

Supporting Information

Machine Learning-Enabled Design and Prediction of Protein Resistance on Self-Assembled Monolayers and Beyond

Yonglan Liu^{1ζ}, Dong Zhang^{1ζ}, Yijing Tang¹, Yanxian Zhang¹, Yung Chang², and Jie Zheng^{1*}

¹Department of Chemical, Biomolecular, and Corrosion Engineering
The University of Akron, Ohio 44325, USA

²Department of Chemical Engineering, R&D Center for Membrane Technology
Chung Yuan Christian University, Taoyuan, Taiwan

ζ The authors contribute equally to this work.

*Corresponding Author: zhengj@uakron.edu

Table S1. Loadings and communalities for the 10 rotated functional groups from factor analysis for the 43 molecular descriptors.

Molecular Descriptors	Functional Group Number										Communality (H^2)
	1	2	3	4	5	6	7	8	9	10	
Me	0.00	0.87	0.05	0.11	0.19	-0.25	0.17	-0.10	-0.01	0.09	0.92
Mp	0.06	-0.05	0.73	-0.05	0.66	0.04	0.03	0.00	-0.04	0.03	0.98
Mi	-0.05	0.03	-0.90	-0.25	-0.16	-0.17	-0.07	-0.07	0.02	0.11	0.95
GD	-0.21	-0.30	-0.38	-0.70	-0.10	-0.17	-0.03	-0.08	-0.12	-0.13	0.84
nTA	0.72	0.48	0.01	0.06	0.05	0.08	-0.04	0.05	0.41	0.16	0.96
RBN	0.22	0.04	-0.17	0.75	-0.02	0.21	0.32	0.21	0.22	-0.04	0.88
RBF	0.01	-0.07	-0.36	0.55	0.10	0.17	0.47	0.20	0.27	-0.10	0.81
nH	0.35	0.17	0.12	0.80	-0.27	0.26	-0.09	0.12	0.05	0.03	0.98
nN	0.58	-0.09	-0.11	-0.04	0.56	-0.10	-0.22	-0.04	0.17	0.44	0.96
N%	0.02	-0.15	-0.22	-0.47	0.75	-0.15	-0.09	-0.09	-0.02	0.20	0.94
O%	-0.02	0.79	0.03	0.36	-0.38	-0.12	0.19	-0.03	0.01	-0.04	0.94
nCIC	-0.13	0.25	0.88	-0.02	-0.17	-0.02	-0.18	0.00	-0.19	-0.03	0.96
Rperim	-0.17	0.33	0.68	0.17	-0.21	-0.03	-0.32	-0.10	-0.36	-0.05	0.93
ARR	-0.07	-0.03	0.93	-0.09	-0.11	-0.06	0.18	0.03	0.11	-0.03	0.95
nCp	0.34	0.27	0.00	0.06	-0.28	0.19	0.59	0.38	-0.34	0.03	0.93
nCs	-0.11	0.51	0.14	0.02	-0.06	0.75	0.01	0.27	-0.04	-0.12	0.95
nCt	0.76	0.01	-0.02	0.13	0.01	0.06	-0.01	0.56	0.01	-0.04	0.93
nRCONHR	0.02	0.08	-0.07	-0.06	0.02	-0.02	0.17	-0.10	-0.03	0.81	0.72
nRCONR2	0.85	-0.02	-0.02	0.11	0.06	-0.06	-0.18	-0.04	0.13	0.23	0.85
nRCO	0.88	-0.03	-0.02	0.01	-0.02	-0.01	0.11	-0.06	-0.10	-0.18	0.84
nRNH2	-0.19	-0.01	-0.05	0.01	-0.06	-0.01	0.83	0.12	-0.28	0.23	0.88
nArNH2	-0.07	0.71	0.42	-0.05	-0.09	0.17	-0.16	-0.03	0.03	-0.13	0.76
nRNHR	0.19	-0.36	-0.19	0.02	0.08	-0.09	-0.73	-0.10	0.27	-0.16	0.86
nRNR2	-0.12	-0.25	-0.02	-0.01	-0.13	-0.09	-0.50	0.20	-0.17	0.13	0.44
nRCN	-0.12	0.00	0.00	-0.13	0.95	0.03	0.01	-0.02	-0.08	-0.11	0.95
nOHp	-0.11	0.56	0.06	0.05	-0.16	0.08	0.04	0.73	-0.12	-0.13	0.94
nROR	-0.28	0.15	-0.02	0.83	-0.16	-0.09	0.05	-0.18	-0.19	-0.17	0.92
C-001	0.88	-0.12	-0.04	-0.03	-0.05	0.26	0.07	-0.13	-0.17	-0.05	0.92
C-002	-0.04	-0.18	-0.05	-0.03	0.07	0.84	0.16	0.42	-0.10	-0.03	0.96
C-003	0.92	0.00	0.01	0.16	0.02	0.00	-0.06	0.03	0.03	0.00	0.88
C-005	0.43	-0.19	-0.13	0.18	-0.10	-0.07	-0.25	-0.12	0.72	0.08	0.89
C-006	-0.05	-0.10	-0.09	0.82	-0.13	-0.22	-0.03	-0.01	-0.39	0.01	0.91
C-041	-0.07	-0.07	-0.04	-0.03	0.00	-0.04	-0.10	-0.03	0.56	-0.02	0.34
H-046	0.33	-0.19	-0.01	0.07	-0.08	0.86	0.07	-0.06	-0.03	0.07	0.91
H-047	0.21	0.01	0.09	0.86	-0.24	-0.21	-0.22	-0.10	0.17	0.03	0.98
H-051	0.33	-0.11	-0.07	-0.29	0.37	0.03	0.19	-0.23	-0.33	-0.26	0.61
H-052	0.00	-0.03	-0.05	0.01	-0.03	0.25	0.09	0.89	-0.03	-0.04	0.86
O-060	-0.07	0.29	0.48	-0.09	-0.15	-0.01	0.15	-0.06	0.43	-0.16	0.58
Ui	0.40	0.14	0.60	-0.21	0.56	-0.04	0.05	-0.10	0.16	0.16	0.97

Hy	-0.18	0.80	-0.01	-0.24	-0.19	0.03	0.27	0.23	-0.16	0.10	0.93
TPSA(NO)	0.18	0.89	0.16	0.25	0.13	0.04	0.13	0.12	-0.07	0.11	0.98
ALOGP	-0.10	-0.68	0.42	-0.15	0.01	0.49	0.02	0.04	0.14	-0.17	0.97
PDI	0.10	-0.63	0.67	0.08	0.08	-0.03	-0.04	-0.11	-0.07	0.22	0.94

Table S2. Regression coefficients of molecular descriptors with absolute coefficient greater than 0.3 for the 10 functional group.

	Functional Group 1	Functional Group 2	Functional Group 3	Functional Group 4		Functional Group 5	
C-003	0.92	nTA	0.48	nArNH2	0.42	O%	0.36
nRCO	0.88	TPSA(NO)	0.89	ALOGP	0.42	H-047	0.86
C-001	0.88	Me	0.87	ARR	0.93	nROR	0.83
nRCONR2	0.85	Hy	0.80	Mi	-0.90	C-006	0.82
nCt	0.76	O%	0.79	nCIC	0.88	nH	0.80
nTA	0.72	nArNH2	0.71	Mp	0.73	RBN	0.75
nN	0.58	ALOGP	-0.68	Rperim	0.68	GD	-0.70
Ui	0.40	Rperim	0.33	PDI	0.67	RBF	0.55
nH	0.35	PDI	-0.63	Ui	0.60	N%	-0.47
H-051	0.33	nCs	0.51	O-060	0.48		
H-046	0.33	nRNHR	-0.36	GD	-0.38		
nCp	0.34	nOHp	0.56	RBF	-0.36		
C-005	0.43						
	Functional Group 6	Functional Group 7	Functional Group 8	Functional Group 9		Functional Group 10	
ALOGP	0.49	Rperim	-0.32	nCt	0.56	nTA	0.41
H-046	0.86	RBN	0.32	C-002	0.42	Rperim	-0.36
C-002	0.84	RBF	0.47	nCp	0.38	O-060	0.43
nCs	0.75	nRNH2	0.83	H-052	0.89	C-006	-0.39
		nRNHR	-0.73	nOHp	0.73	H-051	-0.33
		nCp	0.59			nCp	-0.34
		nRNR2	-0.50			C-005	0.72
						C-041	0.56

Table S3. Summary of percentage of variance of 43 molecular descriptors included in the top 10 functional groups.

Functional Group Number	Percentage of Variance (%)	Cumulative Percentage of Variance (%)
1	13.67	13.67
2	13.58	27.25
3	12.49	39.74
4	11.53	51.27
5	7.95	59.22
6	6.77	65.99
7	6.68	72.67
8	5.77	78.44
9	5.44	83.88
10	3.46	87.34

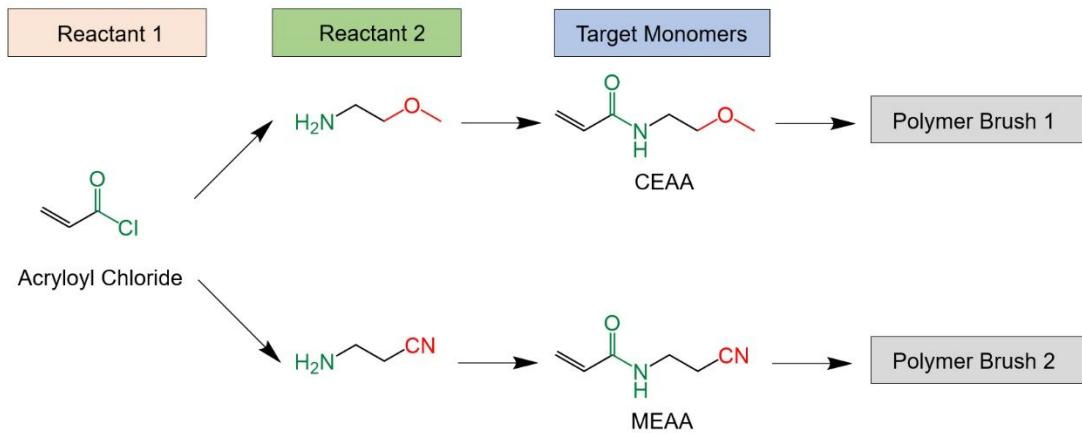


Figure S1. Synthesis procedure of two monomers (CEAA, and MEAA) of designed antifouling materials, which are further used to synthesize polymer brushes.