Supporting Information for:

Topological Prediction of Palladium Coordination Cages

David A. Poole III[a], Eduard O. Bobylev[a], Simon Mathew[a], Joost N.H. Reek[a]*

[a] *Homogeneous, Supramolecular, and Bio-inspired Catalysis group, van 't Hoff Institute for Molecular Science (HIMS), University of Amsterdam (UvA), Science Park 904, 1098 XH Amsterdam, The Netherlands*

Contents

A. Model Dynamic Studies, Parameter Fitting, and Validation ..... - 2 -

B. Topologies and Resulting Energies of Homoleptic Structures & Calculation Details ..... - 7 -

C. Linear Interpolation of Heteroleptic Structures, Detailed Analysis & Calculation Details ..... - 18 -

D. Synthesis and Characterization of Linkers and Complexes ..... - 19 -

E. Synthesis of $Pd_x^LFu_{2x}$ Homoleptic Assemblies ..... - 25 -

F. Synthesis and Characterization of $Pd_x^LEx_{2x}$ Homoleptic Assemblies ..... - 28 -

G. Synthesis and Characterization of $Pd_x^LEn_{2x}$ Homoleptic Assemblies ..... - 31 -

H. Synthesis and Characterization of $Pd_x^LEn_{(2x-n)}^LEx_n$ Heteroleptic Assemblies ..... - 33 -

I: Topological predictions using the Generalized Amber Force Field ..... - 58 -

J: CM reproduction of dihedral Scans ..... - 59 -

A. Model Dynamic Studies, Parameter Fitting, and Validation

A model complex was devised in order to conduct dynamics studies on the individual linkers shown in Figure S1. These models featured the linker of interest coordinated to two tris-pyridyl palladium (II) centers, which adequately reproduced the coordinated species found in palladium coordination cages.

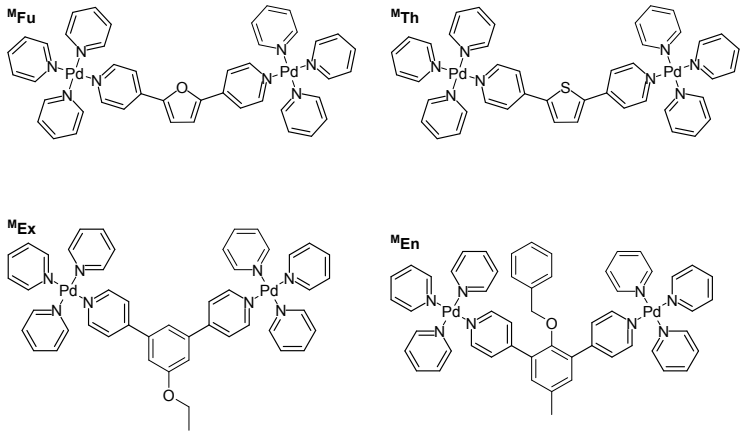


Figure S1. Model systems used for parameter development.

These structure was optimized at a B3LYP/def2-TZV level of theory using *Gaussian 16 rev C*.[8] Charge fitting was conducted using the RESP method employed by *antechamber*.[1] The resulting charge fitting revealed a significant transfer of electron density from the linker to the metal center, resulting in a net charge residing on the linker, ca. 0.87 q_e, as shown in Figure S2.

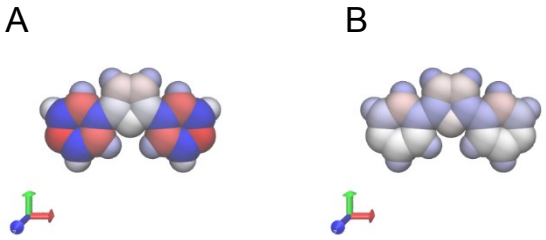
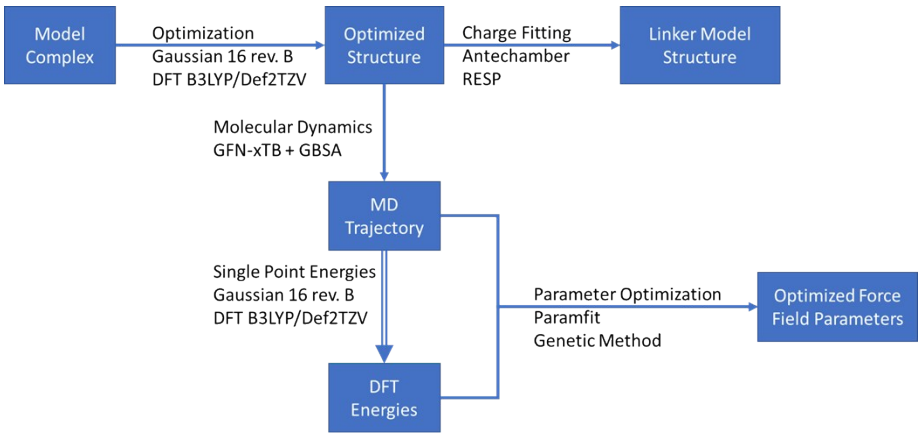


Figure S2. Models of LFu linkers showing the effect of palladium coordination on local charge. A) charge assignment produced from DFT calculations of the neutral linker. B) charge assignment produced from DFT calculations of the coordinated cationic complex.

This apparent charge transfer was found consistently across all models with the same approximate distribution of charge between the linker, pyridyl and palladium groups. In order to produce a uniform force field, these charges were scaled to a common 0.87 q_e on each linker with 0.26 q_e charge remaining on each palladium center, resulting in a net 2.00 q_e charge for

each Pd_1L_2 unit. While previous work has used a non-bonded approach,[2–4] this method does not take into account this observed charge transfer. While these methods are successful in simulating the formation process of coordination assemblies, they are unsuitable when charge interactions are significant. This obliged the use of a bonded model approach to accurately estimate relative free energy of individual topologies.

Parameter development continued following a scheme briefly illustrated below, Scheme S1. Atom types were assigned for these model structures by *antechamber*, though these were subsequently modified to accommodate square planar geometry. The nonbonded parameters were derived from either the GAFF forcefield,[5] for organic atoms, or the parameters defined by Li and coworkers for metal centers.[6] Bond lengths, angles, and dihedrals were set to the average values found across all model structures.



Scheme S1. Method of parameter optimization used in this work.

Molecular dynamics trajectories were then obtained using Grimme's GFN2-xTB for the model complexes.[7] The individual frames of these trajectories were then submitted for single-point energy evaluation using DFT at a B3LYP/def2-TZV theory level as shown by Figure S3. This approach was chosen since the *xTB* is known to produce reasonable geometries for metal complexes with rapid equilibration of its molecular dynamics trajectories, and, the subsequent single point calculations could then be parallelized, resulting in significant time savings.

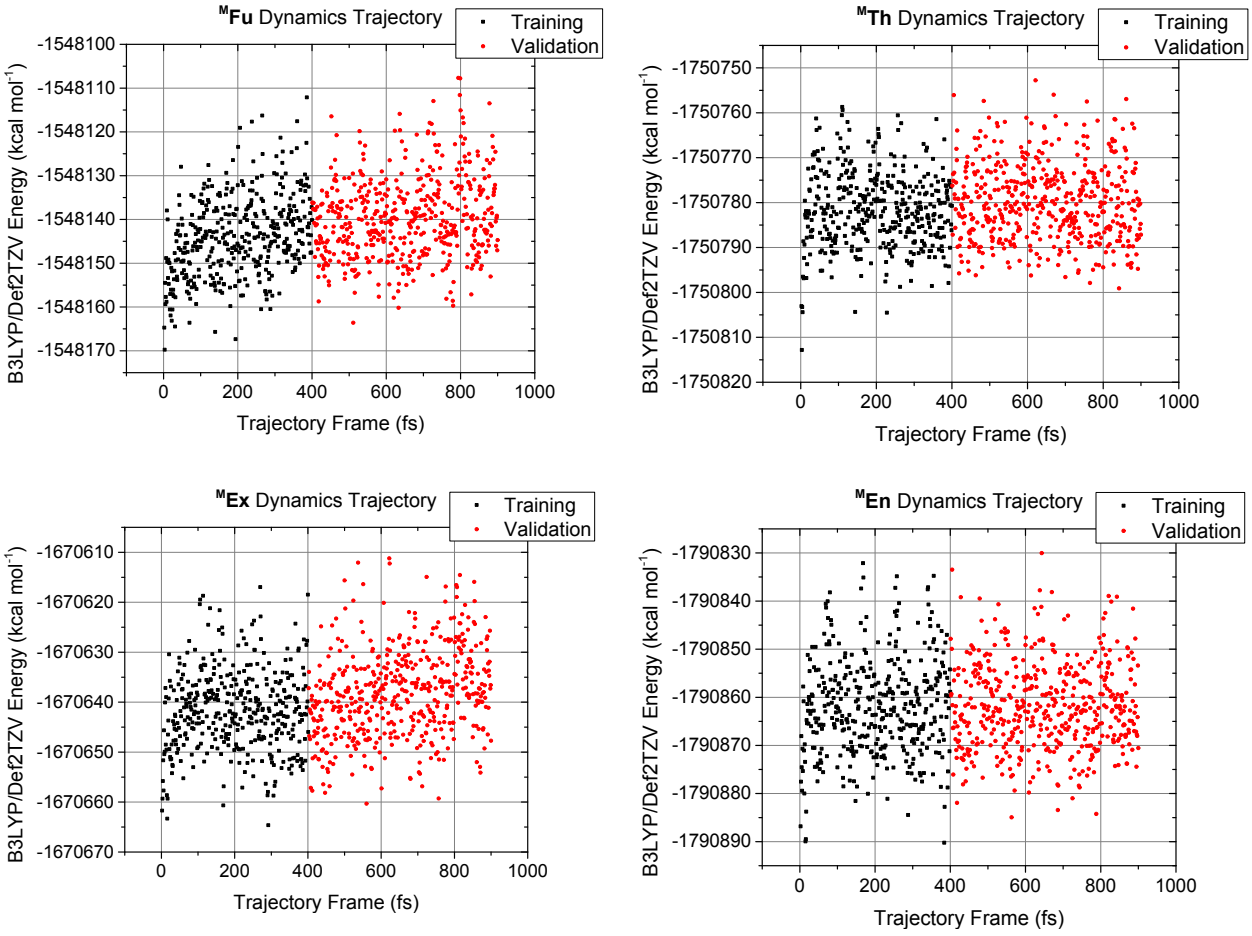
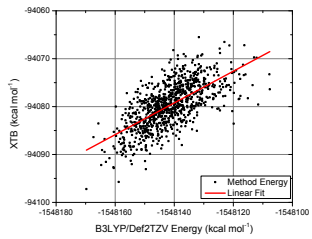
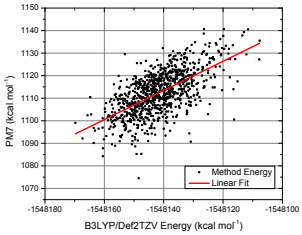
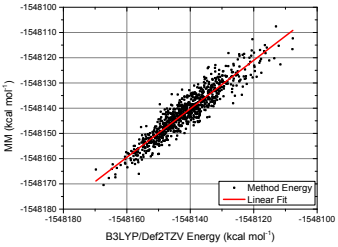
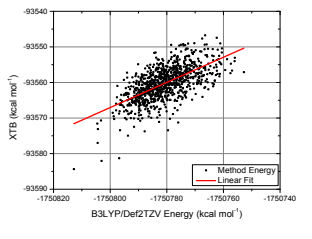
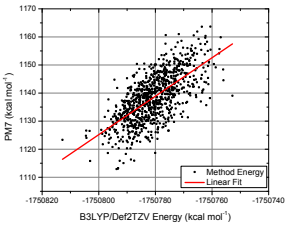
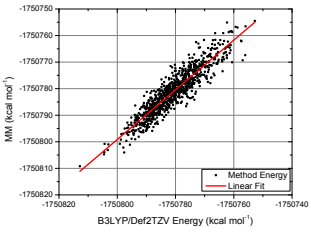
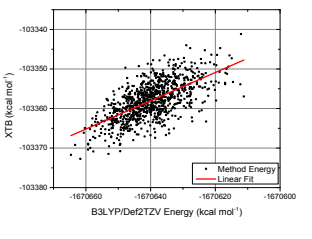
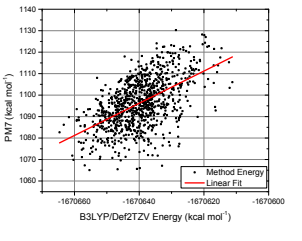
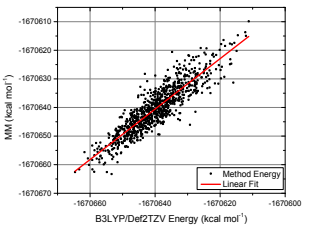
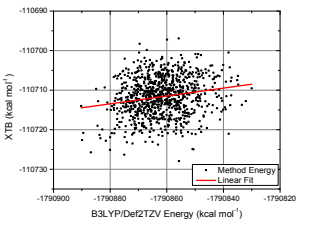
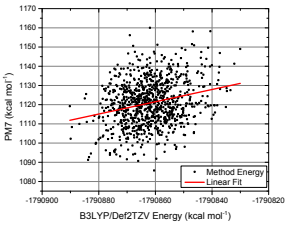
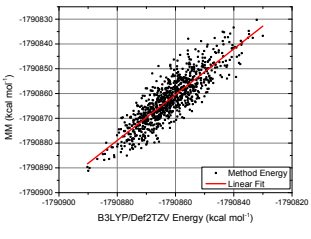


Figure S3. DFT single point energies obtained for xTB derived molecular dynamics trajectories for each model system. The initial 400 trajectory frames and corresponding single point energies were used for parameter fitting (black), the remaining 500 frames and energies were used for validation and method comparison (red).

Forcefield fitting was conducted using the *paramfit* program,[8] implementing the genetic fitting approach described of Betz.[8] The parameters for all four model systems were initially optimized from a random population using the first 100 frames of the trajectory with the DFT computed single point energies. This process was repeated, increasing the number of included frames by 50 until 400 frames and corresponding single point energies. This process was again repeated as a multi-molecule fit using all four model data sets in order to simultaneously fit the common terms. The forcefield terms containing atom types 'nc' and 'nd', corresponding to pyridyl nitrogen atoms, were averaged at this point. Lastly, the offset values were corrected and a final genetic optimization was conducted for each of the four models, optimizing only their unique forcefield terms. The resulting parameters were then evaluated by assessing the energy of the remaining 500 frames and single point energies, the results of which are presented in Figure S3.

Table S1. Comparison and validation results showing the correlated energies obtained for each model trajectory (black dots), with a linear fit of the overall correlation (red line). Validation results (right), show a reliable reproduction of DFT energies using a specifically fit molecular mechanics forcefield. Comparisons were made to both GFN2-xTB (left) and PM7 (center), as two widely employed methods for the study of organometallic complexes.

	GFN2-xTB	PM7	CM
^{M}Fu			
	RMSD = 3.68 kcal mol^{-1} R^2 = 0.48	RMSD = 7.47 kcal mol^{-1} R^2 = 0.42	RMSD = 3.48 kcal mol^{-1} R^2 = 0.88
^{M}Th			
	RMSD = 3.83 kcal mol^{-1} R^2 = 0.44	RMSD = 6.38 kcal mol^{-1} R^2 = 0.47	RMSD = 3.45 kcal mol^{-1} R^2 = 0.85
^{M}Ex			
	RMSD = 3.64 kcal mol^{-1} R^2 = 0.42	RMSD = 9.70 kcal mol^{-1} R^2 = 0.31	RMSD = 3.81 kcal mol^{-1} R^2 = 0.80
^{M}En			
	RMSD = 5.11 kcal mol^{-1} R^2 = 0.08	RMSD = 11.4 kcal mol^{-1} R^2 = 0.08	RMSD = 5.07 kcal mol^{-1} R^2 = 0.76

The results of this validation show that the use of a specially fit molecular mechanics forcefield may achieve similar, and often better, results when compared to state-of-the-art semi-empirical techniques. This is expected, where semi-empirical methods may be generally applied, our forcefield is limited to a narrow class of structures and interactions[10]. Even so, it should be noted that both semi-empirical techniques produce significant errors for ^{M}**Ex** and ^{M}**En**, and caution should be used when applying these techniques for study functionalized or ornamented cages, especially for host-guest interactions. Given the high correlation, low error, and near unit slope observed for the molecular mechanics approach, we do not foresee significant improvements to be made by further fitting, and consider these parameters adequate for estimating the relative energies of the large assemblies.

Table S2: Input parameters used to produce dynamics trajectories (top), conduct DFT optimizations (middle), and genetic optimization (bottom). Single point energy calculations were conducted using identical options, omitting the 'opt' flag for optimization.

```
xtb input.xyzomd --charge 4 --gbsa acetonitrile;
```

```
# opt ub3lyp/def2tzv force integral=grid=ultrafine iop(6/33=2,6/42=6,6/41=10)
```

```
RUNTYE=FIT
COORDINATE_FORMAT=TRAJECTORY
 FUNC_TO_FIT=SUM_SQUARES_AMBER_STANDARD
QM_ENERGY_UNITS=HARTREE
PARAMETERS_TO_FIT=LOAD
PARAMETER_FILE_NAE=fit.list
ALGORITHM=GENETIC
OPTIMIZATIONS=90
MAX_GENERATIONS=500
GENERATIONS_TO_CONV=50
GENERATIONS_TO_SIMPLEX=0
GENERATIONS_WITHOUT_SIMPLEX=5
MUTATION_RAE=0.350000
PARENT_PERCENT=0.200
SEARCH_SPAE=0.15
WRITE_ENERGY=fit.dat
WRITE_FRCMOD=fit.frcmod
```

B. Topologies and Resulting Energies of Homoleptic Structures & Calculation Details

Templates were manually constructed containing the positions of the metal centers and linker pyridyl groups. These templates were based on either reported crystal structures (M_nL_{2n}, n= 3, 4, 6, 12, 24, 30), computational models (M_nL_{2n}, n= 8, 9, 12), known polyhedra (M_nL_{2n}, n= 10, 14, 15, 16, 24), extension of double crown ring polymers (M_nL_{2n}, n= 5, 6, 7), or by modification of existing large structures by removal of a single M_1L_2 unit (M_nL_{2n}, n= 11, 13-15, 17-23, 25-29). Some of topologies of a specific number of metal centers often had alternative topologies, e.g. $M_{12}L_{24}$[3], and these were considered independently.

Models of each topology were then constructed by an in-house automation script using the surprisingly robust *ProFit* software package.[11] These models were then optimized by 750 steps using *sander,* followed by molecular dynamics annealing, and a final structural optimization using *pmemd.cuda*[12] until convergence as shown in Table S3. Optimization was conducted using implicit solvent conditions, using the dielectric constant of acetonitrile (extdiel= 36).

Table S3: Input files used for minimization by CPU (left) and GPU (right) implementations, and molecular dynamics annealing (center). The single threaded sander code is appropriate for minimizing the high-energy state of the initial structure while the GPU implemented pmemd is significantly faster for both annealing and subsequent minimization.

Sander.in	anneal.in	cuda.in
Initial CPU minimization &cntrl imin = 1, ntb = 0, igb = 8, saltcon= 0.001, extdiel= 36, maxcyc = 750, ioutfm = 1, ncyc = 50, cut = 12, ntwr = 10, ntpr = 10, /	MD annealing &cntrl imin = 0, ntb = 0, irest = 0, ntx = 1, igb = 8, saltcon = 0.001, extdiel = 36.0, ntt = 3, gamma_ln= 1.00, tempi = 0.0, temp0 = 300.0, nstlim = 100000, dt = 0.01, cut = 9999, ntpr = 10000, ntwr = 10000, /	Final GPU minimization &cntrl imin = 1, ntb = 0, igb = 8, saltcon= 0.001, extdiel= 36, maxcyc = 500000, ioutfm = 1, ncyc = 1000, cut = 9999, ntwr = 1000, ntpr = 1000, /

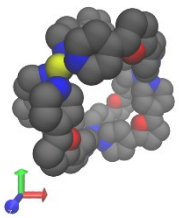
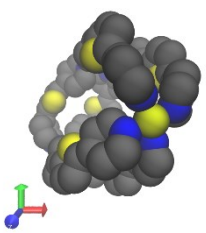
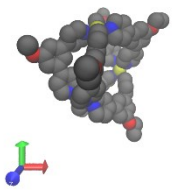
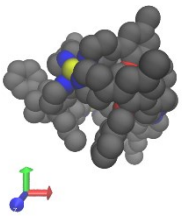
The optimized models were visualized with VMD, and shown below alongside their relative forcefield energy and dynamic radius (Table S4). The probability is computed directly from the forcefield energy of each optimized model using Boltzmann statistics (Equation S1). This model assumes that each topology is an accessible microstate for each linker, ignoring the possible kinetic barriers needed to access a given structure. This assumption is implicit in the process of non-covalent dynamic synthesis which uses a suitably high temperature to overcome the kinetic barriers in self-assembly.

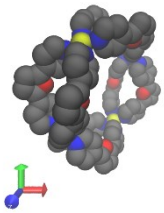
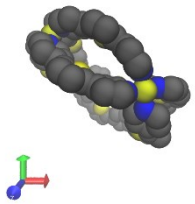
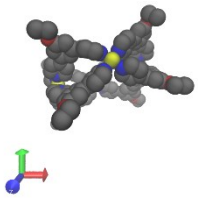
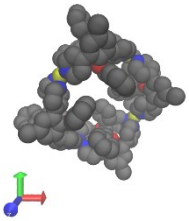
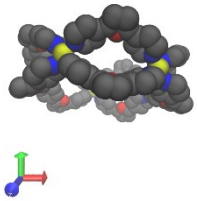
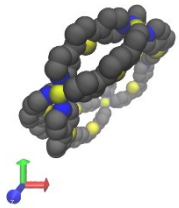
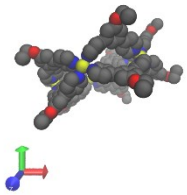
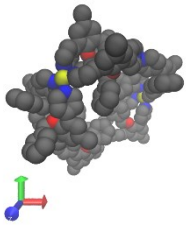
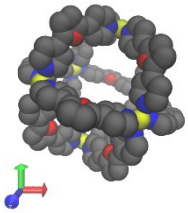
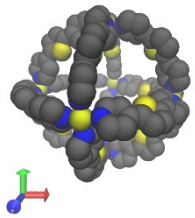
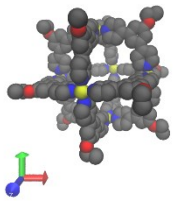
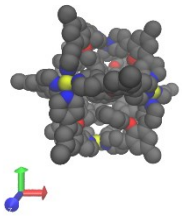
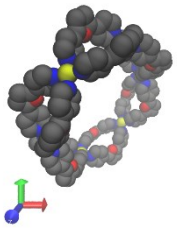
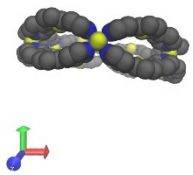
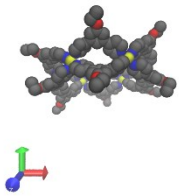
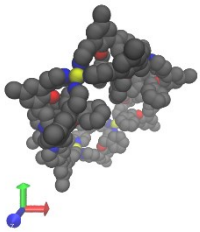
$$P_i = \frac{n_i e^{-\frac{n_i E_i}{k_B T}}}{\sum_j \left(n_j e^{-\frac{n_j E_j}{k_B T}} \right)}$$

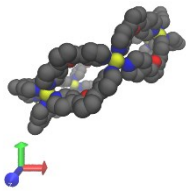
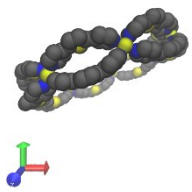
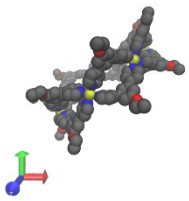
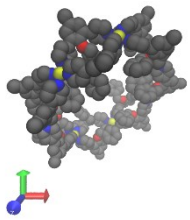
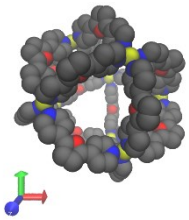
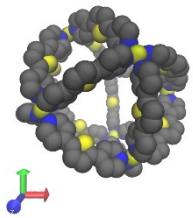
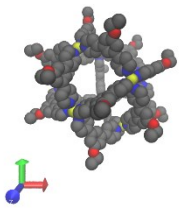
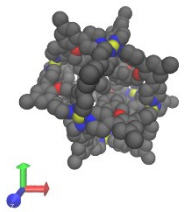
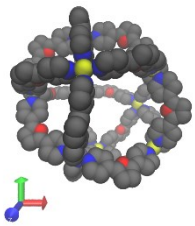
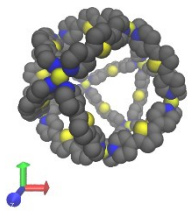
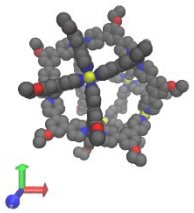
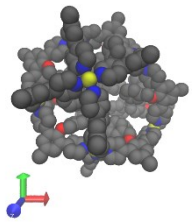
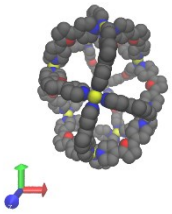
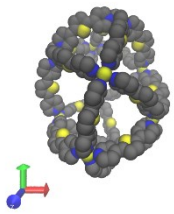
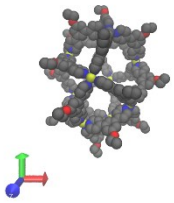
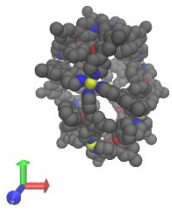
Equation S1. Probability of any topology I is computed directly from the relative energy of the given topology, interpreted as a microstate of the linker's geometry.

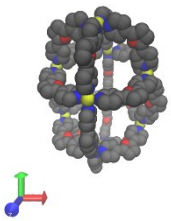
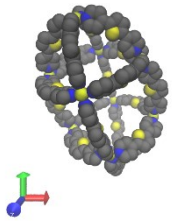
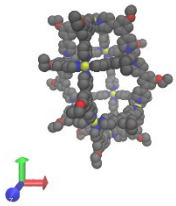
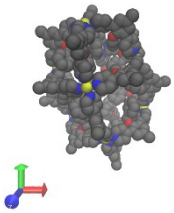
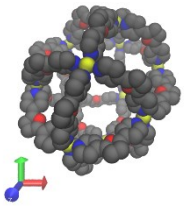
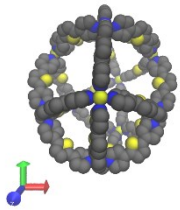
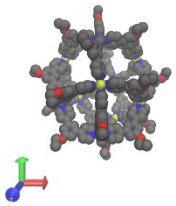
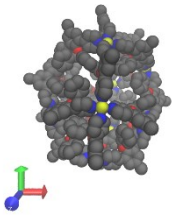
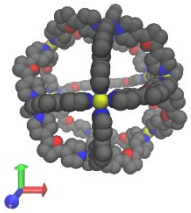
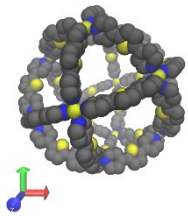
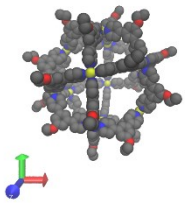
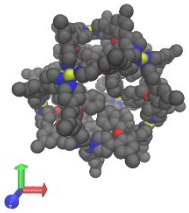
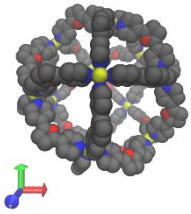
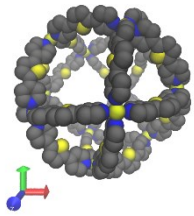
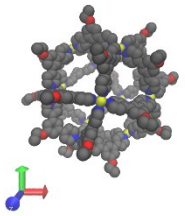
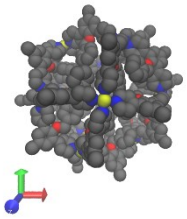
Our results (Table S4) reveal that of the numerous possible topologies, few are preferred— even when these alternative topologies are simple gyrations or extractions from the most preferred topology. Furthermore, favorable minority topologies are found 'remotely' from stable ones, suggesting that the formation process proceeds through a single common intermediate rather than stepwise polymer growth.

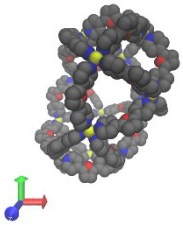
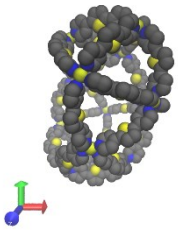
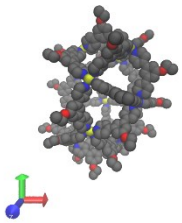
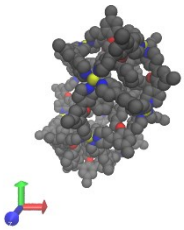
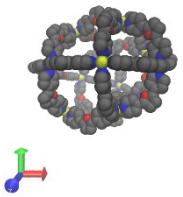
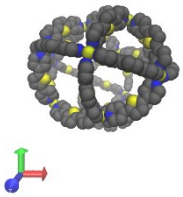
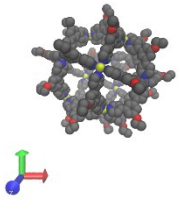
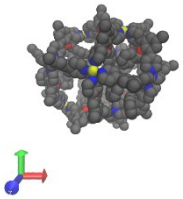
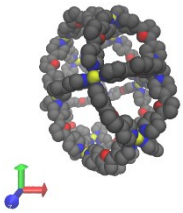
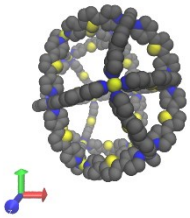
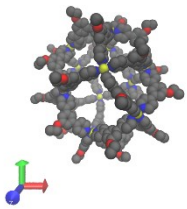
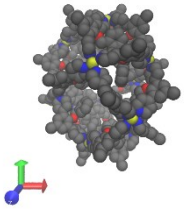
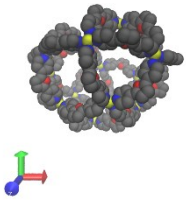
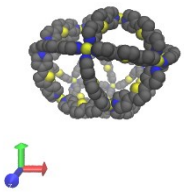
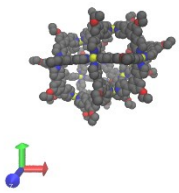
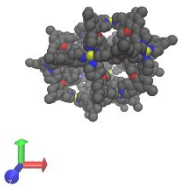
Table S4: Topological prediction results for homoleptic assemblies composed of the studied linkers. Renderings are produced, automatically, showing the van der Waals occupied volumes and configuration colored by element, grey for carbon, red for oxygen, blue for nitrogen, yellow for sulfur and yellow-green for palladium – hydrogen atoms are omitted for clarity. Energy values (E_{rel}) are weighted by the number of metal centers incorporated into the assembly structure (n_i). Hydrodynamic radius (r_H) is derived using a spherical approximation of the computed model volume. [a]S cylindrical or 'double crown ring' structure [b]S Topology produced by excision of M_1L_2 of a larger topology [c] Produced by gyration of an assembly of the same size.

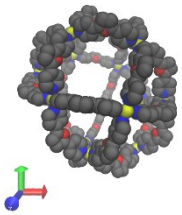
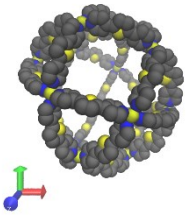
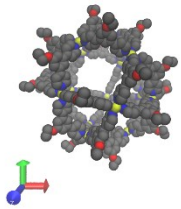
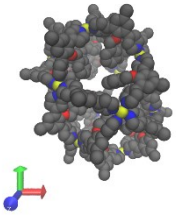
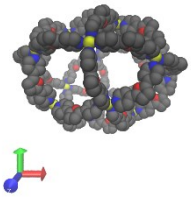
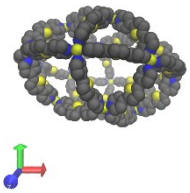
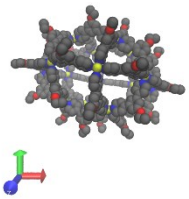
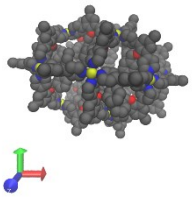
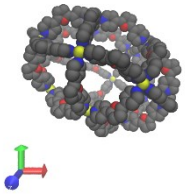
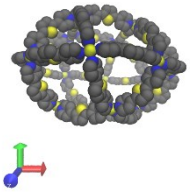
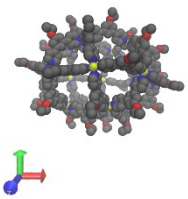
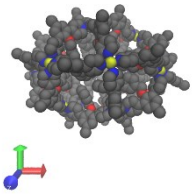
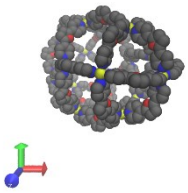
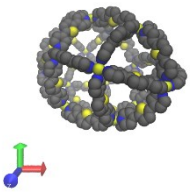
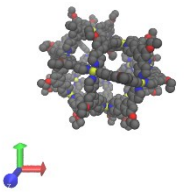
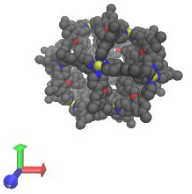
	$Pd_n{}^{L}\mathbf{Fu}_{2n}$	$Pd_n{}^{L}\mathbf{Th}_{2n}$	$Pd_n{}^{L}\mathbf{Ex}_{2n}$	$Pd_n{}^{L}\mathbf{En}_{2n}$
M_3L_6 [a]				
	E_{rel}=69.0 kcal mol^{-1} rH=8.09 Å	E_{rel}=98.0 kcal mol^{-1} rH=8.24 Å	E_{rel}=50.0 kcal mol^{-1} rH=9.81 Å	E_{rel}=68.3 kcal mol^{-1} r_H=-9.16 Å

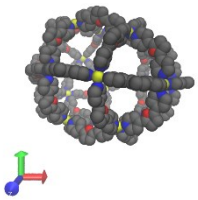
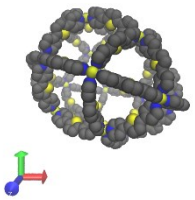
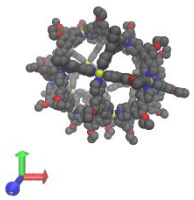
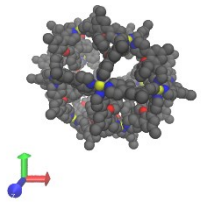
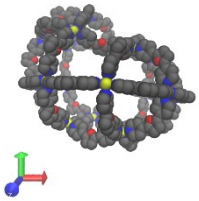
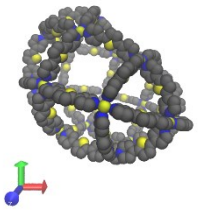
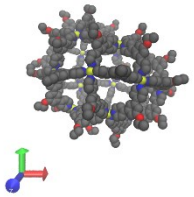
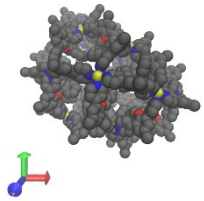
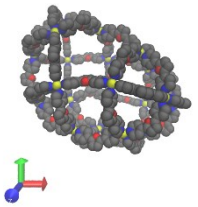
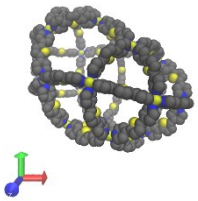
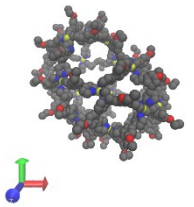
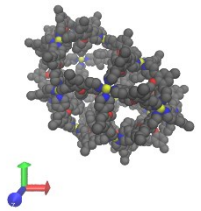
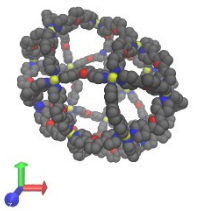
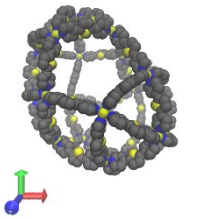
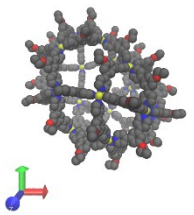
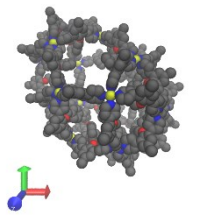
M_4L_8[a]	 E_{rel}=34.5 kcal mol^{-1} r_H=9.27 Å	 E_{rel}=54.7 kcal mol^{-1} r_H=9.37 Å	 E_{rel}=32.3 kcal mol^{-1} r_H=11.37 Å	 E_{rel}=66.2 kcal mol^{-1} r_H=-9.22 Å
M_5L_{10}[a]	 E_{rel}=36.5 kcal mol^{-1} r_H=10.48 Å	 E_{rel}=55.5 kcal mol^{-1} r_H=10.50 Å	 E_{rel}=26.4 kcal mol^{-1} r_H=13.26 Å	 E_{rel}=54.7 kcal mol^{-1} r_H=-9.26 Å
M_6L_{12}	 E_{rel}=8.7 kcal mol^{-1} r_H=11.61 Å	 E_{rel}=28.6 kcal mol^{-1} r_H=11.56 Å	 E_{rel}=26.2 kcal mol^{-1} r_H=14.31 Å	 E_{rel}=30.1 kcal mol^{-1} r_H=-9.30 Å
M_6L_{12}[a]	 E_{rel}=38.7 kcal mol^{-1} r_H=11.19 Å	 E_{rel}=44.7 kcal mol^{-1} r_H=11.23 Å	 E_{rel}=25.1 kcal mol^{-1} r_H=14.35 Å	 E_{rel}=51.7 kcal mol^{-1} r_H=-9.31 Å

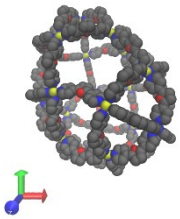
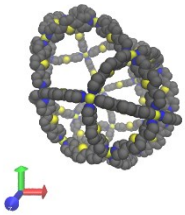
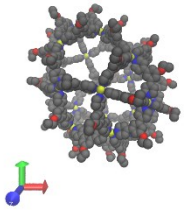
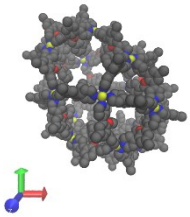
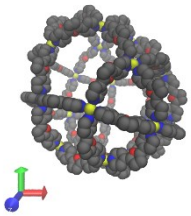
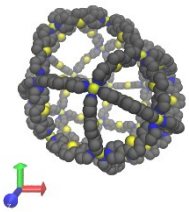
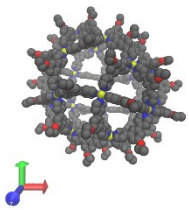
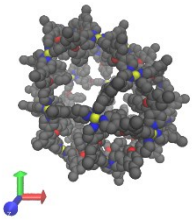
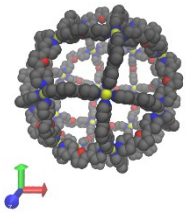
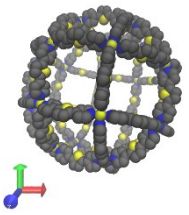
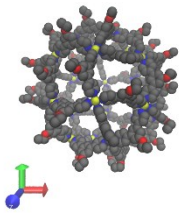
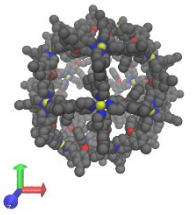
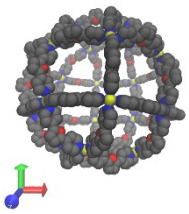
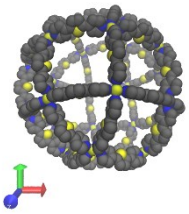
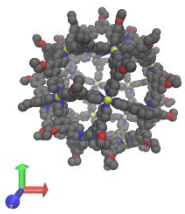
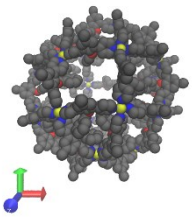
M_7L_{14}[a]	 E_{rel}=57.3 kcal mol^{-1} r_H=11.99 Å	 E_{rel}=42.6 kcal mol^{-1} r_H=11.91 Å	 E_{rel}=25.4 kcal mol^{-1} r_H=15.76 Å	 E_{rel}=45.5 kcal mol^{-1} r_H=-9.34 Å
M_8L_{16}	 E_{rel}=4.4 kcal mol^{-1} r_H=13.31 Å	 E_{rel}=16.6 kcal mol^{-1} r_H=13.36 Å	 E_{rel}=26.0 kcal mol^{-1} r_H=16.45 Å	 E_{rel}=19.0 kcal mol^{-1} r_H=-9.36 Å
M_9L_{18}	 E_{rel}=3.3 kcal mol^{-1} r_H=14.09 Å	 E_{rel}=13.7 kcal mol^{-1} r_H=14.16 Å	 E_{rel}=25.9 kcal mol^{-1} r_H=17.39 Å	 E_{rel}=14.6 kcal mol^{-1} r_H=-9.38 Å
$M_{10}L_{20}$	 E_{rel}=3.7 kcal mol^{-1} r_H=14.82 Å	 E_{rel}=12.3 kcal mol^{-1} r_H=14.92 Å	 E_{rel}=38.7 kcal mol^{-1} r_H=18.36 Å	 E_{rel}=21.5 kcal mol^{-1} r_H=-9.40 Å

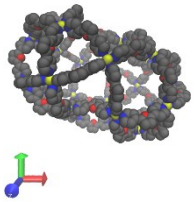
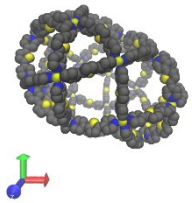
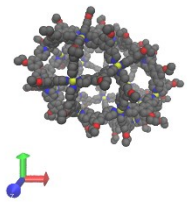
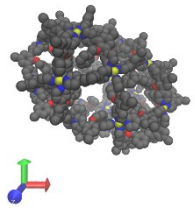
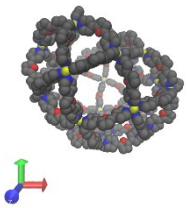
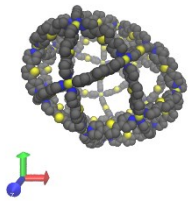
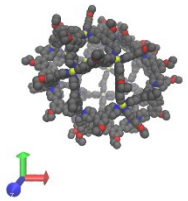
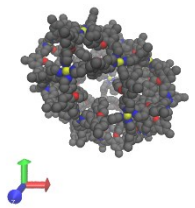
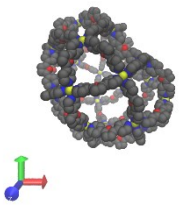
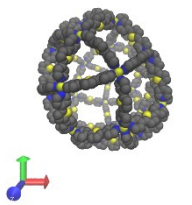
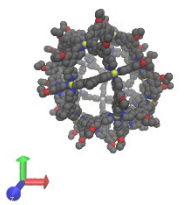
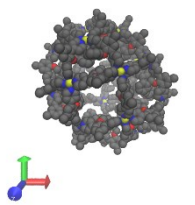
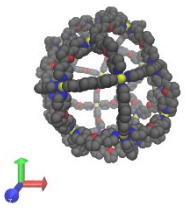
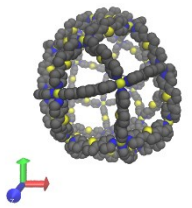
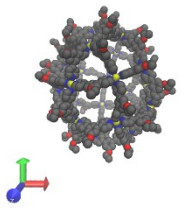
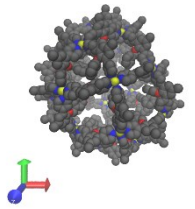
$M_{10}L_{20}$[b]	 E_{rel}=6.3 kcal mol^{-1} r_H=14.76 Å	 E_{rel}=16.5 kcal mol^{-1} r_H=14.83 Å	 E_{rel}=24.4 kcal mol^{-1} r_H=18.55 Å	 E_{rel}=14.3 kcal mol^{-1} r_H=-9.40 Å
$M_{11}L_{22}$[b]	 E_{rel}=2.5 kcal mol^{-1} r_H=15.54 Å	 E_{rel}=9.2 kcal mol^{-1} r_H=15.68 Å	 E_{rel}=28.1 kcal mol^{-1} r_H=19.29 Å	 E_{rel}=12.9 kcal mol^{-1} r_H=-9.42 Å
$M_{12}L_{24}$	 E_{rel}=0.3 kcal mol^{-1} r_H=16.24 Å	 E_{rel}=5.2 kcal mol^{-1} r_H=16.42 Å	 E_{rel}=0.1 kcal mol^{-1} r_H=20.03 Å	 E_{rel}=4.2 kcal mol^{-1} r_H=-9.44 Å
$M_{12}L_{24}$[c]	 E_{rel}=1.2 kcal mol^{-1} r_H=16.24 Å	 E_{rel}=6.4 kcal mol^{-1} r_H=16.41 Å	 E_{rel}=1.4 kcal mol^{-1} r_H=20.00 Å	 E_{rel}=31.7 kcal mol^{-1} r_H=-9.43 Å

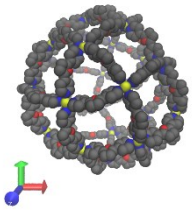
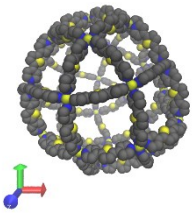
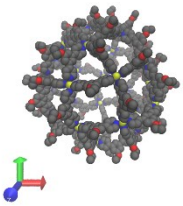
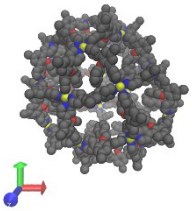
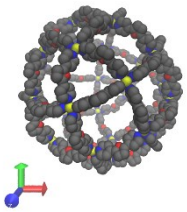
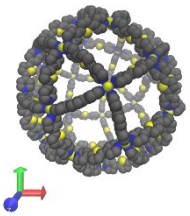
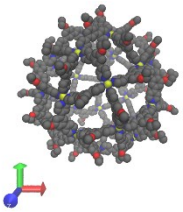
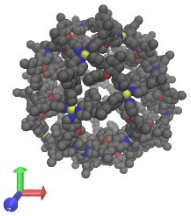
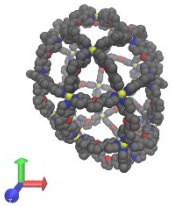
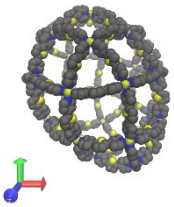
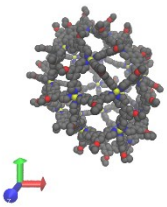
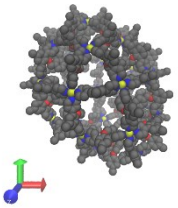
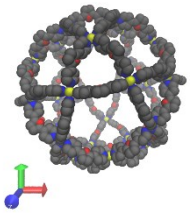
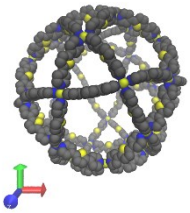
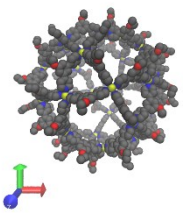
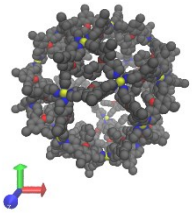
$M_{13}L_{26}$[b]	 E_{rel}=5.6 kcal mol^{-1} r_H=16.56 Å	 E_{rel}=11.9 kcal mol^{-1} r_H=16.72 Å	 E_{rel}=28.8 kcal mol^{-1} r_H=19.89 Å	 E_{rel}=11.8 kcal mol^{-1} r_H=-9.45 Å
$M_{14}L_{28}$[b]	 E_{rel}=6.2 kcal mol^{-1} r_H=17.28 Å	 E_{rel}=8.6 kcal mol^{-1} r_H=17.49 Å	 E_{rel}=25.7 kcal mol^{-1} r_H=21.27 Å	 E_{rel}=18.3 kcal mol^{-1} r_H=-9.47 Å
$M_{14}L_{28}$[b]	 E_{rel}=4.1 kcal mol^{-1} r_H=16.70 Å	 E_{rel}=14.0 kcal mol^{-1} r_H=16.47 Å	 E_{rel}=29.0 kcal mol^{-1} r_H=21.18 Å	 E_{rel}=10.6 kcal mol^{-1} r_H=-9.80 Å
$M_{15}L_{30}$[b]	 E_{rel}=5.7 kcal mol^{-1} r_H=17.83 Å	 E_{rel}=11.6 kcal mol^{-1} r_H=18.00 Å	 E_{rel}=19.0 kcal mol^{-1} r_H=22.11 Å	 E_{rel}=16.4 kcal mol^{-1} r_H=-9.47 Å

$M_{15}L_{30}$[c]	 E_{rel}=1.2 kcal mol^{-1} r_H=18.09 Å	 E_{rel}=3.8 kcal mol^{-1} r_H=18.32 Å	 E_{rel}=3.3 kcal mol^{-1} r_H=22.36 Å	 E_{rel}=25.5 kcal mol^{-1} r_H=-9.49 Å
$M_{16}L_{32}$[c]	 E_{rel}=4.2 kcal mol^{-1} r_H=18.47 Å	 E_{rel}=7.8 kcal mol^{-1} r_H=18.69 Å	 E_{rel}=26.4 kcal mol^{-1} r_H=22.83 Å	 E_{rel}=12.2 kcal mol^{-1} r_H=-9.49 Å
$M_{16}L_{32}$[c]	 E_{rel}=2.2 kcal mol^{-1} r_H=18.51 Å	 E_{rel}=6.4 kcal mol^{-1} r_H=18.76 Å	 E_{rel}=27.3 kcal mol^{-1} r_H=22.68 Å	 E_{rel}=10.7 kcal mol^{-1} r_H=-9.50 Å
$M_{17}L_{34}$[b]	 E_{rel}=3.5 kcal mol^{-1} r_H=18.84 Å	 E_{rel}=6.0 kcal mol^{-1} r_H=19.19 Å	 E_{rel}=29.7 kcal mol^{-1} r_H=23.15 Å	 E_{rel}=12.6 kcal mol^{-1} r_H=-9.50 Å

$M_{18}L_{36}$[b]	 E_{rel}=2.8 kcal mol^{-1} r_H=19.61 Å	 E_{rel}=5.1 kcal mol^{-1} r_H=19.84 Å	 E_{rel}=29.3 kcal mol^{-1} r_H=23.31 Å	 E_{rel}=13.0 kcal mol^{-1} r_H=-9.51 Å
$M_{19}L_{38}$[b]	 E_{rel}=4.2 kcal mol^{-1} r_H=19.93 Å	 E_{rel}=6.0 kcal mol^{-1} r_H=20.25 Å	 E_{rel}=29.7 kcal mol^{-1} r_H=24.08 Å	 E_{rel}=11.9 kcal mol^{-1} r_H=-9.52 Å
$M_{20}L_{40}$[b]	 E_{rel}=4.4 kcal mol^{-1} r_H=20.33 Å	 E_{rel}=6.7 kcal mol^{-1} r_H=20.69 Å	 E_{rel}=29.7 kcal mol^{-1} r_H=24.13 Å	 E_{rel}=10.8 kcal mol^{-1} r_H=-9.53 Å
$M_{21}L_{42}$[b]	 E_{rel}=4.1 kcal mol^{-1} r_H=20.89 Å	 E_{rel}=6.2 kcal mol^{-1} r_H=21.26 Å	 E_{rel}=29.7 kcal mol^{-1} r_H=25.43 Å	 E_{rel}=10.0 kcal mol^{-1} r_H=-9.53 Å

$M_{22}L_{44}$[b]				
	E_{rel}=3.9 kcal mol^{-1} r_H=21.51 Å	E_{rel}=4.5 kcal mol^{-1} r_H=21.90 Å	E_{rel}=29.8 kcal mol^{-1} r_H=25.32 Å	E_{rel}=10.3 kcal mol^{-1} r_H=-9.55 Å
$M_{23}L_{46}$[b]				
	E_{rel}=2.8 kcal mol^{-1} r_H=22.05 Å	E_{rel}=3.6 kcal mol^{-1} r_H=22.42 Å	E_{rel}=28.9 kcal mol^{-1} r_H=27.12 Å	E_{rel}=8.0 kcal mol^{-1} r_H=-9.55 Å
$M_{24}L_{48}$				
	E_{rel}=2.6 kcal mol^{-1} r_H=22.56 Å	E_{rel}=0.5 kcal mol^{-1} r_H=22.99 Å	E_{rel}=28.1 kcal mol^{-1} r_H=27.12 Å	E_{rel}=0.5 kcal mol^{-1} r_H=-9.58 Å
$M_{24}L_{48}$[c]				
	E_{rel}=2.4 kcal mol^{-1} r_H=22.55 Å	E_{rel}=0.6 kcal mol^{-1} r_H=22.99 Å	E_{rel}=28.6 kcal mol^{-1} r_H=26.01 Å	E_{rel}=0.4 kcal mol^{-1} r_H=-9.58 Å

$M_{25}L_{50}$[b]	 E_{rel}=5.4 kcal mol^{-1} r_H=22.46 Å	 E_{rel}=5.3 kcal mol^{-1} r_H=23.06 Å	 E_{rel}=24.7 kcal mol^{-1} r_H=26.79 Å	 E_{rel}=13.7 kcal mol^{-1} r_H=-9.56 Å
$M_{26}L_{52}$[b]	 E_{rel}=5.0 kcal mol^{-1} r_H=23.10 Å	 E_{rel}=4.0 kcal mol^{-1} r_H=23.64 Å	 E_{rel}=25.9 kcal mol^{-1} r_H=26.63 Å	 E_{rel}=12.4 kcal mol^{-1} r_H=-9.58 Å
$M_{27}L_{54}$[b]	 E_{rel}=4.4 kcal mol^{-1} r_H=23.40 Å	 E_{rel}=5.3 kcal mol^{-1} r_H=23.87 Å	 E_{rel}=27.7 kcal mol^{-1} r_H=26.79 Å	 E_{rel}=10.9 kcal mol^{-1} r_H=-9.58 Å
$M_{28}L_{56}$[b]	 E_{rel}=5.7 kcal mol^{-1} r_H=23.87 Å	 E_{rel}=3.5 kcal mol^{-1} r_H=24.50 Å	 E_{rel}=29.1 kcal mol^{-1} r_H=27.94 Å	 E_{rel}=10.0 kcal mol^{-1} r_H=-9.59 Å

$M_{29}L_{58}$[b]				
	E_{rel}=4.4 kcal mol^{-1} r_H=24.52 Å	E_{rel}=2.3 kcal mol^{-1} r_H=25.05 Å	E_{rel}=30.4 kcal mol^{-1} r_H=29.65 Å	E_{rel}=9.9 kcal mol^{-1} r_H=-9.60 Å
$M_{30}L_{60}$				
	E_{rel}=3.8 kcal mol^{-1} r_H=24.98 Å	E_{rel}=1.3 kcal mol^{-1} r_H=25.54 Å	E_{rel}=31.0 kcal mol^{-1} r_H=29.52 Å	E_{rel}=4.8 kcal mol^{-1} r_H=-9.62 Å
$M_{30}L_{60}$[c]				
	E_{rel}=4.5 kcal mol^{-1} r_H=24.85 Å	E_{rel}=2.3 kcal mol^{-1} r_H=25.40 Å	E_{rel}=31.7 kcal mol^{-1} r_H=29.81 Å	E_{rel}=6.7 kcal mol^{-1} r_H=-9.62 Å
$M_{30}L_{60}$				
	E_{rel}=4.6 kcal mol^{-1} r_H=25.38 Å	E_{rel}=2.3 kcal mol^{-1} r_H=25.74 Å	E_{rel}=33.1 kcal mol^{-1} r_H=30.09 Å	E_{rel}=4.5 kcal mol^{-1} r_H=-9.62 Å

C. Linear Interpolation of Heteroleptic Structures, Detailed Analysis & Calculation Details

The topological prediction of heteroleptic assemblies required a new protocol in order to determine the assembly outcomes of arbitrary linker compositions (e.g. an average composition of $Pd_{1.0}{}^{L}\mathbf{Fu}_{0.3}{}^{L}\mathbf{Th}_{1.7}$). This is necessary as many heteroleptic assemblies lack common factors (e.g. $Pd_{11}{}^{L}\mathbf{Fu}_{21}{}^{L}\mathbf{Th}_{1}$ and $Pd_{13}{}^{L}\mathbf{Fu}_{25}{}^{L}\mathbf{Th}_{1}$) and so their relative free energies cannot be compared directly. In this section, we provide a simplified example considering heteroleptic assemblies composed of $^{L}\mathbf{Fu}$ and $^{L}\mathbf{Th}$ forming into $Pd_{12}L_{24}$, $Pd_{15}L_{30}$, or $Pd_{24}L_{48}$ heteroleptic assemblies.

Heteroleptic assemblies of varying composition ($Pd_w{}^{L}\mathbf{Fu}_{(x)}{}^{L}\mathbf{Th}_{(2w-x)}$ where $x \in [0 \ldots 2w]$) were constructed using the same methods in our homoleptic analysis (see Computational Details, Main Text). A total of 30 model structures of heteroleptic assemblies were produced for each composition (x) of each topology (w), with the linkers randomly distributed in the structure. These model structures were then annealed and minimized following the same procedure used for the homoleptic complexes (see Computational Details). The resulting minimized energies were then fit with a linear interpolation as shown in Figure S4.

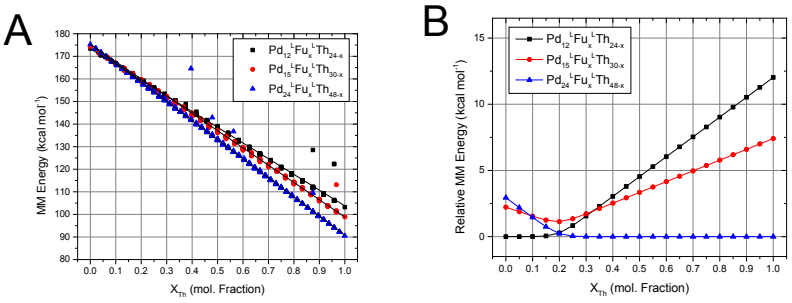


Figure S4. Linear fitting of computed per-linker energies for heteroleptic assemblies of $^{L}\mathbf{Fu}$ *and* $^{L}\mathbf{Th}$*. A, the energies found for each model heteroleptic assembly as described above (*$Pd_w{}^{L}\mathbf{Fu}_{(x)}{}^{L}\mathbf{Th}_{(2w-x)}$ *where* $x \in [0 \ldots 2w]$*, n = 50), and associated linera fit. B, relative per-linker energies for each topology computed by interpolation of the linear function shown in A.*

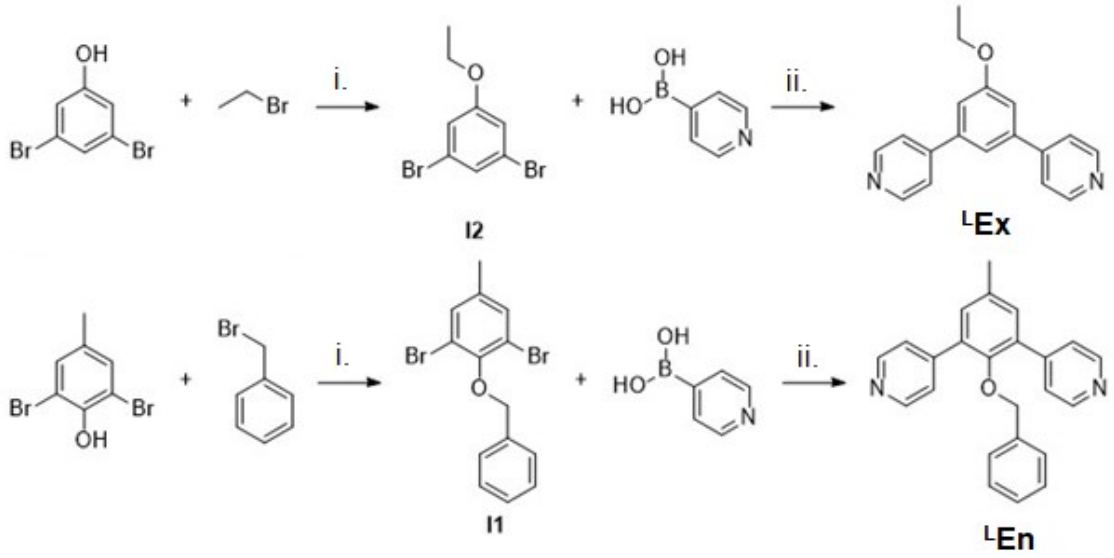
Most energies of the 30 topologies of a given heteroleptic assembly of certain composition ($Pd_w{}^{L}\mathbf{Fu}_{(x)}{}^{L}\mathbf{Th}_{(2w-x)}$) are so close that they are represented by single dots in figure S4, with the exception of a few high-energy outliers. Visual inspection of the high energy outliers show that this is caused by inversion of the linkers. If low-energy outliers were present, they may indicate a particularly favorable arrangement due to a neighboring effect as is found in shape-complementary heteroleptic assemblies. The computed energies form a linear trend with respect to the mole fraction of $^{L}\mathbf{Th}$. The resulting energies were then fit with a linear interpolation (Figure S4A) allowing energies to be estimated for each topology at arbitrary compositions (Figure S4B).

In this case we can clearly see that the primary product topology changes from a $Pd_{12}L_{24}$ to $Pd_{24}L_{48}$ at an average composition of $Pd_{1.0}{}^{L}Fu_{1.6}{}^{L}Th_{0.4}$. This process was applied to the analysis of heteroleptic assemblies of **^{L}Fu**+**^{L}Th** and **^{L}Ex**+**^{L}En** to result in topological predictions inclusive of all topologies (shown in Table S4) were conducted using a Boltzmann fitting (Equation S1) as presented in the main text (Figures 4 and 5 respectively).

D. Synthesis and Characterization of Linkers and Complexes

Table S5. Chemicals and suppliers used for synthetic procedures listed in this section. Unless otherwise noted, all solvents and other chemicals were used as received.

Chemical name	CAS no.	Supplier
3,5-dibromofuran	32460-00-7	Fluorochem Ltd.
4-pyridinylboronic acid	1692-15-5	Fluorochem Ltd.
2,6-dibromo-4-methylphenol	608-33-3	Fluorochem Ltd.
3,5-dibromophenol	626-41-5	Fluorochem Ltd.
Benzyl bromide	100-39-0	Fluorochem Ltd.
Iodoethane	75-03-6	alfa aesar GmbH & Co KG
Pd $(CH_3CN)_4$ $(BF_4)_2$	21797-13-7	Strem
d_3-acetonitrile	2206-26-0	Sigma Aldrich (Merck)
Solvents and misc.		VWR



Scheme S2. Overview for synthesis of ***^{L}Ex*** *(top) and* ***^{L}En*** *(bottom). Conditions: i) K_2CO_3, reflux, 16 h; ii) Pd(dppf)Cl_2, K_2CO_3, dioxane/H_2O (7:1), 95°C, 48 h.*

I2: The intermediate **I2** was synthesized by a modified literature synthesis.[13] Briefly 2,6-dibromophenol (1.5 g, 6 mmol), K_2CO_3 (1 g, 7 mmol) and ethyl bromide (0.7 g, 6.6 mmol) were suspended in DMF (100 mL). The mixture was stirred at room temperature for 16 h. After removal of the volatiles under reduced pressure, the residue was dissolved in Et_2O (300 mL) and washed with $NaOH_{(aq)}$ (1M, 3 × 200 mL). The organic phase was dried (Na_2SO_4), and the volatiles were removed under reduced pressure to yield **I2** colourless oil (1.6 g, 95%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.23 (t, 1H), 6.97 (d, 2H), 3.98 (q, 2H), 1.40 (t, 3H).^{13}C NMR (101 MHz, $CDCl_3$) δ 160.15, 126.21, 123.06, 116.93, 77.35, 77.03, 76.72, 64.16, 14.58.

^{L}Ex: **I2** (1 g, 3.58 mmol), 4-pyridinylboronic acid (1.8 g, 14.33 mmol), and of K_2CO_3 (11 g, 79 mmol) were suspended in a mixture of dioxane (60 mL) and water (10 mL). Nitrogen was bubbled into the solution for 20 min before $Pd(dppf)Cl_2$ (270 mg, 0.358 mmol) was added directly as a solid. The mixture was heated to 95°C for 48 h while stirring. After cooling the mixture to room temperature, the volatiles were removed under reduced pressure. The product was extracted into ethyl acetate (300 mL) and washed with $NaOH_{(aq)}$ (1M, 3 × 200 mL). The organic phase was dried (Na_2SO_4), and the volatiles were removed under reduced pressure. The crude material was purified by column chromatography (SiO_2, 3% MeOH: DCM) to afford **^{L}Ex** (0.61 g, 62%) as an off-white solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ (ppm) = 8.73 (br s, 4H), 7.56 (br s, 4H), 7.45 (s, ^{1}H), 7.23 (s, 2H), 4.20 (m, 2H), 1.50 (m, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ (ppm) = 160.16, 126.23, 123.09, 116.94, 64.17, 14.62.

I1: The intermediate **I1** was synthesized by a modified literature synthesis.[13] 2,6-dibromo-4-methylphenol (3 g, 11.28 mmol), K_2CO_3 (4.7 g, 33.8 mmol), and benzyl bromide (2.1 g, 12.4 mmol) were suspended in acetone (100 mL). The mixture was heated at reflux for 16 hours. After removal of the volatiles under reduced pressure, the residue was dissolved in Et_2O (300 mL) and washed with $NaOH_{(aq)}$ (1M, 3 × 200 mL). The organic phase was dried (Na_2SO_4) and the volatiles were removed under reduced pressure to yield **I1** (3.81 g, 95%) as an off-white solid. The ^{1}H NMR is consistent with the reported spectrum.[13]

^{L}En: **I1** (1.15 g, 3.25 mmol), 4-pyridinylboronic acid (1 g, 8.13 mmol), and K_2CO_3 (6 g, 70 mmol) were suspended in a mixture of dioxane (80 mL) and water (16 mL). Nitrogen was bubbled into the solution for 20 min before the addition of $Pd(dppf)Cl_2$ (240 mg, 0.325 mmol). The mixture was heated to 95°C while stirring for 48 h. After cooling the mixture to room temperature,

the volatiles were removed under reduced pressure. The product was extracted into ethyl acetate (300 mL) and washed with $NaOH_{(aq)}$ (1M, 3 × 200 mL). The organic phase was dried (Na_2SO_4) and the volatiles were removed under reduced pressure. The crude material was purified by column chromatography (SiO_2, 3% MeOH: DCM) to afford (0.8 g, 70%) of **^{L}En** as an off-white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.74 (br s, 4H), 7.58 (br s, 4H), 7.26–7.13 (m, 6H), 6.68 (d, 2H), 4.17 (s, 2H), 2.45 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 151.15, 149.41, 146.20, 135.50, 134.88, 133.65, 131.64, 128.63, 128.27, 128.25, 124.72, 77.43, 77.32, 77.12, 76.80, 75.94, 20.88. HRMS (ESI^+) calc. for $[C_{24}H_{20}N_2O]H^+$ 353.1648, found 353.1549.

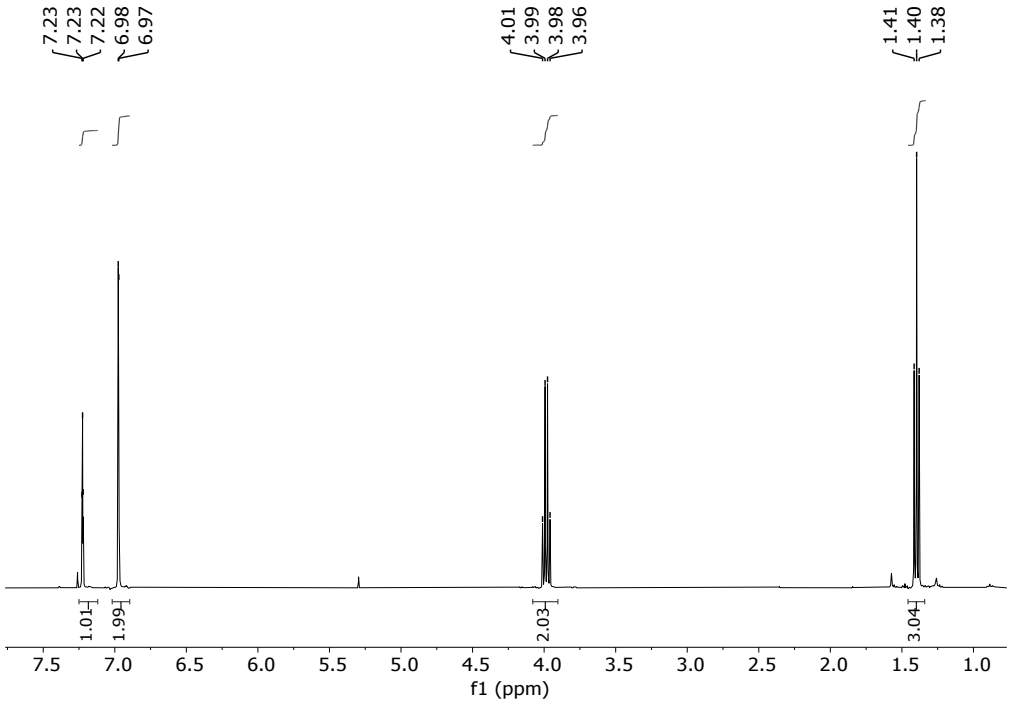


*Figure S5. ^{1}H NMR spectrum of intermediate **I2** ($CDCl_3$, 400 MHz).*

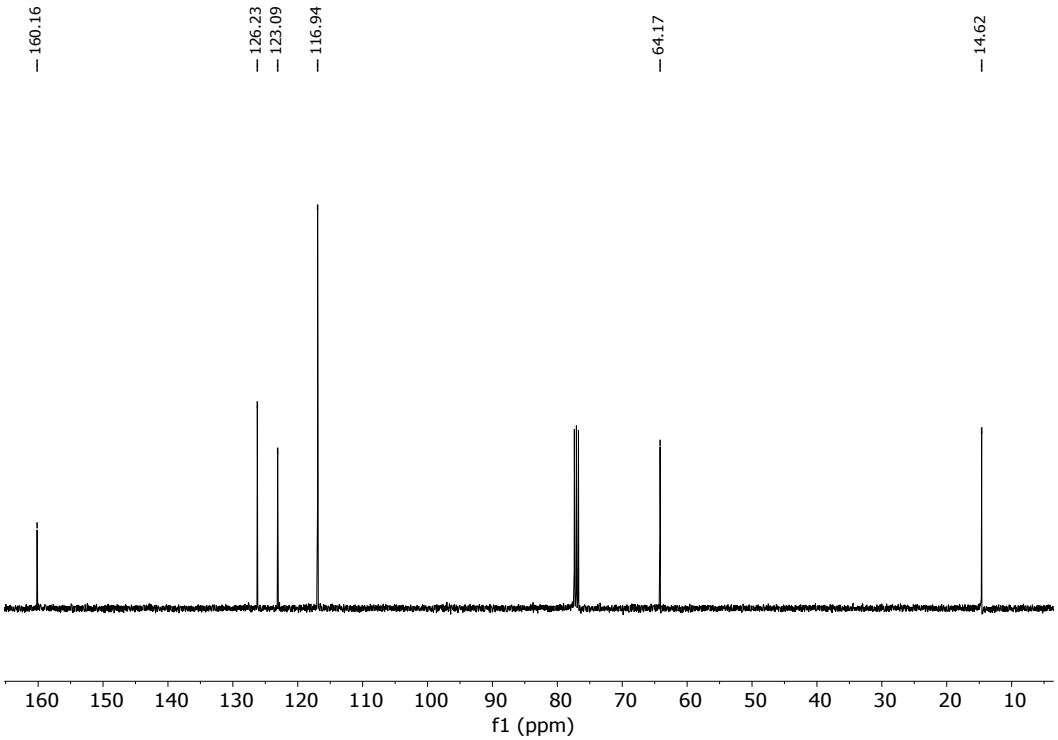


*Figure S6. ^{13}C NMR spectrum of intermediate **I2** ($CDCl_3$, 400 MHz).*

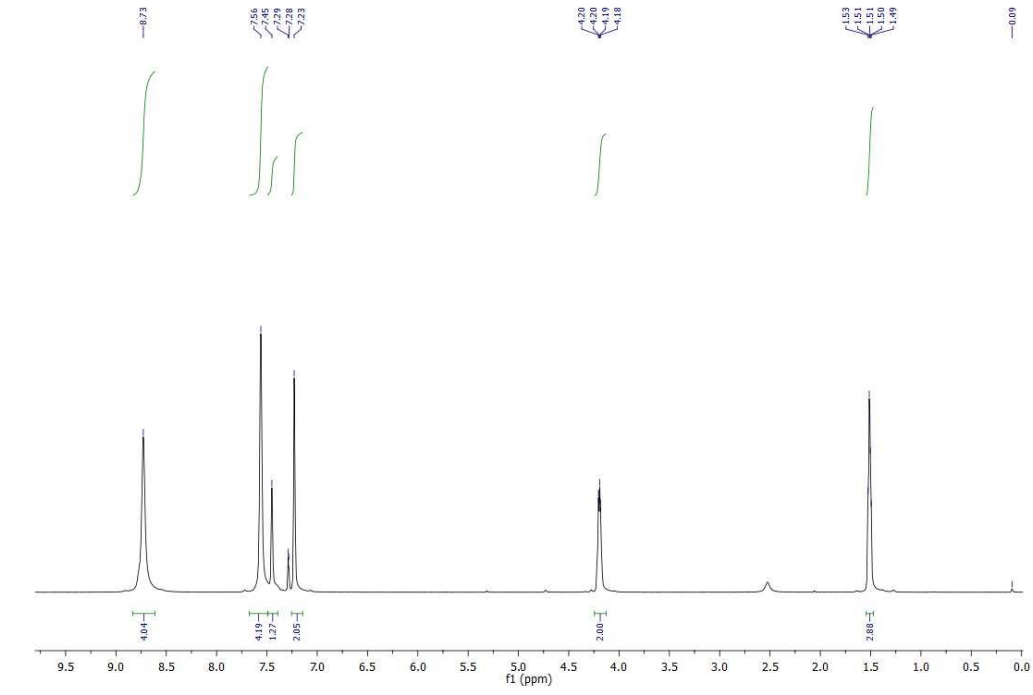


*Figure S7. ^{1}H NMR spectrum of intermediate **^{L}Ex** ($CDCl_3$, 400 MHz).*

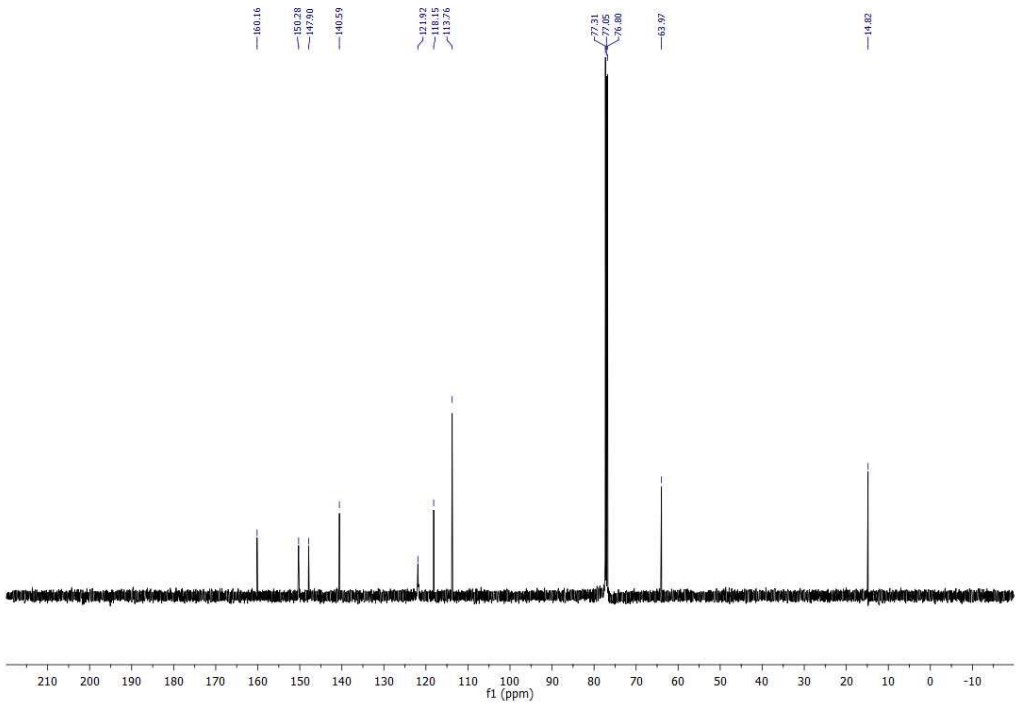


Figure S8. ^{13}CNMR spectrum of intermediate ***L-Ex*** *($CDCl_3$, 400 MHz).*

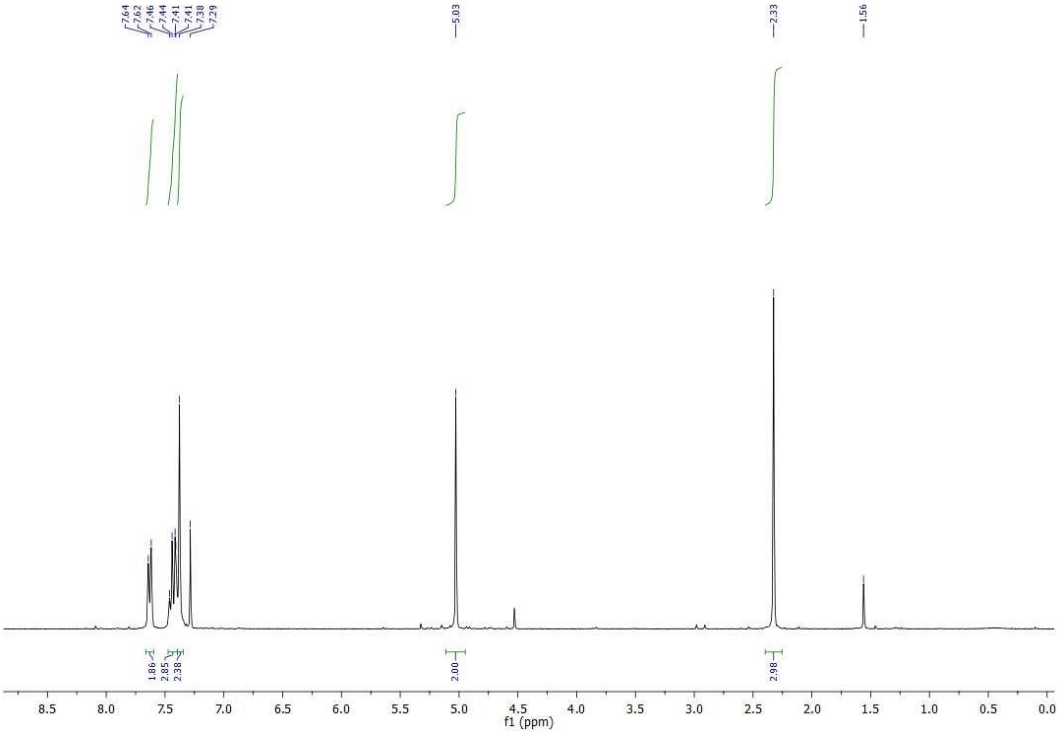


Figure S9. ^{1}H NMR spectrum of intermediate ***I1*** *($CDCl_3$, 400 MHz).*

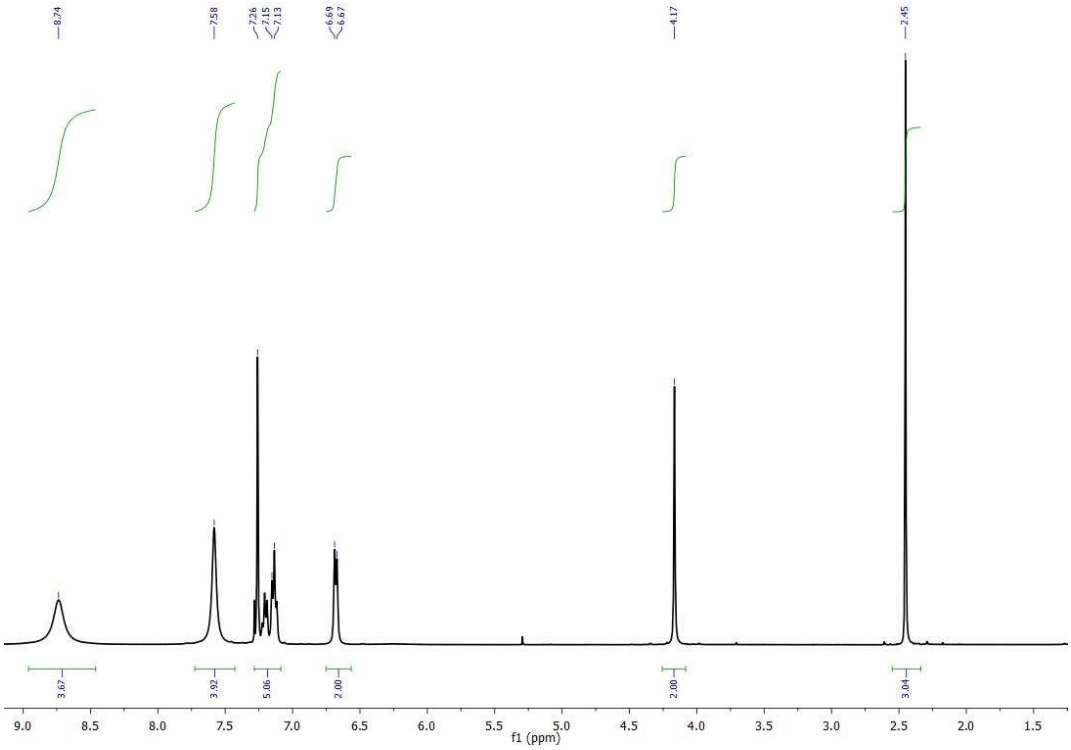


Figure S10. ^{1}H NMR spectrum of intermediate ***^{L}En*** *($CDCl_3$, 400 MHz).*

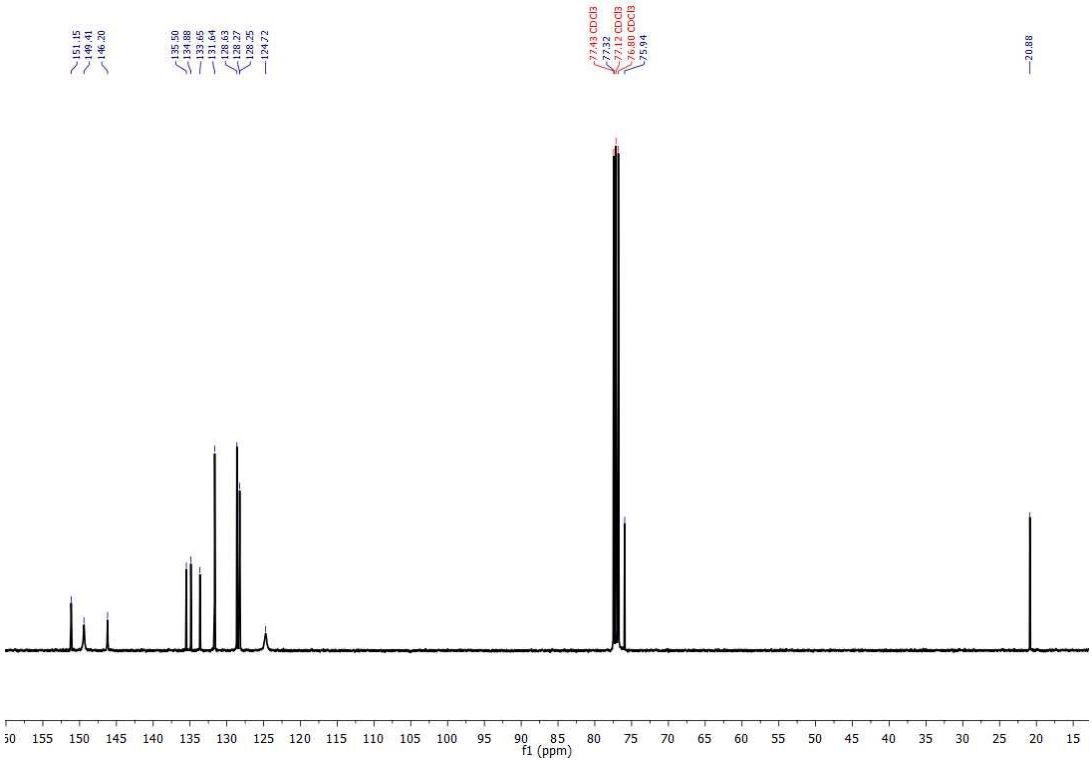


Figure S11. ^{13}C NMR spectrum of intermediate ***^{L}En*** *($CDCl_3$, 400 MHz).*

E. Synthesis of $Pd_x{}^{L}Fu_{2x}$ Homoleptic Assemblies

Stock solutions of $Pd(BF_4)_2$ (50 mM) and **^{L}Fu** (11 mM) were prepared in CD_3CN. A vial was charged with the **^{L}Fu** solution (900 μL) and a small magnetic stir bar. Afterwards, the $Pd(BF_4)_2$ solution (100 μL) was added, and the vial closed with a screw-type cap. The solution was stirred 16 h at room temperature. This solution was then filtered through a Merck 0.45 μm PTFE syringe filter, and analyzed directly afterwards by NMR and ESI–HRMS.

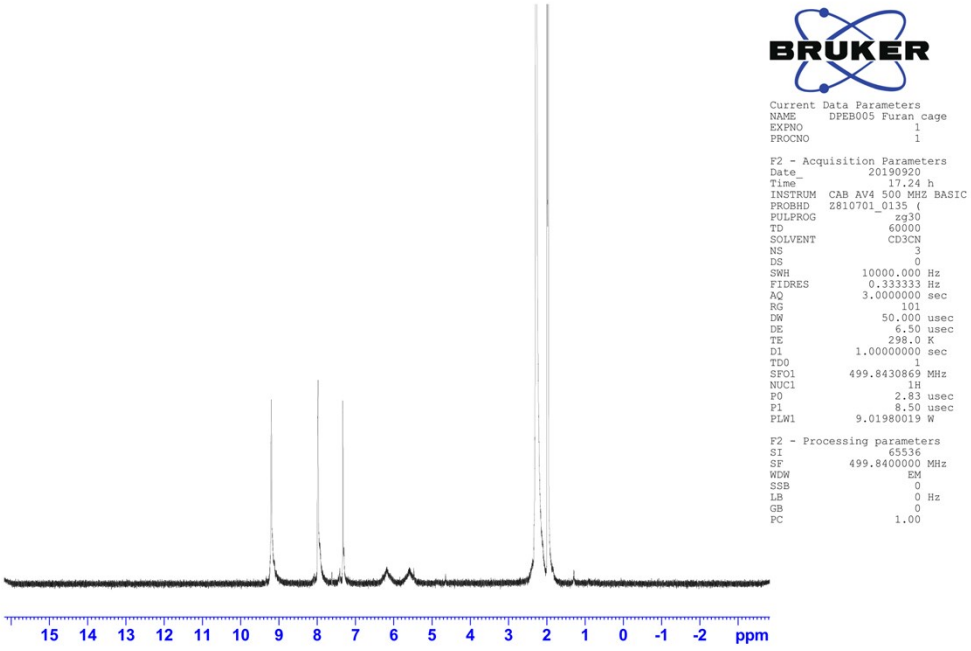
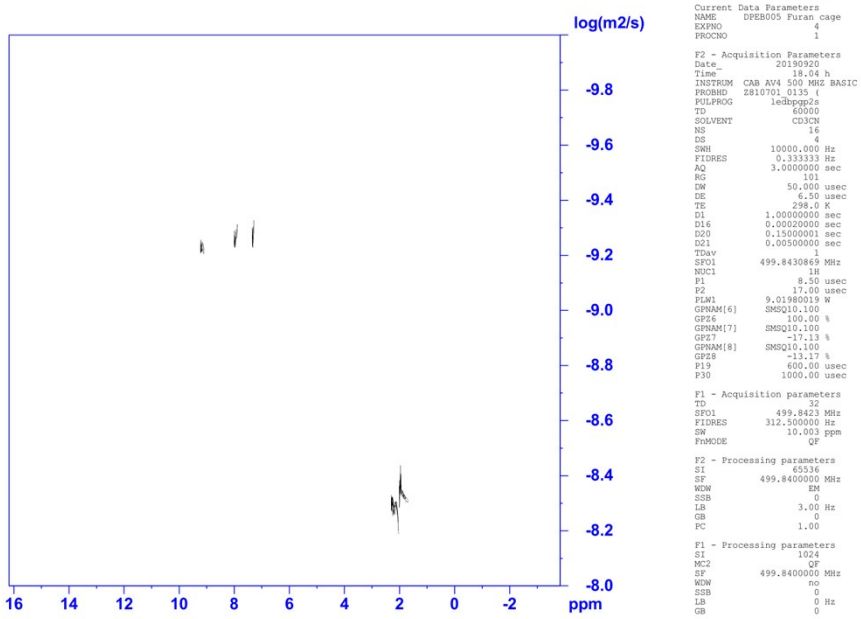


Figure S12. ^{1}H NMR spectrum obtained for Pd assemblies of ***^{L}Fu****. (CD_3CN, 500 MHz).*



Current Data Parameters
NAME DPEB005 Furan cage
EXPNO 4
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190920
Time 18.04 h
INSTRUM CAB AV4 500 MHZ BASIC
PROBHD Z810701_0135 (
PULPROG ledbpgp2s
TD 60000
SOLVENT CD3CN
NS 16
DS 4
SWH 10000.000 Hz
FIDRES 0.333333 Hz
AQ 3.0000000 sec
RG 101
DW 50.000 usec
DE 6.50 usec
TE 298.0 K
D1 1.00000000 sec
D16 0.00020000 sec
D20 0.15000001 sec
D21 0.00500000 sec
TDav 1
SFO1 499.8430869 MHz
NUC1 1H
P1 8.50 usec
P2 17.00 usec
PLW1 9.01980019 W
GPNAM[6] SMSQ10.100
GPZ6 100.00 %
GPNAM[7] SMSQ10.100
GPZ7 -17.13 %
GPNAM[8] SMSQ10.100
GPZ8 -13.17 %
P19 600.00 usec
P30 1000.00 usec

F1 - Acquisition parameters
TD 32
SFO1 499.8423 MHz
FIDRES 312.500000 Hz
SW 10.003 ppm
FnMODE QF

F2 - Processing parameters
SI 65536
SF 499.8400000 MHz
WDW EM
SSB 0
LB 3.00 Hz
GB 0
PC 1.00

F1 - Processing parameters
SI 1024
MC2 QF
SF 499.8400000 MHz
WDW no
SSB 0
LB 0 Hz
GB 0

Figure S13. DOSY NMR obtained for Pd assemblies of ^{L}***Fu***. *(CD_3CN, 500 MHz).*

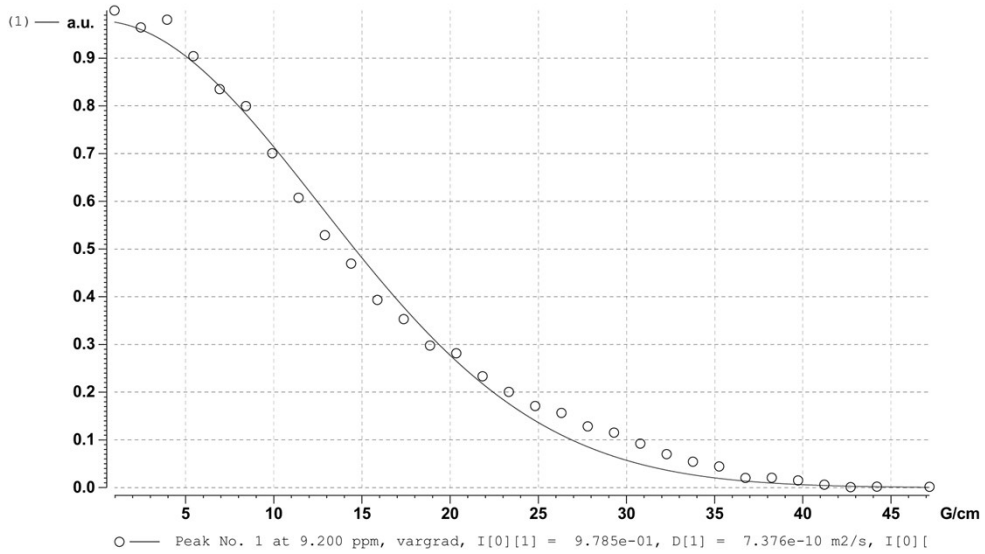


Figure S14. Plot of α-Pyridyl peak (δ ≈ 9.2 ppm, CD_3CN, 500 MHz) integral areas versus gradient strength used for fitting the diffusion decay observed during DOSY analysis of Pd assemblies of ***^{L}Fu****. Exponential function fit computed using the Topspin relaxation module.*

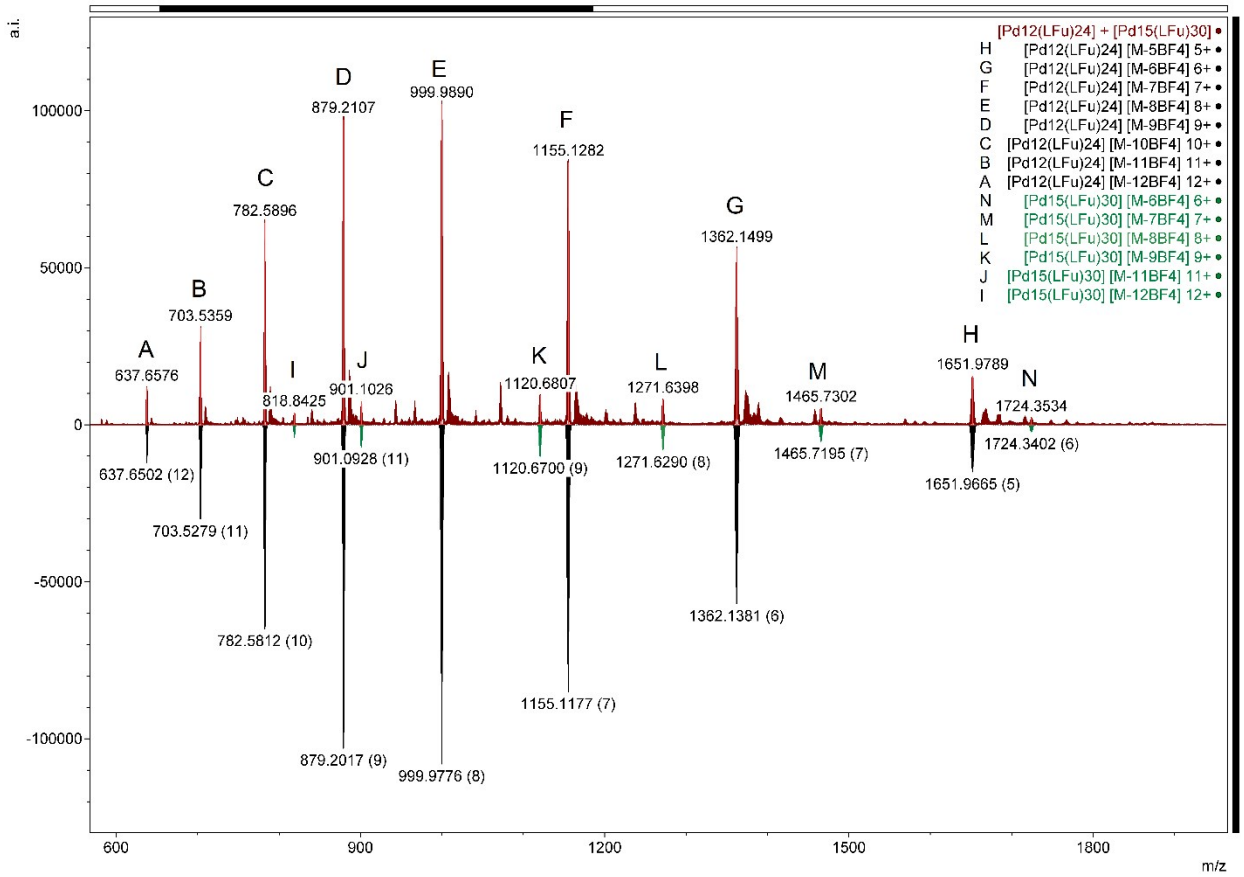


Figure S15. Overview of ESI–HRMS spectra obtained for products derived from self-assembly of 10 mM ^{L}***Fu*** *and 5 mM* $Pd(BF_4)_2(CH_3CN)_4$ *in* CD_3CN*. Observed peaks (red) correspond to simulated peaks for multiple charged species of* Pd_{12}^{L}***Fu***$_{24}$ *and* Pd_{15}^{L}***Fu***$_{30}$ *assemblies respectively. Individual ion spectra shown below.*

F. Synthesis and Characterization of $Pd_x{}^{L}Ex_{2x}$ Homoleptic Assemblies

Stock solutions of $Pd(BF_4)_2$ (50 mM) and **^{L}Ex** (11 mM) were prepared in CD_3CN. a vial was charged with the **^{L}Ex** solution (900 μL), and a small magnetic stir bar, afterwards, the $Pd(BF_4)_2$ solution (100 μL) was added, and the vial closed with a screw-type cap. The solution was stirred 16 h at 50°C. This solution was then filtered through a Merck 0.45 μm PTFE syringe filter, and analyzed by NMR and ESI–HRMS.

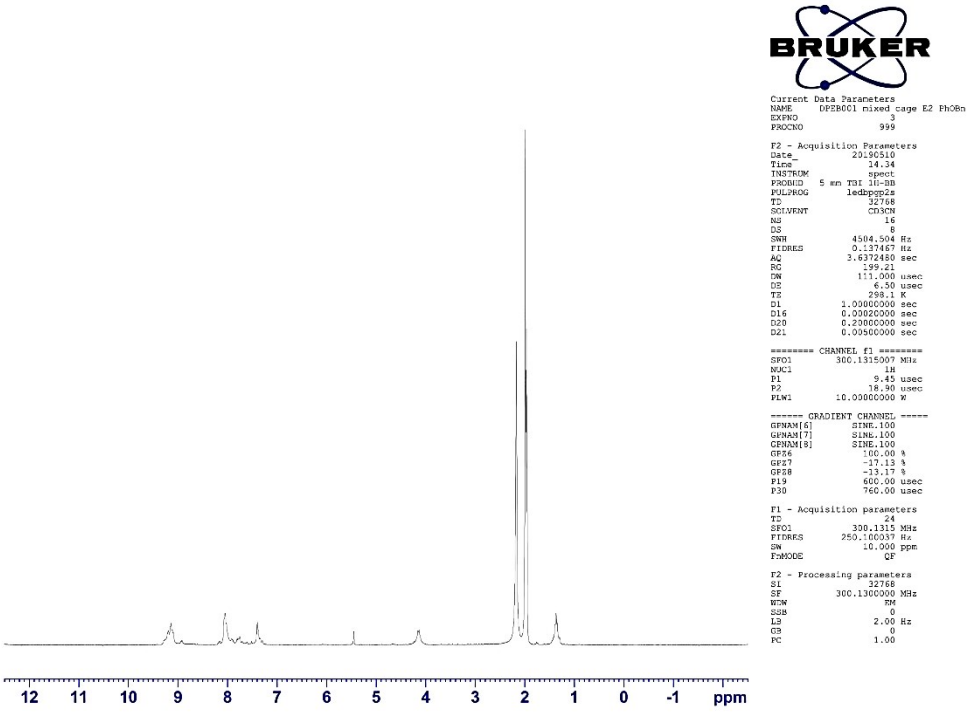


*Figure S16. ^{1}H NMR spectrum obtained for Pd assemblies of **^{L}Ex** (CD_3CN, 300 MHz).*

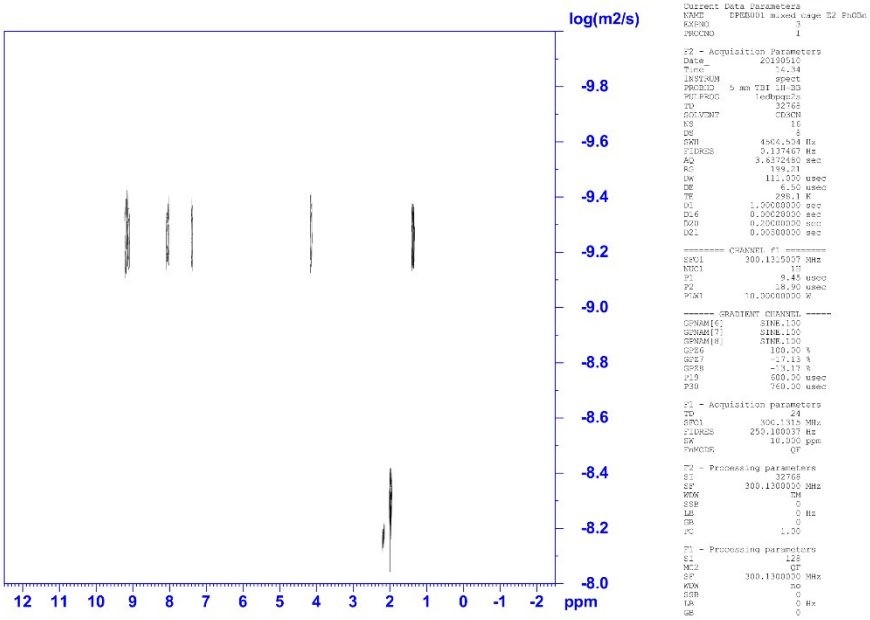


Figure S17. DOSY NMR obtained for Pd assemblies of ^{L}***Ex*** *(CD_3CN, 300 MHz).*

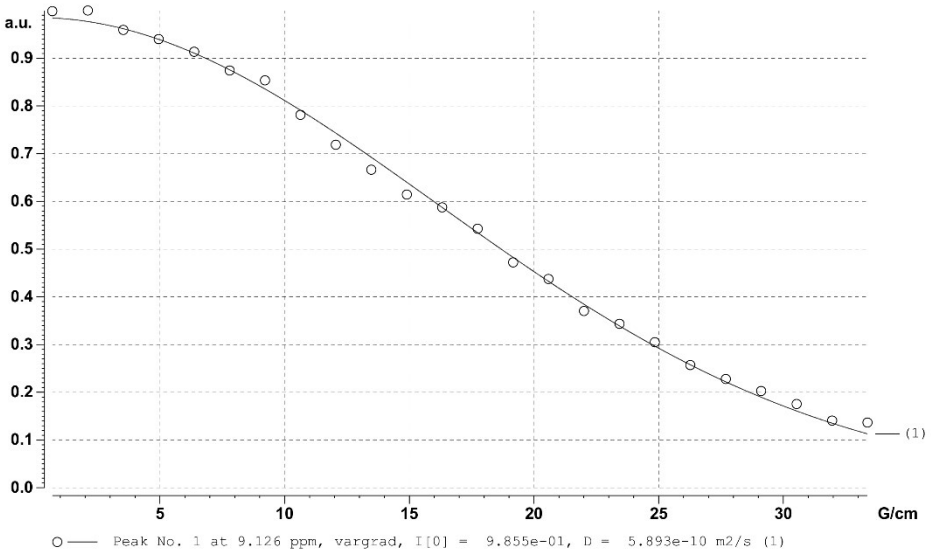


Figure S18. α-Pyridyl peak integrals showing the diffusion decay observed during DOSY analysis of Pd assemblies of ***L*****Ex**. *Exponential function fit computed using the Topspin relaxation module.*

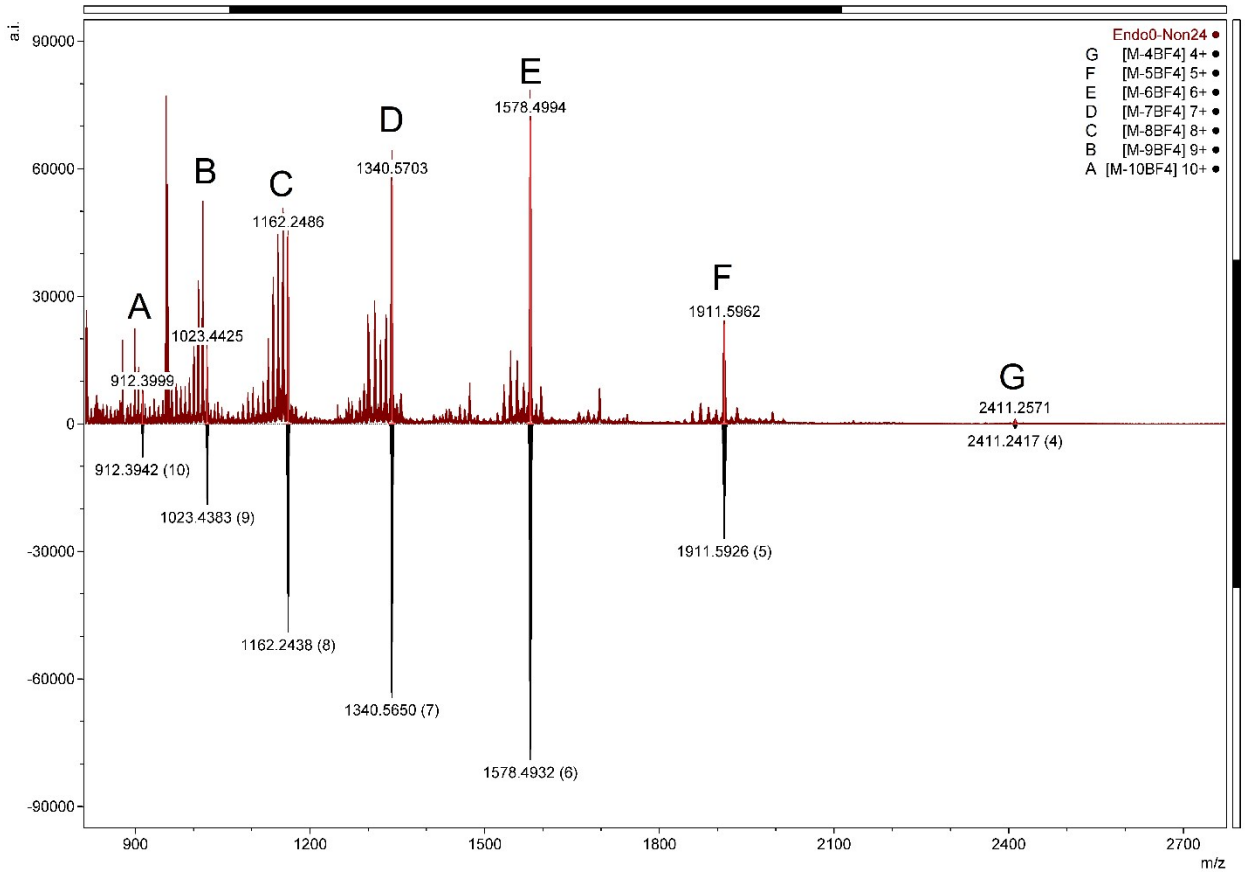


Figure S19. ESI–HRMS spectra obtained from the 12 h self-assembly of 10 mM ^{L}***Ex*** *and 5 mM* $Pd(BF_4)_2$ *in* CD_3CN *at 50°C. Top, the observed mass spectra with annotations indicating the peaks identified as* $Pd_{12}{}^{L}\mathbf{Ex}_{24}(BF_4)_{(24-Z)}{}^{Z+}$ *(red), shown also are simulations of the isotope distribution patterns for the annotated species (black). Peaks were observed for many assembly species at incrementally smaller masses, we attributed these to the ionization loss of ethyl groups during ionization. The intensity of these ionization product peaks increases at higher charge numbers.*

G. Synthesis and Characterization of $Pd_x{}^{L}En_{2x}$ Homoleptic Assemblies

Assemblies were formed under identical conditions as for **^{L}Ex** assemblies.

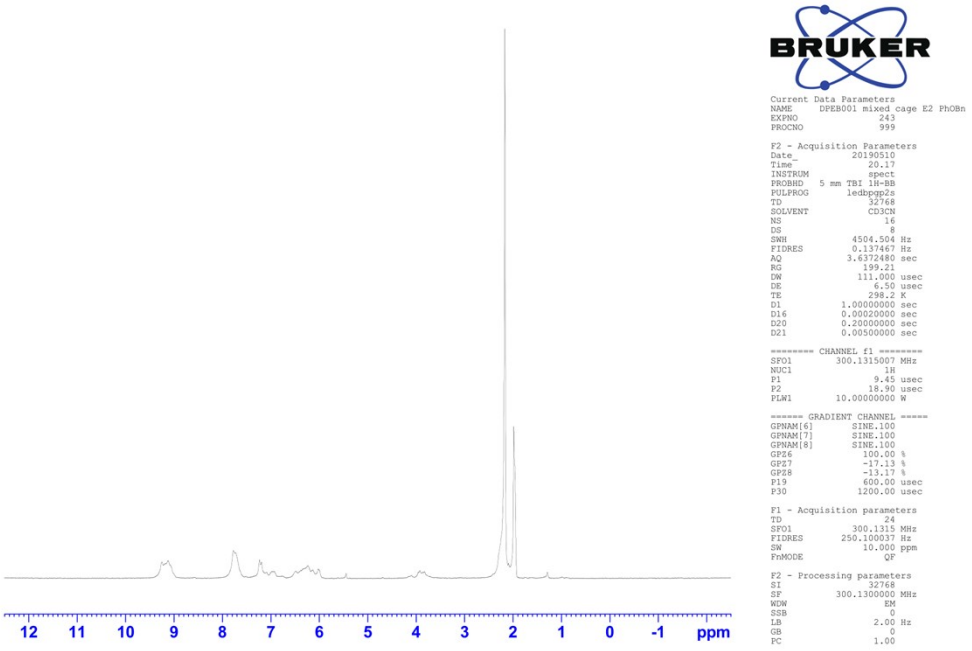
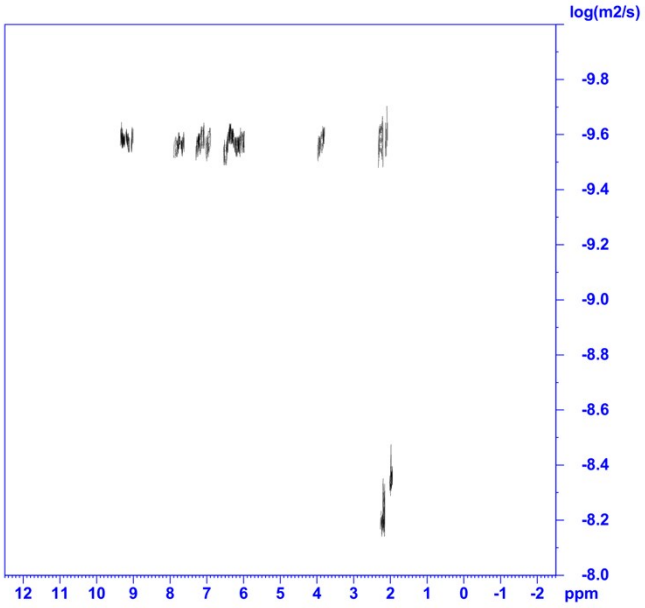
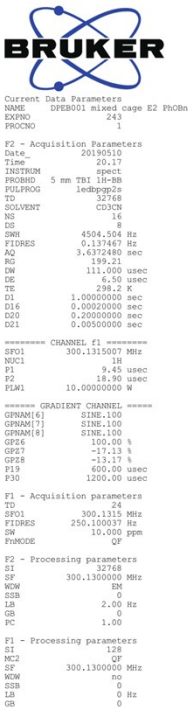


*Figure S20. ^{1}H-NMR spectrum obtained for Pd assemblies of **^{L}En**. (CD_3CN, 300 MHz)*





```
Current Data Parameters
NAME      DFEB001 mixed cage E2 PhOBn
EXPNO                 243
PROCNO                  1

F2 - Acquisition Parameters
Date_           20190510
Time                20.17
INSTRUM             spect
PROBHD   5 mm TBI 1H-BB
PULPROG         ledbpgp2s
TD                  32768
SOLVENT             CD3CN
NS                     16
DS                      8
SWH              4504.504 Hz
FIDRES           0.137467 Hz
AQ               3.6372480 sec
RG                 199.21
DW                111.000 usec
DE                  6.50 usec
TE                 298.2 K
D1            1.00000000 sec
D16           0.00020000 sec
D20           0.20000000 sec
D21           0.00500000 sec

======== CHANNEL f1 ========
SFO1         300.1315007 MHz
NUC1                   1H
P1                   9.45 usec
P2                  18.90 usec
PLW1         10.00000000 W

====== GRADIENT CHANNEL =====
GPNAM[6]          SINE.100
GPNAM[7]          SINE.100
GPNAM[8]          SINE.100
GPZ6               100.00 %
GPZ7               -17.13 %
GPZ8               -13.17 %
P19                600.00 usec
P30               1200.00 usec

F1 - Acquisition parameters
TD                     24
SFO1            300.1315 MHz
FIDRES        250.100037 Hz
SW                10.000 ppm
FnMODE                 QF

F2 - Processing parameters
SI                  32768
SF            300.1300000 MHz
WDW                    EM
SSB                     0
LB                   2.00 Hz
GB                      0
PC                   1.00

F1 - Processing parameters
SI                    128
MC2                    QF
SF            300.1300000 MHz
WDW                    no
SSB                     0
LB                      0 Hz
GB                      0
```

Figure S21. DOSY NMR obtained for Pd assemblies of ^{L}***En***. *(CD_3CN, 300 MHz)*

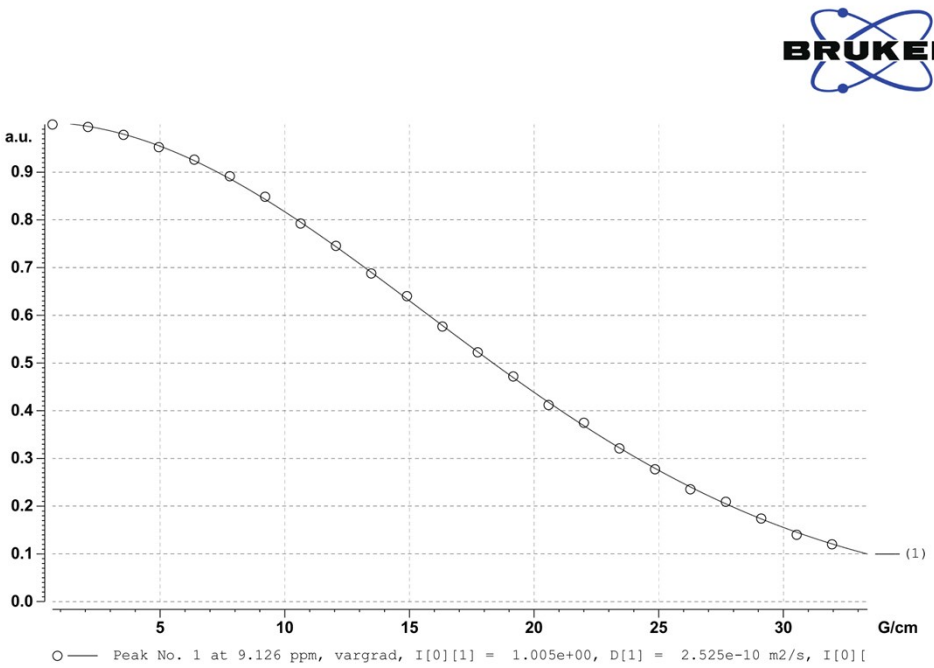


Figure S22. α-Pyridyl integrals showing the diffusion decay observed during DOSY analysis of Pd assemblies of ***L*****Ex**. *Exponential function fit computed using the Topspin relaxation module.*

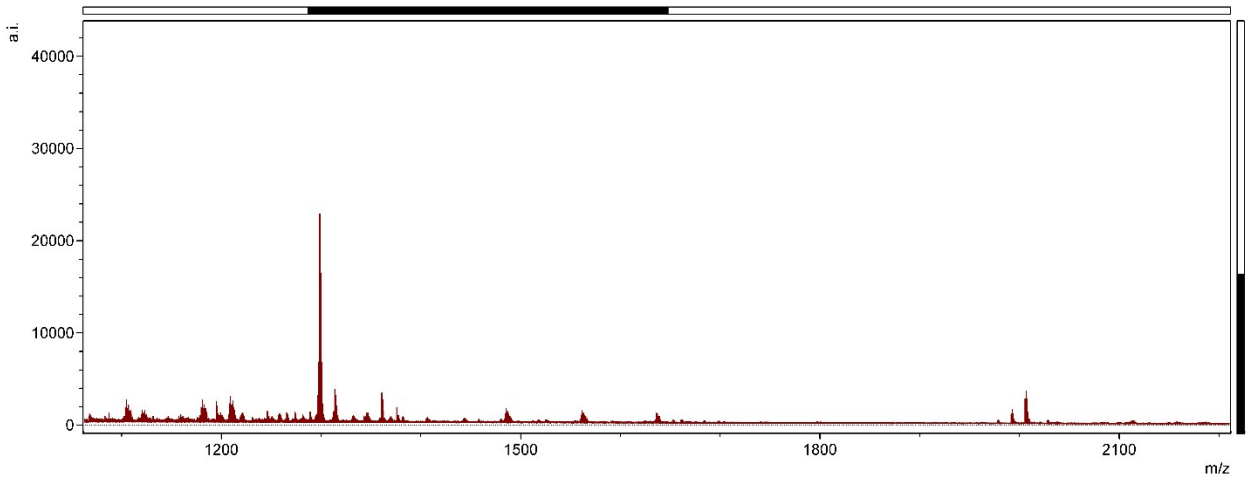


Figure S23. ESI–HRMS spectra obtained from the 12 hour self-assembly of 10 mM ^{L}***En*** *and 5 mM* $Pd(BF_4)_2$ *in* CD_3CN *at 50°C. The observed mass spectra, shown in red, is lacking in peaks corresponding to any assembly formation. Peaks corresponding to free building block, or polymeric species are also absent.*

H. Synthesis and Characterization of $Pd_x{}^{L}En_{(2x-n)}{}^{L}Ex_n$ Heteroleptic Assemblies

Stock solutions of $Pd(BF_4)_2$ (50 mM) and **^{L}Ex** (11 mM), and **^{L}En** (11 mM) were prepared in CD_3CN. A vial was charged with a mixture of the two solutions (900 µL), and a small magnetic stir bar, afterwards the $Pd(BF_4)_2$ solution (100 µL) was added, and the vial closed with a screw-type cap. The solution was stirred 16 h at 50°C. The resulting solution was then filtered through a Merck 0.45 µm PTFE syringe filter, and analyzed directly afterwards by NMR and ESI–HRMS.

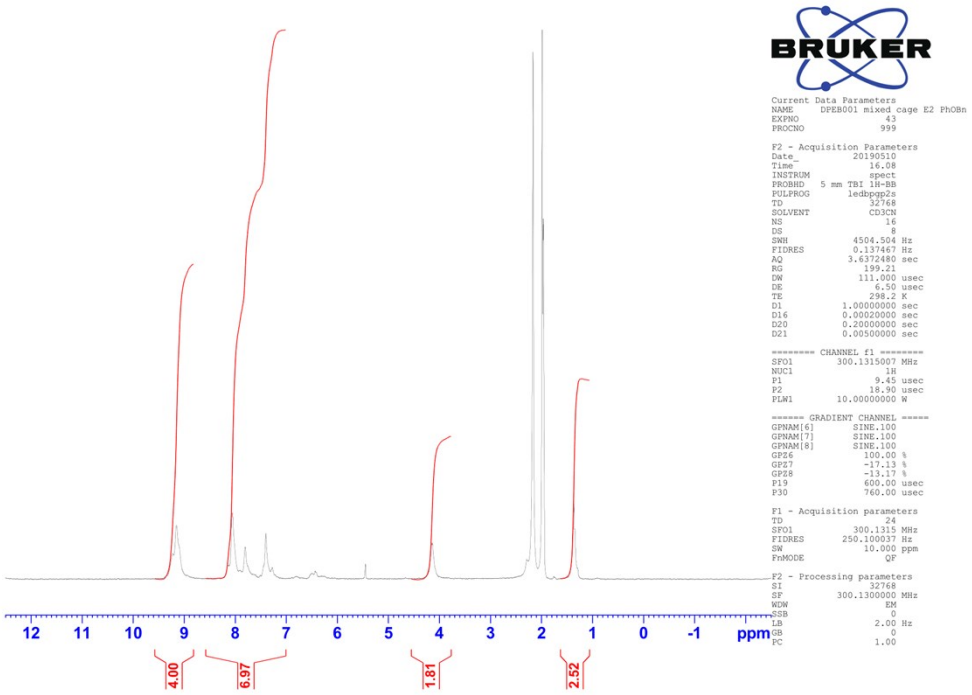


*Figure S24. ^{1}H NMR spectrum obtained for homoleptic Pd assemblies of **^{L}Ex** and **^{L}En**, est. 20.2:3.8 ratio (CD_3CN, 300 MHz).*

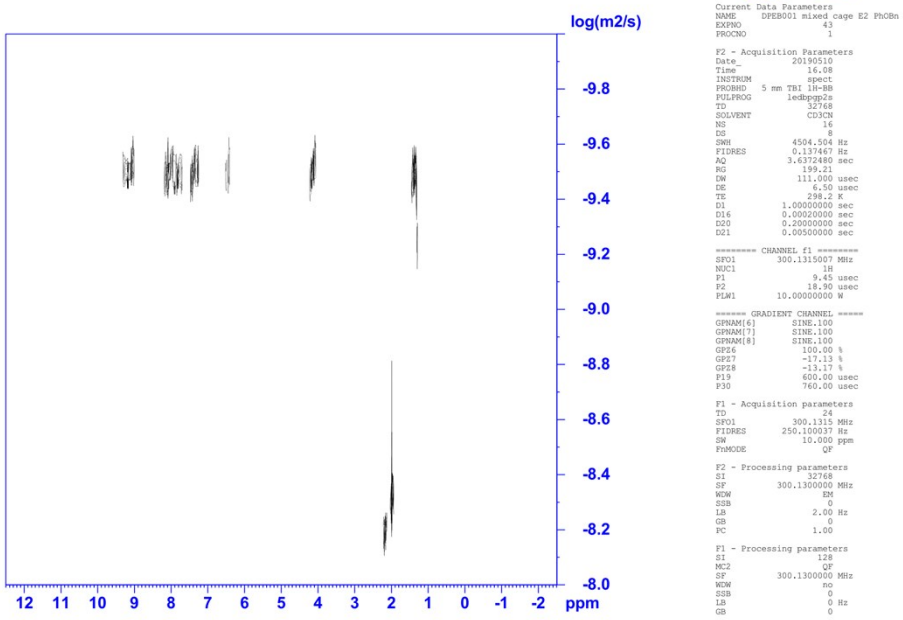


Figure S25. DOSY NMR obtained for Pd assemblies of ^{L}***Ex*** *and* ^{L}***En****, est. 20.2:3.8 ratio (*CD_3CN*, 300 MHz*).

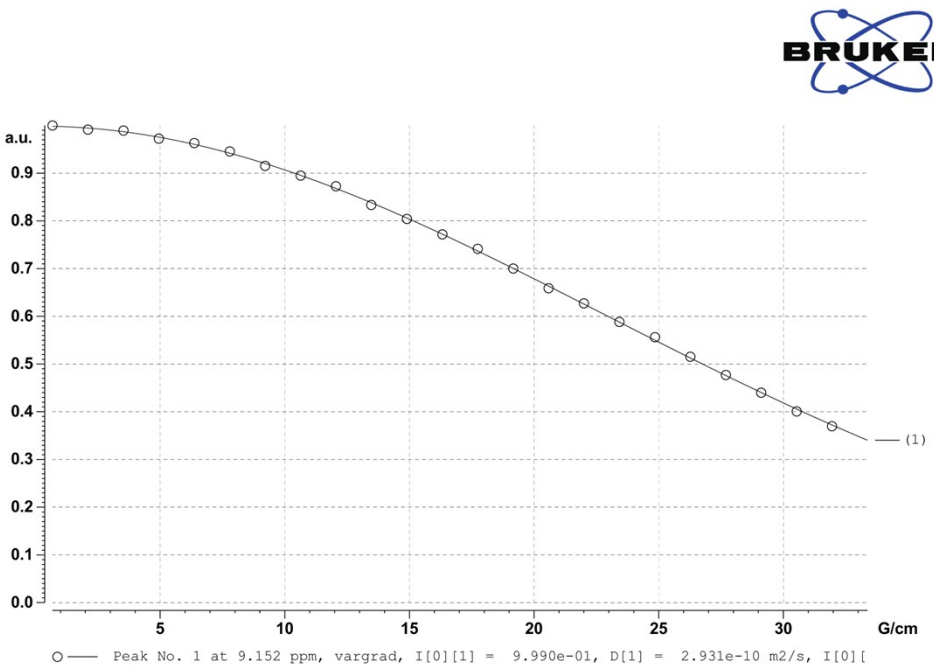


Figure S26. α-Pyridyl peak integrals showing the diffusion decay observed during DOSY analysis of heteroleptic assemblies derived from ***^{L}Ex*** *and* ***^{L}En****, est. 20.2:3.8 ratio. Exponential function fit computed using the Topspin relaxation module.*

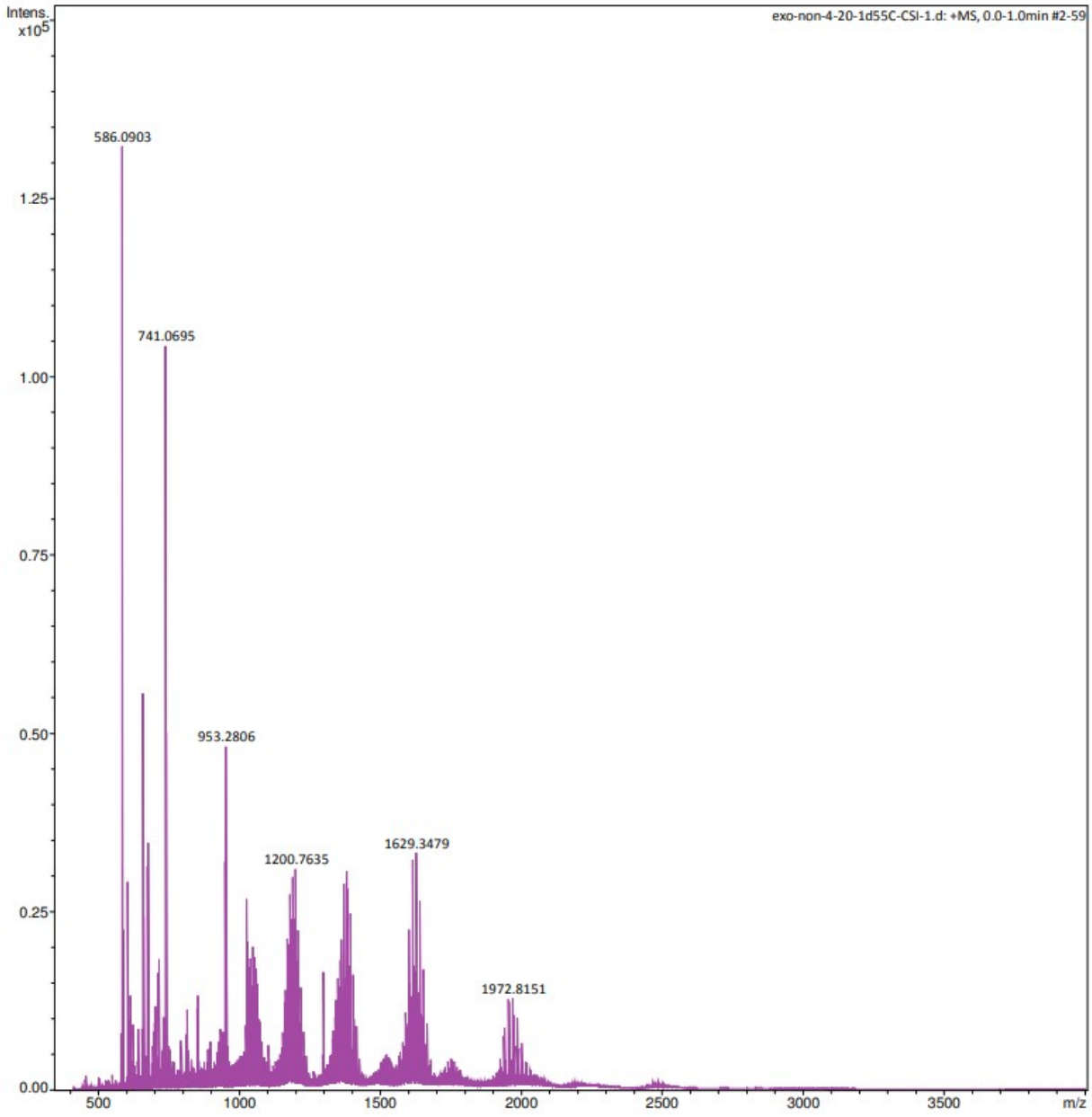


Figure S27. ESI–HRMS spectra of heteroleptic assemblies derived from ***^{L}Ex*** *and* ***^{L}En****, est. 20.2:3.8 ratio showing numerous charged species.*

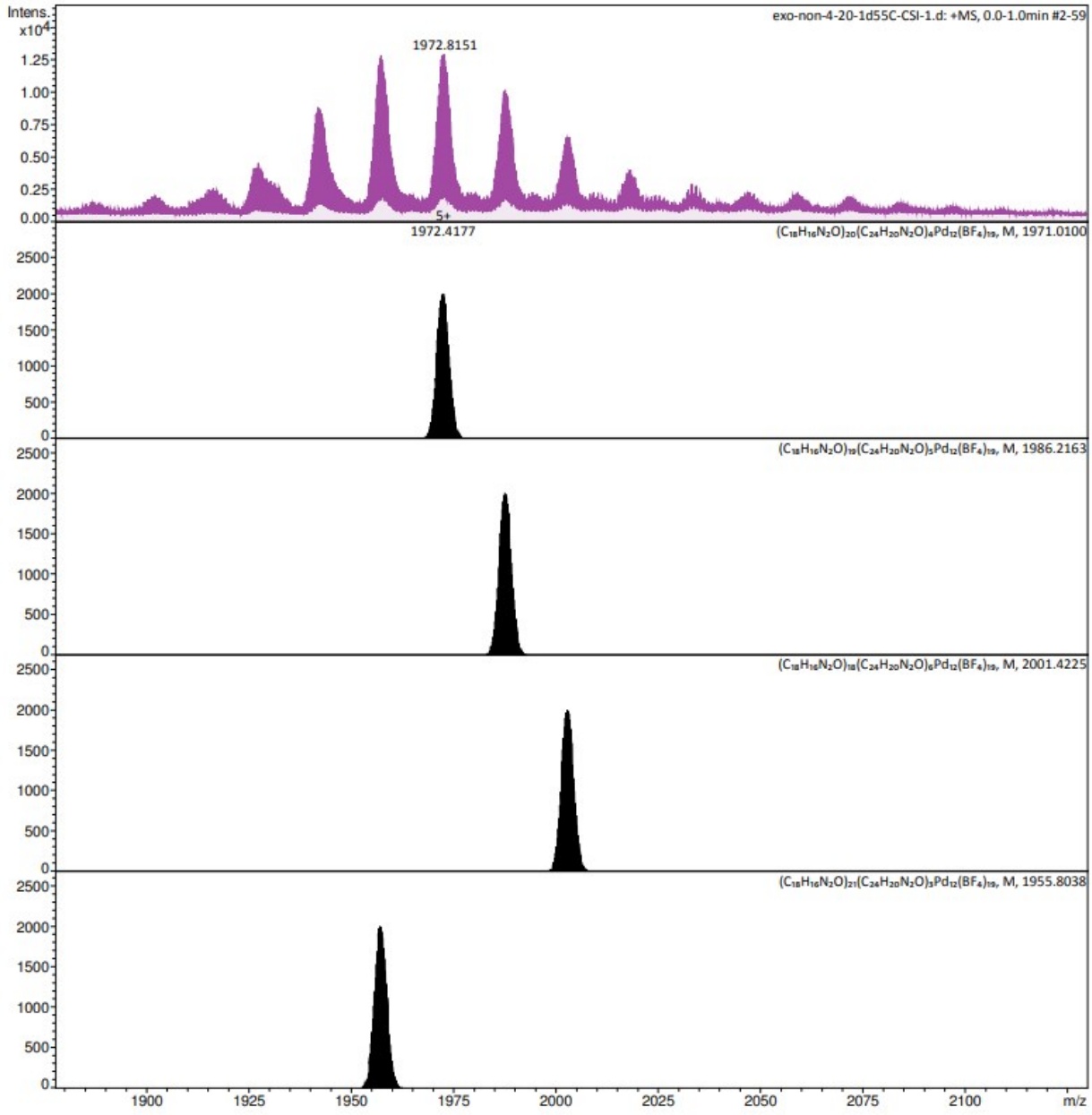


Figure S28. ESI–HRMS spectra of heteroleptic assemblies derived from ^{L}***Ex*** *and* ^{L}***En****, est. 20.2:3.8 ratio, expanded to show peaks of various linker compositions (purple). Mass simulations of* $Pd_{12}{}^{L}\mathbf{Ex}_{(x)}{}^{L}\mathbf{En}_{(24-x)}(BF_4)_{19}{}^{5+}$ *where* $x \in [18\text{-}21]$ *(black).*

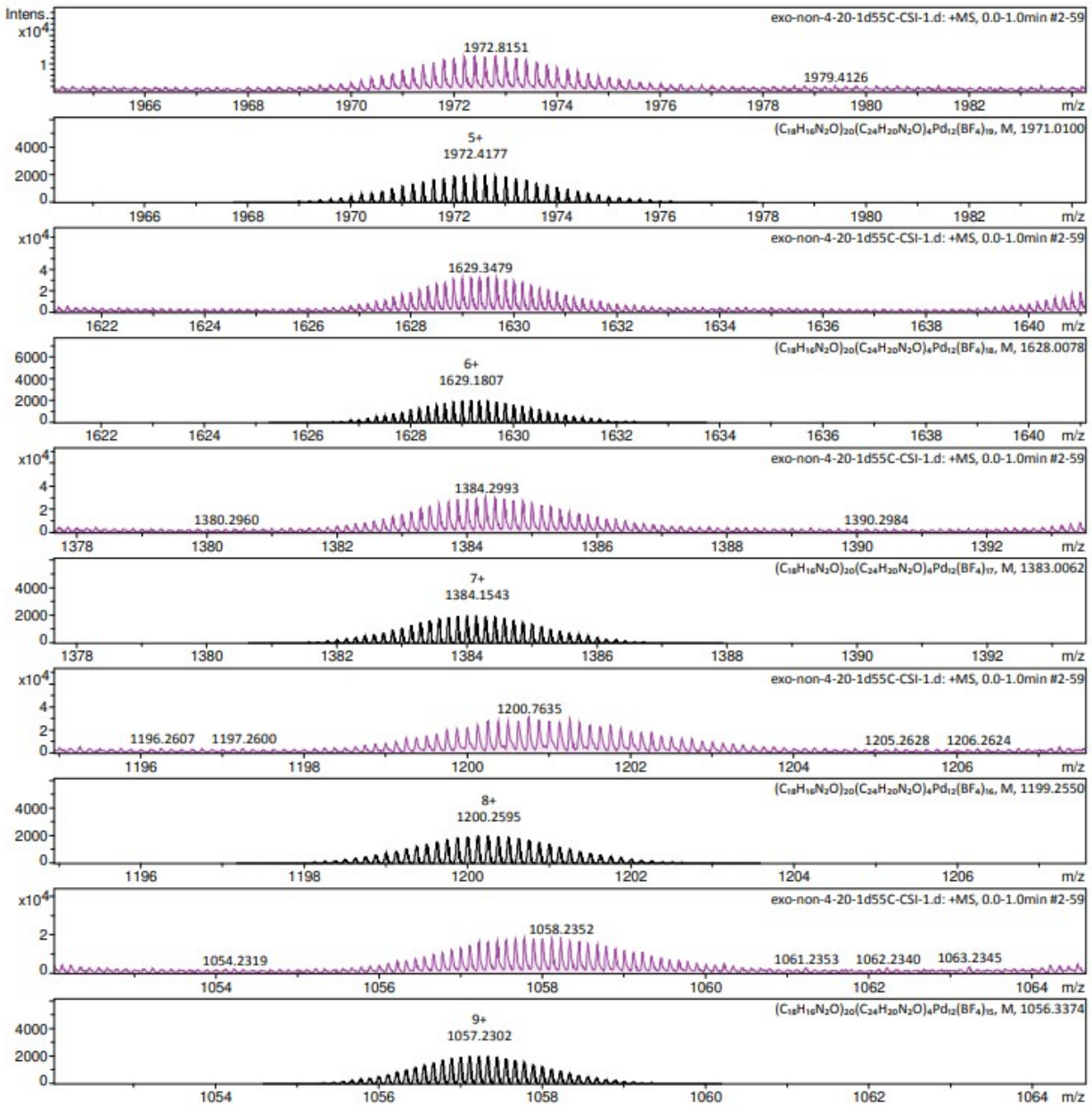


Figure S29. ESI–HRMS spectra of heteroleptic assemblies derived from ^{L}***Ex*** *and* ^{L}***En****, est. 20.2:3.8 ratio, expanded to show multiple charged peaks (purple). Simulations of the isotope distribution patterns for* $Pd_{12}{}^{L}\mathbf{Ex}_{(20)}{}^{L}\mathbf{En}_{(4)}\,(BF_4)_{(24-x)}{}^{x+}$ *where* $x \in [5\text{-}9]$ *(black).*

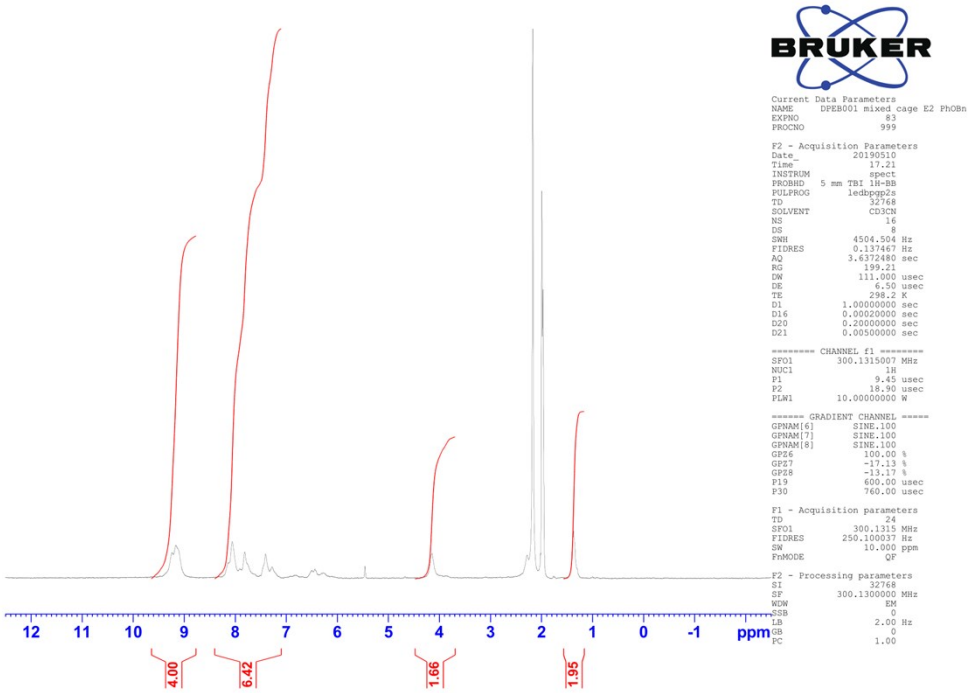


Figure S30. ^{1}H NMR spectrum obtained for homoleptic Pd assemblies of ^{L}***Ex*** *and* ^{L}***En****, est. 15.6:8.4 ratio (CD_3CN, 300 MHz).*

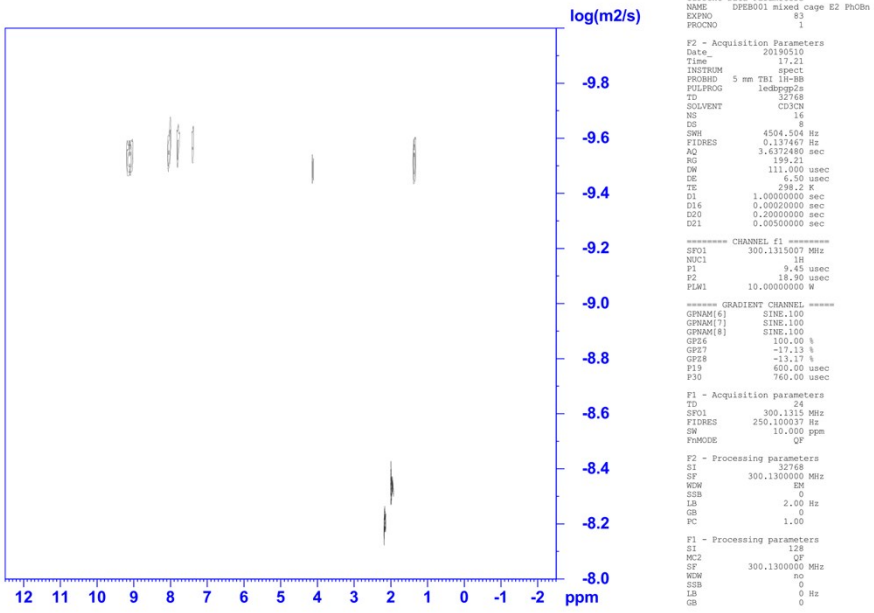


Figure S31. DOSY NMR obtained for Pd assemblies of ^{L}***Ex*** *and* ^{L}***En****, est. 15.6:8.4 ratio. (*CD_3CN*, 300 MHz*)

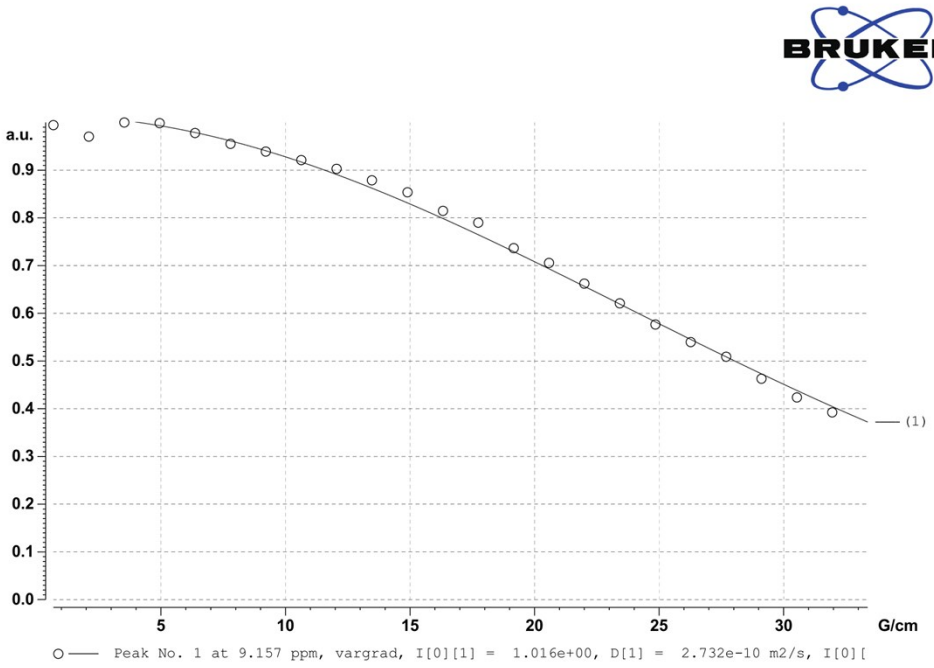


Figure S32. α-Pyridyl peak integrals showing the diffusion decay observed during DOSY analysis of Pd assemblies of ^{L}***Ex*** *and* ^{L}***En****, est. 15.6:8.4 ratio of* ^{L}***En****. Exponential function fit computed using the Topspin relaxation module.*

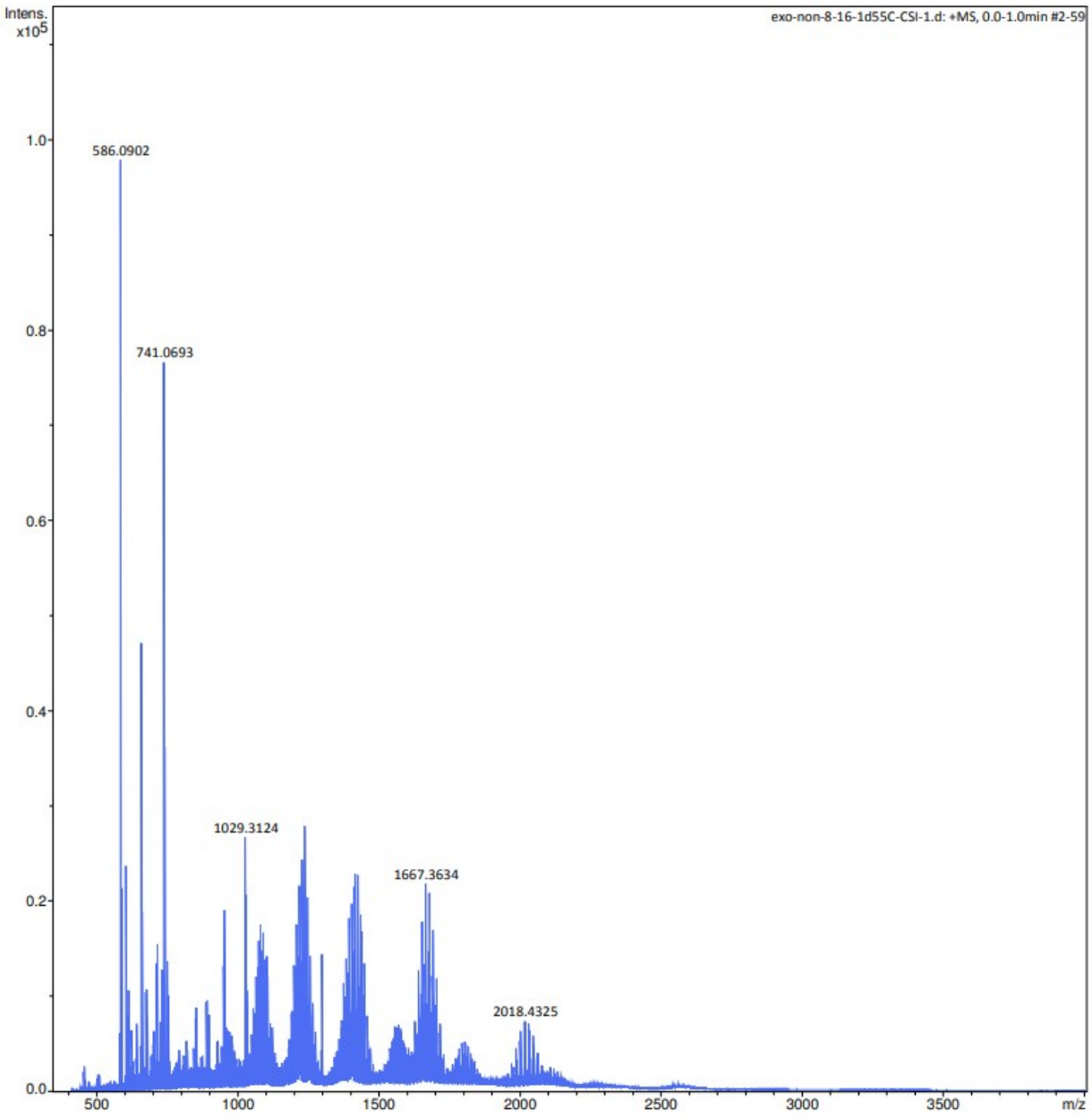


Figure S33. ESI–HRMS spectra of heteroleptic assemblies derived from ***^{L}Ex*** *and* ***^{L}En****, est. 15.6:8.4 ratio showing numerous charged species.*

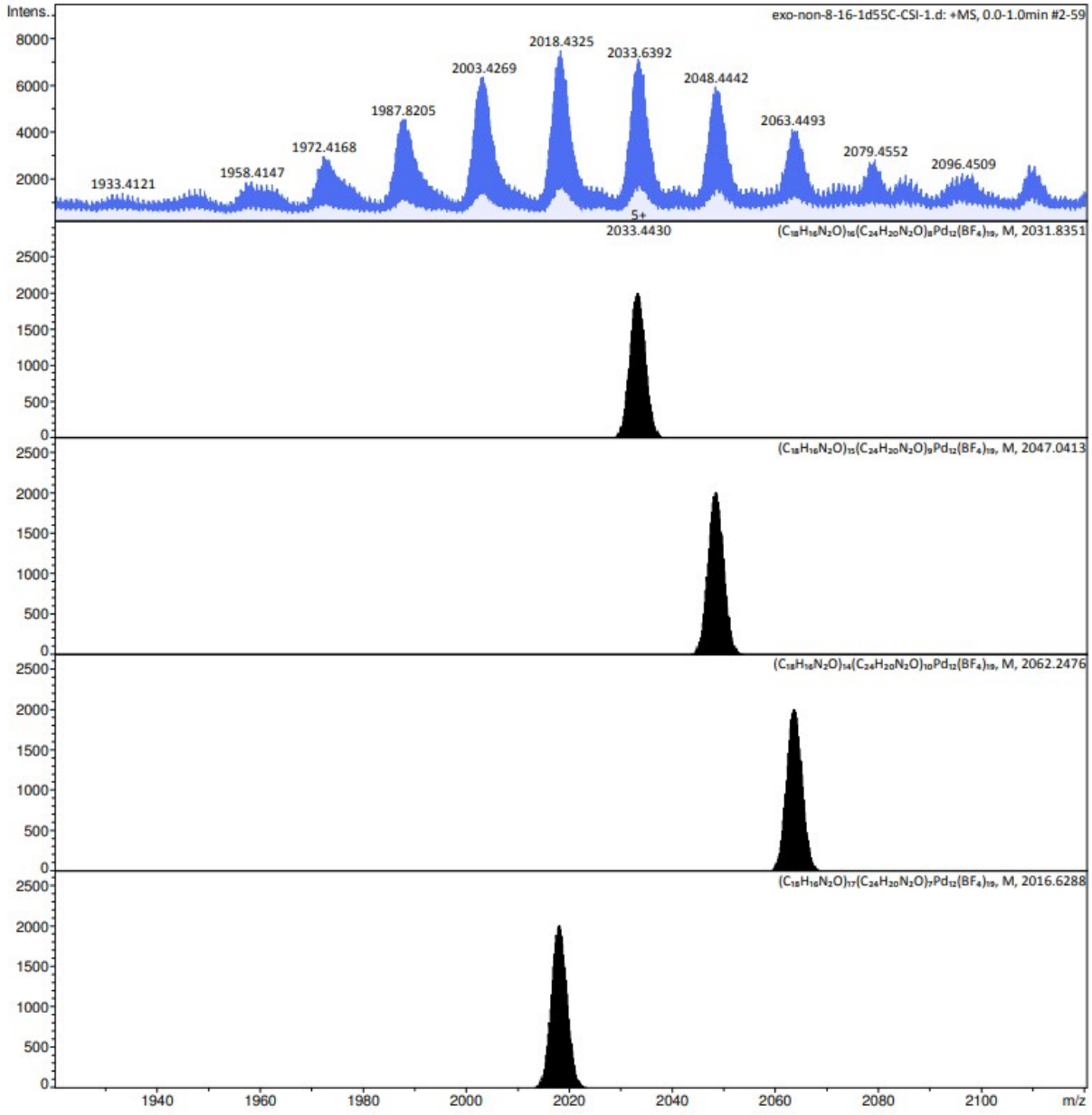


Figure S34. ESI–HRMS spectra of heteroleptic assemblies derived from ^{L}***Ex*** *and* ^{L}***En****, est. 15.6:8.4 ratio, expanded to show peaks of various linker compositions (blue). Mass simulations of* $Pd12^{L}\mathbf{Ex}(x)^{L}\mathbf{En}(24\text{-}x)(BF4)19^{5+}$ *where* $x \in [14\text{-}17]$ *(black).*

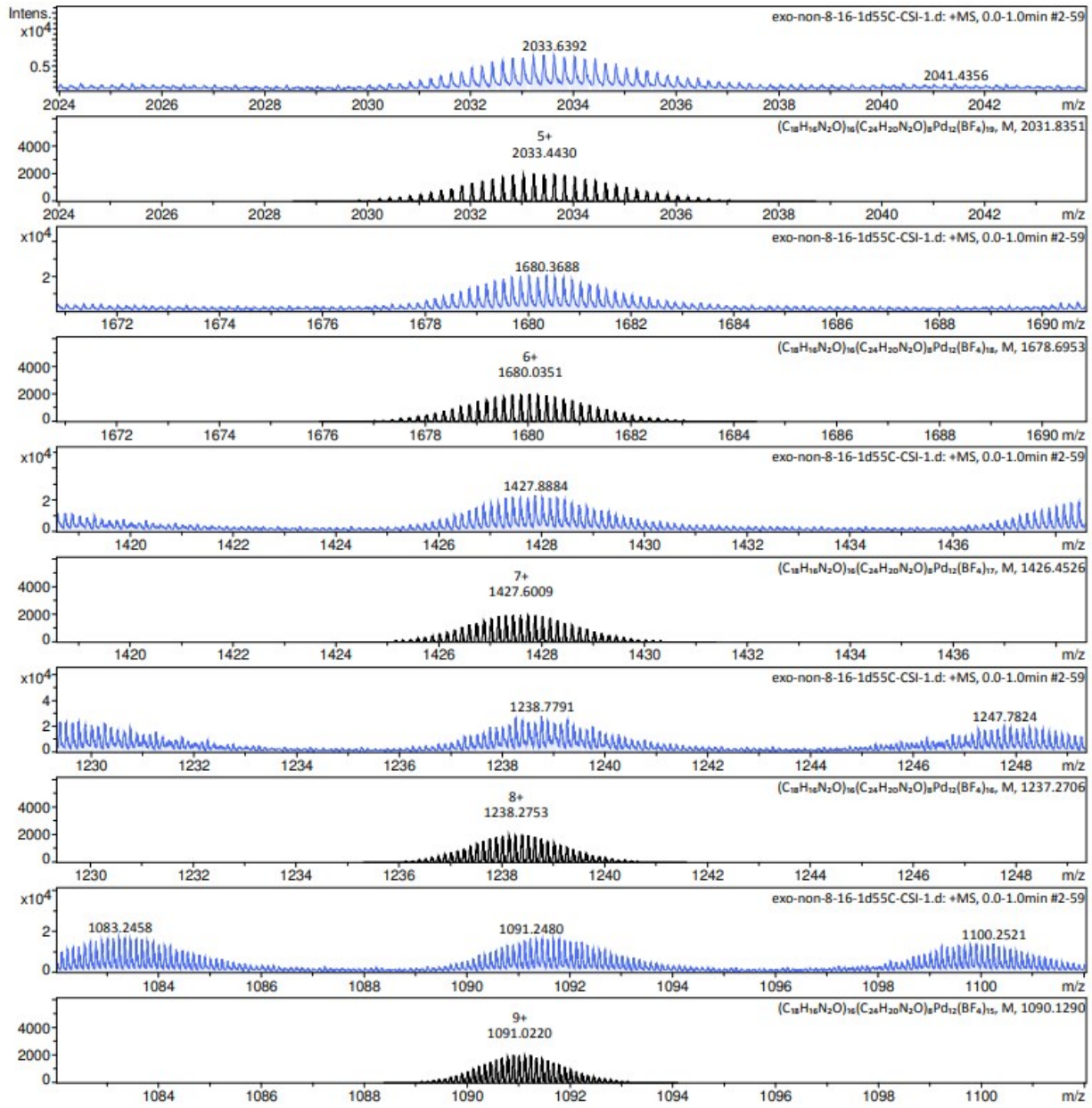


Figure S35. ESI–HRMS spectra of heteroleptic assemblies derived from ***^{L}Ex*** *and* ***^{L}En****, est. 15.6:8.4 ratio, expanded to show multiple charged peaks(blue). Simulations of the isotope distribution patterns for* Pd_{12}***^{L}Ex***$_{(16)}$***^{L}En***$_{(8)}$ $(BF_4)_{(24-x)}^{x+}$ *where* $x \in [5\text{-}9]$.

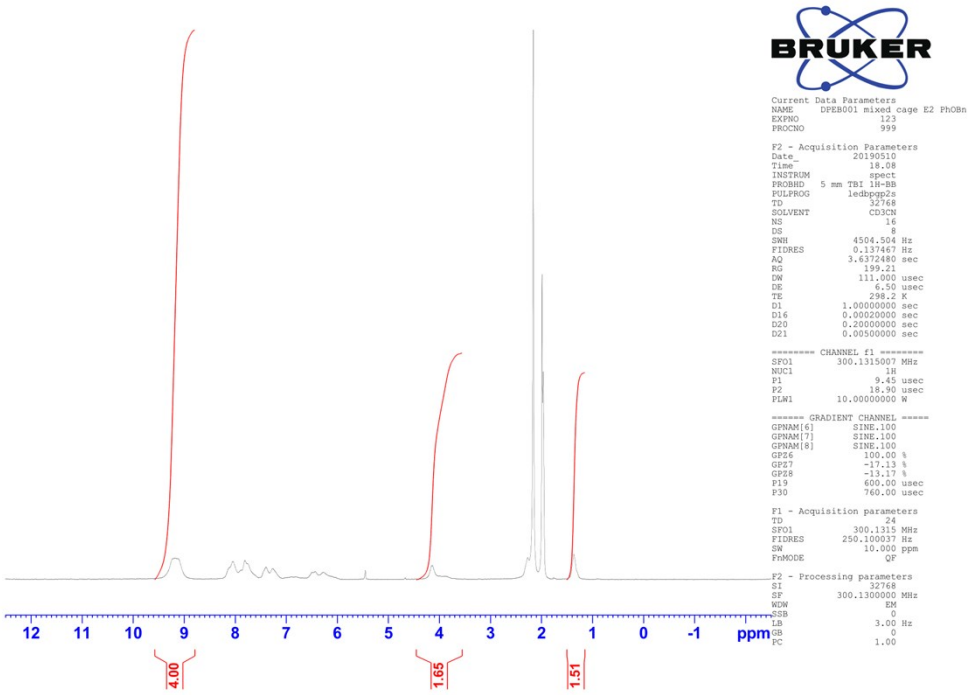
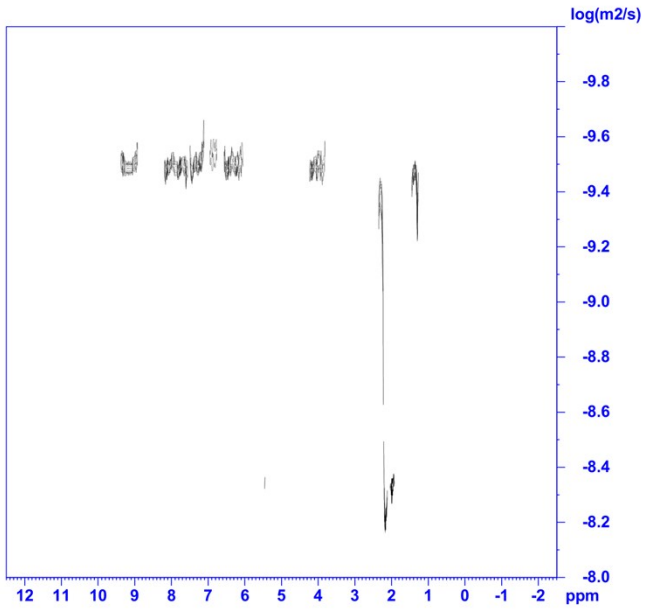


Figure S36. ^{1}H NMR spectrum obtained for homoleptic Pd assemblies of ^{L}***Ex*** *and* ^{L}***En****, est. 12.1:11.9 ratio. (CD_3CN, 300 MHz)*



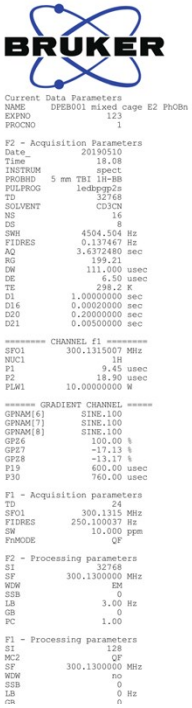


Figure S37. DOSY NMR obtained for Pd assemblies of ^{L}***Ex*** *and* ^{L}***En****, est. 12.1:11.9 ratio. (CD_3CN, 300 MHz)*

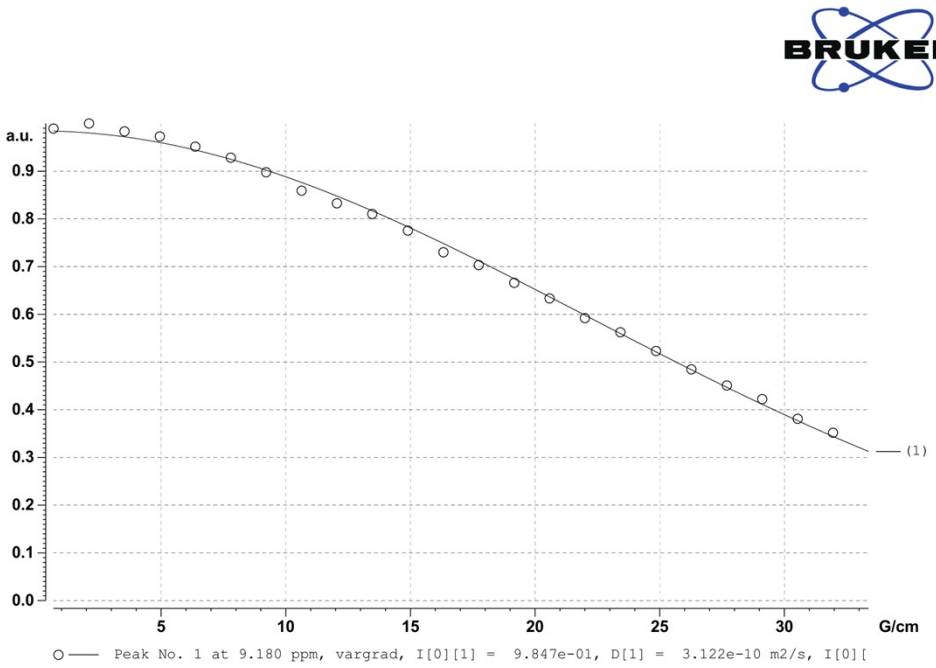


Figure S38. α-Pyridyl peak integrals showing the diffusion decay observed during DOSY analysis of Pd assemblies of ^{L}***Ex*** *and* ^{L}***En****, est. 12.1:11.9 ratio of* ^{L}***En****. Exponential function fit computed using the topspin relaxation module.*

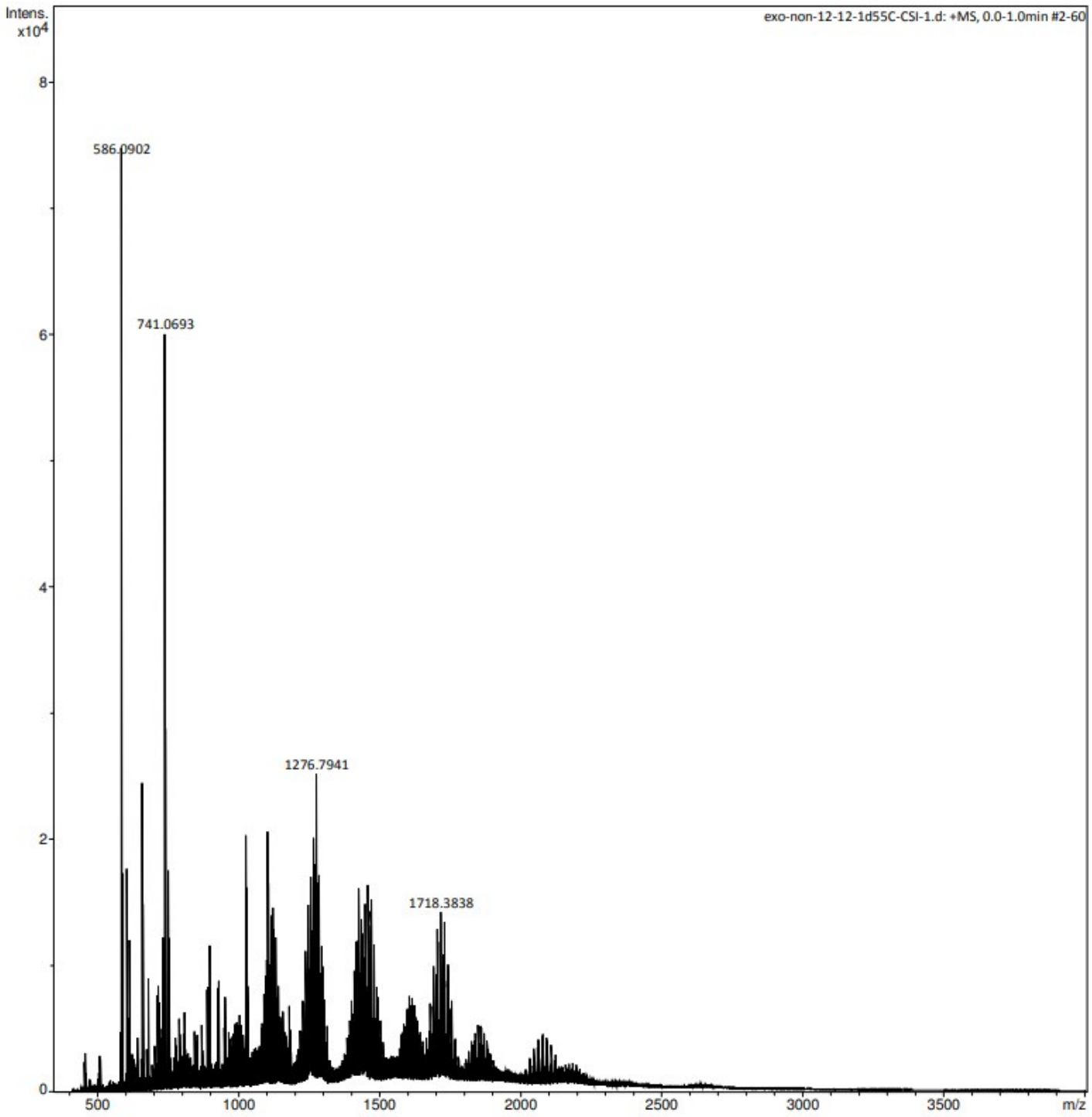


Figure S39. ESI–HRMS spectra of heteroleptic assemblies derived from ^{L}***Ex*** *and* ^{L}***En****, est. 12.1:11.9 ratio showing numerous charged species.*

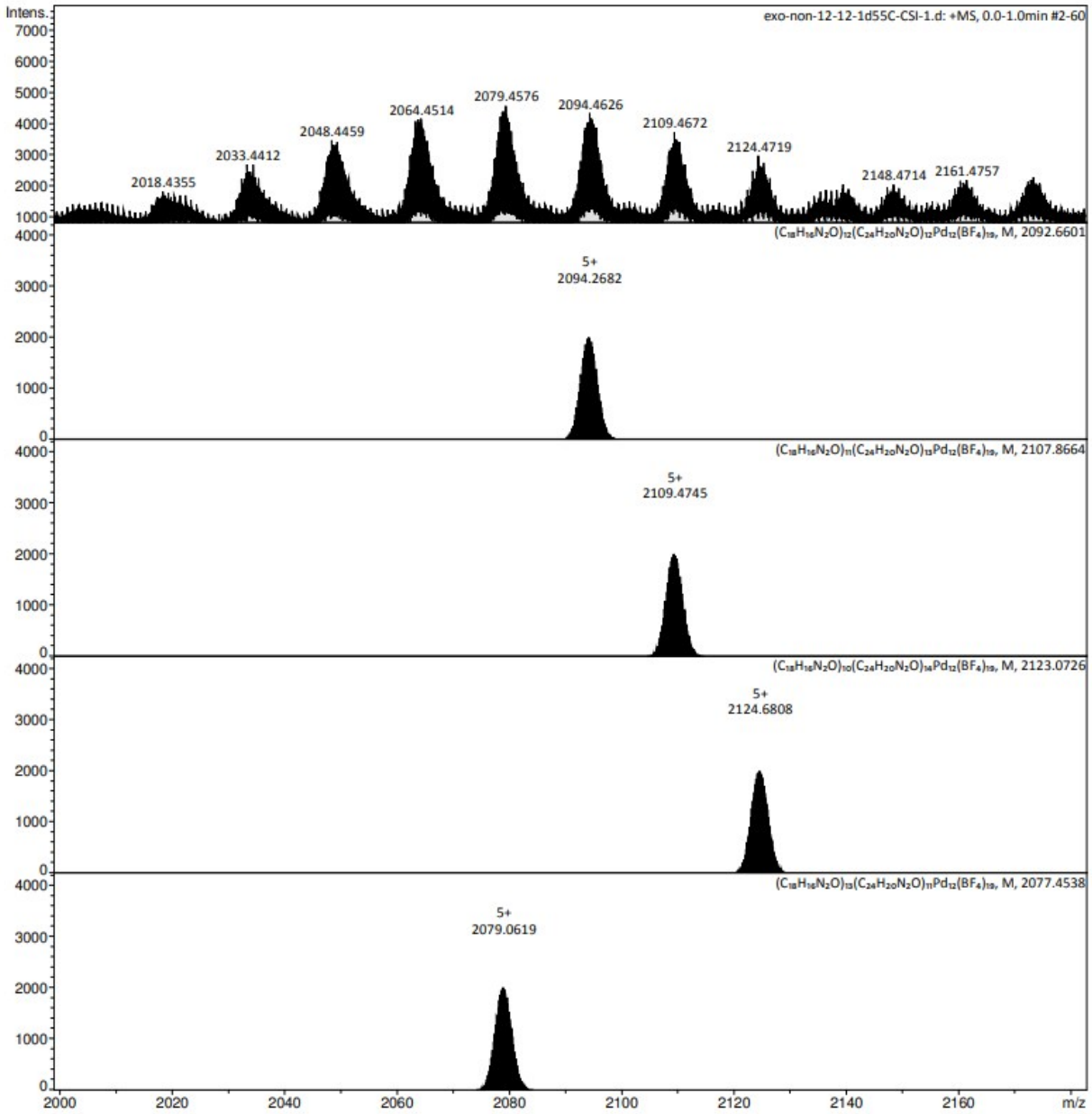


Figure S40. ESI–HRMS spectra of heteroleptic assemblies derived from ***L*Ex*** *and* ***L*En****, est.12.1:11.9 ratio, expanded to show peaks of various linker compositions (top, dark grey). Mass simulations of* Pd_{12}***L*Ex***$_{(x)}$***L*En***$_{(24-x)}(BF_4)_{19}^{5+}$ *where* $x \in [10\text{-}13]$ *(below, black).*

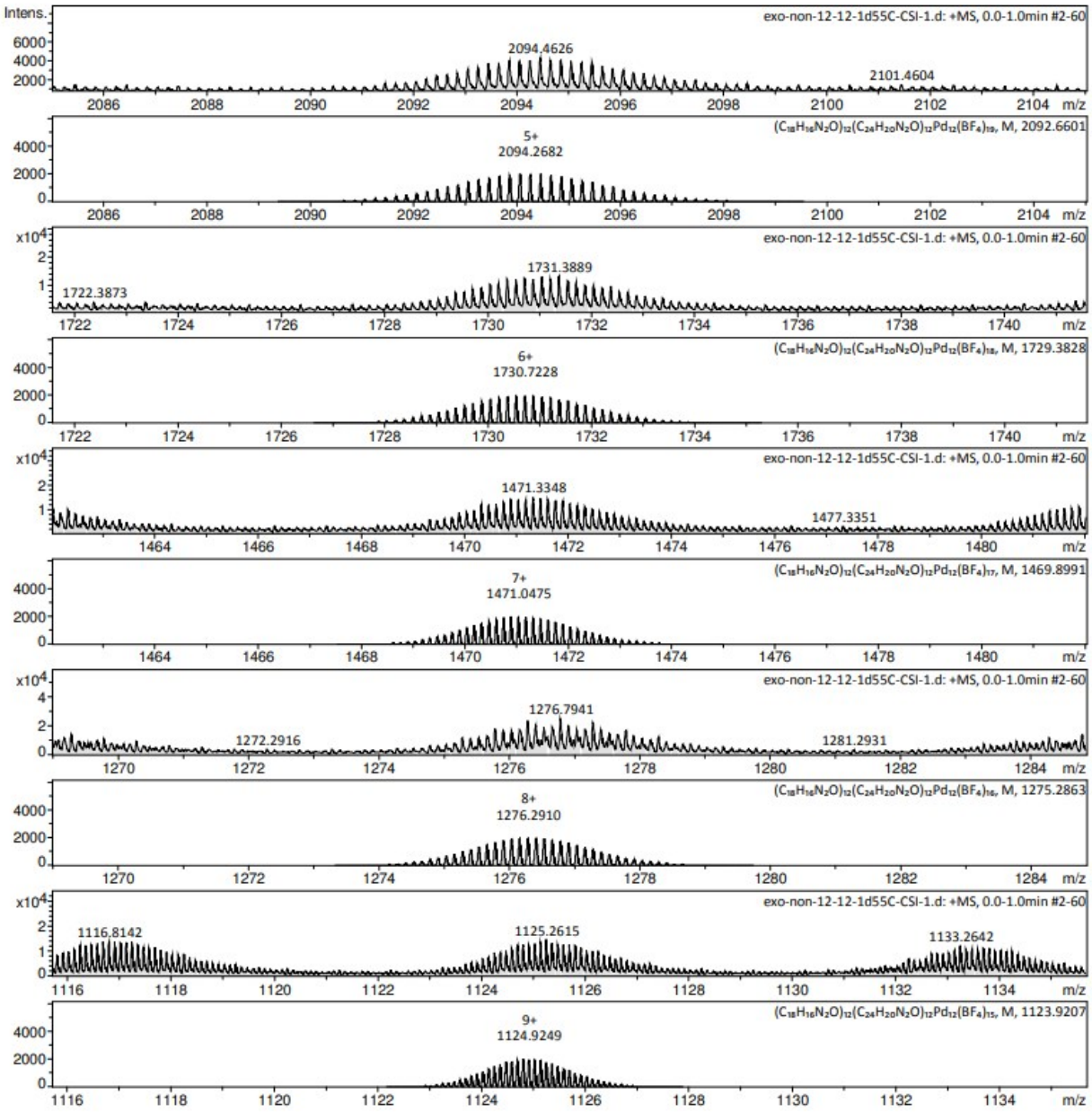
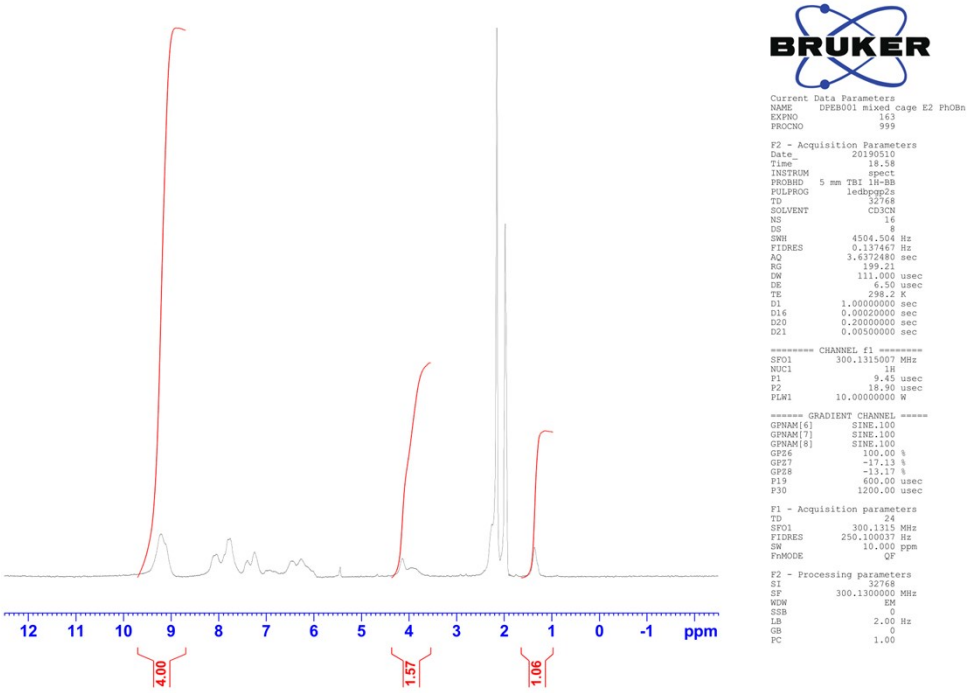


Figure S41. ESI–HRMS spectra of heteroleptic assemblies derived from ^{L}***Ex*** *and* ^{L}***En****, est. 12.1:11.9 ratio, expanded to show multiple charged peaks (alternating first, dark grey). Simulations of the isotope distribution patterns for* $Pd_{12}{}^{L}\mathbf{Ex}_{(12)}{}^{L}\mathbf{En}_{(12)}(BF_4)_{(24-x)}{}^{x+}$ *where* $x \in [5\text{-}9]$ *(alternating second, black).*



Current Data Parameters
NAME DPEB001 mixed cage E2 PhOBn
EXPNO 163
PROCNO 999

F2 - Acquisition Parameters
Date_ 20190510
Time 18.58
INSTRUM spect
PROBHD 5 mm TBI 1H-BB
PULPROG ledbpgp2s
TD 32768
SOLVENT CD3CN
NS 16
DS 8
SWH 4504.504 Hz
FIDRES 0.137467 Hz
AQ 3.6372480 sec
RG 199.21
DW 111.000 usec
DE 6.50 usec
TE 298.2 K
D1 1.00000000 sec
D16 0.00020000 sec
D20 0.20000000 sec
D21 0.00500000 sec

======== CHANNEL f1 ========
SFO1 300.1315007 MHz
NUC1 1H
P1 9.45 usec
P2 18.90 usec
PLW1 10.00000000 W

====== GRADIENT CHANNEL =====
GPNAM[6] SINE.100
GPNAM[7] SINE.100
GPNAM[8] SINE.100
GPZ6 100.00 %
GPZ7 -17.13 %
GPZ8 -13.17 %
P19 600.00 usec
P30 1200.00 usec

F1 - Acquisition parameters
TD 24
SFO1 300.1315 MHz
FIDRES 250.100037 Hz
SW 10.000 ppm
FnMODE QF

F2 - Processing parameters
SI 32768
SF 300.1300000 MHz
WDW EM
SSB 0
LB 2.00 Hz
GB 0
PC 1.00

Figure S42. [1]H NMR spectrum obtained for homoleptic Pd assemblies of ^{L}***Ex*** *and* ^{L}***En****, est. 8.5:15.5 ratio. (*CD_3CN*, 300 MHz)*

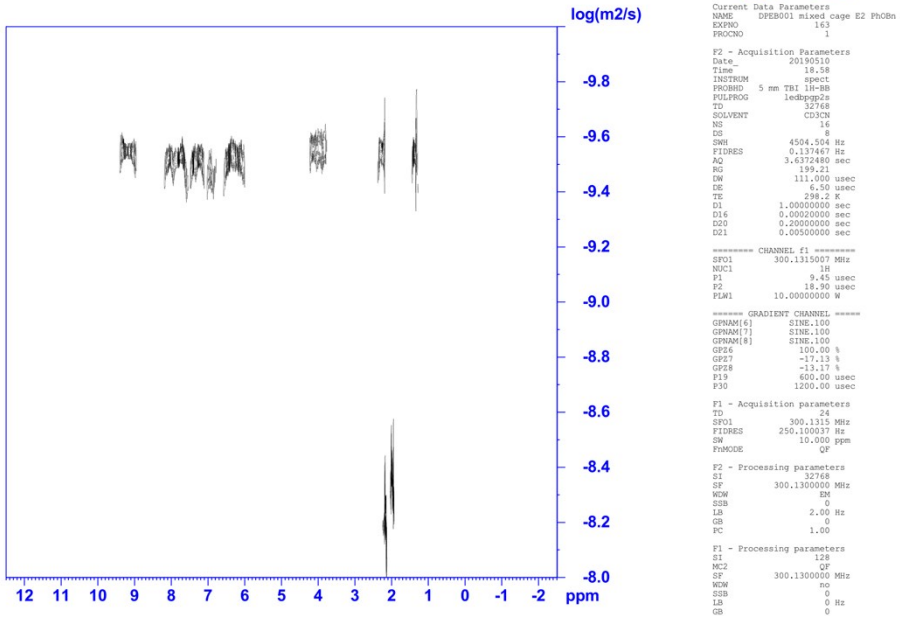


Figure S43. DOSY NMR obtained for Pd assemblies of ^{L}***Ex*** *and* ^{L}***En****, est. 8.5:15.5 ratio. (CD_3CN, 300 MHz)*

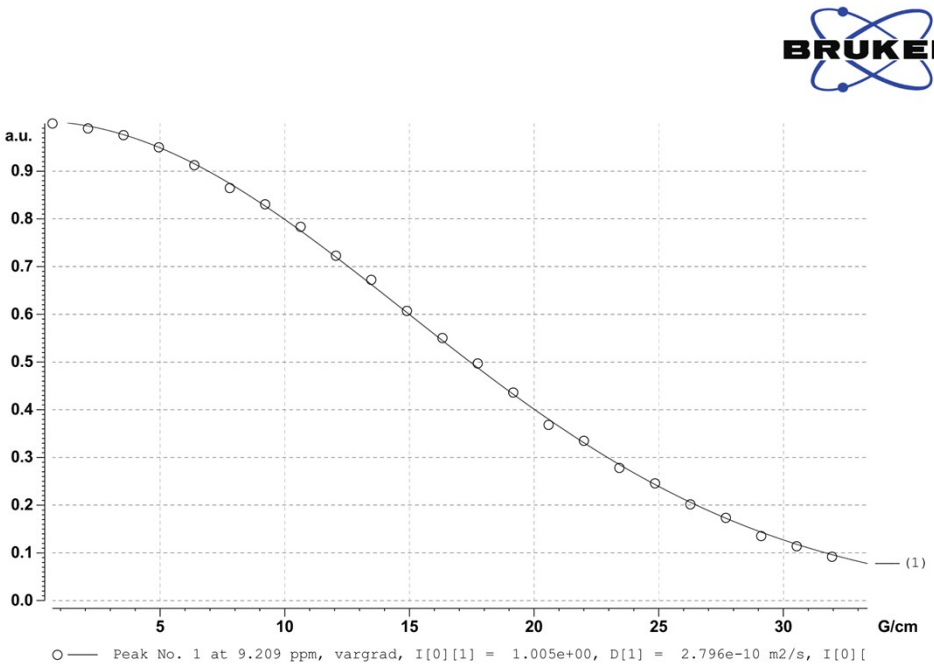


Figure S44. α-Pyridyl peak integrals showing the diffusion decay observed during DOSY analysis of Pd assemblies of ^{L}***Ex*** *and* ^{L}***En****, est. 8.5:15.5 ratio of* ^{L}***En****. Exponential function fit computed using the Topspin relaxation module.*

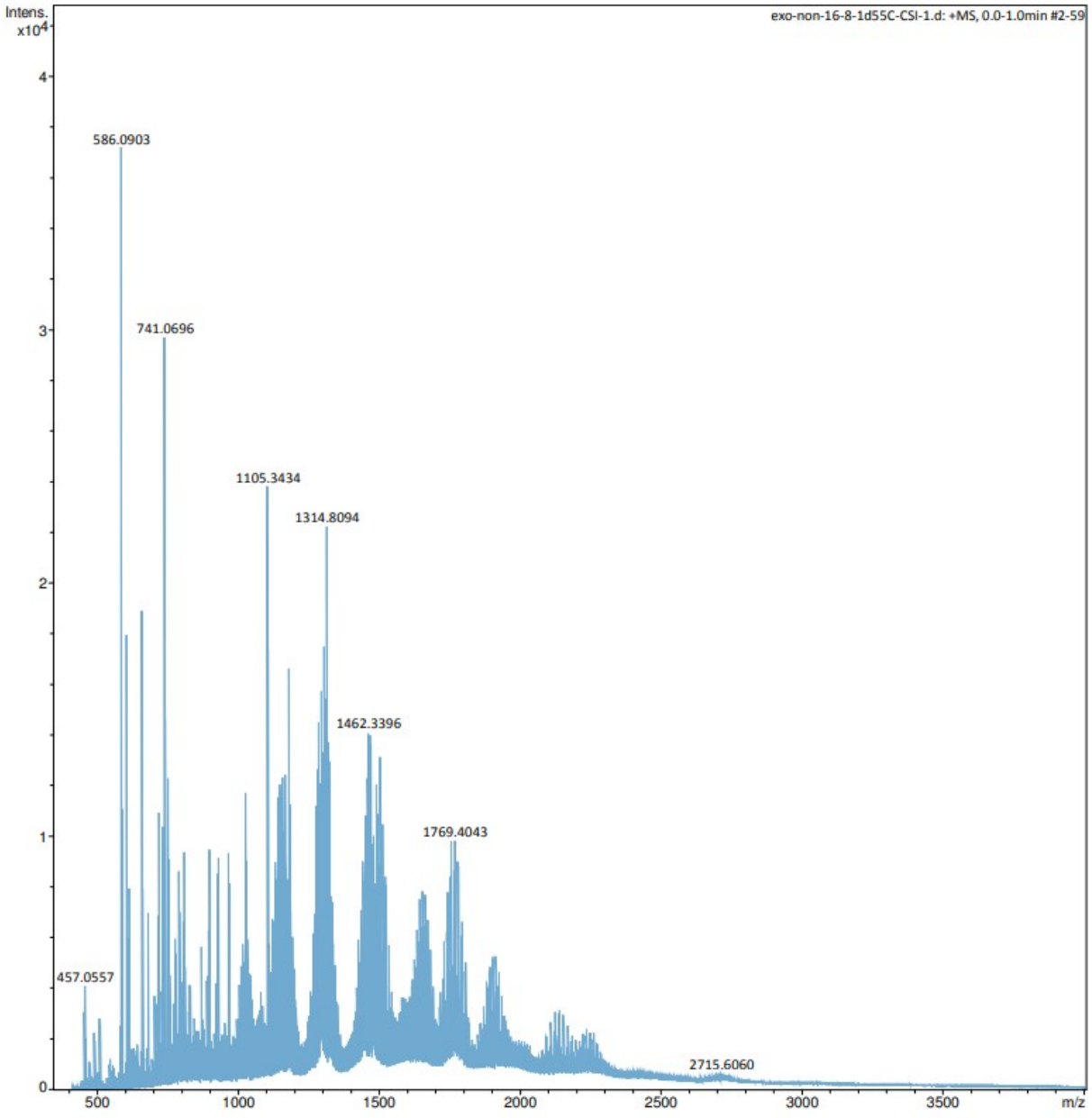


Figure S45. ESI–HRMS spectra of heteroleptic assemblies derived from ***^{L}Ex*** *and* ***^{L}En****, est. 8.5:15.5 ratio showing numerous charged species.*

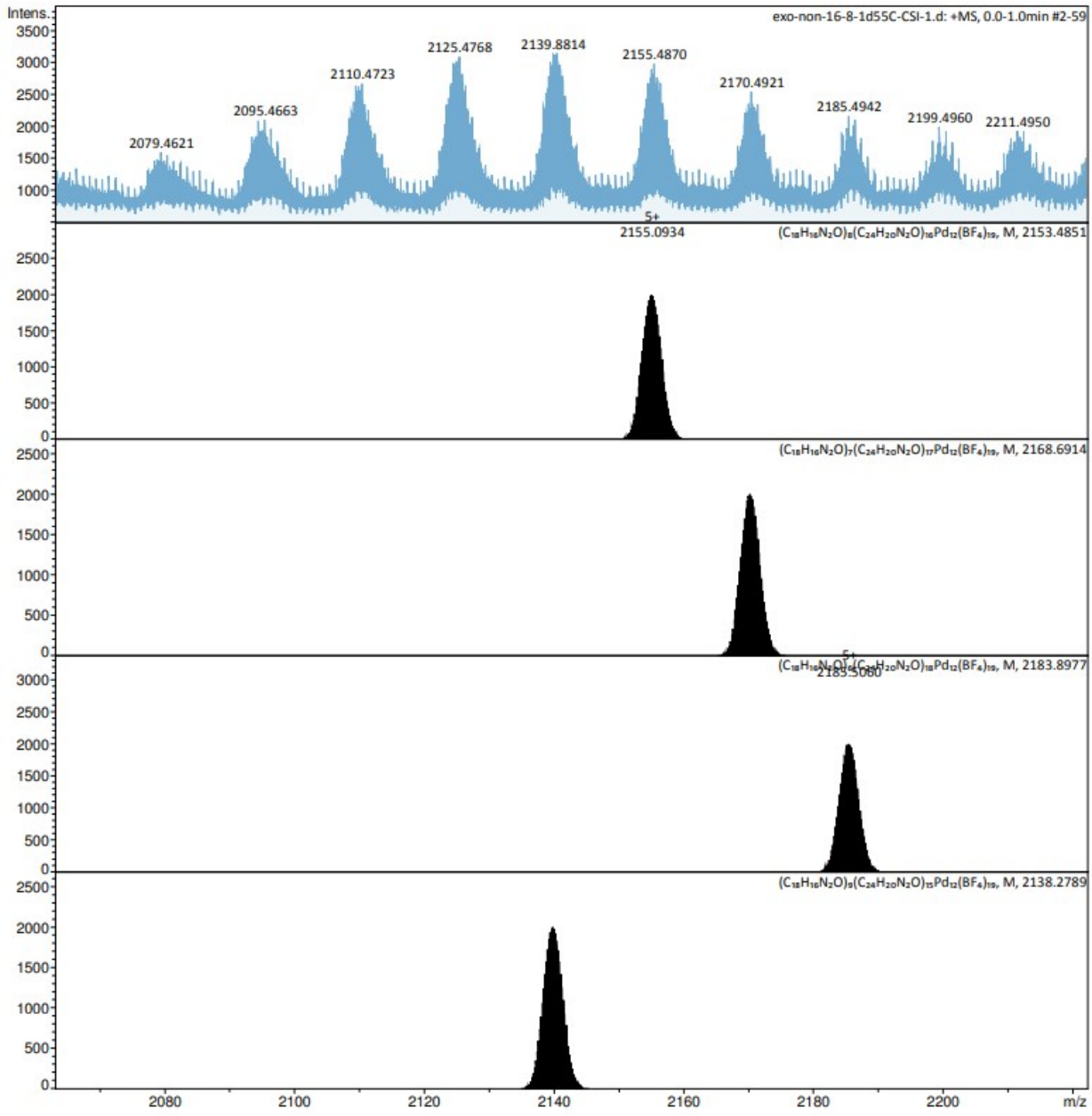


Figure S46. ESI–HRMS spectra of heteroleptic assemblies derived from ^{L}***Ex*** *and* ^{L}***En****, est. 8.5:15.5 ratio, expanded to show peaks of various linker compositions (blue). Mass simulations of* $Pd_{12}{}^{L}\mathbf{Ex}_{(x)}{}^{L}\mathbf{En}_{(24-x)}(BF_4)_{19}{}^{5+}$ *where* $x \in [5\text{-}8]$ *(black).*

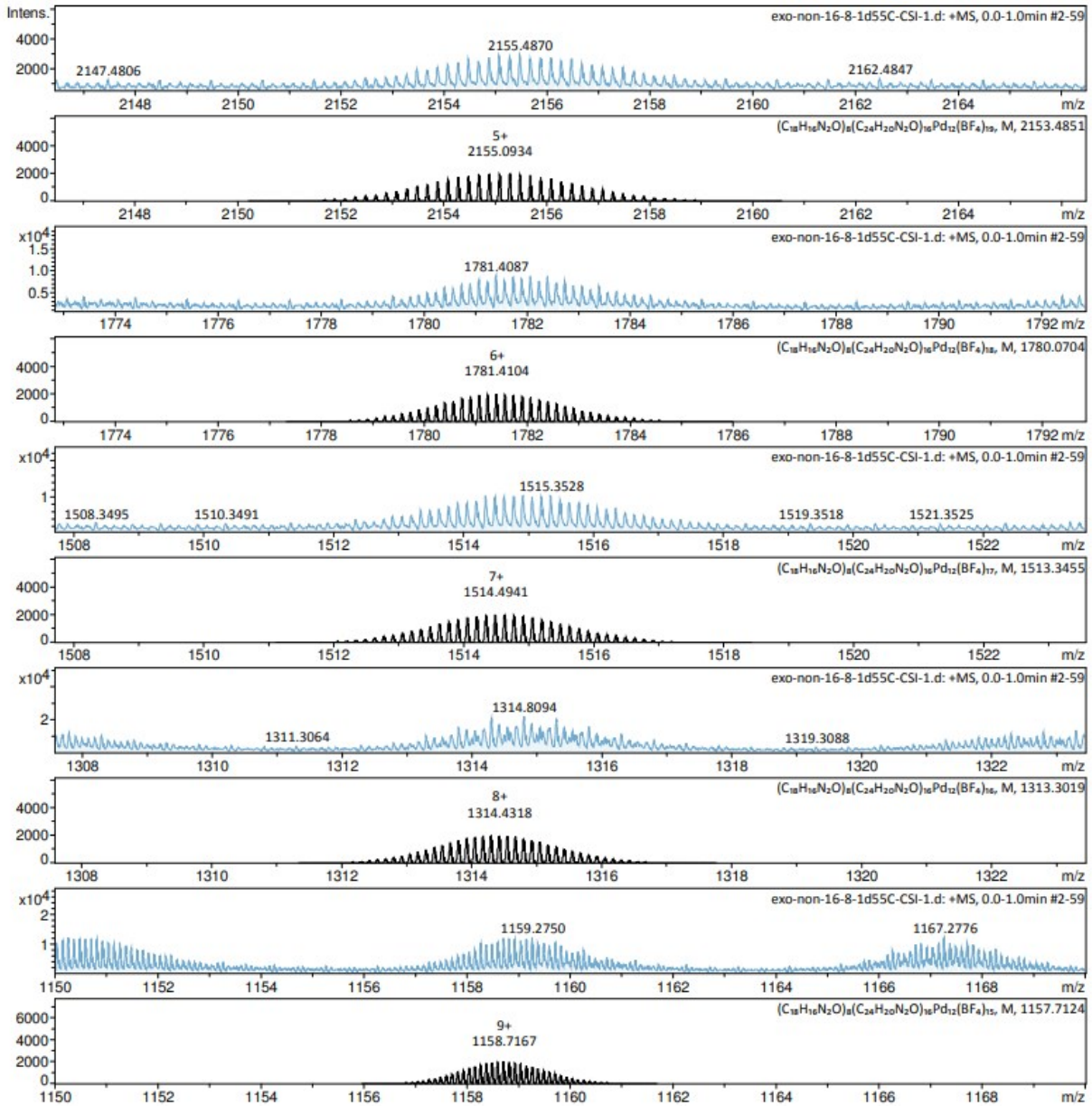
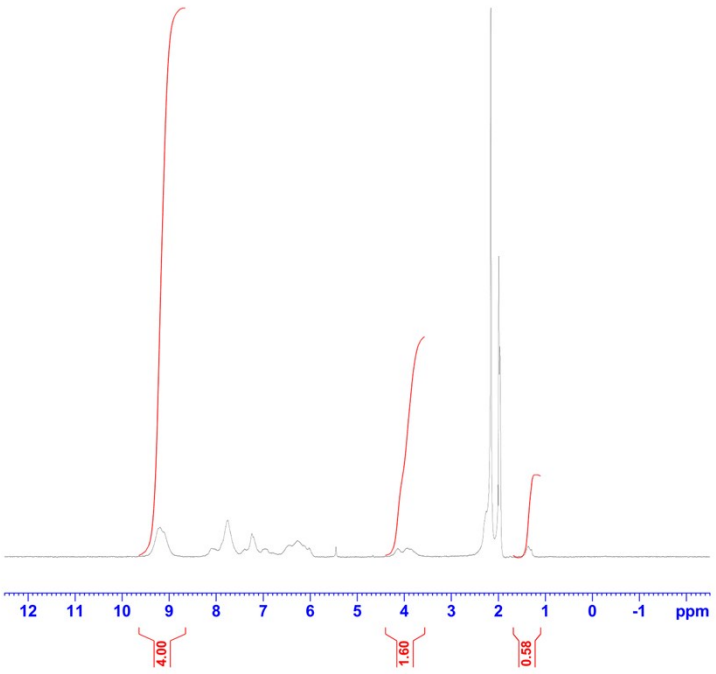
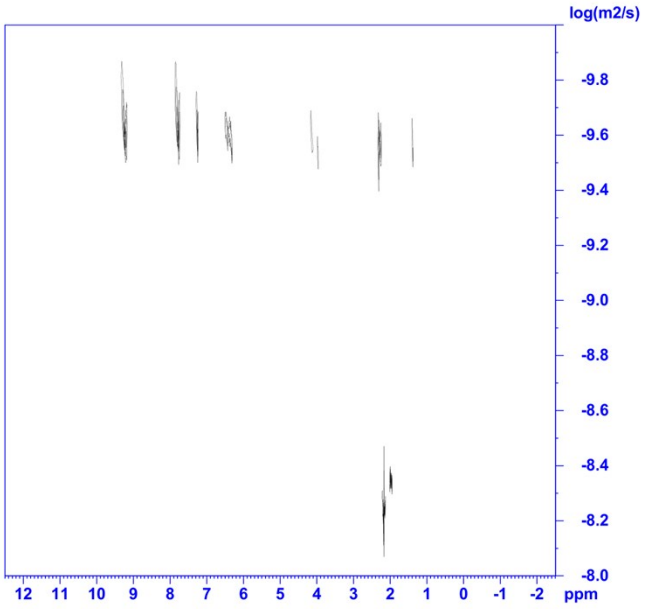


Figure S47. ESI–HRMS spectra of heteroleptic assemblies derived from ^{L}***Ex*** *and* ^{L}***En****, est. 8.5:15.5 ratio, expanded to show multiple charged peaks (blue). Simulations of the isotope distribution patterns for* $Pd_{12}{}^{L}\mathbf{Ex}_{(8)}{}^{L}\mathbf{En}_{(16)}(BF_4)_{(24-x)}{}^{x+}$ *where* $x \in [5\text{-}9]$ *(black).*




```
Current Data Parameters
NAME      DPEB001 mixed cage E2 PhOBn
EXPNO                 203
PROCNO                999

F2 - Acquisition Parameters
Date_              20190510
Time_                 19.39
INSTRUM               spect
PROBHD    5 mm TBI 1H-BB
PULPROG          ledbpgp2s
TD                    32768
SOLVENT               CD3CN
NS                       16
DS                        8
SWH                4504.504 Hz
FIDRES            0.137467 Hz
AQ                3.6372480 sec
RG                   199.21
DW                  111.000 usec
DE                    6.50 usec
TE                  298.2 K
D1                1.00000000 sec
D16               0.00020000 sec
D20               0.20000000 sec
D21               0.00500000 sec

======== CHANNEL f1 ========
SFO1           300.1315007 MHz
NUC1                     1H
P1                    9.45 usec
P2                   18.90 usec
PLW1           10.00000000 W

====== GRADIENT CHANNEL =====
GPNAM[6]          SINE.100
GPNAM[7]          SINE.100
GPNAM[8]          SINE.100
GPZ6                100.00 %
GPZ7                -17.13 %
GPZ8                -13.17 %
P19                 600.00 usec
P30                1200.00 usec

F1 - Acquisition parameters
TD                       24
SFO1              300.1315 MHz
FIDRES           250.100037 Hz
SW                    10.000 ppm
FnMODE                   QF

F2 - Processing parameters
SI                    32768
SF             300.1300000 MHz
WDW                      EM
SSB                       0
LB                     2.00 Hz
GB                        0
PC                     1.00
```

Figure S48. [1]H NMR spectrum obtained for homoleptic Pd assemblies of ***[L]Ex*** *and* ***[L]En****, est. 4.6:19.4 ratio. (CD_3CN, 300 MHz)*



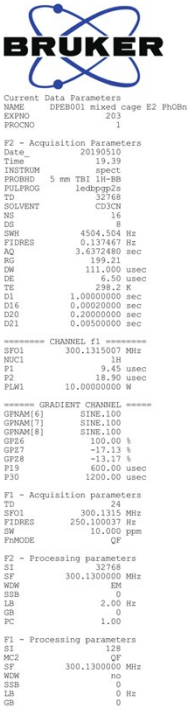


Figure S49. DOSY NMR obtained for Pd assemblies of ***[L]Ex*** *and* ***[L]En****, est. 4.6:19.4 ratio. (CD_3CN, 300 MHz)*

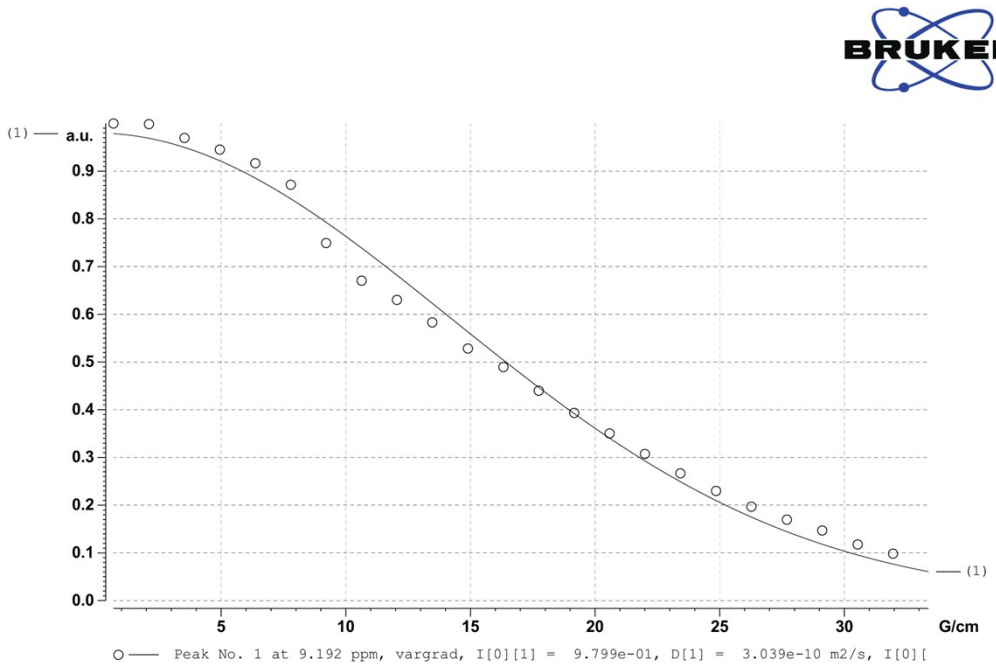


Figure S50. α-Pyridyl peak integrals showing the diffusion decay observed during DOSY analysis of Pd assemblies of ***ᴸEx*** *and* ***ᴸEn****, est. 4.6:19.4 ratio of* ***ᴸEn****. Exponential function fit computed using the Topspin relaxation module.*

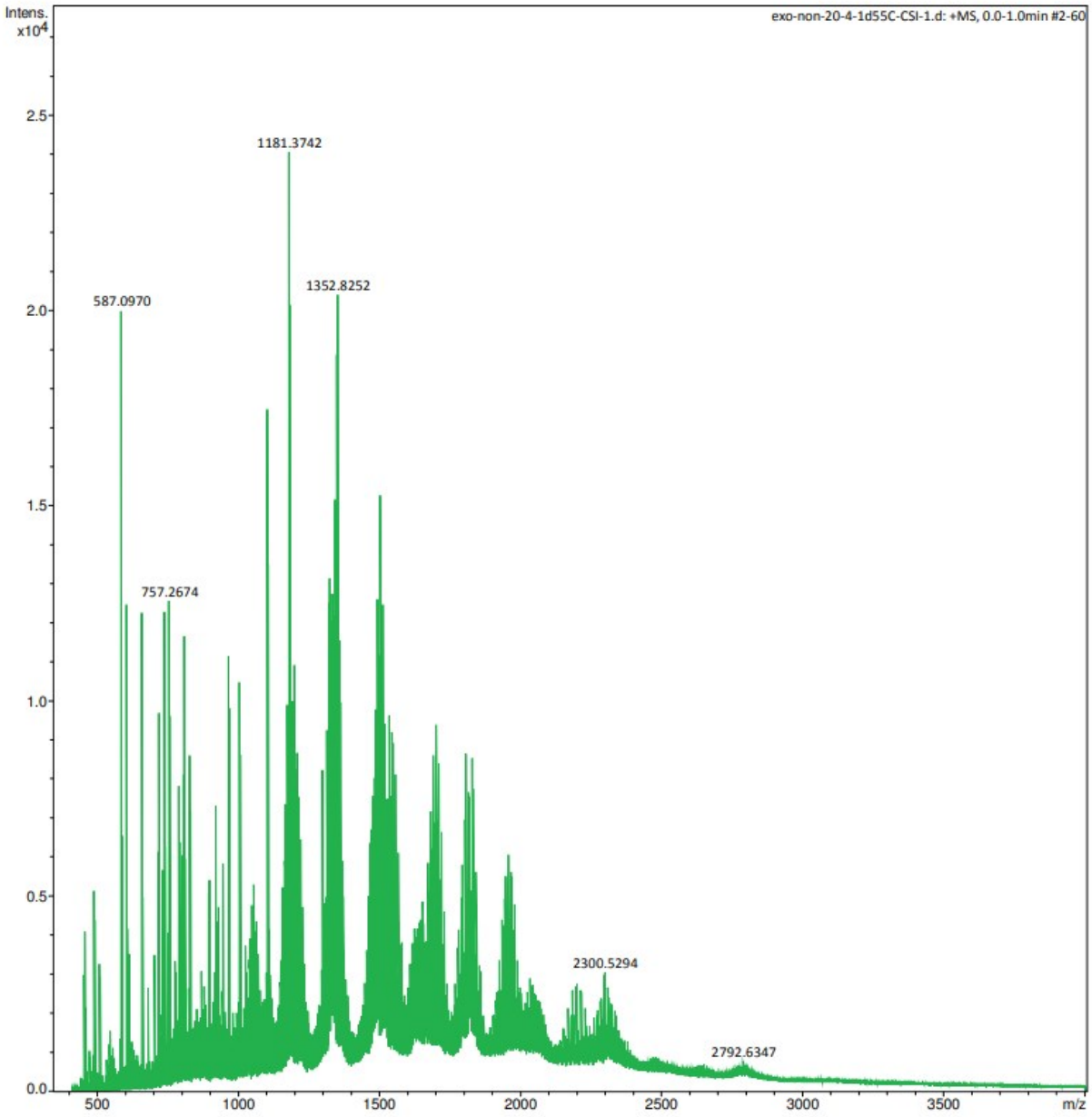


Figure S51. ESI–HRMS spectra of heteroleptic assemblies derived from ^{L}***Ex*** *and* ^{L}***En****, est. 4.6:19.4 ratio showing numerous charged species.*

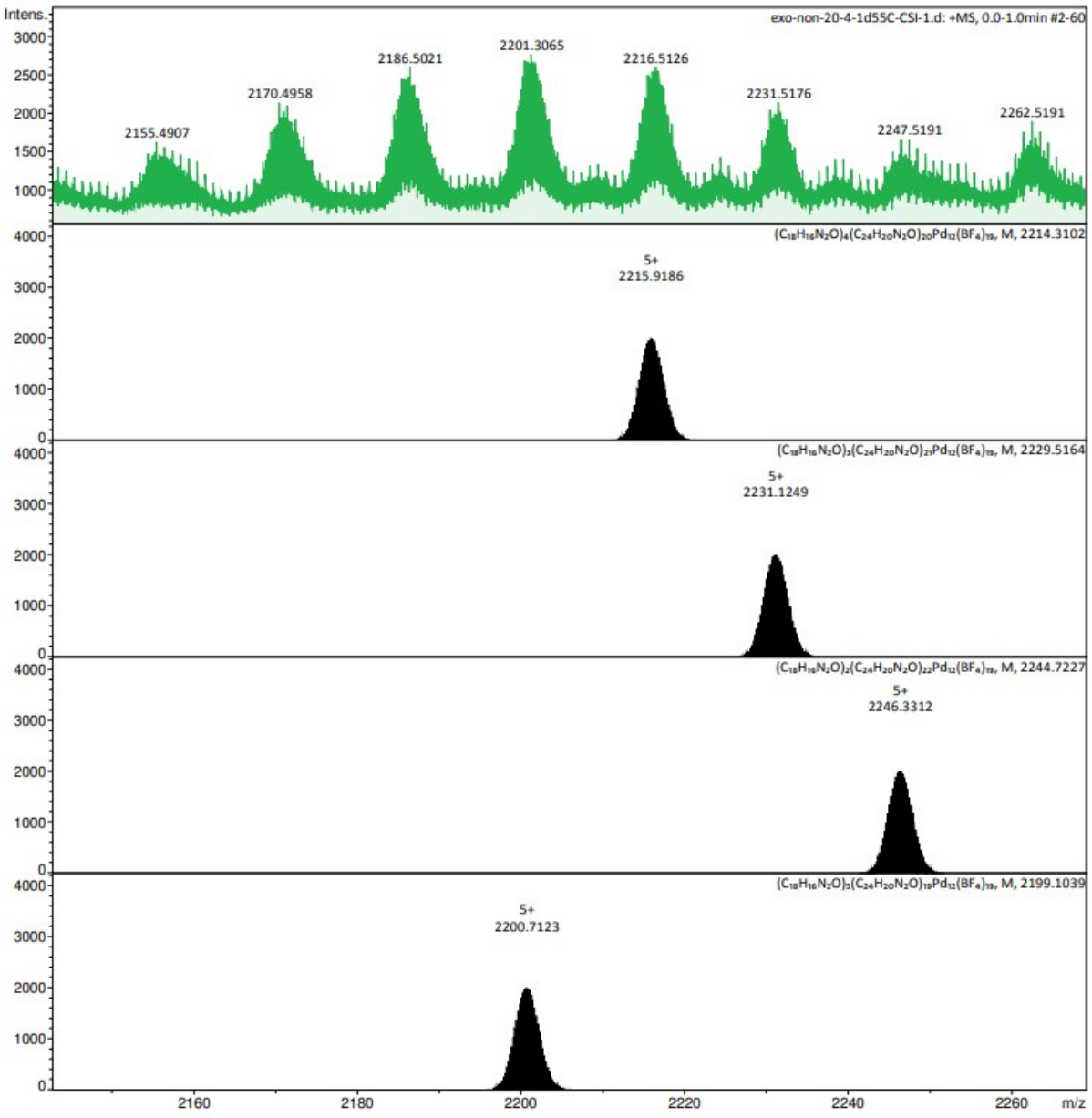


Figure S52. ESI–HRMS spectra of heteroleptic assemblies derived from ***L*Ex*** *and* ***L*En****, est. 4.6:19.4 ratio, expanded to show peaks of various linker compositions (green). Mass simulations of* Pd_{12}***L*Ex***$_{(x)}$***L*En***$_{(24-x)}(BF_4)_{19}^{5+}$ *where* $x \in [2\text{-}5]$ *(black).*

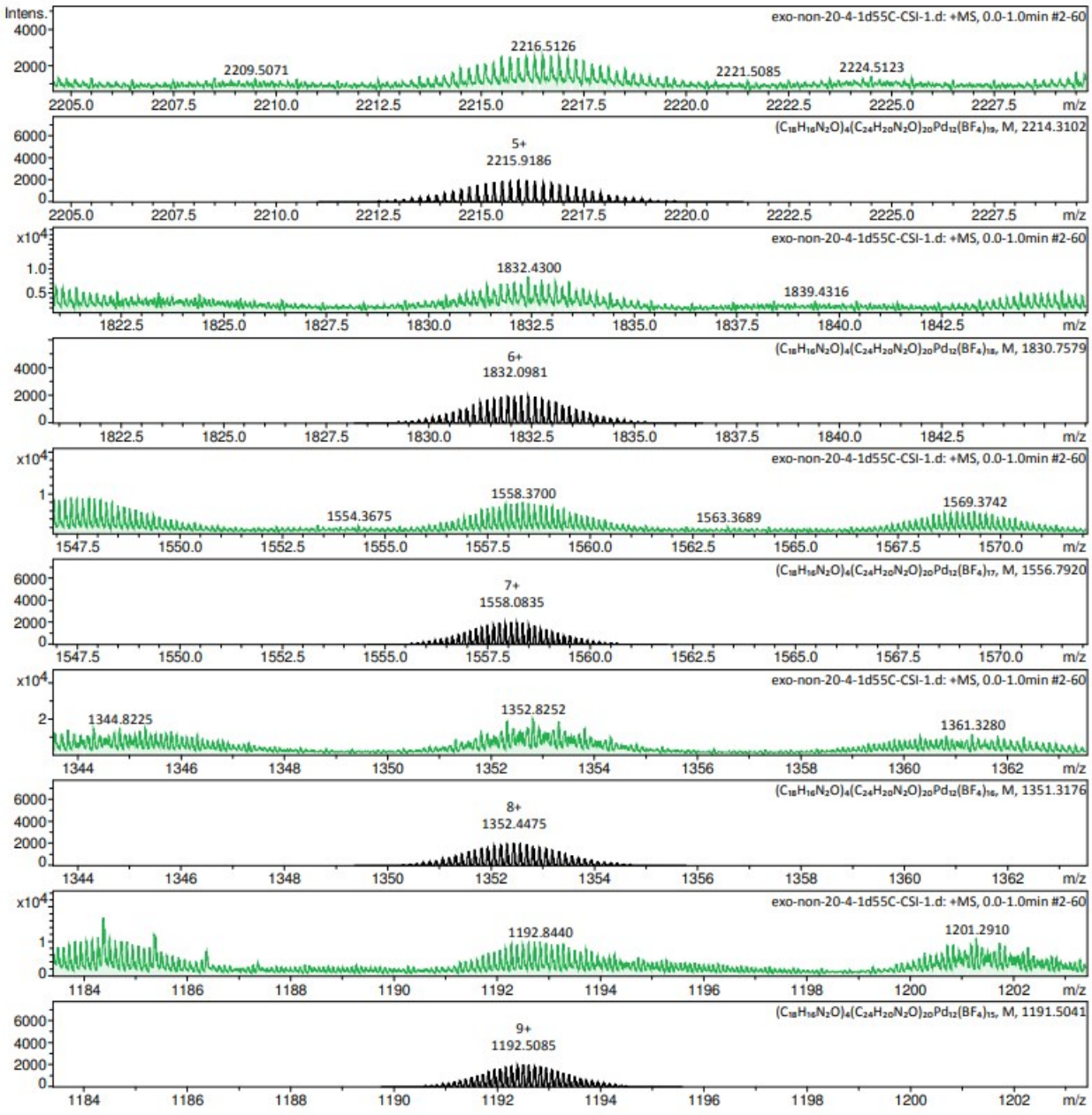


Figure S53. ESI–HRMS spectra of heteroleptic assemblies derived from ***L*Ex** *and* ***L*En**, *est. 4.6:19.4 ratio, expanded to show multiple charged peaks (green). Simulations of the isotope distribution patterns for* Pd_{12}***L*Ex**$_{(4)}$***L*En**$_{(20)}$ $(BF_4)_{(24-x)}^{x+}$ *where* $x \in [5\text{-}9]$ *(black).*

I: Topological predictions using the Generalized Amber Force Field

In the development of models featuring new linkers, it is worthwhile to consider the use of a general approach to describe the organic linkers. To this end we conducted a topological prediction using the Generalized Amber Force Field (GAFF)[5] to describe the organic linker alongside the palladium-pyridyl parameters produced in this study (see Section A). The resulting predictions using this generalized method had little correlation to experimental outcomes. For ^{L}**Fu** and ^{L}**Th**, the predicted topologies were significantly smaller ($Pd_{10}L_{20}$, Pd_8L_{16}). Predictions for ^{L}**Ex** and ^{L}**En** were more accurate, identifying $Pd_{12}L_{24}$ and $Pd_{30}L_{60}$ as the respective major assembly topologies with a number of minor species not experimentally observed. It is clear that GAFF behaves remarkably well for describing poly-aromatic systems, and despite its deficiencies offers an excellent starting point for parameter development.

J: CM reproduction of dihedral scans

Structures from the relaxed dihedral potential energy surfaces, Figure 7, were assessed using our CM parameters as shown below in Figure S55. These potential energy surfaces were determined by single point calculation of the forcefield energy of the DFT relaxed structures.

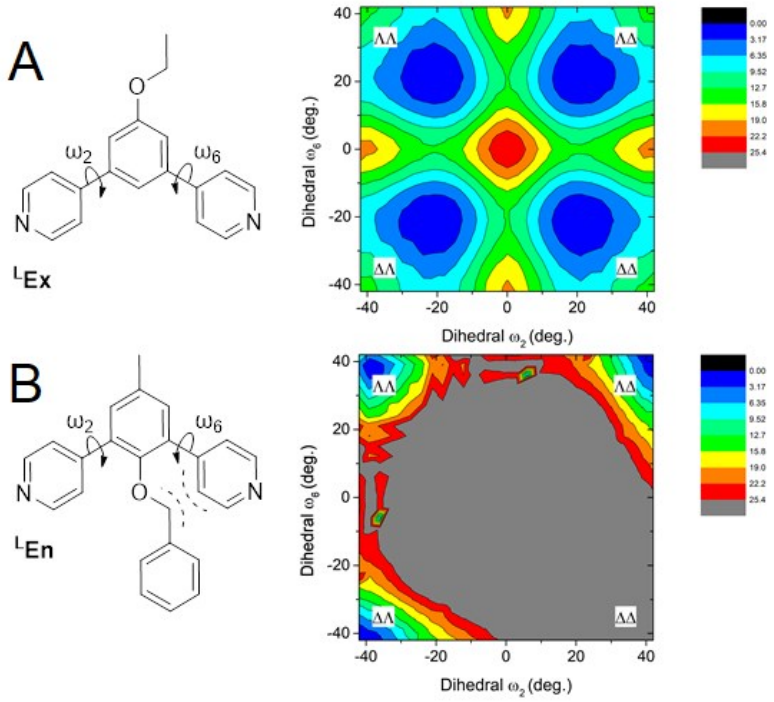


Figure S55. Dihedral angle Potential energy surfaces generated from CM parameters for ***LEx*** *(A), and* ***LEn*** *(B). Left, drawings of each linker showing the dihedral angles and expected steric interactions. Right, CM forcefield energy results based on the rigid single point analysis of structures generated by the relaxed DFT dihedral scan of the two pyridyl-arene dihedrals (ω_2, ω_6) shown in Figure 7. Single point energy computation was completed using our CM parameters, including both the forcefield and charge structures described in Section A for complexes* ***MEx*** *and* ***MEn****. Regions are inaccessible due to their high energy and not detailed further.*

The CM results show similar minima to those found in from the relaxed DFT scans, with a slight difference ~1 kcal mol^{-1} between ΔΛ and ΛΛ barriers. These results reveal a significantly higher dihedral barrier for interconversion between the atropisomers due to the significant steric interactions between the pyridyl and arene groups (Figure S55, left). This may be addressed by performing a relaxed scan, however, this generally highlights the limitations or differences between CM and DFT

approaches for molecular modelling. Ultimately, the high barriers observed make interconversion between the atropoisomers unlikely over the course of structural annealing. Therefore, isomer selection in our models occurs during structural minimization where the coplanar pyridyl groups ($\omega_2 = \omega_6 = 180°$) are rapidly converted to either ΔΛ, ΛΔ, or ΛΛ based on the coordination environment.

M: References

(1) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688. https://doi.org/10.1002/jcc.20290.

(2) Yoneya, M.; Tsuzuki, S.; Yamaguchi, T.; Sato, S.; Fujita, M. Coordination-Directed Self-Assembly of M 12 L 24 Nanocage: Effects of Kinetic Trapping on the Assembly Process. *ACS Nano* **2014**, *8* (2), 1290–1296. https://doi.org/10.1021/nn404595j.

(3) Tachi, Y.; Sato, S.; Yoneya, M.; Fujita, M.; Okamoto, Y. Two Polyhedral Frameworks of an M12L24 Spherical Complex Revealed by Replica-Exchange Molecular Dynamics Simulations. *Chem. Phys. Lett.* **2019**, *714*, 185–189. https://doi.org/10.1016/J.CPLETT.2018.10.059.

(4) Yoneya, M.; Yamaguchi, T.; Sato, S.; Fujita, M. Simulation of Metal–Ligand Self-Assembly into Spherical Complex M 6 L 8. *J. Am. Chem. Soc.* **2012**, *134* (35), 14401–14407. https://doi.org/10.1021/ja303542r.

(5) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *Development and Testing of a General Amber Force Field*; 2004; Vol. 25.

(6) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* **2013**, *9* (6), 2733–2748. https://doi.org/10.1021/ct400146w.

(7) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-XTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15* (3), 1652–1671. https://doi.org/10.1021/acs.jctc.8b01176.

(8) Betz, R. M.; Walker, R. C. Paramfit: Automated Optimization of Force Field Parameters for Molecular Dynamics Simulations. *J. Comput. Chem.* **2015**, *36* (2), 79–87. https://doi.org/10.1002/jcc.23775.

(9) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian16 Revision C.01. 2016.

(10) Holden, D.; Jelfs, K. E.; Cooper, A. I.; Trewin, A.; Willock, D. J. Bespoke Force Field for Simulating the Molecular Dynamics of Porous Organic Cages. *J. Phys. Chem. C* **2012**, *116* (31), 16639–16651. https://doi.org/10.1021/jp305129w.

(11) McLachlan, A. D. Rapid Comparison of Protein Structures. *Acta Crystallogr. Sect. A* **1982**, *38* (6), 871–873. https://doi.org/10.1107/S0567739482001806.

(12) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8* (5), 1542–1555. https://doi.org/10.1021/ct200909j.

(13) Cram, A. J.; Dicker, I. B.; Lauer, M.; Knobler, C. B.; Trueblood, K. N. Host-Guest Complexation. 32. Spherands Composed of Cyclic Urea and Anisyl Units. *J. Am. Chem. Soc.* **1984**, *106* (23), 7150–7167. https://doi.org/10.1021/ja00335a049.