## Analysis and prediction of defects in UV photo-initiated polymer microarrays – Electronic supplementary information

Andrew L. Hook<sup>a</sup>, David J. Scurr<sup>a</sup>, Jonathan C. Burley<sup>a</sup>, Robert Langer<sup>b</sup>, Daniel G. Anderson<sup>b</sup>, Martyn C. Davies<sup>a</sup>, Morgan R. Alexander<sup>a</sup>

<sup>a</sup> Laboratory of Biophysics and Surface Analysis, University of Nottingham, UK, NG7 2RD

<sup>b</sup> David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, USA, 02139

**Table ES11**List of top 20 molecular descriptors and the corresponding regression coefficient (RC) associated with promoting<br/>(positive RC) or preventing (negative RC) a particular defect for PLS models describing the total number of defects. Descriptors<br/>prefixed with ' $\Delta$ ' relate to the difference in the molecular descriptor of the monomer components, whereas non-prefixed descriptors<br/>relate to the average value of the molecular descriptors of the monomer components.

Positive regression coefficient		Negative regression coefficient		
Regression coefficient	Molecular descriptor	Regression coefficient	Molecular descriptor	
0.56	$\Delta khs.ssCH_2$	-1.11	$\Delta C_3 sp2$	
0.52	BCUTw.11	-0.97	∆topoShape	
0.50	$\Delta khs.sCH_3$	-0.38	khs.aaCH	
0.45	C <sub>3</sub> sp2	-0.38	∆BCUTw.1h	
0.45	HybRatio	-0.33	LipinskiFailures	
0.44	$\Delta BCUTw.11$	-0.30	$\Delta VC.4$	
0.33	ΔECCEN	-0.30	∆khs.ssO	
0.32	khs.sCH <sub>3</sub>	-0.26	$\Delta SC.4$	
0.31	VC.4	-0.24	$\Delta VC.3$	
0.30	$\Delta VP.1$	-0.24	C <sub>2</sub> sp3	
0.28	∆bpol	-0.24	SCH.5	
0.26	ΔBCUTp.1h	-0.24	khs.sssCH	
0.25	ΔnAtom	-0.24	VCH.5	
0.25	$\Delta$ WPATH	-0.23	BCUTc.1h	
0.24	$\Delta VP.4$	-0.22	$\Delta ATSc4$	
0.24	$\Delta VP.3$	-0.22	ΔnHBDon	
0.23	∆apol	-0.21	∆khs.dssC	
0.23	ΔKier3	-0.21	$\Delta$ SCH.3	
0.23	BCUTc.11	-0.21	SCH.6	
0.21	VC.3	-0.21	$\Delta$ LipinskiFailures	

Table ESI2	List of top 20 molecular descriptors and the corresponding regression coefficient (RC) associated with promoting
(positive RC) or pro-	eventing (negative RC) a particular defect for PLS models describing each defect. Descriptors prefixed with ' $\Delta$ '
relate to the differe	nce in the molecular descriptor of the monomer components, whereas non-prefixed descriptors relate to the
average value of th	e molecular descriptors of the monomer components.

Non-	Circular	St	oreading		Rough	Chemical	heterogeneity	Ľ	Diameter
Regression	Molecular	Regression	Molecular	Regression	n Molecular	Regression	Molecular	Regression	Molecular
coefficient	descriptor	coefficient	descriptor	coefficient	descriptor	coefficient	descriptor	coefficient	descriptor
0.20	C <sub>3</sub> sp2	0.26	$\Delta khs.ssCH_2$	0.38	ΔFMF	0.16	HybRatio	176	ΔBCUTw.11
0.12	$\Delta khs.ssCH_2$	0.23	HybRatio	0.18	BCUTw.11	0.15	$\Delta C_1 sp3$	145	C <sub>3</sub> sp2
0.10	∆BCUTw.11	0.23	khs.sCH3	0.16	ALogp2	0.14	ΔBCUTp.1h	121	khs.sCH3
0.08	ΔECCEN	0.21	$\Delta khs.sCH_3$	0.15	C <sub>3</sub> sp2	0.11	∆khs.dO	113	HybRatio
0.08	$\Delta VP.4$	0.20	VC.4	0.12	ALogP	0.08	BCUTc.11	111	ΔBCUTp.1h
0.08	$\Delta VP.1$	0.17	C <sub>3</sub> sp2	0.11	WPATH	0.08	∆BCUTw.11	101	BCUTw.11
0.07	$\Delta VP.3$	0.16	∆BCUTw.11	0.11	ΔECCEN	0.08	∆bpol	94	$\Delta khs.sCH_3$
0.06	∆bpol	0.14	BCUTw.11	0.11	$\Delta$ khs.ssCH <sub>2</sub>	0.07	$\Delta AMR$	75	$\Delta khs.ssCH_2$
0.06	khs.dssC	0.13	VC.3	0.10	$\Delta khs.sCH_3$	0.06	ΔnAtom	65	∆bpol
0.06	$\Delta VP.7$	0.12	∆Kier3	0.09	SC.4	0.06	∆apol	65	VC.4
0.06	BCUTw.11	0.12	C <sub>4</sub> sp3	0.09	BCUTc.11	0.06	BCUTw.11	59	$\Delta VP.4$
0.06	$\Delta khs.sCH_3$	0.11	BCUTc.11	0.09	khs.ssssC	0.06	$\Delta$ TopoPSA	59	C <sub>4</sub> sp3
0.06	∆khs.aaCH	0.11	∆bpol	0.08	SC.3	0.05	$\Delta nAtomLAC$	56	$\Delta VP.1$
0.05	khs.sCH₃	0.10	∆BCUTp.1h	0.08	SC.6	0.05	SC.5	56	$\Delta VP.5$
0.05	ΔnAtom	0.10	ΔnAtom	0.08	VC.3	0.05	SC.6	55	ΔnAtom
0.05	khs.ssO	0.10	$\Delta VP.1$	0.08	$\Delta VP.4$	0.05	khs.sssN	54	$\Delta VP.3$
0.05	∆apol	0.09	∆apol	0.08	SC.5	0.05	khs.sCH3	53	∆apol
0.05	$\Delta C_3 sp3$	0.08	$\Delta MLogP$	0.08	VC.4	0.04	$\Delta WTPT.4$	51	$\Delta ATSc3$
0.05	HybRatio	0.08	C <sub>1</sub> sp3	0.08	∆BCUTc.11	0.04	XLogP	51	C <sub>1</sub> sp3
0.05	$\Delta VP.5$	0.08	$\Delta VP.3$	0.08	∆BCUTc.1h	0.04	$\Delta khs.sCH_3$	51	$\Delta$ LipinskiFailures
0.00	40.0	0.42	10.2	0.07	A	0.01		1.42	A
-0.26	$\Delta C_3 sp_2$	-0.43	ΔC <sub>3</sub> sp2	-0.27		-0.21	ΔFMF	-143	
-0.24		-0.24	LipinskiFailures	-0.25	$\Delta C_3 sp_2$	-0.12	$\Delta C_3 sp_2$	-140	$\Delta C_3 sp_2$
-0.12		-0.21		-0.18	$\Delta KRS.dSSC$	-0.11	ADCUT::: 1h	-80	
-0.12	ADCUT-11	-0.16	Alleha and	-0.14	ADCUT: 1h	-0.11	Abbe and CU	-80	
-0.11	ABCUIW.In	-0.14	AKIIS.SSU	-0.14	ΔBCUIW.In	-0.10	AtomoShama	-19	$\Delta VC.5$
-0.09	∆nHB∆on AUD A	-0.12	VCH.5	-0.13		-0.10		-70	Δ50.4
-0.09		-0.11	SCH.5	-0.12	$C_2$ sp3	-0.09	DCUT- 11	-12	AALogp2
-0.07	$\Delta V C.4$	-0.11	ADCUTe 11	-0.12	$\Delta C_1 \text{sps}$	-0.08	BCUIC.In	-08	AAI SCS
-0.07	An Atom I AC	-0.11	ADCUIC.II	-0.11	AVDC 4	-0.08	$\Delta SC.4$	-03	AALOgP
-0.07	ΔΠΑΙΟΠΙLAC	-0.11	ΔBCUIC.III fragC	-0.09	DVPC.4	-0.08	$\Delta C_3 \text{sp5}$	-04	ASC 2
-0.07	C38P3	-0.11	hage CH	-0.09	ASPC 4	-0.08	ASC 2	-02	Alths dCH
-0.07	VCH 5	-0.11	SCU 6	-0.09	$\Delta SFC.4$ WTDT 2	-0.07		-01	Icho doCH
-0.07	Alpha dCH	-0.10	AC an2	-0.08	An AtomIAC	-0.07	ASF.2 EME	-00	C ap2
-0.00	$\Delta \text{KHS.uCH}_2$	-0.10	AVAdiMat	-0.08	khs aaCH	-0.07	AZagreb	-37	C <sub>3</sub> sp <sub>3</sub>
-0.00	khs.sssCH	-0.10	A A TSm2	-0.08	Alcha aE	-0.07		-55	ASD 2
-0.00		-0.10	AnAtomI C	-0.08	AVC 4	-0.00	WTDT A	-51	Akhe seO
-0.00	vr./ fragC	-0.09	AWTPT 4	-0.07	BCUTw 1b	-0.00	w 1F 1.4 kbs aaCU	-51	Akhs desC
-0.00	A A TSe <sup>2</sup>	0.09		-0.07	ATSc/	-0.00	Albe oF	-51	SCH 5
-0.00	AATSe4	-0.09	Δ3Γ.2	-0.07	SCU 6	-0.00	LAKIIS.SI	-49 _/19	SCП.5 VCH 5
-0.05	ΔA1504	-0.09	DAISCI	-0.07	SCH.0	-0.00	KIIS.SSU	-40	VСП.Э

	Monomer content (%) $\sim$ 8 8 8 8 8 8 8 9 0 0		
High	W 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$Ion = CF_{3}^{+}CF^{+}F_{2}^{-}CF^{-}HF_{2}^{-}CF_{3}^{-}C_{4}F_{3}^{-}$ $Ion = C_{4}H_{6}^{+}CH^{+}C_{2}H_{5}^{+}C_{2}H_{5}O^{+}C_{2}H_{3}^{-}C_{5}H_{5}O_{2}^{-}$ $Ion = C_{14}H_{32}NO^{+}$ $Ion = C_{3}H_{8}NO_{2}^{+}C_{5}H_{5}O^{-}$	
	x • • • • • • • • • • • • • • • • • • •	$Ion = CF_3^+ CF^+ F_2^- CF^- HF_2^- CF_3^-$ $Ion = C_3H_5O^+ C_2H_5^+ C_4H_6^+ C_2H_3^-$ $Ion = C_{14}H_{32}NO^+ C_5H_5O_2^-$	
Low	v <b>Construction</b> a	$Ion = C_3H_3O_2^-CH^-C_3H_5^-C_5H_5O^-$ $Ion = CH_3^+C_2H_4O^+C_3H_8^+C_3H_5O^+C_4H_6^+C_2H_5O^+C_5H_5O_2^-$	
	x	$lon = C_2H_3O^+ C_3H_5O^+ C_4H_6^+ C_2H_5O^+ C_2H_5^+ C_5H_5O_2^-$ lon = F_2^+ CF_3^+ CF_+ HF_2^+ C_3H_3O_2^+ CF_3^- C_4F_3^-	
	p • • • • • • • • • • • • • • • • • • •	$lon = \mathbf{F}_{2}^{-}C_{3}F_{3}^{-}C_{4}F_{3}^{-}C_{2}F^{-}CF_{3}^{-}CF^{-}HF_{2}^{-}$ $C_{3}F_{3}^{-}CHF^{+}CF^{+}C_{3}HF_{6}^{+}C_{2}F_{3}^{-}C_{5}F^{-}$ $lon = C_{2}H_{5}^{+}$	80 40 0
	p • • • • • • • • • • • • • • • • • • •	$ \begin{aligned} & \text{Ion} = \mathbf{C}_{6}\mathbf{H}_{15}\mathbf{O}_{3}^{+}  \mathbf{C}_{3}\mathbf{H}_{5}\mathbf{O}^{+} \\ & \text{Ion} = \mathbf{C}_{8}\mathbf{H}_{13}\mathbf{O}^{+}  \mathbf{C}_{5}\mathbf{H}_{5}\mathbf{O}_{2}^{-}  \mathbf{C}_{8}\mathbf{H}_{7}\mathbf{O}^{-}  \mathbf{C}_{4}\mathbf{H}_{5}\mathbf{O}^{+}  \mathbf{C}_{2}\mathbf{H}_{4}\mathbf{O}^{+}  \mathbf{C}_{8}\mathbf{H}_{11}\mathbf{O}^{+} \\ & \mathbf{C}_{3}\mathbf{H}_{5}^{-}  \mathbf{C}_{3}\mathbf{H}_{5}\mathbf{O}^{-}  \mathbf{C}_{5}\mathbf{H}_{7}\mathbf{O}^{-}  \mathbf{C}_{4}\mathbf{H}_{6}\mathbf{O}^{-}  \mathbf{C}_{7}\mathbf{H}_{7}\mathbf{O}_{3}^{-}  \mathbf{C}_{7}\mathbf{H}_{7}\mathbf{O}_{2}^{-} \end{aligned} $	
	r mr	$lon = C_9H_{14}O^+ C_4H_5O^+ C_9H_{11}O^+ C_5H_5O_2^- C_5H_7O^- C_3H_3O^-$ $lon = C_5H_5O^- C_3H_3O_2^- C_2H_3^- C_2H_3O^- C_4H_5^- C_2HO^- C_4H_6O^-$ $lon = PO_2^-$	
	h o o o o o o o	$Ion = C_{14}H_{36}O_2^+ C_2H_3O^+ C_3H_5O^+C_2H_3^- C_5H_5O_2^-$ $Ion = C_8H_7O^- C_4HO^- C_9H_{11}O_2^+ C_8H^-$	80 40 0
	q • • • • • • • • • • • • • • • • • • •	$Ion = C_7 H_{13}O^+ C_7 H_{12}O_3^+ C_{14}H_{30}O_2^- C_5 H_5O_2^-$ $Ion = C_9 H_{11}O_2^+ C_3 H_5O^- C_4 H_6O^- C_5 H_7O^- C_8 H_7O^-$ $Ion = C_{18}H_{26}O_3^+ C_4 H_{10}O_2^+ C_4O_2^-$	ensity 0 0 0 0 0 0 0 0 0 0 0 0 0
	b b b b b b b b b b b b b b b b b b b	$\begin{split} & \text{lon} = \mathbf{C_5}H_5\mathbf{O}^{\cdot} \ C_3H_3O_2^{\cdot} \ C_3H_3^{\cdot} \\ & \text{lon} = \mathbf{C_6}H_7\mathbf{O_4}^{\cdot} \ C_{10}H_{12}O_3^{\cdot} \ C_6HO_2^{\cdot} \ C_{10}H_{17}O_4^{\cdot} \ C_{15}H_5O_6^{\cdot} \ C_7H_4O_3^{\cdot} \\ & \text{lon} = \mathbf{CH_3O^{\cdot}} \end{split}$	
	0 0 0 0 0 0 0 0 k 0 0 0 0 0 0 0 k k k	$lon = C_2 H_3 \cdot C_2 O^{-}$ $lon = C_3 H_3 \cdot C_4 F_7 \cdot C_4 F_5 \cdot C_3 F_2 \cdot C_6 \cdot C_6 HO_2 \cdot C_{14} \cdot C_4 \cdot$ $lon = C_7 H_3 O_3 \cdot$	
		$lon = C_3H_3O_2^{-}C_5H_7O^{-}C_5H_5O^{-}C_8H_7O^{-}$ $lon = C_3H_3^{-}C_4F_5^{-}C_4F_7^{-}C_6H_5^{-}C_6HO_2^{-}C_7H_4O_3^{-}$ $lon = PO_3H^{-}$ $lon = C_3H_3O_2^{-}C_5H_5O_2^{-}C_7H_4O_3^{-}$	
	i i i i i i i i i i i i i i i i i i i	$Ion = C_3H_3O_2 \cdot C_6 \cdot C_2H_3 \cdot C_2HO \cdot C_2 \cdot C_3 \cdot C_4 \cdot Ion = C_3H_4O \cdot C_8H_7O \cdot C_6H_6 \cdot Ion = C_3H_4O_3 \cdot Ion = C_3H_4O_3 \cdot C_6H_6 \cdot Ion = C_3H_4O_3 \cdot Ion = C_3H_4O_3$	
	q b q b q b b b	$Ion = CH_3O^{-}C_7H_{11}O_3^{-}C_6H_5^{-}C_3H_3^{-}C_3H_3O_2^{-}C_5H_5^{-}$ $Ion = C_9H_5O_2^{-}$ $Ion = C_2^{-}C_3H_2^{-}C_3H^{-}C_3^{-}C_4^{-}C_4H_2^{-}C_6H_6^{-}$	
	q b b b c c c c c c c c c c c c c c c c	$Ion = C_8H_5O_3^{-1}$ $Ion = C_4F_3^{-1}C_3F_2^{-1}C_3H_3^{-1}C_6^{-1}C_8H_2F_2^{-1}C_4^{-1}C_4H_3O^{-1}$ $Ion = CI^{-1}$	
	i <b>n</b> i <b>n</b> i <b>n</b> i <b>n</b>	$Ion = C_4H^{-}C_2^{-}C_3^{-}C_4^{-}C_6H^{-}$ $Ion = C_5H_{11}O_3^{-}C_6H_5^{-}C_5H_7O^{-}$ $Ion = C_8H_9O_3^{-}$	
	e <b>e e e e e e e e e e e e e e e e e e </b>	$Ion = C_6H_{11}O_3 \cdot C_5H_7O^{-1}$ $Ion = C_4F_3 \cdot C_3F_2 \cdot C_6H_7 \cdot C_3H_3 \cdot C_3HO_2 \cdot CF^{-1}C_3H_2O_2 \cdot C_8H_2F_2^{-1}$	80 40 0
	t s t s t s s	$Ion = C_7H_7O_3^{-}C_{11}H_9O_2^{-}C_{11}H_8^{-}C_6H_5^{-}C_3H_7O^{-}C_3H_3O_2^{-}C_8H_7O^{-}$ $Ion = C_6F_{11}^{-}C_4F_7^{-}C_6F_7^{-}C_4F_5^{-}C_6F_8^{-}C_3F_5^{-}$ $Ion = C_3HO_2^{-}$	
	f f r r r f f r r r f source boots of r f to the source of r r f to the source of the source of r	$Ion = C_5H_5O^{-}C_8H_9O_3^{-}C_2H_3^{-}C_4H_2^{-}$ $Ion = C_4H_6^{-}C_8H_8O^{-}$ $Ion = C_4^{-}C_6H_3O^{-}C_3H_3^{-}C_6H_5^{-}C_3F_2^{-}CH_3O^{-}CF_2^{-}C_4HO^{-}CF_3^{-}$ $Ion = PO_5^{-}$	
	$\begin{array}{c} \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ \\ \end{array} \\ \hline \text{Monomer content (\%)} \end{array} $		0 50 100 Monomer content (%)

**Fig. ESI1** ToF-SIMS images of each copolymer pair within the odd spot array with various ions mapped. The intensity scale is shown on the left. For each image the intensity was scaled to maximise the contrast for each ion, thus, the scale is arbitrary for each image. All ions that gave a similar image are shown to the right of each image. The intensity of the ion in bold is shown in the chemical maps/graphs.



**Fig. ES12** PLS regression model prediction of the occurrence of defects (a-b) non-circular, (c-d) spreading, (e-f) roughness, and (g-h) chemical heterogeneity. Model calculated for the (a, c, e, g) training set and (b, d, f, h) test set. The number of correct predictions (positive = defect present, negative = defect absent) is (a) 97.7%, (b) 92.9%, (c) 94.6%, (d) 88.1%, (e) 90.7%, (f) 90.5%, (g) 90.7, (h) 81.0%.



**Fig. ESI3** PLS regression model prediction of the (a-b) diameter and (c-d) roughness of polymer spots, calculated for the (a, c) training set and (b, d) test set.  $R^2$  value = (a) 0.42, (b) -0.01, (c) -1.50 and (d) -2.29.



**Fig. ESI4** Root mean square error of cross validation (RMSECV) curves for PLS models. (a) First model produced with 244 x-variables. (b) Second model produced with 166 x-variables. A minimum of the RMSECV curve at the lowest number of latent variables (7) was used to determine the number of latent variables to be used for the formation of a model.