Mapping the properties of bidentate ligands with calculated descriptors (LKB-bid)

Natalie Fey,*[‡] Alexander Koumi,[‡] Andrei V. Malkov,[§] Jonathan D. Moseley,[$] Bao N. Nguyen,[v] Simon N. G. Tyler,[$] Charlotte E. Willans[v]

Electronic Supporting Information

Table of Contents

Computational Details

Density functional theory calculations were performed in Jaguar[1] and used the standard Becke-Perdew (BP86) density functional.[2] The triple-zeta form of the standard Los Alamos ECP basis set (LACV3P) as implemented in Jaguar was used on the transition metal atoms, employing the 6-31G* basis for all other atoms. "Loose" convergence (five times larger than default criteria) was used for optimizations. Test calculations using default criteria did not lead to significant changes in geometries, but were more time-consuming. Calculations were performed on isolated molecules and NBO atomic charges were calculated.[3] Vibrational frequencies were not computed; ligand effects on zero-point energy corrections are likely to be small and have been neglected. Stationary points have not been verified as minima but accidental optimization to transition states seems unlikely for complexes of low symmetry.

Molecular mechanics (MM) conformational searches used the default MMX force field in PCModel.[4] GMMX was used for stochastic conformational searches with default settings unless stated otherwise. 500 Iteration conformational searches were performed on free ligands and ligands bound to the zinc fragment (stop criteria defined as Emin found 10 times and duplicates found 50 times).

Principal component analyses were performed in SIMCA-P+.[5]

References

(1) Schrödinger-LLC, *Jaguar 6.0,* New York, NY, 2005; Schrödinger-LLC, *Jaguar 7.6,* New York, NY, 2009.
(2) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098-3100; Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822-8824; Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406; Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048-5079; Slater, J. C. *Quantum Theory of Molecules and Solids, Vol. 4: The Self-Consistent Field for Molecules and Solids*; McGraw-Hill: New York, 1974.
(3) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Weinhold, F., *NBO 5.0* Madison, 2001.
(4) Gilbert, K., *PCModel* Bloomington, IN, 2004.
(5) Umetrics, *SIMCA P+ 12.0.1.0*, 2009.

All new ligand data, as well as principal component scores and loadings (PCs 1-4), can be found in a separate Excel table.

Description of Computational Workflow

Conformational searches were carried out for free ligands and ligands coordinated to the [$ZnCl_2$] fragments to ensure that substituents adopted sensible conformers. The lowest energy conformer from the molecular mechanics searches was then used as input for DFT calculations. We readily acknowledge that the potential energy surfaces will be different between different levels of theory and that this may introduce some conformational noise. This has been discussed further in prior work (*Organometallics* **2008**, *27*, 1372-1383; *Dalton Trans.* **2013**, *42*, 172-181), and represents a reasonable compromise for such databases. The DFT-optimised free ligand geometry was then adapted as input for coordination of the [AuCl] fragment to each donor atom, at times with minor structural adjustment to accommodate the metal complex, while the DFT-optimised [$ZnCl_2(D_1{\sim}D_2)$] complexes were used to generate a ligand template file. All other starting geometries were based on these templates. To generate these templates, the metal centre was changed to a trigonal planar fragment and "M"-donor distances were set to standard values for different donors (2.28 Å (P), 2.00 Å (C, N), 2.05 Å (O)). These fragments were aligned to a standard orientation and then combined with template files for the [$PdCl_2$] complex and both He_8 adducts. This alignment can be performed in Jaguar through conversion into z-matrix and back to Cartesian coordinates; many other codes for viewing multiple formats of molecular structures will perform a similar alignment as part of this conversion.

Fully optimised ligand and Zn complex geometries for all ligands have been included as xyz coordinates as part of this electronic supporting information.

Note that data for the following ligands has been published previously (*Organometallics* **2008**, *27*, 1372-1383; *ibid.* **2012**, *31*, 5302-5306):

pn02, pn03, pn06, pn08, pn10, pn11, pn13, pn23, pn23, pp01-pp16, pp18-pp28.

In addition, sample input files including all calculation settings are listed here for ligand nn05:

Ligand:

```
&gen
dftname=bp86
basis=6-31g*
numd=5
igeopt=1
iaccg=3
ip11=2
&
&zmat
N -1.193525 -0.091804 -0.134134
N 1.207160 -0.001005 0.134962
C -0.024193 0.768197 0.002201
C -2.424363 0.683820 -0.100786
C -1.127179 -0.901788 -1.341820
C 1.202786 -0.818972 1.339077
C 2.375463 0.866088 0.105217
H 0.044272 1.432519 -0.891615
H -0.142816 1.421469 0.898907
H -2.501957 1.241917 0.860116
H -2.475766 1.408561 -0.946306
H -3.307255 0.006466 -0.150999
H -0.217822 -1.544090 -1.334408
H -1.113896 -0.269687 -2.259883
H -2.002458 -1.588961 -1.390358
```

```
H 0.344891 -1.528511 1.328699
H 2.127783 -1.437819 1.384749
H 1.141640 -0.193739 2.259897
H 2.410243 1.432673 -0.853217
H 3.307305 0.257602 0.152610
H 2.371760 1.588928 0.953913
&
```

Zn complex:

```
&gen
isymm=0
iaccg=3
igeopt=1
numd=5
basis=6-31g*
dftname=bp86
ip11=2
&
&zmat
Zn1 0.229719 -1.079768 0.161516
Cl 0.558730 -2.649796 -1.288874
Cl 0.144862 -2.064287 2.085156
N -0.947052 0.408354 -0.241069
N 1.099609 0.642934 -0.030728
C -0.019361 1.526307 -0.321279
C -1.911817 0.545995 0.841351
C -1.624361 0.139326 -1.502111
C 1.771547 0.968168 1.219901
C 2.059002 0.561498 -1.123561
H -0.188866 2.320509 0.442308
H 0.027598 2.014273 -1.322398
H -2.576230 1.425644 0.677374
H -1.388998 0.670152 1.816583
H -2.539927 -0.370580 0.917173
H -2.272304 0.995672 -1.800365
H -2.253924 -0.775195 -1.414446
H -0.882626 -0.046223 -2.311576
H 1.049319 0.949622 2.067176
H 2.241146 1.977793 1.172470
H 2.560595 0.214022 1.441369
H 2.545073 1.547821 -1.305269
H 1.555690 0.233247 -2.060983
H 2.846598 -0.190593 -0.890250
&
&atomic
atom   basis
Zn1    lacv3p
&
&nbo
&
```

Aligned ligand

```
&zmat
X22     0.0000000000000   0.0000000000000   0.0000000000000
H23     0.0000000000000   0.0000000000000   1.8000002365880
N1      1.2186682530345   0.0000000000000  -1.5859954316576
N2     -1.2175919847126  -0.0000000000000  -1.5866536354112
C3      0.0004613130294  -0.3114826026797  -2.3874355859076
C4      2.2436849011447  -1.0674931155787  -1.6985690154401
C5      1.8091775830353   1.3214962824709  -1.9118968755389
```

```
C6     -2.2411781795875  -1.0691018899829  -1.6959755491544
C7     -1.8105253167141   1.3193839691186  -1.9159426237250
H8      0.0015908875490  -1.3926244325052  -2.5958219091444
H9      0.0010008771100   0.2333253690327  -3.3545254774338
H10     2.6167224915107  -1.1753444290240  -2.7383239291961
H11     1.8169060281297  -2.0197178318019  -1.3448746124746
H12     3.0930556958116  -0.8103647529240  -1.0451033370450
H13     2.1838959541816   1.3512820077724  -2.9564043397152
H14     2.6459249549741   1.5165226587667  -1.2227034879721
H15     1.0683121448681   2.1197367770073  -1.7579156919741
H16    -1.8125572969802  -2.0205522533281  -1.3423019485369
H17    -2.6156817507455  -1.1795469583322  -2.7349872226275
H18    -3.0899682656800  -0.8122018985245  -1.0416884490906
H19    -2.1886506552332   1.3443802670221  -2.9593596532925
H20    -1.0700260559313   2.1190962426515  -1.7681355217922
H21    -2.6451236912067   1.5166085429484  -1.2247429411686
&
```

Pd complex

```
&gen
dftname=bp86
basis=6-31g*
numd=5
igeopt=1
iaccg=3
isymm=0
ip11=2
&
&zmat
Pd1    -0.0000000000000   0.0000000000000  -0.0350590000000
Cl2    -1.6192750000000   0.0000000000000   1.5842150000000
Cl3     1.6192750000000   0.0000000000000   1.5842150000000
N1      1.2186682530345   0.0000000000000  -1.5859954316576
N2     -1.2175919847126  -0.0000000000000  -1.5866536354112
C3      0.0004613130294  -0.3114826026797  -2.3874355859076
C4      2.2436849011447  -1.0674931155787  -1.6985690154401
C5      1.8091775830353   1.3214962824709  -1.9118968755389
C6     -2.2411781795875  -1.0691018899829  -1.6959755491544
C7     -1.8105253167141   1.3193839691186  -1.9159426237250
H8      0.0015908875490  -1.3926244325052  -2.5958219091444
H9      0.0010008771100   0.2333253690327  -3.3545254774338
H10     2.6167224915107  -1.1753444290240  -2.7383239291961
H11     1.8169060281297  -2.0197178318019  -1.3448746124746
H12     3.0930556958116  -0.8103647529240  -1.0451033370450
H13     2.1838959541816   1.3512820077724  -2.9564043397152
H14     2.6459249549741   1.5165226587667  -1.2227034879721
H15     1.0683121448681   2.1197367770073  -1.7579156919741
H16    -1.8125572969802  -2.0205522533281  -1.3423019485369
H17    -2.6156817507455  -1.1795469583322  -2.7349872226275
H18    -3.0899682656800  -0.8122018985245  -1.0416884490906
H19    -2.1886506552332   1.3443802670221  -2.9593596532925
H20    -1.0700260559313   2.1190962426515  -1.7681355217922
H21    -2.6451236912067   1.5166085429484  -1.2247429411686
&
&atomic
atom   basis
Pd1    lacv3p
&
&nbo
```

```
&
```

He8-wedge

```
&gen
igeopt=1
dftname=bp86
basis=6-31g*
numd=5
iaccg=3
ip11=2
isymm=0
&
&zmat
N1      1.2186682530345#   0.0000000000000#  -1.5859954316576#
N2     -1.2175919847126#  -0.0000000000000#  -1.5866536354112#
C3      0.0004613130294  -0.3114826026797  -2.3874355859076
C4      2.2436849011447  -1.0674931155787  -1.6985690154401
C5      1.8091775830353   1.3214962824709  -1.9118968755389
C6     -2.2411781795875  -1.0691018899829  -1.6959755491544
C7     -1.8105253167141   1.3193839691186  -1.9159426237250
H8      0.0015908875490  -1.3926244325052  -2.5958219091444
H9      0.0010008771100   0.2333253690327  -3.3545254774338
H10     2.6167224915107  -1.1753444290240  -2.7383239291961
H11     1.8169060281297  -2.0197178318019  -1.3448746124746
H12     3.0930556958116  -0.8103647529240  -1.0451033370450
H13     2.1838959541816   1.3512820077724  -2.9564043397152
H14     2.6459249549741   1.5165226587667  -1.2227034879721
H15     1.0683121448681   2.1197367770073  -1.7579156919741
H16    -1.8125572969802  -2.0205522533281  -1.3423019485369
H17    -2.6156817507455  -1.1795469583322  -2.7349872226275
H18    -3.0899682656800  -0.8122018985245  -1.0416884490906
H19    -2.1886506552332   1.3443802670221  -2.9593596532925
H20    -1.0700260559313   2.1190962426515  -1.7681355217922
H21    -2.6451236912067   1.5166085429484  -1.2247429411686
He      0.000#   2.28#     0.000#
He      0.000#  -2.28#     0.000#
He      1.612#   0.00#     1.612#
He     -1.612#   0.00#     1.612#
He      0.806#   1.14#     0.806#
He     -0.806#   1.14#     0.806#
He     -0.806#  -1.14#     0.806#
He      0.806#  -1.14#     0.806#
&
```

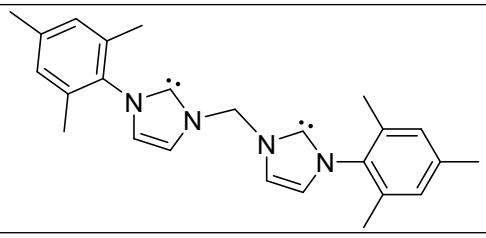
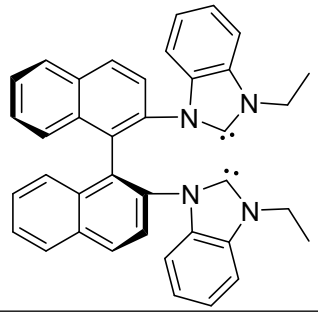
nHe8

```
&gen
igeopt=1
dftname=bp86
basis=6-31g*
numd=5
iaccg=3
ip11=2
isymm=0
&
&zmat
X1 0.0# 0.0#  0.0#
N2      1.2186682530345   0.0000000000000  -1.5859954316576
N3     -1.2175919847126  -0.0000000000000  -1.5866536354112
```

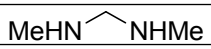
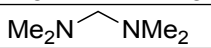
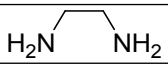
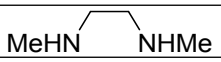
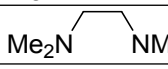
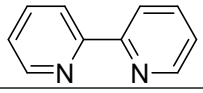
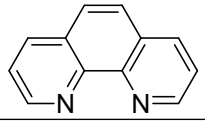
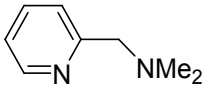
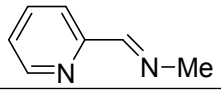
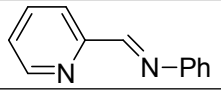
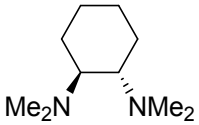
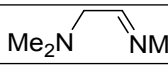
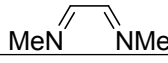
```
C3      0.0004613130294  -0.3114826026797  -2.3874355859076
C4      2.2436849011447  -1.0674931155787  -1.6985690154401
C5      1.8091775830353   1.3214962824709  -1.9118968755389
C6     -2.2411781795875  -1.0691018899829  -1.6959755491544
C7     -1.8105253167141   1.3193839691186  -1.9159426237250
H8      0.0015908875490  -1.3926244325052  -2.5958219091444
H9      0.0010008771100   0.2333253690327  -3.3545254774338
H10     2.6167224915107  -1.1753444290240  -2.7383239291961
H11     1.8169060281297  -2.0197178318019  -1.3448746124746
H12     3.0930556958116  -0.8103647529240  -1.0451033370450
H13     2.1838959541816   1.3512820077724  -2.9564043397152
H14     2.6459249549741   1.5165226587667  -1.2227034879721
H15     1.0683121448681   2.1197367770073  -1.7579156919741
H16    -1.8125572969802  -2.0205522533281  -1.3423019485369
H17    -2.6156817507455  -1.1795469583322  -2.7349872226275
H18    -3.0899682656800  -0.8122018985245  -1.0416884490906
H19    -2.1886506552332   1.3443802670221  -2.9593596532925
H20    -1.0700260559313   2.1190962426515  -1.7681355217922
H21    -2.6451236912067   1.5166085429484  -1.2247429411686
He      0.000#   2.28#     0.000#
He      0.000#  -2.28#     0.000#
He      1.612#   0.00#     1.612#
He     -1.612#   0.00#     1.612#
He      0.806#   1.14#     0.806#
He     -0.806#   1.14#     0.806#
He     -0.806#  -1.14#     0.806#
He      0.806#  -1.14#     0.806#
&
&coord
X1  N2 # 2.
X1  N3 # 2.
&
```

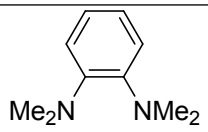
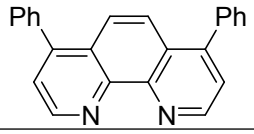
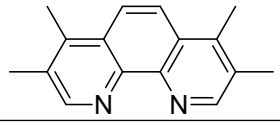
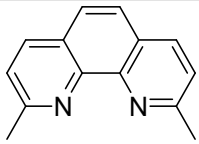
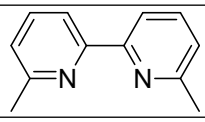
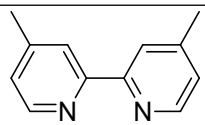
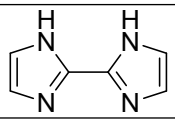
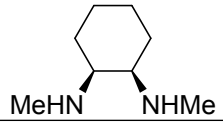
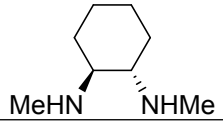
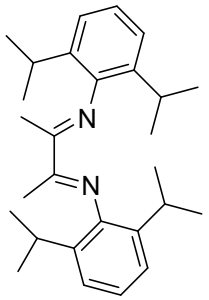
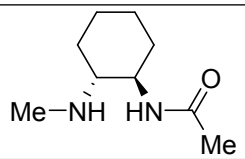
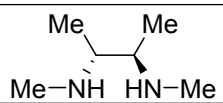
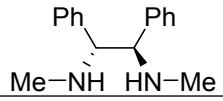
Table S1: Ligand structures and labelling.
(Green highlights for ligands used as reference)

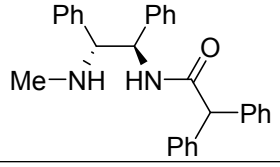
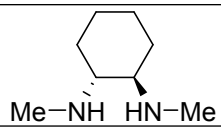
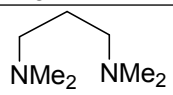
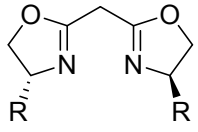
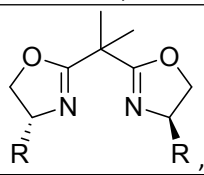
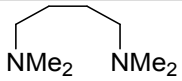
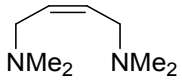
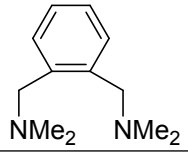
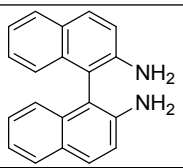
a) CC (reference: cn01)

cc01	
cc02	

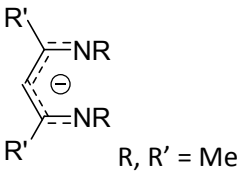
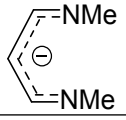
b) NN (reference: nn05)

nn01	
nn02	
nn03	
nn04	
nn05	
nn06	
nn07	
nn08	
nn09	
nn10	
nn11	
nn12	
nn13	
nn14	

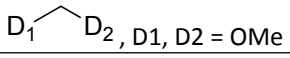
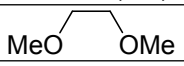
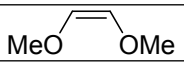
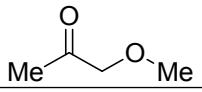
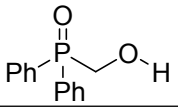
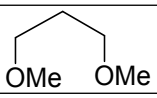
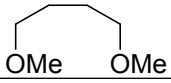
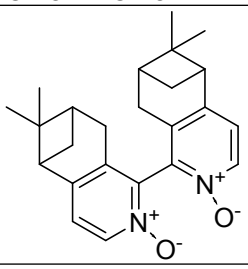
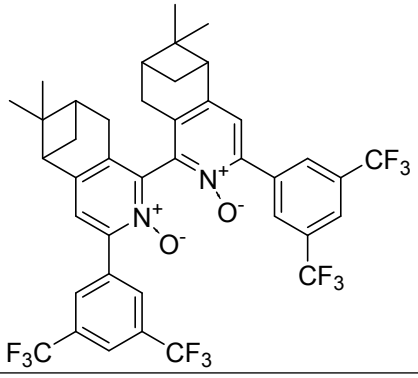
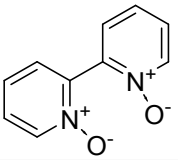
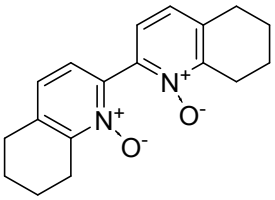
nn15	
nn16	
nn17	
nn18	
nn19	
nn20	
nn21	
nn22	
nn23	
nn24	
nn25	
nn26	
nn27	
nn28	

nn29	
nn30	
nn31	
nn32	 , R=Me
nn33	as ldnn045, R=Ph
nn34	as ldnn045, R=iPr
nn35	 , R = Me
nn36	as ldnn048, R = Ph
nn37	as ldnn048, R = iPr
nn38	
nn39	
nn40	
nn41	

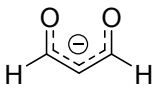
c) NN, charged (reference: nn_ch02)

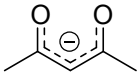
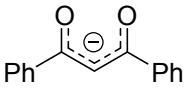
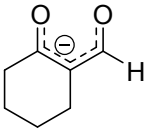
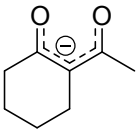
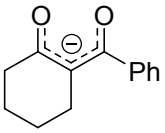
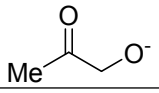
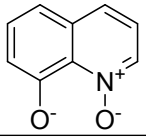
nn_ch02	 R, R’ = Me
nn_ch03	as ldnn051, R’ = Me, R = Ph
nn_ch01	

d) OO (reference: oo02)

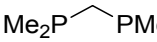
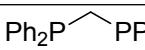
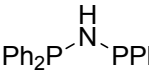
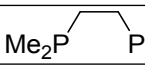
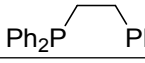
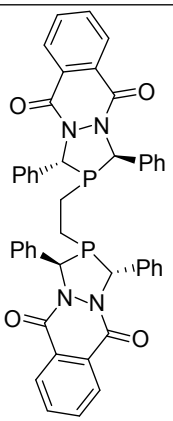
oo01	
oo02	
oo03	
oo04	
oo05	
oo06	
oo07	
oo08	
oo09	
oo10	
oo11	

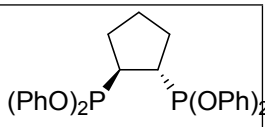
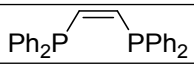
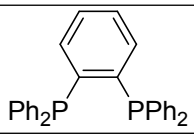
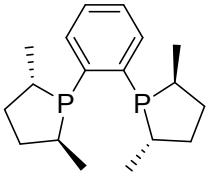
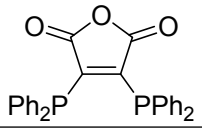
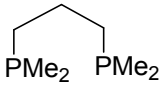
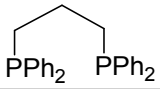
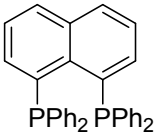
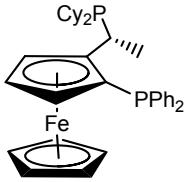
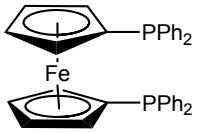
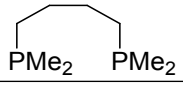
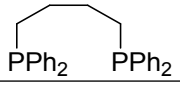
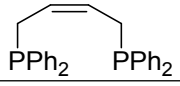
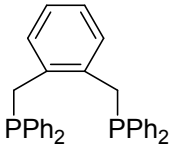
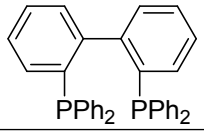
e) OO charged (reference: oo_ch02)

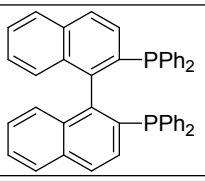
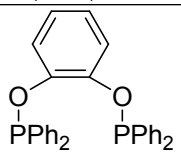
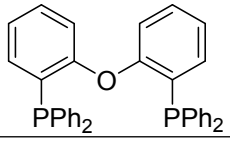
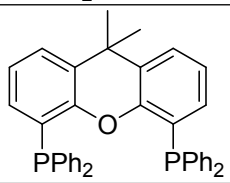
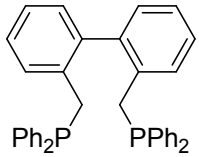
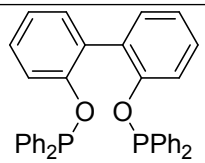
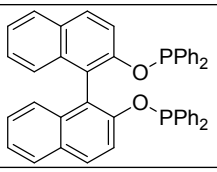
oo_ch01	

oo_ch02	
oo_ch03	
oo_ch04	
oo_ch05	
oo_ch06	
oo_ch07	
oo_ch08	

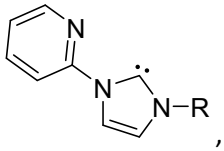
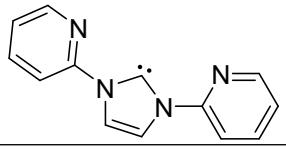
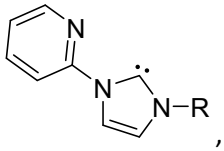
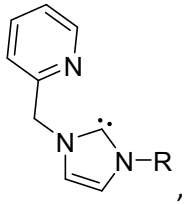
f) PP (reference: pp04)

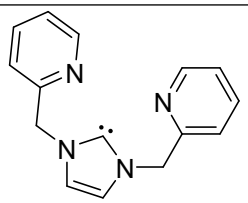
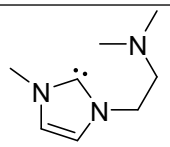
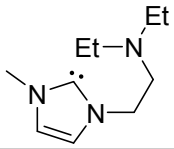
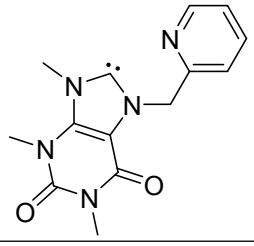
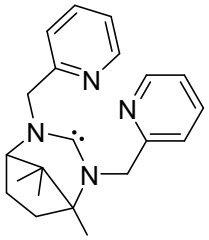
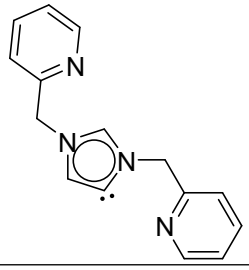
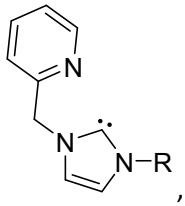
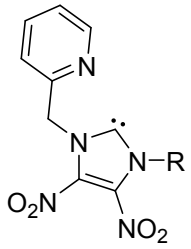
pp01	
pp02	
pp03	
pp04	
pp05	
pp06	

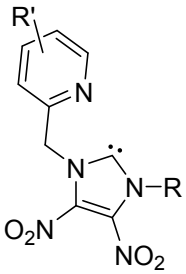
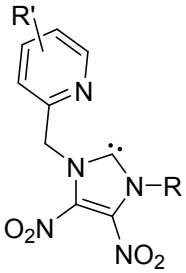
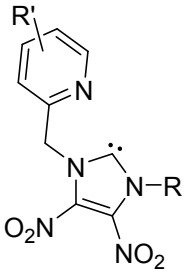
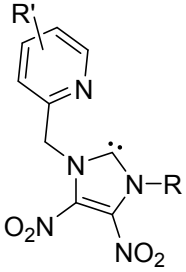
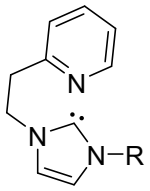
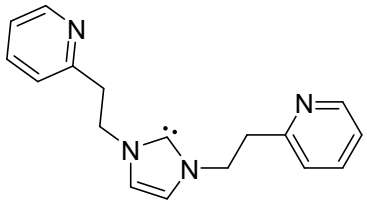
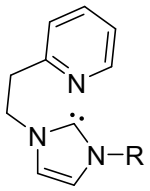
pp07	
pp08	
pp09	
pp10	
pp11	
pp12	
pp13	
pp14	
pp15	
pp16	
pp17	
pp18	
pp19	
pp20	
pp21	

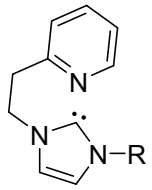
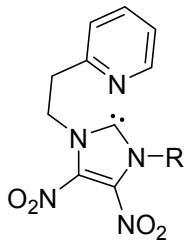
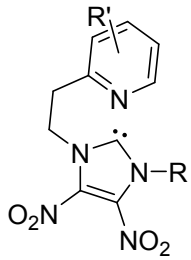
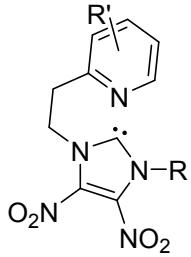
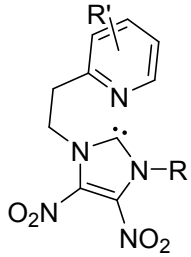
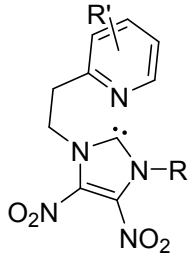
pp22	
pp23	
pp24	
pp25	
pp26	
pp27	
pp28	

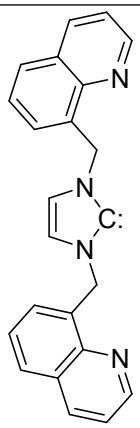
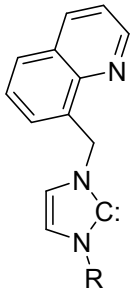
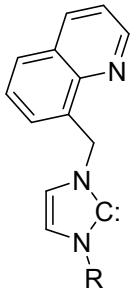
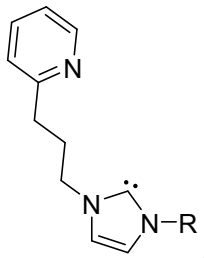
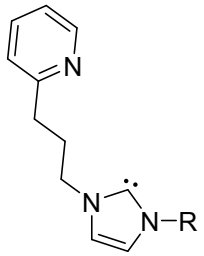
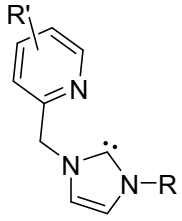
g) CN (reference: cn01)

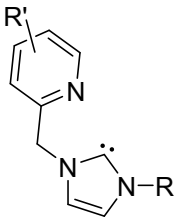
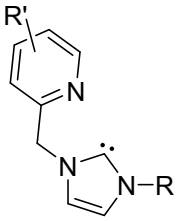
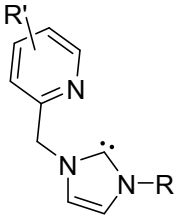
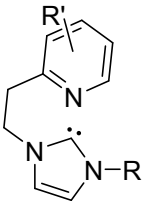
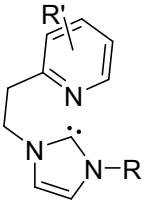
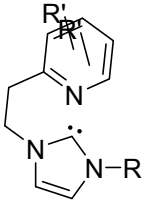
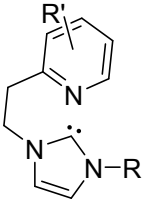
cn01	 , R = Me
cn02	
cn03	 , R = Me
cn04	 , R = Me

cn05	
cn06	
cn07	
cn08	
cn09	
cn10	
cn11	 , R = Mes
cn12	 , R = Mes

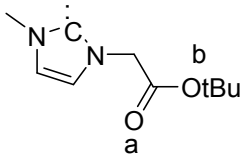
cn13	 , R’ = o-NO_2, R = Mes
cn14	 , R’ = p-NO_2, R = Mes
cn15	 , R’ = o-Me, R = Mes
cn16	 , R’ = p-Me, R = Mes
cn17	 , R = Me
cn18	
cn19	 , R = Mes

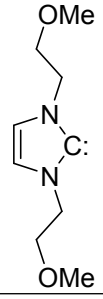
cn20	 , R = 2,6-iPr_2-C_6H_3
cn21	 , R = Mes
cn22	 , R’ = o-NO_2, R = Mes
cn23	 , R’ = p-NO_2, R = Mes
cn24	 , R’ = o-Me, R = Mes
cn25	 , R’ = p-Me, R = Mes

cn26	
cn27	 R = Me
cn28	 , R = Mes
cn29	 , R = Me
cn30	 , R = Mes
cn31	 R’= o-Me, R = Mes

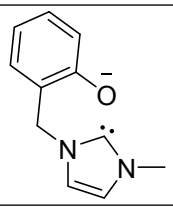
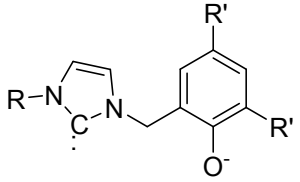
cn32	 , R’ = pMe, R = Mes
cn33	 R’= o-NO_2, R = Mes
cn34	 R’= p-NO_2, R = Mes
cn35	 , R’ = o-Me, R = Mes
cn36	 , R’ = p-Me, R = Mes
cn37	 , R’ = o-NO_2, R = Mes
cn38	 , R’ = p-NO_2, R = Mes

h) CO (reference: co02)

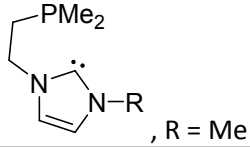
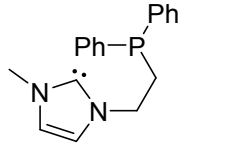
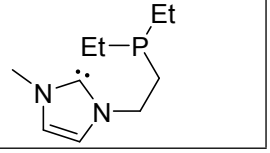
co01	

co02	

i) CO charged (reference co_ch01)

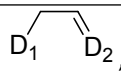
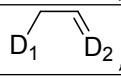
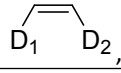
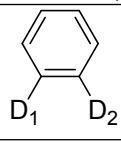
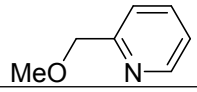
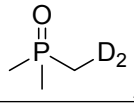
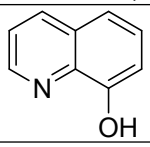
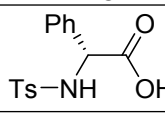
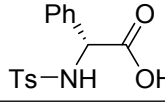
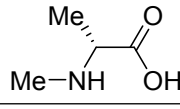
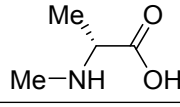
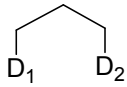
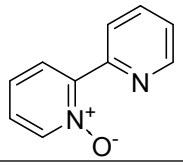
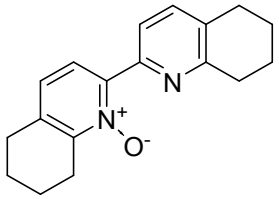
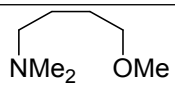
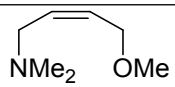
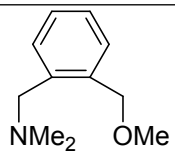
co_ch01	
co_ch02	 , R = Me, R’ = H
co_ch03	as co004, R = iPr
co_ch04	as co004, R = tBu
co_ch05	as co004, R = Mes
co_ch06	as co004, R = Me, R’ = Me
co_ch07	as co004, R = Me, R’ = CF_3

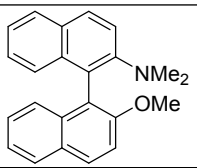
j) CP (reference: cp01)

cp01	 , R = Me
cp02	
cp03	

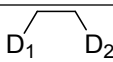
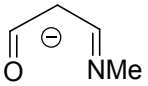
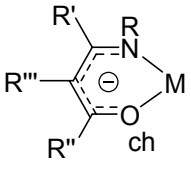
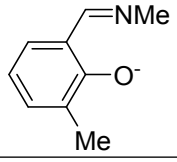
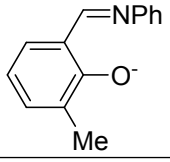
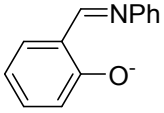
k) NO (reference: no02)

no01	 , D1 NMe_2, D2 OMe
no02	 , D1 NMe_2, D2 OMe

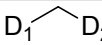
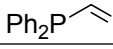
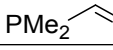
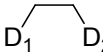
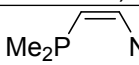
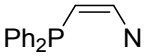
no03	 , D1 NMe_2, D2 O
no04	 , D1 = OMe, D2 = NMe
no05	 , D1 = NMe_2, D2 = OMe
no06	 , D1 NMe_2, D2 OMe
no07	
no08	 , D2 = NMe_2
no09	
no10	 (N/OH)
no11	 (N/O=)
no12	 (N/OH)
no13	 (N/O=)
no14	 , D1 NMe_2, D2 OMe
no15	
no16	
no17	
no18	
no19	

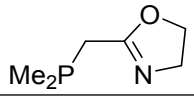
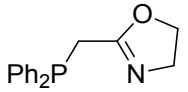
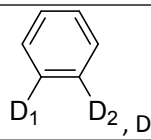
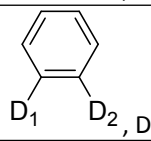
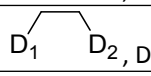
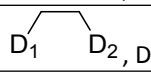
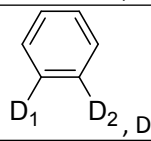
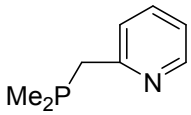
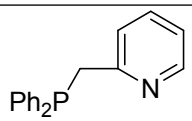
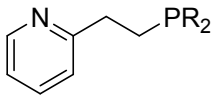
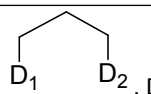
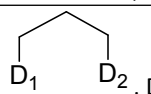
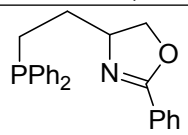
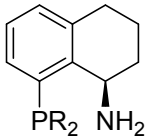
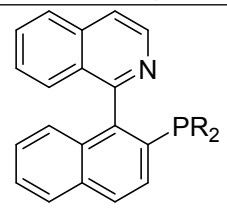
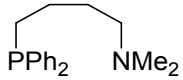
no20	

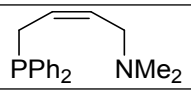
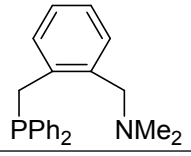
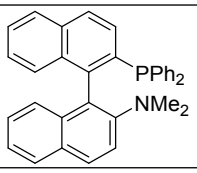
l) NO charged (reference: no_ch01)

no_ch01	 , D1 NMe_2, D2 O^-
no_ch02	
no_ch03	 , R', R'' – Me, R'''=H, R = Me
no_ch04	as ldno023, R = Ph
no_ch05	
no_ch06	
no_ch07	

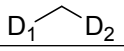
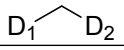
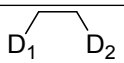
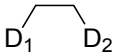
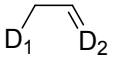
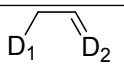
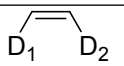
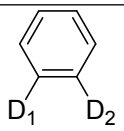
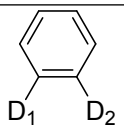
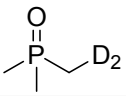
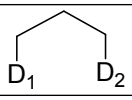
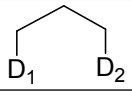
m) PN (reference: pn05)

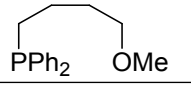
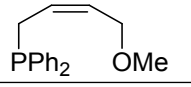
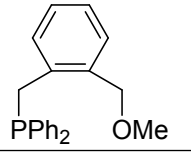
pn01	 , D1 =PMe_2, D2 = NMe_2
pn02	 , D1 = PPh_2, D2 = NMe_2
pn03	
pn04	
pn05	 , D1 PMe_2, D2 = NMe_2
pn06	 , D1 PPh_2, D2 = NMe_2
pn07	
pn08	
pn09	 , D1 = PMe_2, D2 = NMe
pn10	 , D1 = PPh_2, D2 = NMe
pn11	 , D1 = PPh_2, D2 = NPh

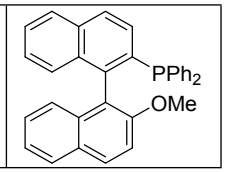
pn12	
pn13	
pn14	 , D1 = PMe_2, D2 = NMe_2
pn15	 , D1 = PPh_2, D2 = NMe_2
pn16	 , D1 = PPh_2, D2 = NH_2
pn17	 , D1 = PPh_2, D2 = NHMe
pn18	 , D1 = PPh_2, D2 = NH_2
pn19	
pn20	
pn21	 , R = Ph
pn22	 , D1 = PMe_2, D2 = NMe_2
pn23	 , D1 = PPh_2, D2 = NMe_2
pn24	
pn25	 , R = Ph
pn26	 , R = Ph
pn27	

pn28	
pn29	
pn30	

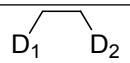
n) PO (reference: po03)

		D1	D2
po01		PMe_2	OMe
po02		PPh_2	OMe
po03		PMe_2	OMe
po04		PPh_2	OMe
po05		PMe_2	O
po06		PPh_2	O
po07		PPh_2	OMe
po08		PMe_2	OMe
po09		PPh_2	OMe
po10			PMe_2
po11		PMe_2	OMe
po12		PPh_2	OMe

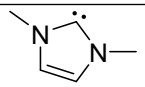
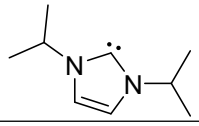
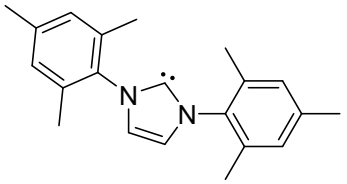
po13	
po14	
po15	

po16	

o) PO charged

		D1	D2
po_ch_01		PMe_2	O^-

p) Monodentates (same or mixed)

mcc01	
mcc02	
mnn01	
mcc03	
mcn01	mcc01 + mnn01
mcn02	mcc02 + mnn01
mcn03	mcc03 + mnn01

PCA Diagnostics and Details

a) All ligands included

Table S2: Model overview (224 ligands, 18 descriptors)

Component	R^2X	R^2X(cum)	Eigenvalue	Q^2	Q^2(cum)
0	Cent.				
1	0.359	0.359	6.45	0.239	0.239
2	0.156	0.514	2.81	0.0825	0.302
3	0.126	0.640	2.26	0.148	0.406

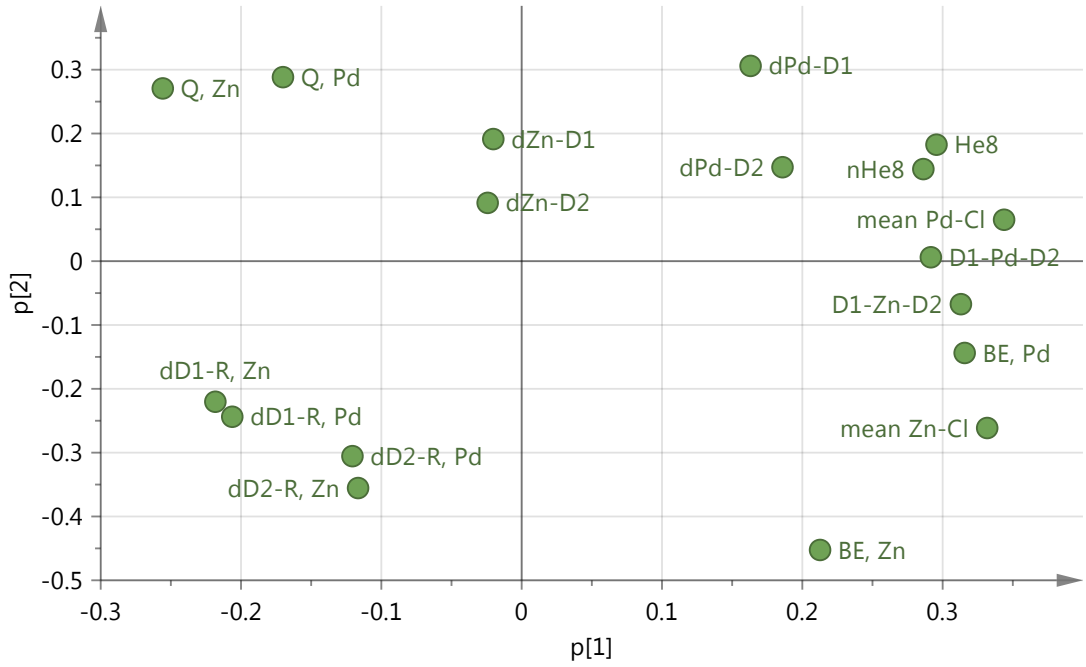


Figure S1: Descriptor loadings plot (see spreadsheet for data)

Detailed PCA plots

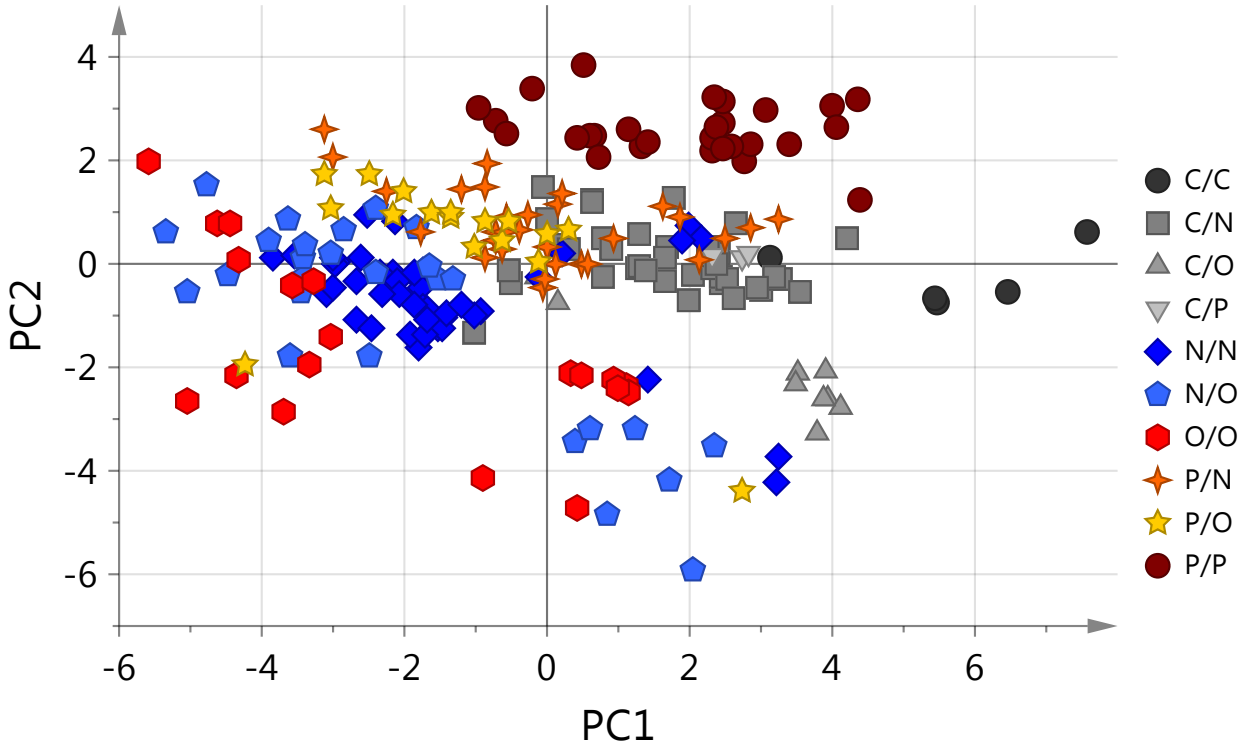


Figure S2: Larger version of Figure 1.

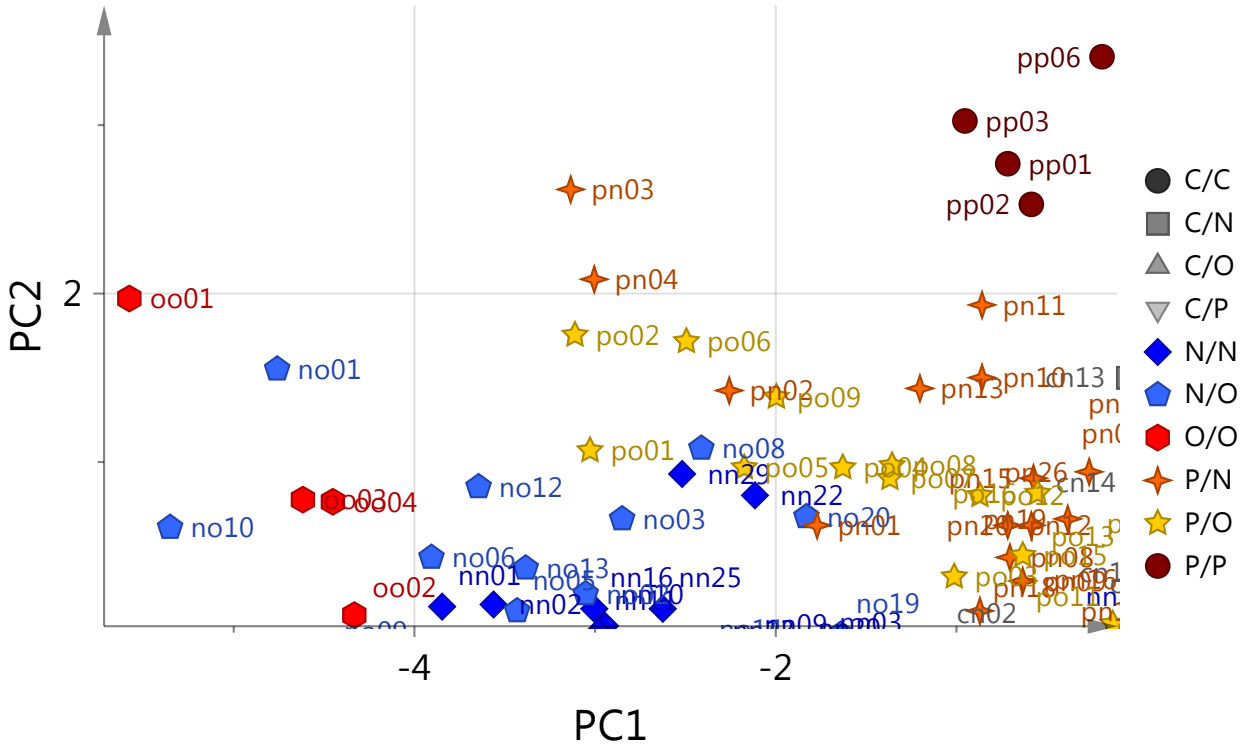


Figure S3: North West quadrant of Figure 1, all labels.

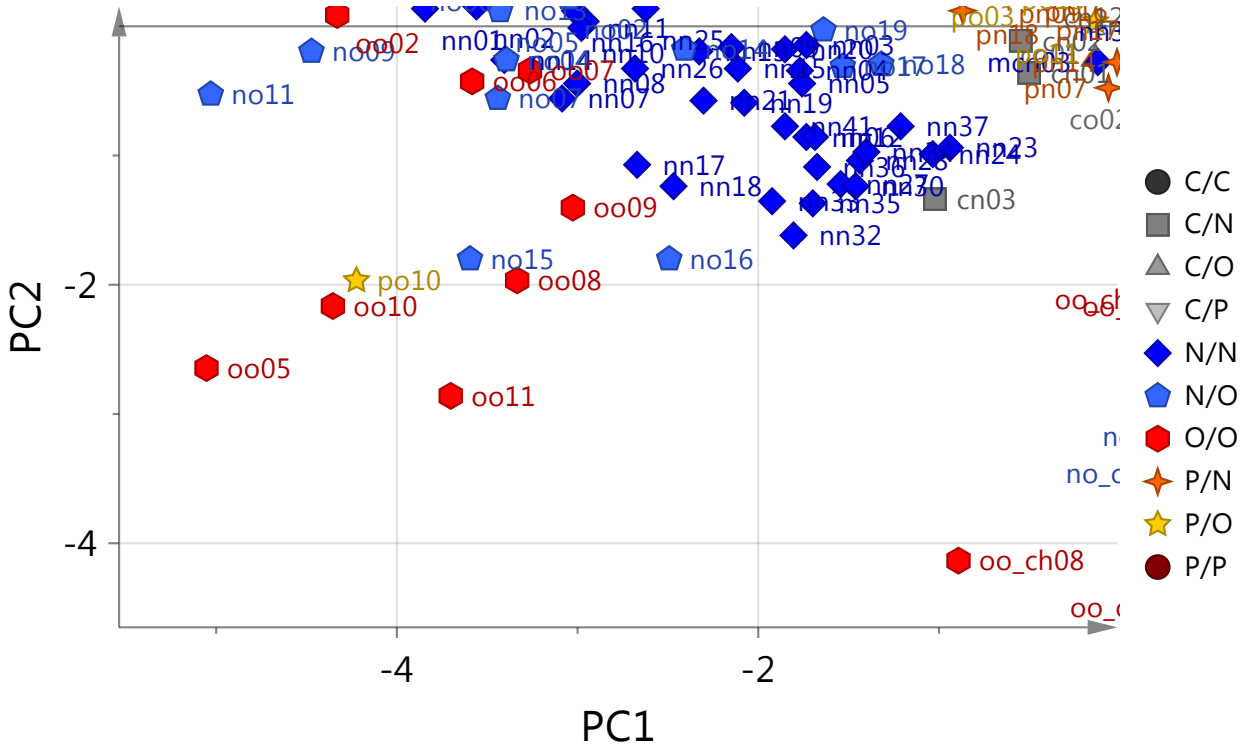


Figure S4: South West quadrant of Figure 1, all labels.

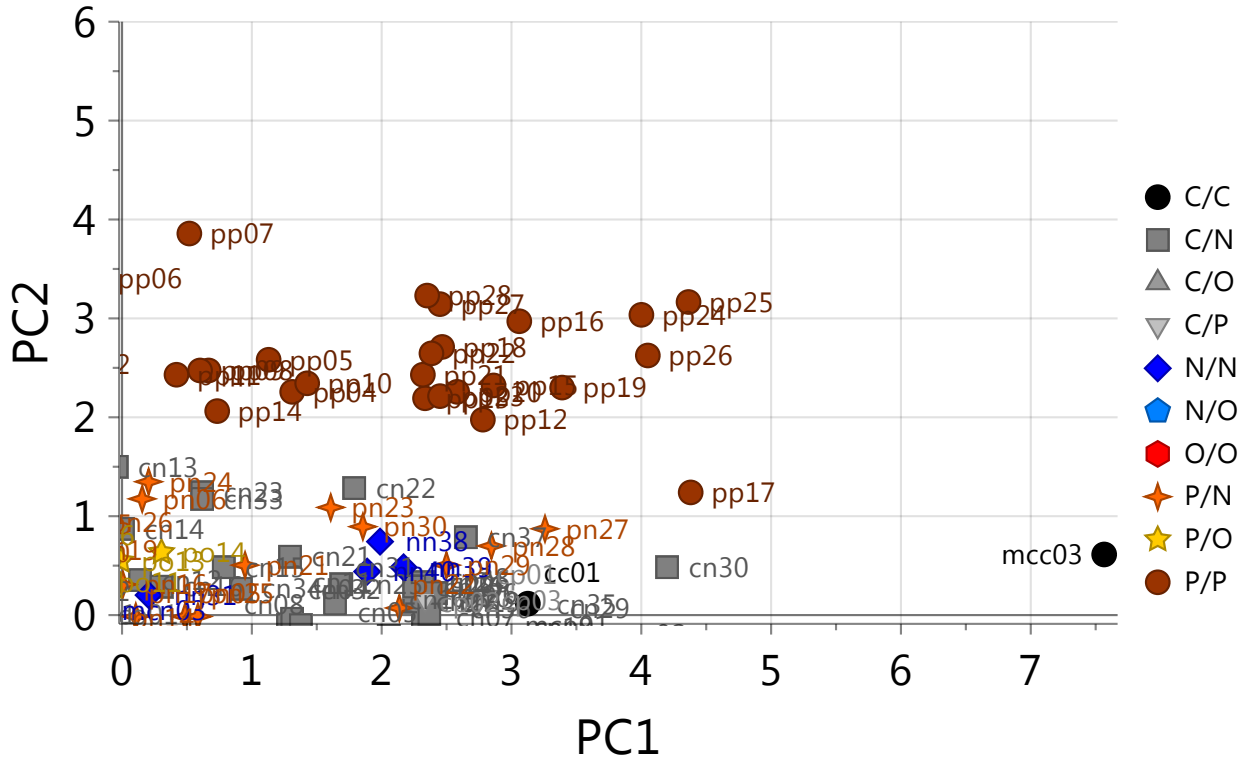


Figure S5: North East quadrant of Figure 1, all labels.

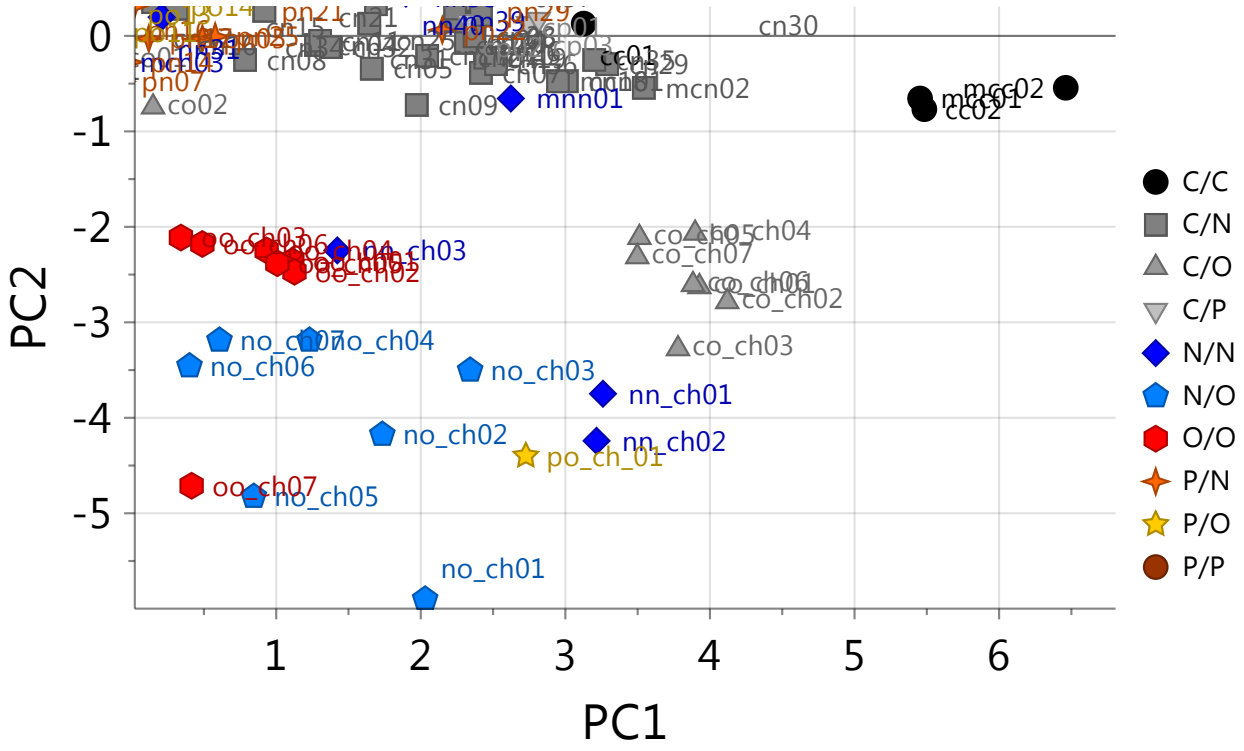


Figure S6: South East quadrant of Figure 1, all labels.

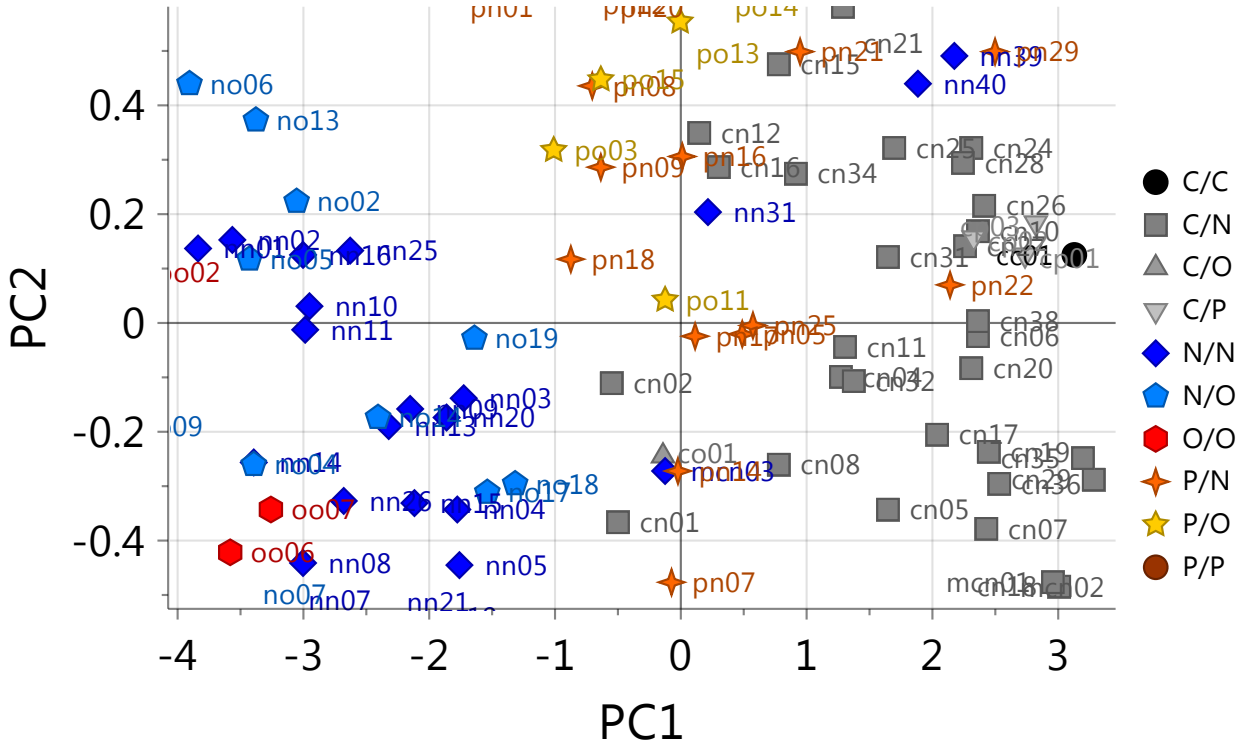
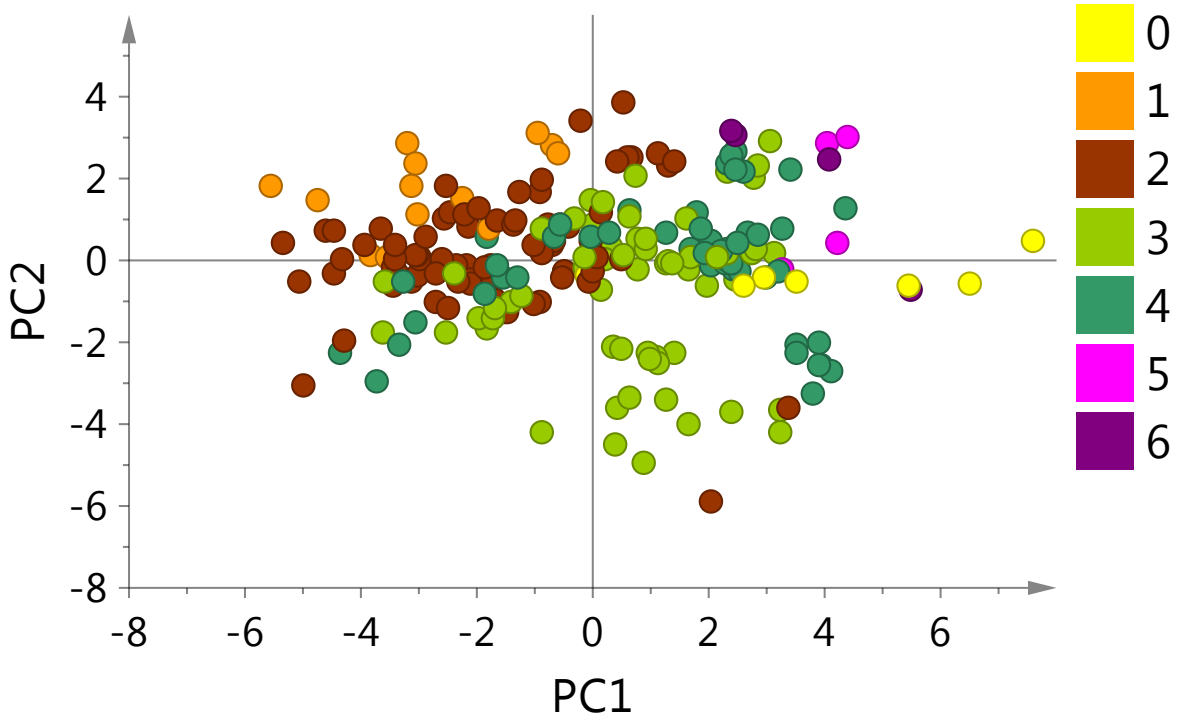


Figure S7: Zoomed in on equatorial region of Figure 1, all labels.

a)



b)

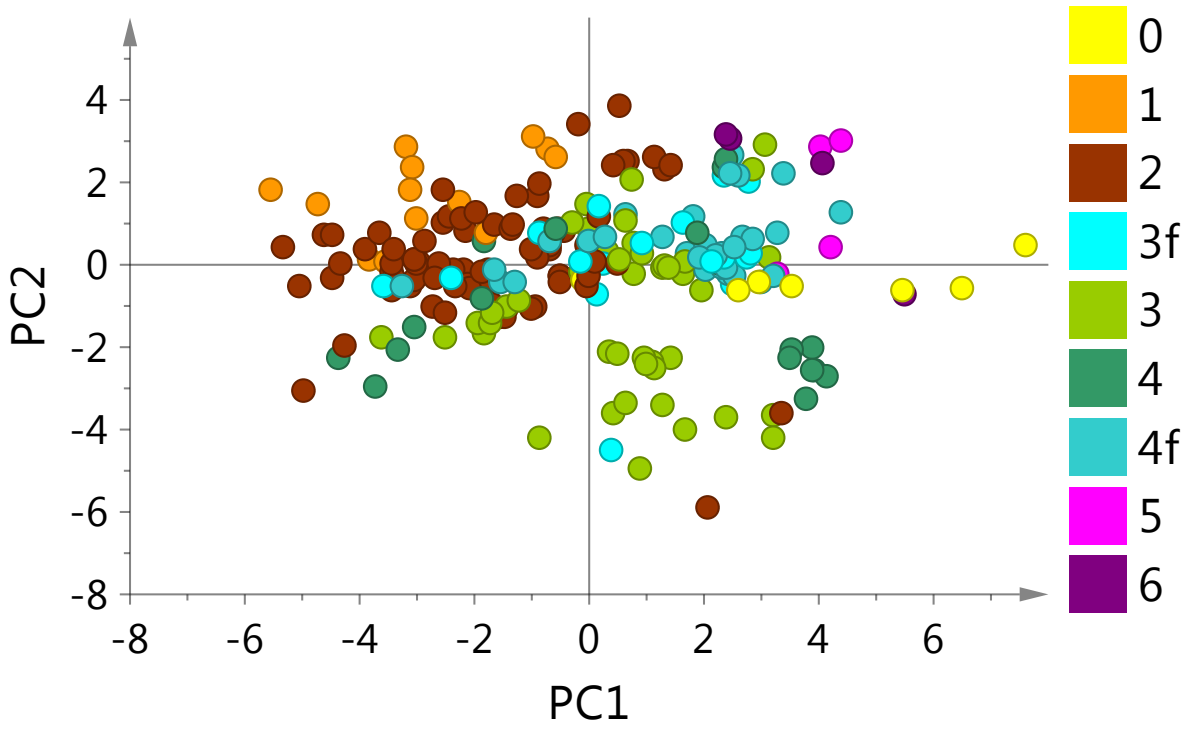


Figure S8: a) Inset of Figure 1, map coloured by bridge length. b) Additional coding according to flexibility of bridge with 3f and 4f bridges likely to be more flexible.

Further discussion of PCA results

While PCA is useful for dimension reduction, 18 descriptors will give rise to 18 PCs, albeit of less and less importance to capturing the variation in the dataset. In the present case, PCs 1 and 2 capture 52% of the variation, with PC3 adding a further 13%. While we have focussed on the first two PCs to aid visualisation, it can be instructive to inspect PC3 and a 3D plot of PCs (Figure S9) can be seen below. In the orientation shown, this further emphasises the separation of different ligands by this analysis approach.

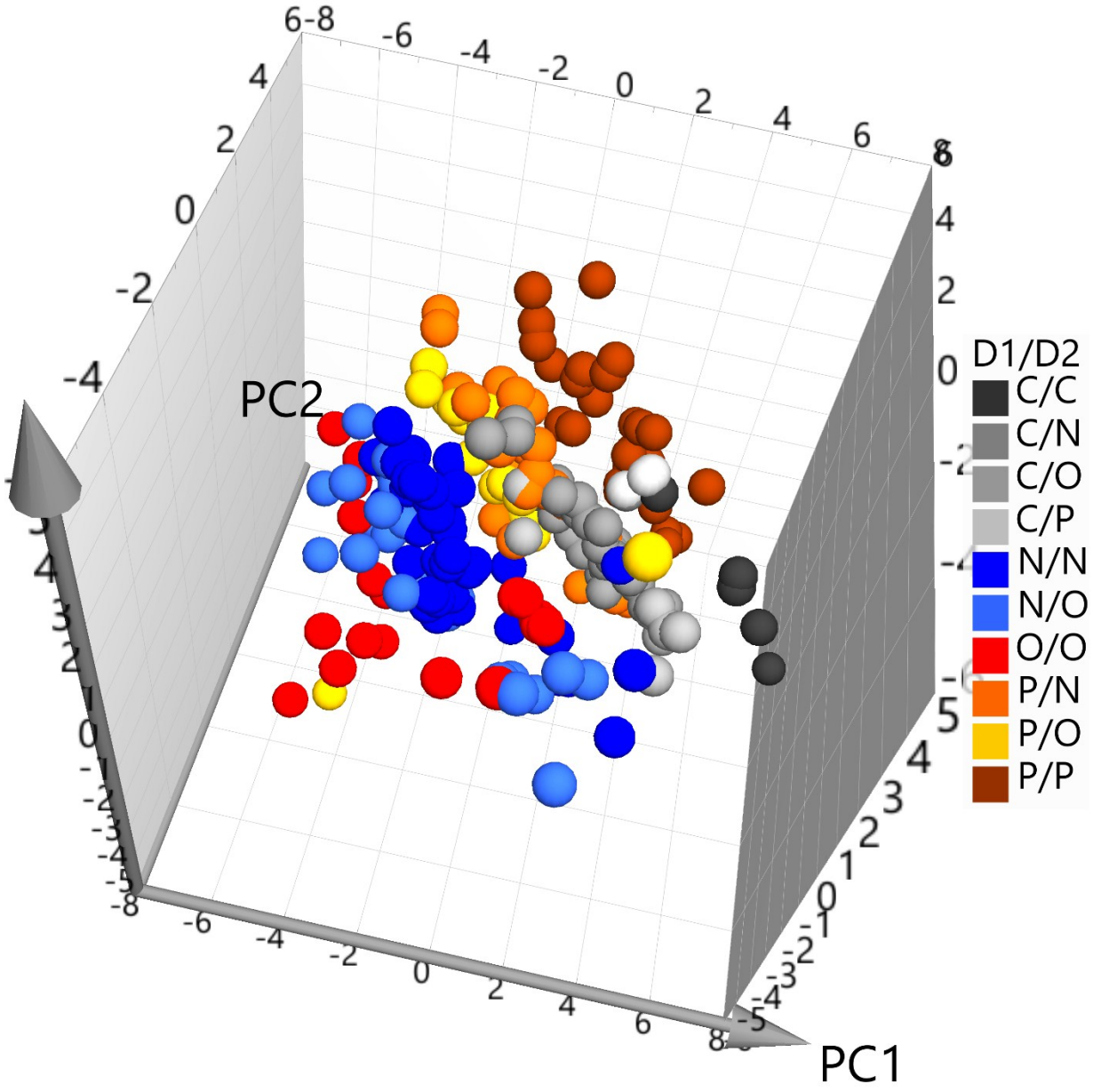


Figure S9: Principal component score plot showing PCs 1-3 for LKB-bid, all ligands.

For the full ligand set, PC3 loads steric and bite angle descriptors quite highly (see Figure S10 and spreadsheet for loadings), along with structural changes around the donor atoms for both metal complexes and the binding energy to the palladium fragment. This suggests a steric component for this PC, albeit one that is not separated well from electronic effects.

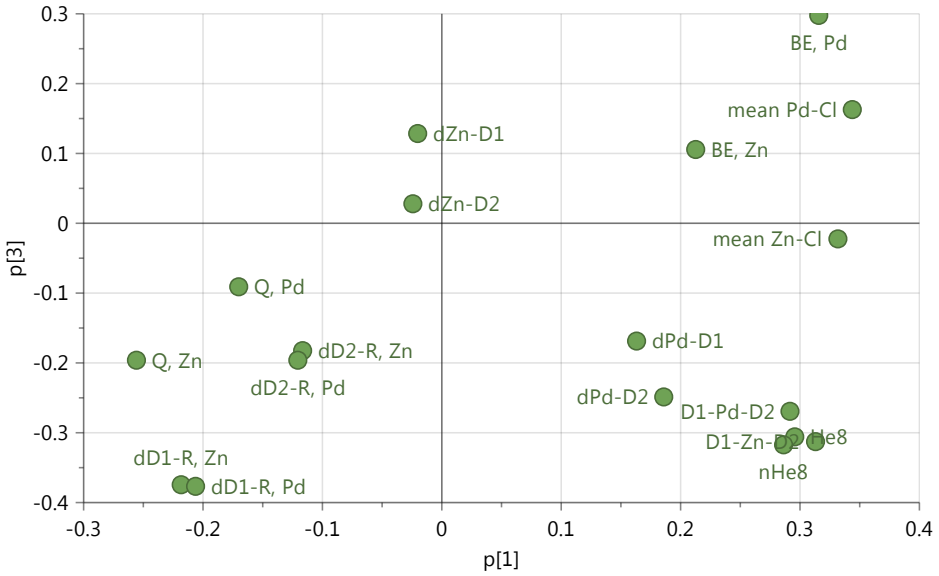


Figure S10: Descriptor loadings for PCs 1 and 3.

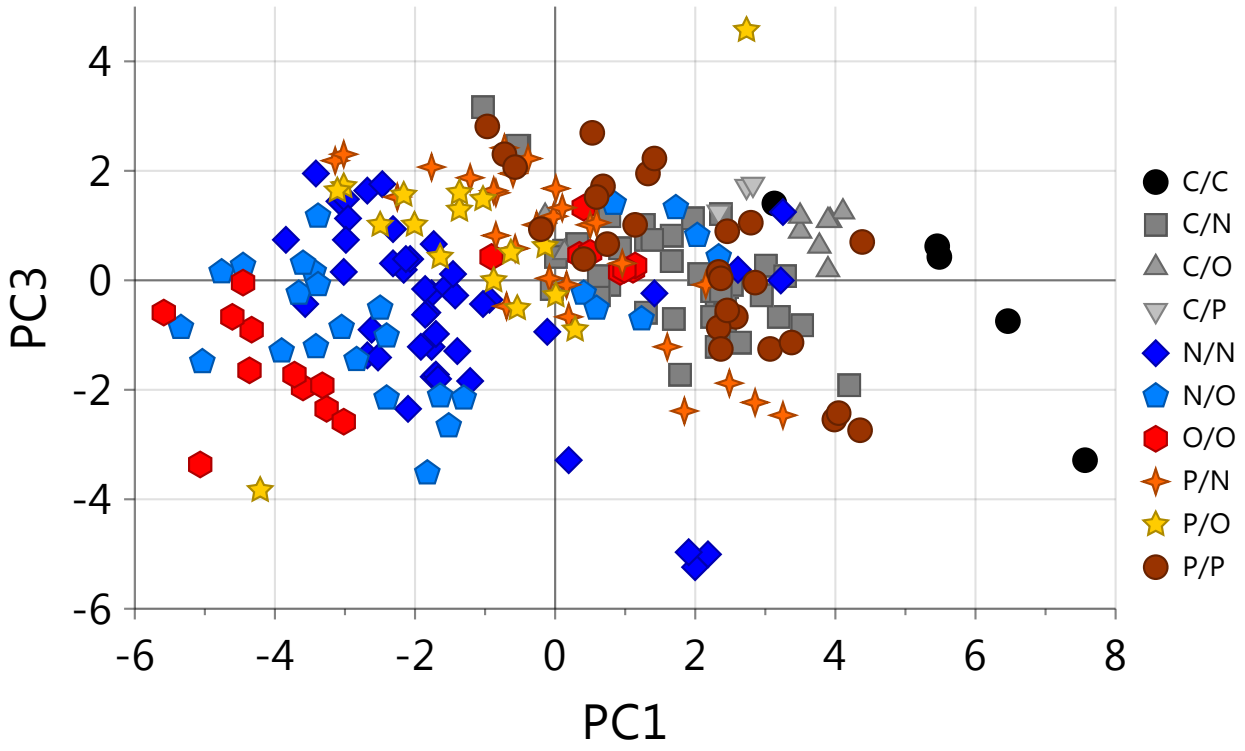


Figure S11: Principal component score plot for PCs 1 and 3.

Figure S11 shows a score plot of PCs 1 and 3. This analysis of PC3 suggests that an improved resolution of ligand donor atoms, particularly for NN, NO and OO donor ligands can be achieved by consideration of the third PC, but that these continue to overlap/lie close in chemical space, suggesting that their properties are quite similar. Ligand classes are less well-defined at higher values of PC1, where charged ligands lie alongside other ligand classes. Overall, while we have focussed our discussion in the manuscript predominantly on PCs1 and 2, it appears prudent to consider PC3 as well in experimental design and data analysis.

PCA excluding charged donors

Table S3: Model overview (198 ligands, 18 descriptors)

Component	R^2X	R^2X(cum)	Eigenvalue	Q^2	Q^2(cum)
0	Cent.				
1	0.4	0.4	7.2	0.323	0.323
2	0.143	0.543	2.58	0.127	0.409
3	0.102	0.645	1.83	0.0377	0.431

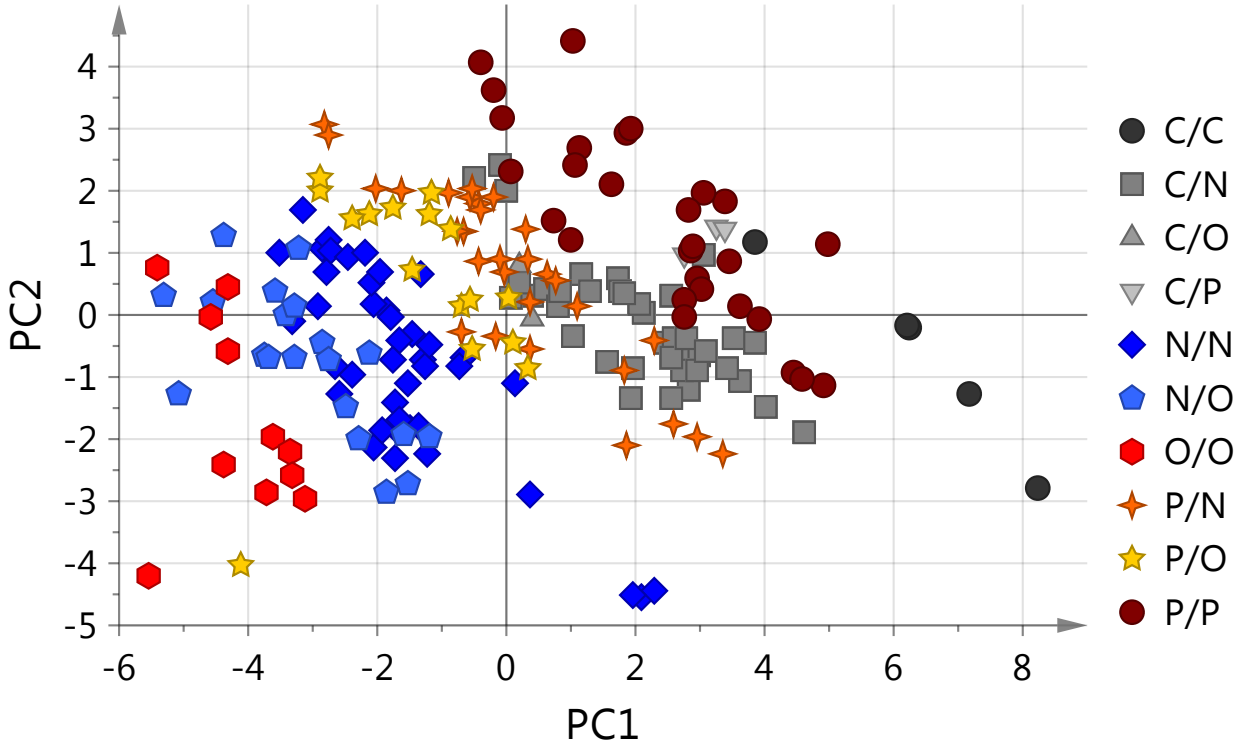


Figure S12: Principal component score plot (PCs 1 and 2) for all uncharged ligands (198), coloured by donor atoms.

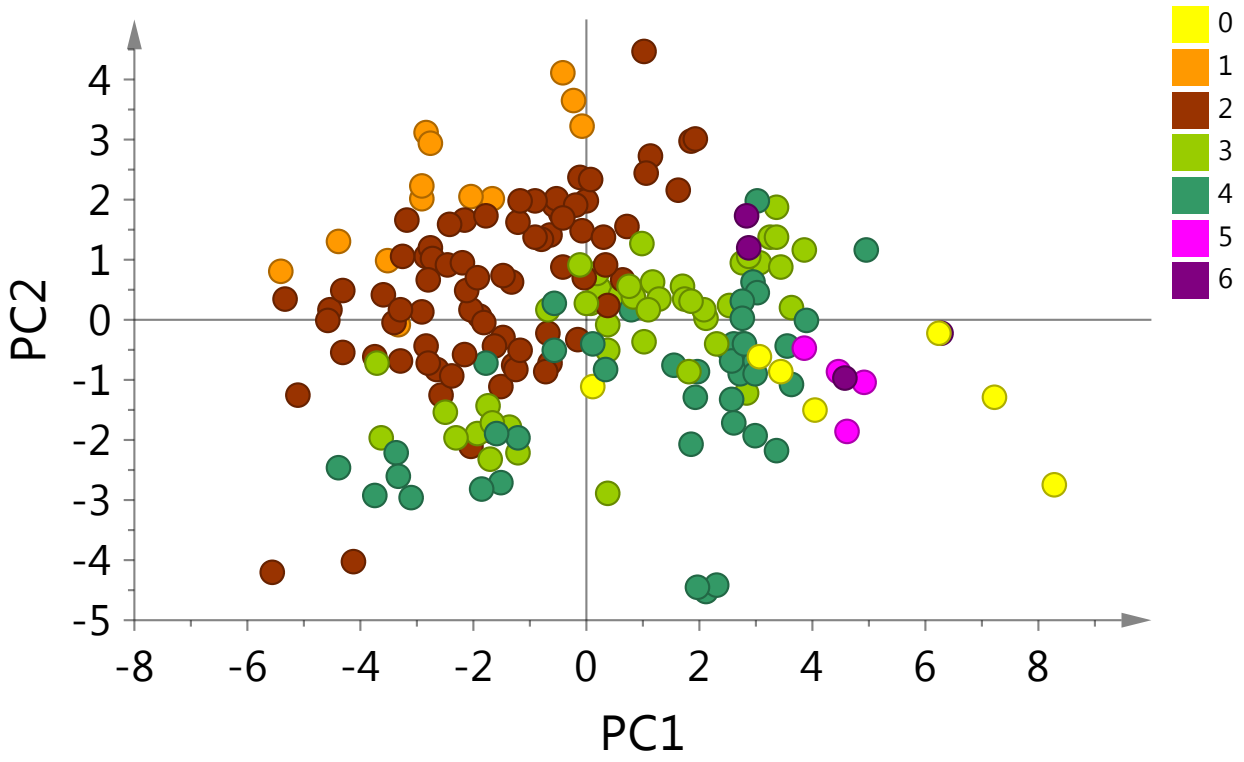


Figure S13: Principal component score plot (PCs 1 and 2) for all uncharged ligands (198), coloured by bridge length

PCA Diagnostics and Details

C, N and O donors, excluding all P-donors

Table S4: Model overview (146 ligands, 18 descriptors)

Component	R^2X	R^2X(cum)	Eigenvalue	Q^2	Limit	Q^2(cum)
0	Cent.					
1	0.433	0.433	7.8	0.342	0.0591	0.342
2	0.135	0.568	2.42	0.0722	0.0621	0.389
3	0.111	0.679	1.99	0.0678	0.0654	0.431
4	0.0787	0.757	1.42	-0.027	0.0691	0.415

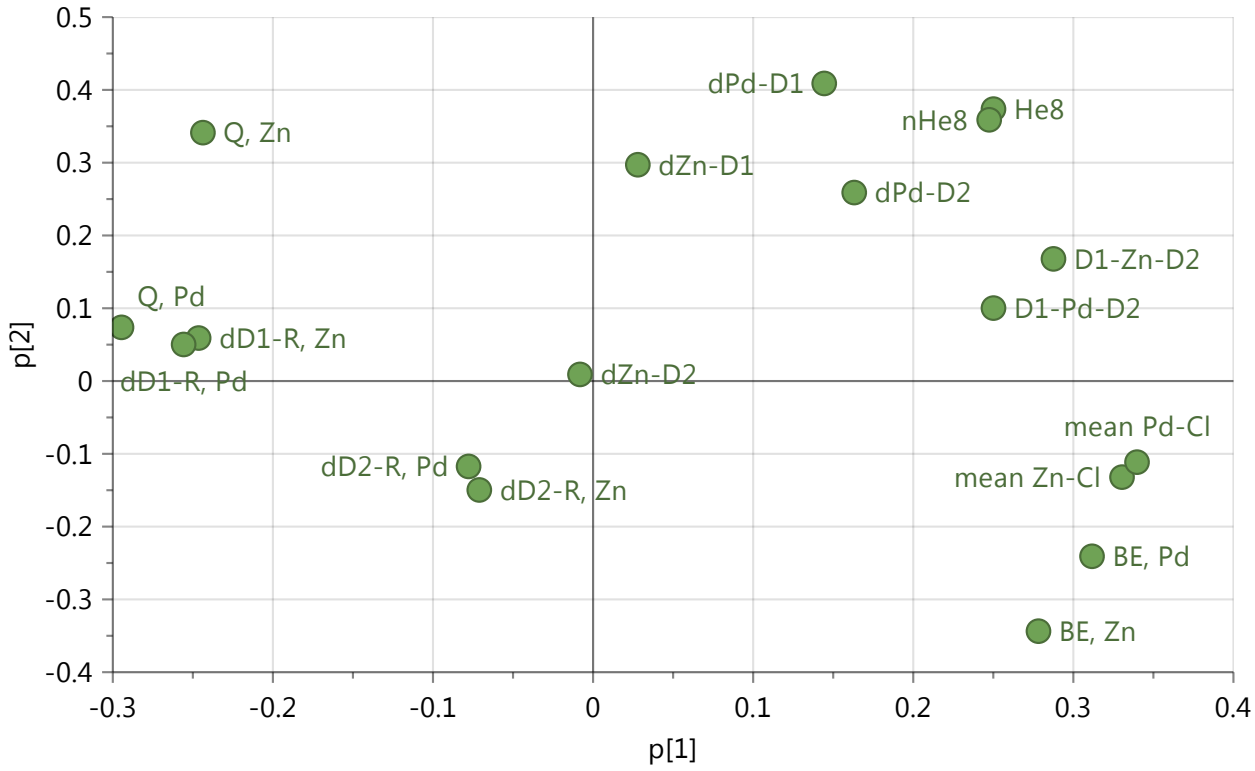


Figure S14: Loadings plot PCs 1 and 2.

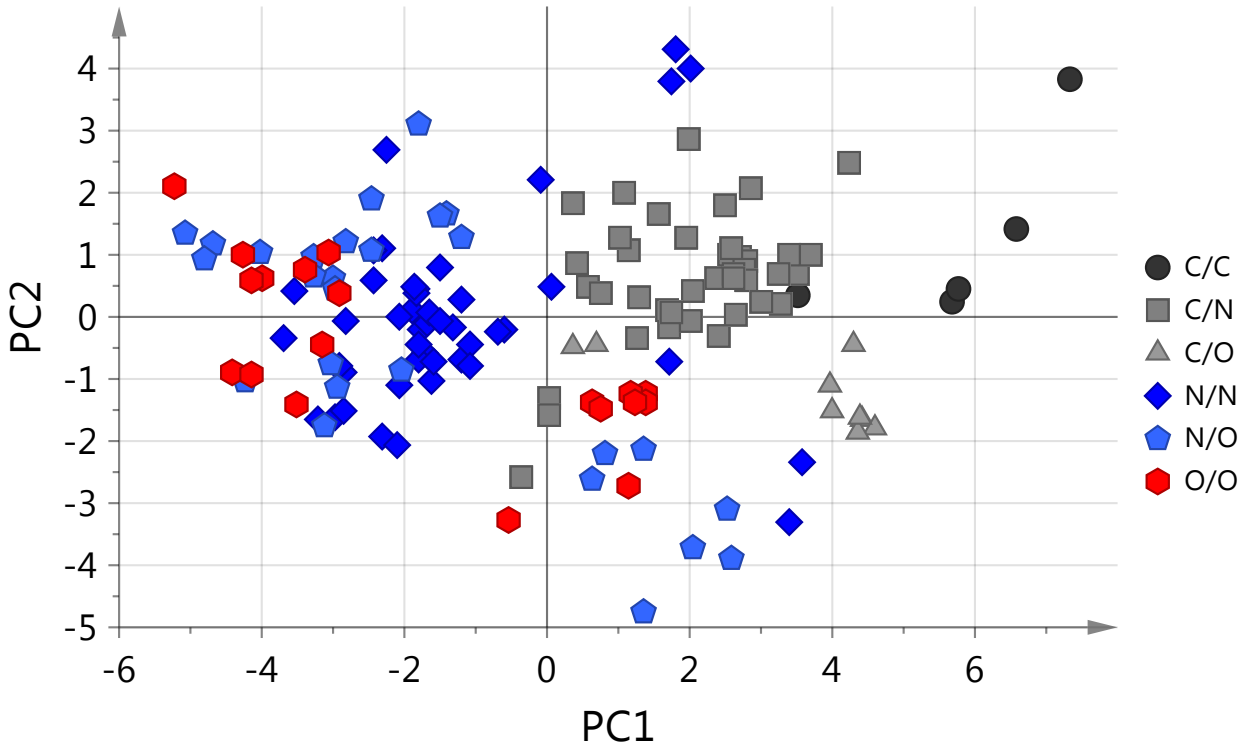


Figure S15: Figure 2 enlarged.

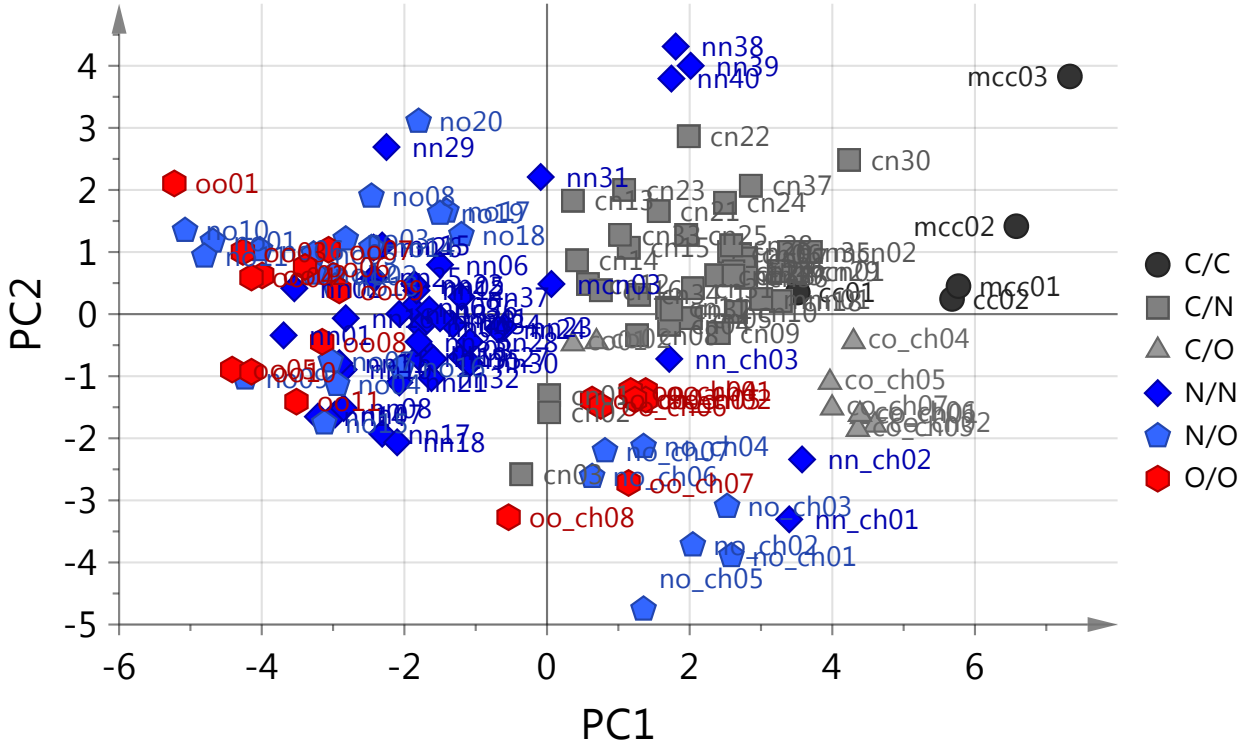


Figure S16: Overview with ligand numbers.

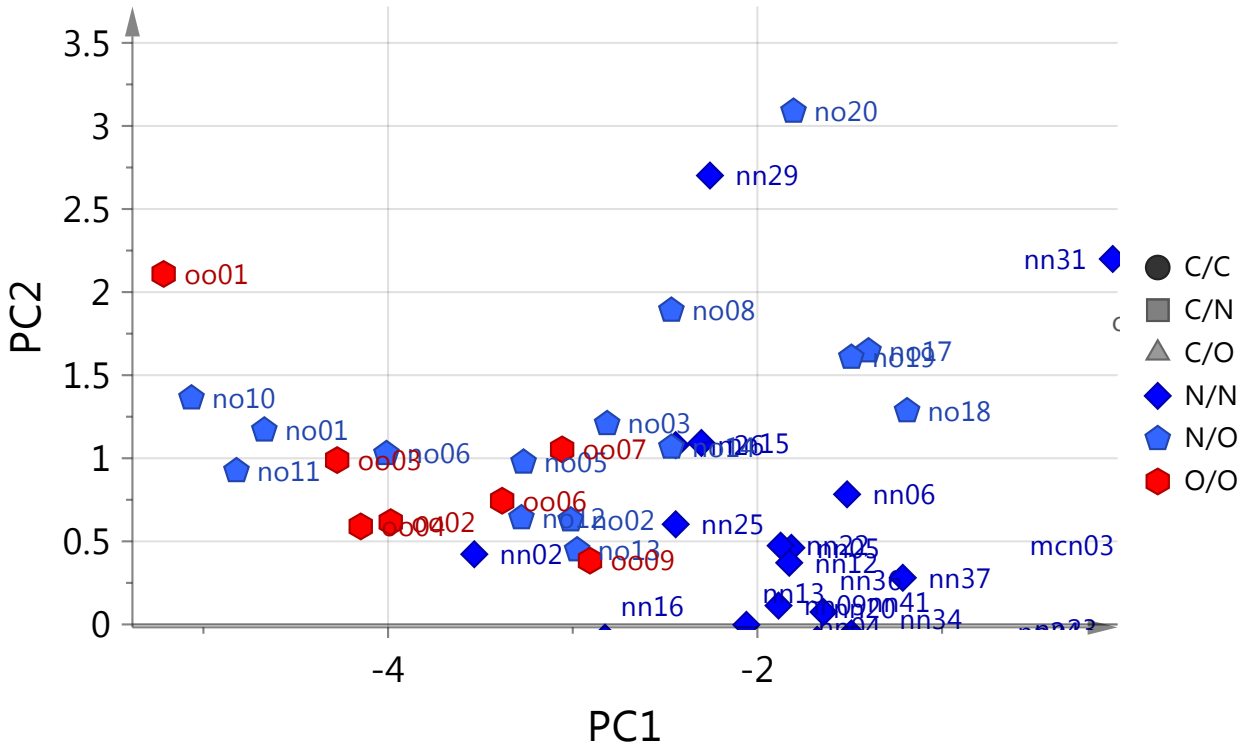


Figure S17: North West Quadrant of Fig. 2, all labels.

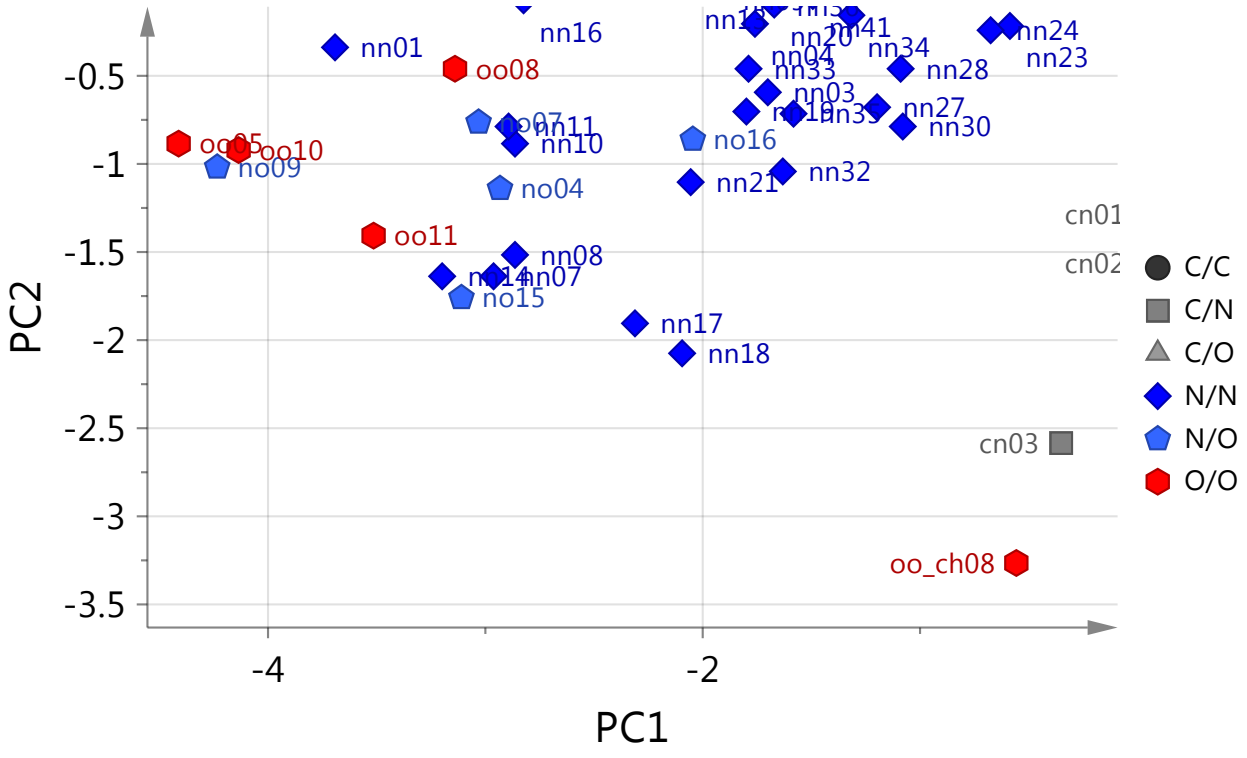


Figure S18: South West Quadrant of Fig. 2, all labels.

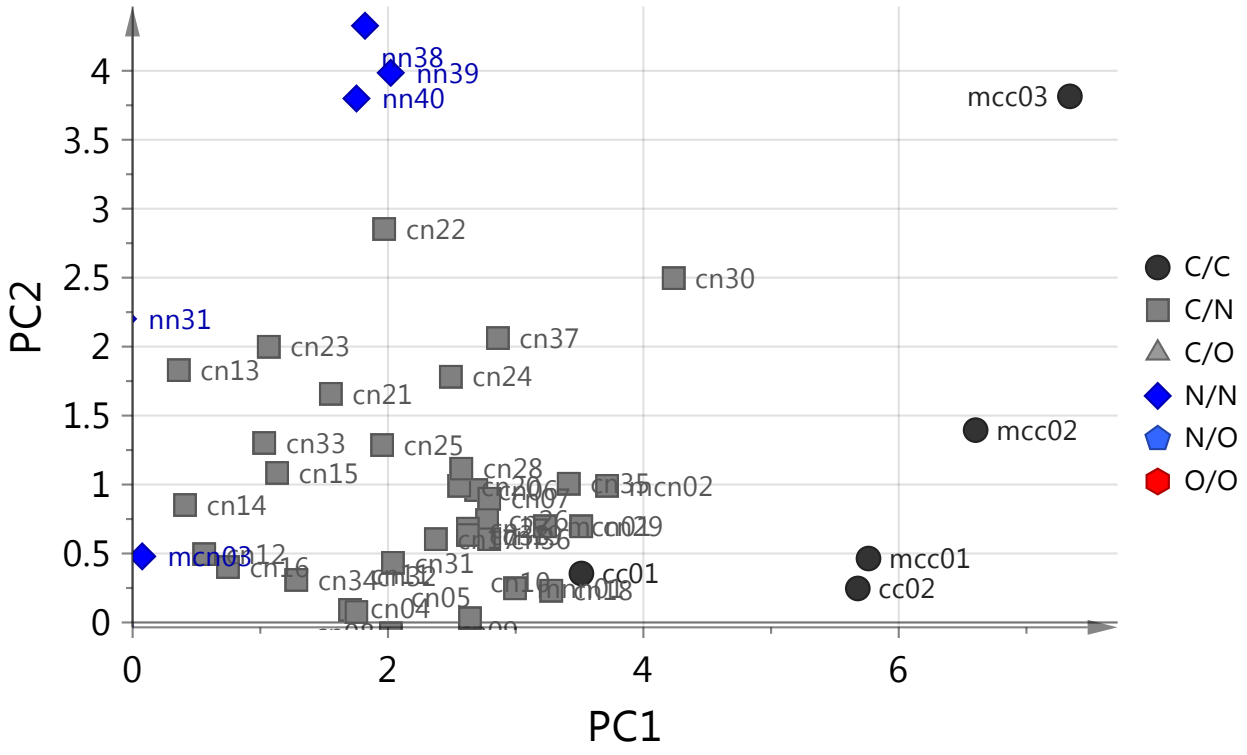


Figure S19: North East Quadrant of Fig. 2, all labels.

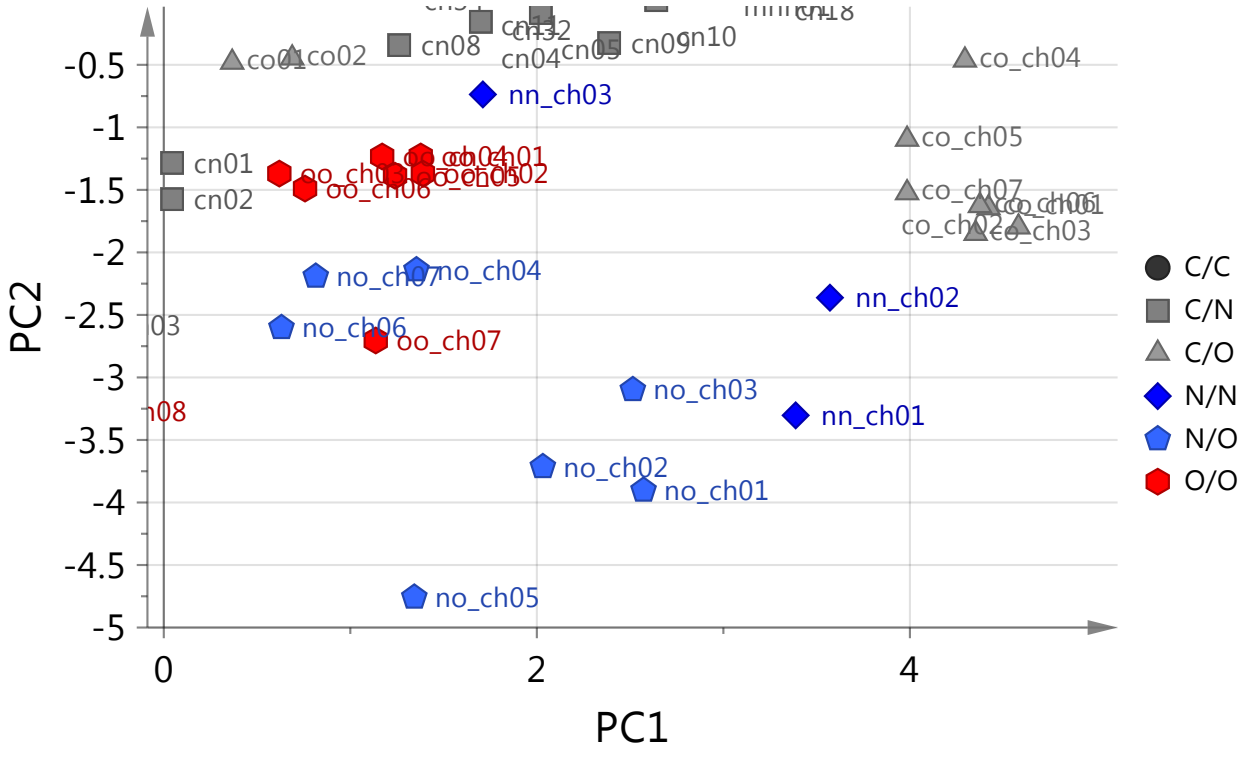


Figure S20: South East Quadrant of Fig. 2, all labels.

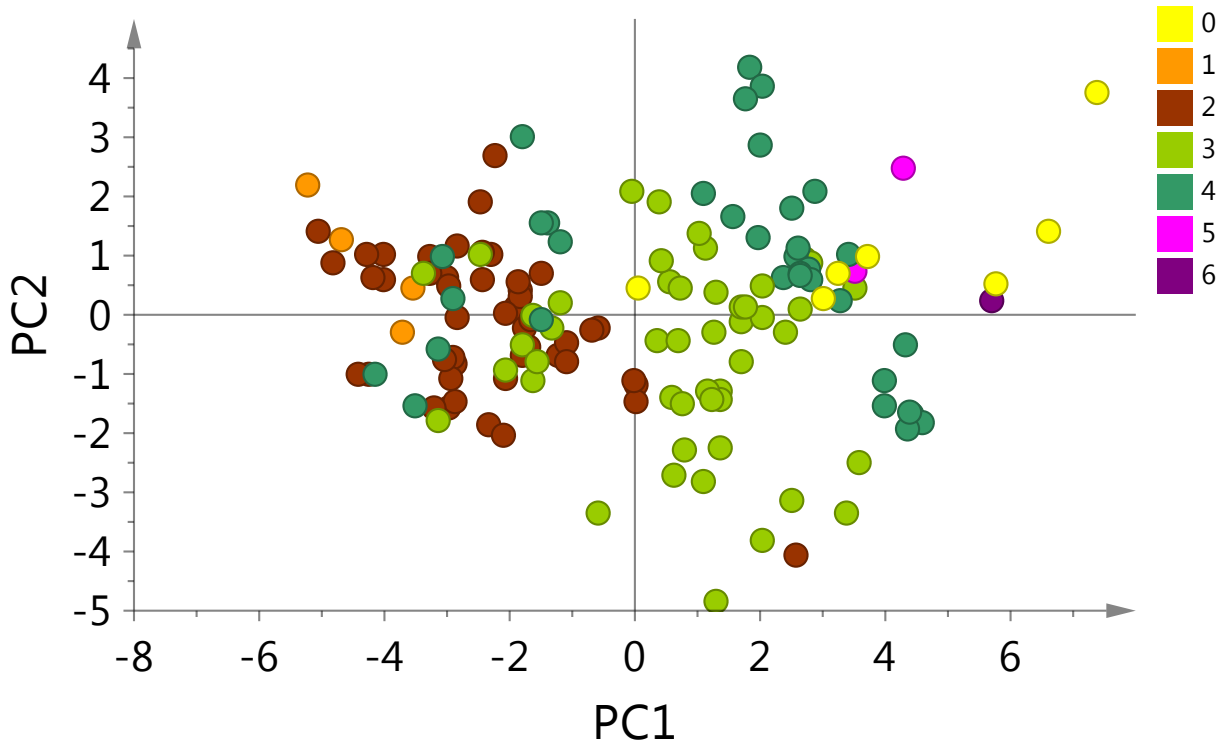


Figure S21: Figure 2 coloured by bridge length.

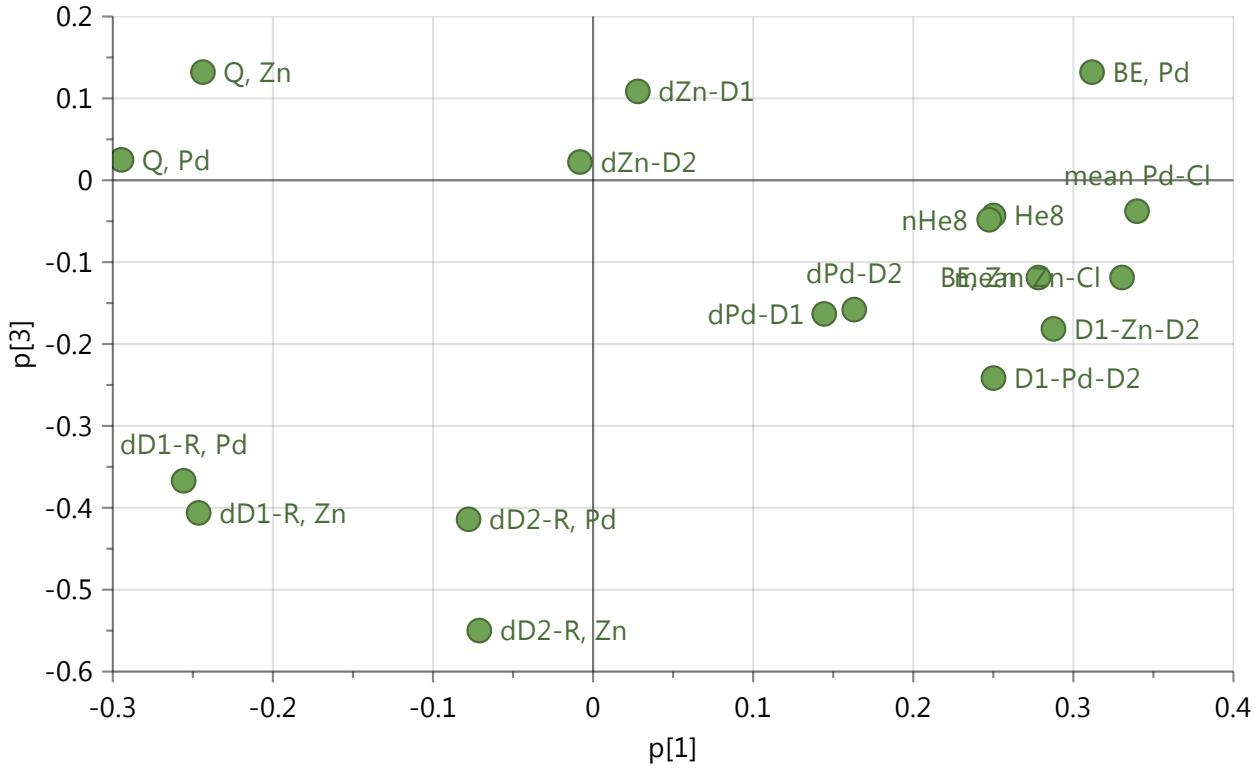


Figure S22: Loadings plot, PC1 and 3, C, N and O donor ligands.

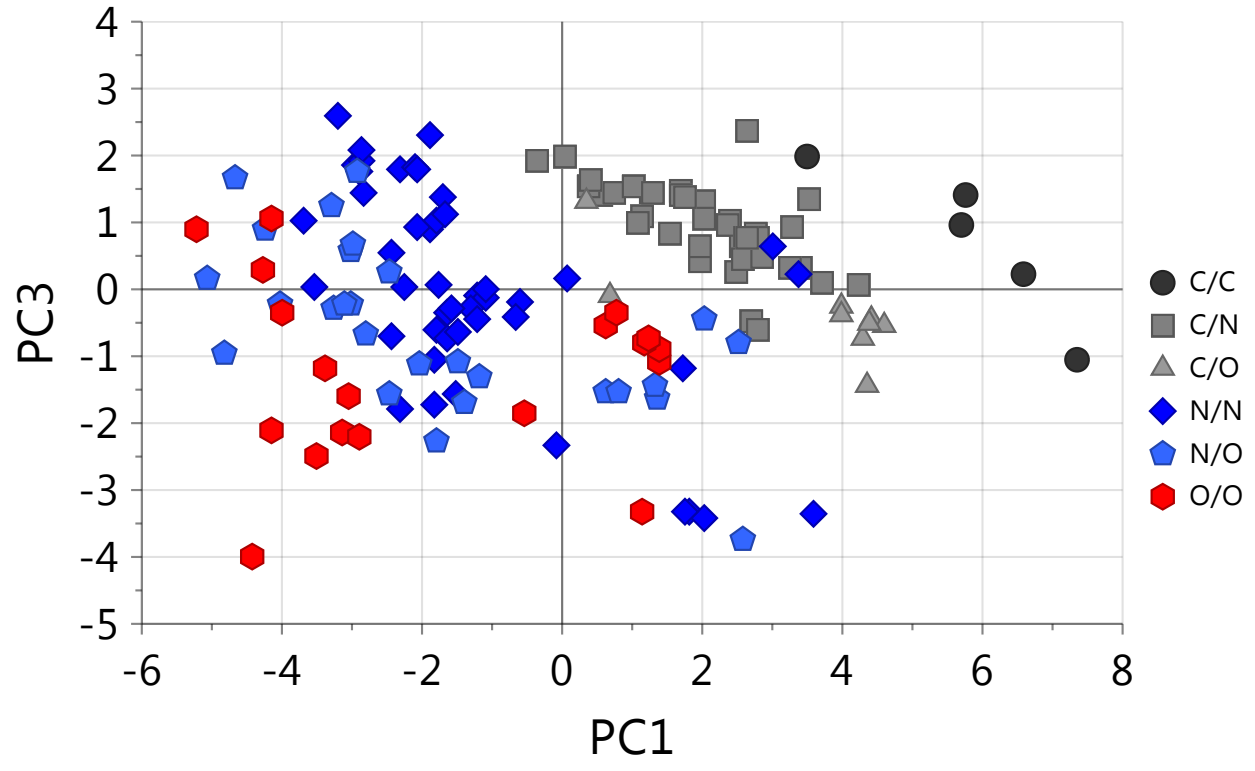


Figure S23: Principal component score plot for PCs 1 and 3, C, N and O donor ligands.

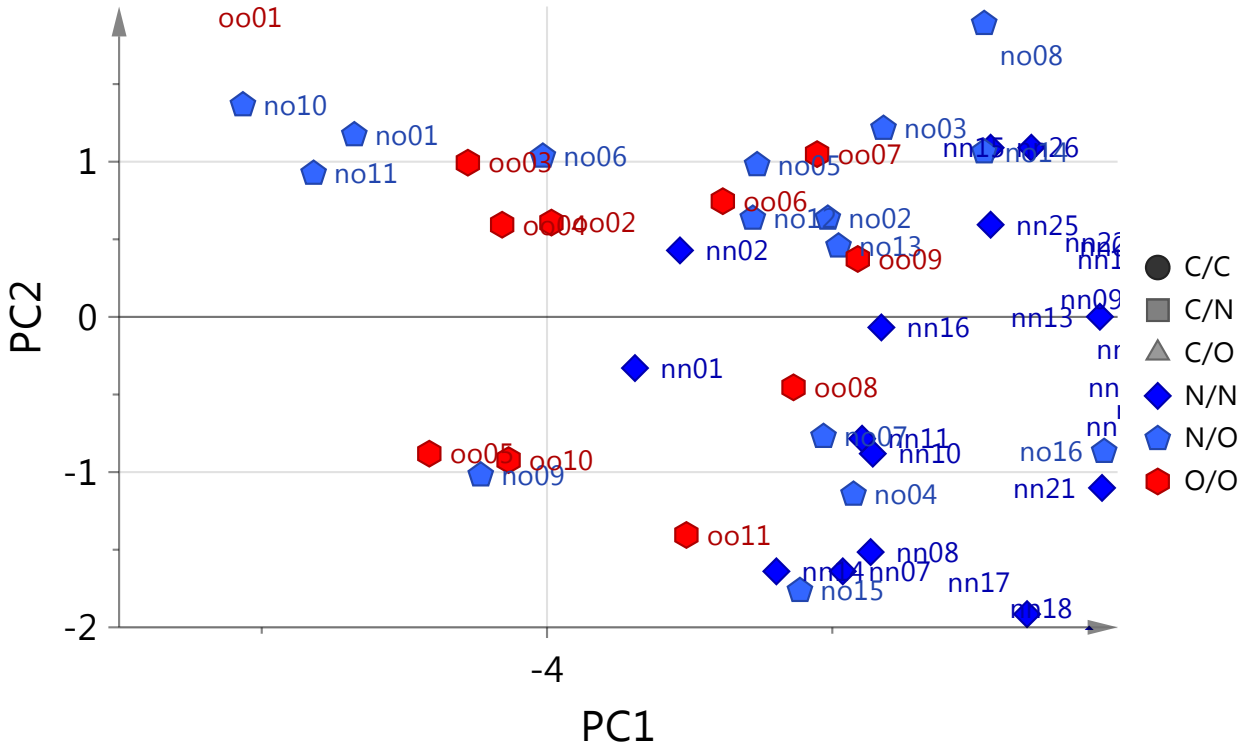


Figure S24: Enlarged version of Figure 3.

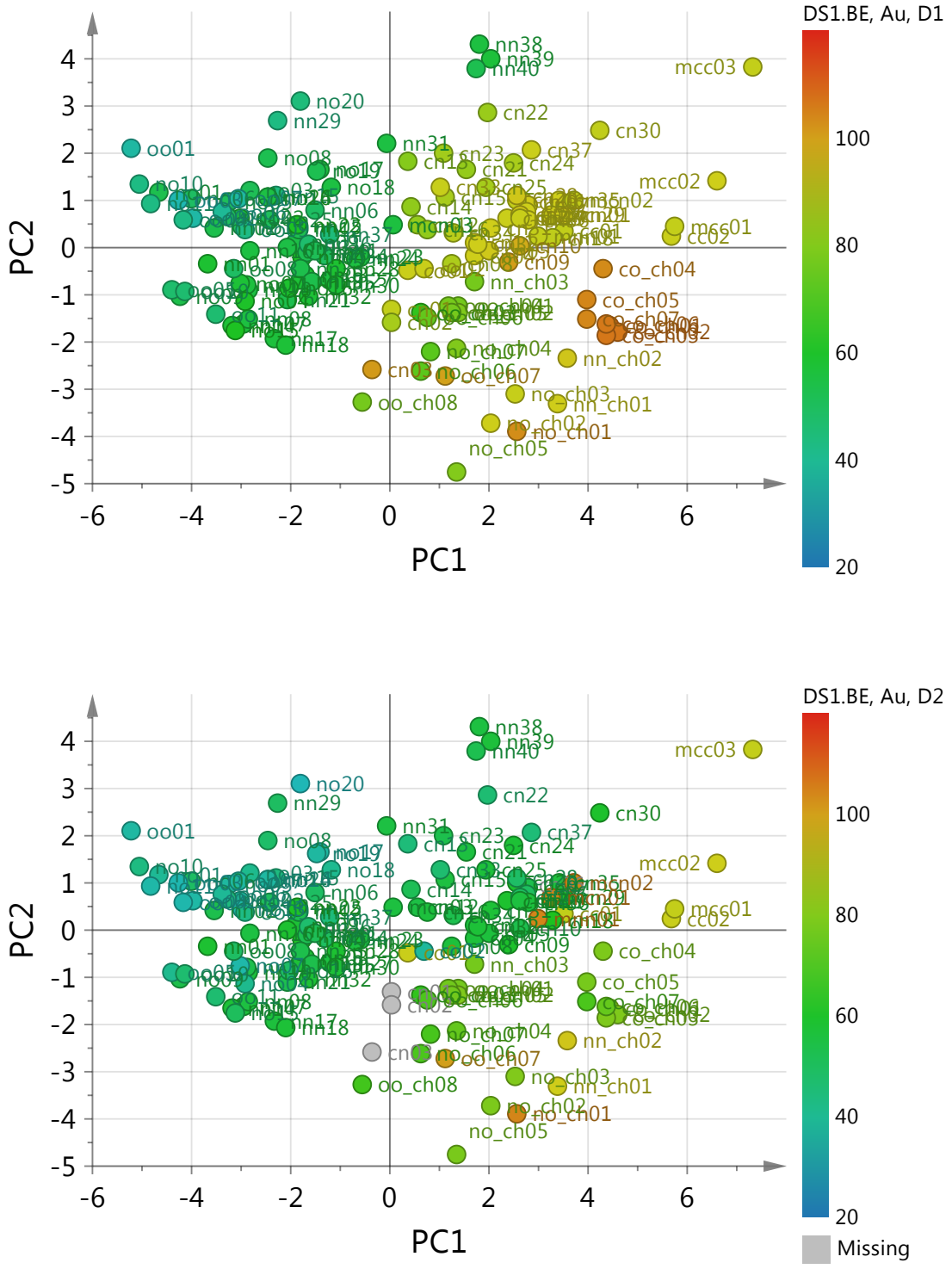


Figure S25: Figure 4 with all labels.

Transferability to other metal complexes

Our work to date, albeit dominated by P,P donor ligands, suggests that the full set of parameters is likely to have some transferability to other coordination environments, such as square-based pyramidal, trigonal bipyramidal, piano stool, and octahedral, while the expansion to sample these environments directly with additional parameters will introduce computational challenges (evaluating equatorial-equatorial vs. equatorial-axial coordination in trigonal bipyramidal geometries (and, perhaps more difficult, in square-based pyramidal ones, as well as their distortion towards trigonal bipyramidal), steric clashes in octahedral complexes affecting conformer preferences/convergence and the interactions/meshing between the ligands of interest and the remaining coordination sphere for all) that are not necessarily justified by achieving improved models. As such, the square planar complex seems to capture the strongest electronic interactions in octahedral complexes (*trans* influence) while the two steric parameters (He_8_wedge and nHe_8) capture steric clashes that would arise from *cis* ligands in this geometry, rigidly in the former and allowing for ligand flexibility in response to such clashes in the latter. Between them, some of the features of a more hindered coordination environment are thus captured. The $ZnCl_2$ complex does not exert a strong coordination preference from the metal centre and sees only limited interactions with the chloride ligands. It should thus be a better model for unhindered coordination environments. Between them, then, some transferability can be achieved.