Supplementary Information -A Materials Discovery Framework Based on Conditional Generative Models Applied to the Design of Polymer Electrolytes

Khajeh et al.

TABLE S1: Hyperparameter selection for the GPT model. [1]

Model	Hyperparameters	Tested values	Optimal values	# of weights
GPT	step	1000, 2000, ..., 10000	6000	~0.12M
	model architecture	"gpt2", "gpt-mini", "gpt-nano"	"gpt-nano"	
	temperature	0.1, 0.5, 1.0	1.0	

TABLE S2: Grid search results for the GPT model (part 1). [1]

Hyperparameters			Evaluation metric					
epoch	temperature	architecture	validity	novelty	uniqueness	synthesizability	diversity	similarity
1000	0.1	gpt2	0.905	0.760	0.540	0.947	0.685	0.265
1000	0.5	gpt2	0.815	0.825	0.955	0.943	0.714	0.269
1000	1.0	gpt2	0.695	0.930	1.000	0.744	0.745	0.259
2000	0.1	gpt2	0.97	0.815	0.460	0.986	0.682	0.267
2000	0.5	gpt2	0.92	0.755	0.980	0.895	0.715	0.270
2000	1.0	gpt2	0.785	0.930	0.995	0.762	0.745	0.260
3000	0.1	gpt2	0.98	0.675	0.325	0.778	0.627	0.280
3000	0.5	gpt2	0.95	0.785	0.925	0.884	0.691	0.276
3000	1.0	gpt2	0.79	0.915	1.000	0.716	0.744	0.260
4000	0.1	gpt2	0.925	0.740	0.395	0.938	0.636	0.277
4000	0.5	gpt2	0.925	0.725	0.940	0.919	0.693	0.277
4000	1.0	gpt2	0.815	0.835	1.000	0.746	0.736	0.264
5000	0.1	gpt2	0.995	0.625	0.315	0.837	0.641	0.267
5000	0.5	gpt2	0.93	0.730	0.910	0.873	0.707	0.269
5000	1.0	gpt2	0.82	0.835	1.000	0.809	0.746	0.259
6000	0.1	gpt2	0.925	0.790	0.545	0.957	0.655	0.276
6000	0.5	gpt2	0.93	0.730	0.960	0.900	0.705	0.276
6000	1.0	gpt2	0.83	0.870	1.000	0.829	0.732	0.267
7000	0.1	gpt2	0.885	0.540	0.310	0.706	0.659	0.275
7000	0.5	gpt2	0.915	0.730	0.900	0.771	0.697	0.276
7000	1.0	gpt2	0.82	0.915	1.000	0.755	0.735	0.265
8000	0.1	gpt2	0.87	0.830	0.535	0.904	0.666	0.264
8000	0.5	gpt2	0.83	0.790	0.980	0.862	0.717	0.269
8000	1.0	gpt2	0.63	0.890	1.000	0.721	0.750	0.257
9000	0.1	gpt2	0.85	0.795	0.705	0.989	0.659	0.252
9000	0.5	gpt2	0.78	0.860	0.980	0.839	0.716	0.268
9000	1.0	gpt2	0.665	0.905	0.990	0.717	0.754	0.254
10000	0.1	gpt2	0.76	0.660	0.475	0.825	0.652	0.254
10000	0.5	gpt2	0.755	0.850	0.965	0.888	0.720	0.257
10000	1.0	gpt2	0.54	0.935	1.000	0.726	0.745	0.259

TABLE S3: Grid search results for the GPT model (part 2). [1]

1000	0.1	gpt-mini	0.99	0.925	0.110	0.750	0.637	0.267
1000	0.5	gpt-mini	0.945	0.735	0.860	0.908	0.703	0.269
1000	1.0	gpt-mini	0.83	0.910	1.000	0.818	0.736	0.262
2000	0.1	gpt-mini	1.0	0.655	0.260	0.969	0.687	0.255
2000	0.5	gpt-mini	0.98	0.655	0.950	0.910	0.713	0.269
2000	1.0	gpt-mini	0.93	0.830	0.990	0.855	0.747	0.258
3000	0.1	gpt-mini	1.0	0.415	0.085	0.833	0.636	0.277
3000	0.5	gpt-mini	0.99	0.620	0.885	0.909	0.712	0.267
3000	1.0	gpt-mini	0.965	0.765	0.990	0.757	0.745	0.260
4000	0.1	gpt-mini	1.0	0.300	0.190	0.941	0.647	0.261
4000	0.5	gpt-mini	0.995	0.620	0.910	0.889	0.713	0.261
4000	1.0	gpt-mini	0.955	0.690	0.990	0.828	0.747	0.259
5000	0.1	gpt-mini	1.0	0.350	0.205	0.900	0.654	0.278
5000	0.5	gpt-mini	1.0	0.475	0.940	0.945	0.707	0.272
5000	1.0	gpt-mini	0.995	0.650	0.985	0.825	0.740	0.262
6000	0.1	gpt-mini	1.0	0.250	0.165	0.929	0.687	0.264
6000	0.5	gpt-mini	1.0	0.440	0.905	0.940	0.715	0.271
6000	1.0	gpt-mini	0.98	0.565	1.000	0.862	0.742	0.262
7000	0.1	gpt-mini	1.0	0.260	0.270	0.947	0.703	0.268
7000	0.5	gpt-mini	1.0	0.425	0.925	0.913	0.699	0.276
7000	1.0	gpt-mini	1.0	0.500	0.975	0.860	0.735	0.264
8000	0.1	gpt-mini	1.0	0.160	0.235	1.000	0.634	0.273
8000	0.5	gpt-mini	0.995	0.260	0.875	0.898	0.720	0.265
8000	1.0	gpt-mini	0.965	0.480	0.990	0.820	0.738	0.261
9000	0.1	gpt-mini	0.955	0.215	0.315	1.000	0.683	0.273
9000	0.5	gpt-mini	1.0	0.265	0.870	0.940	0.717	0.269
9000	1.0	gpt-mini	0.975	0.400	0.985	0.867	0.739	0.263
10000	0.1	gpt-mini	1.0	0.040	0.215	0.667	0.723	0.253
10000	0.5	gpt-mini	1.0	0.135	0.900	0.889	0.714	0.267
10000	1.0	gpt-mini	0.985	0.310	0.980	0.847	0.738	0.264
1000	0.1	gpt-nano	0.055	1.000	0.110	0.600	0.616	0.280
1000	0.5	gpt-nano	0.72	0.900	0.920	0.681	0.695	0.276
1000	1.0	gpt-nano	0.6	0.965	0.995	0.619	0.728	0.264
2000	0.1	gpt-nano	0.97	0.465	0.180	0.818	0.605	0.278

TABLE S4: Grid search results for the GPT model (part 3). [1]

2000	0.5	gpt-nano	0.835	0.795	0.930	0.874	0.681	0.280
2000	1.0	gpt-nano	0.76	0.925	0.995	0.691	0.729	0.266
3000	0.1	gpt-nano	1.0	0.965	0.095	0.941	0.643	0.268
3000	0.5	gpt-nano	0.94	0.770	0.905	0.908	0.690	0.274
3000	1.0	gpt-nano	0.77	0.915	1.000	0.745	0.729	0.266
4000	0.1	gpt-nano	1.0	0.170	0.095	0.857	0.651	0.275
4000	0.5	gpt-nano	0.99	0.735	0.930	0.957	0.698	0.275
4000	1.0	gpt-nano	0.825	0.850	1.000	0.896	0.732	0.265
5000	0.1	gpt-nano	1.0	0.500	0.175	1.000	0.673	0.259
5000	0.5	gpt-nano	0.985	0.710	0.910	0.953	0.710	0.268
5000	1.0	gpt-nano	0.89	0.855	0.995	0.832	0.736	0.264
6000	0.1	gpt-nano	1.0	0.400	0.150	0.857	0.655	0.265
6000	0.5	gpt-nano	0.985	0.730	0.940	0.933	0.700	0.275
6000	1.0	gpt-nano	0.94	0.910	1.000	0.788	0.732	0.266
7000	0.1	gpt-nano	1.0	0.660	0.065	0.857	0.660	0.277
7000	0.5	gpt-nano	0.995	0.710	0.885	0.913	0.705	0.272
7000	1.0	gpt-nano	0.915	0.810	0.995	0.736	0.744	0.260
8000	0.1	gpt-nano	1.0	0.370	0.075	0.667	0.616	0.271
8000	0.5	gpt-nano	0.995	0.665	0.865	0.885	0.701	0.268
8000	1.0	gpt-nano	0.91	0.795	0.995	0.773	0.738	0.263
9000	0.1	gpt-nano	1.0	0.415	0.140	1.000	0.625	0.260
9000	0.5	gpt-nano	0.975	0.745	0.935	0.912	0.708	0.270
9000	1.0	gpt-nano	0.935	0.825	1.000	0.796	0.746	0.259
10000	0.1	gpt-nano	1.0	0.400	0.090	0.800	0.643	0.260
10000	0.5	gpt-nano	0.985	0.725	0.915	0.939	0.707	0.268
10000	1.0	gpt-nano	0.93	0.850	1.000	0.840	0.734	0.265

TABLE S5: Average of computed ion transport properties of generated polymers in the 1^st iteration

Monomer SMILES	Density (gr/cm3)	Li Diffusivity (cm2/s)	TFSI Diffusivity (cm2/s)	Cluster-Nernst-Einstein Conductivity (S/cm)	Nernst-Einstein Conductivity (S/cm)	Cluster-Nernst-Einstein Transference Number	Nernst-Einstein Transference Number
ONCCOC	1.31	2.80E-07	3.59E-07	2.04E-03	2.69E-03	4.21E-01	4.40E-01
OCCOC	1.30	1.94E-07	2.49E-07	1.51E-03	1.86E-03	4.17E-01	4.38E-01
NCCOCCN	1.23	9.15E-08	1.17E-07	7.28E-04	8.41E-04	4.27E-01	4.36E-01
OCCCN	1.23	1.81E-07	1.88E-07	8.09E-04	1.47E-03	4.61E-01	4.90E-01
SCCOCCCN	1.28	1.79E-07	2.17E-07	9.91E-04	1.61E-03	4.32E-01	4.53E-01
CNC	1.15	4.40E-08	6.25E-08	3.21E-04	4.07E-04	4.00E-01	4.13E-01
NCCNCCN	1.20	4.30E-08	5.71E-08	3.36E-04	3.95E-04	4.20E-01	4.27E-01
OCCOCCN	1.27	2.19E-07	2.70E-07	1.49E-03	2.01E-03	4.33E-01	4.51E-01
OCCOCCO	1.30	2.90E-07	3.50E-07	9.08E-04	2.69E-03	1.49E-01	4.53E-01
SCCNCCN	1.30	6.06E-08	9.35E-08	5.16E-04	6.54E-04	3.85E-01	3.99E-01
OCC(C)N	1.21	1.10E-07	1.58E-07	5.58E-04	1.07E-03	3.26E-01	4.14E-01
C(C)N	1.16	3.92E-08	4.68E-08	5.41E-05	3.12E-04	-1.13E-01	4.55E-01
OCCOCCCN	1.23	1.76E-07	2.15E-07	1.07E-03	1.58E-03	4.24E-01	4.50E-01
COCCCN	1.20	1.08E-07	1.30E-07	6.65E-04	9.23E-04	4.34E-01	4.52E-01
OCCNCCN	1.23	8.65E-08	1.23E-07	6.68E-04	8.41E-04	3.90E-01	4.10E-01
CSCCN	1.33	2.17E-07	1.81E-07	8.02E-04	1.76E-03	4.82E-01	5.43E-01
SCCC	1.28	3.31E-07	2.10E-07	3.73E-05	2.26E-03	2.39E+00	6.12E-01
C(C)OCCN	1.19	6.45E-08	1.16E-07	4.15E-04	6.83E-04	2.86E-01	3.64E-01
OCCCOCCN	1.23	1.92E-07	2.84E-07	1.45E-03	1.92E-03	3.79E-01	4.02E-01
NOCCCNCN	1.27	6.11E-08	6.51E-08	3.26E-04	5.28E-04	4.63E-01	4.86E-01
COCCN	1.25	1.37E-07	1.35E-07	8.96E-04	1.10E-03	4.93E-01	5.02E-01
C(C)SCCN	1.26	1.58E-07	1.40E-07	3.52E-04	1.19E-03	3.26E-01	5.32E-01
NCCCN	1.18	4.84E-08	6.96E-08	3.92E-04	4.56E-04	4.08E-01	4.16E-01
SCCN	1.41	1.16E-07	1.18E-07	7.01E-04	1.07E-03	4.68E-01	4.95E-01
C	0.97	4.80E-07	3.63E-07	0.00E+00	2.67E-03		5.72E-01
OCCOC	1.30	1.94E-07	2.49E-07	1.51E-03	1.86E-03	4.17E-01	4.38E-01
COCC	1.17	2.28E-07	3.42E-07	1.47E-03	2.21E-03	3.34E-01	3.98E-01
SCCOCCN	1.34	1.49E-07	2.29E-07	1.23E-03	1.64E-03	3.80E-01	3.97E-01
CC(N)C	1.16	4.26E-08	3.57E-08	2.28E-04	2.89E-04	5.27E-01	5.41E-01
OCCOCCOC	1.28	1.92E-07	2.56E-07	1.48E-03	1.88E-03	4.03E-01	4.29E-01
CC(=O)	1.47	1.65E-08	1.44E-08	7.07E-05	1.44E-04	3.84E-01	5.32E-01
SC	1.55	2.44E-07	1.20E-07	9.14E-05	1.80E-03	3.06E-01	6.70E-01
OCCOC(O)	1.48	1.18E-07	6.65E-08	7.40E-04	9.01E-04	6.14E-01	6.36E-01

TABLE S6: Average of computed ion transport properties of generated polymers in the $2^n d$ Iteration

Monomer SMILES	Density (gr/cm3)	Li Diffusivity (cm2/s)	TFSI Diffusivity (cm2/s)	Cluster-Nernst-Einstein Conductivity (S/cm)	Nernst-Einstein Conductivity (S/cm)	Cluster-Nernst-Einstein Transference Number	Nernst-Einstein Transference Number
NCCN	1.24	5.58E-08	6.29E-08	4.30E-04	4.87E-04	4.62E-01	4.74E-01
COCCOCCN	1.25	1.46E-07	1.75E-07	9.72E-04	1.32E-03	4.31E-01	4.53E-01
OCCCOC	1.23	1.79E-07	3.04E-07	1.61E-03	1.92E-03	3.61E-01	3.71E-01
SCC	1.37	4.38E-07	2.15E-07	1.79E-04	2.92E-03	5.89E-01	6.68E-01
SCCCOCCN	1.28	1.37E-07	1.80E-07	7.75E-04	1.28E-03	3.97E-01	4.37E-01
ONCCCN	1.29	7.84E-08	6.28E-08	3.61E-04	5.89E-04	5.42E-01	5.58E-01
C(N)CCN	1.26	2.71E-08	2.05E-08	1.36E-04	1.88E-04	5.53E-01	5.72E-01
SCCOC	1.39	3.56E-07	3.87E-07	1.50E-03	3.40E-03	2.98E-01	4.79E-01
OCCCON	1.32	2.00E-07	2.64E-07	1.52E-03	1.96E-03	4.09E-01	4.31E-01
OCCSCCOC	1.33	1.67E-07	2.61E-07	1.06E-03	1.89E-03	3.15E-01	3.89E-01
OCCSCCN	1.33	2.67E-07	2.35E-07	9.63E-04	2.17E-03	4.59E-01	5.33E-01
OOCCOCCN	1.34	2.05E-07	2.62E-07	1.31E-03	2.03E-03	4.00E-01	4.35E-01
OCCCC	1.12	3.53E-07	3.53E-07	5.89E-04	2.60E-03	7.37E-02	4.99E-01
OCCCCON	1.27	1.82E-07	2.00E-07	9.32E-04	1.57E-03	4.32E-01	4.76E-01
C(C)OC	1.11	1.30E-07	1.94E-07	4.33E-04	1.14E-03	1.43E-01	4.05E-01
OCCCNCCN	1.20	8.50E-08	1.10E-07	5.92E-04	7.71E-04	4.13E-01	4.34E-01
CCOCOCCN	1.24	1.31E-07	2.17E-07	1.23E-03	1.41E-03	3.65E-01	3.77E-01
NCCOC	1.25	1.21E-07	1.49E-07	8.97E-04	1.10E-03	4.43E-01	4.51E-01
SCCOCC	1.30	4.23E-07	3.55E-07	8.84E-04	3.22E-03	1.93E-01	5.41E-01
ONCCOCCN	1.31	1.43E-07	1.62E-07	1.03E-03	1.31E-03	4.54E-01	4.69E-01
OCNCN	1.36	2.31E-07	8.94E-08	5.87E-04	1.40E-03	7.34E-01	7.21E-01
CSCCOCCN	1.29	1.49E-07	2.47E-07	1.01E-03	1.61E-03	3.39E-01	3.78E-01
OCCCOCC	1.20	1.63E-07	2.59E-07	1.27E-03	1.65E-03	3.72E-01	3.89E-01
CNCC	1.10	5.23E-08	7.82E-08	2.62E-04	4.58E-04	3.79E-01	4.01E-01
CNCCOCCN	1.21	8.29E-08	1.04E-07	4.72E-04	7.50E-04	4.16E-01	4.47E-01
NCCOCCOC	1.25	1.35E-07	1.80E-07	9.78E-04	1.29E-03	4.17E-01	4.30E-01
SCCSCCN	1.39	1.44E-07	1.32E-07	5.54E-04	1.20E-03	4.61E-01	5.22E-01
ONCOCCOC	1.34	2.06E-07	1.85E-07	1.04E-03	1.71E-03	5.08E-01	5.27E-01
CCNCOCCN	1.21	1.01E-07	1.22E-07	7.17E-04	8.89E-04	4.49E-01	4.58E-01
SCCNCCOC	1.29	9.27E-08	1.55E-07	7.55E-04	1.01E-03	3.45E-01	3.76E-01
OCCOCCSC	1.34	1.91E-07	3.11E-07	1.21E-03	2.23E-03	2.86E-01	3.81E-01

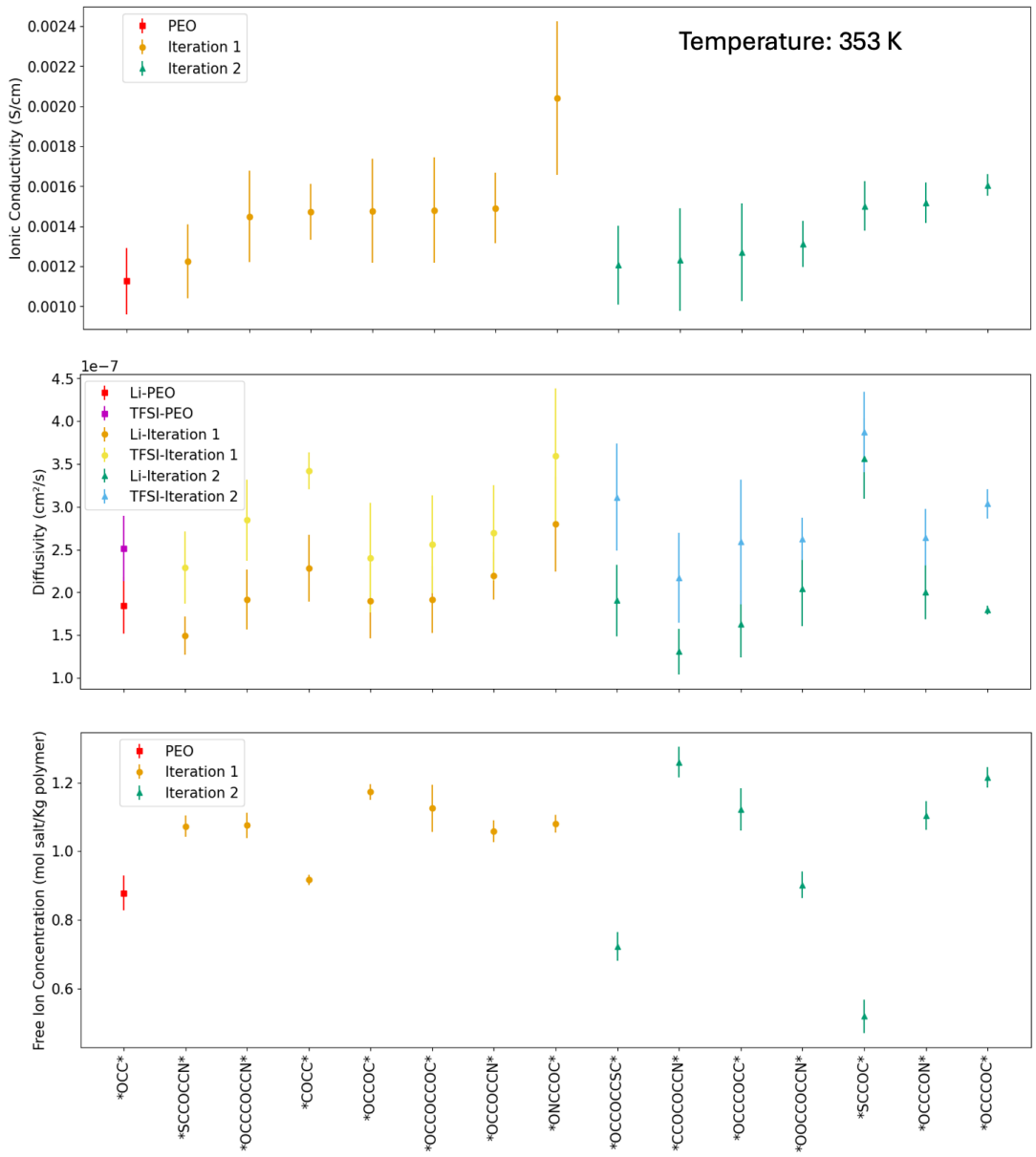

FIG. S1: Comparison of the ionic conductivity, ions' diffusivity, and free ion concentration between PEO and the generated polymers with higher conductivity. The first point in each plot represents the ion transport properties of PEO. All reported data in this graph are computed from molecular dynamics (MD) simulations.

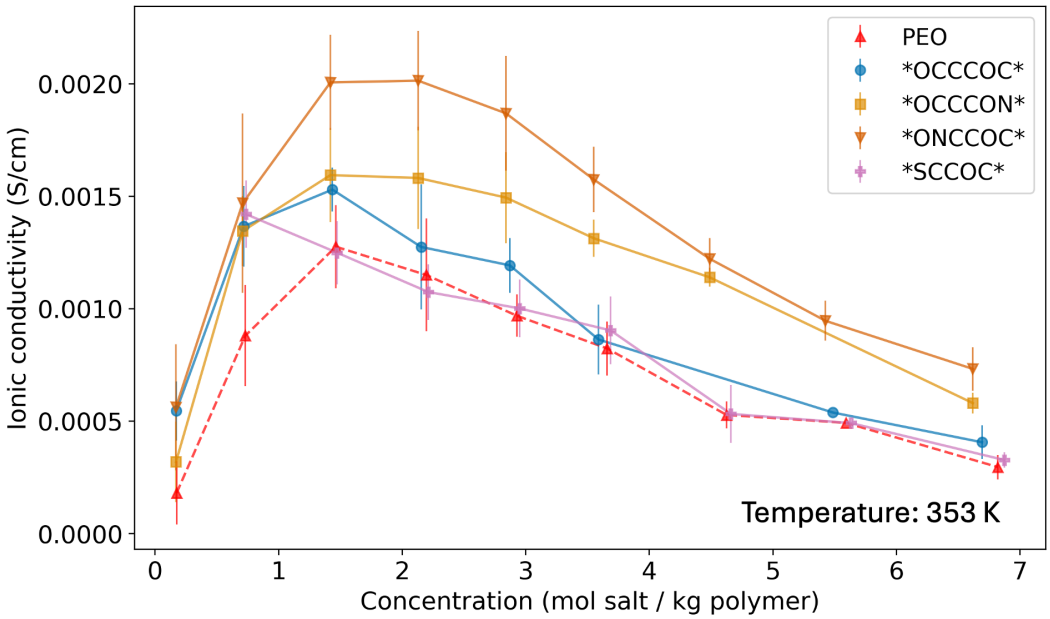

FIG. S2: Effect of salt concentration on ionic conductivity captured in MD simulations. The position of maximum ionic conductivity depends on the polymer structure and occurs at different salt molalities for different polymers.

[1] Zhenze Yang, Weike Ye, Xiangyun Lei, Ha-Kyung Kwon, Daniel Schweigert, and Arash Khajeh. De novo design of polymer electrolytes with high conductivity using gpt-based and diffusion-based generative models. Unpublished manuscript, 2023.