G_{sp}$ (eV)	7.3486	6.8208
$G_{pp}$ (eV)	11.7557	11.1947
$G_{p2}$ (eV)	10.8072	10.3120
$H_{sp}$ (eV)	1.1367	1.1808
$a_1$ ( $\text{\AA}$ )	1.5017	1.5134
$b_1$ ( $\text{\AA}$ )	5.9011	6.5824
$c_1$ ( $\text{\AA}$ )	1.7107	1.6951
$a_2$ ( $\text{\AA}$ )	-1.5058	-1.4883
$b_2$ ( $\text{\AA}$ )	6.0047	6.3962
$c_2$ ( $\text{\AA}$ )	1.7161	1.6990
$\alpha$ ( $\text{\AA}^{-1}$ )	5.0000	2.4212
$\rho$ ( $\text{\AA}$ )	0.0000	0.0709

**Table S2: Vibrational Frequency of the Maximum Peak in the Simulated IR Spectra for Two Independent MD Trajectories<sup>a</sup>**

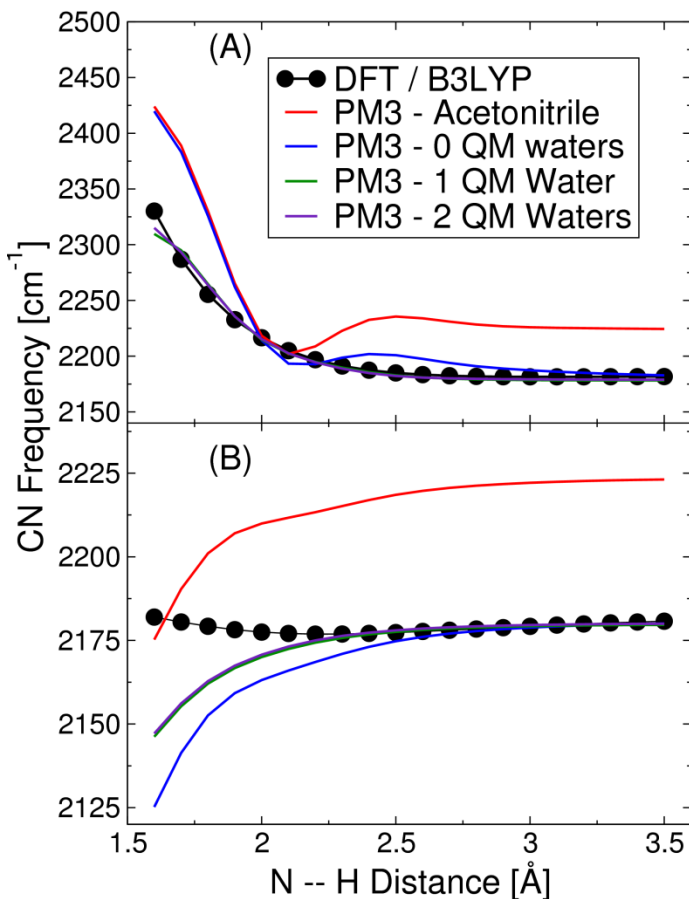
M116C-CN			
	Deprotonated Site <sup>b</sup>	Trajectory 1	Trajectory 2
No Ligand	Tyr57	2162.1	2162.3
EQU	EQU	2164.1	2164.2
EQU	Tyr16	2165.1	2164.4
F86C-CN			
	Deprotonated Site	Trajectory 1	Trajectory 2
No Ligand	Tyr57	2186.8	2187.2
EQU	EQU	2187.8	2188.0
EQU	Tyr16	2188.5	2188.3

<sup>a</sup> Units are in  $\text{cm}^{-1}$ .

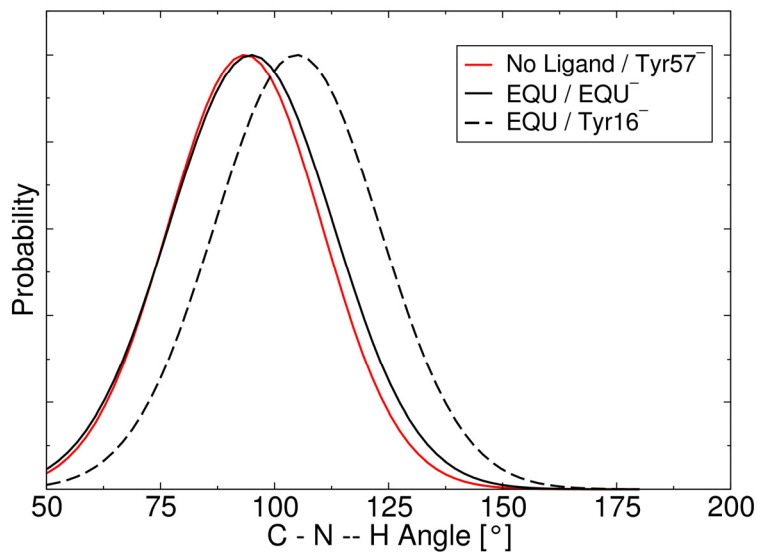
<sup>b</sup> The location of the negatively charged oxygen site to form either a tyrosinate residue or the deprotonated EQU ligand.



**Figure S1:** Correlation between the PM3 and DFT CN vibrational stretching frequencies of methyl thiocyanate in water for a set of 500 configurations. These calculations used the optimized PM3 parameters derived in this work and the DFT B3LYP/6-311++G(d,p) level of theory. The methyl thiocyanate and the nearest water molecule are treated quantum mechanically, and the remaining water molecules are treated as classical point charges. The dotted line represents perfect agreement between the PM3 and DFT calculated frequencies.

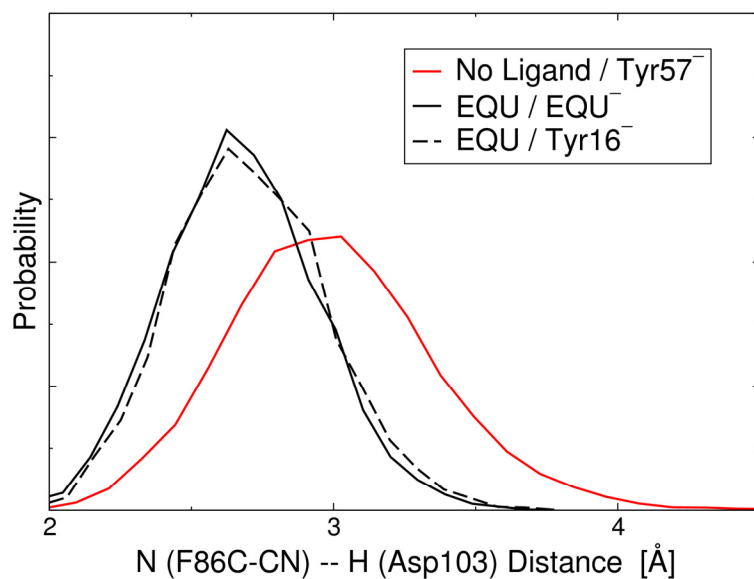


**Figure S2:** Calculated CN frequencies for the gas phase methyl thiocyanate-water dimer treated fully quantum mechanically using the following methods: DFT B3LYP/6-311++G(d,p) (black circles), PM3 with parameters developed for acetonitrile by Corcelli and coworkers<sup>33</sup> (red), and PM3 from the present work obtained with no waters (blue), one water (green), and two waters (violet) included in the QM region during the fitting procedure. The CN frequency is calculated as a function of the N--H distance with a CNH angle of (A) 90° and (B) 180°. This figure is the analog of Figure 4B in the main paper for different CNH angles.



**Figure S3:** Distribution of the CNH angle between the nitrile probe and the hydrogen-bonded water molecule as determined from the MD simulations of the M116C-CN D40N KSI system with no ligand bound and Tyr57 deprotonated (solid red line), with EQU bound in its anionic form (solid black line), and with EQU bound in its neutral form and Tyr16 deprotonated (dashed black line). This figure is the analog of Figure 6 in the main paper and corresponds to the second trajectory presented in Table S2.





**Figure S4:** Distribution of the distance between the amide hydrogen of the Asp103 backbone and the nitrogen of the F86C-CN probe as determined from the MD simulations of the F86C-CN D40N KSI system with no ligand bound and Tyr57 deprotonated (solid red line), with EQU bound in its anionic form (solid black line), and with EQU bound in its neutral form and Tyr16 deprotonated (dashed black line). This figure is the analog of Figure 8 in the main paper and corresponds to the second trajectory presented in Table S2.

### Structure Used for Potential Energy Scans in Figure 4

The potential energy scans presented in Figures 4 and S2 were performed for a methyl thiocyanate-water dimer. The geometry of the methyl thiocyanate was fixed at the geometry optimized with the previously described extension of the Amber99SB forcefield. The geometry of the water molecule was fixed at the TIP3P water geometry,  $R(\text{O-H}) = 0.96\text{\AA}$  and  $\angle(\text{H-O-H}) = 104.35^\circ$ . The Cartesian coordinates for the methyl thiocyanate-water dimer with fixed monomer geometries and relative coordinates optimized at the DFT B3LYP/6-311++G(d,p) level of theory are as follows:

C	0.000000	0.000000	0.000000
S	0.000000	0.000000	1.837226
C	1.653612	0.000000	2.166570
N	2.799384	-0.004736	2.344008
H	-1.033402	0.001948	-0.350250
H	0.513804	0.892920	-0.349283
H	0.509934	-0.894445	-0.347758
O	5.832525	0.001453	2.023890
H	4.878668	-0.000869	2.149765
H	6.190964	-0.011674	2.916663

The DFT B3LYP/6-311++G(d,p) energy of this complex is -607.463482181 Hartrees.