nature computational science

Article

https://doi.org/10.1038/s43588-024-00616-5

Directional multiobjective optimization of metal complexes at the billion-system scale

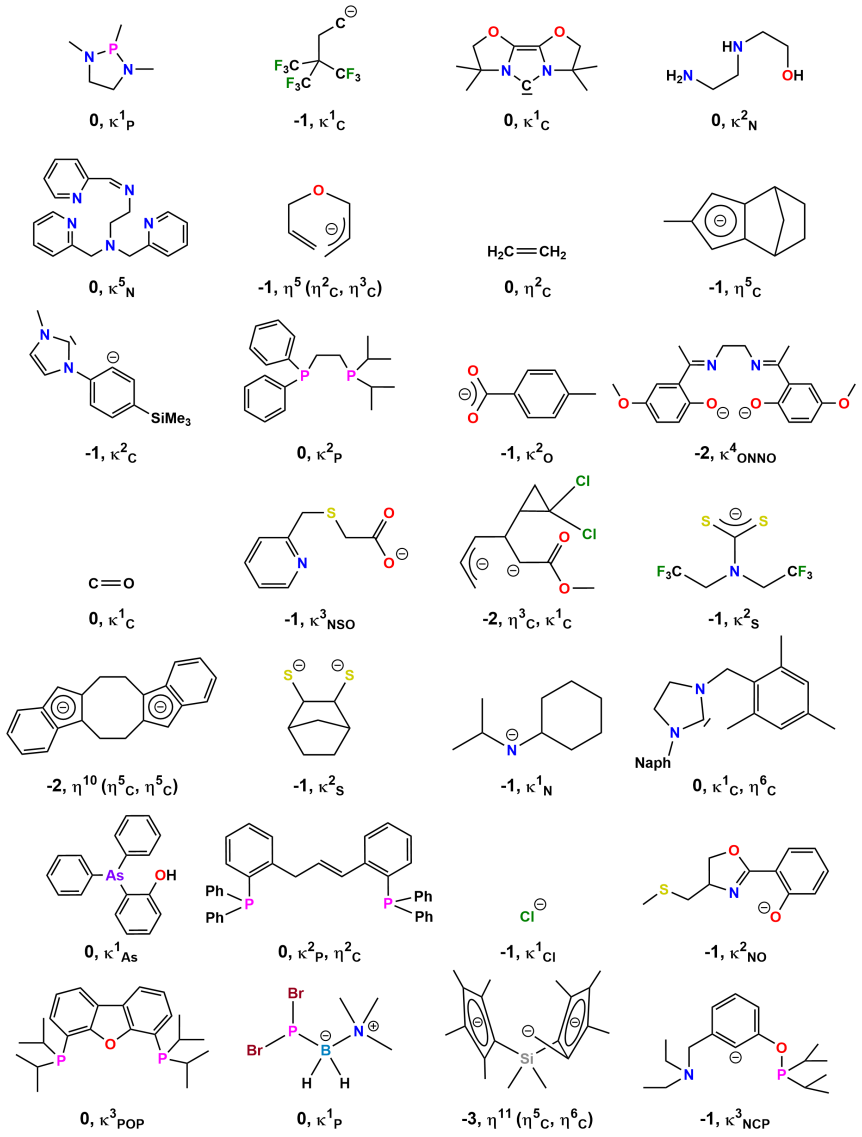
In the format provided by the authors and unedited

Table of Contents

Supplementary Section 1. tmQMg-L dataset

Random ligand samples

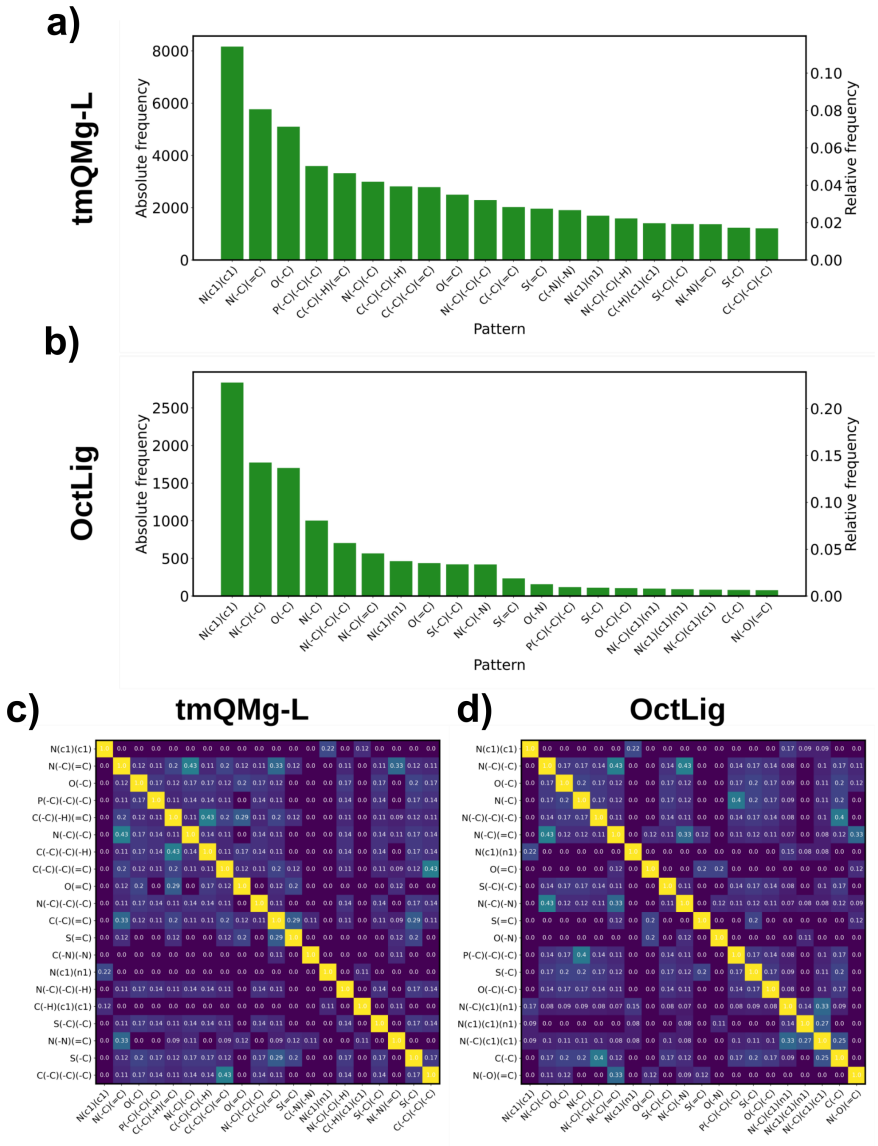
The random sampling over the tmQMg-L ligand dataset shown in Supplementary Figure 1 contains twenty-eight ligands belonging to these twelve different categories: phosphines, alkyls, carbenes, chelating amines, allyls, olefins, arenes, carboxylates, amidos, arsinines, halides, and pincers.



Supplementary Figure 1: Structural diversity in the tmQMg-L dataset. Random sample of 28 ligands, each labeled with its charge, followed by the denticity (κ^n_E) and hapticity (η^m_E), where n and m are the number of metal-bound atoms of element(s) E, respectively.

Metal coordination environments

The histograms in Supplementary Figure 2 show the absolute and relative frequencies of the SMARTS patterns representing the 20 most common coordination environments in the tmQMg-L and OctLig[1] ligand datasets, as well as the corresponding heatmaps of the pairwise Tanimoto coefficients. For tmQMg-L, the five most common environments, in decreasing order of popularity, can be assigned to aromatic N-heterocyles, imines, alcohols (and alkoxides), aliphatic phosphines, and olefine ligands. Most of these environments also appear in the OctLig dataset, though with significantly different relative frequencies in most cases. Whereas the Tanimoto heatmaps suggest that both datasets are highly diverse, with most coefficients taking values < 0.2, the SMARTS histograms show that the tmQMg-L dataset is more balanced, with a softer decay of the frequencies, which, in absolute terms, are generally significantly larger than those of the smaller OctLig dataset.



Supplementary Figure 2: Chemical diversity in the tmQMg-L and OctLig datasets. Frequencies of the 20 most popular coordination environments in the tmQMg-L (a) and OctLig (b) datasets, and heatmaps showing the associated pairwise Tanimoto coefficients for the tmQMg-L (c) and OctLig (d) datasets. The coordination environments are encoded with SMARTS patterns, in which the first atom is the metal-bound.

Special cases in charge and metal coordination assignment

In the manual inspection of the 500 ligands selection, we found only one systematic error in the assignment of the charges from the NBO Lewis structure; *i.e.* several η^2 and η^4 ligands containing the neutral $(R)_2C{=}C(R)_2$ and $(R)_2C{=}C(R){-}C(R){=}C(R)_2$ moieties were incor-

rectly assigned a charge of –2e. The manual inspection of 100 randomly selected examples containing these two moieties revealed that 92% and 96% of them, respectively, had to be corrected to a neutral charge. This correction was systematically applied to all tmQMg-L ligands falling into these two categories. The ligands for which the bonding to the metal center was dominated by 3C NBOs were excluded.

It should be noted that, depending on several factors, like the synthesis of the TMCs and the nature of their metal centers, there are ligands for which the charge can take different values and the coordination to the metal can occur in different modes. For example, regarding the charge, the O_2 ligand can be either neutral (*i.e.* dioxygen ligand) or anionic (*i.e.* superoxo ${O_2}^{-1}$ and peroxo ${O_2}^{-2}$). Regarding the coordination mode, whereas the superoxo often coordinates in κ^1 fashion, the peroxo prefers the η^2. Further, these two factors are not necessarily related. For example, aliphatic ligands with a single carboxylate functional group will have a unique charge of -1e, and yet they may coordinate to the metal in either κ^1 or κ^2 fashion. For this type of ligands, the tmQMg-L dataset may contain either one or multiple charge and coordination variants. Using the fingerprints provided, these ligands can be easily distinguished and found in the dataset and, if needed for a given task, their charge and coordination mode can be further diversified.

Stereoelectrochemical fingerprints

Supplementary Table 1: Systematic list of all features included in the stereoelectrochemical fingerprints of the tmQMg-L ligand dataset. NBO = Natural Bond Orbital theory; NatQG = Natural Quantum Graph;[2] DFT1 = PBE/def2SVP optimization; DFT2 = PBE0/def2TZVP single-point.

Chemical properties. General				
Label	Definition	Units	Method/Software	Structure(s)
F	Chemical formula	—	Hill system	—
M	Molecular mass	Da	—	—
q_L	Molecular charge	e	NBO	—
N_A	Total number of atoms	—	—	—
$N_{A,E}$	Number of atoms by element	—	—	—
N_e	Number of electrons	—	—	—
Pop	Occurrences in tmQMg	—	—	$L_{\in TMC}$
Chemical properties. Coordination mode				
Label	Definition	Units	Method/Software	Structure
N_{MB}	Number of metal-bound atoms	—	NatQG	$L_{\in TMC}$
κ^n	Denticity order	—	NatQG	$L_{\in TMC}$
κ^n_E	κ^n by element	—	NatQG	$L_{\in TMC}$
η^n	Hapticity order	—	NatQG	$L_{\in TMC}$
η^n_E	η^n by element	—	NatQG	$L_{\in TMC}$
Cheminformatics descriptors				
$SMILES$	SMILES string	—	RDKit	—
MFP	Morgan fingerprints	—	RDKit	—
$logP$	Octanol/water partition coefficient	—	RDKit	—
$N_{R,Al}$	Num. of aliphatic rings	—	RDKit	—
$N_{R,Ar}$	Num. of aromatic rings	—	RDKit	—
$N_{R,Sat}$	Num. of saturated rings	—	RDKit	—
N_{RB}	Num. of rotatable bonds	—	RDKit	—
Electronic properties				
Label	Definition	Units	Method/Software	Structure
α	Polarizability	Bohr^3	DFT1	L_{free}
μ	Dipole moment	D	DFT1, DFT2	L_{free}, $L^*_{\in TMC}$
ν	Largest vibrational frequency	cm^{-1}	DFT1	L_{free}
ϵ	HOMO-LUMO gap	Ha	DFT1, DFT2	L_{free}, $L^*_{\in TMC}$
E^{MB}_{HOMO}	Metal-bound HOMO energy	Ha	DFT1, DFT2	L_{free}, $L^*_{\in TMC}$
S^{MB}_{HOMO}	Metal-bound HOMO symmetry	—	DFT1, DFT2	L_{free}, $L^*_{\in TMC}$
E^{MB}_{LUMO}	Metal-bound LUMO energy	Ha	DFT1, DFT2	L_{free}, $L^*_{\in TMC}$
S^{MB}_{LUMO}	Metal-bound LUMO symmetry	—	DFT1, DFT2	L_{free}, $L^*_{\in TMC}$
Steric properties				
Label	Definition	Units	Method/Software	Structure
V	Molecular volume	Å^3	DFT1, DFT2	L_{free}, $L^*_{\in TMC}$
$\frac{I_1}{I_3}, \frac{I_2}{I_3}$	Principal moments of inertia ratios[3]	—	DFT1, DFT2	L_{free}
A^{SAS}	Solvent accessible surface area (SAS)[4,5]	Å^2	Morfeus	L_{free}, $L^*_{\in TMC}$
V^{SAS}	Volume within SAS[4,5]	Å^3	Morfeus	L_{free}, $L^*_{\in TMC}$
V_{Bur}	Buried volume[6]	%	Morfeus	$L^*_{\in TMC}$
θ°	Exact cone angle[7,8]	°	Morfeus	$L^*_{\in TMC}$
Ω	Solid angle[9,10]	sr	Morfeus	$L^*_{\in TMC}$
Θ	Solid cone angle[9,10]	°	Morfeus	$L^*_{\in TMC}$
G	G parameter[9,10]	—	Morfeus	$L^*_{\in TMC}$

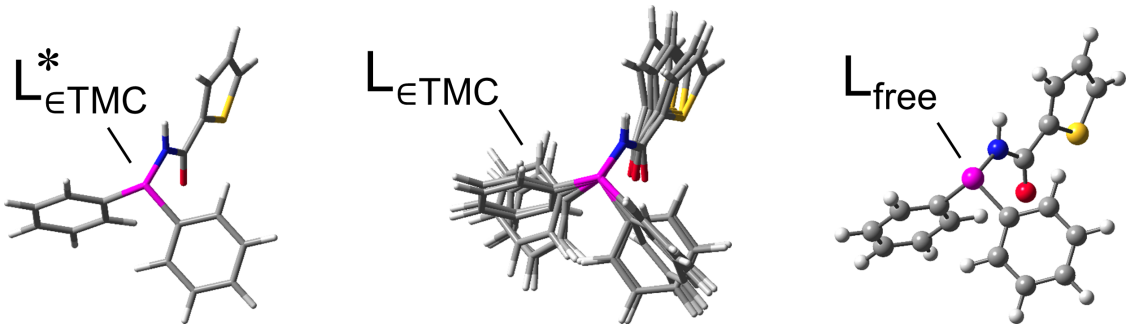
RDKit sanitization

When loading molecules into RDKit from a SMILES string, a "sanitization" process is performed to ensure chemical correctness before properties are computed. This includes a check for the explicit valence of each atom of the molecule, which fails if a deviation from the values allowed is detected.[11] In the tmQMg-L ligand dataset, however, these constraints are often times violated due to the nature of the ligands and the fact that they are extracted from TMCs; *e.g.* elements from groups 13–17 can be hypervalent in their lowest-energy configuration. Therefore, the default RDKit sanitization was deactivated and a custom procedure was adopted. In order to allow for the violation of the valency constraint, we made use of the individual sanitization flags provided by RDKit. The following code snippet shows our custom sanitization procedure:

```
from rdkit import Chem
# read molecule to RDKit
mol = Chem.MolFromSmiles(smi, sanitize=False)
# custom sanitization
mol.UpdatePropertyCache(strict=False)
Chem.SanitizeMol(
    mol,
    Chem.SanitizeFlags.SANITIZE_SETAROMATICITY|
    Chem.SanitizeFlags.SANITIZE_SETCONJUGATION|
    Chem.SanitizeFlags.SANITIZE_SETHYBRIDIZATION|
    Chem.SanitizeFlags.SANITIZE_CLEANUP
)
```

where `smi` refers to the variable containing the SMILES string of the ligand.

Ligand geometry categories



Supplementary Figure 3: Ligand geometry categories. The $L^*_{\in TMC}$, $L_{\in TMC}$, and L_{free} geometry sets for an example phosphine ligand in tmQMg-L. Element color code: Violet (P), blue (N), grey (C), red (O), yellow (S), and white (H).

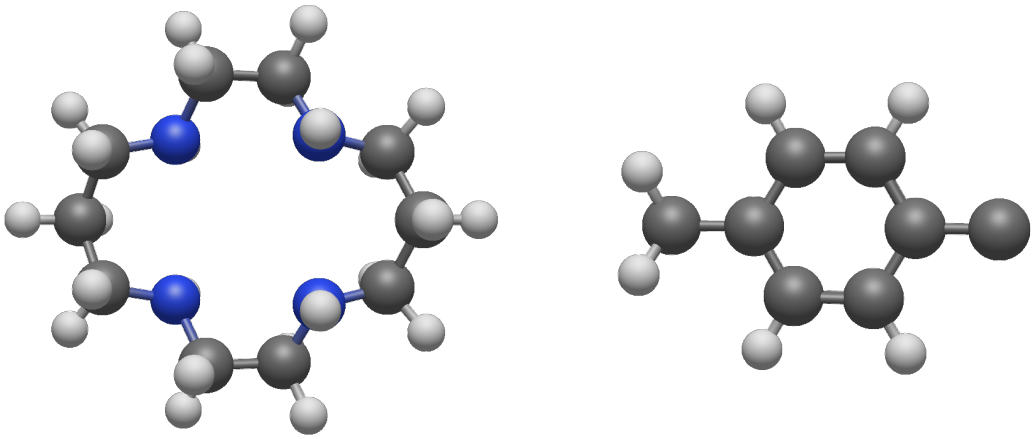
Steric descriptor failures

In total, there were 12 ligands for which the calculation of the steric descriptors with the MORFEUS package[12] failed. More precisely, the calculation of the exact cone angle failed in these cases:

- Seven monodentate ligands which fully surrounded the metal center. This should yield a cone angle of 360 degrees but caused a mathematical error during the calculation.
- Five monodentate ligands for which a linear coordination geometry left all atoms, except the metal-bound, outside of the sphere defined by the covalent radii of the metal center, also causing an error in the calculation of the cone angle.

By definition, in the latter case, the calculation of the solid angle, solid cone angle, and G parameter also failed. The corresponding cells in the descriptors *.csv* files were left empty. Examples of these structures are shown in Supplementary Figure 4.

Furthermore, the calculation of principal moments of inertia with Gaussian[13] could not be extracted due to a bug in the printing of the data. In these cases their ratios could not be computed and were left blank in the dataset.



(a) Surrounded metal center **(b)** Linear coordination

Supplementary Figure 4: Non-defined cone angles. Examples of ligands for which the calculation of the exact cone angle failed. Element color code: Blue (N), grey (C), and white (H).

Definition of unique ligands

We define unique ligands in terms of their chemical graph and metal-coordination mode. This means that one chemical graph can give rise to multiple unique ligands if it occurs with different coordination modes in the tmQMg dataset. We determined this by looping over all extracted ligands and checking for graph isomorphism between them. If the chemical graphs of two ligands are isomorphic to each other, they correspond to the same structural isomer and we subsequently compare their coordination modes in order to determine whether they are identical. Ligands with isomorphic graphs *and* the same coordination mode are considered duplicates of each other. Noteworthy, chemical graph isomorphism does not distinguish stereoisomers. However, these isomers, *e.g.* cis/trans configurations, tend to yield different coordination modes (*e.g.* a cis bidentate ligand becoming monodentate in the trans form), which are differentiated by our algorithm. For conformers and, in general, any collection of isomorphic graphs having the same coordination mode, we took the ligand yielding the lowest energy. Overall, our approach strictly excluded any ligand duplicates in the final tmQMg-L dataset.

Weisfeiler-Lehman graph hashes

The Weisfeiler-Lehman graph hashes[14] were computed for all ligands using the Python package `networkx`.[15] Four variations were computed differing in what graph information was considered for generating the hashes:

- **Plain:** Only the connectivity was taken into account.
- **Atom attribution:** Connectivity as well as the atom elements were taken into account.
- **Bond attribution:** Connectivity as well as the bond orders were taken into account.
- **Atom and bond attribution:** Connectivity as well as both the atom elements and bond orders were taken into account.

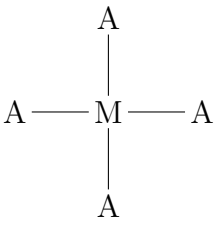
All hashes were computed with 3 neighbor aggregation iterations, meaning that neighbors up to 3 bonds away were considered.

Supplementary Section 2. 1.37M chemical space

ML_4 square planar TMC combinatorics

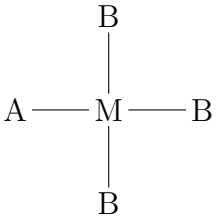
For a square planar geometry with four monodentate ligands, there are seven different coordination patterns. Each pattern gives rise to a different number of unique TMCs depending on the number of ligands, n. Some of the patterns exhibit equivalencies due to symmetry that need to be accounted for. Additionally, in this work, we restricted the overall charge of the TMCs to the range $\{-1, 0, 1\}$, which further reduces the number of possible complexes. We considered only neutral and monoanionic ligands and the +II oxidation state for the palladium center. Therefore, we express the number of unique TMCs per pattern in terms of the number of neutral (n_n) and anionic (n_a) ligands.

Below, we list the number of unique TMCs considering the aforementioned constraints for each pattern. Furthermore, the associated symmetries and the resulting reduction in the number unique TMCs are also presented.



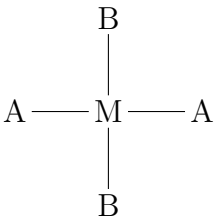
- 0 unique members as the charge constraint is violated for all possible arrangements.
- No equivalencies.

Supplementary Figure 5: AAAA pattern. Square planar geometry with all four ligands (A) equal in the $M(A)_4$ formulation. M = Metal center.



- $2n_n n_a$ unique members.
- no equivalencies.

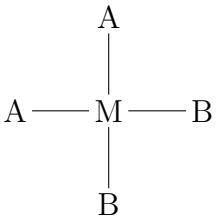
Supplementary Figure 6: ABBB pattern. Square planar geometry combining two different ligands (A and B) in the $M(A)(B)_3$ formulation. M = Metal center.



- $2n_n n_a$ unique members.
- 1 equivalency:
  - (rotation, 2) XYXY = YXYX

  $\rightarrow$ corresponds to a prefactor of $\frac{1}{2}$

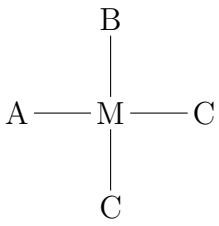
Supplementary Figure 7: ABAB pattern. Trans square planar geometry combining two different ligands (A and B) in the $M(A)_2(B)_2$ formulation. M = Metal center.



- $2n_n n_a$ unique members.
- 1 equivalency:
  - (rotation, 2) XXYY = YYXX

  $\rightarrow$ corresponds to a prefactor of $\frac{1}{2}$

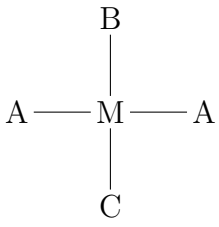
Supplementary Figure 8: AABB pattern. Cis square planar geometry combining two different ligands (A and B) in the $M(A)_2(B)_2$ formulation. M = Metal center.



- $3n_n n_a(n_a-1) + 3n_a n_n(n_n-1)$ unique members.
- 1 equivalency:
  - (reflection, 2) XYZZ = YXZZ

  $\rightarrow$ corresponds to a prefactor of $\frac{1}{2}$

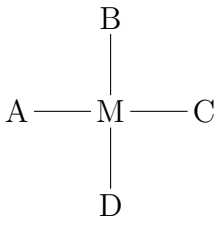
Supplementary Figure 9: ABCC pattern. Cis square planar geometry combining three different ligands (A, B, and C) in the M(A)(B)$(C)_2$ formulation. M = Metal center.



- $3n_n n_a(n_a-1) + 3n_a n_n(n_n-1)$ unique members.
- 1 equivalency:
  - (rotation, 2) XYXZ = XZXY

  $\rightarrow$ corresponds to a prefactor of $\frac{1}{2}$

Supplementary Figure 10: ABAC pattern. Trans square planar geometry combining three different ligands (A, B, and C) in the M$(A)_2$(B)(C) formulation. M = Metal center.



- $4n_n(n_n-1)(n_n-2)n_a + 4n_a(n_a-1)(n_a-2)n_n + 6n_n(n_n-1)n_a(n_a-1)$ unique members.
- 3 equivalencies:
  - (rotation, 4) WXYZ = XYZW
  - (reflection, 2) WXYZ = ZYXW

  $\rightarrow$ corresponds to a prefactor of $\frac{1}{4\cdot 2}$

Supplementary Figure 11: ABCD pattern. Square planar geometry combining four different ligands (A, B, C and D) in the M(A)(B)(C)(D) formulation. M = Metal center.

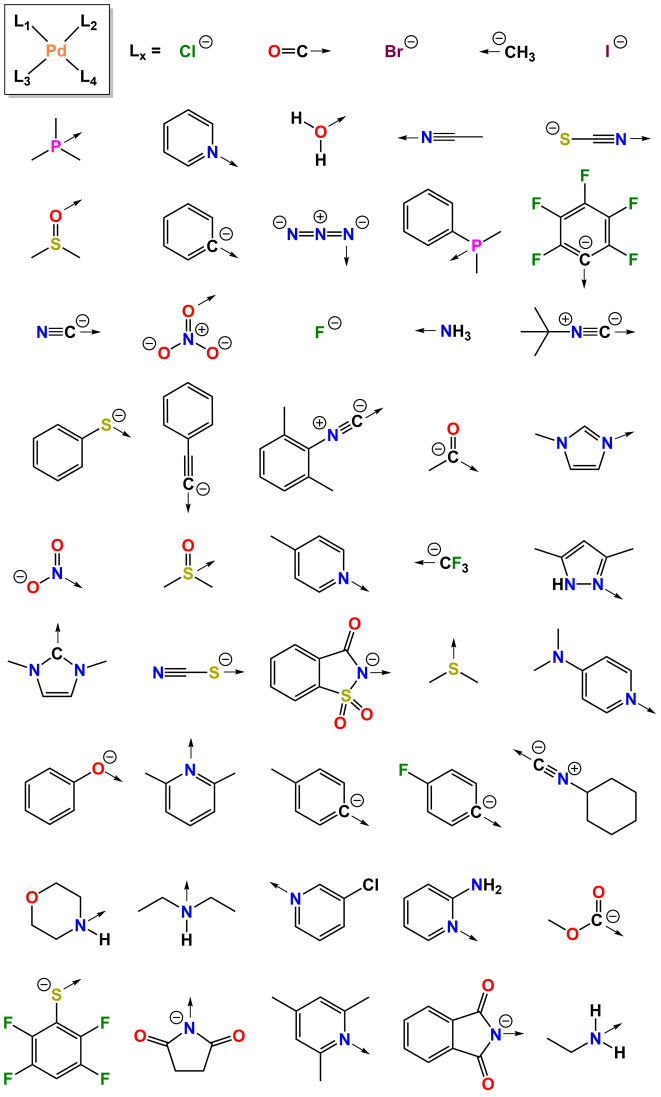
Regarding the distinction between rotation and reflection symmetries, we have to consider that the ligands used are not necessarily symmetric. This means that both the ordering of the ligands in the TMC and their position relative to each other are relevant. In other words, in the ABCD pattern, the member WXYZ might not be equivalent to the member ZYXW if some of the ligands {W,X,Y,Z} are non-symmetric. To account for this, we excluded all reflection symmetries. This changes the symmetry pre-factors of the ABCC and ABCD patterns to 1 and $\frac{1}{4}$, respectively.

After adding the formulas of the seven patterns and simplifying the resulting expression, the number of unique TMCs in this systems is given by

$$n_a n_n^3 + n_a^3 n_n + \frac{3}{2} n_a^2 n_n^2 + \frac{1}{2} n_a n_n \tag{1}$$

50-ligand pool

The 50 ligands used in the generation of the 1.37M chemical space are shown in Supplementary Figure 12.



Supplementary Figure 12: The 50-ligand pool. Ligands appear ordered by popularity from top-left to bottom-right. In the polyatomic ligands, the arrow signals the metal-coordinating atom.

Supplementary Section 3. Lighthouse algorithm

The Lighthouse functionality of the PL-MOGA was implemented with the ZEROMASK function shown in algorithm 1,

Algorithm 1 PL-MOGA ZEROMASK function

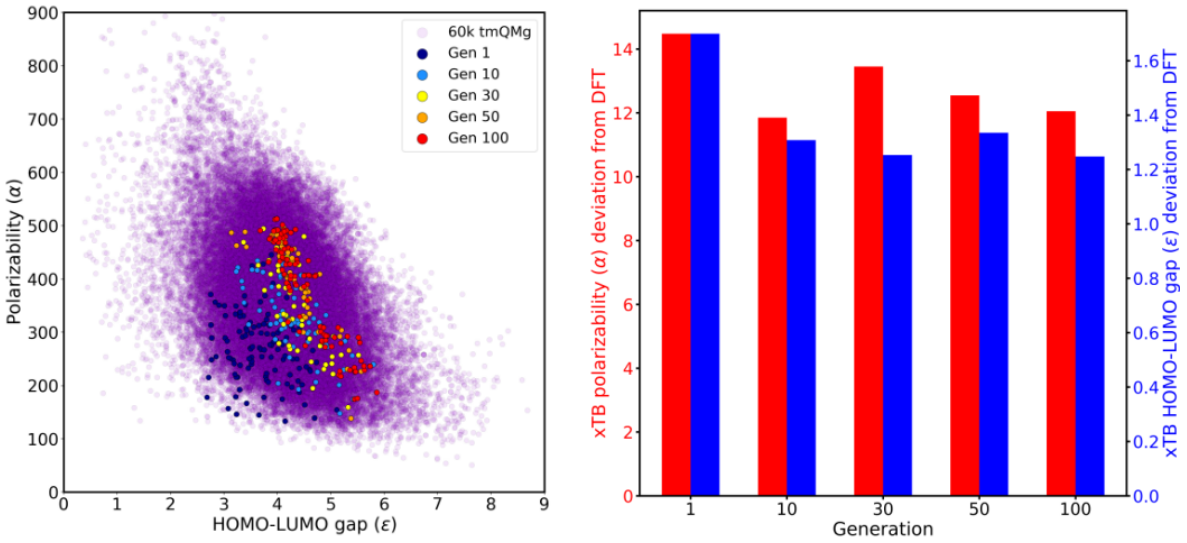
```
function ZEROMASK(x, X, T, S)
    for t in T do
        m_t ← population X_t median
        if x_t < s_t · m_t then
            return 0
        end if
    end for
    return x
end function
```

where x denotes the fitness vector of any query TMC, X denotes the list of fitness vectors of the current population, T denotes the list of t indices of the selected targets, and S denotes the list of corresponding factors s_t used to scale the population median of each target (m_t). ZEROMASK returns the fitness either unchanged or transformed to zero, assuming that this is its lowest possible value for any target. This function is only valid when $dim(T) = dim(S)$; *i.e.* one scaling factor must be provided for each selected target. Numerical experiments showed that the scaling factors should be within the continuous $[0, 1]$ range, where zero corresponds to no masking applied. Scaling values larger than 1 yielded high thresholds with which large portions of the population were mapped to zero fitness, significantly hindering evolution. Further, compared to a GA using a weighted-sum fitness (see Supplementary Section 7), the PL-MOGA generally performed at a higher level, finding more dominating points within more diverse Pareto front solutions, with optimization scope control, and without requiring prior knowledge on the numerical limits of the objectives.

Supplementary Section 4. Estimation of population diversity

In order to quantify diversity, we compute the Tanimoto coefficients of each TMC relative to all other TMCs in any given population and average them to obtain a mean estimate of chemical similarity. Averaging again over the total number of TMCs give us a global Tanimoto coefficient measuring the diversity of the populations evolved in each MOGA generation.

The Tanimoto coefficients are computed with Morgan fingerprints (MFP) derived from SMILES strings. The problem of generating MFPs for TMCs is circumvented by using a string representation consisting of the concatenation of all ligand SMILES with the **.** symbol, which accounts for non-covalent-bound fragments. This representation is permutation-invariant relative to the order in which the ligand SMILES are concatenated and excludes the troublesome fingerprinting of SMILES strings containing metal–ligand bonds. Since the metal is Pd in all TMCs, this approach is still able to capture all relevant chemical (dis-)similarities. The TMC MFPs are generated with a size of 2^{15} bits.



Supplementary Figure 13: DFT fitness. Left-hand side: MOGA (α,ϵ) optimization in the 1.37M space plotted against the 60K tmQMg dataset. All data points were computed at the DFT(PBE/def2SVP//PBE0/def2TZVP) level of theory. Right-hand side: Evolution of the deviation of the GFN2-xTB α and ϵ values relative to the DFT over the MOGA generations. In both plots, the α and ϵ units are Bohr^3 and eV, respectively.

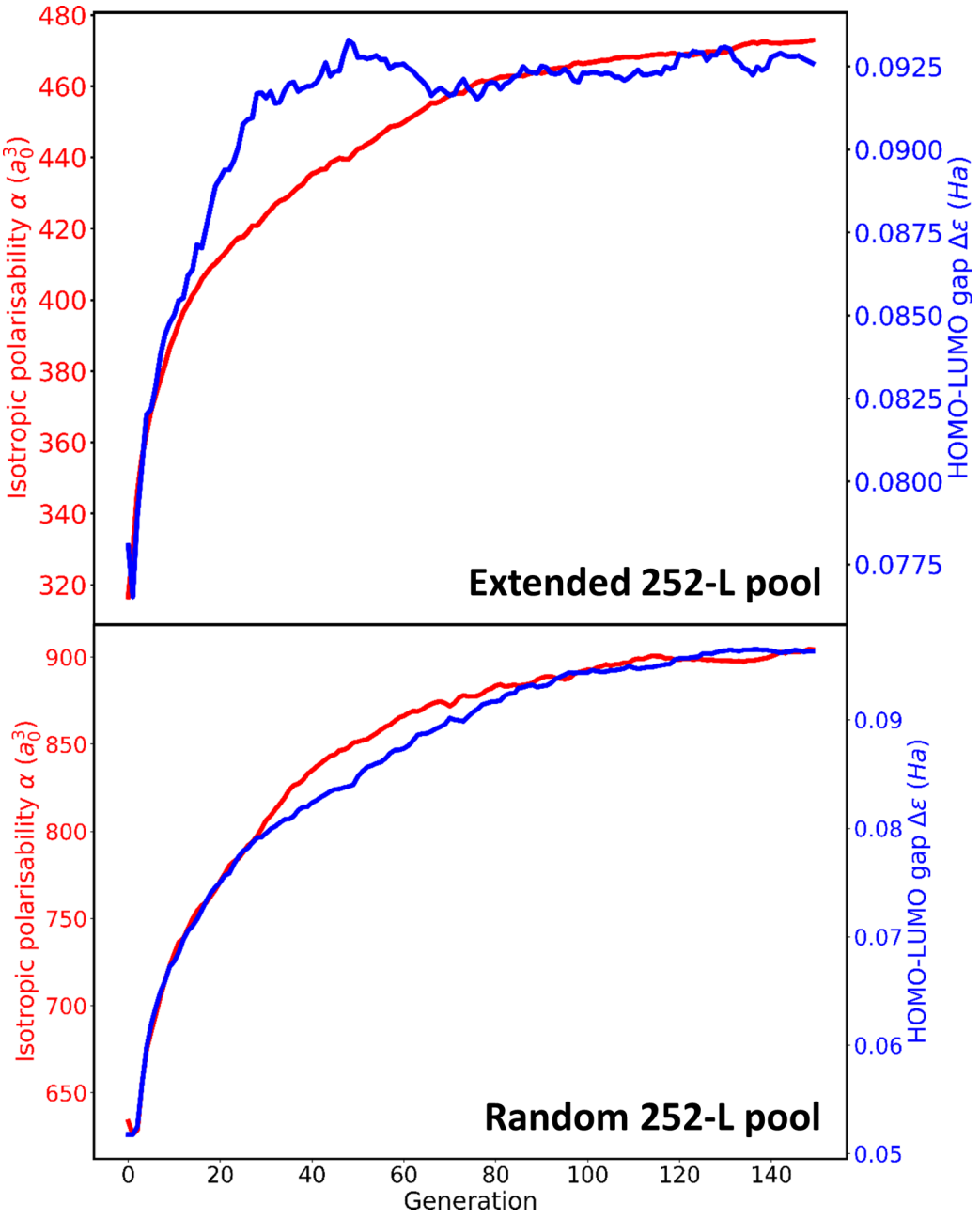
Supplementary Section 5. DFT fitness

The evolution of the hits with the GFN2-xTB fitness was also analyzed after recomputing the target properties with a DFT method. We extracted all TMCs for which the calculation of α and ϵ with xTB was successful in generations 1, 10, 30, 50, and 100. For these TMCs (99% of the total), we recomputed both properties at the PBE/def2SVP (α) and PBE0/def2TZVP (ϵ) levels. The results were plotted against the 60K tmQMg dataset, for which these properties were also available at the same DFT levels (Supplementary Figure 13). This plot showed that 1) there was clear progress towards the DFT tmQMg Pareto front, though less stable than with xTB, in line with the significant differences between these two methods, and 2) there was a gap between the TMC hits and the DFT tmQMg Pareto front, in line with the different sizes of the associated fragment pools; *i.e.*: 1 *vs.* 30, for the metal center, and 50 *vs.* 30K for the ligands. We hypothesized that the average deviation between DFT and xTB would increase by approaching the Pareto front but the opposite trend was seen (Supplementary Figure 13), likely due to the TMCs being constrained into a smaller (α,ϵ) region. These results suggested that an xTB-guided MOGA followed by a DFT recalculation of selected solutions may constitute a robust computational protocol.

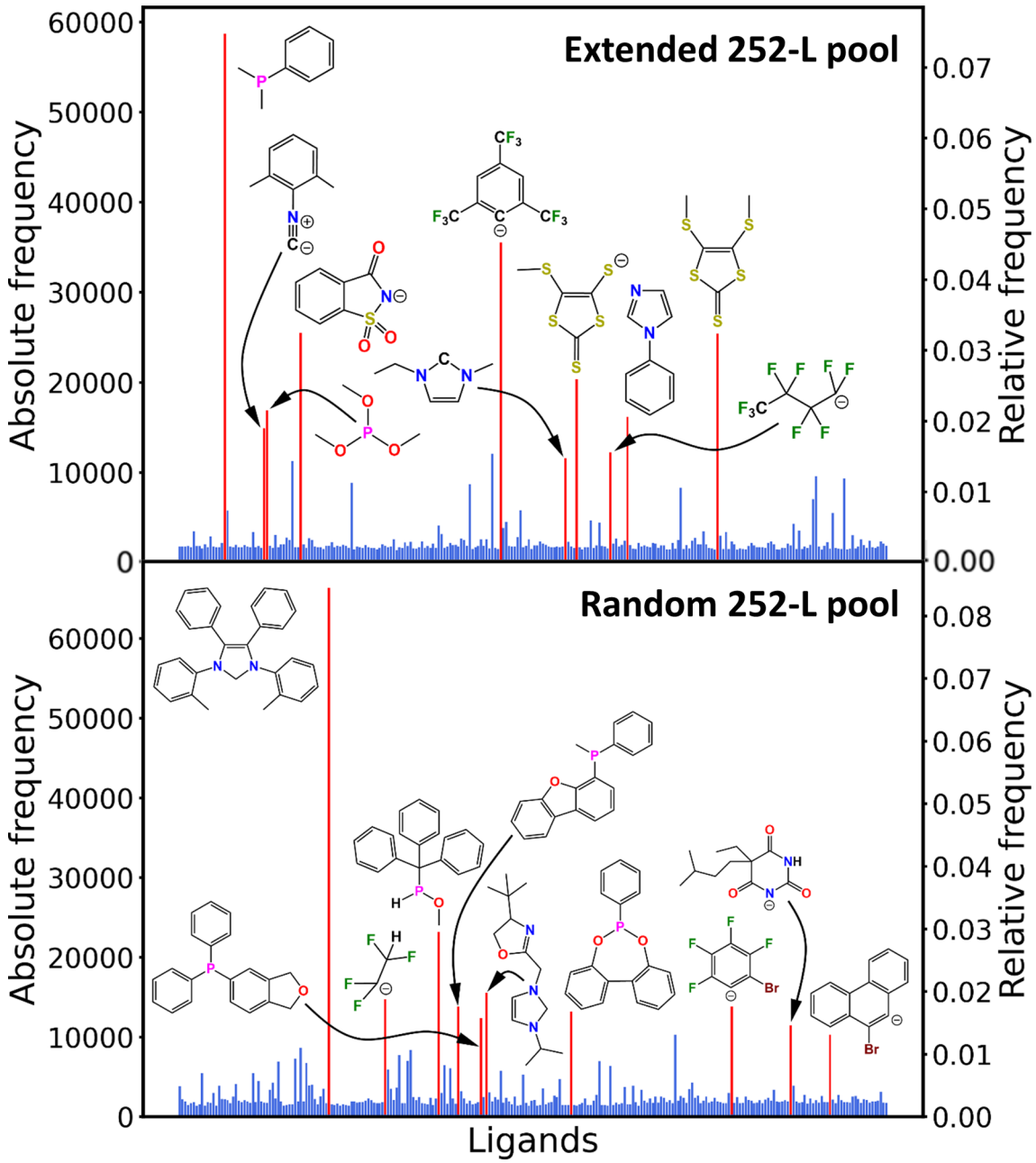
Supplementary Section 6. Convergence and diversity in the billion spaces

For the PL-MOGA optimizations in the billion spaces, we also assessed the convergence of the properties (Supplementary Figure 14), the ligand frequency use (Supplementary Figure 15), and the chemical diversity of the generations evolved (Supplementary Figures 16 and 17) in a way analogous to that followed in the benchmarks taking the 1.37M as the reference. The geometries of four selected TMC hits evolved with the random 252-ligand pool, which, in general, had complex structures with large ligands, were optimized at the DFT(PBE/def2SVP) yielding energy minima with all-real vibrational frequencies in all cases

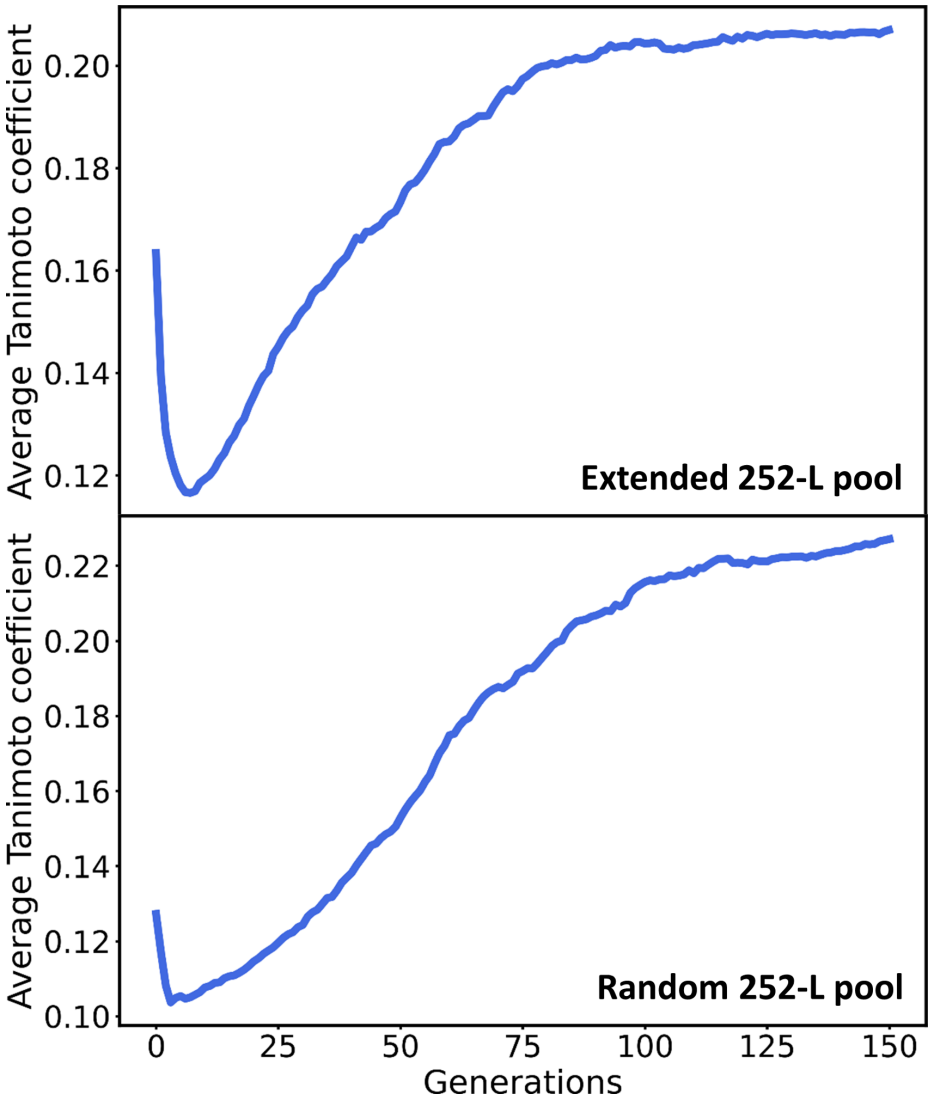
(Supplementary Figure 18). Further, optimization repetitions were also carried out to assess the robustness of the algorithm (Supplementary Figure 19).



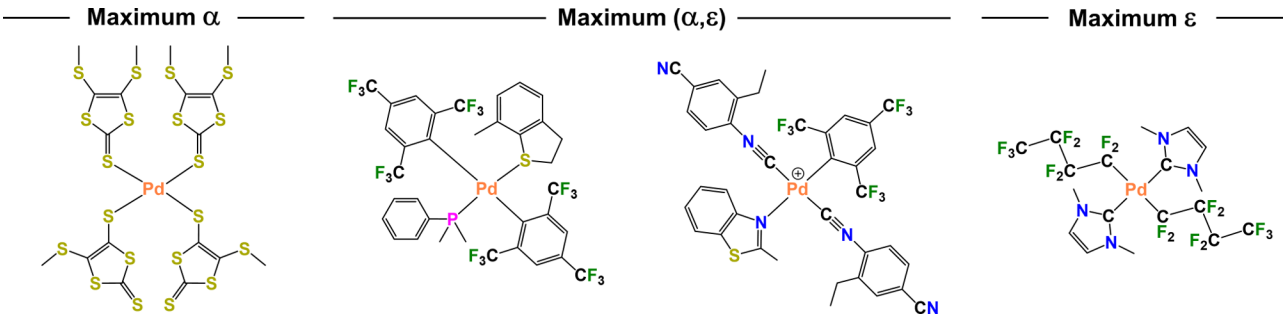
Supplementary Figure 14: Billion scale properties convergence. Evolution of α and ϵ in the unmasked PL-MOGA optimizations carried out in the implicit billion spaces.



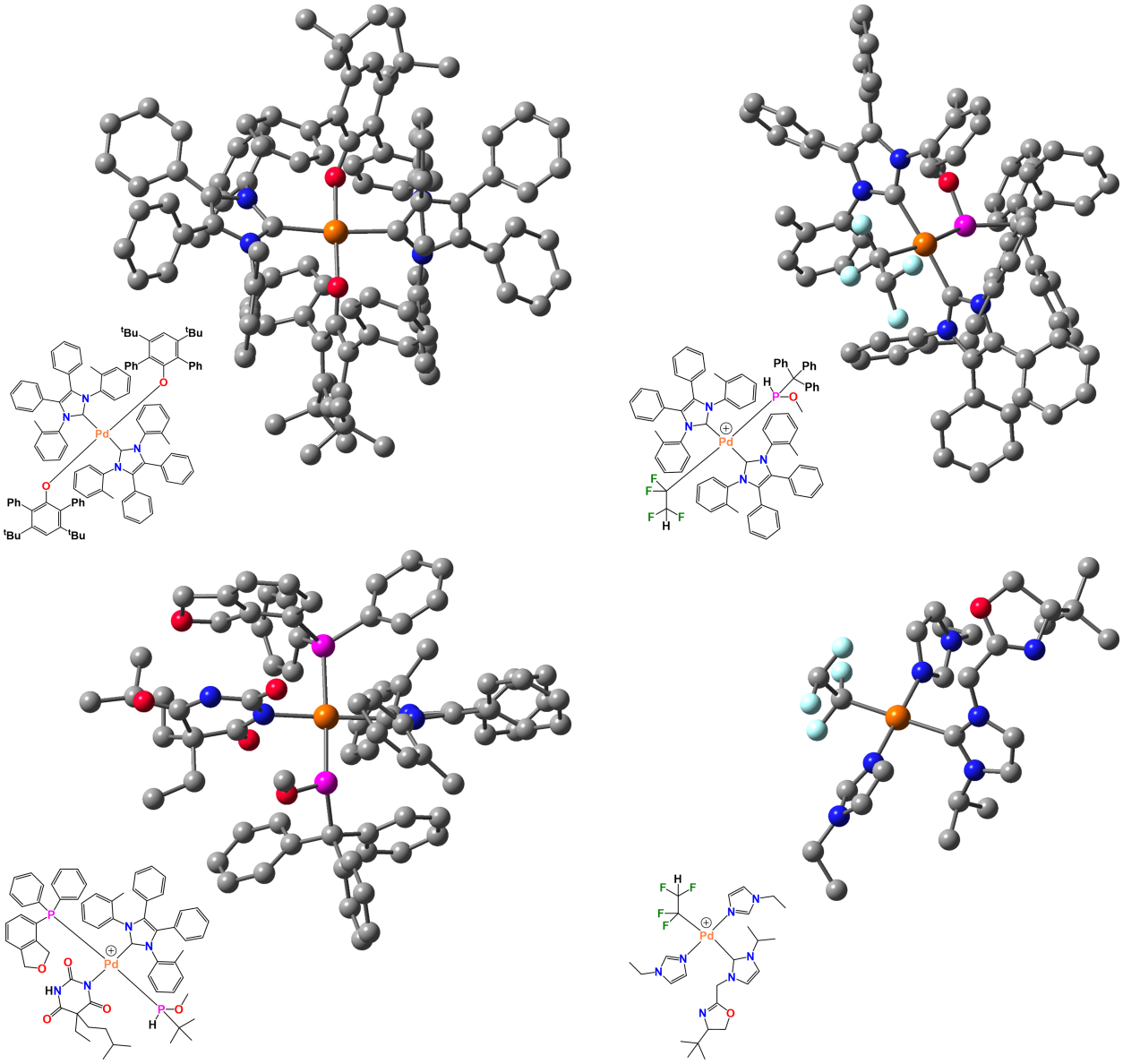
Supplementary Figure 15: Billion scale ligand use frequencies. Ligand use frequencies in the unmasked PL-MOGA optimizations carried out in the implicit billion spaces.



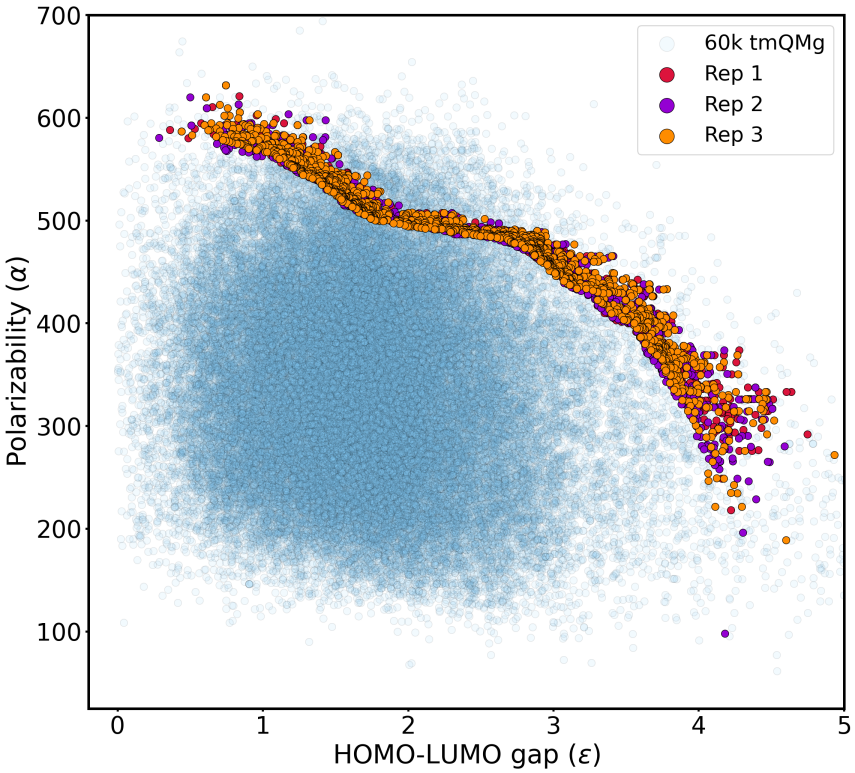
Supplementary Figure 16: Diversity in the billion scale optimizations. Convergence of the average Tanimoto coefficient in the unmasked PL-MOGA optimizations carried out in the implicit billion spaces.



Supplementary Figure 17: Pareto solutions from billion scale optimization. TMC hits from the unmasked PL-MOGA leveraging the extended 252-ligand pool.



Supplementary Figure 18: DFT geometries from billion scale optimization. Geometries fully optimized at the PBE/def2SVP level of selected TMC hits evolved in the unmasked PL-MOGA using the random 252-ligand pool. Each 3D structure is presented together with its 2D chemical drawing. Element color code: Violet (P), blue (N), grey (C), red (O), green (F), and orange (Pd).



Supplementary Figure 19: Optimization dependence on initial population. PL-MOGA repetitions from three different random initial populations in the extended 252-ligand pool billion space.

Supplementary Section 7. Weighted-sum benchmark

For the sake of benchmarking the PL-MOGA method, we used the weighted-sum approach to multiobjective optimization with genetic algorithms (WS-MOGA) as baseline. The calculations were carried out for the reference 1.37M space. In the WS-MOGA, the fitness (F) is calculated as a weighted sum of the different optimization targets:

$$F = \sum_i \beta_i f_i \tag{2}$$

where β_i denotes the weighting factor of target f_i. By choosing different weighting factors, the algorithm can be tuned to prioritize the optimization of one target over the others. In the 1.37M space, the fitness can therefore be expanded as the weighted sum of the polarizability (f_α) and the HOMO-LUMO gap (f_ϵ) with independent weighting factors (β_α and β_ϵ):

$$F = \beta_\alpha f_\alpha + \beta_\epsilon f_\epsilon \tag{3}$$

The weighted-sum approach has the effect of turning the multiobjective optimization problem into a single objective optimization of a composite fitness function. Therefore, Pareto-based ranking and selection schemes are not necessary and standard GA methods can be employed. The parent selection, which in the PL-MOGA was based on the non-dominated front ranking of individuals, becomes a simple roulette wheel selection with probabilities proportional to the fitness values. Similarly, the survivor selection simply takes the n best individuals and discards the rest, where n is the population size of the generations evolved.

A relevant issue with this weighted-sum fitness approach is that the different optimization targets (y) might be of vastly different scales making it difficult to choose appropriate weighting factors. In fact, this is the case for the polarizability and the HOMO-LUMO gap. Therefore, the values were individually normalized to the range $[0, 1]$ using min-max normalization:

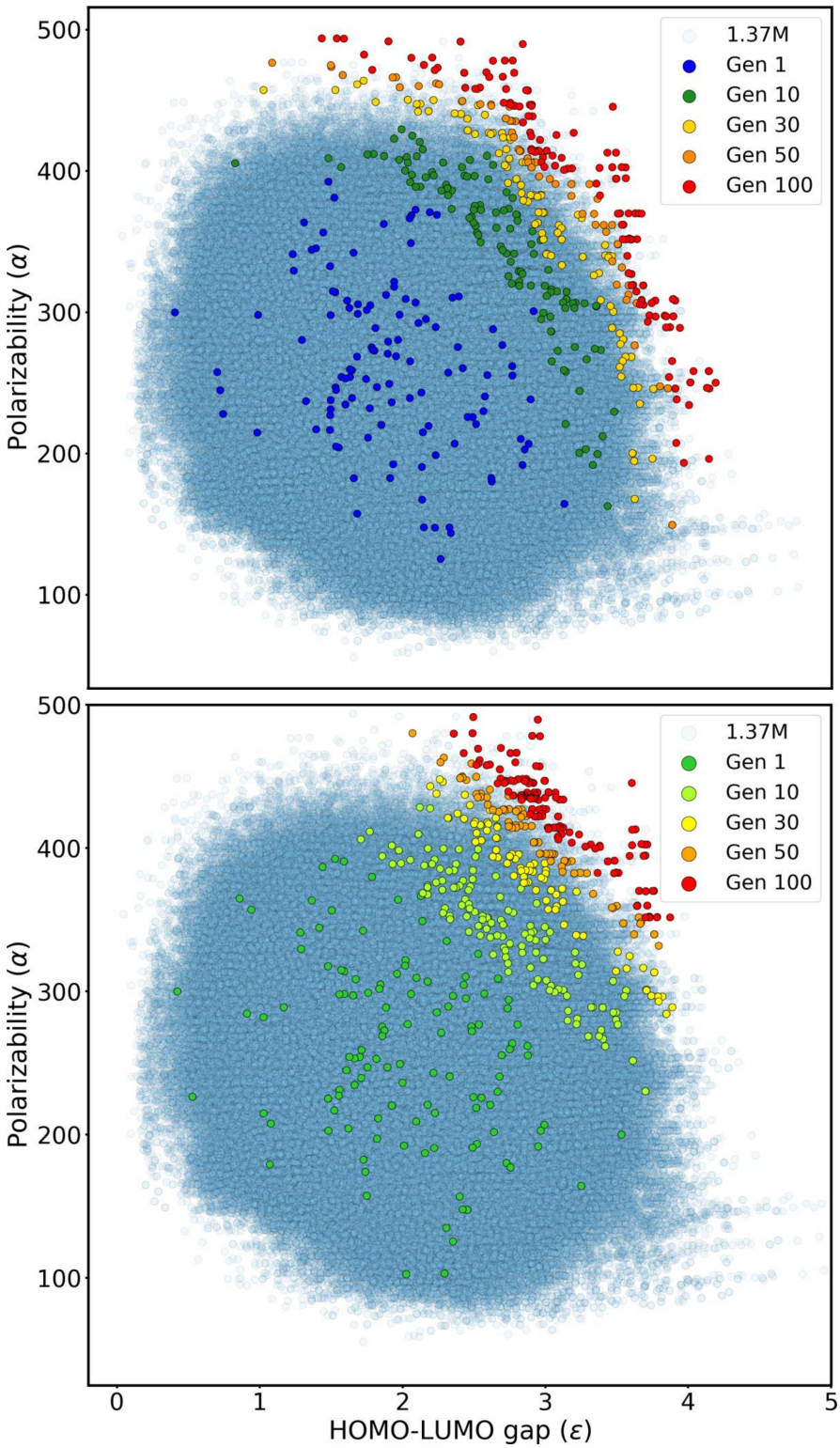
$$y' = \frac{y - y_{min}}{y_{max} - y_{min}} \tag{4}$$

where y' denotes the normalized target and the y_{min} and y_{max} values have been obtained from the calculated ground truth of the 1.37M space. Noteworthy, the ground truth is normally unknown in real-world design tasks, thus making the normalization of the optimization targets a challenging task. This prior knowledge is not required in the implementation and application of the PL-MOGA algorithm.

The largest difference in performance between the two algorithms was found when the WS-MOGA was used for a balanced search in which both coefficients were equal to one; *i.e.* $(\beta_\alpha, \beta_\epsilon) = (1.0, 1.0)$. The comparison of the evolution scatter plots in Supplementary Figure 20 shows that the WS-MOGA optimization resolved a narrower region of the Pareto front, resulting in poorer performance in terms of both the number of dominating points found

(8 versus 18) and the diversity of the final generation (TC = 0.44 versus 0.33). Another important limitation of the WS-MOGA algorithm relative to the PL-MOGA is the lack of scope control. For example, the separate $(\beta_\alpha, \beta_\epsilon) = (1.0, 1.0)$ and $(0.1, 0.1)$ optimizations produced exactly the same results and, therefore, it was not possible to direct a multiobjective optimization towards a central region of the Pareto front meanwhile fine-tuning its scope.

For non-balanced WS-MOGA optimizations using the $(\beta_\alpha, \beta_\epsilon)$ coefficients $(0.8, 0.2)$ and $(0.2, 0.8)$, the differences relative to the PL-MOGA extreme runs discussed in the manuscript were smaller and yet the latter algorithm performed slightly better in most cases. Whereas for α, the WS-MOGA found less dominating points (7 versus 9) with lower diversity (TC = 0.57 versus 0.56), for ϵ, the WS-MOGA found more dominating points (7 versus 5) though with lower diversity (TC = 0.39 versus 0.37).



Supplementary Figure 20: Weighted-sum benchmark. PL-MOGA (top) and weighted-sum (bottom) optimizations within the 1.37M space using (0.0, 0.0) scaling factors and (1.0, 1.0) weights, respectively. Units: Bohr^3 (α) and eV (ϵ).

References

(1) Duan, C.; Ladera, A. J.; Liu, J.; Taylor, M. G.; Ariyarathna, I. R.; Kulik, H. J. Exploiting Ligand Additivity for Transferable Machine Learning of Multireference Character across Known Transition Metal Complex Ligands. *J. Chem. Theory Comput.* **2022**, *18*, 4836–4845.

(2) Kneiding, H.; Lukin, R.; Lang, L.; Reine, S.; Pedersen, T. B.; De Bin, R.; Balcells, D. Deep learning metal complex properties with natural quantum graphs. *Digital Discovery* **2023**, *2*, 618–633.

(3) Sauer, W. H.; Schwarz, M. K. Molecular shape diversity of combinatorial libraries: a prerequisite for broad bioactivity. *J. Chem. Inf. Model.* **2003**, *43*, 987–1003.

(4) Shrake, A.; Rupley, J. A. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.* **1973**, *79*, 351–371.

(5) Eisenhaber, F.; Lijnzaad, P.; Argos, P.; Sander, C.; Scharf, M. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem.* **1995**, *16*, 273–284.

(6) Falivene, L.; Credendino, R.; Poater, A.; Petta, A.; Serra, L.; Oliva, R.; Scarano, V.; Cavallo, L. SambVca 2. A web tool for analyzing catalytic pockets with topographic steric maps. *Organometallics* **2016**, *35*, 2286–2293.

(7) Tolman, C. A. Phosphorus ligand exchange equilibriums on zerovalent nickel. Dominant role for steric effects. *J. Am. Chem. Soc.* **1970**, *92*, 2956–2965.

(8) Bilbrey, J. A.; Kazez, A. H.; Locklin, J.; Allen, W. D. Exact ligand cone angles. *J. Comput. Chem.* **2013**, *34*, 1189–1197.

(9) Bilbrey, J. A.; Kazez, A. H.; Locklin, J.; Allen, W. D. Exact Ligand Solid Angles. *J. Chem. Theory Comput.* **2013**, *9*, 5734–5744.

(10) Guzei, I. A.; Wendt, M. An improved method for the computation of ligand steric effects based on solid angles. *Dalton Trans.* **2006**, 3991–3999.

(11) Landrum, G. et al. rdkit/rdkit: 2023_03_1 (Q1 2023) Release. 2023; `https://doi.org/10.5281/zenodo.7880616`.

(12) Jorner, K.; Turcani, L. kjelljorner/morfeus: v0.7.2. 2022; `https://doi.org/10.5281/zenodo.7017599`.

(13) Frisch, M. J. et al. Gaussian˜16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.

(14) Shervashidze, N.; Schweitzer, P.; Van Leeuwen, E. J.; Mehlhorn, K.; Borgwardt, K. M. Weisfeiler-lehman graph kernels. *Journal of Machine Learning Research* **2011**, *12*.

(15) Hagberg, A.; Swart, P.; S Chult, D. *Exploring network structure, dynamics, and function using NetworkX*; 2008.