Leveraging Triphenylphosphine-Containing Polymers to Explore Design Principles for Protein-Mimetic Catalysts

Matthew A. Sanders, Supraja S. Chittari, Jack R. Foley, William M. Swofford, Bridgette M. Elder, Abigail S. Knight*

Department of Chemistry, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, United States

Email: aknight@unc.edu

Abstract: Complex interactions between non-coordinating residues are significant yet commonly overlooked components of macromolecular catalyst function. While these interactions have been demonstrated to impact binding affinities and catalytic rates in metalloenzymes, the roles of similar structural elements in synthetic polymeric catalysts remain underexplored. Using a model Suzuki-Miyuara cross-coupling reaction, we performed a series of systematic studies to probe the interconnected effects of metal-ligand cross-links, electrostatic interactions, and local rigidity in polymer catalysts. To achieve this, a novel bi-functional triphenylphosphine acrylamide (BisTPPAm) monomer was synthesized and evaluated alongside an analogous monofunctional triphenylphosphine acrylamide (TPPAm). In model copolymer catalysts, increased initial reaction rates were observed for copolymers untethered by Pd complexation (BisTPPAm-containing) as compared to the Pd-cross-linked catalysts (TPPAm-containing). Further, incorporating local rigidity through secondary structure-like and electrostatic interactions revealed nonmonotonic relationships between composition and reaction rate, demonstrating the potential for tunable behavior through secondary sphere interactions. Finally, through rigorous cheminformatics featurization strategies and statistical modeling, we quantitated relationships between chemical descriptors of the substrate and reaction conditions on catalytic performance. Collectively, these results provide insights into relationships between composition, structure, and function of protein-mimetic catalytic copolymers.

Introduction

The hierarchical structure of metalloenzymes allows for precise spacing, density, and thus accessibility of metal centers, which leads to substrate selectivity and efficient reactivity. Beyond the coordinating ligands, non-ligand residues in the secondary coordination sphere provide structural elements that fine tune enzyme function.[1–3] Native proteins, shaped by evolution, possess refined structures optimized for catalysis in ambient aqueous conditions; however, they often unfold when exposed to unnatural environments common to industrial processes.[4] Further, it can be challenging to rationally tune the selectivity and reactivity to target non-native substrates. Conversely, synthetic macromolecules retain function in broader environmental conditions, are capable of protein-like folding, and have easily tunable compositions, enabling the development of structure-function relationships for targeted catalytic outcomes.[5,6] Metal-mediated reactions have been performed with a range of synthetic macromolecular catalysts, such as metal organic frameworks and heterogeneous cross-linked polymer catalysts.[7–10] While these methods offer effective catalysis on a variety of substrates, systematically tuning second-sphere interactions is a significant challenge, making the development of design rules that incorporate non-ligand chemical components non-trivial.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

To develop materials with fine-tuned structures analogous to metalloenzymes, investigations into homogenous polymeric catalysts have increased, as these harness solvent-excluded reactive sites improving catalyst efficiency and introducing the potential for selectivity.[11–15] Frequently, these materials utilize folded single-chain nanoparticle assemblies imbued with catalytic moieties.[16–18] Building on these efforts, we recently reported a systematic study of macromolecular catalysts using triphenylphosphine-coordinated palladium and a well-characterized Suzuki-Miyaura cross-coupling reaction as a model system to explore structure-function design principles.[19] While we described key relationships between macromolecule composition and catalytic function, key classes of protein-mimetic secondary-sphere interactions such as electrostatic interactions and local rigidity analogous to secondary structure were unexplored.[20,21] Further, the triphenylphosphine moieties inherently cross-linked the macromolecular catalysts, hindering the independent exploration of metal-ligand and non-covalent cross-links.

To address this gap, we report herein the synthesis of triphenylphosphine-containing polymer catalysts that offer orthogonal control of metal-driven cross-links and protein-mimetic secondary-sphere interactions. To achieve this, we leveraged ligands that generate either metal-cross-linked or untethered polymers, as well as pendent moieties that induce local rigidity and electrostatic interactions. We determined that polymer features primarily drive reaction efficacy in traditional organic solvent mixtures, and that the reaction yield was substrate-driven in polar, environmentally benign solvents. Therefore, we further studied how chemical features of substrates (e.g., hydrophobicity, charge distribution, and steric bulk) impact reaction yields. Through a comprehensive evaluation of molecular featurization strategies (e.g., density functional theory simulations) and statistical models, we identified the relative importance of a variety of chemical features describing target substrates and their reactivity with the synthesized polymer catalysts. The analyses described herein expand the structure-function toolbox for polymer catalyst design, continuing to bridge the gap between the complex structure of metalloenzymes and the versatility of small molecule catalysts.

Results and Discussion

Bi-functional triphenylphosphine acrylamide monomer (BisTPPAm) synthesis and copolymerization

Triphenylphosphine is a versatile ligand in transition metal catalysis, and thus has been leveraged previously in polymer catalysts through direct modification of poly(styrene) and solution-phase systems that use post-polymerization modifications.[17,22,23] We recently developed a triphenylphosphine acrylamide (TPPAm) monomer that can be copolymerized with other compatible monomers and can perform Suzuki-Miyaura cross-couplings when complexed to palladium.[19] These materials contain metal-ligand cross-links, which invariably impact the conformation of the polymer chains in solution (Figure 1a).[11]

To target polymer catalysts where metal-ligand interactions could form without intrachain cross-links, we synthesized a bi-functional triphenylphosphine acrylamide (BisTPPAm; Figure 1b). Mirroring the synthesis of TPPAm (Figures S01-S08), an acyl chloride substitution of acryloyl chloride with a di-Boc-protected triamine was followed by Boc deprotection and amide coupling to the NHS ester of 4-(diphenylphosphino)benzoic acid (Figures 1b, S09-S13) yielding BisTPPAm.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

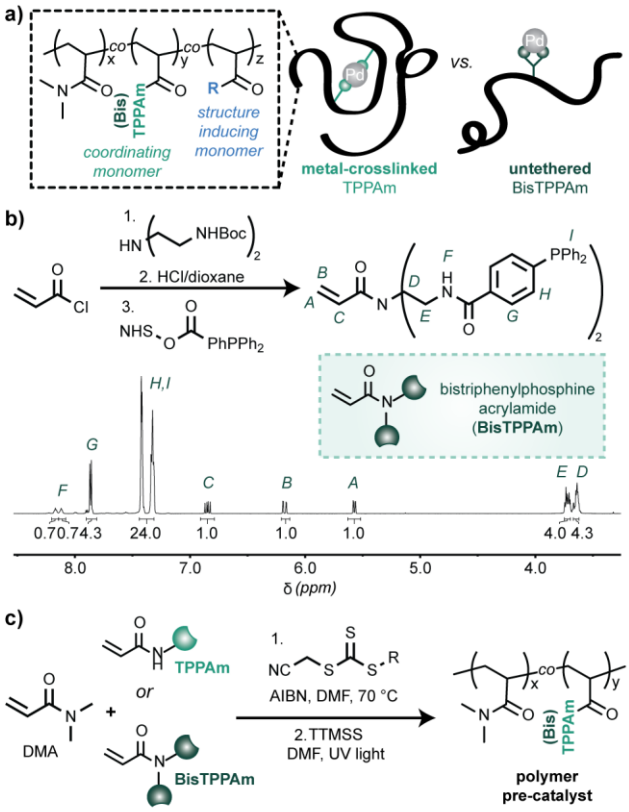

Figure 1. Bis(triphenylphosphine) acrylamide monomer (BisTPPAm) synthesis and copolymerization. a) Schematic illustrating the copolymer catalysts containing triphenylphosphine acrylamide (TPPAm) and BisTPPAm. b) Schematic of BisTPPAm synthesis and purified ^{1}H-NMR spectrum. c) Polymerization schematic for dimethylacrylamide (DMA) copolymers with each coordinating monomer, with R corresponding to a dodecyl carbon chain. Polymerization is followed by chain end capping with tris(trimethylsilyl) silane (TTMSS).

The rate of catalysis of BisTPPAm- and TPPAm-containing copolymers was initially compared through statistical copolymers synthesized using reversible addition fragmentation chain-transfer (RAFT) polymerization with dimethylacrylamide (DMA), an inert monomer (Figure 1c).[24] This polymerization method yielded polymers of controlled molecular weight (M_n), low dispersity (Đ), and equivalent rates of incorporation of both monomers indicative of statistical copolymers (Figures S14-S17, Table 1). Copolymers with molecular weights of 25 kDa were synthesized with TPPAm (5 mol%) or BisTPPAm (2.5 mol%, an equivalent phosphine percentage) to facilitate catalysis while limiting the impact of the ligand monomer on polymer conformation. Following polymer synthesis, the chain ends of these polymers were proton capped using tris(trimethylsilyl)silane and UV light in order to remove trithiocarbonates capable of forming palladium complexes (Figure S18).[25,26] To form the catalysts, the polymers were complexed to palladium in dilute solution (1 mg/mL) using dichloro(1,5-cyclooctadiene)palladium as a metal source. The complexed polymers were purified via precipitation in cold diethyl ether, and coordination of triphenylphosphine to the metal species was confirmed by ^{31}P-NMR (Figure S19-

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

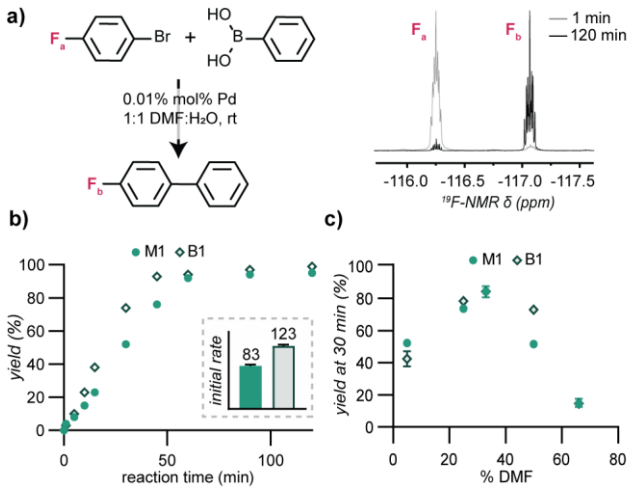

Figure 2. Comparison of the reactivity of polymer catalysts containing TPPAm (**M1**) and BisTPPAm (**B1**). a) Schematic of the model Suzuki-Miyaura cross-coupling reaction and ^{19}F-NMR spectra at 1 min and 120 min of reaction time. b) Plot of yield vs. reaction time comparing **M1** (filled) and **B1** (unfilled) in 1:1 DMF:H_2O. The calculation of initial reaction rate is inset. c) Comparison of **M1** (filled) and **B1** (unfilled) copolymer catalysts at different ratios of DMF:H_2O. Error bars correspond to the standard deviation (n=3). All reactions were run with 4-bromofluorobenzene (0.3M) at ambient temperature and atmosphere with 0.01 mol% Pd, and all time points were separate reactions.

S20). Dynamic light scattering (DLS) confirmed limited metal-induced multichain aggregates indicating that metal cross-links were intramolecular (Figure S21).

The Suzuki-Miyaura cross-coupling of bromofluorobenzene and phenylboronic acid was selected as a model reaction for this system,[27] as the conversion can be monitored using the fluorine as a probe for ^{19}F-NMR (Figure 2a).[28] For catalyst loading, we assumed one equivalent of Pd for every two triphenylphosphines (i.e., two TPPAm monomers per Pd and one BisTPPAm monomer per Pd). The BisTPPAm catalyst (**B1**) has an increased initial reaction rate (123 μM/sec) as compared to the previously characterized TPPAm-containing catalyst (**M1**, 83 μM/sec), in a 1:1 mixture of DMF:H_2O under ambient temperature and atmosphere (Figures 2b, S22, Table S2).[19] We hypothesize that this is due to the higher local phosphine concentration around the metal center inherent to BisTPPAm, which increases the likelihood that the proximal phosphines will be binding to the same Pd ion. This allows open coordination sites to be closed more rapidly than when the phosphines are further apart.[21,29,30]

As copolymers containing BisTPPAm lack the metal-induced cross-linking inherent to TPPAm, we probed how changing the solvent mixture would impact the relative reactivity. As the DMF content of the DMF:H_2O mixtures was decreased, the catalysts achieve similar yields at an early time point (30 min, Figure 2c); however, under aqueous conditions (i.e., 5% DMF), the TPPAm-containing catalyst **M1** achieves higher conversion than **B1 (**Figures 2c, S23**)**. We hypothesize that the TPPAm-containing catalyst **M1** forms a solvophobic pocket driven by the metal-ligand cross-links that aids solubilizing the relatively hydrophobic bromofluorobenzene, as evidenced by the increased yield. Similar reactivity between catalysts was observed in additional polar solvent mixtures (ethanol and acetone, 1:1 with H_2O), whereas THF led to negligible yields, likely due to limited solubility of the catalysts in the ethereal solvent (Figure S24).

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Ethanol/water mixtures were of particular interest due to their industrial implementation as environmentally benign solvents.[31,32] Both polymers behaved analogously in the 1:1 ethanol/water mixtures over extended timepoints, suggesting that catalyst identity plays less of a role in this solvent and that these reactions are more substrate controlled (Figure S25). Thus, an increase in local phosphine density can both increase initial reaction rates yet lower overall yields in aqueous solvents in the absence of metal-driven cross-links. In summary, local phosphine density, polymer conformation, and solvent environment concurrently impact catalytic efficacy.

Polymer	*Catalytic moiety (mol%)*	*Additional monomer (mol%)*	*Additional monomer (wt%)*	*Monomer feed ratio (mol%)*	*Measured monomer ratio (mol%)*	*MW (kDa)*	$Đ_{DMF}$[a]
M1	5% TPPAm	-	-	95:5	95:5	23	1.16
B1	2.5% BisTPPAm	-	-	97.5:2.5	97:3	25	1.12
M-2.5(+)	5% TPPAm	2.5% APTAC	5% APTAC	92.5:5:2.5	92:5:3	25	-
M-5(+)	5% TPPAm	5% APTAC	9% APTAC	90:5:5	90:5:5	26	-
M-10(+)	5% TPPAm	10% APTAC	15% APTAC	85:5:10	86:5:9	26	-
M-20(+)	5% TPPAm	20% APTAC	31% APTAC	75:5:20	73:6:21	26	-
M-35(+)	5% TPPAm	35% APTAC	46% APTAC	60.5:35	62:5:33	24	-
M-45(+)	5% TPPAm	47.5% APTAC	60% APTAC	47.5:5:47.5	47:5:48	24	-
B-2.5(+)	2.5% BisTPPAm	2.5% APTAC	5% APTAC	95:2.5:2.5	94:3:3	25	-
B-5(+)	2.5% BisTPPAm	5% APTAC	7% APTAC	92.5:2.5:5	93:3:4	26	-
B-10(+)	2.5% BisTPPAm	10% APTAC	13% APTAC	87.5:2.5:10	89:3:8	26	-
B-20(+)	2.5% BisTPPAm	20% APTAC	27% APTAC	77.5:2.5:20	79:3:18	25	-
B-35(+)	2.5% BisTPPAm	35% APTAC	44% APTAC	62.5:2.5:35	64:3:33	24	-
B-47.5(+)	2.5% BisTPPAm	47.5% APTAC	56% APTAC	50:2.5:47.5	52:3:45	25	-
M-2.5FF	5% TPPAm	2.5% FF	9% FF	92.5:5:2.5	91:6:3	24	1.17
M-5FF	5% TPPAm	5% FF	16% FF	90:5:5	90:5:5	24	1.14
M-10FF	5% TPPAm	10% FF	27% FF	85:5:10	84:6:10	24	1.22
M-9wBAA	5% TPPAm	6.5% BAA	9% BAA	88.5:5:6.5	87:7:6	24	1.20
M-16wBAA	5% TPPAm	12% BAA	16% BAA	83:5:12	82:6:12	23	1.21
M-27wBAA	5% TPPAm	21% BAA	27% BAA	74:5:21	74:5:21	24	1.22
B-2.5FF	2.5% BisTPPAm	2.5% FF	9% FF	95:2.5:2.5	94:3:3	24	1.15
B-5FF	2.5% BisTPPAm	5% FF	17% FF	92.5:2.5:5	91:3:6	23	1.16
B-10FF	2.5% BisTPPAm	10% FF	28% FF	87.5:2.5:10	86:3:11	24	1.15

Table 1. ^{1}H-NMR and DMF SEC characterization of polymer catalysts. Remaining monomer mol% is composed of dimethylacrylamide (DMA) for all polymers. Monomer abbreviations: (3-acrlyimidopropyl)trimethylammonium chloride (APTAC), di(phenylalanine) (FF), *N*-benzyl acrylamide (BAA). [a]Catalysts synthesized with APTAC were not compatible with DMF SEC characterization. Naming convention is as follows: M-2.5(+) = (M – TPPAm; 2.5 – mol% additional monomer; (+) – additional monomer identity).

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Impact of protein-mimetic structural elements on catalyst efficacy

The structural diversity of metalloenzyme catalysts leads to unique reactivity and selectivity compared to small molecule catalysts.[4] Using this as inspiration, we sought to understand the impact of the addition of structural elements (i.e., β-sheet-like and electrostatic interactions) on the observed catalytic profile. A di(phenylalanine) acrylamide monomer (FF, Figure 3a, S26-S28, Scheme S1) has been previously observed to imbue local (secondary-like) and global (tertiary and quaternary-like) structure into polymer systems,[20,33] and these structural elements can impact functionality, such as metal binding.[34] Although di(phenylalanine)-driven assembly is primarily achieved in aqueous media, assembly of di(phenylalanine) dipeptides has been observed in mixtures of organic solvents including DMF.[35,36] This suggests that FF could drive intramolecular interactions within polymer catalysts in aqueous:organic solvent mixtures. We hypothesized that the addition of FF would impact the relative rigidity of these polymer catalysts and therefore their catalytic profiles.

A series of TPPAm- and BisTPPAm-containing catalysts were synthesized with increasing FF content (2.5-10 mol%) to drive local interactions while maintaining solubility (**M-2.5FF, M-5FF, M-10FF, B-2.5FF, B-5FF, B-10FF** Figures S29-S32, Table 1). Dynamic light scattering of these polymeric materials revealed limited aggregation, suggesting interactions driven by FF are predominantly intramolecular (Figure S33-S34). At a constant Pd loading (0.01 mol%), increasing FF content of the macromolecular catalyst led to a slight elevation in the initial reaction rates for

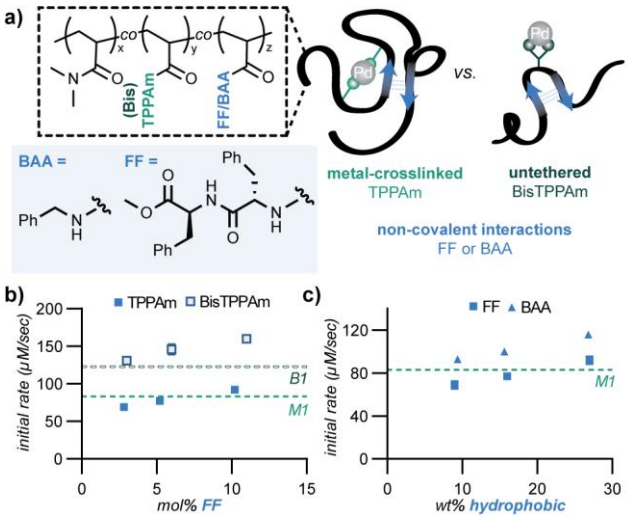

Figure 3. Comparison of hydrophobic monomers imbuing non-covalent interactions. a) Schematic illustrating polymers including local rigidity induced by a di(phenylalanine) acrylamide monomer (FF) and benzyl acrylamide (BAA). b) Plot of initial reaction rates for FF-containing polymers with TPPAm (filled, **M-2.5FF**, **M-5FF**, **M-10FF**) and BisTPPAm (unfilled, **B-2.5FF**, **B-5FF**, **B-10FF**). Dotted lines correspond to initial rates of parent polymers (**M1**, **B1**). c) Plot of initial reaction rates comparing TPPAm-containing polymers copolymerized with FF (square, **M-2.5FF**, **M-5FF**, **M-10FF**) and BAA (triangle, **M-9wBAA**, **M-16WBAA**, **M-27wBAA**). Dotted line corresponds to initial rate of parent polymer **M1**. All reactions were run with 4-bromofluorobenzene (0.3M) at ambient conditions with 0.01 mol% Pd, and all time points were separate reactions. Error bars correspond to the standard error of the rate calculation.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

both TPPAm- and BisTPPAm-containing copolymers (Figure 3b, S35-S36, Table S2). Interestingly, **M-2.5FF** displayed an initial rate decrease with a low incorporation ratio of FF (69 μM/sec) as compared to **M1** (83μM/sec), whereas **B-2.5FF** did not relative to **B1** (131 μM/sec vs 123μM/sec).

To probe this difference, we synthesized TPPAm-containing copolymers with benzyl acrylamide (BAA), a hydrophobic monomer lacking ordered hydrogen-bonds at weight percentages equivalent to the FF-containing copolymers (**M-9wBAA, M-16wBAA, M-27wBAA**; Figure S37, Table 1). These polymeric catalysts similarly demonstrated limited aggregation by DLS (Figure S38). We observed an analogous rate increase with increasing hydrophobicity for BAA-containing copolymers, but not a decrease in initial reaction rate compared to the parent polymer **M1** at low wt% incorporation (93 μM/sec **M-9wBAA** vs. 83 μM/sec **M1**) (Figure 3c, S39, Table S2). This suggests that the increased rigidity of FF-containing copolymers causes a slight decrease in reaction rate at low incorporation of FF. We hypothesize this is due to a decrease in the rate at which phosphines can close open coordination sites on the metal due to increased rigidity in FF polymers. This will not impact the analogous BisTPPAm-containing catalysts, due to their increased local phosphine density. While both FF and BAA confine the polymer chain, increasing initial reaction rates, FF further limits flexibility, diversifying the dynamics of the reactive sites.

Complementary to local structure, charged moieties have been attributed to both aiding in ligand solubility[37] and changing ligand conformation.[38,39] Specifically, cationic additives such as quaternary ammoniums have been shown to stabilize transition states in Suzuki-Miyaura cross-couplings.[40,41] To probe the impact of positively charged residues on the polymer catalysts, we synthesized a series of copolymers containing (3-acrylamidopropyl)triethylammonium chloride (APTAC) with both TPPAm (**M-2.5(+), M-5(+), M-10(+), M-20(+), M-35(+), M-47.5(+)**) and BisTPPAm (**B-2.5(+), B-5(+), B-10(+), B-20(+), B-35(+), B-47.5(+)**) (Figures 4a, S40-S41, Table 1). The incorporation of APTAC ranged from low loadings of 2.5 to 47.5 mol%, as higher incorporation ratios could not be polymerized due to mismatched solubility of the phosphine monomers and the high concentration of APTAC. Interestingly, these polymers displayed some aggregation observed by DLS at higher cationic monomer incorporation percentages (Figure S42-S43). Similar behavior has been observed for other polymer catalysts, where aggregation is observed at high concentrations, as DLS is necessarily performed at 10x excess of catalytic concentrations, yet single chain behavior is observed at reaction concentrations.[42] The polymers synthesized herein are similar in molecular weight and catalyst concentration, suggesting that while small multi-chain aggregates may form during the cross-coupling reactions, they are likely not the primary morphology in solution.

The initial reaction rates of APTAC-containing copolymers displayed a nonmonotonic dependence on APTAC incorporation ratio (Figures 4b, S44, Table S2). For TPPAm-containing copolymers, we observed that at low charge incorporations the initial reaction rate increased (**M-2.5(+)**, **M-5(+)**), followed by a dip at moderate charge incorporation (**M-10(+)**, **M-20(+)**), and an increase for densely charged catalysts (**M-35(+)**, **M-47.5(+)**). We hypothesize that at low incorporation ratios, the positive charge is favorably interacting with the reaction intermediates, while minimally effecting the solution conformation of the polymers. However, in the moderately charged species, the charged moieties arrange at the solvent-interface, potentially disrupting the metal cross-links. As the charge density continues to increase, the positive charges are forced into proximity with the catalytic sites, resulting in an increase in initial reaction rate. Similar behavior, where forced

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

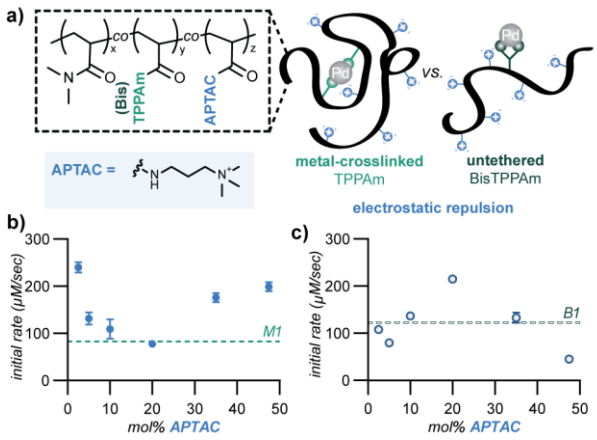

Figure 4. Comparison of electrostatic interactions within TPPAm-containing and BisTPPAm-containing catalysts. a) Schematic of polymer structures containing electrostatic interactions through the addition of (3-acrylimidopropyl) trimethylammonium chloride (APTAC). b) Plot of initial reaction rates for APTAC containing polymers with TPPAm (**M-2.5(+)**, **M-5(+)**, **M-10(+)**, **M-20(+)**, **M-35(+)**, **M-47.5(+)**). Dotted line corresponds to initial rate of parent polymer **M1**. c) Plot of initial rates of reaction for APTAC containing polymers with BisTPPAm (**B-2.5(+)**, **B-5(+)**, **B-10(+)**, **B-20(+)**, **B-35(+)**, **B-47.5(+)**). Dotted line corresponds to initial rate of parent polymer **B1**. Error bars correspond to the standard error of the rate calculation. All reactions were run in 1:1 DMF:H_2O, with 0.3M 4-bromofluorobenzene at room temperature and ambient conditions with 0.01mol% Pd loading, and all time points were separate reactions.

proximity of charged moieties causes increase in rate of reaction, has been seen in small molecule dendritic cationic phosphorous ligands.[43,44]

Interestingly, the BisTPPAm-containing catalysts displayed opposite trends than the analogous TPPAm-containing catalysts (Figures 4c, S45, Table S2). We hypothesize that at low incorporation ratios, there is not a sufficient proximity of the charged moieties to the catalytic sites to stabilize the reaction intermediates without the cross-links provided by TPPAm (**B-2.5(+)**, **B-5(+)**). As the incorporation ratio increases, proximity is achieved enabling intermediate stabilization (**B-10(+)**, **B-20(+)**). Without metal-ligand cross-links, at high charge densities the reactivity decreases (**B-35(+)**, **M-47.5(+)**). Similar charge densities have been demonstrated to increase chain repulsion, expanding the size of single chain nanoparticle assemblies.[45,46] Thus, we hypothesize the reactivity decreases as the polymer extends due to electrostatic repulsions, similar to a polyelectrolyte.

Through these data we have demonstrated that cationic monomer density significantly impacts copolymer reactivity, and that the reactivity is intertwined with additional polymer structural elements. At the extremes, low and high charge densities lead to increased reaction rates for metal cross-linked polymers and moderate charge densities lead to increased reaction rates for polymers lacking metal-driven cross-links. The incorporation of structure-inducing monomers has expanded the tunability of these polymer catalysts, leading to a broad range of accessible reactivity.

Predicting reactivity of diverse substrates in environmentally benign solvent mixtures

As we had identified that reactions in environmentally benign ethanol/water mixtures were substrate-controlled, we sought to identify chemical features of the substrate that modulated the

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

reactivity in our polymeric catalyst system. We therefore selected an expanded substrate scope of boronic acids and aryl halides to include substrates with varied sterics surrounding the reactive site (o-tolylboronic acid, 2-6-dimethylphenylboronic acid, and 2-bromo-4-fluorotoluene), substrate bulk (1-naphthylboronic acid, 1-pyreneboronic acid, and 2-bromo-6-fluoronaphthalene), and electronics (4-hydroxyphenylboronic acid, 4-pyridineboronic acid, and 4-methoxyphenylboronic acid; Figure 5a). Each aryl halide contains a fluorine probe to facilitate yield monitoring by ^{19}F-NMR. Two model polymer catalysts (**M1** and **B1**) were selected, as we hypothesized that these catalysts had the potential to vary in reactivity due to the presence and lack of metal cross-links. A gradient of ethanol/water mixtures ranging from 0-90% ethanol were selected as the solvent due to the growing popularity of ethanol/water mixtures as environmentally benign solvent alternatives.[31,32] With these conditions, 144 total reactions were performed; the final yield of each reaction (16 h, ambient temperature and atmosphere) was calculated by comparing the integrations of the fluorine signals. The resulting dataset contained yields from >90% to <5%, indicating that the substrate structure and solvent selection significantly impacts reactivity (Figure S46-S48). The reaction features, such as steric bulk and hydrophobicity, were selected due to their intuitive relationship with reactivity; however, extracting the relative impact of each of these features is challenging based on intuition alone.

Thus, we sought to quantify the impact of reaction features on yield through the construction of a predictive model and dataset visualization. This enables quantitative insight on which chemical features play a role in reactivity for this system. The performance of statistical models depends strongly on the selection of chemical descriptors and model type, and it is generally a significant challenge to know a priori which combination of descriptors and model will be optimal for a specific dataset.

We first benchmarked our dataset against a variety of featurization strategies, including cheminformatics approaches [RDkit, an open-source python package and density functional theory (DFT) calculations using Gaussian[47]], as well as statistical models spanning both linear and non-linear regressions. Through RDKit, over 200 molecular descriptors and extended connectivity fingerprints (ECFPs, capturing chemical topology) were calculated for each substrate using canonical SMILES as the input.[48] Given the large number of resulting features obtained, we considered strategies to reduce the dimensionality of model inputs. These strategies included data-driven methods, such as least absolute shrinkage and selection operator (LASSO) feature selection[49] and principal component analysis (PCA), as well as selection of descriptors based on chemical intuition. Concurrently, using DFT calculations, we extracted additional intuitive molecular descriptors (Table S3-S4): electronic parameters (natural bond orbital (NBO) charges, the dipole moment describing charge separation, and polarizability describing induced dipoles) and steric descriptors (B_{min} and B_{max} width parameters, and L length parameter from Sterimol).[50] To represent each of the two polymeric catalysts (**M1** and **B1),** we opted for label encoding, where each catalyst is labeled numerically rather than based on chemical composition.[51] Thus, for each reaction, the input feature vector was a concatenation of features of the aryl halide and boronic acid substrates, as well as the percentage of water as a cosolvent and the label-encoded catalyst.

We then fit selected feature sets to three types of supervised learning models – ensemble approaches (gradient boosting regressor, random forest regressor), a supported vector machine, and a linear regression – selected for differing abilities to capture nonlinearity and high-dimensional effects. The model performance for each regressor was captured by the root mean square error (RMSE; Table S5). This systematic analysis revealed the highest performing feature

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

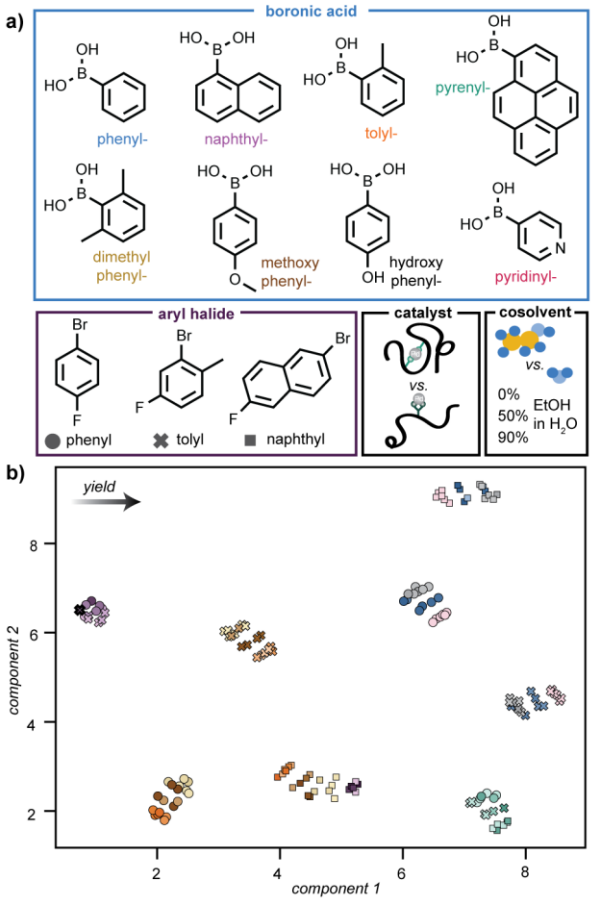

Figure 5. Reaction scope and chemical space representation of substrates. a) Reaction variables evaluated including substrates (boronic acid and aryl halides), copolymer catalysts, and solvent. b) Uniform manifold approximation and projection (UMAP) visualization of the chemical space of aryl halide and boronic acid chemical descriptors (color – boronic acid identity, shape – aryl halide identity) with color intensity-coded yields (darker color indicates higher yield).

set to be LogP from RDKit in addition to parameters extracted from DFT calculations fit to the gradient boosting regressor.

Visualization of the dataset aids in understanding the chemical space in which reactions are being evaluated. Given the success of the above featurization strategy, we then used it to visualize the entire 144-reaction dataset. Visualization axes are determined independent of yield, which offers an opportunity to visualize the chemical space being probed blind to the outcome. We used a uniform manifold approximation and projection (UMAP), a nonlinear dimension reduction strategy preferred for its ability to preserve global and local structure within a dataset.[52] Analogous to how a globe unfolds into a world map, UMAP projects a high-dimensional feature space into a low-dimensional representation. This visualization reveals the relationships between various coupling partners (Figure 5b). For example, boronic acids with similar chemical functionality (e.g., dimethylphenyl and tolylboronic acids) often end up in similar clusters based on aryl halide choice, and bulkier boronic acids (e.g., 1-pyreneboronic acid) appear segregated regardless of the aryl halide being coupled.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

This visualization also offers an opportunity to superimpose the yield, as represented by the color intensity in Figure 5b. For example, phenylboronic acid is consistently the highest yielding substrate, regardless of the location on the chemical space, which we hypothesize is due to the general stability of this boronic acid to side reactions.[53] Pyridineboronic acid, despite being chemically similar to phenylboronic acid, had low yields in each evaluated condition, as heteroaromatic substrates are challenging coupling partners due to probable interactions of the heteroatoms with the metal center.[54] Sterics around the reactive boronic acid also contribute to reactivity, with the chemically similar dimethylphenyl and tolylboronic acids achieving different yields due to the decreased accessibility of the dimethylphenyl reactive site.[55]

While the UMAP visualization guides intuitive relationships between the evaluated substrates, in order to identify quantitative relationships between the reaction yield and chemical features, a predictive algorithm is required (Figure 6a). A predictive algorithm was developed by fitting with a gradient boosting machine to the previous feature set, which showed a moderately predictive model (RMSE = 8.86; Figure 6b). Encouraged by this performance, we further interrogated the model to extract and interpret relative feature importance by calculating shapely additive explanations (SHAP) values, which have been previously used to quantitate relationships between descriptors in a diverse set of models (Figure 6c).[56] Solvent is the strongest predictor, as yield decreases with increased water content due to poor substrate solubility. The next strongest descriptors are features of charge delocalization surrounding the boronic acid moiety (NBO charge at the carbon adjacent to the boron and molecule polarization). Chemical intuition also supports the importance of these features, as charge delocalization increases the substrate nucleophilicity.[57] Finally, one of the steric bulk parameters for boronic acid (B_{max}) is significant given the importance of the accessibility of the boron site.[55] We had also hypothesized that differences in polymer structure may impact reactivity; however, the SHAP analysis classified the polymeric catalyst as minimally important. We hypothesize that further changes to the solution-phase polymer structure, such as covalent cross-links generating more compact structures, may lead to further differentiation. We look to further screen, rigorously featurize, and model the effects of polymer composition and structure on chemical reaction yield in future efforts.

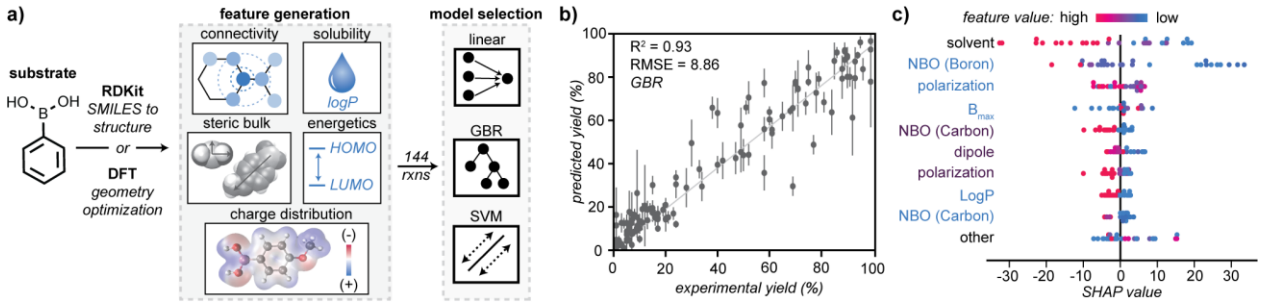

Figure 6. Quantitative characterization of features describing reaction substrates, catalysts, and solvent. a) Schematic of protocols evaluated for featurization of the aryl halide and boronic acid substrates in addition to linear and non-linear regressors evaluated (gradient boosted regressor = GBR, supported vector machine = SVM). b) Fitting of the GBR and model performance (i.e., RMSE). c) Shapely additive explanations (SHAP) analysis where red and blue datapoints indicate the feature value (red – high, blue – low). Feature labels in blue correspond to boronic acid features, purple corresponds to aryl halide features, and black corresponds to other reaction features.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

In summary, we have described a strategy to extract quantitative structure-function relationships in complex reaction mixtures and demonstrated the utility in a model polymeric catalyst system. We connected properties of reaction conditions and substrate chemistries to reaction performance. We anticipate this series of tools to visualize, featurize, and perform regression modelling will translate to other experimentally tractable, yet limited datasets at the interface of organic and polymer chemistries. This workflow may be used to elucidate relevant chemical features underlying reactivity and facilitate the iterative design of higher yielding reactions, in addition to optimization for alternative desired outcomes, such as selectivity and solvent compatibility.

Conclusion

We have expanded our platform of palladium-complexed macromolecular catalysts to uncover structure-function relationships that include ligand density, secondary-like local structure, and electrostatic interactions. We observed an increased reaction rate for bi-functional phosphines that are unlikely to generate metal cross-linked polymer catalysts. We also demonstrate that changes in catalyst flexibility, due to the introduction of hydrophobicity, lead to increases in reactivity; and nonmonotonic changes in reaction rate are observed with variation in charge density. These non-coordinating secondary sphere interactions are vital to protein folding, and we demonstrate that these interactions analogously play a critical role in shaping the reactivity of protein-mimetic synthetic catalysts. Finally, we identified that reactivity is substrate-driven in polar, environmentally benign solvents, and we developed a data-driven workflow to quantify the relative impacts of specific substrate parameters.

Synthetic macromolecules offer nearly limitless chemical functionality, making them prime targets for catalyst development. By incorporating additional protein-mimetic complexities, including multiple charges, covalent and non-covalent cross-linking strategies, and additional changes to the dynamics of the reactive site,[58,59] we anticipate further refinement of the reactivity, yielding efficient, selective, and functional materials in diverse environments. Furthermore, the versatility of phosphine-based ligands allows for complexation of a variety of metals,[60,61] enabling a multitude of macromolecular catalysts for a variety of reaction types and conditions. These materials begin to bridge the gap between enzymes and small molecules, combining the capability for complex folding with the vast chemical diversity beyond that of biological monomers.

Acknowledgements

This material is based upon work supported by the U.S, Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0021295. We acknowledge support from the UNC Macromolecular Interactions Facility in DLS experiments. The facility is supported in part by the National Cancer Institute of the National Institutes of Health under award number P30CA016086. S.S.C. acknowledges the National Science Foundation Graduate Research Fellowship (DGE-2040435). We acknowledge the UNC Department of Chemistry Mass Spectrometry Core Laboratory (HRMS is supported by the National Science Foundation under Grant No. CHE-1726291), and the NMR core laboratory staff Drs. Marc ter Horst, Nuwanthika Dilrukshi, and Luke Fulton, supported by a National Science Foundation award number CHE-1828183. We also thank the UNC Research Computing group for providing the computational resources and support towards these results, the Gagne and Johnson research groups at UNC in addition to Prof. Meredith Borden for valuable conversations.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

References

(1) Chen, A. Y.; Adamek, R. N.; Dick, B. L.; Credille, C. V.; Morrison, C. N.; Cohen, S. M. Targeting Metalloenzymes for Therapeutic Intervention. *Chem. Rev.* **2019**, *119* (2), 1323–1455. https://doi.org/10.1021/acs.chemrev.8b00201.

(2) Zhao, M.; Wang, H.-B.; Ji, L.-N.; Mao, Z.-W. Insights into Metalloenzyme Microenvironments: Biomimetic Metal Complexes with a Functional Second Coordination Sphere. *Chem. Soc. Rev.* **2013**, *42* (21), 8360. https://doi.org/10.1039/c3cs60162e.

(3) Van Stappen, C.; Deng, Y.; Liu, Y.; Heidari, H.; Wang, J.-X.; Zhou, Y.; Ledray, A. P.; Lu, Y. Designing Artificial Metalloenzymes by Tuning of the Environment beyond the Primary Coordination Sphere. *Chem. Rev.* **2022**, *122* (14), 11974–12045. https://doi.org/10.1021/acs.chemrev.2c00106.

(4) Rosati, F.; Roelfes, G. Artificial Metalloenzymes. *ChemCatChem* **2010**, *2* (8), 916–927. https://doi.org/10.1002/cctc.201000011.

(5) Amamoto, Y. Data-Driven Approaches for Structure-Property Relationships in Polymer Science for Prediction and Understanding. *Polym. J.* **2022**, *54* (8), 957–967. https://doi.org/10.1038/s41428-022-00648-6.

(6) Barbee, M. H.; Wright, Z. M.; Allen, B. P.; Taylor, H. F.; Patteson, E. F.; Knight, A. S. Protein-Mimetic Self-Assembly with Synthetic Macromolecules. *Macromolecules* **2021**, *54* (8), 3585–3612. https://doi.org/10.1021/acs.macromol.0c02826.

(7) Dong, Y.; Bi, J.; Zhang, S.; Zhu, D.; Meng, D.; Ming, S.; Qin, K.; Liu, Q.; Guo, L.; Li, T. Palladium Supported on N-Heterocyclic Carbene Functionalized Hydroxyethyl Cellulose as a Novel and Efficient Catalyst for the Suzuki Reaction in Aqueous Media. *Appl. Surf. Sci.* **2020**, *531*, 147392. https://doi.org/10.1016/j.apsusc.2020.147392.

(8) Diyali, N.; Rasaily, S.; Biswas, B. Metal–Organic Framework: An Emergent Catalyst in C–N Cross-Coupling Reactions. *Coord. Chem. Rev.* **2022**, *469*, 214667. https://doi.org/10.1016/j.ccr.2022.214667.

(9) Wei, Y. S.; Zhang, M.; Zou, R.; Xu, Q. Metal-Organic Framework-Based Catalysts with Single Metal Sites. *Chem. Rev.* **2020**, *120* (21), 12089–12174. https://doi.org/10.1021/acs.chemrev.9b00757.

(10) Dong, Y.; Zhou, Z.; Wang, Y.; Li, X.; Li, T.; Ren, Y.; Hu, W.; Zhang, L.; Zhang, X.; Wei, C. Palladium Supported on Pyrrole Functionalized Hypercrosslinked Polymer: Synthesis and Its Catalytic Evaluations towards Suzuki-Miyaura Coupling Reactions in Aqueous Media. *J. Mol. Liq.* **2022**, *368*, 120679. https://doi.org/10.1016/j.molliq.2022.120679.

(11) Verde-Sesto, E.; Arbe, A.; Moreno, A. J.; Cangialosi, D.; Alegría, A.; Colmenero, J.; Pomposo, J. A. Single-Chain Nanoparticles: Opportunities Provided by Internal and External Confinement. *Mater. Horiz.* **2020**, *7* (9), 2292–2313. https://doi.org/10.1039/D0MH00846J.

(12) Chen, J.; Garcia, E. S.; Zimmerman, S. C. Intramolecularly Cross-Linked Polymers: From Structure to Function with Applications as Artificial Antibodies and Artificial Enzymes. *Acc. Chem. Res.* **2020**, *53* (6), 1244–1256. https://doi.org/10.1021/acs.accounts.0c00178.

(13) Rothfuss, H.; Roesky, P. W.; Barner-kowollik, C. Single-Chain Nanoparticles as Catalytic Nanoreactors. **2018**. https://doi.org/10.1021/jacs.8b02135.

(14) Mundsinger, K.; Izuagbe, A.; Tuten, B. T.; Roesky, P. W.; Barner-Kowollik, C. Single Chain Nanoparticles in Catalysis. *Angew. Chem. Int. Ed.* **2023**, e202311734. https://doi.org/10.1002/anie.202311734.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

(15) Pomposo, J. A.; Perez-Baena, I.; Lo Verso, F.; Moreno, A. J.; Arbe, A.; Colmenero, J. How Far Are Single-Chain Polymer Nanoparticles in Solution from the Globular State? *ACS Macro Lett.* **2014**, *3* (8), 767–772. https://doi.org/10.1021/mz500354q.

(16) Chen, J.; Wang, J.; Bai, Y.; Li, K.; Garcia, E. S.; Ferguson, A. L.; Zimmerman, S. C. Enzyme-like Click Catalysis by a Copper-Containing Single-Chain Nanoparticle. *J. Am. Chem. Soc.* **2018**, *140* (42), 13695–13702. https://doi.org/10.1021/jacs.8b06875.

(17) Sathyan, A.; Croke, S.; Pérez-López, A. M.; de Waal, B. F. M.; Unciti-Broceta, A.; Palmans, A. R. A. Developing Pd(II) Based Amphiphilic Polymeric Nanoparticles for pro-Drug Activation in Complex Media. *Mol. Syst. Des. Eng.* **2022**, *7* (12), 1736–1748. https://doi.org/10.1039/D2ME00173J.

(18) Deng, L.; Sathyan, A.; Adam, C.; Unciti-Broceta, A.; Sebastian, V.; Palmans, A. R. A. Enhanced Efficiency of Pd(0)-Based Single Chain Polymeric Nanoparticles for *in Vitro* Prodrug Activation by Modulating the Polymer's Microstructure. *Nano Lett.* **2024**, *24* (7), 2242–2249. https://doi.org/10.1021/acs.nanolett.3c04466.

(19) Sanders, M. A.; Chittari, S. S.; Sherman, N.; Foley, J. R.; Knight, A. S. Versatile Triphenylphosphine-Containing Polymeric Catalysts and Elucidation of Structure–Function Relationships. *J. Am. Chem. Soc.* **2023**, *145* (17), 9686–9692. https://doi.org/10.1021/jacs.3c01092.

(20) Warren, J. L.; Dykeman-Bermingham, P. A.; Knight, A. S. Controlling Amphiphilic Polymer Folding beyond the Primary Structure with Protein-Mimetic Di(Phenylalanine). *J. Am. Chem. Soc.* **2021**, *143* (33), 13228–13234. https://doi.org/10.1021/jacs.1c05659.

(21) Clevenger, A. L.; Stolley, R. M.; Aderibigbe, J.; Louie, J. Trends in the Usage of Bidentate Phosphines as Ligands in Nickel Catalysis. *Chem. Rev.* **2020**, *120* (13), 6124–6196. https://doi.org/10.1021/acs.chemrev.9b00682.

(22) Artar, M.; Terashima, T.; Sawamoto, M.; Meijer, E. W.; Palmans, A. R. A. Understanding the Catalytic Activity of Single-chain Polymeric Nanoparticles in Water. *J. Polym. Sci. Part Polym. Chem.* **2014**, *52* (1), 12–20. https://doi.org/10.1002/pola.26970.

(23) Guinó, M.; Hii, K. K. (Mimi). Applications of Phosphine-Functionalised Polymers in Organic Synthesis. *Chem Soc Rev* **2007**, *36* (4), 608–617. https://doi.org/10.1039/B603851B.

(24) Perrier, S. *50th Anniversary Perspective*: RAFT Polymerization—A User Guide. *Macromolecules* **2017**, *50* (19), 7433–7447. https://doi.org/10.1021/acs.macromol.7b00767.

(25) Fackler, J. P.; Seidel, W. C. Sulfur Ligand Complexes. IX. Reactions of Metal Xanthates and Their Derivatives. The Formation of Bisphosphine-Dithiocarbonate and -Trithiocarbonate Complexes of Palladium(II) and Platinum(II). *Inorg. Chem.* **1969**, *8* (8), 1631–1639. https://doi.org/10.1021/ic50078a012.

(26) Jones, P. G.; Sheldrick, G. M.; Usón, R.; Forniés, J.; Usón, M. A. Dithiocarbonate and Trithiocarbonate Complexes of Palladium(II); Crystal Structure of Pd(Ph3P)2(CS20) CH2C12. *Z. Für Naturforschung B* **1983**, *38* (4), 449–453. https://doi.org/10.1515/znb-1983-0410.

(27) Miyaura, Norio.; Suzuki, Akira. Palladium-Catalyzed Cross-Coupling Reactions of Organoboron Compounds. *Chem. Rev.* **1995**, *95* (7), 2457–2483. https://doi.org/10.1021/cr00039a007.

(28) Zientek, N.; Laurain, C.; Meyer, K.; Kraume, M.; Guthausen, G.; Maiwald, M. Simultaneous 19F–1H Medium Resolution NMR Spectroscopy for Online Reaction Monitoring. *J. Magn. Reson.* **2014**, *249*, 53–62. https://doi.org/10.1016/j.jmr.2014.10.007.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

(29) Adnan, R. H.; Madridejos, J. M. L.; Alotabi, A. S.; Metha, G. F.; Andersson, G. G. A Review of State of the Art in Phosphine Ligated Gold Clusters and Application in Catalysis. *Adv. Sci.* **2022**, *9* (15), 2105692. https://doi.org/10.1002/advs.202105692.

(30) Krogstad, D. A.; Young, V. G.; Pignolet, L. H. Synthesis and Characterization of a Series of Diphosphine Ligated Pt□Au Cluster Compounds. *Inorganica Chim. Acta* **1997**, *264* (1–2), 19–32. https://doi.org/10.1016/S0020-1693(97)05587-4.

(31) Polshettiwar, V.; Decottignies, A.; Len, C.; Fihri, A. Suzuki-Miyaura Cross-Coupling Reactions in Aqueous Media: Green and Sustainable Syntheses of Biaryls. *ChemSusChem* **2010**, *3* (5), 502–522. https://doi.org/10.1002/cssc.200900221.

(32) Hooshmand, S. E.; Heidari, B.; Sedghi, R.; Varma, R. S. Recent Advances in the Suzuki–Miyaura Cross-Coupling Reaction Using Efficient Catalysts in Eco-Friendly Media. *Green Chem.* **2019**, *21* (3), 381–405. https://doi.org/10.1039/C8GC02860E.

(33) Farazi, S.; Stenzel, M. H.; Chapman, R. Confinement of Folding Motifs within Central Blocks Improves Single Chain Polymer Nanoparticle Folding. *Polym. Chem.* **2024**, *15* (4), 332–340. https://doi.org/10.1039/D3PY01166F.

(34) Dykeman-Bermingham, P. A.; Bogen, M. P.; Chittari, S. S.; Grizzard, S. F.; Knight, A. S. Tailoring Hierarchical Structure and Rare Earth Affinity of Compositionally Identical Polymers via Sequence Control. *J. Am. Chem. Soc.* **2024**, *146* (12), 8607–8617. https://doi.org/10.1021/jacs.4c00440.

(35) Casanovas, J.; Mayans, E.; Díaz, A.; Gil, A. M.; Jiménez, A. I.; Cativiela, C.; Puiggalí, J.; Alemán, C. Amyloid Fibrils from Organic Solutions of an Amphiphilic Dipeptide. *Chem. Commun.* **2019**, *55* (59), 8556–8559. https://doi.org/10.1039/C9CC04139G.

(36) Mayans, E.; Ballano, G.; Sendros, J.; Font-Bardia, M.; Campos, J. L.; Puiggalí, J.; Cativiela, C.; Alemán, C. Effect of Solvent Choice on the Self-Assembly Properties of a Diphenylalanine Amphiphile Stabilized by an Ion Pair. *ChemPhysChem* **2017**, *18* (14), 1888–1896. https://doi.org/10.1002/cphc.201700180.

(37) DeVasher, R. B.; Spruell, J. M.; Dixon, D. A.; Broker, G. A.; Griffin, S. T.; Rogers, R. D.; Shaughnessy, K. H. Experimental and Computational Study of Steric and Electronic Effects on the Coordination of Bulky, Water-Soluble Alkylphosphines to Palladium under Reducing Conditions: Correlation to Catalytic Activity. *Organometallics* **2005**, *24* (5), 962–971. https://doi.org/10.1021/om049241w.

(38) Moore, L. R.; Western, E. C.; Craciun, R.; Spruell, J. M.; Dixon, D. A.; O'Halloran, K. P.; Shaughnessy, K. H. Sterically Demanding, Sulfonated, Triarylphosphines: Application to Palladium-Catalyzed Cross-Coupling, Steric and Electronic Properties, and Coordination Chemistry. *Organometallics* **2008**, *27* (4), 576–593. https://doi.org/10.1021/om7008606.

(39) Snelders, D. J. M.; van der Burg, C.; Lutz, M.; Spek, A. L.; van Koten, G.; Klein Gebbink, R. J. M. Structure–Activity Relationships of Oligocationic, Ammonium-Functionalized Triarylphosphines as Ligands in the Pd-Catalyzed Suzuki–Miyaura Reaction. *ChemCatChem* **2010**, *2* (11), 1425–1437. https://doi.org/10.1002/cctc.201000232.

(40) Castanet, A.-S.; Colobert, F.; Desmurs, J.-R.; Schlama, T. Biaryl Synthesis via Suzuki Coupling Promoted by Catalytic Amounts of Quaternary Ammonium Salts. *J. Mol. Catal. Chem.* **2002**, *182–183*, 481–487. https://doi.org/10.1016/S1381-1169(01)00489-7.

(41) Bedford, R. B.; Blake, M. E.; Butts, C. P.; Holder, D. The Suzuki Coupling of Aryl Chlorides in TBAB–Water Mixtures. *Chem. Commun.* **2003**, No. 4, 466–467. https://doi.org/10.1039/b211329e.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

(42) M. Xiong, T.; S. Garcia, E.; Chen, J.; Zhu, L.; J. Alzona, A.; C. Zimmerman, S. Enzyme-like Catalysis by Single Chain Nanoparticles That Use Transition Metal Cofactors. *Chem. Commun.* **2022**, *58* (7), 985–988. https://doi.org/10.1039/D1CC05578J.

(43) Snelders, D. J. M.; Kreiter, R.; Firet, J. J.; Van Koten, G.; Klein Gebbink, R. J. M. Fast Suzuki–Miyaura Cross-Coupling Reaction with Hexacationic Triarylphosphine Bn- *Dendriphos* as Ligand. *Adv. Synth. Catal.* **2008**, *350* (2), 262–266. https://doi.org/10.1002/adsc.200700366.

(44) Snelders, D. J. M.; Van Koten, G.; Klein Gebbink, R. J. M. Hexacationic *Dendriphos* Ligands in the Pd-Catalyzed Suzuki−Miyaura Cross-Coupling Reaction: Scope and Mechanistic Studies. *J. Am. Chem. Soc.* **2009**, *131* (32), 11407–11416. https://doi.org/10.1021/ja904042h.

(45) Wen, J.; Zhang, J.; Zhang, Y.; Yang, Y.; Zhao, H. Controlled Self-Assembly of Amphiphilic Monotailed Single-Chain Nanoparticles. *Polym. Chem.* **2014**, *5* (13), 4032. https://doi.org/10.1039/c4py00100a.

(46) Mortimer, D. A. Synthetic Polyelectrolytes—A Review. *Polym. Int.* **1991**, *25* (1), 29–41. https://doi.org/10.1002/pi.4990250107.

(47) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; 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, J. A. Jr.; 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. Gaussian 16, 2016. https://gaussian.com/.

(48) Gregory Landrum. RDkit: Open-Source Cheminformatics Software, 2006. https://www.rdkit.org (accessed 2023-11-02).

(49) Muthukrishnan, R.; Rohini, R. LASSO: A Feature Selection Technique in Predictive Modeling for Machine Learning. In *2016 IEEE International Conference on Advances in Computer Applications (ICACA)*; IEEE: Coimbatore, India, 2016; pp 18–20. https://doi.org/10.1109/ICACA.2016.7887916.

(50) Santiago, C. B.; Guo, J.-Y.; Sigman, M. S. Predictive and Mechanistic Multivariate Linear Regression Models for Reaction Development. *Chem. Sci.* **2018**, *9* (9), 2398–2412. https://doi.org/10.1039/C7SC04679K.

(51) Day, E. C.; Chittari, S. S.; Cunha, K. C.; Zhao, J.; Dodds, J. N.; Davis, D. C.; Baker, E. S.; Berlow, R. B.; Shea, J.-E.; Kulkarni, R. U.; Knight, A. S. A High-Throughput Workflow to Analyze Sequence-Conformation Relationships and Explore Hydrophobic Patterning in Disordered Peptoids. ChemRxiv October 4, 2023. https://doi.org/10.26434/chemrxiv-2023-b84wl.

(52) McInnes, L.; Healy, J.; Saul, N.; Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **2018**, *3* (29), 861. https://doi.org/10.21105/joss.00861.

(53) Kuivila, H. G.; Reuwer Jr., J. F.; Mangravite, J. A. Electrophilic Displacement Reactions: XV. Kinetics and Mechanism of the Base-Catalyze Protodeboronation of Areneboronic Acids. *Can. J. Chem.* **1963**, *41* (12), 3081–3090. https://doi.org/10.1139/v63-451.

(54) Zhao, D.; You, J.; Hu, C. Recent Progress in Coupling of Two Heteroarenes. *Chem. – Eur. J.* **2011**, *17* (20), 5466–5492. https://doi.org/10.1002/chem.201003039.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

(55) Yin, J.; Rainka, M. P.; Zhang, X.-X.; Buchwald, S. L. A Highly Active Suzuki Catalyst for the Synthesis of Sterically Hindered Biaryls: Novel Ligand Coordination. *J. Am. Chem. Soc.* **2002**, *124* (7), 1162–1163. https://doi.org/10.1021/ja017082r.

(56) Karpovich, C.; Pan, E.; Jensen, Z.; Olivetti, E. Interpretable Machine Learning Enabled Inorganic Reaction Classification and Synthesis Condition Prediction. *Chem. Mater.* **2023**, *35* (3), 1062–1079. https://doi.org/10.1021/acs.chemmater.2c03010.

(57) Thomas, A. A.; Zahrt, A. F.; Delaney, C. P.; Denmark, S. E. Elucidating the Role of the Boronic Esters in the Suzuki–Miyaura Reaction: Structural, Kinetic, and Computational Investigations. *J. Am. Chem. Soc.* **2018**, *140* (12), 4401–4416. https://doi.org/10.1021/jacs.8b00400.

(58) Azuma, Y.; Terashima, T.; Sawamoto, M. Self-Folding Polymer Iron Catalysts for Living Radical Polymerization. *ACS Macro Lett.* **2017**, *6* (8), 830–835. https://doi.org/10.1021/acsmacrolett.7b00498.

(59) Hanlon, A. M.; Lyon, C. K.; Berda, E. B. What Is Next in Single-Chain Nanoparticles? *Macromolecules* **2016**, *49* (1), 2–14. https://doi.org/10.1021/acs.macromol.5b01456.

(60) Knöfel, N. D.; Rothfuss, H.; Willenbacher, J.; Barner-Kowollik, C.; Roesky, P. W. Platinum(II)-Crosslinked Single-Chain Nanoparticles: An Approach towards Recyclable Homogeneous Catalysts. *Angew. Chem. Int. Ed.* **2017**, *56* (18), 4950–4954. https://doi.org/10.1002/anie.201700718.

(61) Pittman, C. U.; Smith, L. R.; Hanes, R. M. Catalytic Reactions Using Polymer-Bound vs. Homogeneous Complexes of Nickel, Rhodium, and Ruthenium. *J. Am. Chem. Soc.* **1975**, *97* (7), 1742–1748. https://doi.org/10.1021/ja00840a021.

https://doi.org/10.26434/chemrxiv-2024-mjxf4 **ORCID:** https://orcid.org/0000-0002-6107-9352 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0