Supplementary Material

Chemical basis for the selectivity of the von Hippel-Lindau tumor suppressor pVHL for prolyl-hydroxylated HIF-1α

Christopher J. R. Illingworth, Christoph Loenarz, Christopher J. Schofield, and Carmen Domene

QM calculations

Small molecule calculations

In order to evaluate the performance of different QM methods to reproduce the experimentally observed conformational preferences for C-4 substituted prolyl variants in solution, studies were then carried out on *N*-acetyl proline methyl ester with and without hydroxyl and fluoro substituents at C-4 (Figure 2a, Table S1; a positive difference corresponds to the C^4-*exo* conformation as having the lower energy).

Modeled using the 6-31G* basis set, the B3LYP method was found to be the lowest level of theory to indicate the correct C^4-*endo*/*exo* preferences for all five of the molecules. Thus, the *trans*-4-hydroxyproline and *trans*-4-fluoroproline variants exhibited a preference for the C^4-*exo* conformation, albeit with only a slight preference for the former (0.04 kcal/mol), while *cis*-4-hydroxyproline and *cis*-4-fluoroproline variants and proline exhibited a preference for the *endo* conformation. Improving the size of the basis set used with this method to 6-311++G** increased the C^4-*endo*/*exo* difference in *trans*-4-hydroxyproline. However, a substantial difference between the conformations in this variant is only observed with the MP2 level of theory, where the *exo* conformation is clearly preferred (~0.5 kcal/mol). As with the B3LYP functional, increasing the size of the basis set leads to a small increase (0.08 kcal/mol) in the energy difference.

Calculations were repeated for the same molecules in an implicit aqueous solvent (Table S2). Similar preferences were obtained for all of the variants except for proline itself, where the *exo* conformation

had lower energy than the C^4-*endo* form (a difference of 0.19 kcal/mol at B3LYP/6-31G*). We suggest that this preference could be related to the solvent model. NMR calculations on *N*-acetyl proline methyl ester in dioxane solution indicate a preference for an *endo* conformation(*1*), while an aqueous solvent model was used for calculation. A noticeable difference was evident in the behavior of *trans*-4-hydroxyproline in solvent, with a measured preference of close to 1 kcal/mol for the C^4-*exo* conformation.

The above calculations indicate that of the methods tested, MP2 methods are the best suited to reproduce the experimentally observed preferential C^4-*exo* conformation in *trans*-4-hydroxyproline for free peptides in solution, suggesting the importance of electron correlation in obtaining accurate values. However, the B3LYP/6-31G* level of theory was initially chosen for use as the highest feasible level of theory for which geometry optimizations of the entire system (>250 atoms) could be modeled. Despite limitations in reproducing the correct conformation of *trans*-4-hydroxyproline, this reproduces the correct *in vacuo* conformational preferences for the other proline substitutions, while remaining within the bounds of computational feasibility.

Solvation effects

Our calculations on interaction energy described in the main body were carried out *in vacuo*, thus neglected effects such as desolvation. It would be expected that the free energy of desolvation of the proline residue would be altered by the substitution of a hydroxyl group. In order to derive an approximation of the effect of desolvation on the calculated interaction energies, free energy calculations were performed on the *N*-acetyl proline methyl ester derivatives used for the calculations described above. Using Gaussian03, prolyl and *trans*-4-hydroxyprolyl variants of the residue were optimized both *in vacuo* and in an IEFPCM implicit solvent at the B3LYP/6-31G* level of theory, performing frequency calculations to derive approximate free energies. The model energies calculated in implicit solvent were -10.26 kcal/mol lower than *in vacuo* for the prolyl variant, and -16.04 kcal/mol

lower than *in vacuo* for the *trans*-4-hydroxyprolyl variant. This equates to a difference in the desolvation free energy of 5.78 kcal/mol in favor of the prolyl variant, which would lead to a corresponding reduction in the calculated difference between the interaction energies given for the systems in the main body of the paper. Note that with calculations taking into account desolvation, the *exo*-conformation of *trans*-4-hydroxyproline is predicted to be the most stable conformer by all three calculation methods.

Table S1. Differences in the predicted gas phase energies of the C^4-*endo* and *exo* conformations of C-4 substituted variants of *N*-acetyl proline methyl ester, measured in vacuum using a range of basis sets and levels of theory**.** A positive difference indicates a preference for the C^4-*exo* conformation.

	Difference in predicted gas phase energies [*exo-endo*] (kcal/mol)						
	HF 3-21G	**HF 6-31G***	**HF 6-311++G****	**B3LYP 6-31G***	**B3LYP 6-311++G****	**MP2 6-31G***	**MP2 6-311++G****
***trans*-4-hydroxyproline**	-0.55	-0.01	+0.08	+0.04	+0.08	+0.47	+0.55
***cis*-4-hydroxyproline**	-2.86	-0.53	-0.04	-1.10	-0.49	-2.24	-1.55
***trans*-4-fluoroproline**	-0.58	+0.7	+0.90	+0.84	+0.99	+1.06	+0.99
***cis*-4-fluoroproline**	-1.60	-1.04	-0.68	-1.37	-1.15	-2.16	-1.53
proline	-0.69	-0.25	-0.15	-0.20	-0.25	-0.49	-0.46

Table S2. Predicted differences in energy between C^4-*endo* and *exo* conformations of variants of *N*-acetyl proline methyl ester, measured at different levels of theory and basis sets, using the IEFPCM implicit aqueous solvent method.

	Level of theory		
	B3LYP/6-31G*	**B3LYP/6-311+G(2p,d)**	**MP2/6-31G***
***trans*-4-hydroxyproline**	+0.94	+1.02	+1.07
***cis*-4-hydroxyproline**	-0.67	-0.15	-0.66

***trans*-4-fluoroproline**	+1.99	+2.18	+2.08
***cis*-4-fluoroproline**	-1.65	-1.56	-1.50
proline	+0.19	+0.09	+0.26

Table S3. Calculated interaction energies with VCB for CODD containing different variants of prolyl-564_{CODD}. Calculated Interaction energies are given for the systems in which solvent waters were included. Interaction energies are given for the 3-residue fragment (containing $P564_{CODD}$ and adjacent residues $A563_{CODD}$ and $Y565_{CODD}$), and for $P564_{CODD}$ in isolation. The only significant difference to the obtained results calculated without the solvent is for the case of *cis*-4-hydroxyproline in the C^4-*endo* conformation, where the hydroxyl group of the $Y98_{VCB}$ residue forms a hydrogen bond with a solvent molecule rather than the carbonyl oxygen of $P564_{CODD}$, reducing the interaction energy by ~6 kcal/mol.

	Interaction energy (kcal/mol)	
	3-residues	1-residue
	A563-P564-Y565	P564
***trans*-4-hydroxyprolyl (*exo*)**	-82.52	-30.14
***trans*-4-hydroxyprolyl (*endo*)**	-78.54	-22.74
***cis*-4-hydroxyprolyl (*exo*)**	-71.38	-25.66
***cis*-4-hydroxyprolyl (*endo*)**	-54.64	-8.64
***trans*-4-fluoroprolyl (*exo*)**	-73.94	-20.60
***cis*-4-fluoroprolyl (*exo*)**	-69.41	-16.30
prolyl (*exo*)	-66.56	-10.60
prolyl (*endo*)	-65.31	-11.70

Table S4. Decomposition of calculated interaction energies for the three-residue fragment A563-P564-Y565 of CODD for different variants and C^4-*endo*/*exo* conformations of the $P564_{CODD}$ residue, from calculations at the B3LYP/6-31G* level of theory, for systems in which the solvent waters were included. Residues which make the largest contribution to the difference in the calculated interaction between different prolyl variants ($Y98_{VCB}$, $S111_{VCB}$ and $H115_{VCB}$) are in bold.

	Interaction Energy (kcal/mol)									
$P564_{CODD}$ variant	**W88**	**F91**	**Y98**	**I109**	**H110**	**S111**	**Y112**	**H115**	**W117**	**Water**
t-4-Hyp (*exo*)	0.58	-0.81	**-12.02**	-1.09	-14.39	**-5.84**	0.04	**-9.74**	-0.47	-41.21
t-4-Hyp (*endo*)	-2.74	-0.75	**-12.36**	-1.09	-13.25	**-4.01**	0.64	**-1.23**	-3.06	-41.22
c-4-Hyp (*exo*)	1.02	-0.65	**-7.44**	-1.77	-10.61	**-3.99**	-5.04	**-10.15**	-0.70	-31.60
c-4-Hyp (*endo*)	0.33	-0.14	**-6.40**	-1.76	-10.11	**1.18**	-4.6	**-1.56**	0.18	-28.82
t-4-Flp (*exo*)	-1.30	-0.76	**-11.94**	-0.98	-14.90	**-4.79**	0.20	**-0.50**	-0.38	-41.16
c-4-Flp (*exo*)	-1.87	-0.74	**-11.76**	-1.02	-13.79	**-2.75**	0.67	**-1.41**	-0.92	-41.27
Pro (*exo*)	-0.26	-0.83	**-11.19**	-1.17	-14.36	**0.75**	0.59	**-1.88**	0.30	-43.22
Pro (*endo*)	-0.12	-0.79	**-10.87**	-1.18	-13.83	**1.12**	0.66	**-1.70**	0.47	-41.73
t-4-Hyp HIP (*exo*)	-1.76	-0.77	**-11.66**	-1.14	-13.27	**-4.36**	0.54	**-12.99**	-2.82	-41.06
t-4-Flp HIP (*exo*)	-1.82	-0.72	**-7.46**	-0.27	-11.12	**-2.25**	-0.16	**-8.35**	-0.05	-26.38
Pro HIP (*exo*)	-0.67	-0.96	**-11.18**	-1.19	-14.33	**0.67**	0.69	**2.36**	0.48	-41.81

The largest contributions to selectivity are in the $Y98_{VCB}$, $S111_{VCB}$, and $H115_{VCB}$ residues. $H115_{VCB}$ is in the HIE state unless stated otherwise.

Table S5. Decomposition of calculated interaction energies for the three-residue fragment A563-P564-Y565 of CODD for different variants and C^4-*endo*/*exo* conformations of the $P564_{CODD}$ residue, from calculations at the B3LYP/6-31G* level of theory, for systems in which the solvent water molecules were included. Residues which make the largest contribution to the difference in calculated interaction between different prolyl variants ($Y98_{VCB}$, $S111_{VCB}$ and $H115_{VCB}$) are in bold.

	Interaction Energy (kcal/mol)									
$P564_{CODD}$ variant	**W88**	**F91**	**Y98**	**I109**	**H110**	**S111**	**Y112**	**H115**	**W117**	**Water**
t-4-Hyp (*exo*)	0.71	0.02	**-9.06**	0.12	-2.31	**-6.92**	-0.62	**-8.31**	-0.50	-1.00
c-4-Hyp (*endo*)	-2.55	0.06	**-9.55**	0.06	-1.45	**-4.67**	-0.16	**-0.29**	-3.31	-0.86
c-4-Hyp (*exo*)	0.97	0.10	**-6.06**	0.03	-3.25	**-4.59**	-2.11	**-8.78**	-0.89	-2.23
c-4-Hyp (*endo*)	0.37	0.05	**-5.71**	0.03	-2.97	**1.05**	-1.05	**-0.79**	-2.00	0.22
t-4-Flp (*exo*)	-1.13	0.04	**-8.86**	0.22	-2.87	**-5.66**	-0.51	**0.31**	-0.37	-1.02
c-4-Flp (*exo*)	-2.07	0.03	**-8.66**	0.09	-1.88	**-3.39**	-0.12	**-0.49**	-1.18	-1.11
Pro (*exo*)	-0.13	0.02	**-8.34**	0.06	-2.24	**0.15**	-0.05	**-0.96**	0.29	-1.72
Pro (*endo*)	-0.15	0.02	**-8.43**	0.02	-1.84	**0.65**	-0.04	**-0.75**	0.45	-1.01
t-4-Hyp HIP (*exo*)	-1.07	0.03	**-8.76**	0.12	-1.67	**-5.01**	-0.25	**-15.80**	-2.71	-0.61
t-4-Flp HIP (*exo*)	-0.76	0.06	**-5.64**	0.30	-2.65	**-3.09**	-0.45	**-10.99**	0.01	1.17
Pro HIP (*exo*)	-0.24	-0.03	**-8.35**	0.05	-2.28	**0.11**	-0.02	**0.77**	0.47	-1.16

Legend: The largest contributions to selectivity are in the $Y98_{VCB}$, $S111_{VCB}$, and $H115_{VCB}$ residues. $H115_{VCB}$ is in the HIE state unless stated otherwise.

Table S6. Calculated interaction energies for the three-residue CODD fragment (A563-P564-Y565$_{CODD}$), and for the single-residue P564$_{CODD}$, calculated for structures re-optimized at the HF/6-31G* level of theory. The H115$_{VCB}$ residue was in the HIE protonation state. Energies are given for both the systems with solvent water molecules included and for the systems with the water molecules excluded.

	Optimization//Energies			
	Waters included		**Waters excluded**	
	Interaction energy (kcal/mol)			
	3-residues A563-P564-Y565	1-residue P564	3-residues A563-P564-Y565	1-residue P564
***trans*-4-hydroxyprolyl (*exo*)**	-82.52	-30.14	-40.74	-23.47
***cis*-4-hydroxyprolyl (*exo*)**	-71.38	-25.66	-32.89	-20.12
***trans*-4-fluoroprolyl (*exo*)**	-73.94	-20.60	-29.67	-15.31
***cis*-4-fluoroprolyl (*exo*)**	-69.41	-16.30	-26.52	-12.86
prolyl (*exo*)	-66.56	-10.60	-21.83	-8.41

Table7S. Calculated interaction energies for different variants of HIFα hydroxyprolyl with electron correlation effects in the prolyl ring either included or excluded from the model**.**

System	**Single-residue Interaction Energy NO electron correlation (kcal/mol)**	**Single-residue Interaction Energy MP2 correlation in P564 (kcal/mol)**
trans-4-hydroxyprolyl (*exo*)	-23.47	-23.82
trans-4-hydroxyprolyl (*endo*)	-17.74	-16.24
cis-4-hydroxyprolyl (exo)	-20.12	-19.87
cis-4-hydroxyprolyl (*endo*)	-6.61	-4.86
Prolyl (*exo*)	-8.41	-8.91

Prolyl (*endo*)	-8.10	-7.78

References

1. DeRider, M. L., Wilkens, S. J., Waddell, M. J., Bretscher, L. E., Weinhold, F., Raines, R. T., and Markley, J. L. (2002) Collagen stability: Insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations, *J Am Chem Soc 124*, 2497-2505.