Electronic Supplementary Material (ESI) for Polymer Chemistry.

A Predictive Machine-Learning Model for Kinetic Rate Coefficients in Radical Polymerization

Emma Van de Reydt,[a] Noam Marom,[b] James Saunderson,[b] Mario Boley[c] and Tanja Junkers[a,*]

[a.] *Polymer Reaction Design Group, School of Chemistry, Monash University 19 Rainforest Walk, Building 23, Clayton, Vic 3800 (Australia) E-mail: tanja.junkers@monash.edu.*

[b.] *Electrical and Computer Systems Engineering, Monash University, 14 Alliance Lane, Building 72, Clayton, Vic 3800 (Australia).*

[c.] *Faculty for Information Technology, Monash University, 20 Exhibition Walk, Clayton, Vic 3800 (Australia)*

Training data overview

Table S1: Alphabetic list of monomers considered with their abbreviation and SMILES notation.

Monomer name	Abbreviation	SMILES
1,3-Butadiene	**BuDE**	C=CC=C
2-(Hexylcarbamoyloxy)ethyl acrylate	**HCEA**	C=CC(OCCOC(NCCCCCC)=O)=O
2-(Hexylcarbamoyloxy)isopropyl acrylate	**HCPA**	C=C(OC(NCCCCCC)=O)C(OC(C)C)=O
2-(Phenylcarbamoyloxy)ethyl acrylate	**PhCEA**	C=CC(OCCOC(NC1=CC=CC=C1)=O)=O
2-(Phenylcarbamoyloxy)isopropyl acrylate	**PhCPA**	C=C(OC(NC1=CC=CC=C1)=O)C(OC(C)C)=O
2-ethylhexyl acrylate	**EHA**	C=CC(OC[C@H](CC)CCCC)=O
2-ethylhexyl methacrylate	**EHMA**	CC(C(OC[C@H](CC)CCCC)=O)=C
2-hydroxypropyl methacrylate	**HPMA**	CC(C(OC[C@H](O)C)=O)=C
2-propylheptyl acrylate	**PHA**	C=CC(OC[C@H](CCC)CCCCC)=O
Acrylonitrile	**CAN**	C=CC#N
Behenyl acrylate	**BeA**	C=CC(OCCCCCCCCCCCCCCCCCCCCCC)=O
Behenyl methacrylate	**BeMA**	CC(C(OCCCCCCCCCCCCCCCCCCCCCC)=O)=C
Benzyl acrylate	**BnA**	C=CC(OCC1=CC=CC=C1)=O
Benzyl methacrylate	**BzMA**	CC(C(OCC1=CC=CC=C1)=O)=C
Butyl methacrylate	**BMA**	CC(C(OCCCC)=O)=C
Cyclohexyl methacrylate	**CHMA**	CC(C(OC1CCCCC1)=O)=C
Dodecyl methacrylate	**DMA**	CC(C(OCCCCCCCCCCCC)=O)=C
Ethoxyethyl acrylate	**EEA**	C=CC(OCCOCC)=O
Ethyl methacrylate	**EMA**	CC(C(OCC)=O)=C
Glycidyl methacrylate	**GMA**	CC(C(OC[C@@H]1CO1)=O)=C
Henicosyl acrylate	**C21A**	C=CC(OCCCCCCCCCCCCCCCCCCCCC)=O
Heptadecyl acrylate	**C17A**	C=CC(OCCCCCCCCCCCCCCCCC)=O
Hydroxyethyl methacrylate	**HEMA**	CC(C(OCCO)=O)=C
Isobornyl acrylate	**iBoa**	C=CC(O[C@H]1C[C@@H]2CC[C@@]1(C)C2(C)C)=O
iso-bornyl methacrylate	**iBoMA**	CC(C(O[C@H]1C[C@@H]2CC[C@@]1(C)C2(C)C)=O)=C
iso-butyl methacrylate	**iBMA**	CC(C(OCC(C)C)=O)=C
iso-decyl methacrylate	**iDeMA**	CC(C(OCCCCCCCC(C)C)=O)=C
iso-nonyl acrylate	**INA-A**	C=CC(OCCCCCCC(C)C)=O
Methacrylic acid	**MAA**	CC(C(O)=O)=C
Methyl acrylate	**MA**	C=CC(OC)=O
Methyl methacrylate	**MMA**	CC(C(OC)=O)=C
n-Butyl acrylate	**BA**	C=CC(OCCCC)=O
n-Pentyl Methacrylate	**PnMA**	CC(C(OCCCCC)=O)=C
N-vinyl formamide	**NVF**	O=CNC=C
N-Vinyl Pyrrolidone	**NVP**	O=C1N(C=C)CCC1
Propylheptyl methacrylate	**PHMA**	CC(C(O[C@@H](CCC)CCCCC)=O)=C
Stearyl acrylate	**SA**	C=CC(OCCCCCCCCCCCCCCCCCC)=O
Stearyl methacrylate	**SMA**	CC(C(OCCCCCCCCCCCCCCCCCC)=O)=C
Styrene	**Sty**	C=CC1=CC=CC=C1
tert-butyl acrylate	**tBA**	C=CC(OC(C)(C)C)=O
Vinyl acetate	**VAc**	CC(OC=C)=O

Table S2: List of monomers classified by type (acrylates, H-bonding monomers, methacrylates and other) with the activation energy E_a and pre-exponential factor A. Using the Arrhenius equation, the natural logarithm of the rate constant $\ln(k_p)$ was calculated for four different temperatures.

	Monomer	E_a / J·mol^{-1}	A / s^{-1}	$\ln(k_p$ / L·mol^{-1}·s^{-1}) T = 25˚C	$\ln(k_p$ / L·mol^{-1}·s^{-1}) T = 50˚C	$\ln(k_p$ / L·mol^{-1}·s^{-1}) T = 75˚C	$\ln(k_p$ / L·mol^{-1}·s^{-1}) T = 100˚C	Source
Acrylates	**BA**	17.90	22100000	9.69	10.25	10.73	11.14	[1, 2]
	BeA	13.02	5350000	10.24	10.65	10.99	11.30	[2]
	BnA	16.12	12800000	9.86	10.36	10.80	11.17	[3]
	C17A	14.66	8150000	10.00	10.46	10.85	11.19	[2]
	C21A	12.99	3220000	9.74	10.15	10.50	10.80	[2]
	EEA	13.80	6300000	10.09	10.52	10.89	11.21	[2]
	EHA	15.80	9100000	9.65	10.14	10.57	10.93	[2]
	iBoA	15.35	4810000	9.19	9.67	10.08	10.44	[3]
	INA-A	16.54	13500000	9.75	10.26	10.70	11.09	[2]
	MA	17.30	14100000	9.48	10.02	10.48	10.89	[4]
	PHA	16.41	10500000	9.55	10.06	10.50	10.88	[2]
	SA	16.93	18600000	9.91	10.44	10.89	11.28	[2]
	***t*BA**	18.90	22100000	9.69	10.25	10.73	11.14	[5]
H-Bonding monomers	**HCEA**	13.30	6600000	10.34	10.75	11.11	11.42	[6]
	HCPA	14.10	6600000	10.01	10.45	10.83	11.16	[6]
	HEMA	21.90	8880000	7.16	7.85	8.43	8.94	[7]
	HPMA	20.80	3510000	6.68	7.33	7.88	8.37	[7]
	PhCEA	14.30	12000000	10.53	10.98	11.36	11.69	[6]
	PhCPA	14.20	4900000	9.68	10.12	10.50	10.83	[6]
Methacrylates	**BeMA**	20.52	2510000	6.46	7.10	7.65	8.12	[8]
	BMA	22.90	3801894	5.91	6.63	7.24	7.77	[9]
	BzMA	22.90	6760830	6.49	7.20	7.82	8.35	[10]
	CHMA	23.00	6309573	6.38	7.10	7.71	8.24	[10]
	DMA	21.00	2511886	6.27	6.92	7.48	7.97	[9]
	EHMA	21.60	2390000	5.97	6.65	7.22	7.72	[7]
	EMA	23.40	4073803	5.78	6.51	7.14	7.68	[9]
	GMA	22.90	5011872	6.19	6.90	7.52	8.05	[10]
	iBMA	21.80	2640000	5.99	6.67	7.25	7.76	[7]
	iBoMA	23.10	6165950	6.32	7.04	7.65	8.19	[10]
	iDeMA	21.60	2390000	5.97	6.65	7.22	7.72	[7]
	MMA	22.36	2673006	5.78	6.48	7.07	7.59	[11]
	PHMA	21.72	2830000	6.09	6.77	7.35	7.85	[8]
	PnMA	23.80	6000000	6.01	6.75	7.38	7.94	[12]
	SMA	21.49	3450000	6.39	7.06	7.63	8.13	[8]
Other	**BuDE**	35.70	80500000	3.81	4.91	5.87	6.70	[7]
	CAN	15.40	1790000	8.19	8.67	9.08	9.43	[13]
	MAA	16.10	380000	6.35	6.86	7.29	7.66	[7]
	NVF	19.50	6400000	7.81	8.41	8.93	9.39	[14]
	NVP	17.60	25700000	9.96	10.51	10.98	11.39	[15]
	Sty	32.51	42657952	4.45	5.47	6.34	7.09	[16]
	VAc	20.40	13500000	8.19	8.83	9.37	9.84	[17]

Table S3: List with monomer parameters classified by type (acrylates, H-bonding monomers, methacrylates and other). The molecular weight *MW*, dipole moment *DP*, boiling point *BP* and the Gibbs free energy *GFE* were predicted with chemdraw. The other parameters were researched from literature. The A_value is the substituent effect on the C2 carbon of the polymerising double bond, the R_value is the substituent effect on the moiety chain. H_acc and H_don are the H-bonding effects, A1* and A2* are the dissociation constants. Colour highlights indicate which features were used in the final predictions of k_p.

	Monomer	MW / g·mol^{-1}	DP	BP / K	GFE / J	A_value [18]	R_value [18]	H_acc [19]	H_don [19]	A1* [20]	A2*
Acrylates	**BA**	128.17	2.27	397	-273.05	0.00	11.28	0	-0.5	-2.8	1
	BeA	380.60	2.28	691	-124.33	0.00	12.67	0	-0.5	-2.8	1
	BnA	162.18	2.40	506	-138.22	0.00	13.08	-2.6	0	-2.8	1
	C17A	310.50	3.17	633	-166.43	0.00	12.67	0	-0.5	-2.8	1
	C21A	366.60	2.76	679	-132.75	0.00	12.71	0	-0.5	-2.8	1
	EEA	144.17	3.91	442	-380.89	0.00	13.43	-5.1	0	-2.8	1
	EHA	184.27	2.96	490	-244.65	0.00	12.46	0	-0.5	-2.8	1
	iBoA	208.30	2.37	516	-142.37	0.00	12.42	0	-0.5	-2.8	1
	INA-A	198.30	2.50	509	-236.23	0.00	12.71	0	-0.5	-2.8	1
	MA	86.09	2.23	349	-301.15	0.00	12.71	0	-0.5	-2.8	1
	PHA	212.33	2.50	527	-227.81	0.00	12.44	0	-0.5	-2.8	1
	SA	324.50	2.67	644	-158.01	0.00	12.71	0	-0.5	-2.8	1
	***t*BA**	128.17	2.27	397	-273.05	0.00	11.28	0	-0.5	-2.8	1
H-Bonding monomers	**HCEA**	243.30	1.74	585	-454.77	0.00	13.79	-7.7	-4.1	-2.8	1
	HCPA	257.33	4.94	587	-457.34	1.91	12.15	-16.8	-4.1	-2.8	1
	HEMA	130.14	3.68	472	-429.68	-1.55	13.23	-5.5	-3.5	-2.8	647
	HPMA	144.17	3.43	473	-423.70	-1.55	13.00	-5.5	-3.5	-2.8	647
	PhCEA	235.24	1.07	602	-342.36	0.00	13.70	-11	-3.7	-2.8	1
	PhCPA	249.26	4.29	603	-344.93	2.51	12.15	-13.1	-3.7	-2.8	1
Methacrylates	**BeMA**	394.70	1.65	699	-124.46	-1.55	12.67	0	-0.5	-2.8	647
	BMA	142.20	1.71	436	-276.02	-1.55	13.24	0	-0.5	-2.8	647
	BzMA	176.21	1.71	519	-138.35	-1.55	13.08	-2.6	0	-2.8	647
	CHMA	168.23	1.27	486	-234.73	-1.55	12.29	0	-0.5	-2.8	647
	DMA	254.41	1.50	579	-208.66	-1.55	12.71	0	-0.5	-2.8	647
	EHMA	198.30	1.59	504	-244.78	-1.55	12.46	0	-0.5	-2.8	647
	EMA	114.12	1.70	391	-292.86	-1.55	12.58	0	-0.5	-2.8	647
	GMA	142.15	4.10	445	-309.81	-1.55	13.12	-5.1	0	-2.8	647
	iBMA	142.20	1.72	423	-278.46	-1.55	12.50	0	-0.5	-2.8	647
	iBoMA	222.32	1.85	529	-142.50	-1.55	12.42	0	-0.5	-2.8	647
	iDeMA	226.35	1.61	539	-227.94	-1.55	12.67	0	-0.5	-2.8	647
	MMA	100.12	1.64	367	-301.28	-1.55	12.71	0	-0.5	-2.8	647
	PHMA	226.35	1.51	539	-227.94	-1.55	12.09	0	-0.5	-2.8	647
	PnMA	156.2	1.65	457	-267.60	-1.55	12.75	0	-0.5	-2.8	647
	SMA	338.6	1.96	653	-158.14	-1.55	12.71	0	-0.5	-2.8	647
Other	**BuDE**	54.09	0.00	289	97.06	0.00	0.00	0	0	0	-
	CAN	53.06	3.94	368	133.98	0.00	14.33	-4.7	-4.7	-3.4	10
	MAA	86.09	1.78	433	-343.24	-1.55	12.36	0	-5	-2.8	-
	NVF	71.08	3.98	377	-9.33	-1.55	3.48	-6.9	-1.6	-3.4	-
	NVP	111.14	3.95	493	70.61	0.00	-0.33	-5.5	-4.1	-3.4	-
	Sty	104.15	0.26	420	155.31	0.00	1.76	-2.6	0	-8.4	15
	VAc	86.09	1.54	349	-301.15	0.00	3.90	-2.9	0	-5.9	-

Table S4: List with monomer parameters from Chemspider, generated using the ACD/Labs Percepta Platform – PhysChem Module. The selected parameters are the boiling point *BP*, the refraction index *RI*, the density ρ, the partition coefficient ACDlogP, the polarisation, the surface tension and the vapor pressure.

	Monomer	BP / °C	RI	ρ / g·mL^{-1}	ACDlogP	Polarisation	Surface tension	Vapor pressure
Acrylates	**BA**	145.0	1.418	0.9	2.39	14.3	26.7	4.8
	BeA	455.7	1.455	0.9	11.95	47.3	31.3	0
	BnA	228.7	1.517	1.1	2.27	18.5	36.1	0.1
	C17A	385.6	1.451	0.9	9.3	38.1	30.8	0
	C21A	442.3	1.455	0.9	11.42	45.5	31.2	0
	EEA	174.6	1.420	1	1.02	15	28.4	1.2
	EHA	216.0	1.434	0.9	4.33	21.6	28	0.1
	iBoA	244.5	1.491	1	4.22	23.7	33	0
	INA-A	247.3	1.437	0.9	4.86	23.4	28.4	0
	MA	80.2	1.390	0.9	0.79	8.8	23.5	86.3
	PHA	266.1	1.439	0.9	5.39	25.3	28.7	0
	SA	400.2	1.452	0.9	9.83	40	30.9	0
	***t*BA**	133.0	1.418	0.9	2.02	14.3	25.1	8.6
H-Bonding monomers	**HCEA**	357.3	1.456	1	3.43	25.6	34.6	0
	HCPA	366.2	1.456	1	3.78	27.4	33.8	0
	HEMA	189.0	1.443	1.1	0.5	13	33.7	0.2
	HPMA	189.0	1.458	1.1	1.39	14	36.1	0.6
	PhCEA	218.8	1.444	1	0.85	14.8	32.3	0
	PhCPA	322.2	1.557	1.2	2.65	24.7	45.9	0
Methacrylates	**BeMA**	331.7	1.549	1.2	3	26.5	43.7	0
	BMA	468.7	1.455	0.9	12.5	49.1	31	0
	BzMA	160.0	1.423	0.9	2.94	16	26.4	2.4
	CHMA	247.0	1.512	1	2.82	20.2	34.8	1
	DMA	210.0	1.459	1	3.4	18.9	30.5	0.2
	EHMA	322.7	1.445	0.9	7.19	30.7	29.6	0
	EMA	234.8	1.436	0.9	4.88	23.4	27.7	0.1
	GMA	120.5	1.410	0.9	1.88	12.4	24.7	15.2
	iBMA	155.0	1.421	0.9	2.76	16	25.4	3.1
	iBoMA	263.1	1.488	1	4.77	25.4	32.4	0
	iDeMA	283.9	1.441	0.9	5.94	27	28.4	0
	MMA	100.3	1.400	0.9	1.35	10.5	23.5	36.9
	PHMA	-	-	-	-	-	-	-
	PnMA	191.3	1.427	0.9	3.47	17.9	27	0.5
	SMA	414.3	1.452	0.9	10.38	41.8	30.6	0
Other	**BuDE**	-4.4	1.389	0.6	1.86	7.9	15.9	2101
	CAN	77.3	1.385	0.8	0.19	6.2	25	97.1
	MAA	160.5	1.430	1	0.83	8.7	30.9	1.2
	NVF	184.3	1.406	0.9	-0.18	7.7	25	0.7
	NVP	217.6	1.593	1.1	0.37	13.1	52.4	0.1
	Sty	145.2	1.558	0.9	2.7	14.7	31	6.2
	Vac	72.5	1.390	0.9	0.73	8.8	23.5	118.5

Table S5: Results of figure 3Czc

	Monomer	Abbreviation	Experimental $\ln(k_p$ / L·mol^{-1}·s$^{-1})$ T = 25°C	Predicted $\ln(k_p$ / L·mol^{-1}·s$^{-1})$ T = 25°C	Relative error [experimental – predicted] $\ln(k_p$ / L·mol^{-1}·s$^{-1})$ T = 25°C
Acrylates	n-Butyl acrylate	**BA**	9.689923	9.458881	0.231
	Stearyl acrylate	**SA**	9.908823	9.842755	0.066
	2-Propylheptyl acrylate	**PHA**	9.546813	9.659156	-0.11
	Methyl acrylate	**MA**	9.482579	9.381649	0.101
	iso-Nonyl acrylate	**INA-A**	9.745663	9.606882	0.139
	Isobornyl acrylate	**iBoA**	9.193703	9.757866	-0.56
	tert-Butyl acrylate	***t*BA**	9.689923	9.630068	0.06
	Ethoxyethyl acrylate	**EEA**	10.088888	9.873488	0.215
	Henicosyl acrylate	**C21A**	9.744492	9.925102	-0.18
	Heptadecyl acrylate	**C17A**	9.999434	9.827086	0.172
	Benzyl acrylate	**BnA**	9.861884	9.677131	0.185
	Behenyl acrylate	**BeA**	10.240103	9.960121	0.28
	2-Ethylhexyl acrylate	**EHA**	9.649756	9.605180	0.045
H-Bonding monomers	2-(Phenylcarbamoyloxy)isopropyl acrylate	**PhCPA**	9.676210	9.593760	0.082
	2-(Phenylcarbamoyloxy)ethyl acrylate	**PhCEA**	10.531536	10.621373	-0.09
	2-hydroxypropyl methacrylate	**HPMA**	6.679599	6.503095	0.177
	2-(Hexylcarbamoyloxy)isopropyl acrylate	**HCPA**	10.014403	10.193138	-0.18
	2-(Hexylcarbamoyloxy)ethyl acrylate	**HCEA**	10.337119	10.430283	-0.09
	Hydroxyethyl methacrylate	**HEMA**	7.164720	6.324925	0.84
Methacrylates	iso-bornyl methacrylate	**iBoMA**	6.315358	6.401280	-0.09
	Stearyl methacrylate	**SMA**	6.385194	6.046577	0.339
	n-Pentyl methacrylate	**PnMA**	6.006353	5.948854	0.057
	Methyl methacrylate	**MMA**	5.777652	6.199693	-0.42
	iso-decyl methacrylate	**iDeMA**	5.973810	6.055019	-0.08
	iso-butyl methacrylate	**iBMA**	5.991465	5.969650	0.022
	Butyl methacrylate	**BMA**	5.913503	5.983812	-0.07
	Ethyl methacrylate	**EMA**	5.780744	6.162680	-0.38
	2-ethylhexyl methacrylate	**EHMA**	5.973810	6.246495	-0.27
	Dodecyl methacrylate	**DMA**	6.265301	6.178625	0.087
	Cyclohexyl methacrylate	**CHMA**	6.378426	6.244365	0.134
	Benzyl methacrylate	**BzMA**	6.487684	6.518645	-0.03
	Behenyl methacrylate	**BeMA**	6.458338	6.504005	-0.05
	Glycidyl methacrylate	**GMA**	6.188264	6.324925	-0.14
Other	Styrene	**Sty**	4.454347	4.599294	-0.14
	Acrylonitrile	**CAN**	8.185071	8.590136	-0.41

Table S6: List with monomer collected experimental parameters from Chemspider, including the boiling point *BP*, refraction index *RI* and the density ρ. The A2* dissociation constant parameter is included as well. Monomers that had a value for all of four categories were highlighted in bold and coloured according to their coloration used Figure 4.

	Monomer	BP / ˚C	RI	ρ / g·mL^{-1}	A2*
Acrylates	**BA**	**145**	**1.418**	**0.895**	**1**
	BeA	-	-	0	1
	BnA	111	-	1.06	1
	EEA	174	-	-	1
	iBoa	119	1.476	-	1
	MA	**80.5**	**1.403**	**0.956**	**1**
	***t*BA**	**121**	**1.411**	**0.883**	**1**
H-bonding monomers	**HEMA**	**250**	**1.453**	**1.071**	**647**
Methacrylates	**BzMA**	**162**	**1.424**	**0.894**	**647**
	CHMA	**231**	**1.514**	**1.04**	**647**
	DMA	**72**	**1.4577**	**0.97**	**647**
	EHMA	**142**	**1.445**	**0.873**	**647**
	GMA	**118**	**1.413**	**0.917**	**647**
	iBMA	**155**	**1.42**	**0.889**	**647**
	MMA	**100**	**1.414**	**0.939**	**647**
	SMA	-	-	0.86	647
Other	**BuDE**	-4.5	1.4292	0.62	**-**
	CAN	**77**	**1.391**	**0.806**	**10**
	MAA	163	1.431	1.015	**-**
	NVF	210	1.494	1.014	-
	NVP	92	1.512	1.044	-
	Sty	**145**	**1.546**	**0.906**	**15**
	VAc	72	1.395	0.932	**-**

Table S7: Coefficients of the model with acrylates and methacrylates (Standart deviation of feature, ridge regression fit coefficient and relative importance)

Parameters	Standard Deviation	Coefficients	Importance
Polarizability	808.4313	0.00268	2.166596
Dissociation constant 2	328.9271	-0.00555	1.825545
Molecular weight	89.20704	0.00149	0.132918
Effect of H-donor	1.477476	-0.06087	0.089934
Inductive effect of tailgroup	0.384931	0.00162	0.000624
Effect of H-acceptor	0.181007	0.01008	0.001825
Inductive effect	0.790796	0.00001	7.91E-06
Dissociation constant 1	9.05E-16	0	0

Table S8: Coefficients of the model with acrylates, methacrylates and H-bonding monomers (Standart deviation of feature, ridge regression fit coefficient and relative importance)

Parameters	Standard Deviation	Coefficients	Importance
Dissociation constant 2	327.8574	-0.00636	2.085173
Polarizability	10.96917	0.10146	1.112932
Molecular weight	83.51009	-0.01179	0.984584
Effect of H-acceptor	4.292892	-0.106	0.455047
Inductive effect	1.058983	-0.37847	0.400793
Effect of H-donor	1.323033	-0.07107	0.094028
Inductive effect of tailgroup	0.468596	0.08725	0.040885
Dissociation constant 1	1.35E-15	0	0

Table S9: Coefficients of the model with acrylates, methacrylates and others (Standart deviation of feature, ridge regression fit coefficient and relative importance)

Parameters	Standard Deviation	Coefficients	Importance
Dissociation constant 2	327.7552	-0.00526	1.723992
Molecular weight	92.44408	0.01385	1.280351
Polarizability	11.97967	-0.08162	0.977781
Inductive effect of tailgroup	2.09159	0.17573	0.367555
Dissociation constant 1	1.041882	0.10051	0.10472
Effect of H-acceptor	1.65423	0.02534	0.041918
Effect of H-donor	0.819197	0.02161	0.017703
Inductive effect	0.789821	-0.00071	0.000561

Table S10: Coefficients of the model with acrylates, methacrylates, H-bonding monomers and others (Standart deviation of feature, ridge regression fit coefficient and relative importance)

Parameters	Standard Deviation	Coefficients	Importance
Dissociation constant 2	325.9089	-0.00676	2.203144
Molecular weight	86.79396	0.01508	1.308853
Polarizability	11.15569	-0.10046	1.120701
Dissociation constant 1	0.94902	1.09465	1.038845
Inductive effect	1.037659	-0.57508	0.596737
Inductive effect of tailgroup	1.934587	-0.13211	0.255578
Effect of H-donor	1.438212	-0.06727	0.096749
Effect of H-acceptor	4.185084	-0.00045	0.001883

Regression analysis

Each experiment, where possible, was done trying three different methods: linear regression, cross validated Ridge regression and Lasso (Least Absolute Shrinkage and Selection Operator) regression. The term linear regression and multivariant linear regression is in here used interchangeably since multiple parameters were combined. Simply said, multivariate linear regression is a combination of several linear regressions on independent variables for one dependant variable. They are the most straightforward and least complex regressions. A linear model can be written as:

$$y_{ii} = \beta_0 + \beta_{ii1}XX_{ii1} + \beta_{ii2}XX_{ii2} + \cdots + \beta_{iiii}XX_{iiii} + \varepsilon\varepsilon_{ii}, \qquad \text{for for i = 1 , … , n and j = 1 , … , p} \qquad (2)$$

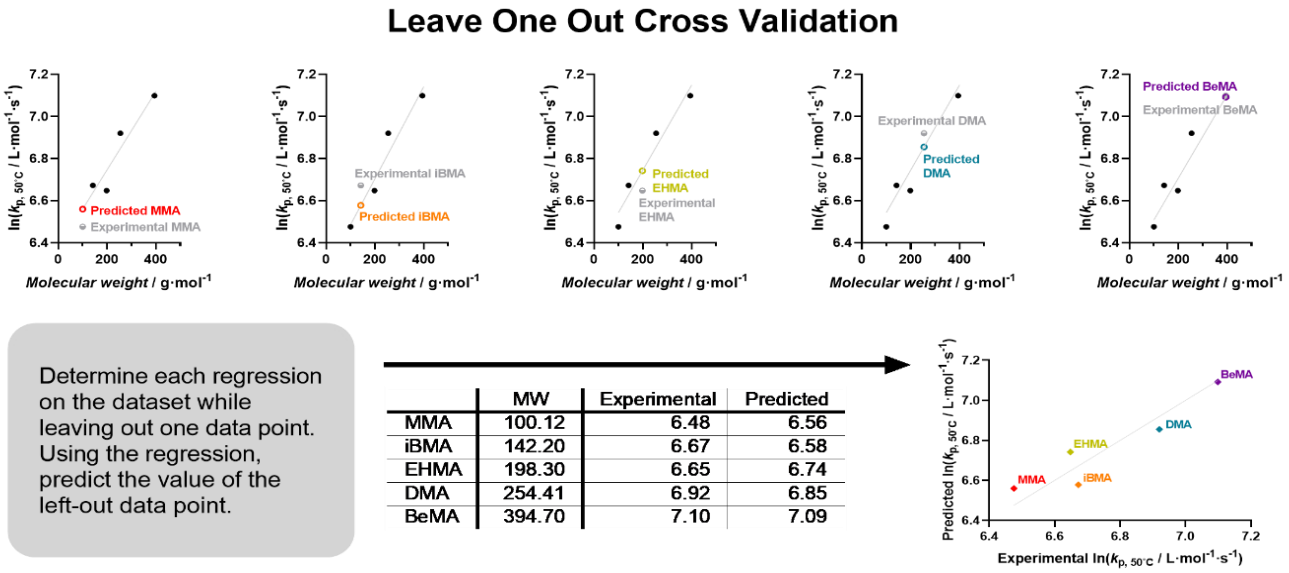
The linear regression used in this project utilises ordinary least squares to estimate the values of β_i, thereby minimising the sum of the squared residuals. However, the simplicity of the regression can result in over-generalisation and over-fitting. Both Ridge and Lasso regression are simple techniques to reduce model complexity.

Ridge regression is similar to linear regression but adds a penalty, also known as the penalty term λ, equivalent to the square of the magnitude of the coefficients. Thus, the coefficients are smaller, reducing the overall complexity. [21]

$$\min(\sum_{ii=1}^{M}(y_{ii} - \hat{y}_{ii})^2 + \lambda \sum_{ii=0}^{p} w_{ii}^2) = \min(\sum_{ii=1}^{M}(y_{ii} - \sum_{ii=0}^{p} w_{ii} * x_{iiii})^2 + \lambda \sum_{ii=0}^{p} w_{ii}^2), \qquad (3)$$

Lasso regression is like Ridge regression; however, the cost function instead minimises the absolute sum of coefficients. Thus, in contrast to Ridge regression, Lasso leaves out complete parameters, which is useful for feature selection.

$$\min(\sum_{ii=1}^{M}(y_{ii} - \hat{y}_{ii})^2 + \lambda \sum_{ii=0}^{p} w_{ii}^2) = \min(\sum_{ii=1}^{M}(y_{ii} - \sum_{ii=0}^{p} w_{ii} * x_{iiii})^2 + \lambda \sum_{ii=0}^{p} |w_{ii}|), \qquad (4)$$



Supplemental regression analysis plots

All K_p and ln(K_p) data [Molecular weight, a_value]

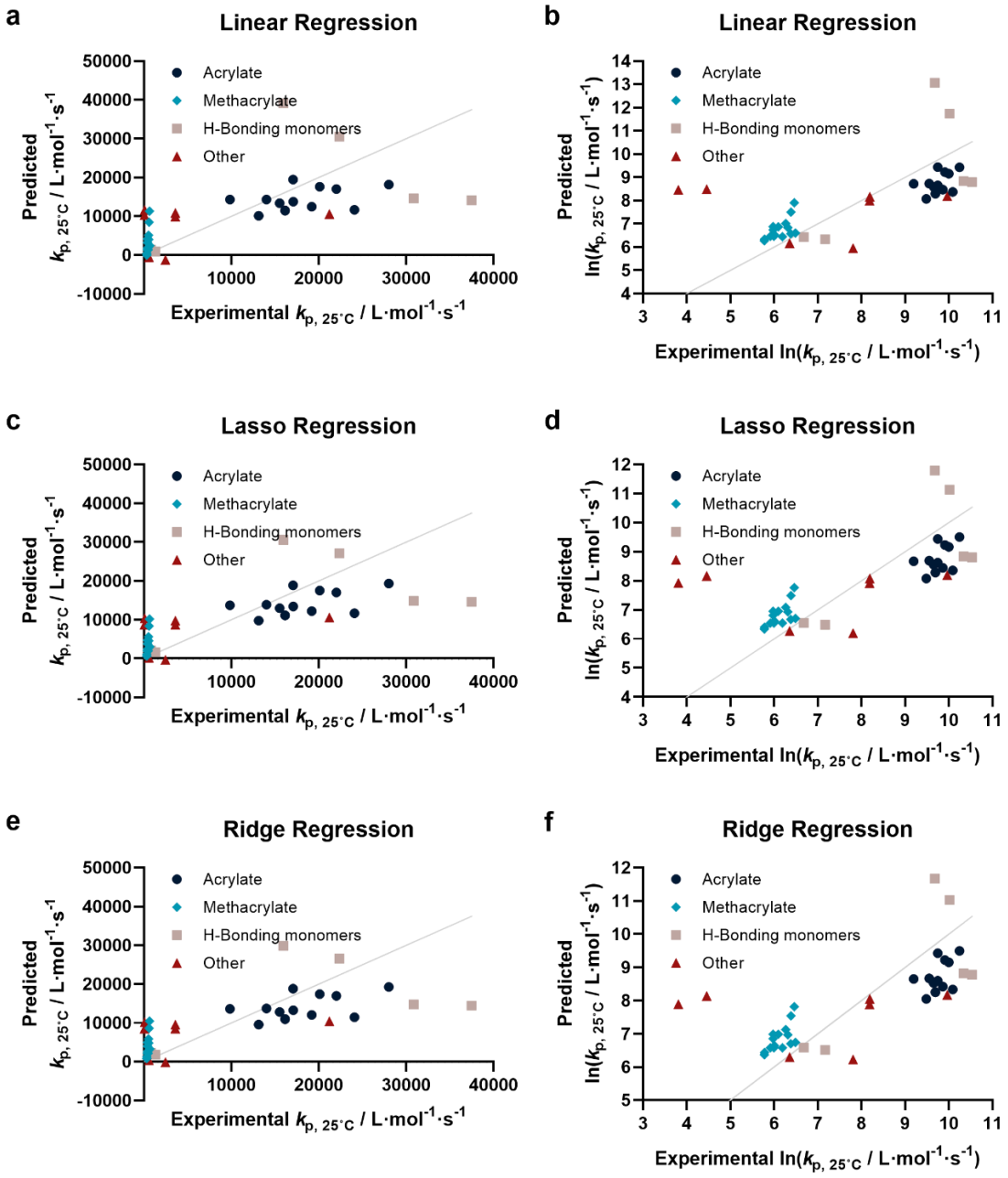


Figure 1: Correlation plots of the predicted propagation rate constant (k_p) and the natural logarithm of the predicted propagation rate constant (ln(k_p)) at 25˚C versus their experimental values for different regressions. The Linear (a and b), Lasso (c and d) and Ridge (e and f) regression were done using all available data (n = 41). The predictions were generated with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers) as parameters. The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	7.892E3	1.470	7.085E3	1.330	7.114E3	1.334
R2 Value	0.434	0.414	0.544	0.521	0.0540	0.518
Variance Predictions	6.865E7	2.167	5.053E7	1.645	4.782E7	1.563
Variance Residuals	6.380E7	2.213	5.145E7	1.813	5.187E7	1.824

All available K_p and ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants, Chemspider parameters (Polarizability)]

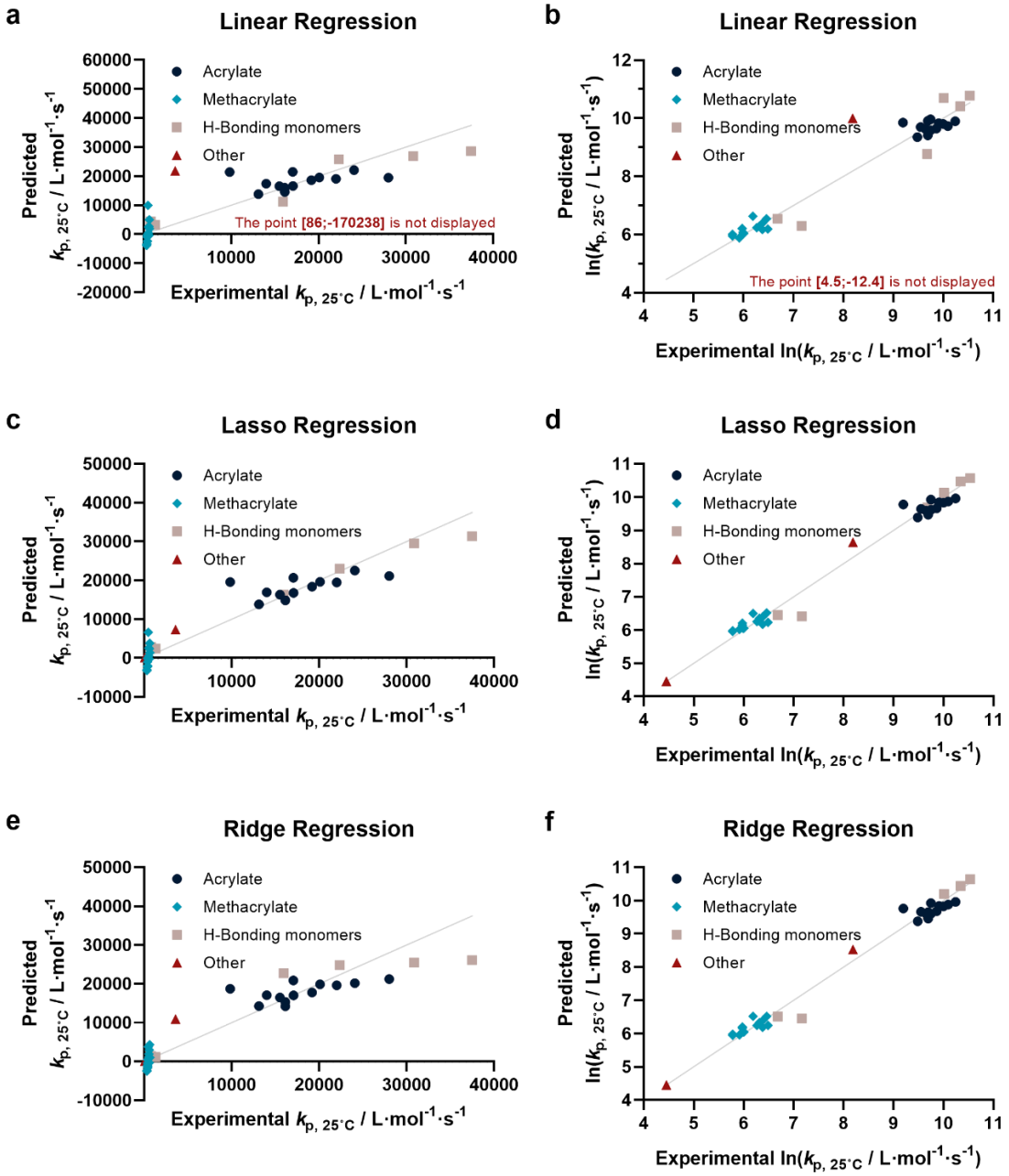
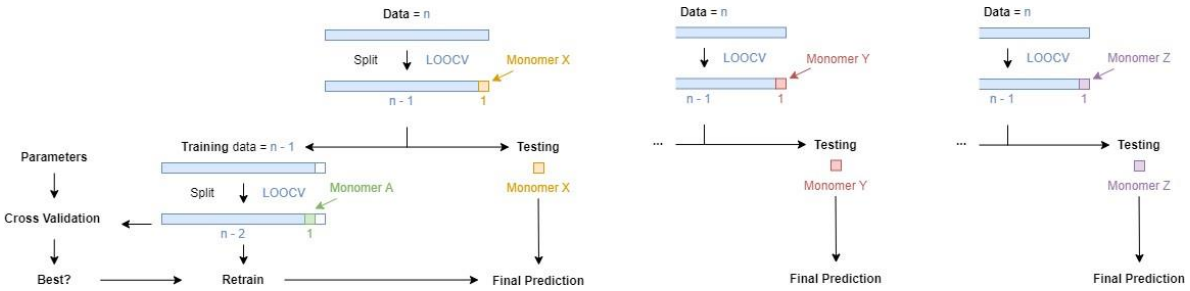


Figure 2: Correlation plots of the predicted propagation rate constant (k_p) and the natural logarithm of the predicted propagation rate constant (ln(k_p)) at 25˚C versus their experimental values for different regressions. The Linear (a and b), Lasso (c and d) and Ridge (e and f) regression were done using all available data (n = 35). The predictions were generated with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, and the dissociation constants as parameters. The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	2.924E4	2.889	3.113E3	0.235	3.784E3	0.223
R2 Value	-6.420	-1.350	0.916	0.984	0.876	0.986
Variance Predictions	1.010E9	15.005	1.039E8	3.471	9.580E7	3.484
Variance Residuals	8.595E8	8.386	9.977E6	0.057	1.474E7	0.051

Part 2: Tests with cross validation



All ln(K_p) data [Molecular weight]

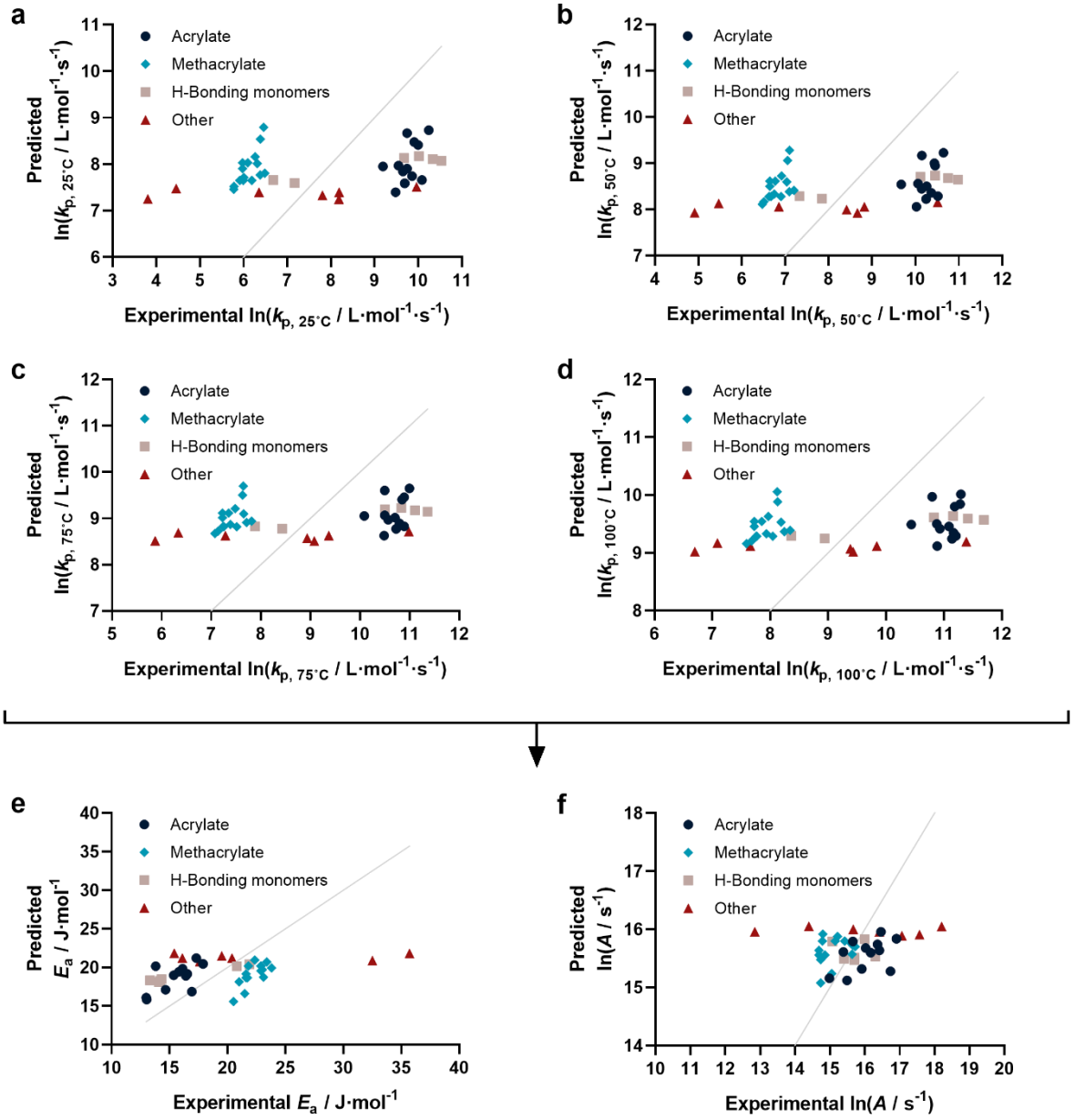


Figure 3: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 41) with the molecular weight as parameter. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	1.839	1.722	1.623	1.540	4.408	0.932
R2 Value	0.083	0.076	0.069	0.063	0.152	0.030
Variance Predictions	3.468	0.123	0.095	0.073	2.687	0.065
Variance Residuals	1.722	3.039	2.700	2.431	19.912	0.891

All ln(K_p) data [Molecular weight, a_value]

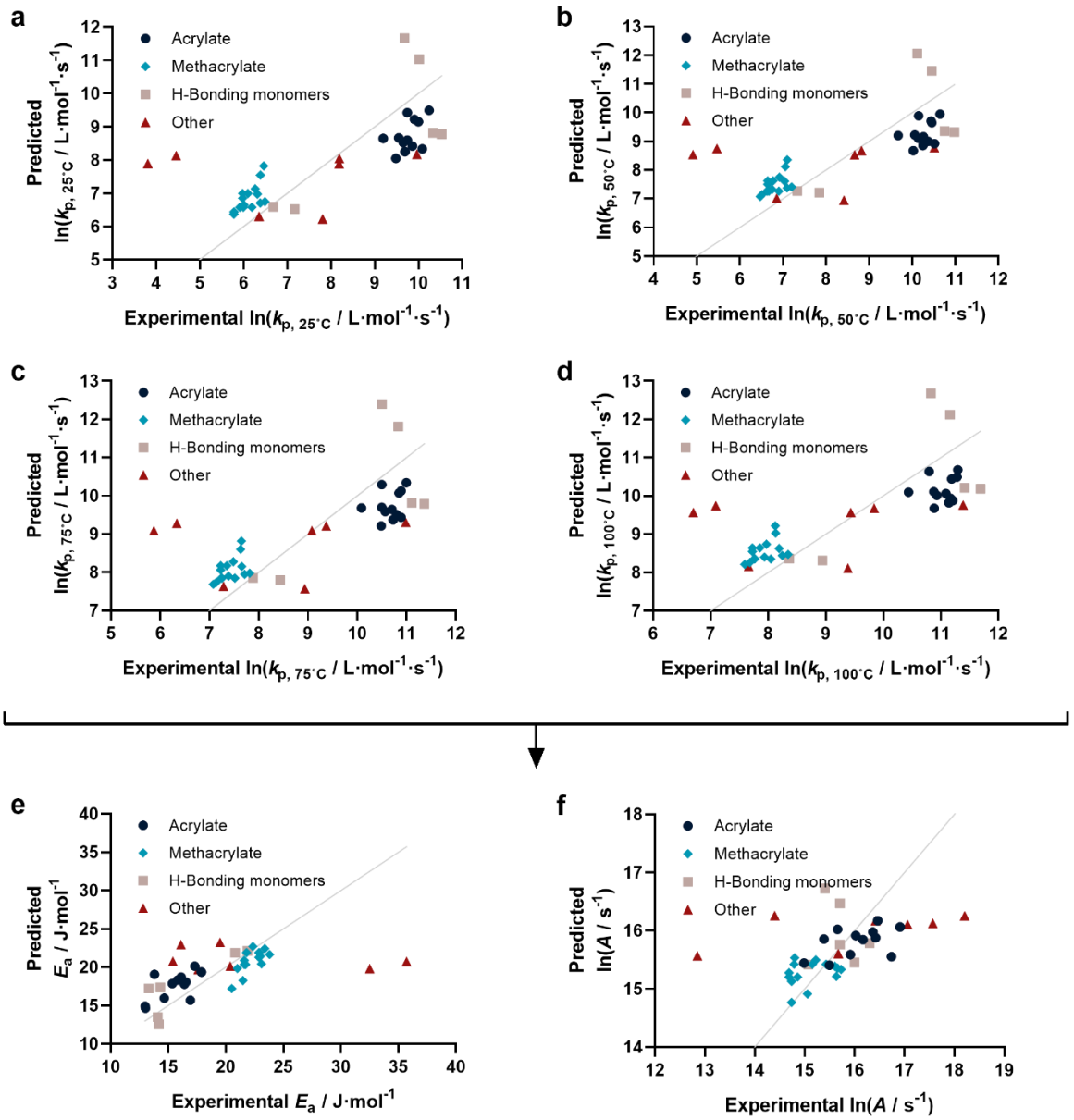


Figure 4: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 41) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers) as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	1.334	1.232	1.146	1.076	3.914	0.817
R2 Value	0.518	0.527	0.536	0.543	0.332	0.255
Variance Predictions	1.561	1.377	1.228	1.106	6.943	0.183
Variance Residuals	1.825	1.555	1.348	1.186	15.703	0.684

All ln(K_p) data [Molecular weight, Inductive effects (a_value, R_Value)]

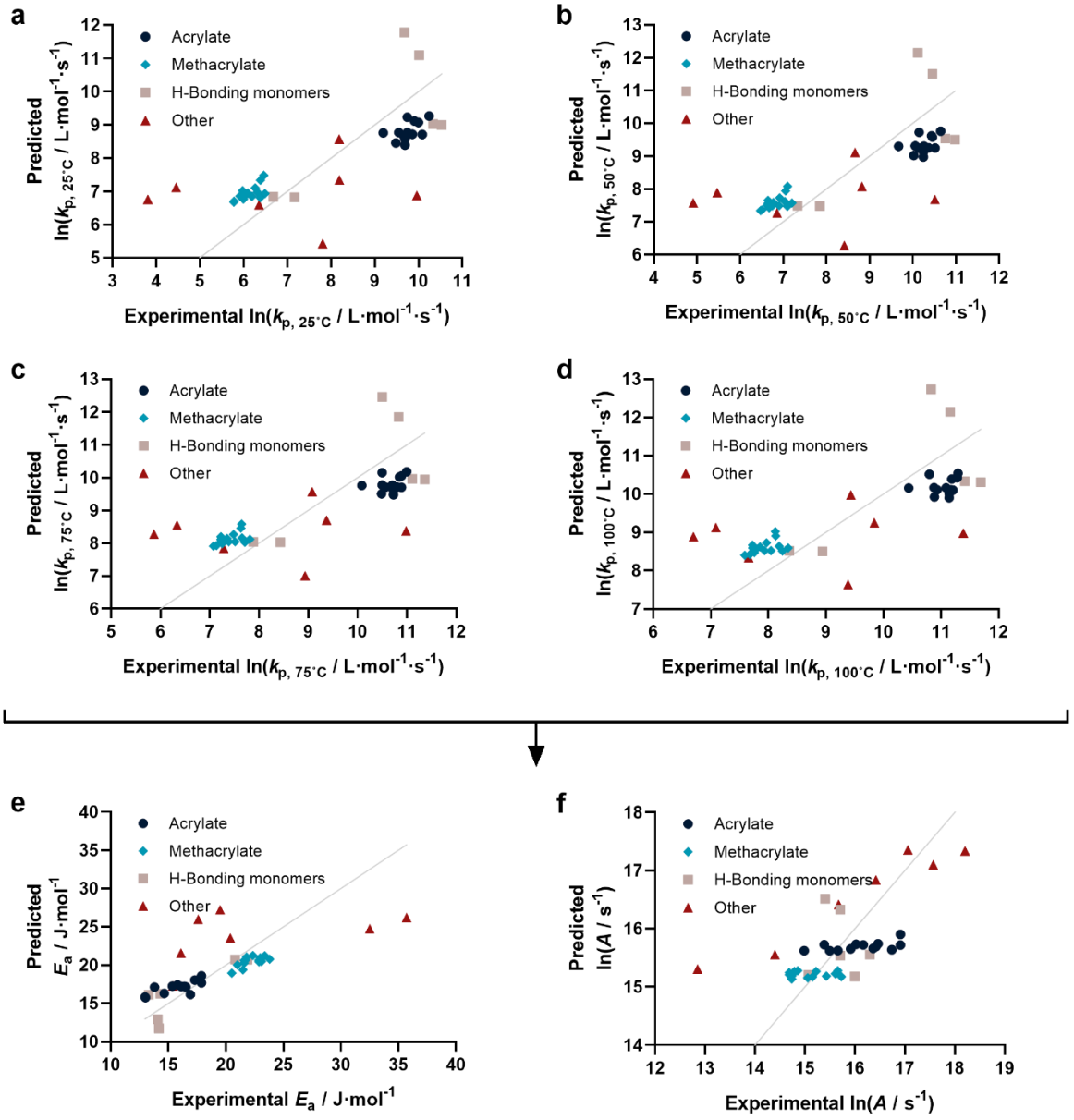


Figure 5: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 41) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers) and inductive effect of the tail group as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figurea	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	1.256	1.171	1.100	1.041	3.271	0.708
R2 Value	0.572	0.572	0.572	0.572	0.533	0.441
Variance Predictions	1.647	1.424	1.249	1.111	10.839	0.363
Variance Residuals	1.618	1.407	1.241	1.110	10.970	0.513

The ln(K_p) data of all acrylates and methacrylates [Molecular weight, Inductive effects (a_value, R_Value)]

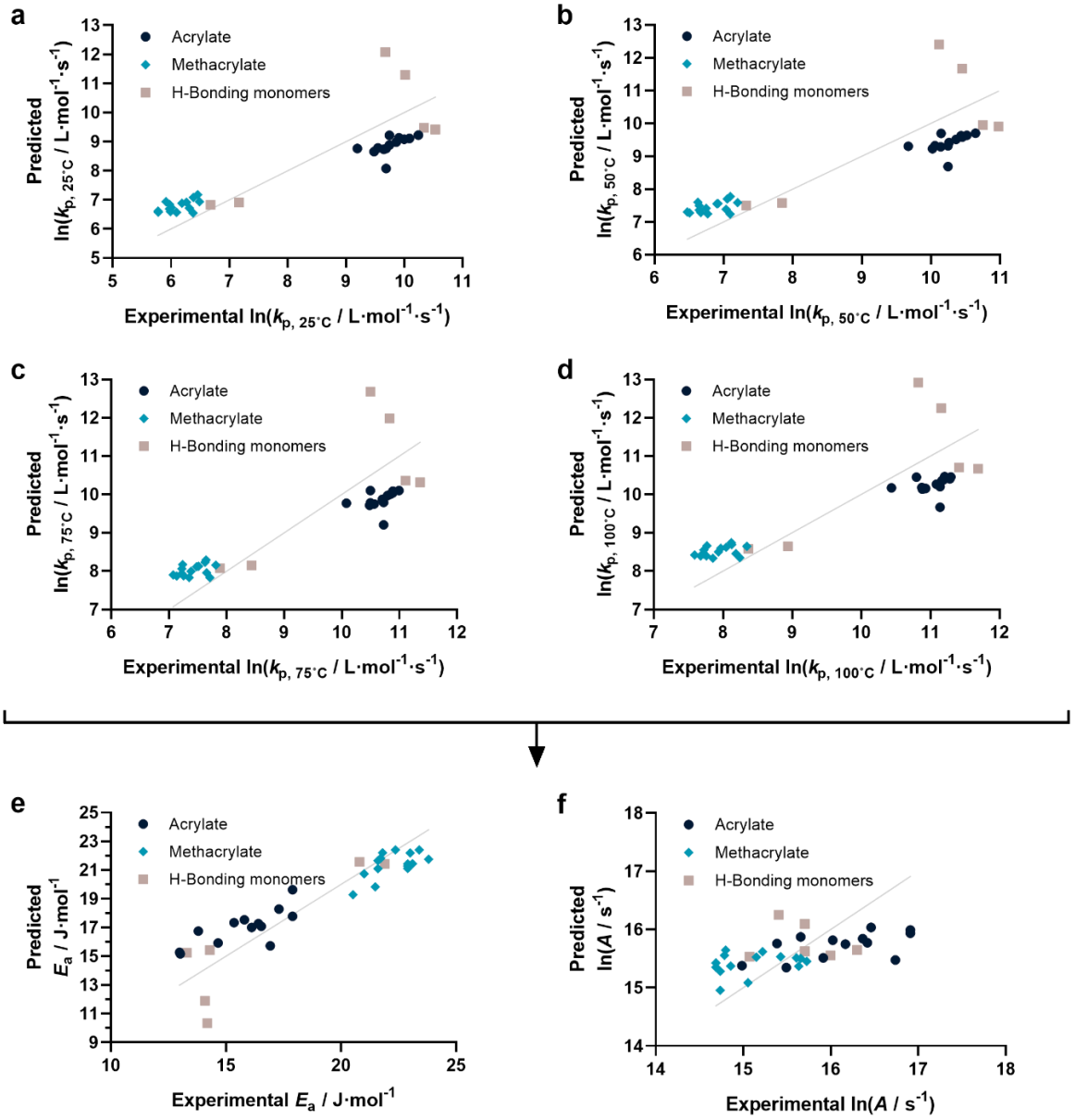


Figure 6: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all acrylates and methacrylates (n = 34) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers) and inductive effect of the tail group as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.899	0.861	0.829	0.803	1.546	0.545
R2 Value	0.761	0.753	0.745	0.736	0.822	0.301
Variance Predictions	1.998	1.740	1.532	1.357	10.012	0.078
Variance Residuals	0.834	0.764	0.708	0.664	2.461	0.306

The data ln(K_p) of acrylates and methacrylates without H-binding monomers [Molecular weight, Inductive effects (a_value, R_Value)]

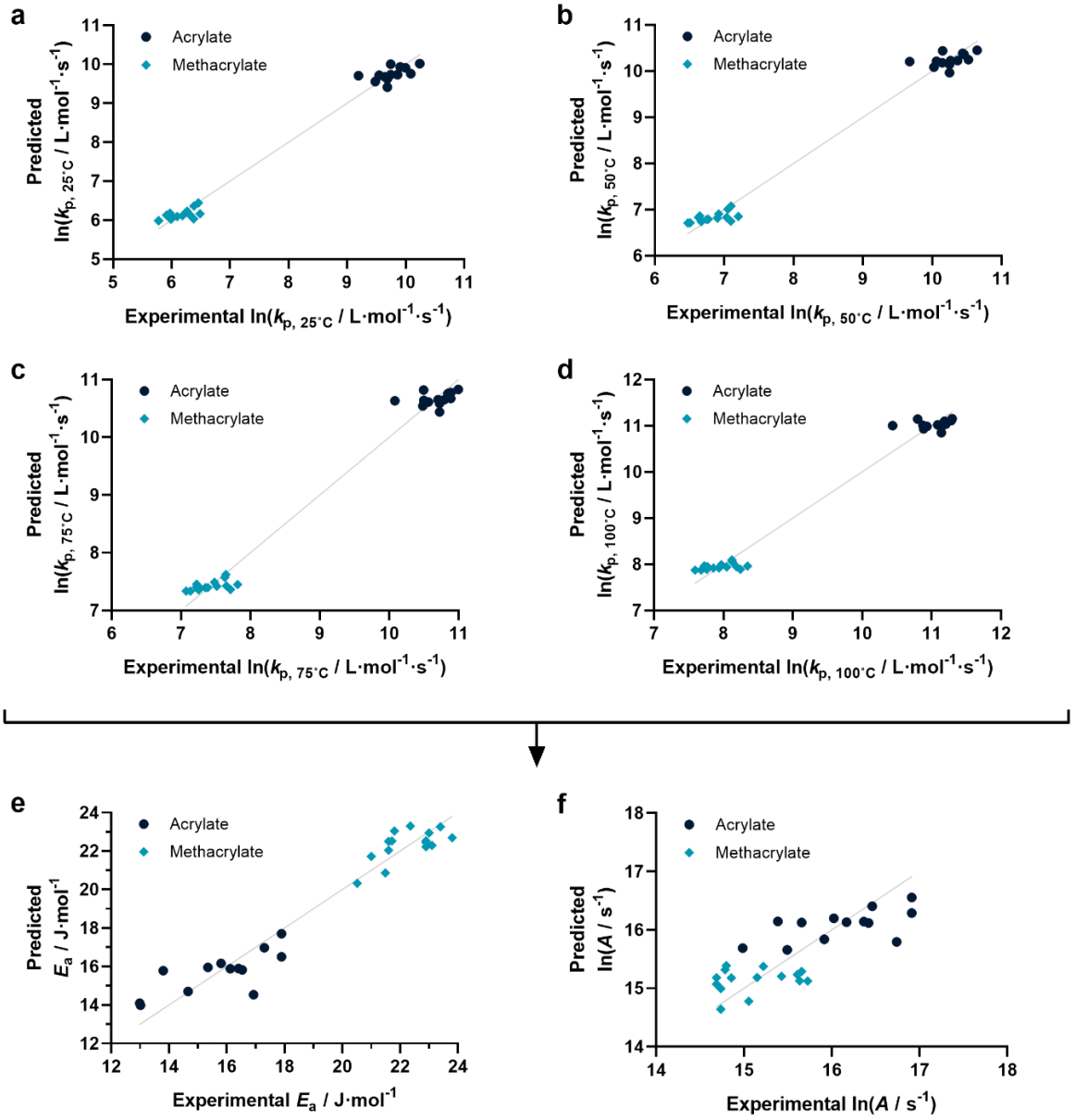


Figure 7: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all acrylates and methacrylates without H-binding monomers (n = 28) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers) and inductive effect of the tail group as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.198	0.202	0.209	0.216	0.907	0.429
R2 Value	0.988	0.986	0.984	0.981	0.933	0.618
Variance Predictions	3.273	2.905	2.607	2.362	12.203	0.284
Variance Residuals	0.041	0.042	0.045	0.049	0.853	0.190

All ln(K_p) data [Molecular weight, a_value, R_value, Chemdraw (dipole moment, boiling point, melting point, Gibbs Free Energy)]

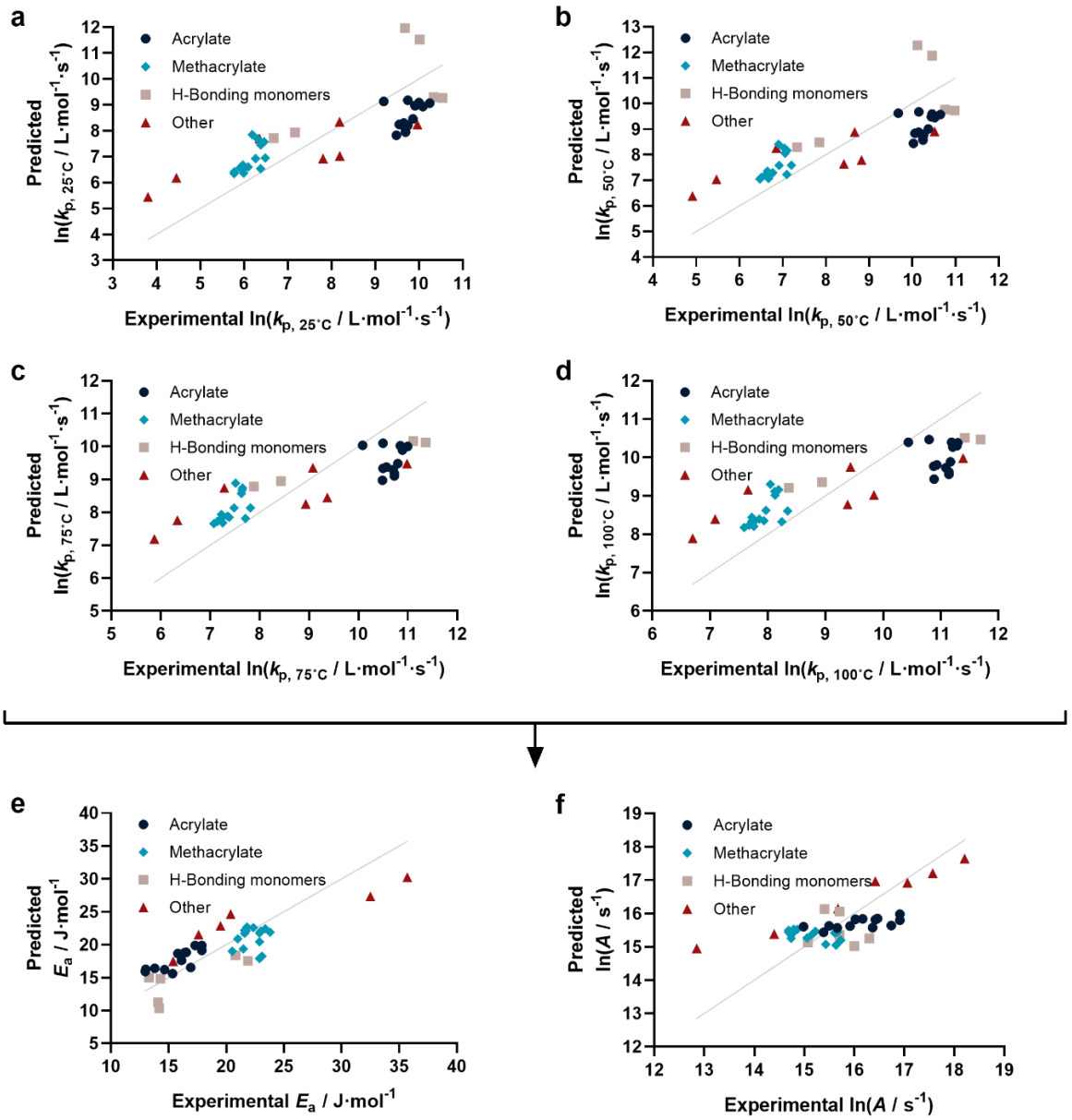


Figure 8: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 41) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, dipole moment, boiling point, melting point and Gibbs Free Energy as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	1.167	1.100	1.044	0.998	2.667	0.690
R2 Value	0.631	0.623	0.615	0.607	0.690	0.487
Variance Predictions	1.775	1.504	1.292	1.124	14.206	0.345
Variance Residuals	1.396	1.241	1.117	1.020	7.289	0.471

The ln(K_p) data of all acrylates and methacrylates [Molecular weight, a_value, R_value, Chemdraw (dipole moment, boiling point, melting point, Gibbs Free Energy)]

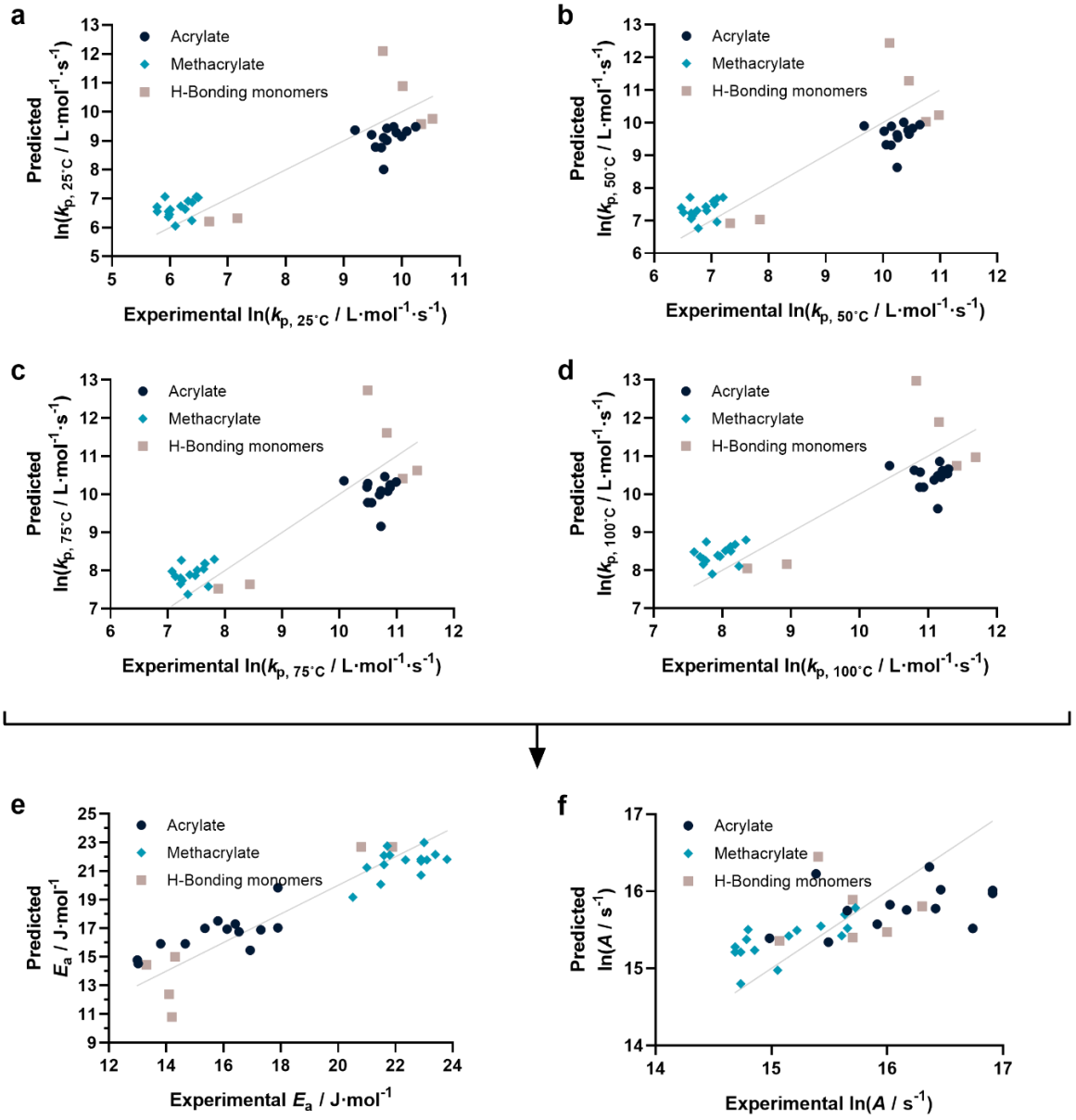


Figure 9: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all acrylates and methacrylates (n = 34) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, dipole moment, boiling point, melting point and Gibbs Free Energy as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.806	0.773	0.745	0.722	1.397	0.512
R2 Value	0.808	0.801	0.794	0.787	0.855	0.384
Variance Predictions	2.434	2.125	1.877	1.677	11.592	0.132
Variance Residuals	0.670	0.615	0.572	0.537	2.012	0.270

The ln(K_p) data of acrylates and methacrylates without H-binding monomers [Molecular weight, a_value, R_value, Chemdraw (dipole moment, boiling point, melting point, Gibbs Free Energy)]

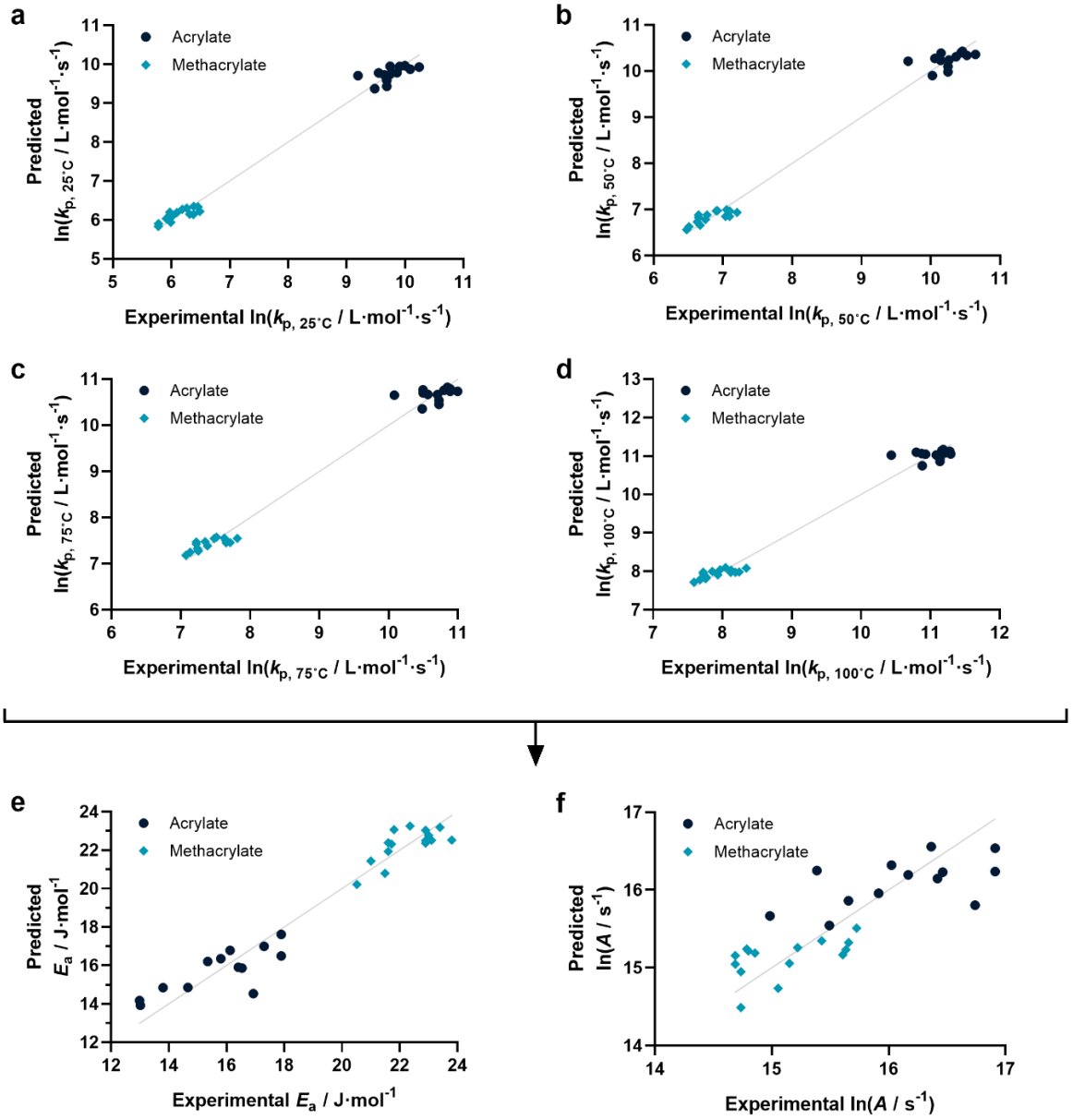


Figure 10: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all acrylates and methacrylates without H-binding monomers (n = 28) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, dipole moment, boiling point, melting point and Gibbs Free Energy as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.181	0.186	0.192	0.200	0.849	0.403
R2 Value	0.990	0.988	0.986	0.984	0.941	0.661
Variance Predictions	3.277	2.915	2.621	2.376	12.065	0.320
Variance Residuals	0.034	0.036	0.038	0.042	0.748	0.169

All ln(K_p) data [Molecular weight, a_value, R_value, H-bonding (H-donor, H-acceptor)]

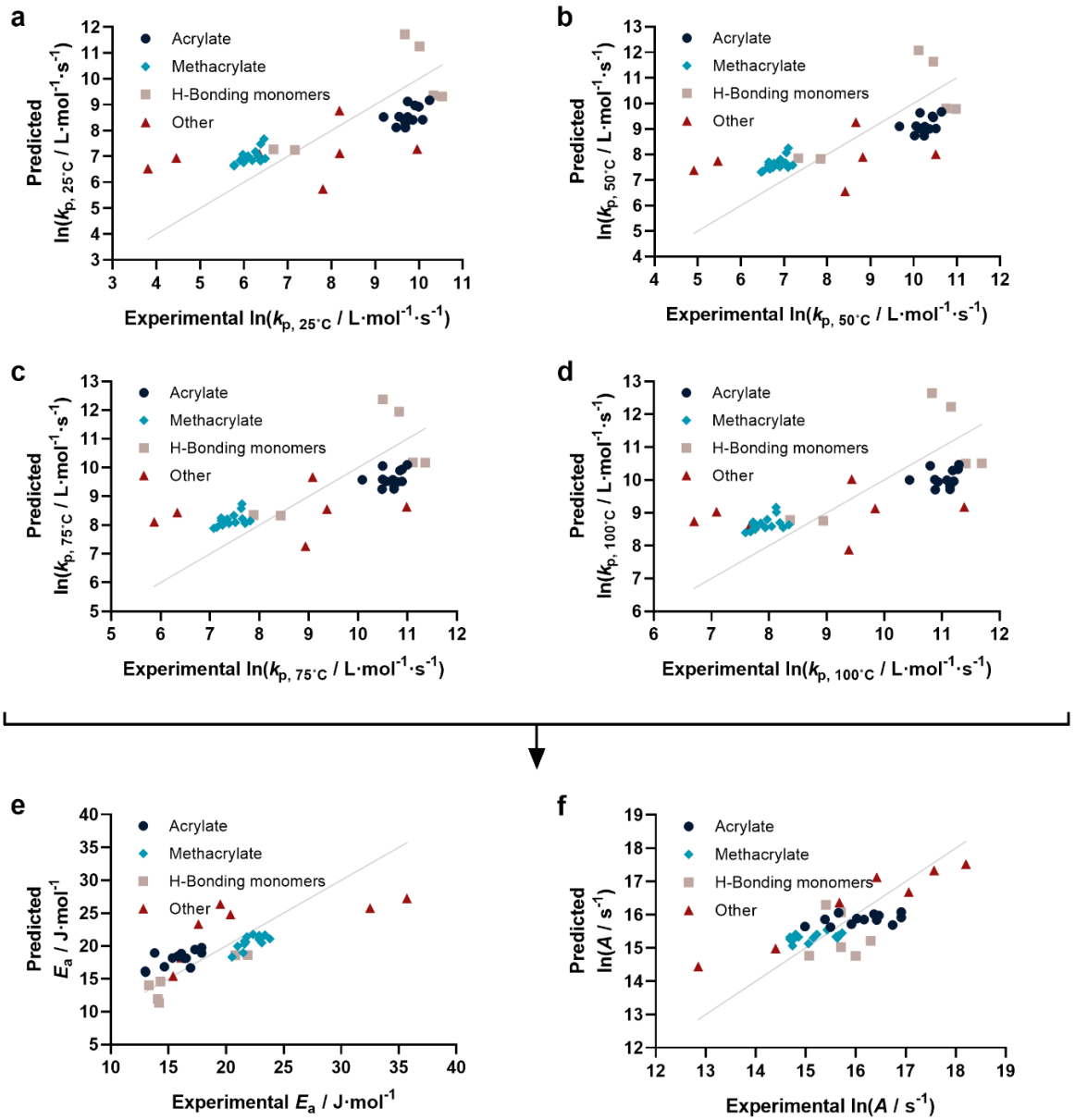


Figure 11: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 41) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	1.272	1.189	1.120	1.061	3.053	0.617
R2 Value	0.561	0.559	0.557	0.555	0.593	0.574
Variance Predictions	1.509	1.287	1.126	0.994	11.659	0.434
Variance Residuals	1.658	1.449	1.285	1.154	9.554	0.391

All ln(K_p) data [Molecular weight, a_value, R_value, Chemdraw parameters, H-bonding parameters]

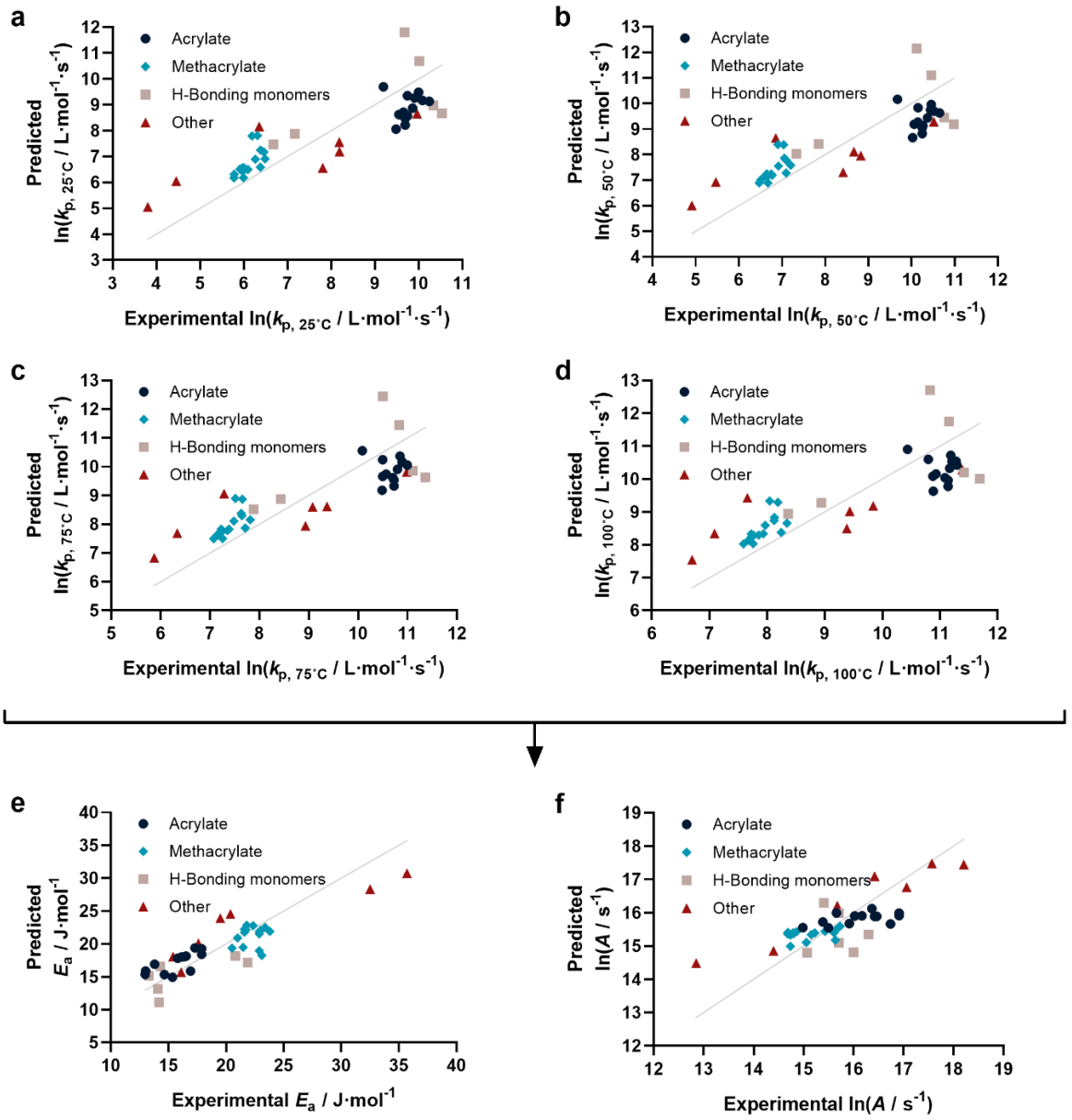


Figure 12: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 41) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, dipole moment, boiling point, melting point, Gibbs Free Energy, the effect of H-donors and H-acceptors as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	1.048	0.986	0.936	0.893	2.421	0.606
R2 Value	0.703	0.697	0.691	0.685	0.744	0.590
Variance Predictions	1.901	1.622	1.402	1.230	14.638	0.426
Variance Residuals	1.125	0.998	0.898	0.818	6.006	0.377

All ln(K_p) data [Molecular weight, a_value, R_value, Dissociation constants (A1* and A2*)]

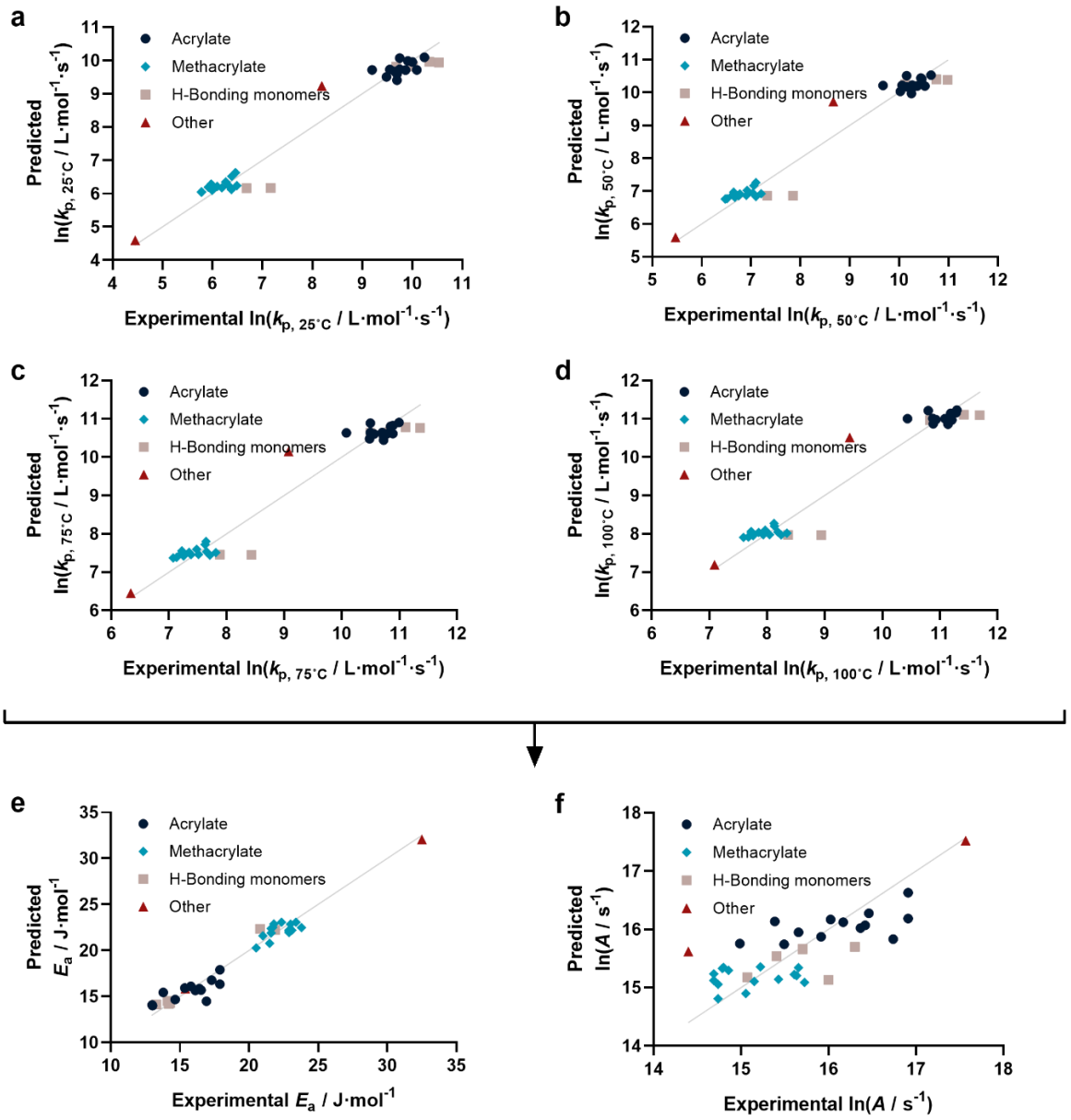


Figure 13: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 36) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group and the dissociation constants as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.343	0.341	0.342	0.344	0.874	0.472
R2 Value	0.967	0.963	0.957	0.952	0.958	0.595
Variance Predictions	3.383	2.947	2.599	2.318	17.495	0.305
Variance Residuals	0.121	0.120	0.121	0.122	0.786	0.229

All ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants]

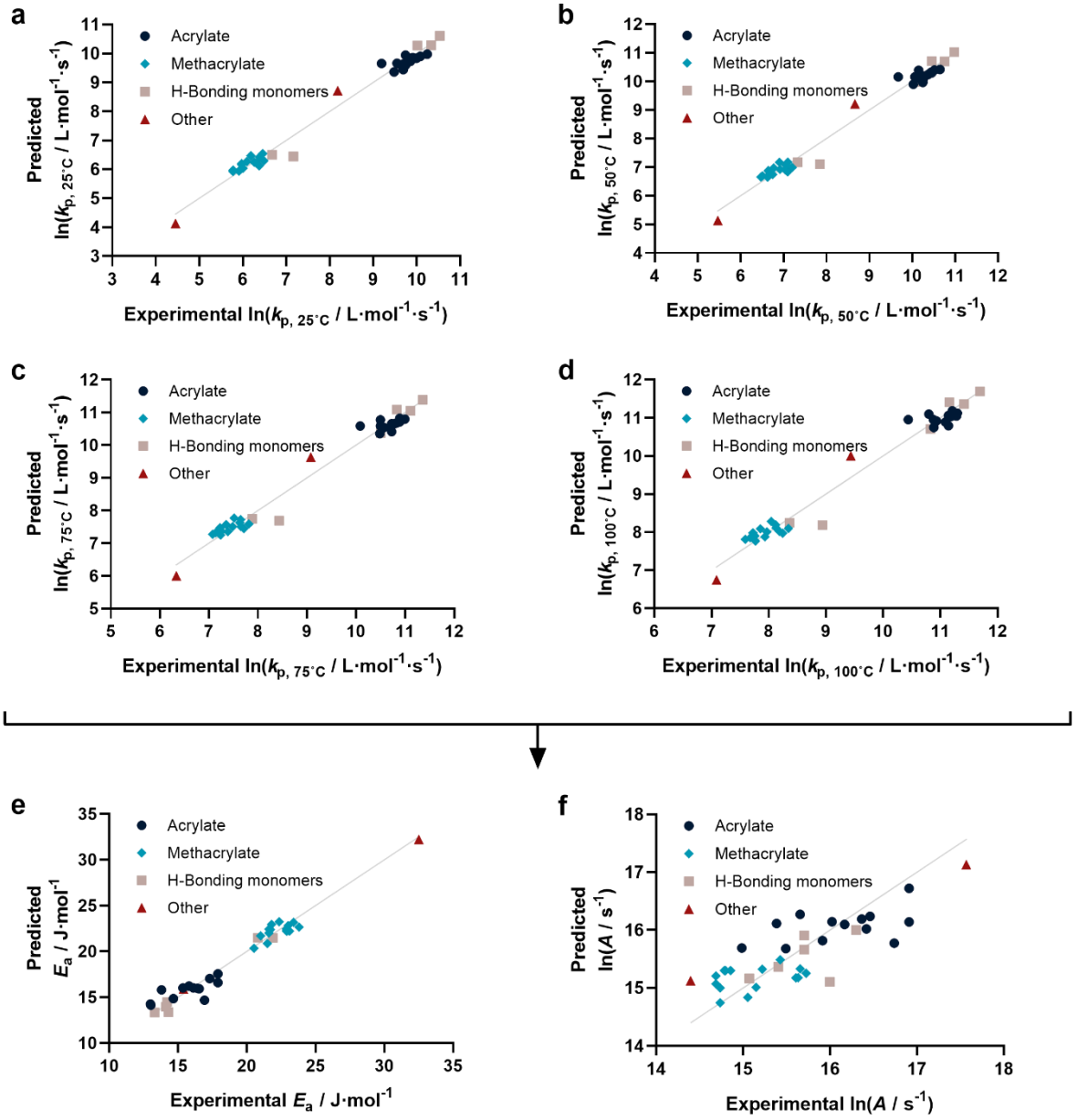


Figure 14: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 36) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, and the dissociation constants as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.232	0.238	0.245	0.253	0.824	0.439
R2 Value	0.985	0.982	0.978	0.974	0.963	0.649
Variance Predictions	3.540	3.088	2.727	2.435	17.788	0.300
Variance Residuals	0.055	0.058	0.062	0.066	0.698	0.198

All ln(K_p) data [Molecular weight, a_value, R_value, Chemdraw parameters, H-bonding parameters, Dissociation constants]

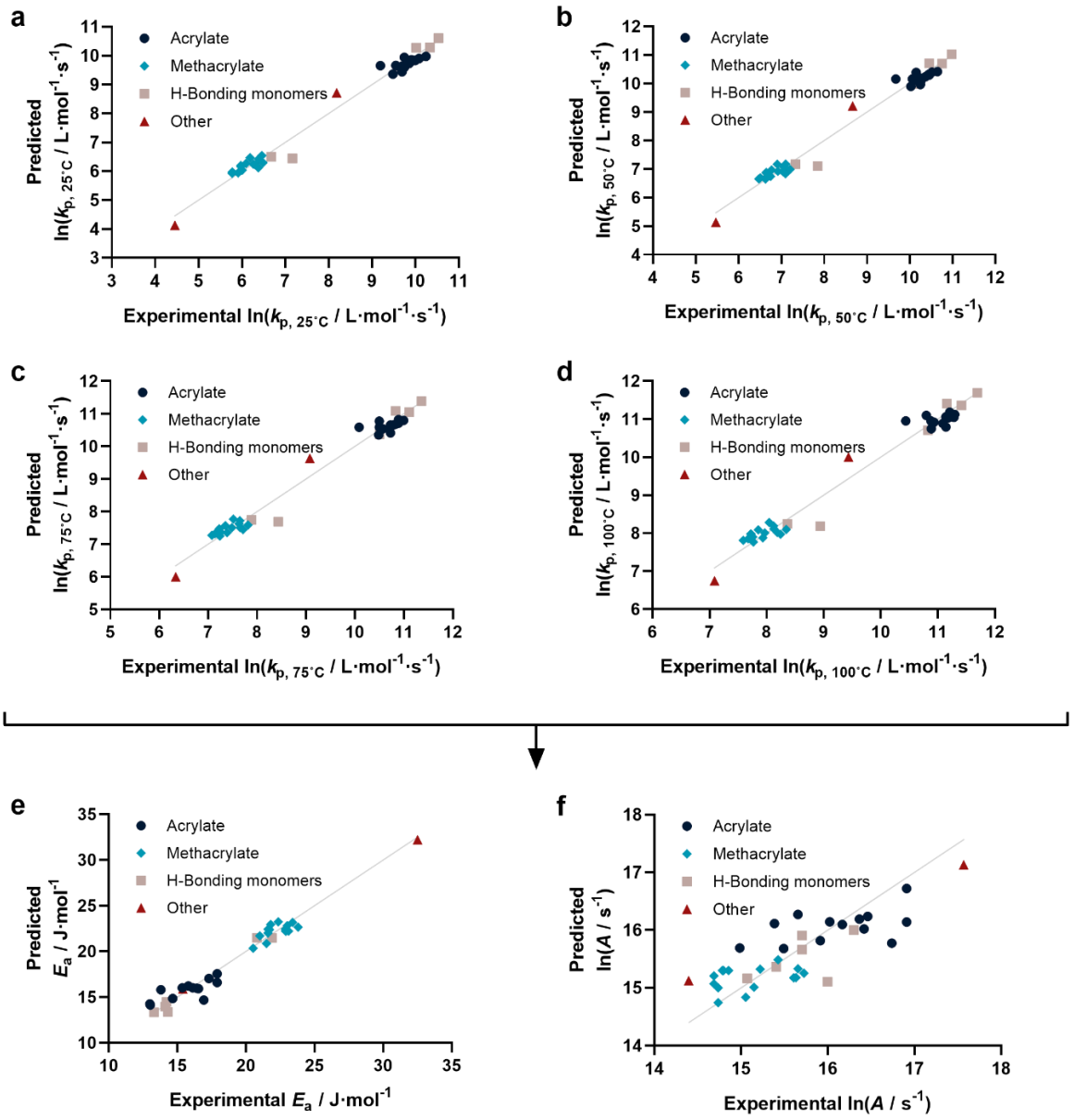


Figure 15: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 36) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, dipole moment, boiling point, melting point, Gibbs Free Energy, the effect of H-donors and H-acceptors, and the dissociation constants as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figuree	Figure f
RMSE	0.172	0.178	0.185	0.193	0.792	0.386
R2 Value	0.992	0.990	0.988	0.985	0.965	0.729
Variance Predictions	3.497	3.057	2.706	2.421	17.734	0.398
Variance Residuals	0.031	0.032	0.035	0.038	0.645	0.153

All ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants, Chemspider parameters (predicted boiling point, refractive index and density)]

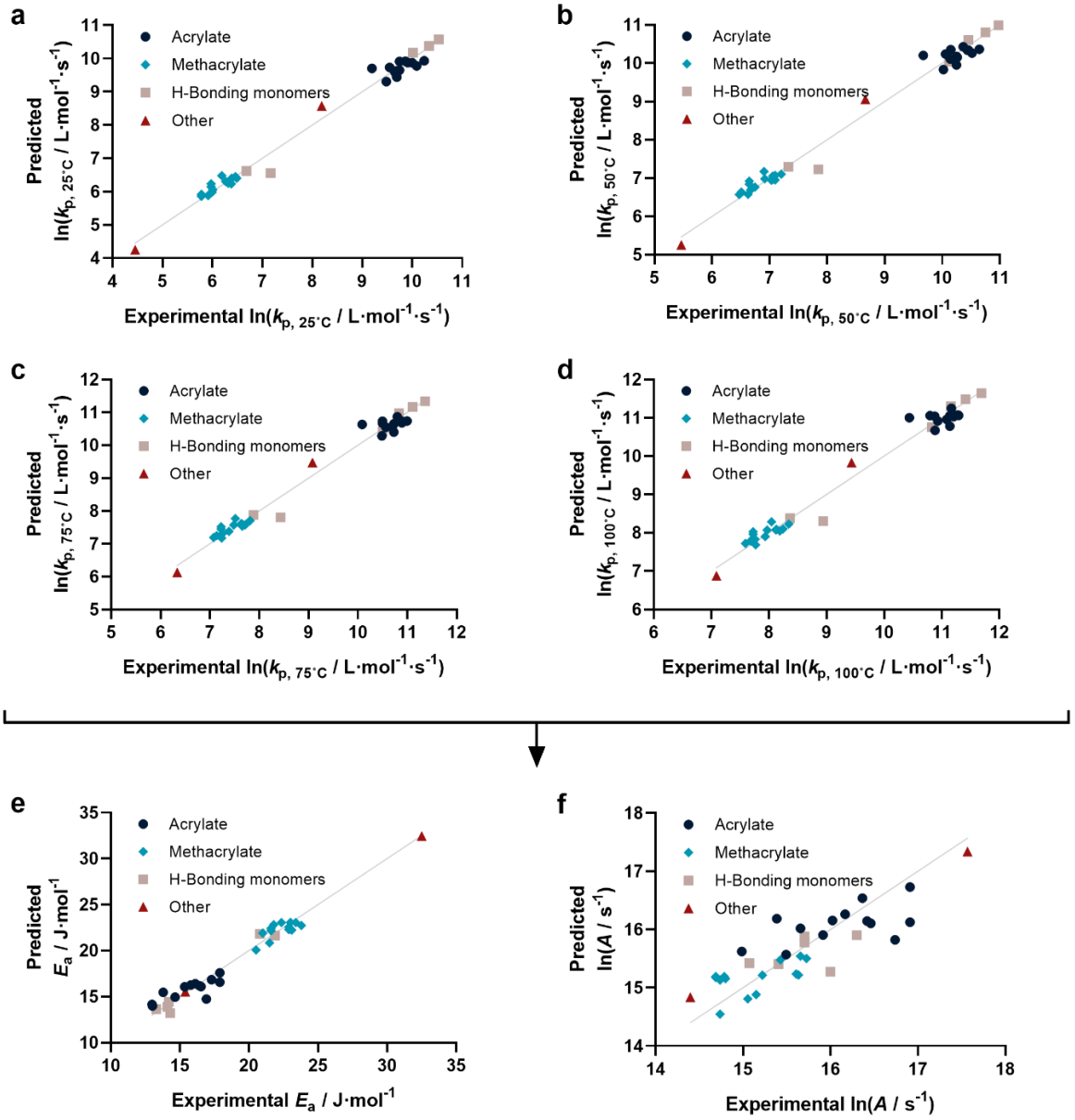


Figure 16: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 35) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, the dissociation constants as parameters, predicted boiling point, predicted refractive index and predicted density as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.207	0.210	0.216	0.222	0.907	0.429
R2 Value	0.987	0.986	0.983	0.980	0.933	0.618
Variance Predictions	3.532	3.079	2.718	2.426	12.206	0.284
Variance Residuals	0.044	0.045	0.048	0.051	0.853	0.190

All available ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants, Chemspider parameters (experimental boiling point, refractive index and density)]

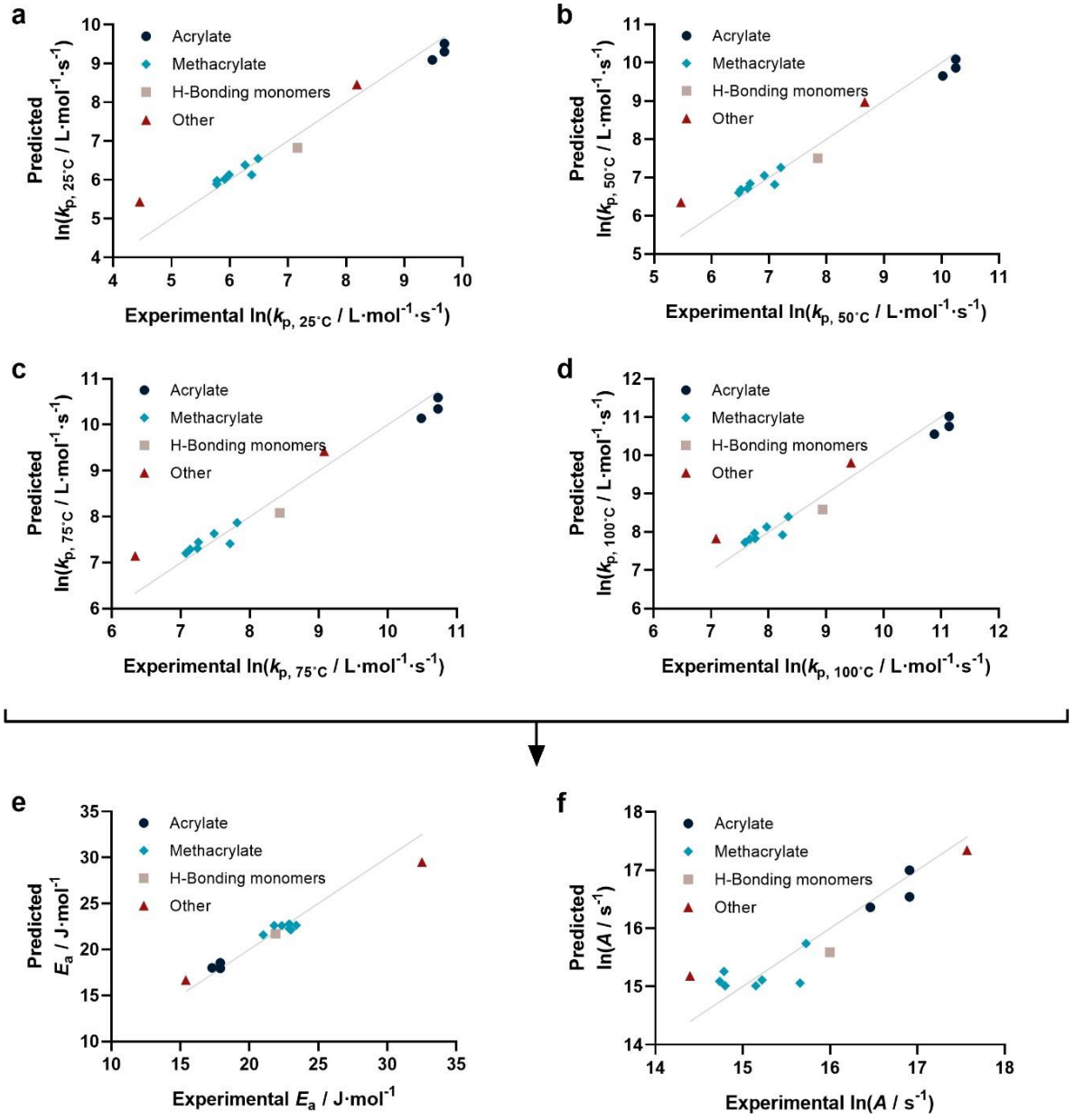


Figure 17: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 13) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, the dissociation constants, experimental boiling point, experimental refractive index and experimental density as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.355	0.339	0.325	0.316	1.050	0.369
R2 Value	0.953	0.951	0.950	0.948	0.933	0.851
Variance Predictions	1.991	1.769	1.600	1.459	10.617	0.673
Variance Residuals	0.136	0.123	0.114	0.107	1.186	0.148

All available ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants, Chemspider parameters (ACDlogP)]

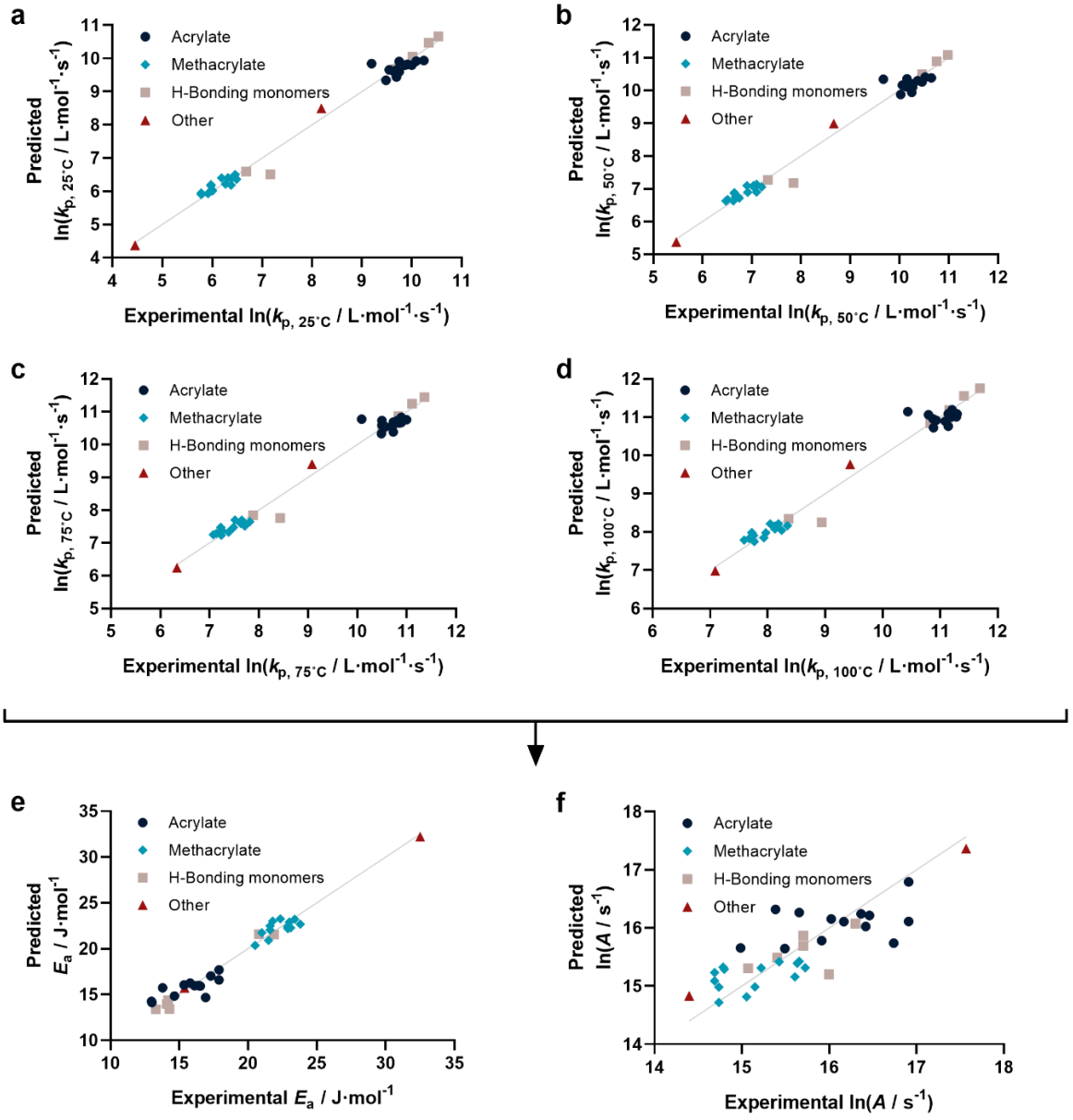


Figure 18: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 35) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, the dissociation constants and ACDlogP as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.212	0..218	0.225	0.233	0.823	0.423
R2 Value	0.987	0.985	0.982	0.978	0.963	0.674
Variance Predictions	3.512	3.062	2.704	2.413	18.061	0.343
Variance Residuals	0.046	0.049	0.052	0.056	0.698	0.184

All available ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants, Chemspider parameters (Polarizability)]

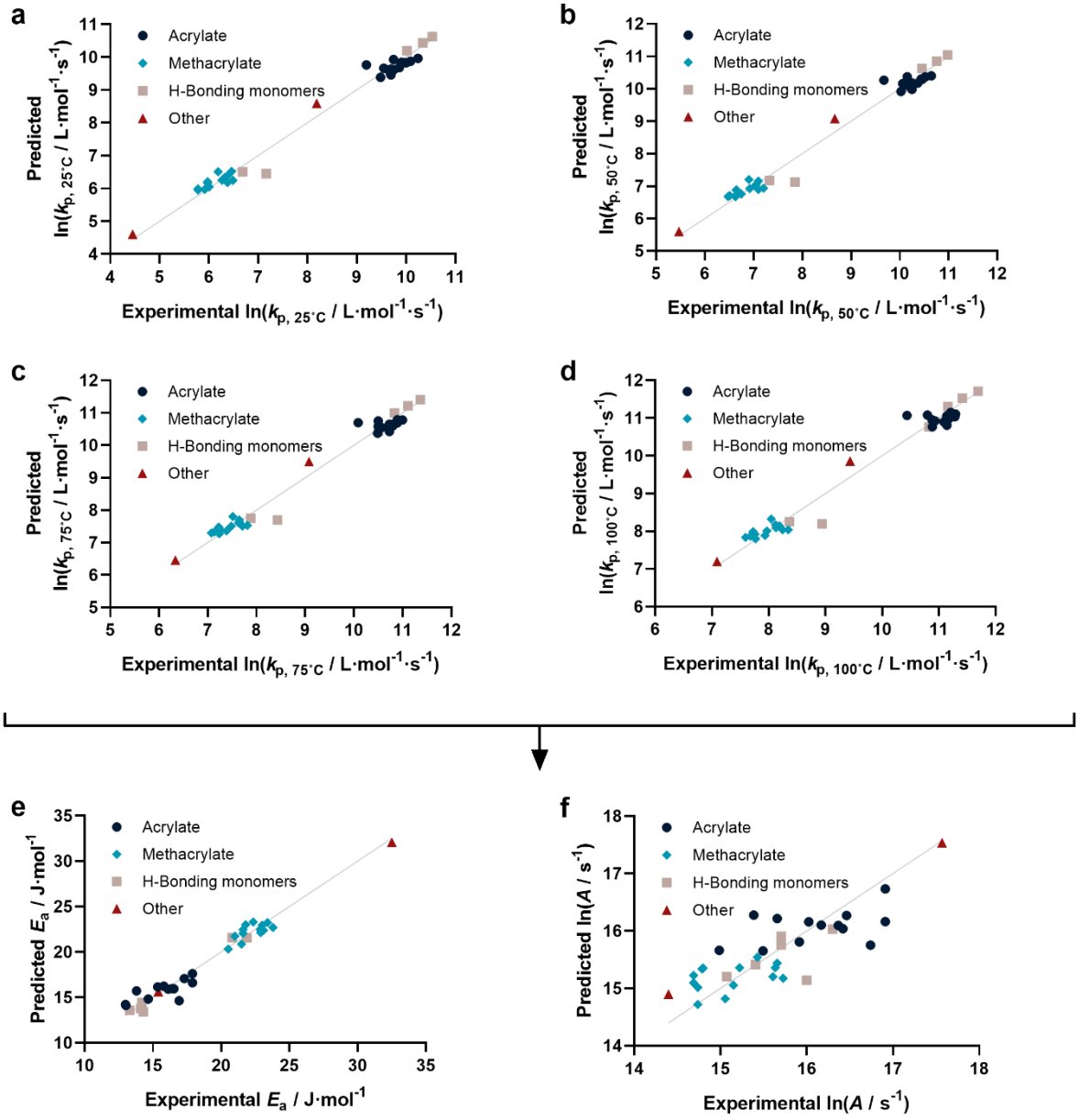


Figure 19: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 35) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, the dissociation constants and polarizability as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.227	0.232	0.238	0.245	0.823	0.427
R2 Value	0.985	0.983	0.979	0.976	0.963	0.666
Variance Predictions	3.454	3.010	2.658	2.368	18.064	0.349
Variance Residuals	0.053	0.055	0.058	0.062	0.701	0.188

All available ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants, Chemspider parameters (Surface tension)]

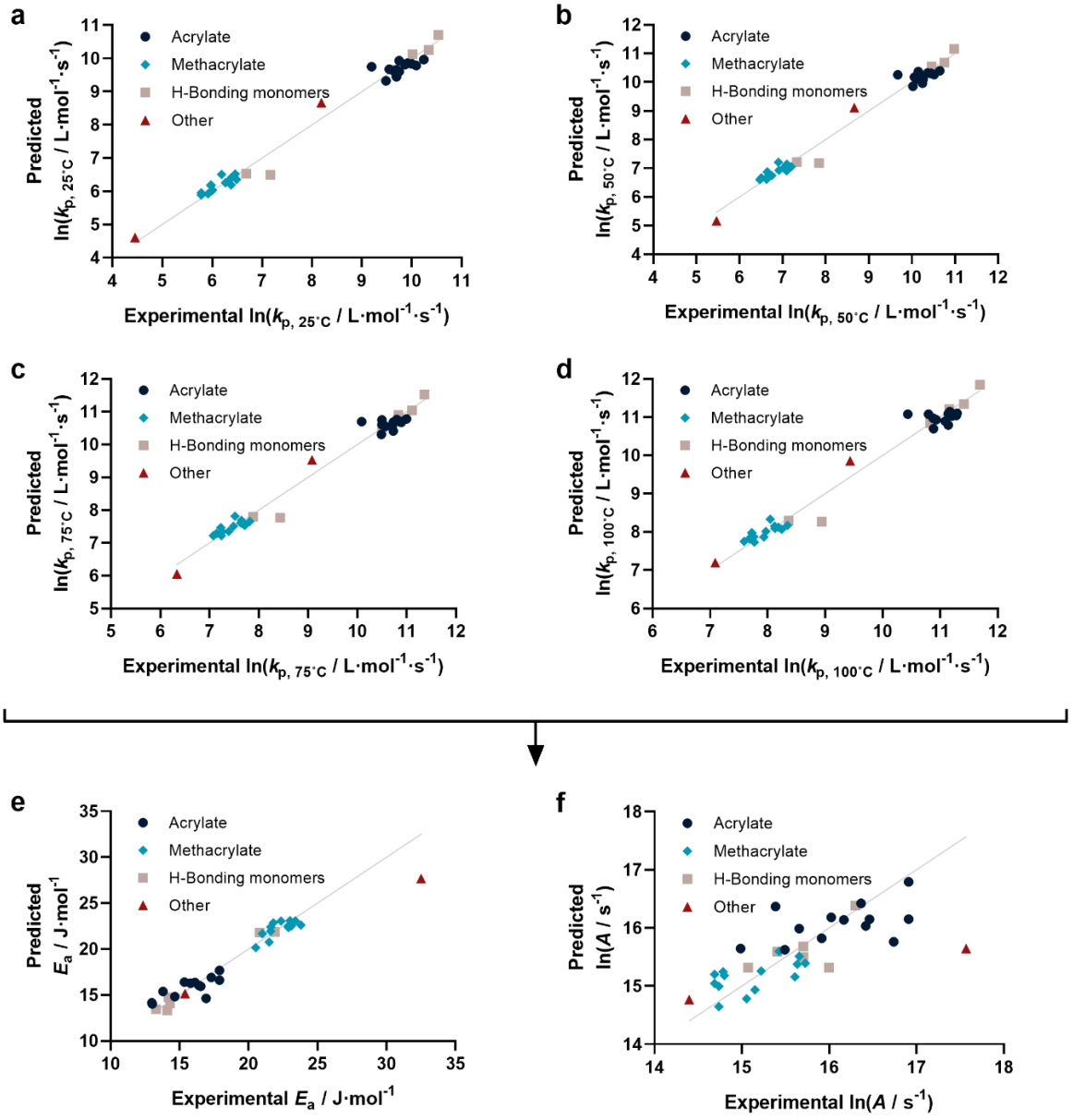


Figure 20: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 35) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, the dissociation constants and surface tension as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.222	0.226	0.231	0.236	1.139	0.512
R2 Value	0.986	0.983	0.981	0.977	0.930	0.521
Variance Predictions	3.450	3.090	2.725	2.430	15.167	0.276
Variance Residuals	0.050	0.053	0.055	0.058	1.316	0.267

All available ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants, Chemspider parameters (Vapor pressure)]

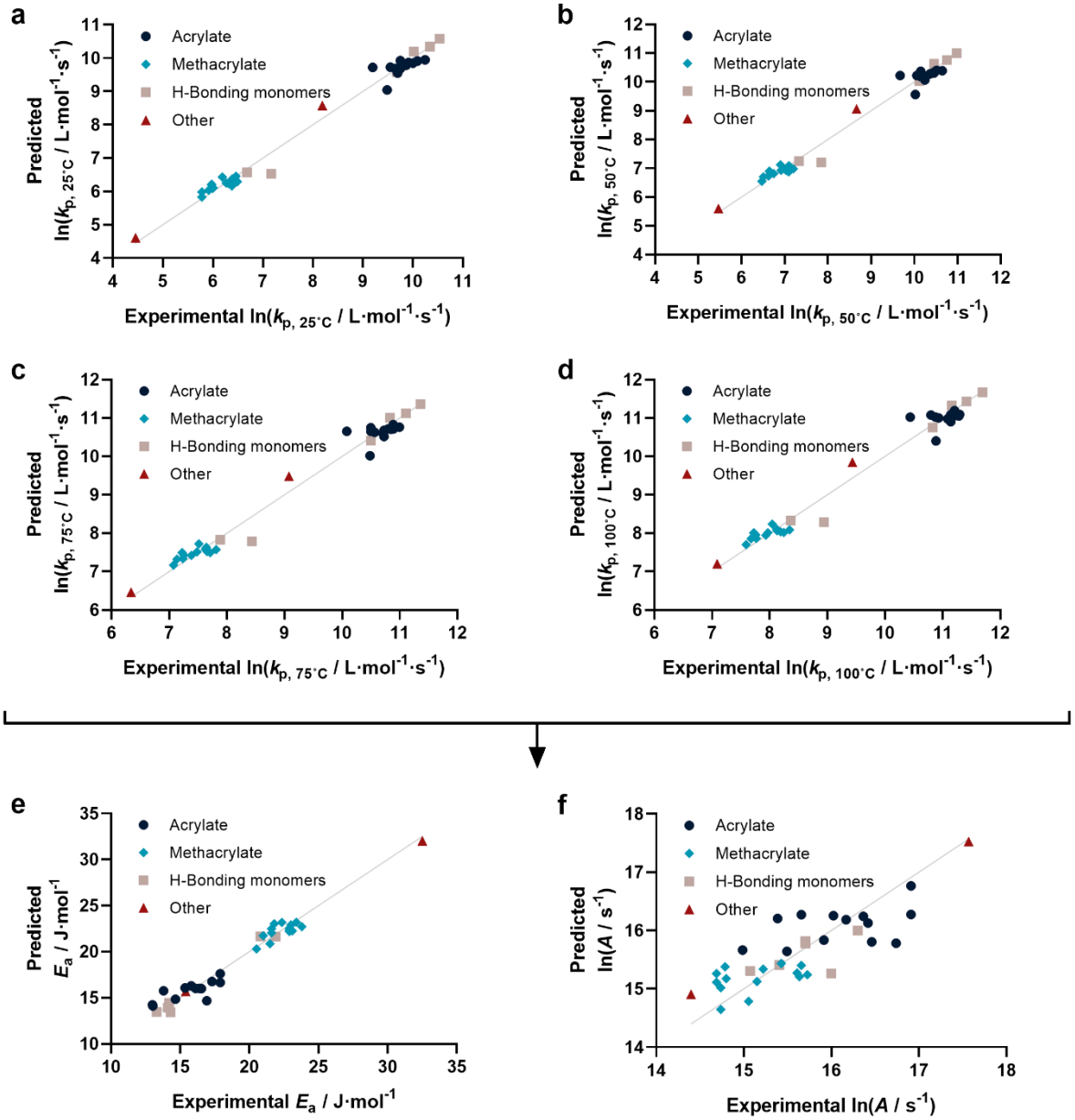


Figure 21: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 35) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, the dissociation constants and vapor pressure as parameters. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.219	0.224	0.230	0.237	0.825	0.421
R2 Value	0.987	0.984	0.981	0.977	0.963	0.676
Variance Predictions	3.455	3.012	2.660	2.374	17.927	0.355
Variance Residuals	0.049	0.052	0.054	0.058	0.701	0.183
RMSE	0.062	0.064	0.061	0.062	0.527	0.320
R2 Value	0.999	0.998	0.998	0.998	0.983	0.888
Variance Predictions	2.605	2.286	2.045	1.854	16.162	0.674
Variance Residuals	0.004	0.004	0.004	0.004	0.299	0.109

All available ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants, Chemspider parameters (Polarizability)]; using a weight of 0.5 for selected data

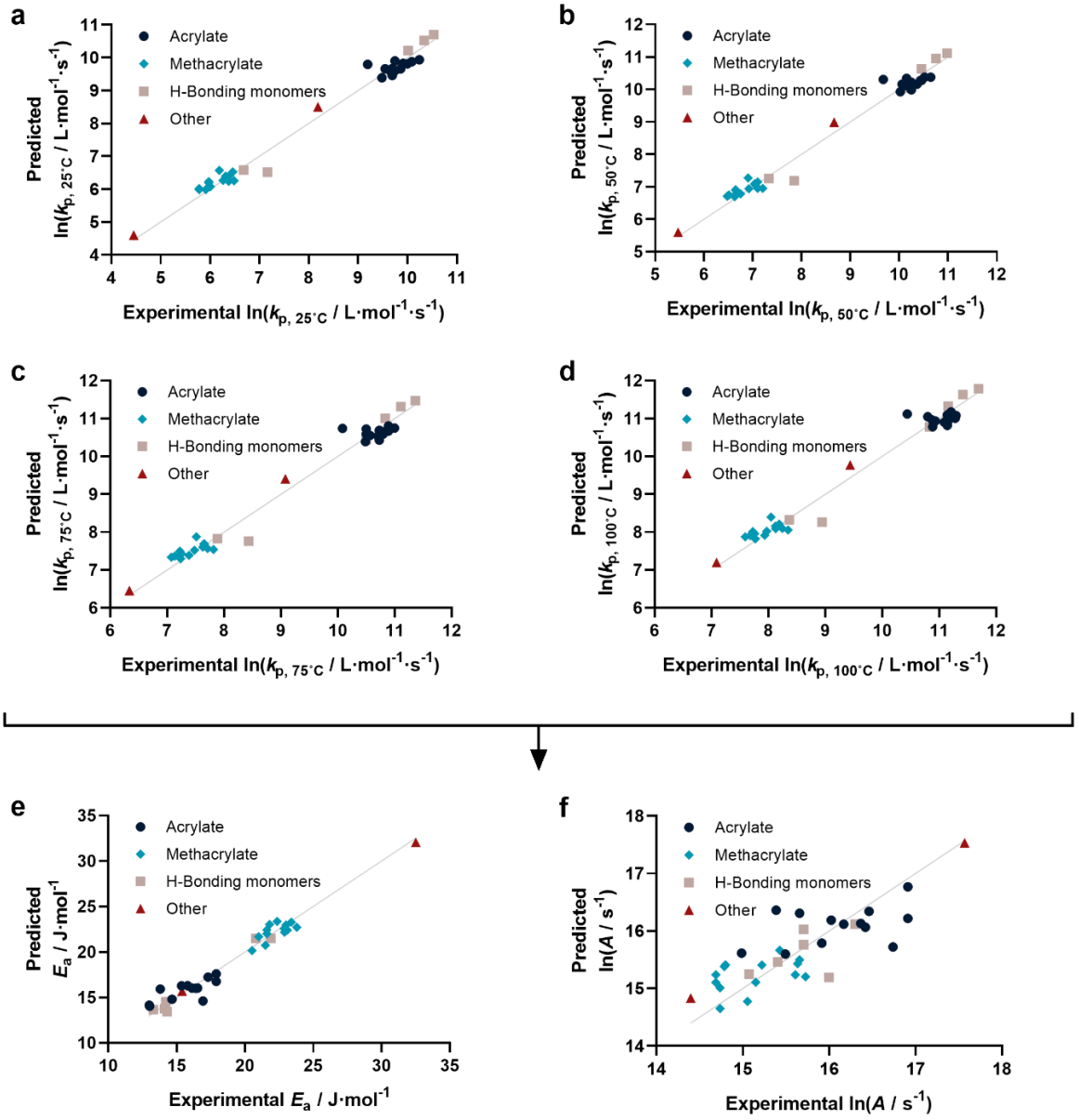


Figure 22: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 35) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, the dissociation constants and polarizability as parameters. A selection of monomers (n = 13) is given a weight of 1 while the rest (n = 22) has a weight of 0.5. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	0.228	0.233	0.240	0.247	0.832	0.432
R2 Value	0.985	0.983	0.979	0.975	0.963	0.660
Variance Predictions	3.149	2.977	2.629	2.348	17.910	0.366
Variance Residuals	0.053	0.055	0.058	0.062	0.713	0.191

All available ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants, Chemspider parameters (Polarizability)]; using a weight of 0 for selected data

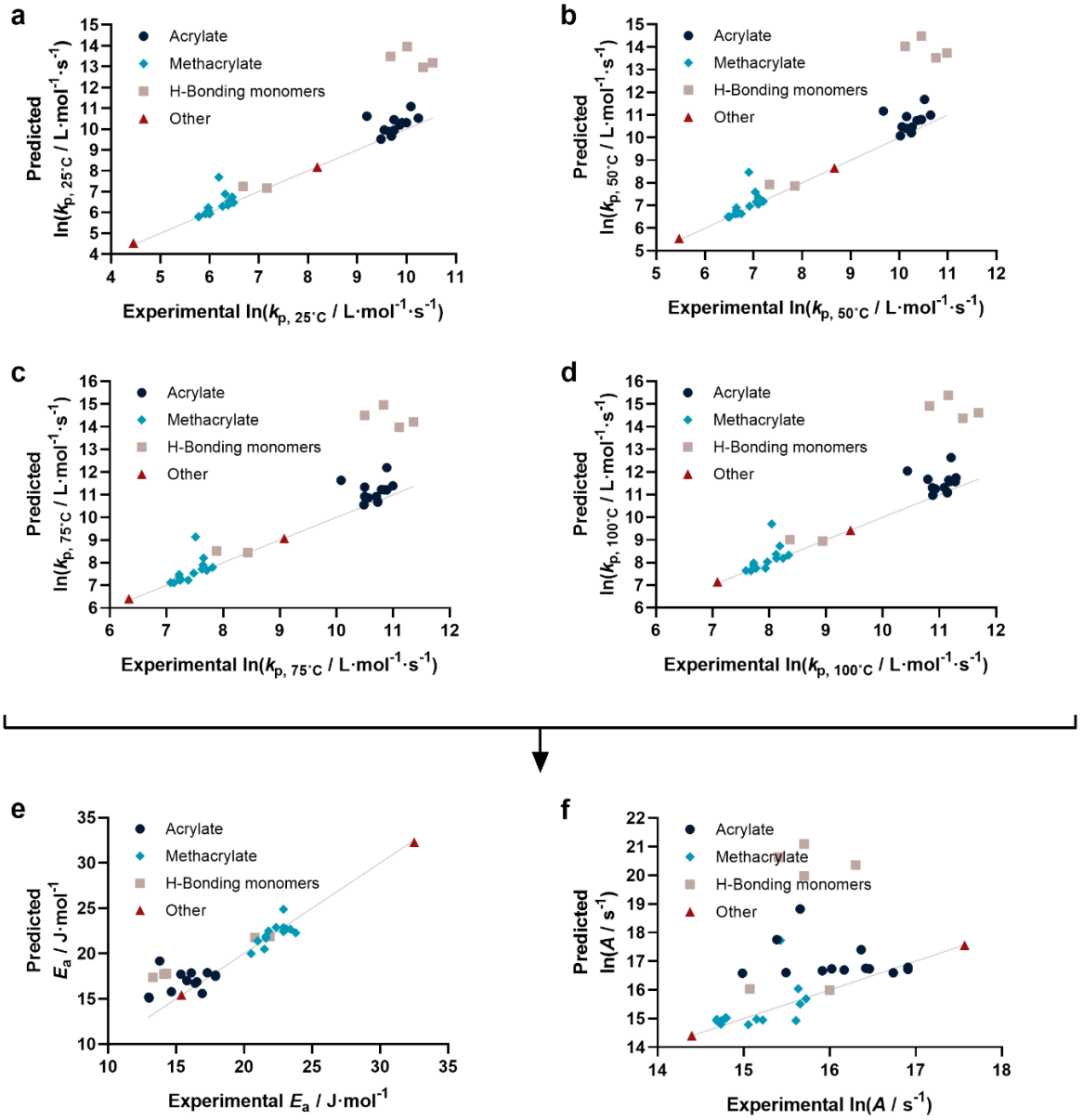


Figure 23: Correlation plots of the natural logarithm of the predicted propagation rate constant (ln(k_p)) versus the experimental ln(k_p) for different temperatures: 25˚C, 50˚C, 75˚C and 100˚C for a, b, c and d respectively. The predictions were generated via a cross validated Ridge regression using all available data (n = 35) with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, the dissociation constants and polarizability as parameters. A selection of monomers (n = 13) is given a weight of 1 while the rest (n = 22) has a weight of 0.5. From the predicted ln(k_p) results, a prediction for the activation energy (E_a) and the natural logarithm of the pre-exponential factor (ln(A)) was made using the Arrhenius plot and formula. The predictions were plotted against their experimental value (e and f). The metrics of each figure are annotated in the table below.

	Figure a	Figure b	Figure c	Figure d	Figure e	Figure f
RMSE	1.209	1.255	1.296	1.332	1.829	1.857
R2 Value	0.588	0.492	0.389	0.280	0.819	-5.294
Variance Predictions	6.356	5.959	5.637	5.374	12.590	3.149
Variance Residuals	1.109	1.190	1.266	1.337	2.700	2.594

Arrhenius plots of four selected monomers, values determined by a cross validated Ridge regression on All available ln(K_p) data [Molecular weight, a_value, R_value, H-bonding parameters, Dissociation constants, Chemspider parameters (Polarizability)]

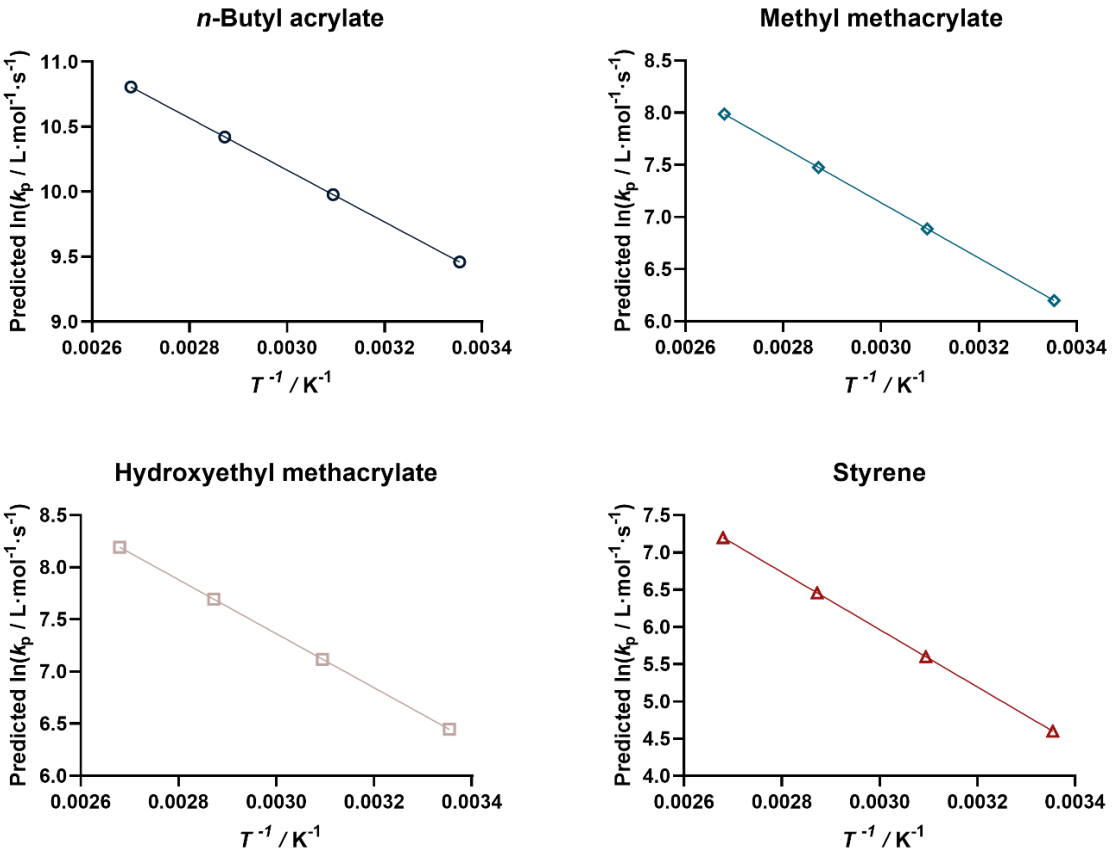


Figure 24: Arrhenius plots of n-Butyl acrylate, methyl methacrylate, hydroxyethyl methacrylate and styrene. The predicted ln(k_p) values were determined by a cross validated Ridge regression on all available ln(K_p) data with the molecular weight and a distinction between the type of monomer (blue = acrylate, teal = methacrylate, grey = H-binding monomers and red = other monomers), inductive effect of the tail group, the effect of H-donors and H-acceptors, the dissociation constants and polarizability as parameters

References

1. Asua, J.M., et al., *Critically Evaluated Rate Coefficients for Free-Radical Polymerization, 5.* Macromolecular Chemistry and Physics, 2004. **205**(16): p. 2151-2160.
2. Kockler, K.B., et al., *Determining Free-Radical Propagation Rate Coefficients with High-Frequency Lasers: Current Status and Future Perspectives.* Macromol Rapid Commun, 2016. **37**(2): p. 123-34.
3. Kockler, K.B., et al., *No Apparent Correlation of kp with Steric Hindrance for Branched Acrylates.* Macromolecular Chemistry and Physics, 2015. **216**(14): p. 1573-1582.
4. Barner-Kowollik, C., et al., *Critically evaluated rate coefficients in radical polymerization – 7. Secondary-radical propagation rate coefficients for methyl acrylate in the bulk.* Polymer Chemistry, 2014. **5**(1): p. 204-212.
5. Dervaux, B., et al., *Propagation rate coefficients of isobornyl acrylate, tert-butyl acrylate and 1-ethoxyethyl acrylate: A high frequency PLP-SEC study.* Journal of Polymer Science Part A: Polymer Chemistry, 2009. **47**(23): p. 6641-6654.
6. Barner-Kowollik, C., et al., *Detailed investigation of the propagation rate of urethane acrylates.* Polymer Chemistry, 2010. **1**(4): p. 470-479.
7. Beuermann, S. and M. Buback, *Rate coefficients of free-radical polymerization deduced from pulsed laser experiments.* Progress in Polymer Science, 2002. **27**(2): p. 191-254.
8. Haehnel, A.P., et al., *Global Trends for kp? Expanding the Frontier of Ester Side Chain Topography in Acrylates and Methacrylates.* Macromolecules, 2013. **46**(1): p. 15-28.
9. Beuermann, S., et al., *Critically evaluated rate coefficients for free-radical polymerization, 3. Propagation rate coefficients for alkyl methacrylates.* Macromolecular Chemistry and Physics, 2000. **201**(12): p. 1355-1364.
10. Beuermann, S., et al., *Critically Evaluated Rate Coefficients for Free-Radical Polymerization, 4.* Macromolecular Chemistry and Physics, 2003. **204**(10): p. 1338-1350.
11. Beuermann, S., et al., *Critically evaluated rate coefficients for free-radical polymerization, 2.. Propagation rate coefficients for methyl methacrylate.* Macromolecular Chemistry and Physics, 1997. **198**(5): p. 1545-1560.
12. Nitschke, A., et al., *Investigation into the Kinetics of n-Pentyl Methacrylate Radical Polymerization.* Macromolecular Chemistry and Physics, 2020. **221**(1): p. 1900345.
13. Junkers, T., S.P.S. Koo, and C. Barner-Kowollik, *Determination of the propagation rate coefficient of acrylonitrile.* Polymer Chemistry, 2010. **1**(4): p. 438-441.
14. Stach, M., et al., *Free-Radical Propagation Kinetics of N-Vinyl Formamide in Aqueous Solution Studied by PLP–SEC.* Macromolecular Chemistry and Physics, 2010. **211**(5): p. 580-593.
15. Stach, M., et al., *Propagation Rate Coefficient for Radical Polymerization of N-Vinyl Pyrrolidone in Aqueous Solution Obtained by PLP−SEC.* Macromolecules, 2008. **41**(14): p. 5174-5185.
16. Buback, M., et al., *Critically evaluated rate coefficients for free-radical polymerization, 1. Propagation rate coefficient for styrene.* Macromolecular Chemistry and Physics, 1995. **196**(10): p. 3267-3280.
17. Barner-Kowollik, C., et al., *Critically Evaluated Rate Coefficients in Radical Polymerization – 8. Propagation Rate Coefficients for Vinyl Acetate in Bulk.* Macromolecular Chemistry and Physics, 2017. **218**(1): p. 1600357.
18. Hansch, C. and T. Fujita, *p-σ-π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure.* Journal of the American Chemical Society, 1964. **86**(8): p. 1616-1626.
19. Santos-Martins, D. and S. Forli, *Charting Hydrogen Bond Anisotropy.* Journal of chemical theory and computation, 2020. **16**(4): p. 2846-2856.
20. Brocks, J.J., et al., *Estimation of Bond Dissociation Energies and Radical Stabilization Energies by ESR Spectroscopy.* The Journal of Organic Chemistry, 1998. **63**(6): p. 1935-1943.

21. Hoerl, A.E. and R.W. Kennard, *Ridge Regression: Biased Estimation for Nonorthogonal Problems.* Technometrics, 1970. **12**(1): p. 55-67.