Computational and Machine Learning Exploration of Pd-C Photocleavage in T-shaped Organopalladium Complexes
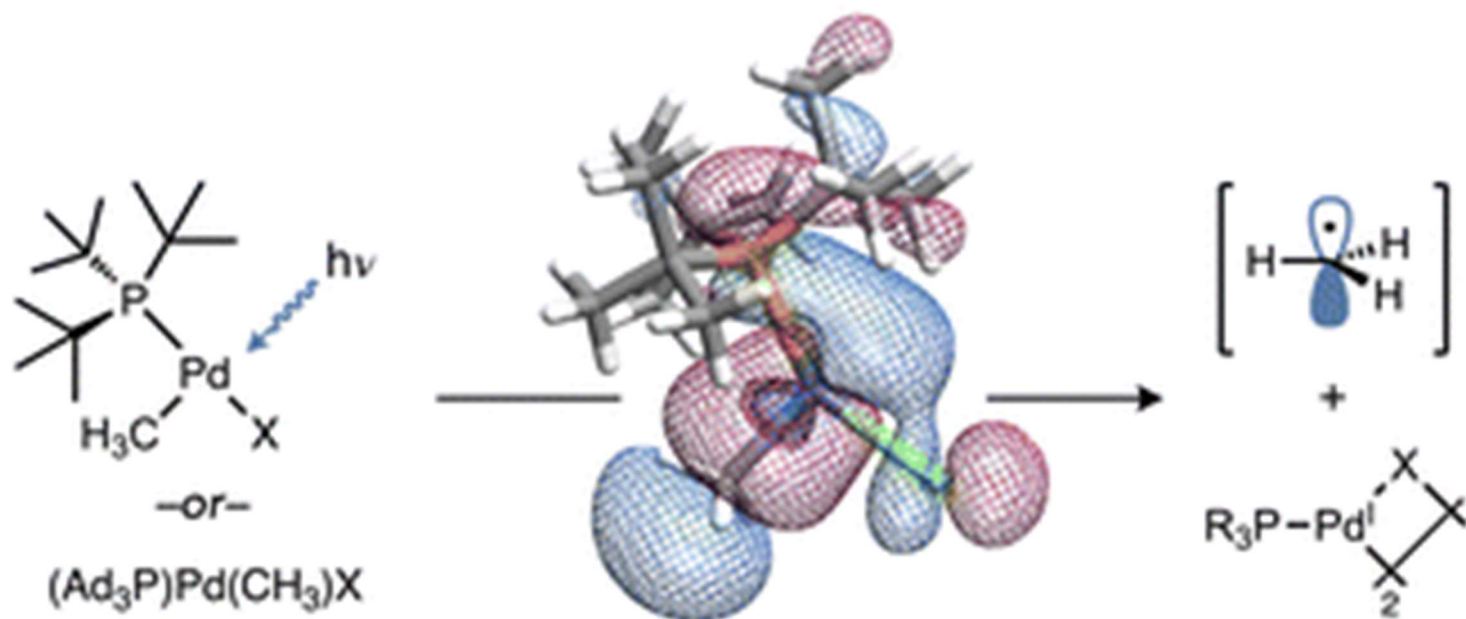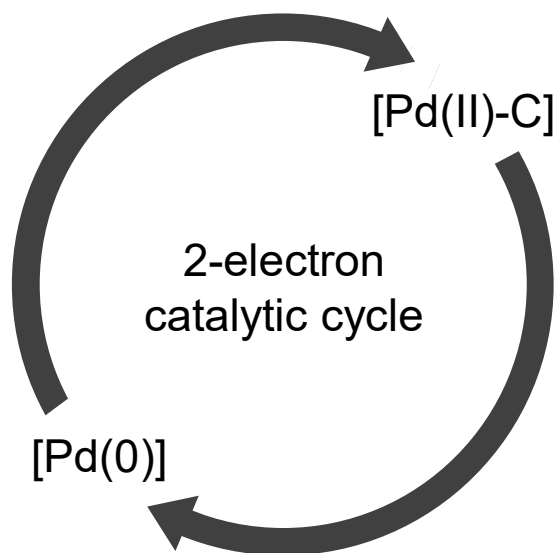

Peter M. Waddell, Ph.D.
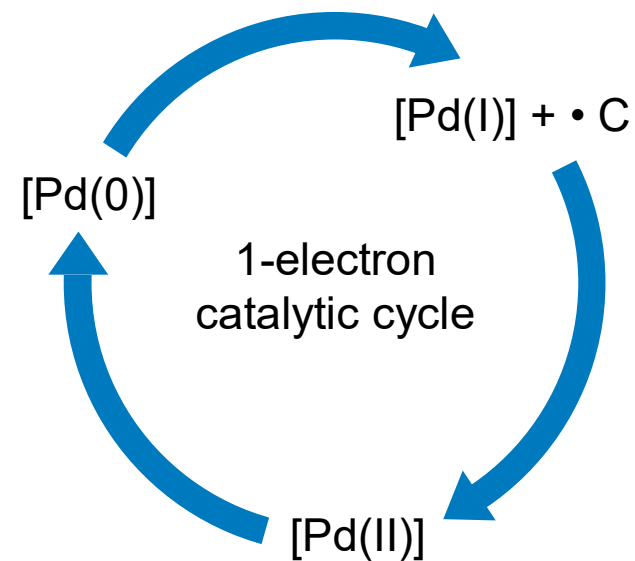
T-SHAPED COMPLEX Pd-C PHOTOCLEAVAGE WITH VISIBLE LIGHT
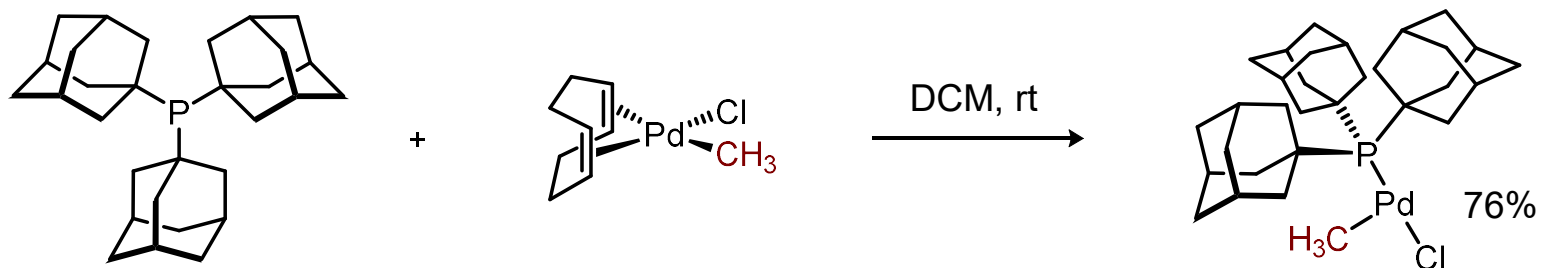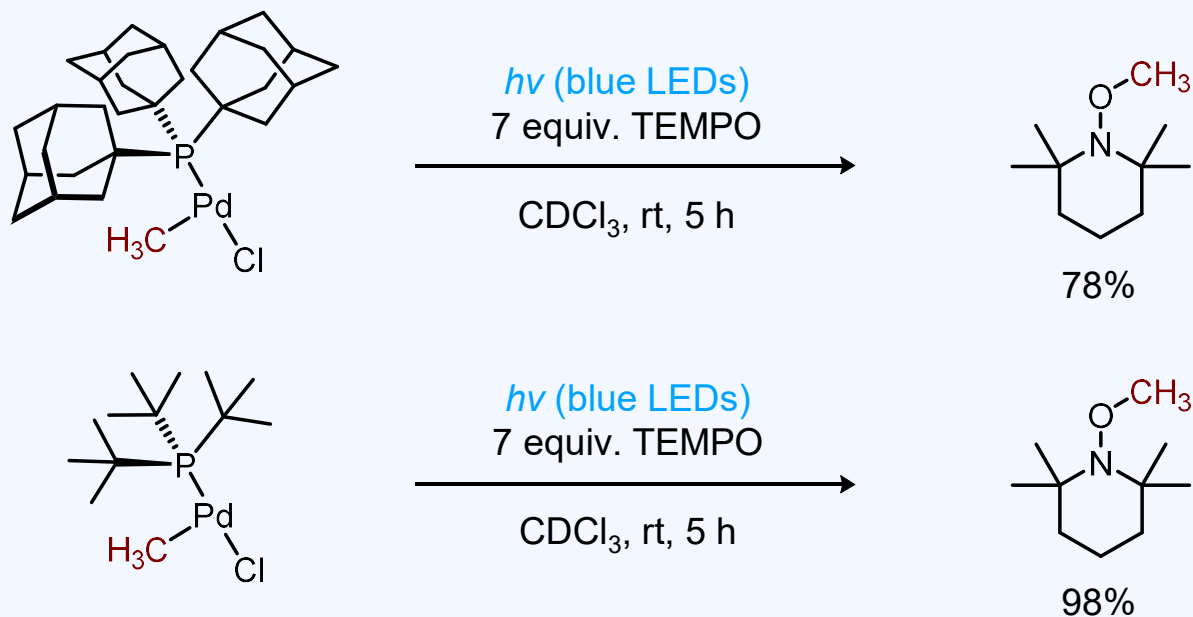

Do these T-shaped complexes exhibit any photoreactivity?

Visible light-induced bond weakening triggers SET chemistry in the most synthetically versatile metal and oxidation state (d^8, Pd^{II})!
Previously, this reactivity was much more limited to high-energy UV light.

Pd-C PHOTOCLEAVAGE WITH VISIBLE LIGHT
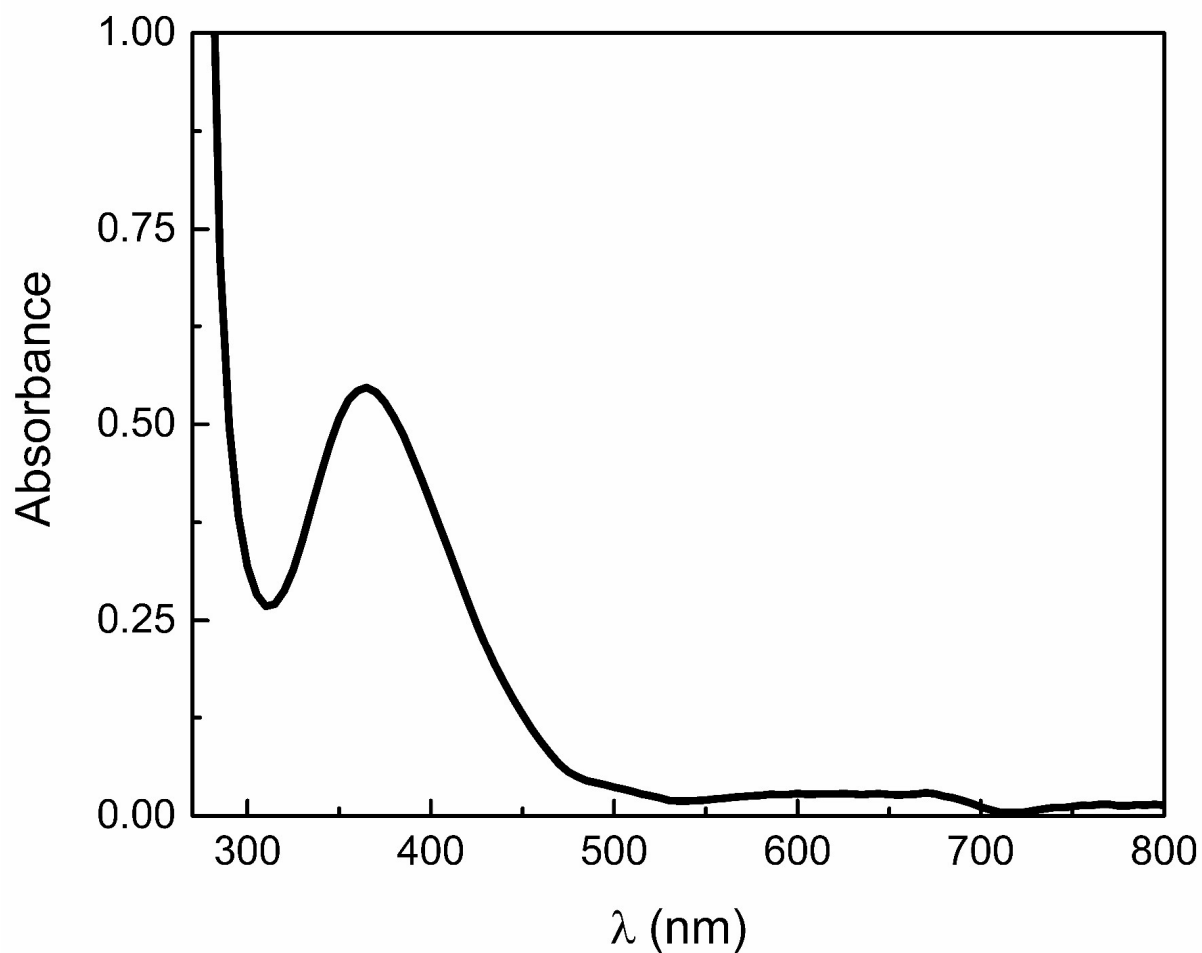
Absorption spectrum of the T-shaped complex shows the key peak for the transition that leads to Pd-C cleavage, with a λ_{max} just under 400 nm.

HOMO/LUMO ANALYSIS

with Lei Tian

LUMO: $d_{x^2-y^2}/\sigma^*$ ⎯⎯

Gap: 2.72 eV
(456 nm)
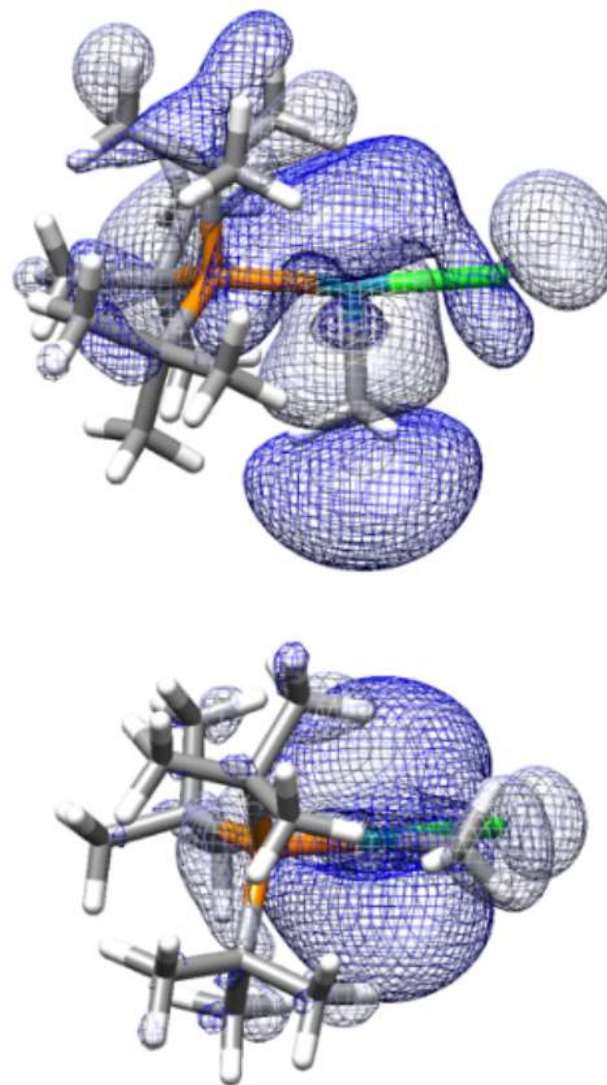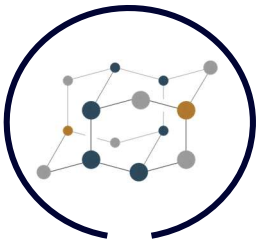
HOMO: d_{z^2} ⎯⎯

LUMO shows Pd-C σ* character, ***possibly accessible with <u>visible light</u> due to the absence of a ligand** trans* **to C in the fourth site of the square plane.***

Carrow *Chem. Sci.* **2023**, *14*, 14217.
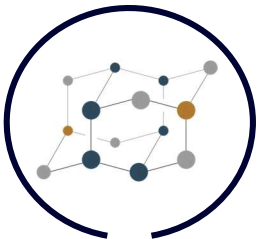
COMPUTATIONAL STUDY OBJECTIVES

- How do the steric and electronic properties of the ancillary ligand affect the UV-vis absorption spectrum of their organopalladium complexes? What characteristics cause red or blue shifts in the λ_{max} of the key peak? What is the lowest energy light that might be able to access this reactivity?

- <u>Hypothesis:</u> stronger bonding interaction in the fourth site will cause a blue shift in λ_{max}.

- Can we build a model with machine learning techniques to predict the spectral profile of a complex based only on information from its ancillary ligand?

- How conserved is the Pd-C photocleavage reactivity among a broad set of organophosphorus ligands?
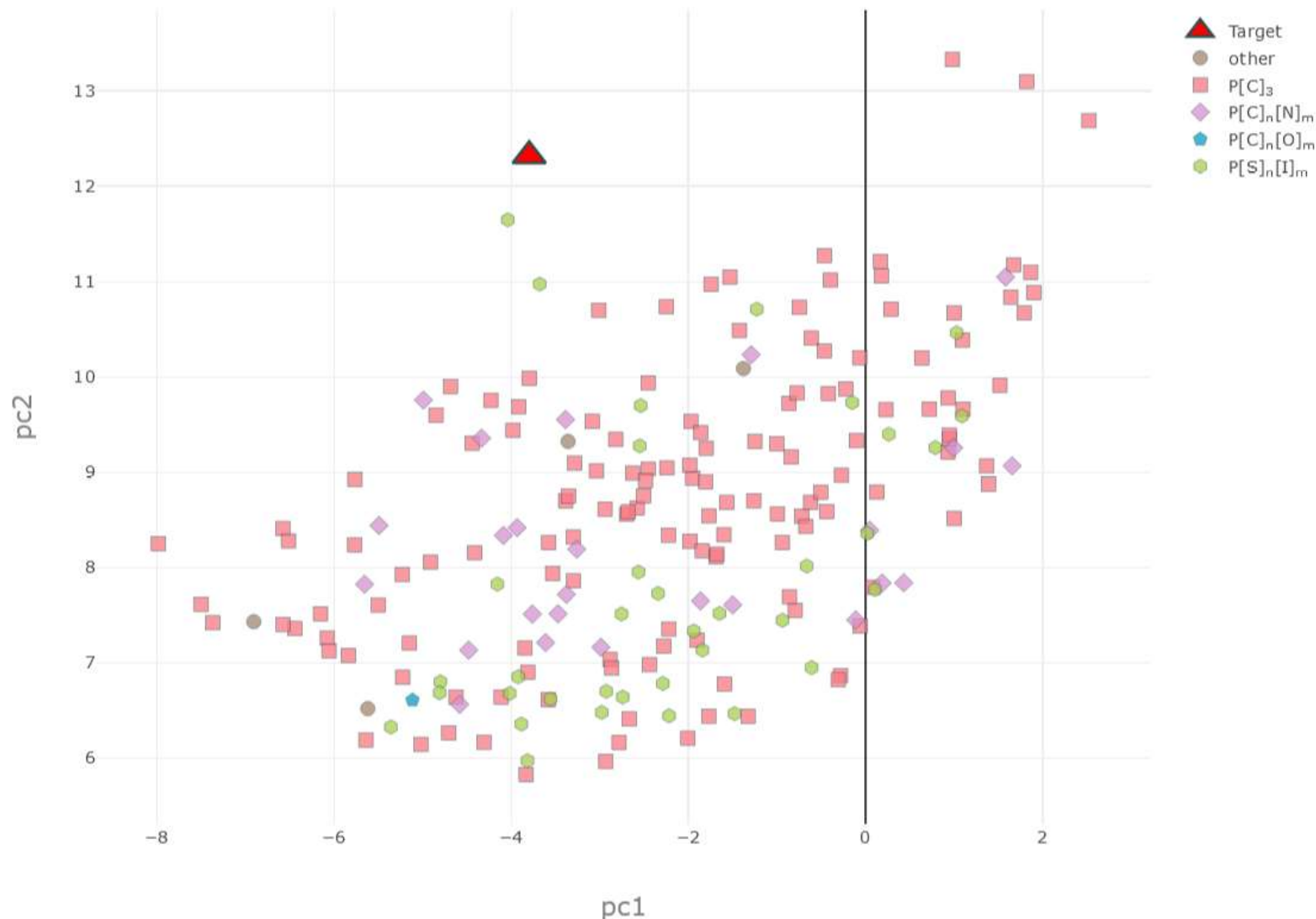
EXPLORING PHOSPHINE LIGAND SPACE WITH KRAKEN LIBRARY
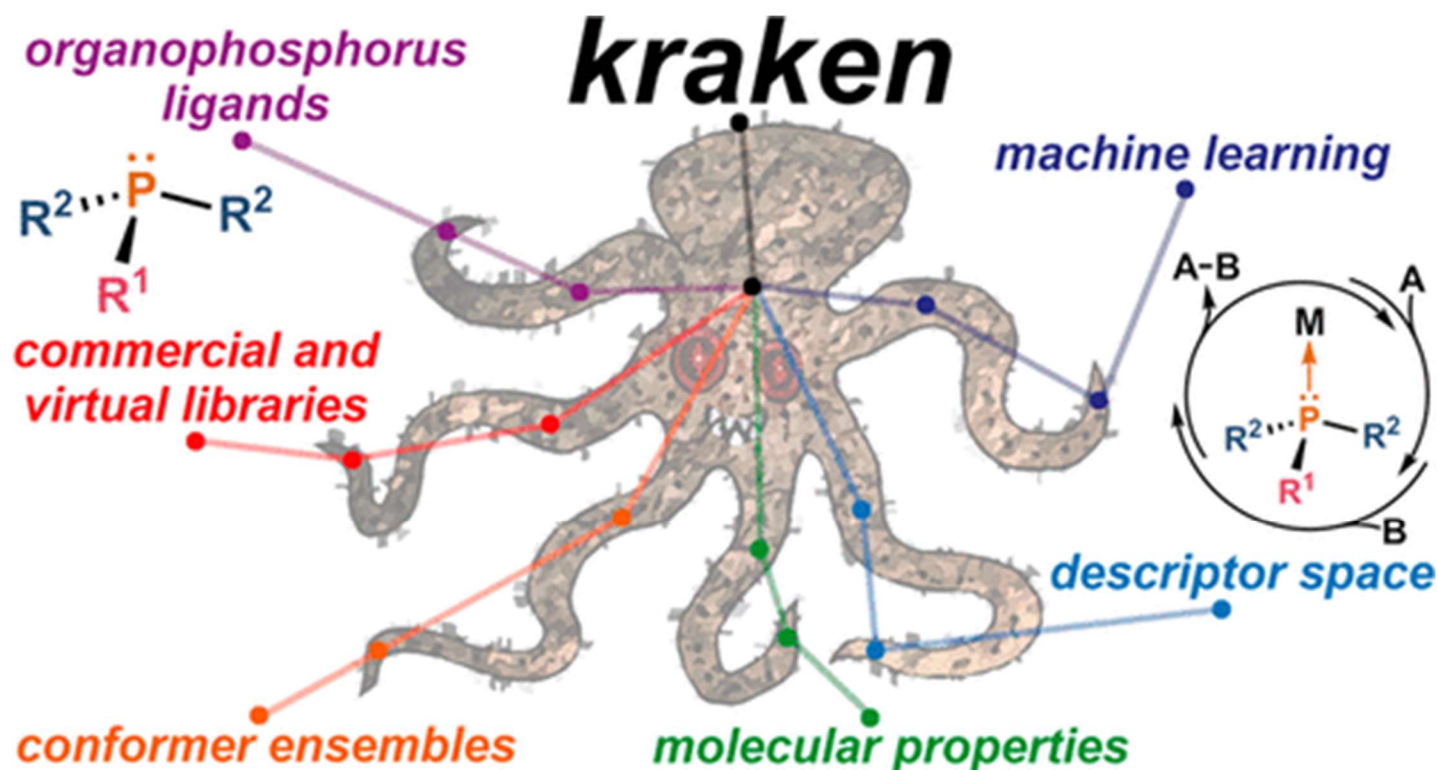
The Kraken library of phosphine ligands is a perfect starting point to explore the chemical space of bulky organophosphorus ligands.

Gensch, Sigman, Aspuru-Guzik *J. Am. Chem. Soc.* **2022**, *144*, 1205.

EXPLORING PHOSPHINE LIGAND SPACE WITH KRAKEN LIBRARY

Legend:
- Target
- other
- $P[C]_3$
- $P[C]_n[N]_m$
- $P[C]_n[O]_m$
- $P[S]_n[I]_m$

To start, I used the Kraken library to find the 210 closest (in principal component space) organophosphorus ligands to our target (tri-*t*-butylphosphine)

# EXPLORING PHOSPHINE LIGAND SPACE WITH KRAKEN LIBRARY
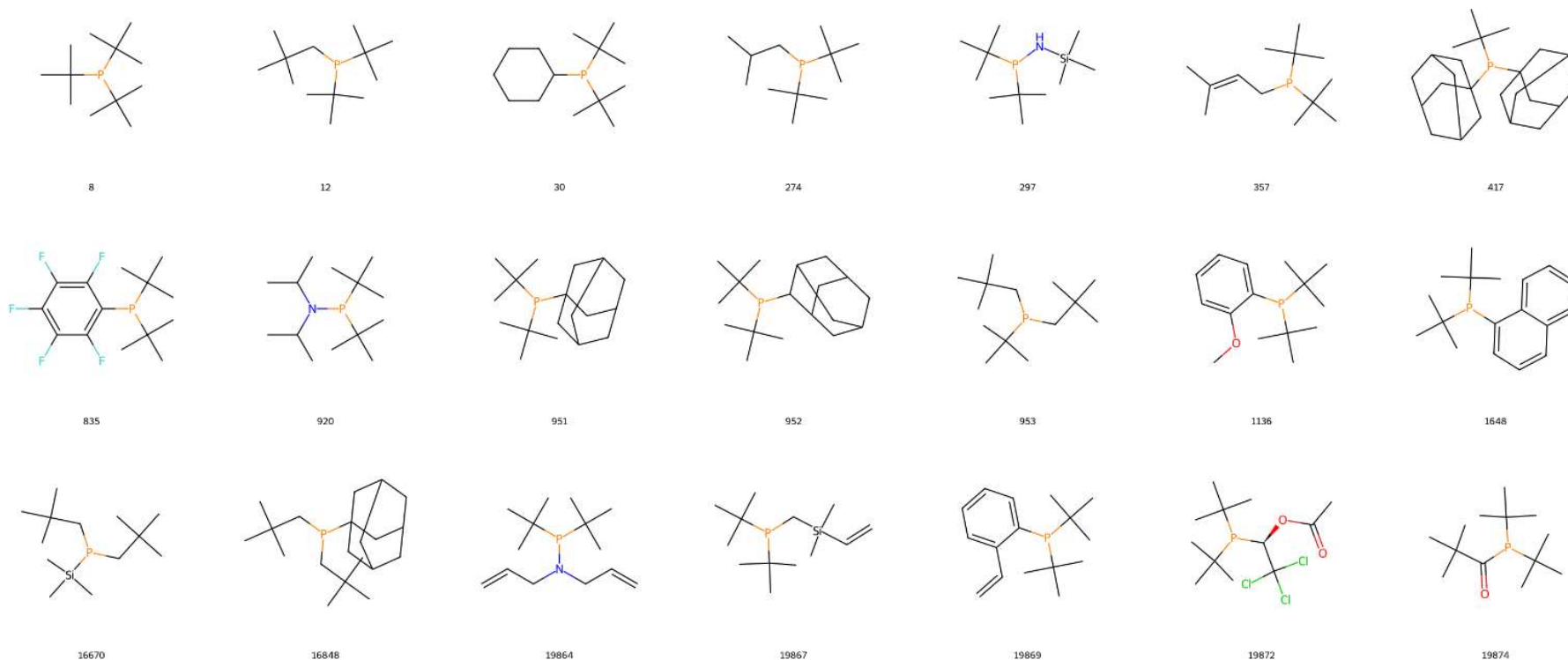
In order to ensure that the organophosphorus ligands would form T-shaped (monoligated) Pd complexes, I restricted this set to the ligands with Boltzmann-average $\%V_{bur}$ of 68% or more, very close to the value for tri-t-butylphosphine.

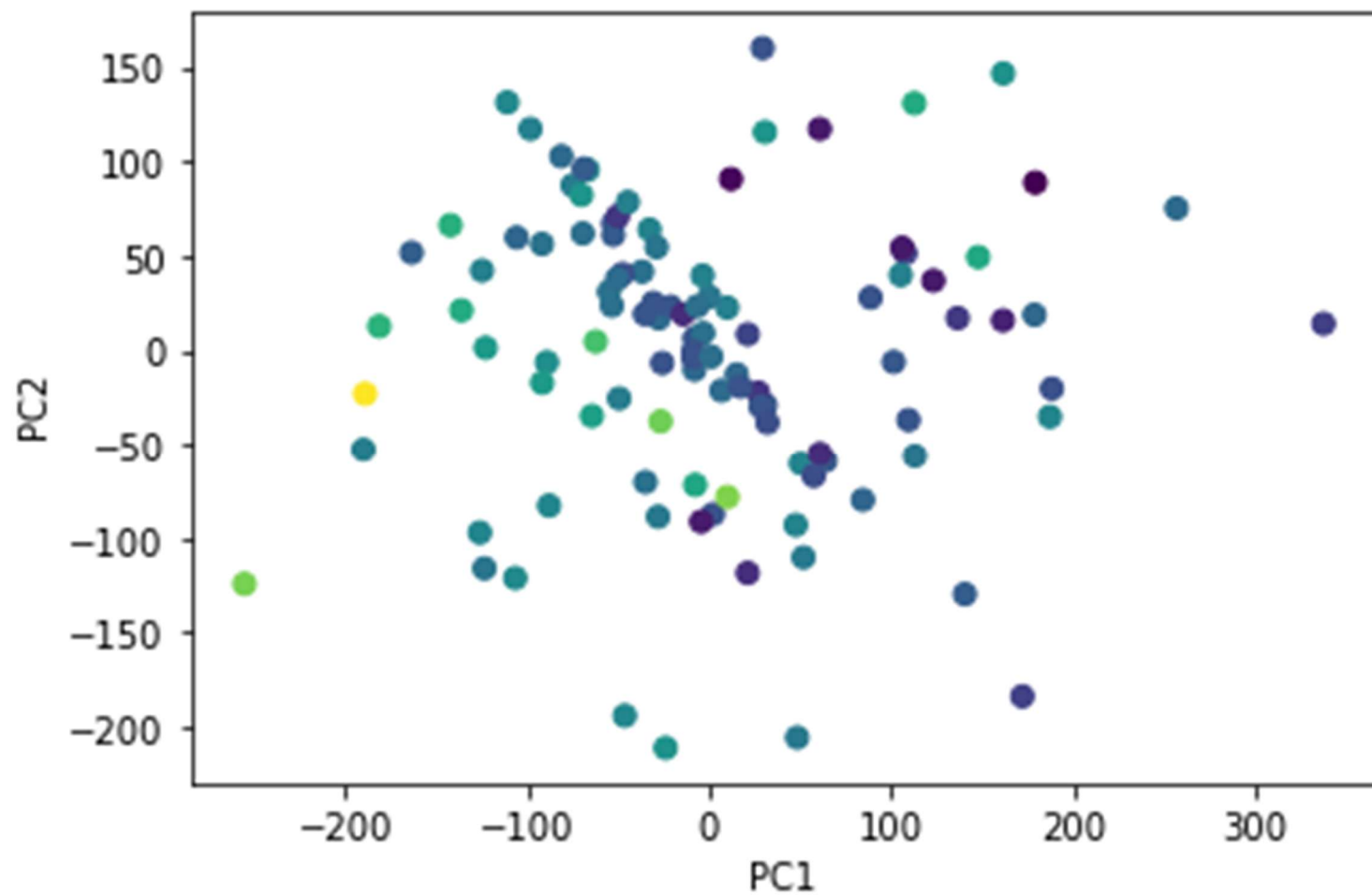This gave a final set of **107 ligands** (small subset pictured).
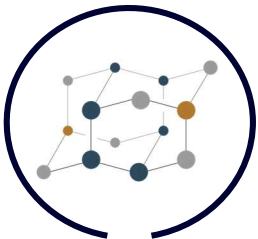
LIGAND SET PRINCIPAL COMPONENT ANALYSIS
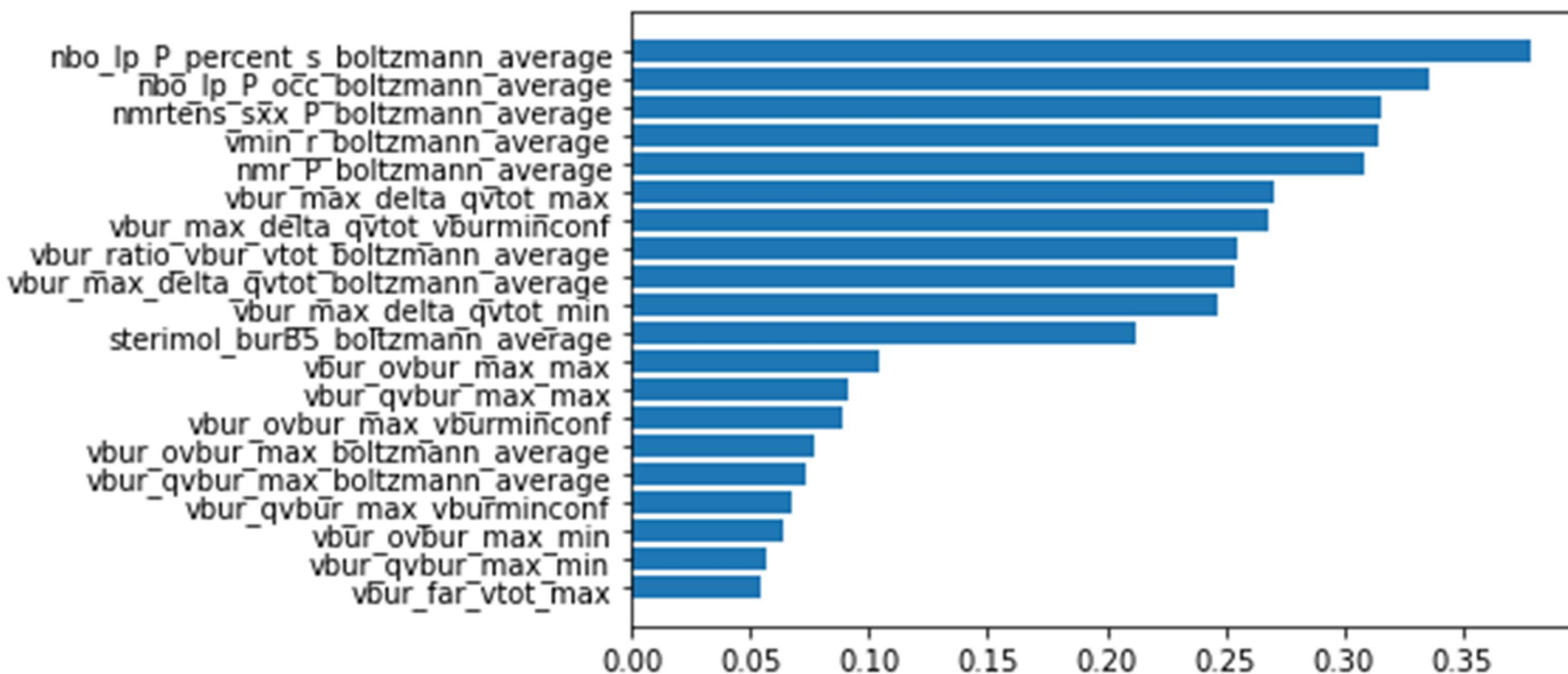
Here are the twenty most important features for the first principal component. Notably, electronic features such as Boltzmann-average NBO P lone pair %s and NMR values rank most highly.
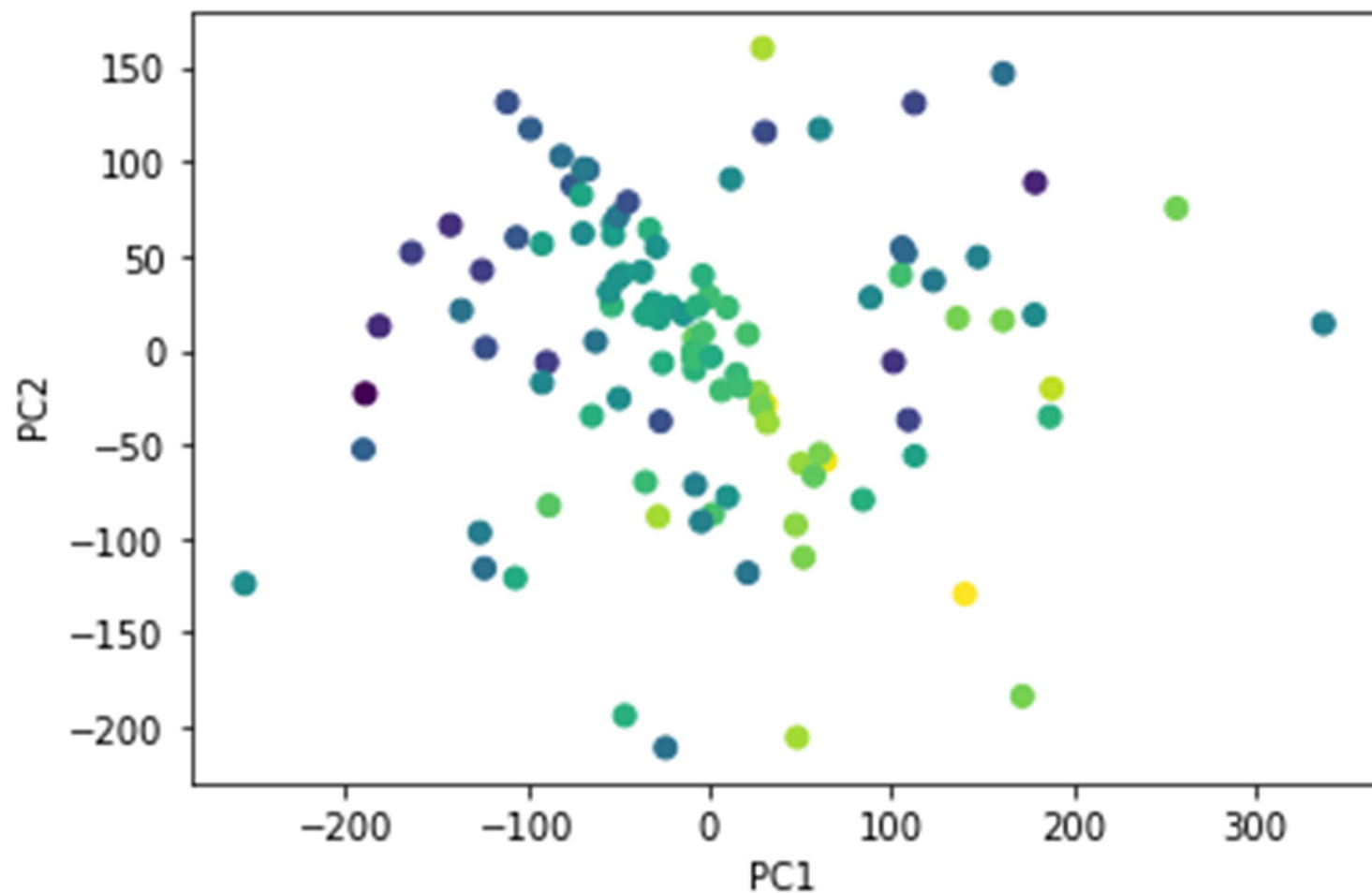
LIGAND SET PRINCIPAL COMPONENT ANALYSIS

PCA on our ligand set, colored according to Boltzmann-average NBO P lone pair %s.

LIGAND SET PRINCIPAL COMPONENT ANALYSIS

PCA on our ligand set, colored according to highest total volume difference in %V_{bur} between two neighboring quadrants.

COMPUTATIONAL WORKFLOW

- Gather and prune ligand set from Kraken database

- Generate SMILES strings for methylpalladium chloride T-shaped complexes for each ligand

- Use RDKit to run conformer searches for each complex

- Perform geometry optimizations based on the **_lowest energy conformer only_** for each complex (note that currently I am neglecting all other conformers at this time, which I intend to address later with more compute time)

- Perform TDDFT calculations and simulate the UV-vis spectra for each complex, determine the λ_{max} for the relevant transition

- Find trends in λ_{max} and ligand/complex properties

- Use machine learning to build and validate models to predict λ_{max} for other organophosphorus ligands, extract insights

- See the repo at: https://github.com/pmwaddell/pd-c-photochem-ML for the source code, all steps are automated with scripts

DFT: LEVEL OF THEORY, GEOMETRY OPTIMIZATION

Literature X-ray structure

DFT geometry optimized structure
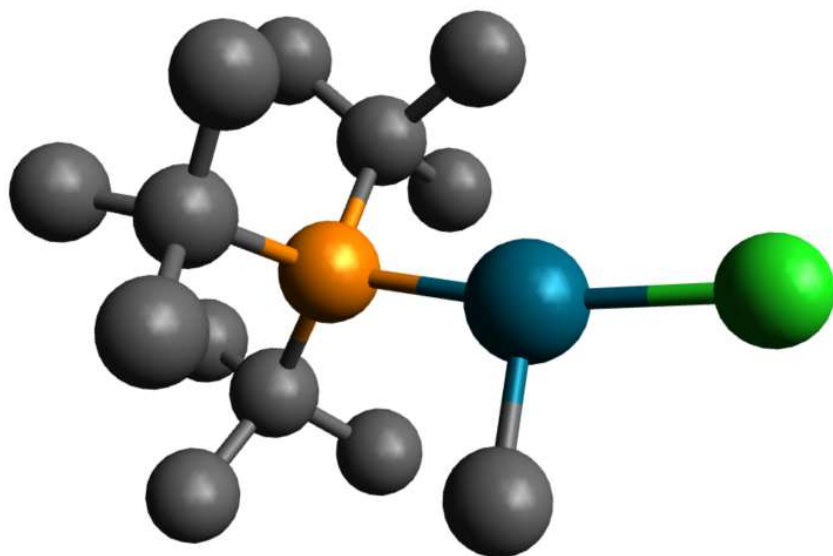

From Nozaki *Organometallics* **2006**, *25*, 4588.
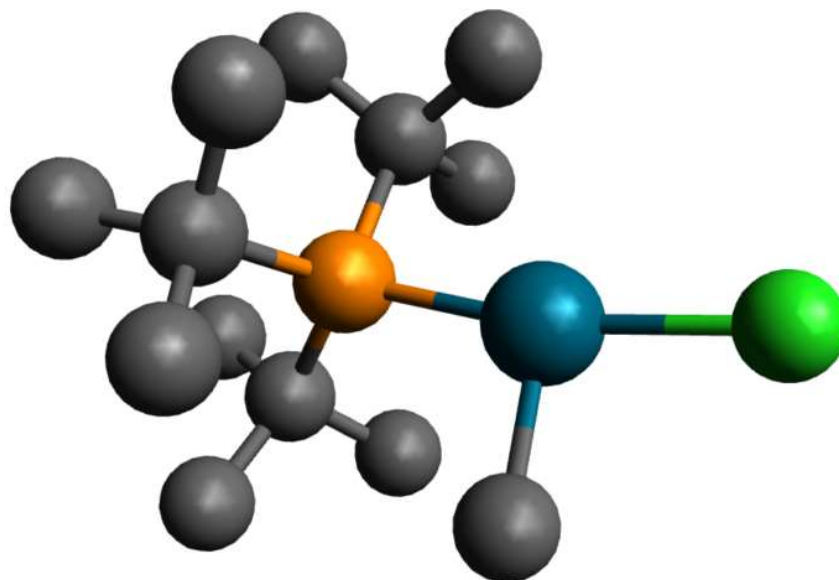
(hydrogens omitted)

After screening various basis sets and functionals for geometry optimization, **B3LYP-D3/def2-TZVP with CPCM(CHCl₃)** was found to give good agreement (key bond lengths within 0.05 Å, angles within 1°) with the reported X-ray structure while being relatively time-efficient.
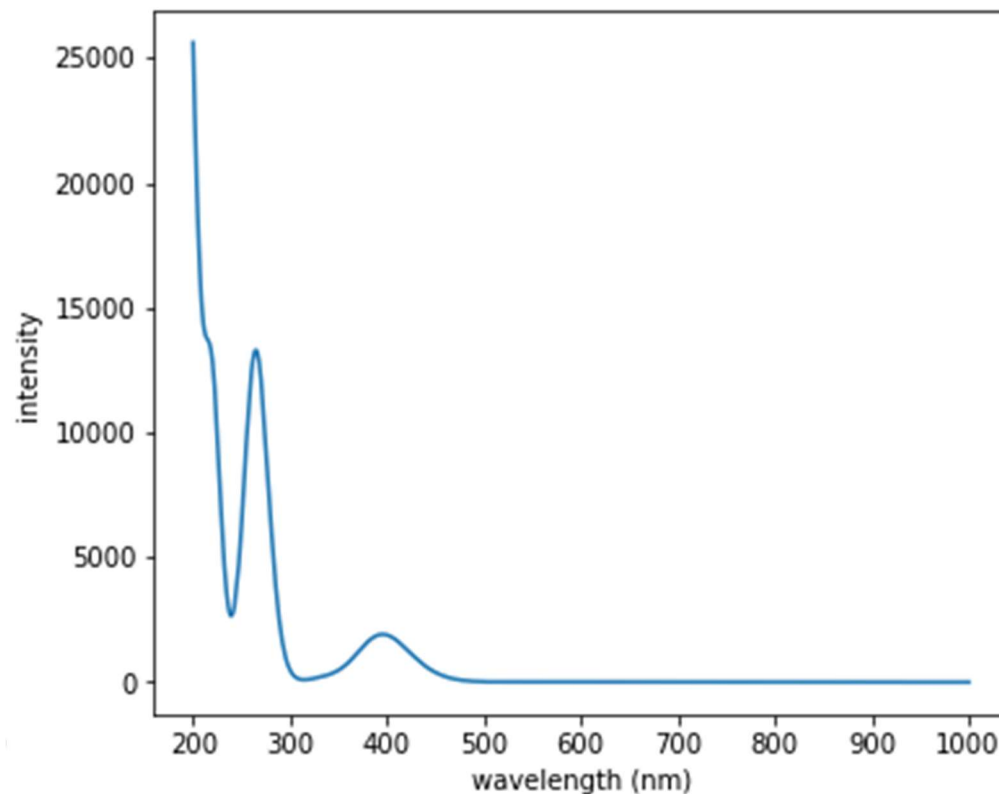
DFT: LEVEL OF THEORY, TDDFT

TDDFT UV-vis spectrum
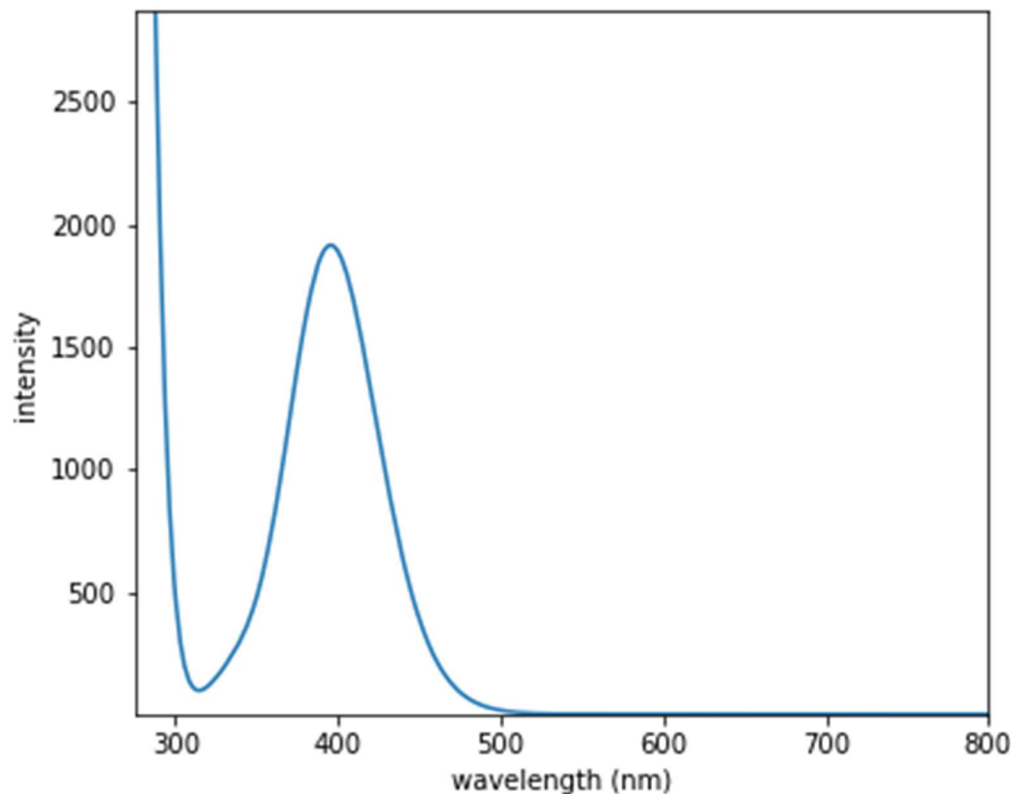
TDDFT UV-vis spectrum, full

For TDDFT calculations, agreement with the experimental UV-vis absorption spectrum was used to benchmark performance. Similarly, **B3LYP/def2-TZVPP with CPCM(CHCl$_3$)** was found to give the closest agreement.

DFT: LEVEL OF THEORY, TDDFT

TDDFT UV-vis spectrum
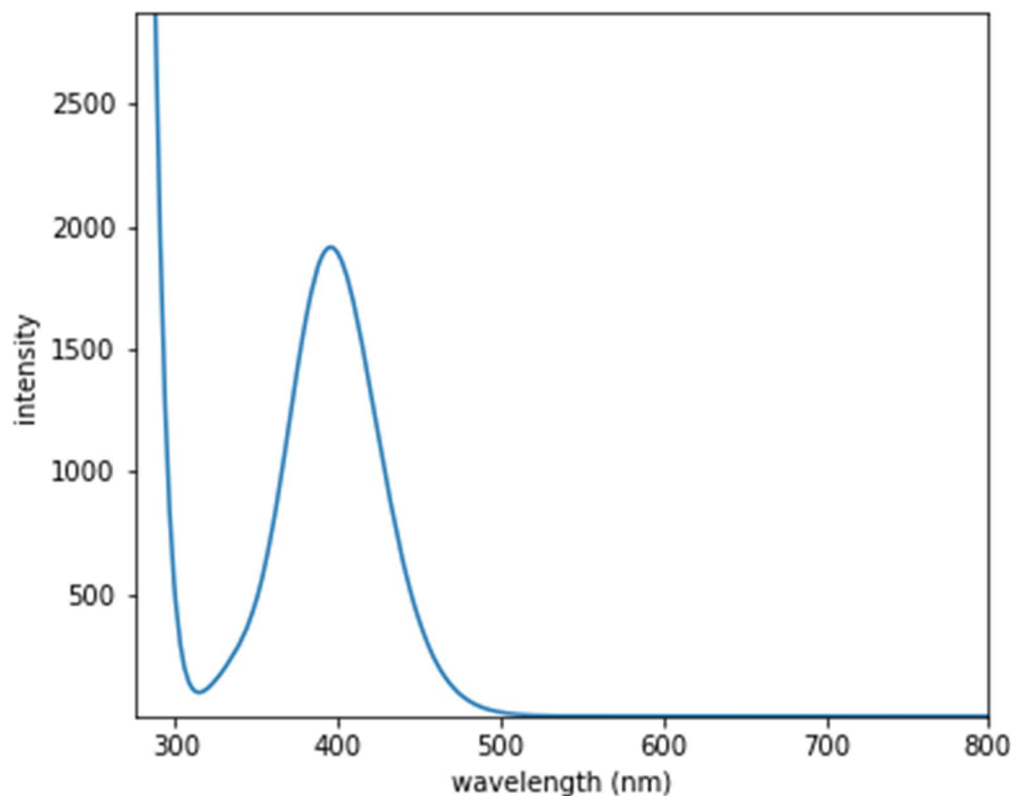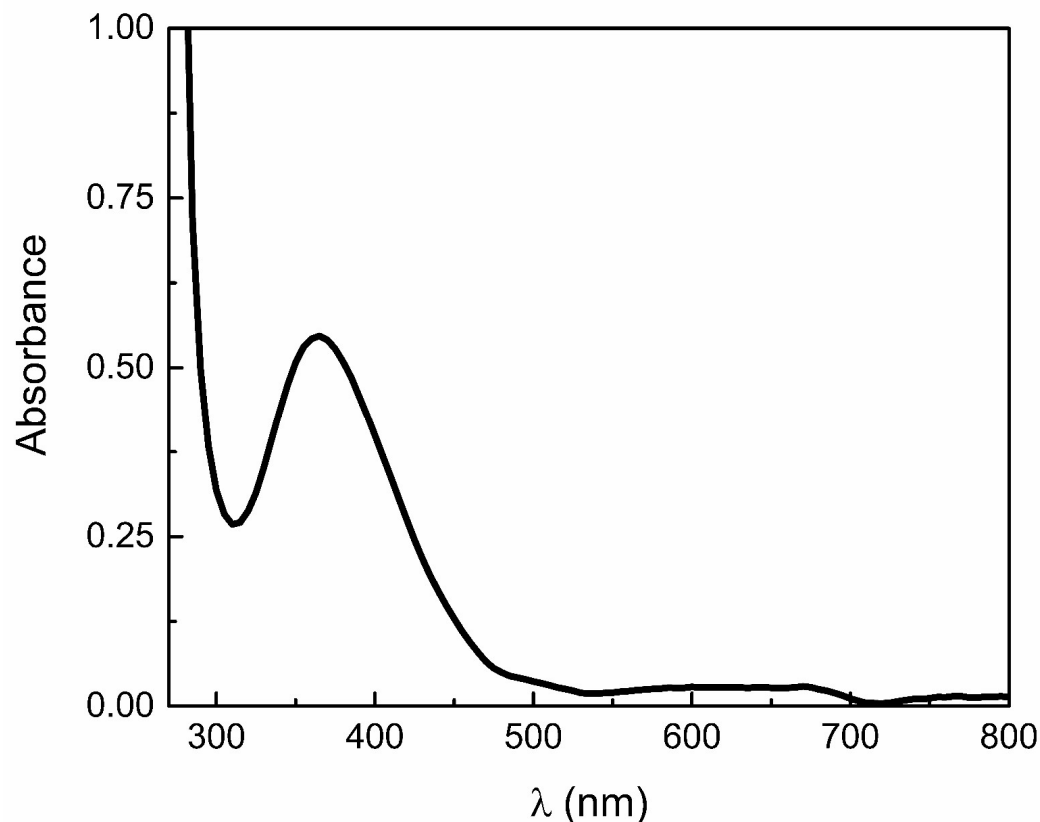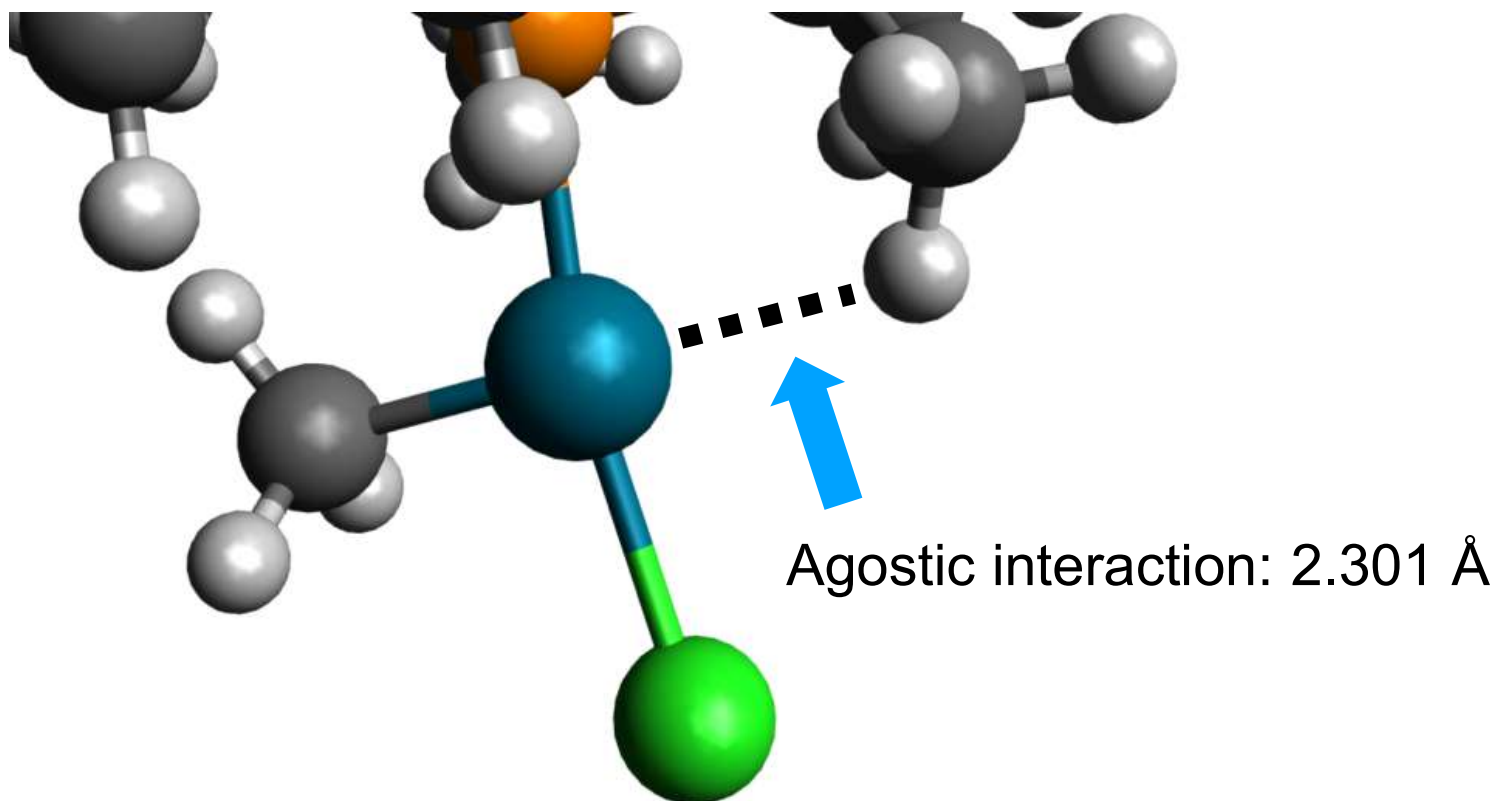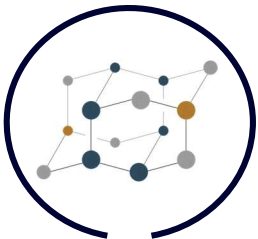
experimental UV-vis spectrum

For TDDFT calculations, agreement with the experimental UV-vis absorption spectrum was used to benchmark performance. Similarly, **B3LYP/def2-TZVPP with CPCM(CHCl$_3$)** was found to give the closest agreement.

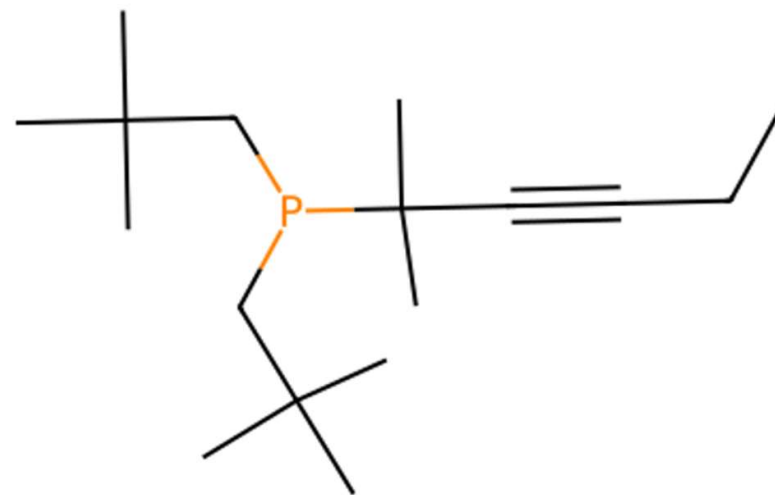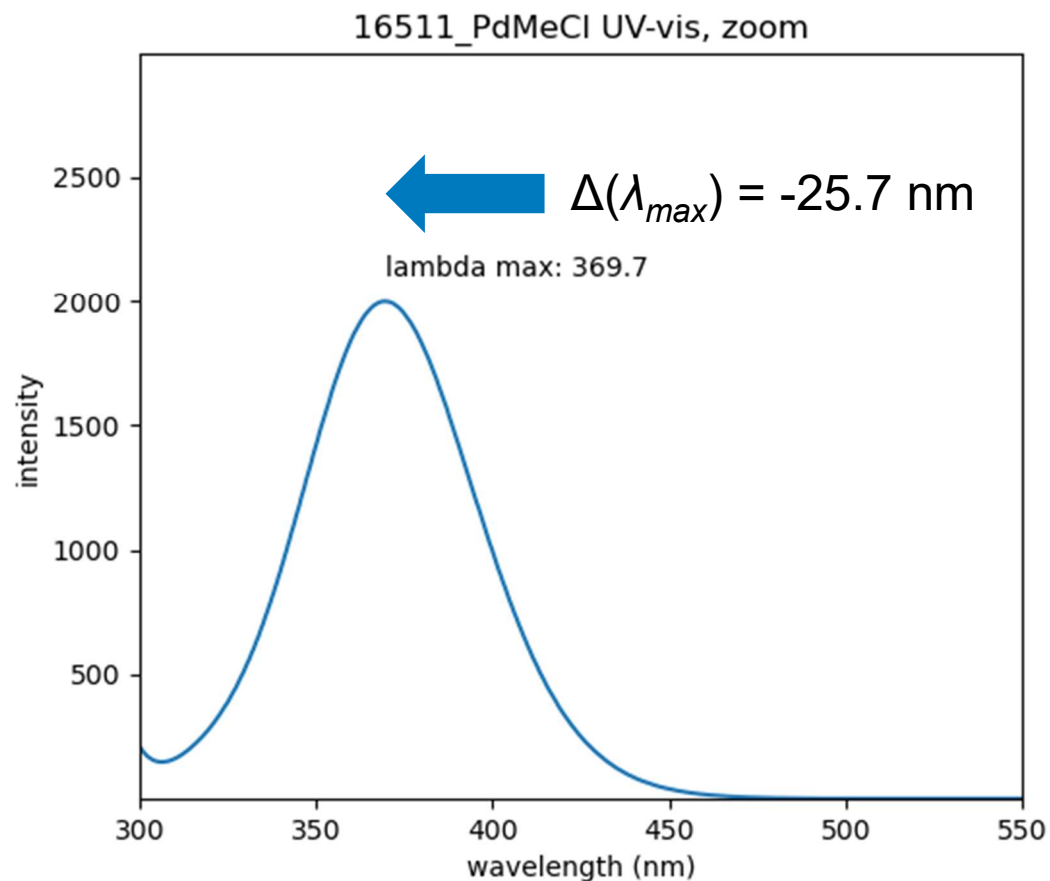Agostic interaction: 2.301 Å

While we refer to these complexes as T-shaped, they often feature an agostic interaction in the fourth site. The distance between Pd and H correlates with the strength of the interaction.

My hypothesis is that the lack of a strong bonding interaction in this site lowers the LUMO (σ^*) energy. Thus, λ_{max} should correlate with agostic interaction strength.

16511_PdMeCl UV-vis, zoom

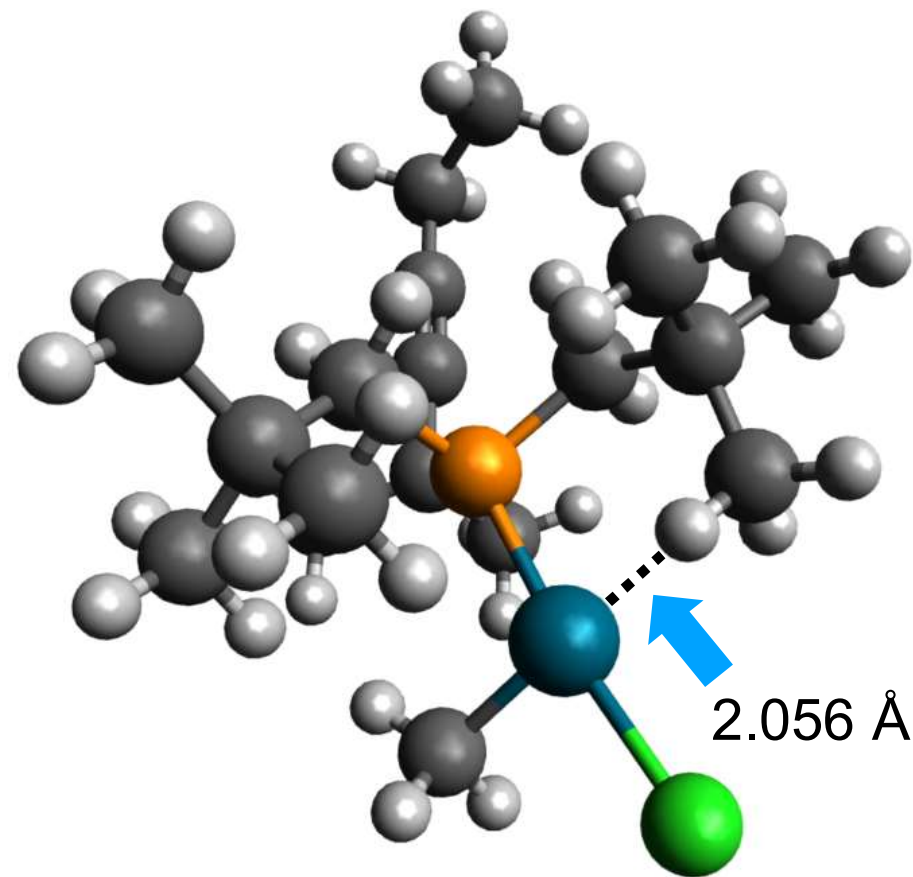$\Delta(\lambda_{max}) = -25.7$ nm

lambda max: 369.7

Ligand 16511 (Kraken designation)

As an example, this ligand features neopentyl groups which allow the H to come in closer proximity to Pd. As a result, the stronger agostic interaction leads to a λ_{max} blueshift of over 25 nm!

16511_PdMeCl UV-vis, zoom

$\Delta(\lambda_{max}) = -25.7$ nm

lambda max: 369.7

2.056 Å
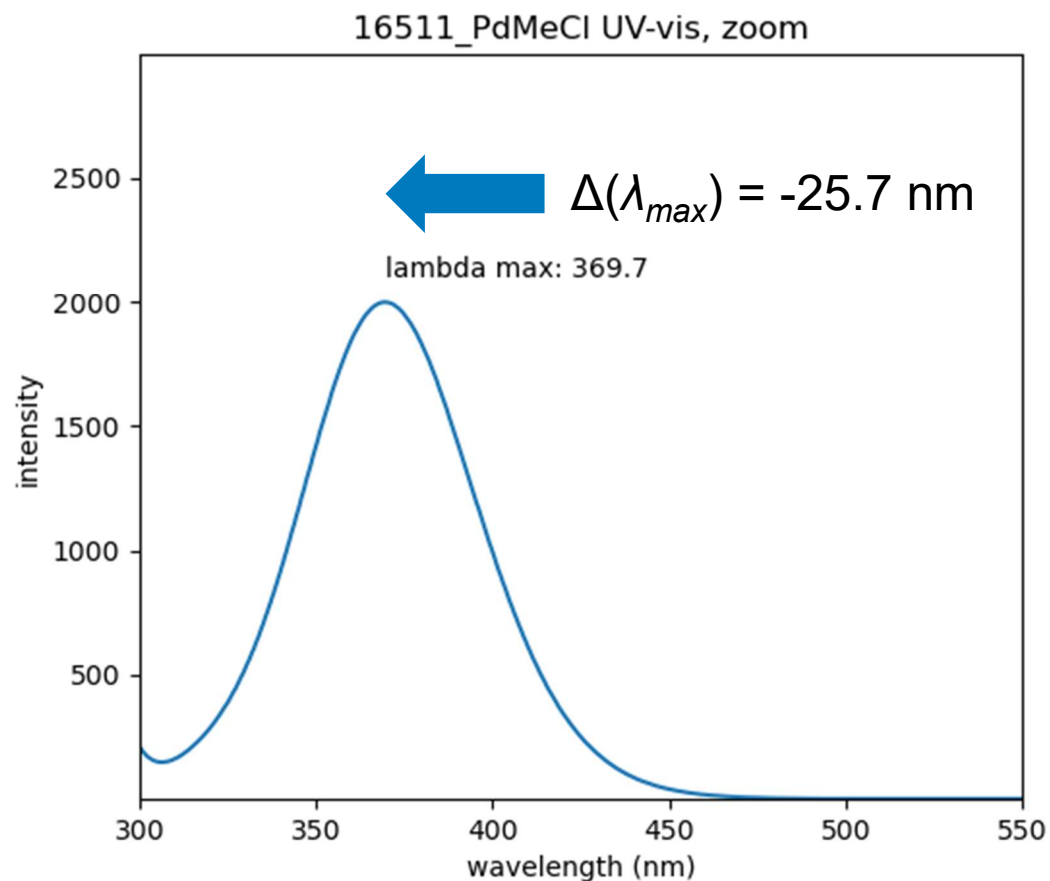
As an example, this ligand features neopentyl groups which allow the H to come in closer proximity to Pd. As a result, the stronger agostic interaction leads to a λ_{max} blueshift of over 25 nm!
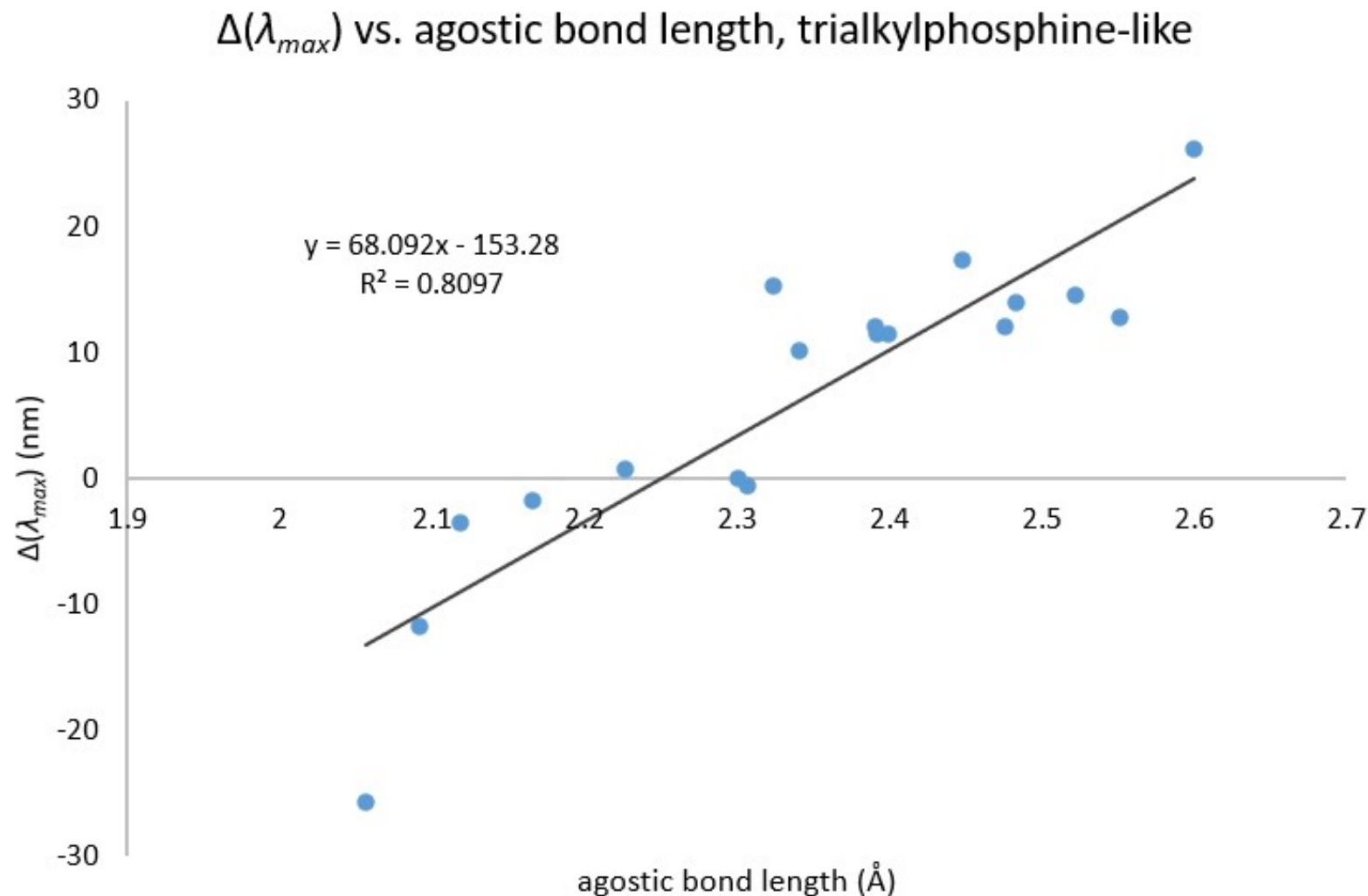
INITIAL OBSERVATIONS: TRENDS IN λMAX, AGOSTIC INTERACTIONS

$\Delta(\lambda_{max})$ vs. agostic bond length, trialkylphosphine-like

$y = 68.092x - 153.28$
$R^2 = 0.8097$

Across a series of trialkylphosphine-like ligands (also including Pd-Si and non-aryl P-Csp^2), the length of the agostic bond correlates well with the change in λ_{max} relative to tri-*t*-butylphosphine (which we will generally refer to as $\Delta(\lambda_{max})$!

$\Delta(\lambda_{max})$ vs. agostic bond length

$y = 73.347x - 168.33$
$R^2 = 0.4429$
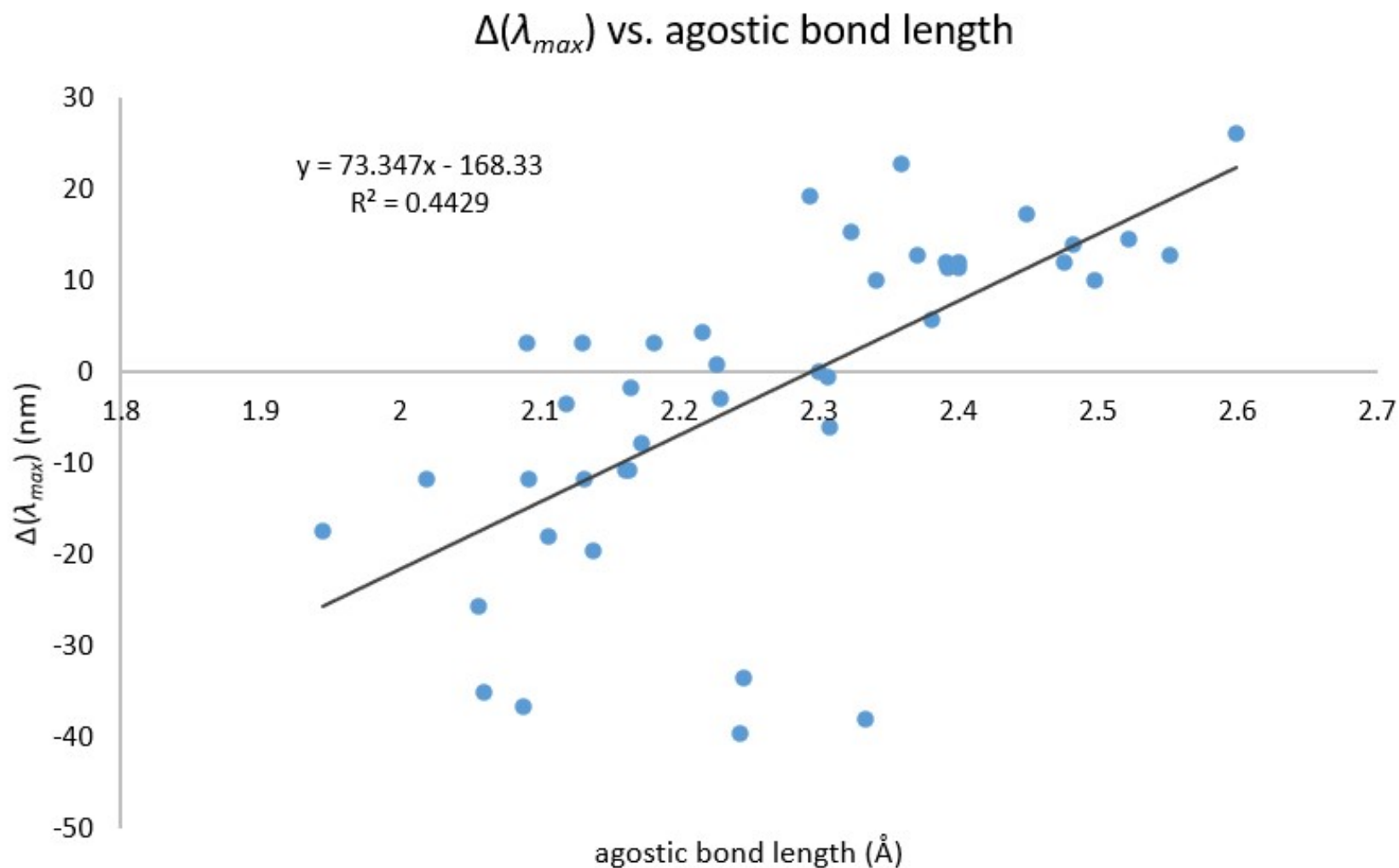
Even across all ligands that apparently have an agostic interaction, the correlation is still pretty good.

$\Delta(\lambda_{max})$ vs. agostic bond length

$y = 73.347x - 168.33$
$R^2 = 0.4429$
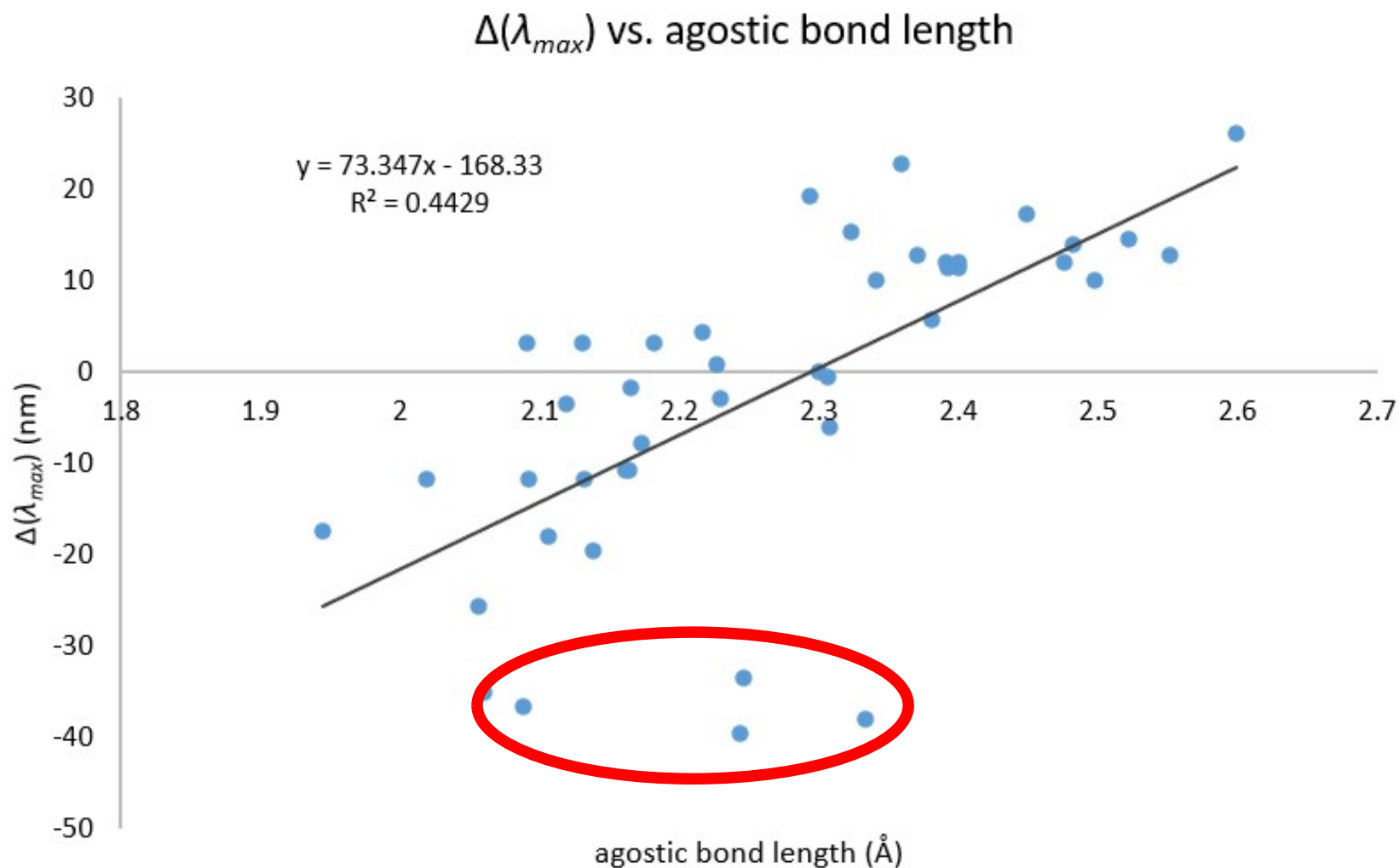
Several examples stand out as being quite badly predicted: can we figure out why?

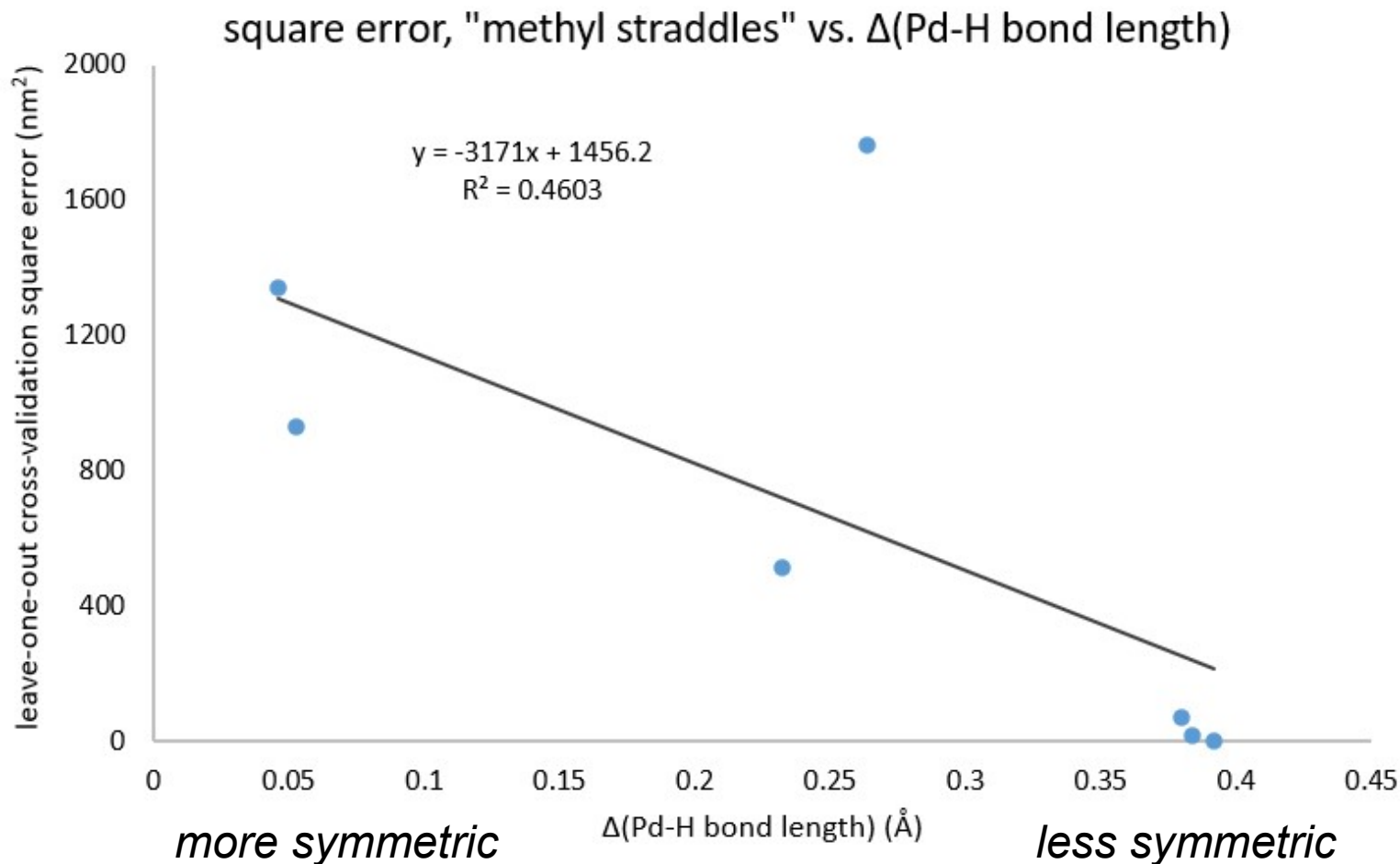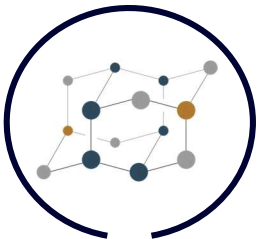**AGOSTIC INTERACTIONS, SYMMETRIC "METHYL STRADDLES"**
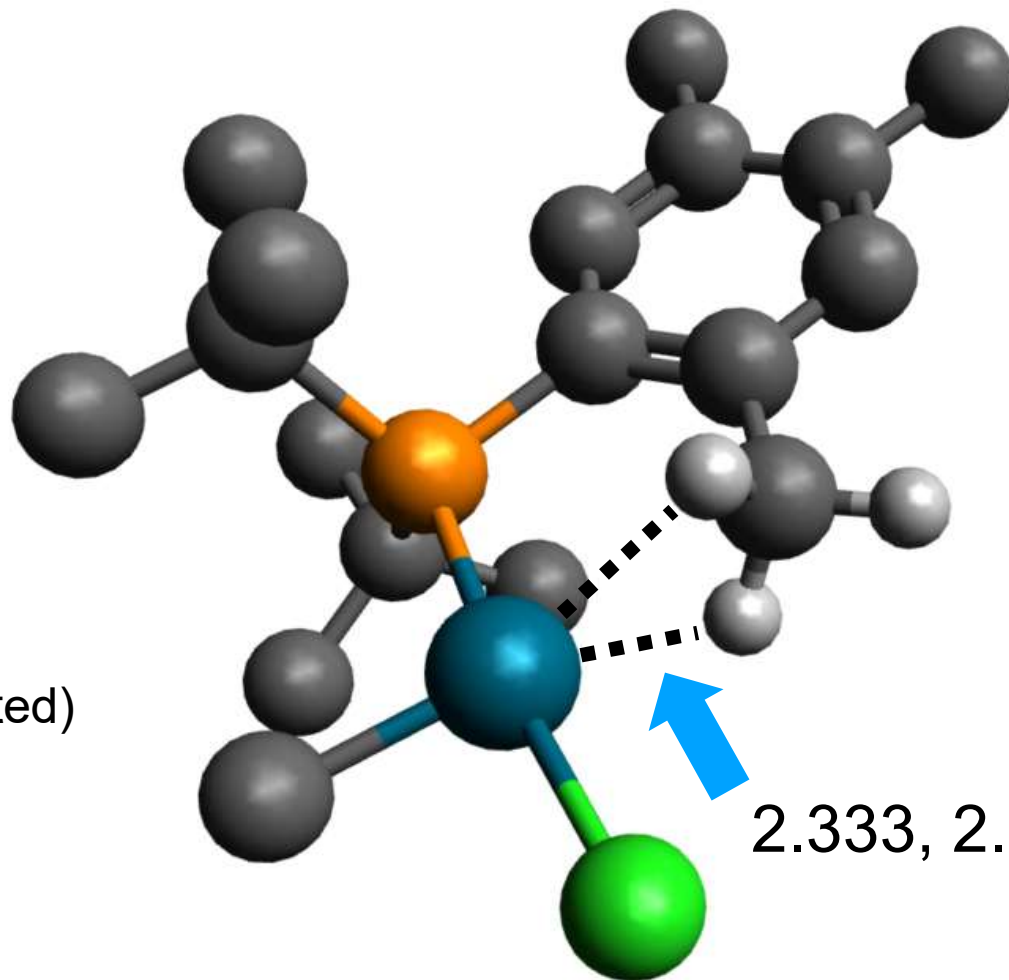
(all other Hs omitted)

2.333, 2.070 Å

Several of these ligands feature a methyl group which appears to almost symmetrically "straddle" the metal center. I suspect that, in these cases, both Hs engage in an agostic interaction, making it stronger than would be predicted by this kind of model, and thus leading to a larger blueshift in λ_{max}.

AGOSTIC INTERACTIONS, SYMMETRIC "METHYL STRADDLES"

square error, "methyl straddles" vs. Δ(Pd-H bond length)

$y = -3171x + 1456.2$
$R^2 = 0.4603$

more symmetric   Δ(Pd-H bond length) (Å)   *less symmetric*

(y-axis: leave-one-out cross-validation square error (nm²))

Indeed, the square error from leave-one-out cross-validation with this simple linear model using only the *minimum* Pd-H distance correlates well with Δ(Pd-H) between the methyl group's two closest Hs (lower values mean more symmetric straddling). So, these cases with symmetric straddles are not very well predicted, while more asymmetric straddles are more akin to the other agostic interactions in the dataset.
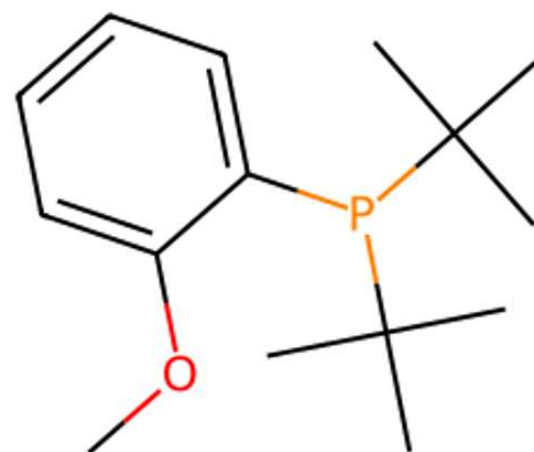
$\Delta(\lambda_{max})$ vs. agostic bond length, no symmetric "methyl straddles"

$y = 70.337x - 158.21$
$R^2 = 0.6781$
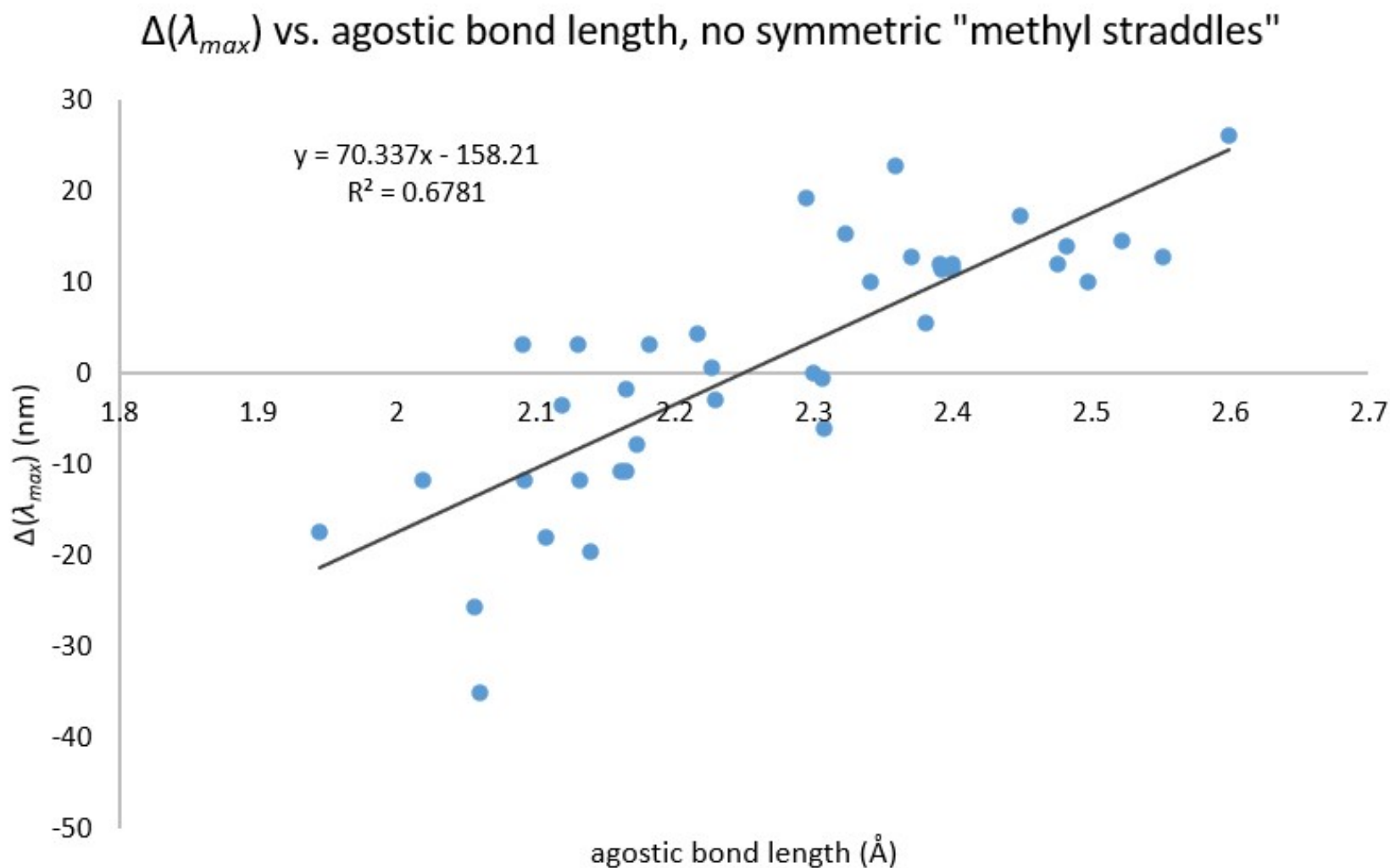
If these cases with symmetric "methyl straddles" (Δ(Pd-H) < 0.3 Å between both Hs of the straddling methyl group) are removed, the R^2 improves significantly of course.
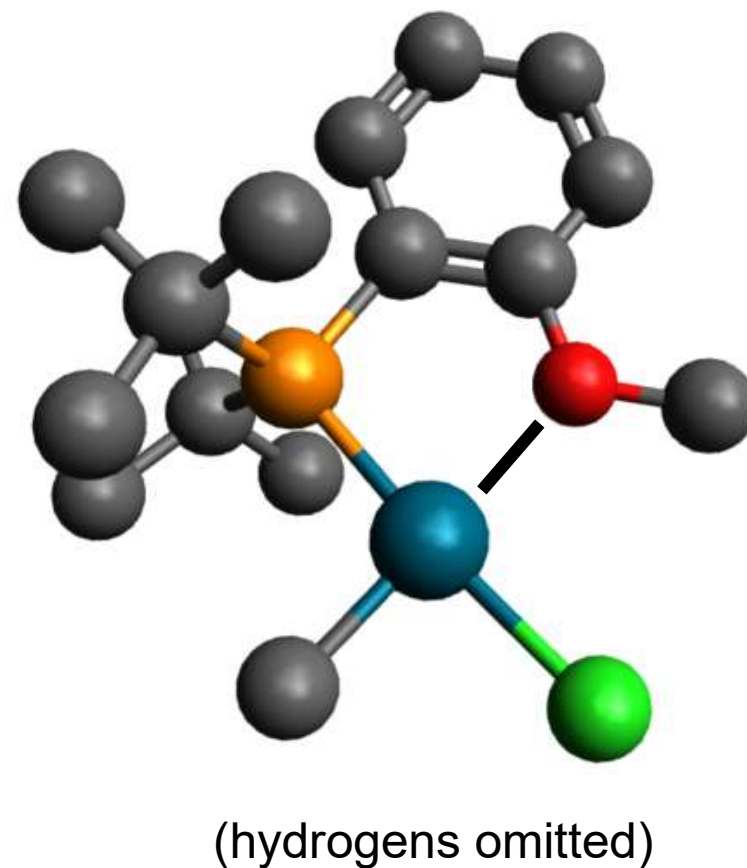
INITIAL OBSERVATIONS: TRENDS IN λMAX, CHELATION

1136_PdMeCl UV-vis, zoom

$\Delta(\lambda_{max})$ = -52.9 nm

lambda max: 342.5

intensity

wavelength (nm)
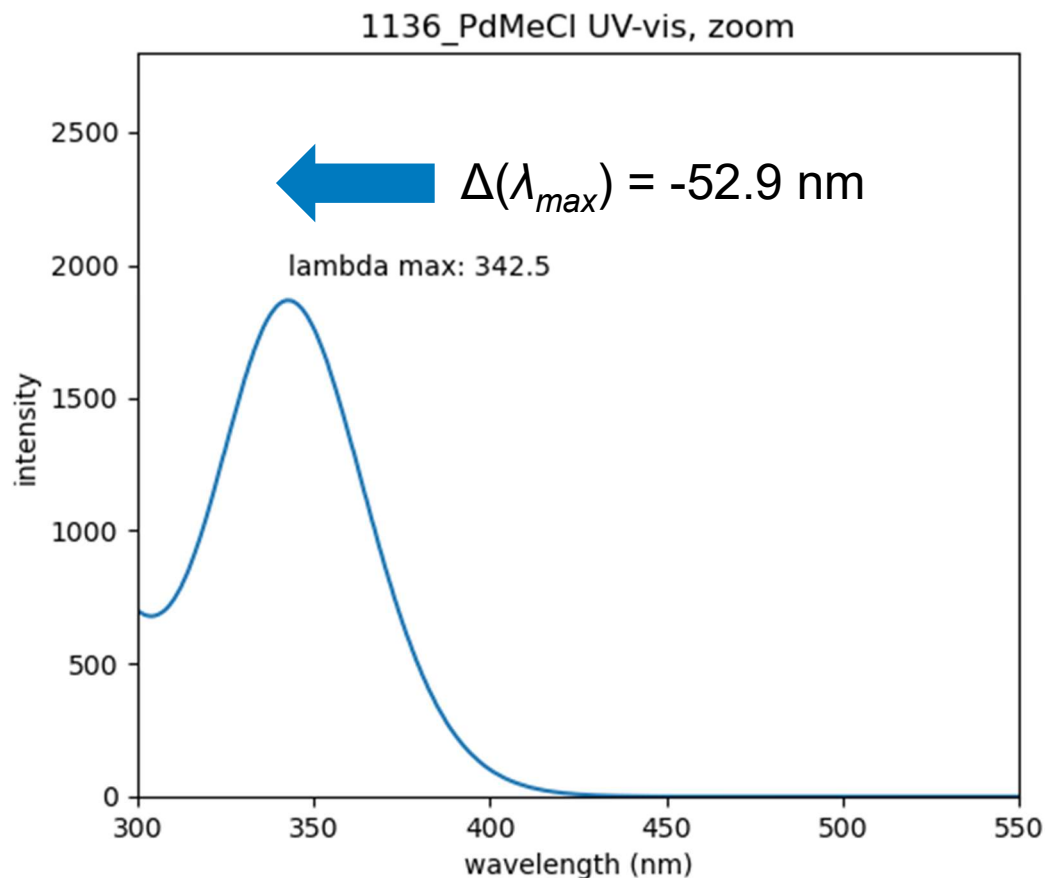
Ligand 1136

Furthermore, for ligands which have pendant heteroatoms that allow chelation, very large λ_{max} blueshifts are observed, again consistent with my hypothesis. Indeed, these complexes are in fact square planar rather than T-shaped.

1136_PdMeCl UV-vis, zoom

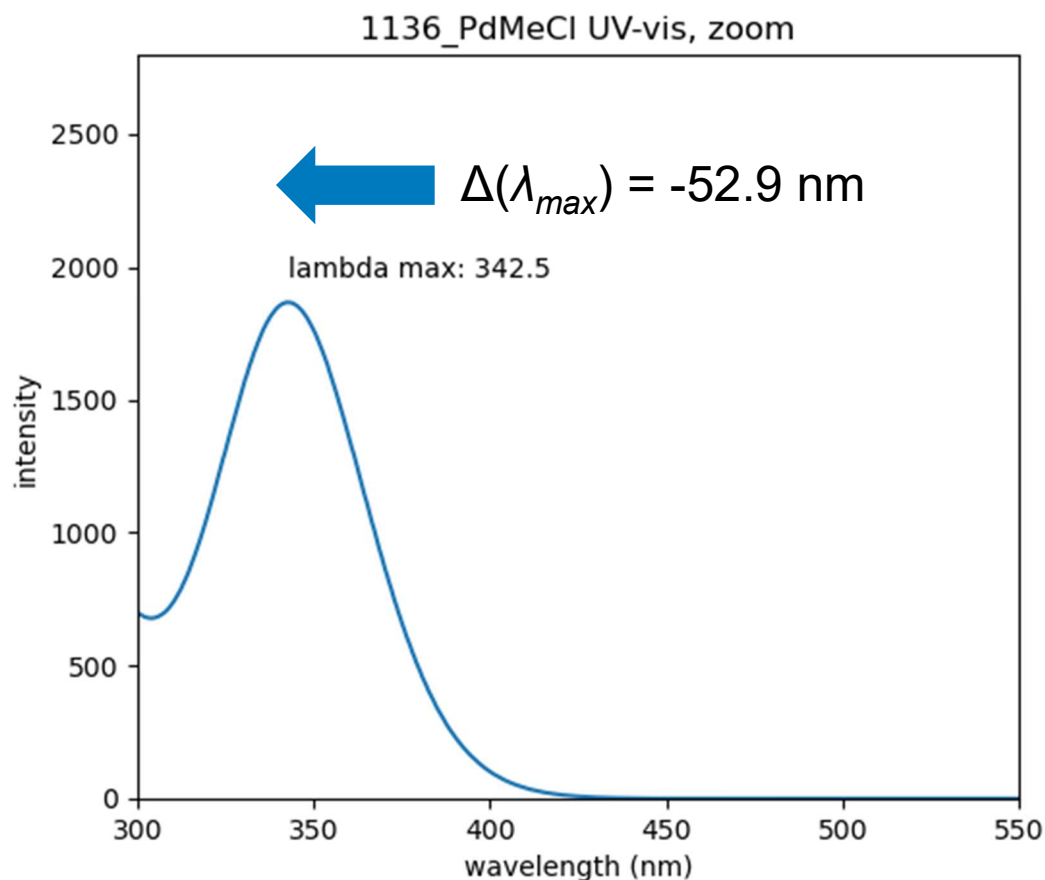$\Delta(\lambda_{max}) = -52.9$ nm
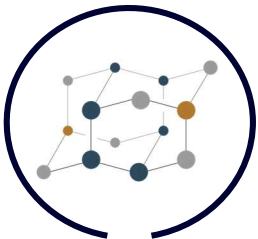
lambda max: 342.5

(hydrogens omitted)

Furthermore, for ligands which have pendant heteroatoms that allow chelation, very large λ_{max} blueshifts are observed, again consistent with my hypothesis. Indeed, these complexes are in fact square planar rather than T-shaped.
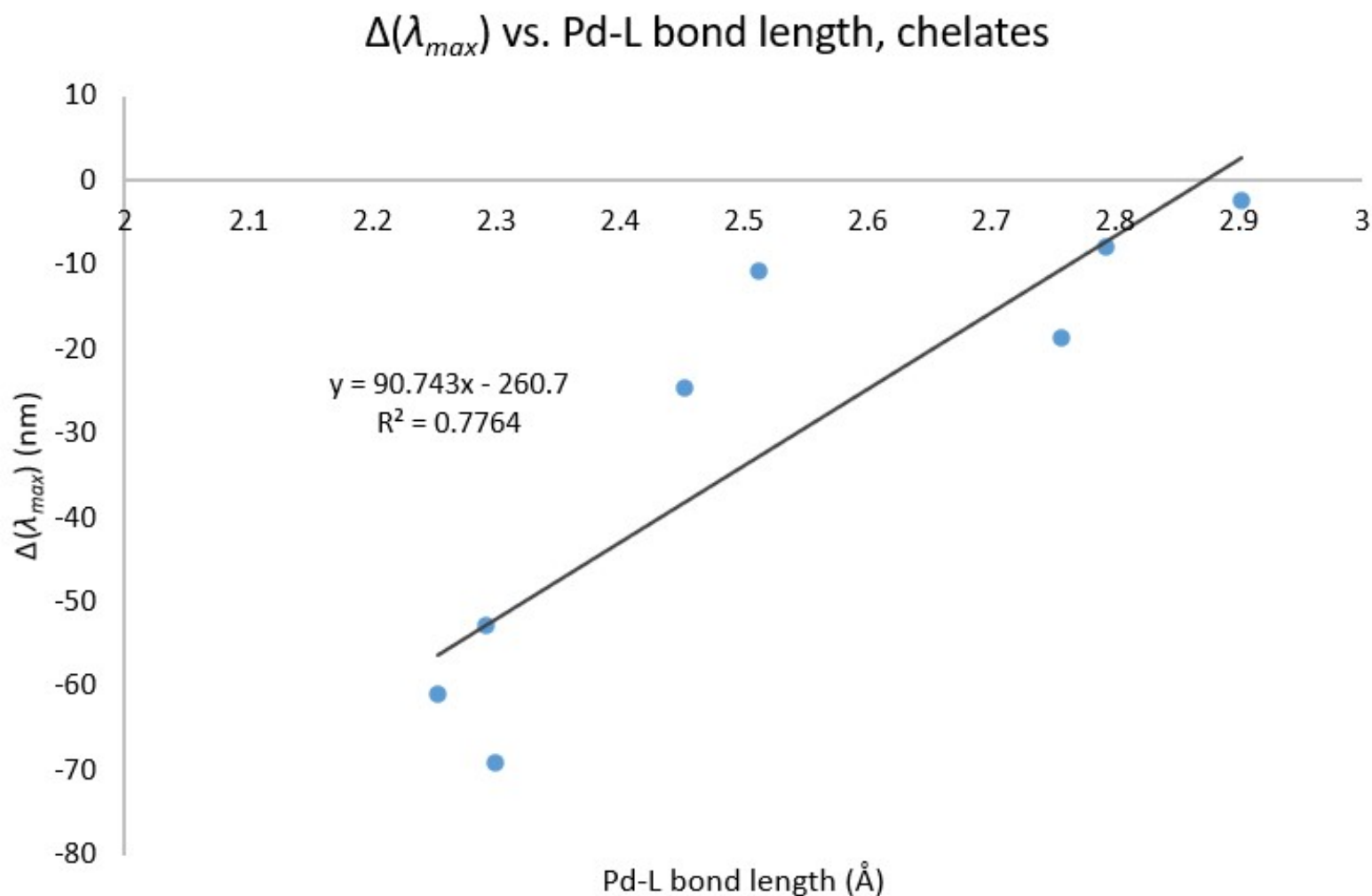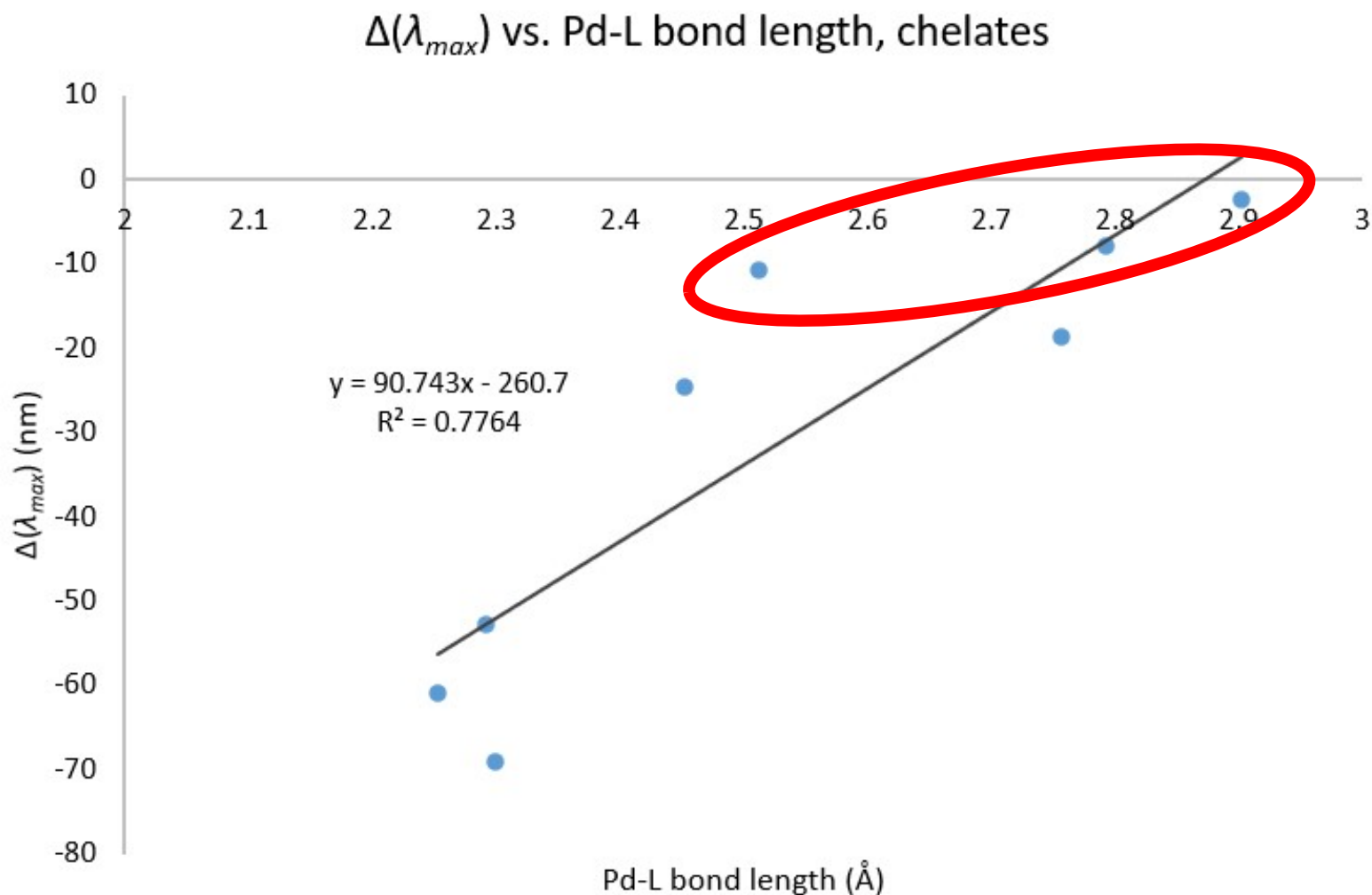
INITIAL OBSERVATIONS: TRENDS IN λMAX, CHELATION

$\Delta(\lambda_{max})$ vs. Pd-L bond length, chelates

$y = 90.743x - 260.7$
$R^2 = 0.7764$

Pd-L bond length (Å)

$\Delta(\lambda_{max})$ (nm)

For ligands that form chelates with pendant L-type donors, the distance between Pd and the pendant donor atom correlates pretty well with λ_{max}, even across different types of donors (N, O, S, Cl).
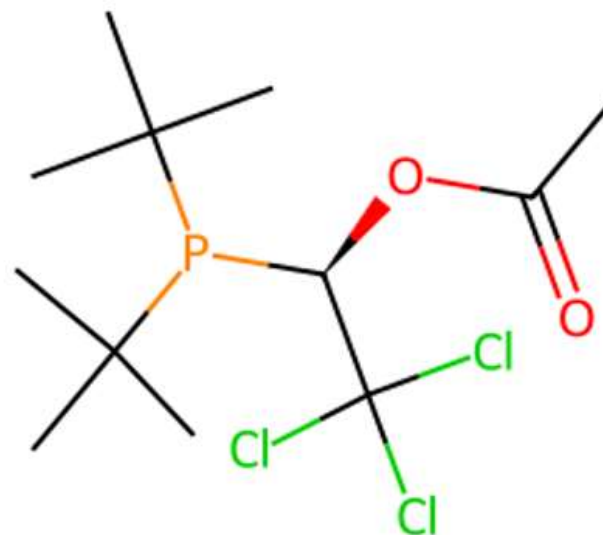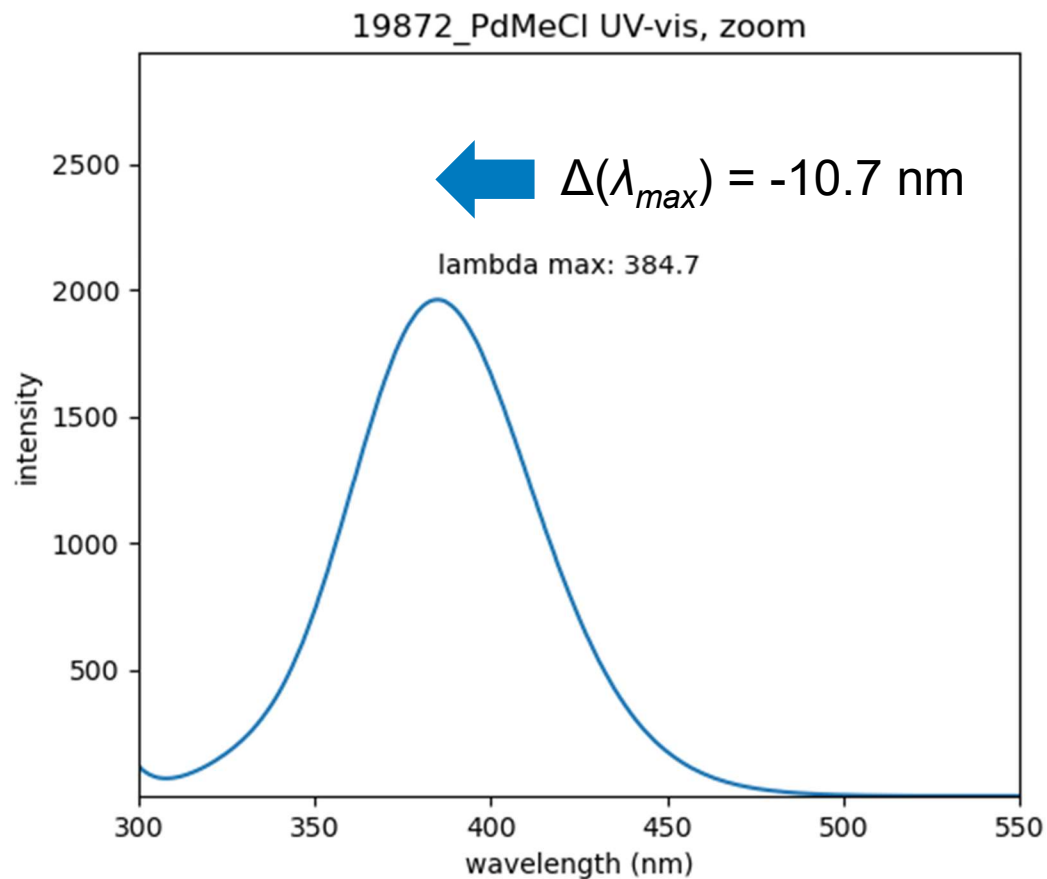
INITIAL OBSERVATIONS: TRENDS IN λMAX, CHELATION

$\Delta(\lambda_{max})$ vs. Pd-L bond length, chelates

$y = 90.743x - 260.7$

$R^2 = 0.7764$

Pd-L bond length (Å)

Interestingly, several of the apparently weaker chelates show a minimal blueshift in λ_{max}! This suggests that Pd-C photocleavage could be accessible with visible light from *square planar complexes*!
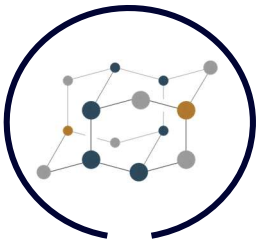
INITIAL OBSERVATIONS: TRENDS IN λMAX, CHELATION

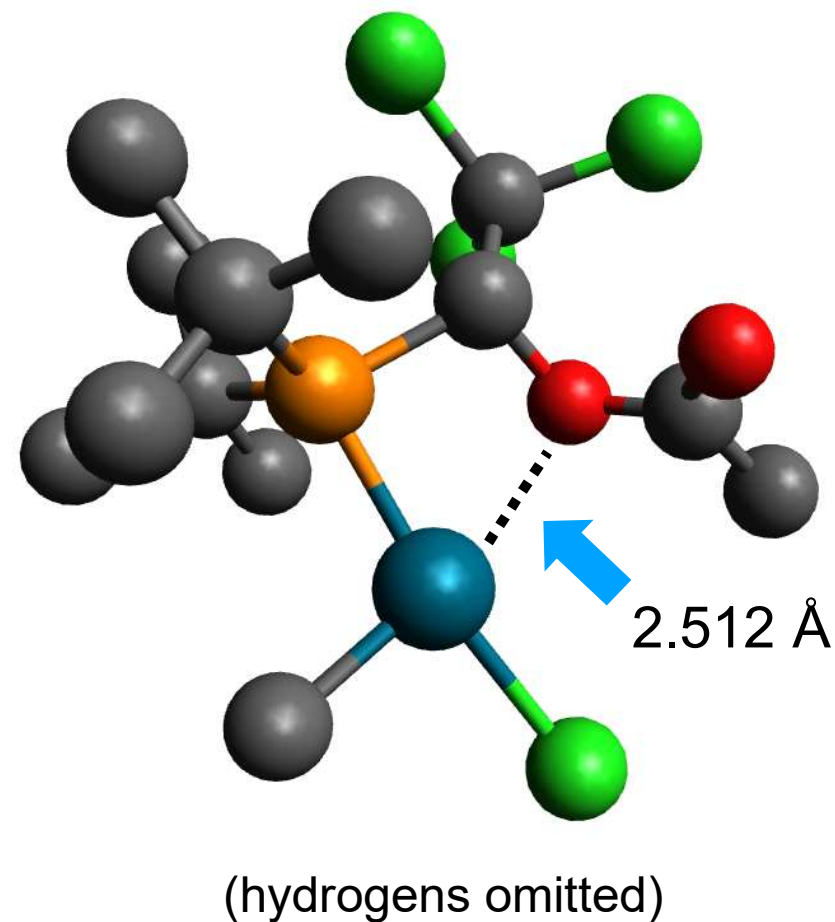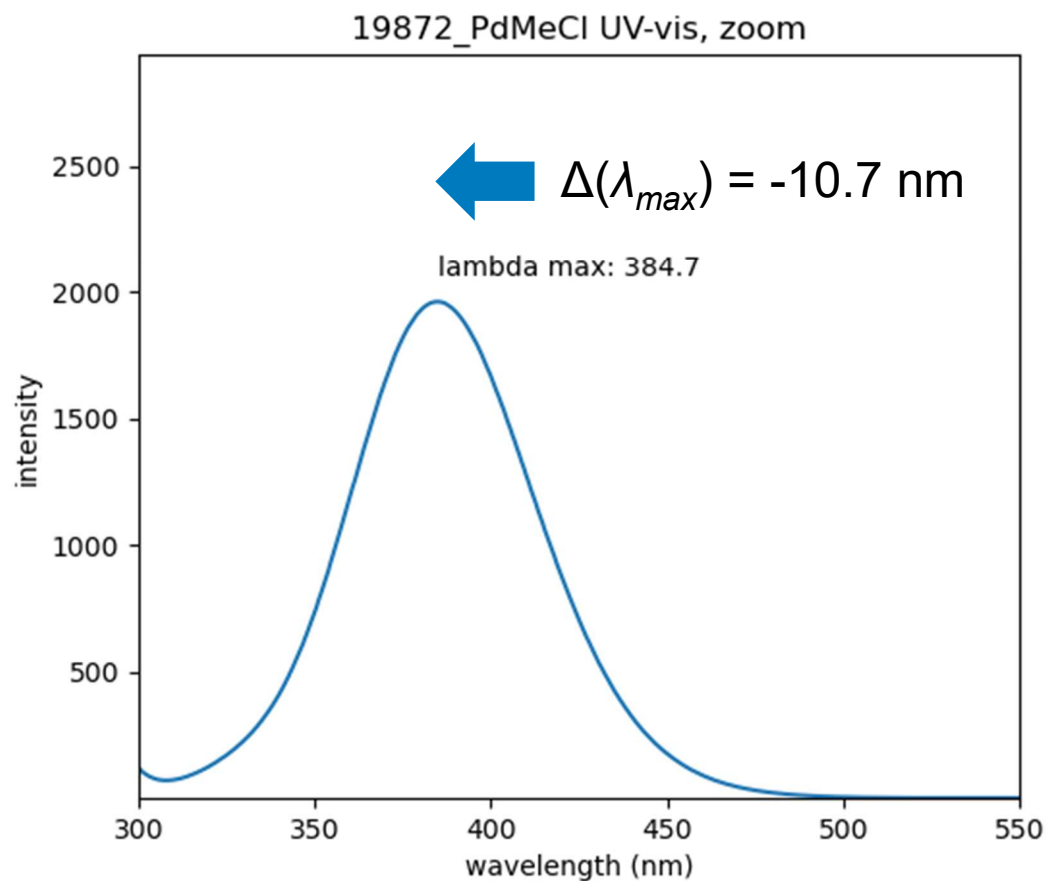19872_PdMeCl UV-vis, zoom

$\Delta(\lambda_{max})$ = -10.7 nm

lambda max: 384.7

Ligand 19872

For example, consider this complex with a pendant ester has a λ_{max} which is blueshifted by only about 11 nm compared to tri-*t*-butylphosphine.

19872_PdMeCl UV-vis, zoom

$\Delta(\lambda_{max}) = -10.7$ nm

lambda max: 384.7
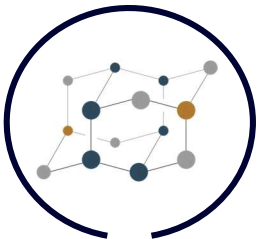
2.512 Å

(hydrogens omitted)

For example, consider this complex with a pendant ester has a λ_{max} which is blueshifted by only about 11 nm compared to tri-*t*-butylphosphine.
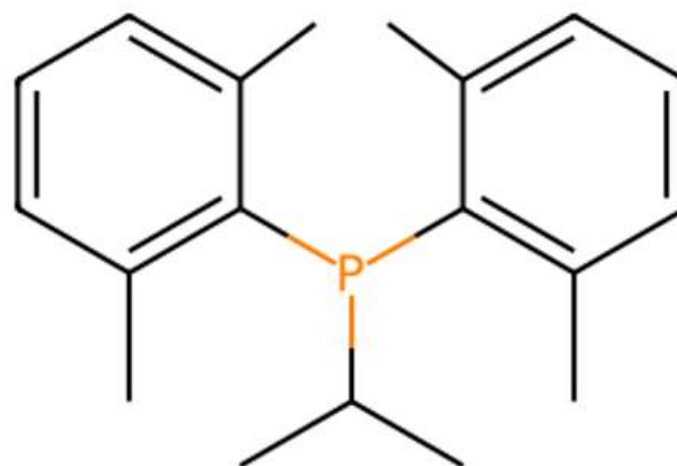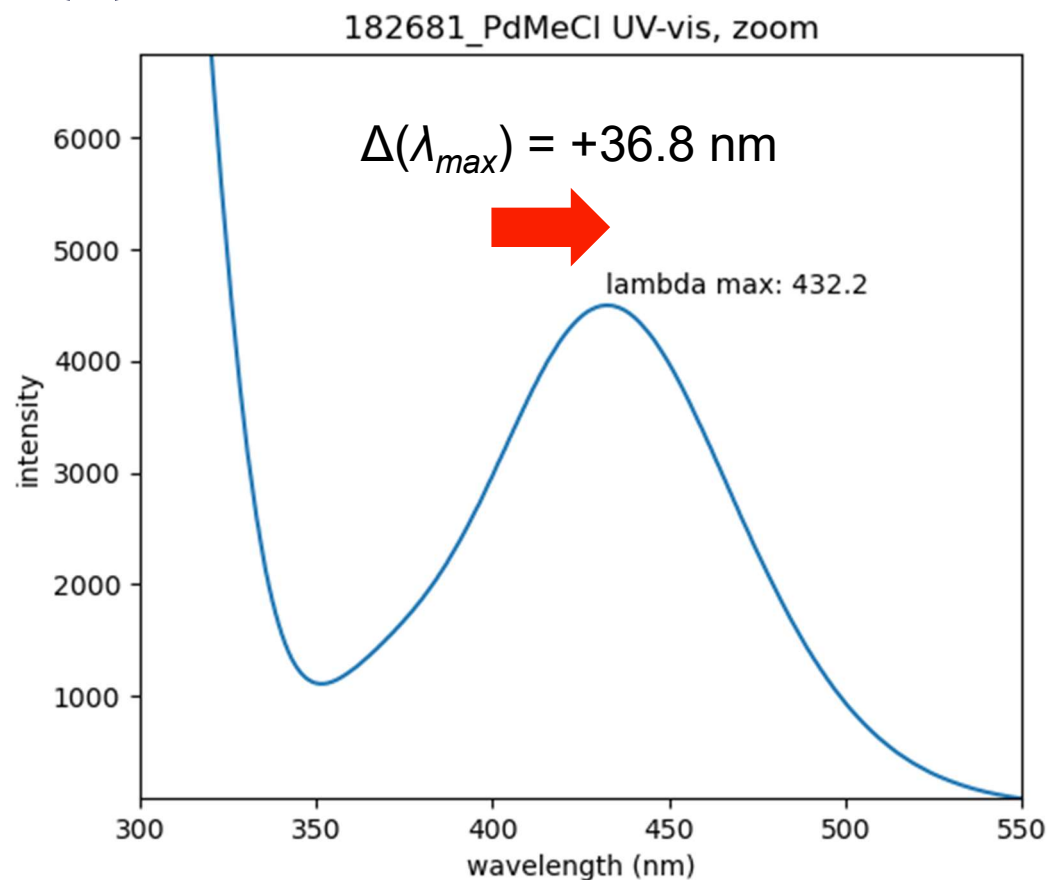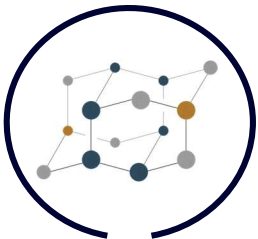
INITIAL OBSERVATIONS: TRENDS IN λMAX, OPEN SITE

182681_PdMeCl UV-vis, zoom

$\Delta(\lambda_{max}) = +36.8$ nm

lambda max: 432.2

Ligand 182681

In some cases, the favored ligand conformation appears to leave the fourth site open entirely. In this limiting case, λ_{max} tends to be around 420-430 nm. However, in the condensed phase, this site would be open to occupation by solvent or e.g. bridging Cl ligand, so they may not be a good choice in practice for triggering this photochemistry with low energy light.

INITIAL OBSERVATIONS: TRENDS IN λMAX, OPEN SITE

182681_PdMeCl UV-vis, zoom

$\Delta(\lambda_{max}) = +36.8$ nm

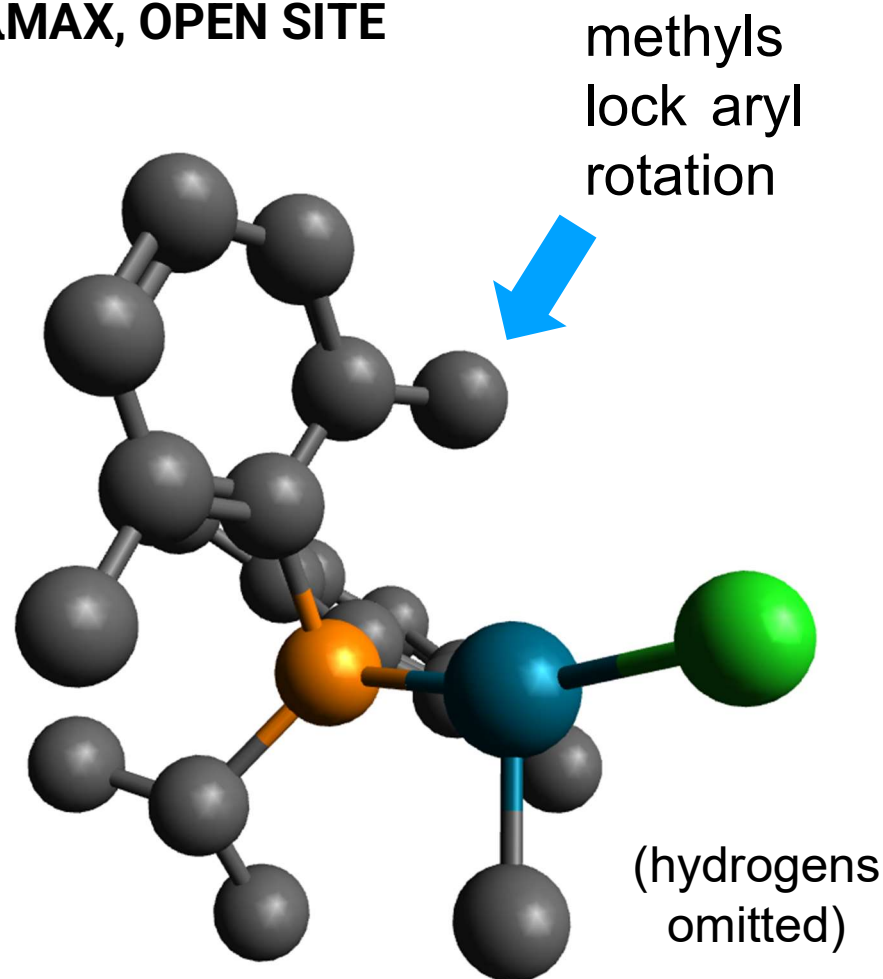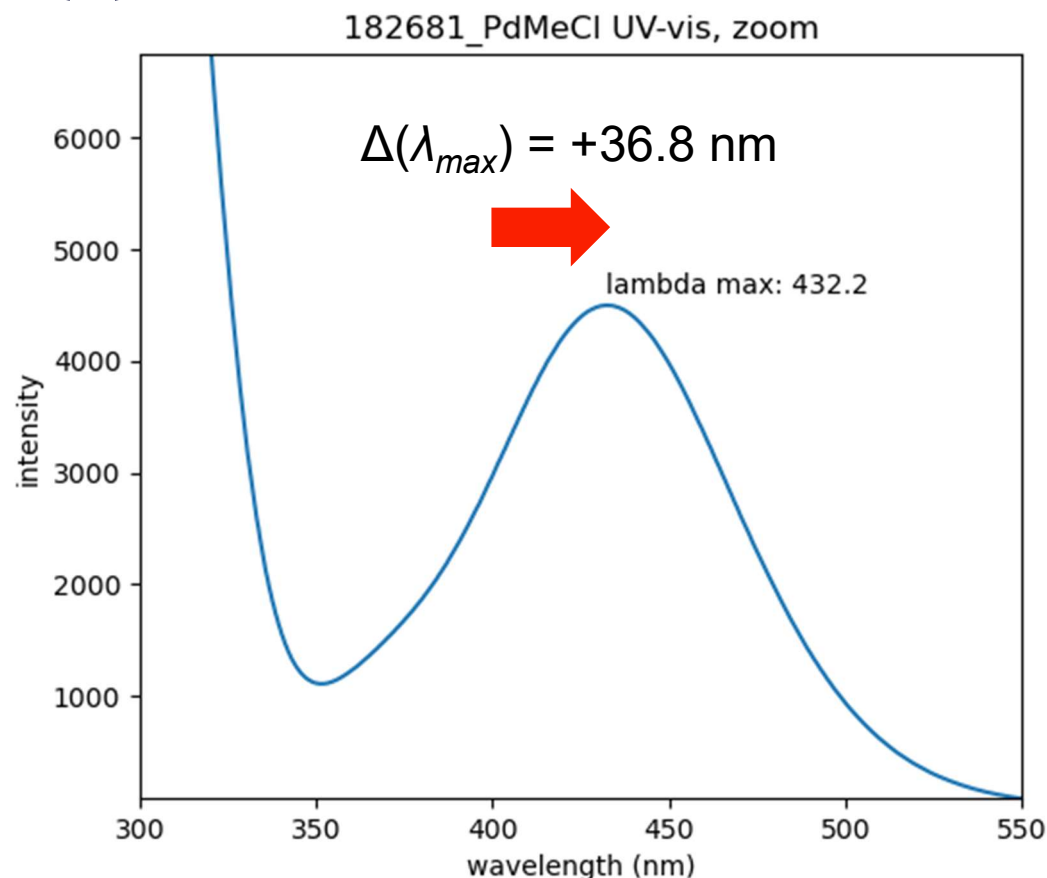lambda max: 432.2

methyls lock aryl rotation

(hydrogens omitted)

In some cases, the favored ligand conformation appears to leave the fourth site open entirely. In this limiting case, λ_{max} tends to be around 420-450 nm. However, in the condensed phase, this site would be open to occupation by solvent or e.g. bridging Cl ligand, so they may not be a good choice in practice for triggering this photochemistry with low energy light.
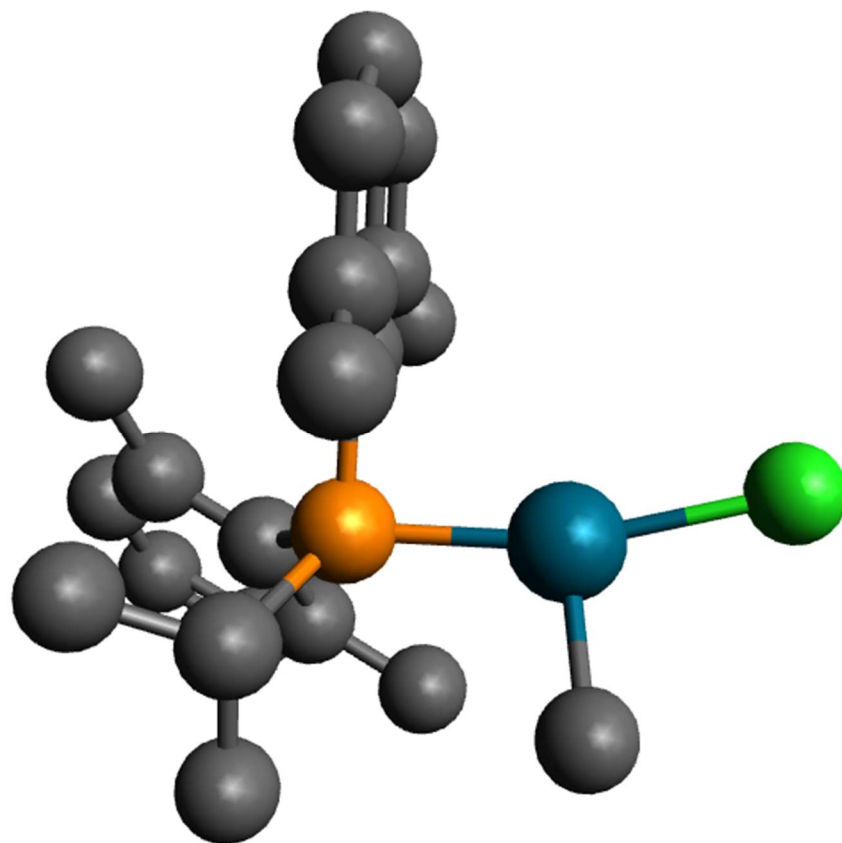
182681_PdMeCl UV-vis, zoom

$\Delta(\lambda_{max}) = +36.8$ nm

lambda max: 432.2

In some cases, the favored ligand conformation appears to leave the fourth site open entirely. In this limiting case, λ_{max} tends to be around 420-430 nm. However, in the condensed phase, this site would be open to occupation by solvent or e.g. bridging Cl ligand, so they may not be a good choice in practice for triggering this photochemistry with low energy light.
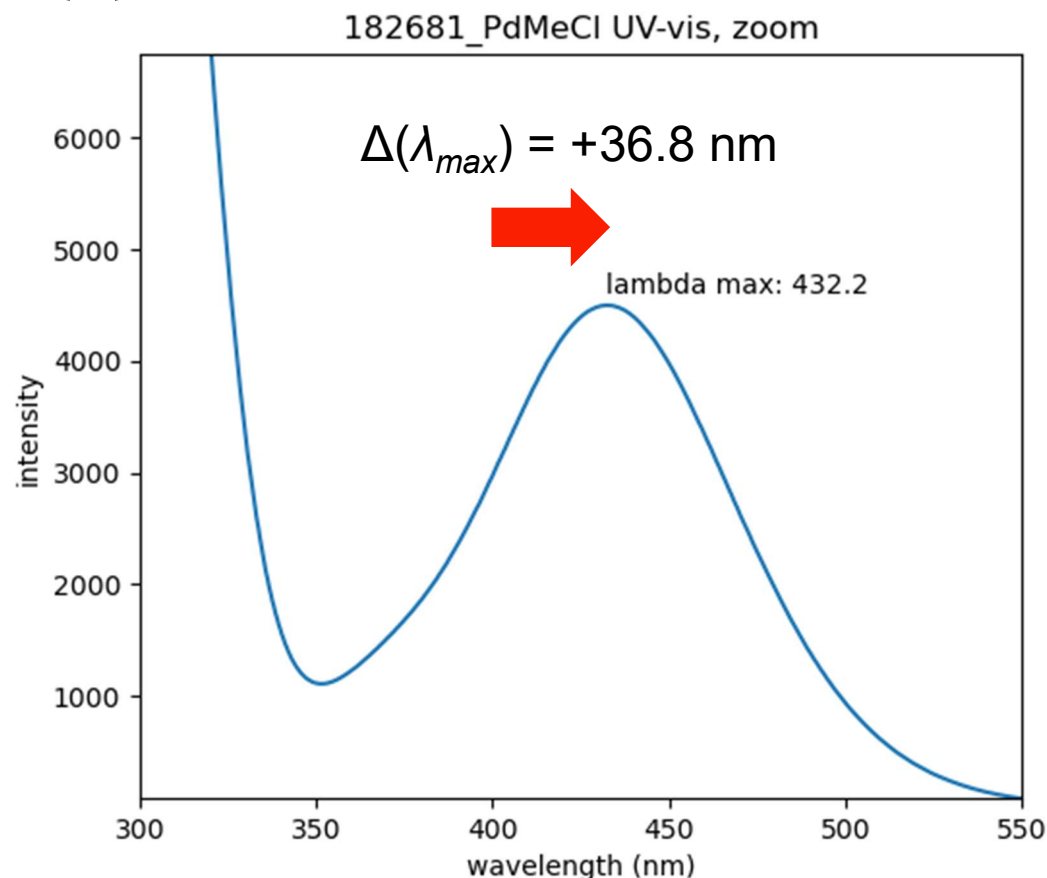
So, my hypothesis that the strength of the bonding interaction in the fourth site strongly influences the position of the peak in the absorption spectrum corresponding to the HOMO-LUMO electronic transition (and thus, the wavelength of light that can be used to cause Pd-C homolytic cleavage) seems to be very well validated.

Now, what other factors can we find that influence λ_{max}?
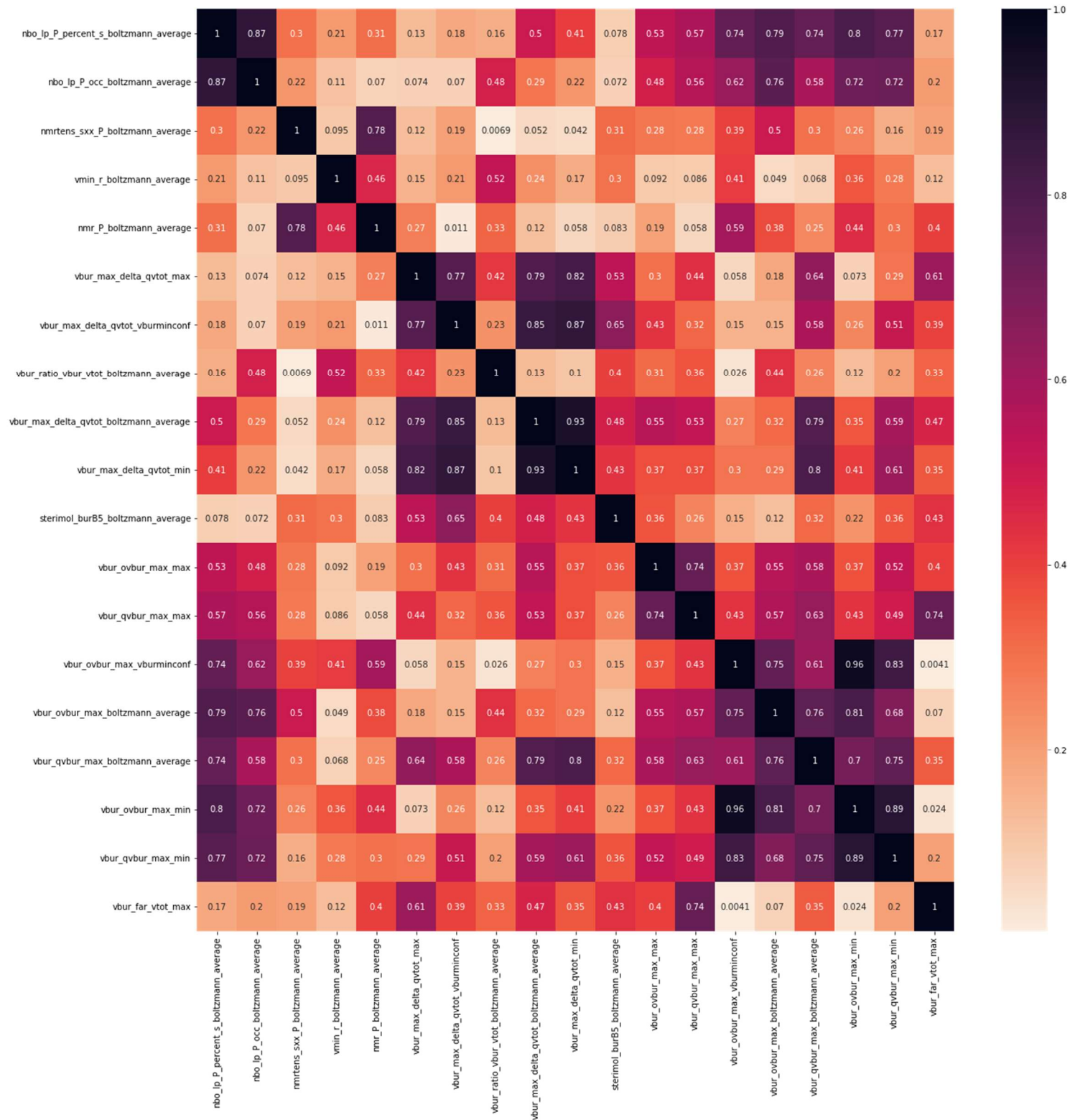
Comments on electronic effects, P-C vs. P-N vs. P-O, aryl, etc.

MACHINE LEARNING MODELS TO PREDICT λMAX: ELASTIC NET

- Our case is highly dimensional: each ligand has 192 features, while we have at most around 100 data points.

- To begin with, let us restrict ourselves to the top 20 features from our principal component analysis from earlier.

- Furthermore, there is significant multicollinearity in our feature set (see next slide)

- **Results to come!!**

For simplicity, the absolute correlation is shown here, with darker meaning more highly correlated.

To be continued…