Supplementary information

Characterization of degradation products of regorafenib by LC-QTOF-MS and NMR: Investigation of rearrangement and odd-electron ion formation during collision-induced dissociations under ESI-MS/MS

Shandilya Mahamuni Baira,[a] G. Srinivasulu,[a] Rakesh Nimbalkar,[b] Prabha Garg,[b] R. Srinivas,[ac] M.V.N Kumar Talluri*[a]

[a] Department of Pharmaceutical Analysis, National Institute of Pharmaceutical Education & Research, Balanagar, Hyderabad, Telangana, India

[b] Department of Pharmacoinformatics, National Institute of Pharmaceutical Education & Research, S.A.S Nagar, Mohali, Punjab, India

[c] NCMS, CSIR-Indian Institute of Chemical Technology, Hyderabad, Telangana, India

*[a] Corresponding author: Dr. M.V.N. Kumar Talluri, Department of Pharmaceutical Analysis, National

Institute of Pharmaceutical Education & Research (NIPER), Hyderabad, Telangana, India

Tel: +91-40-23074750, Ext: 2012, Fax: +91-40-23073751.

E-mail: narendra.talluri@gmail.com, narendra@niperhyd.a

Captions to Figures of Supplementary Information

Figures

Tables

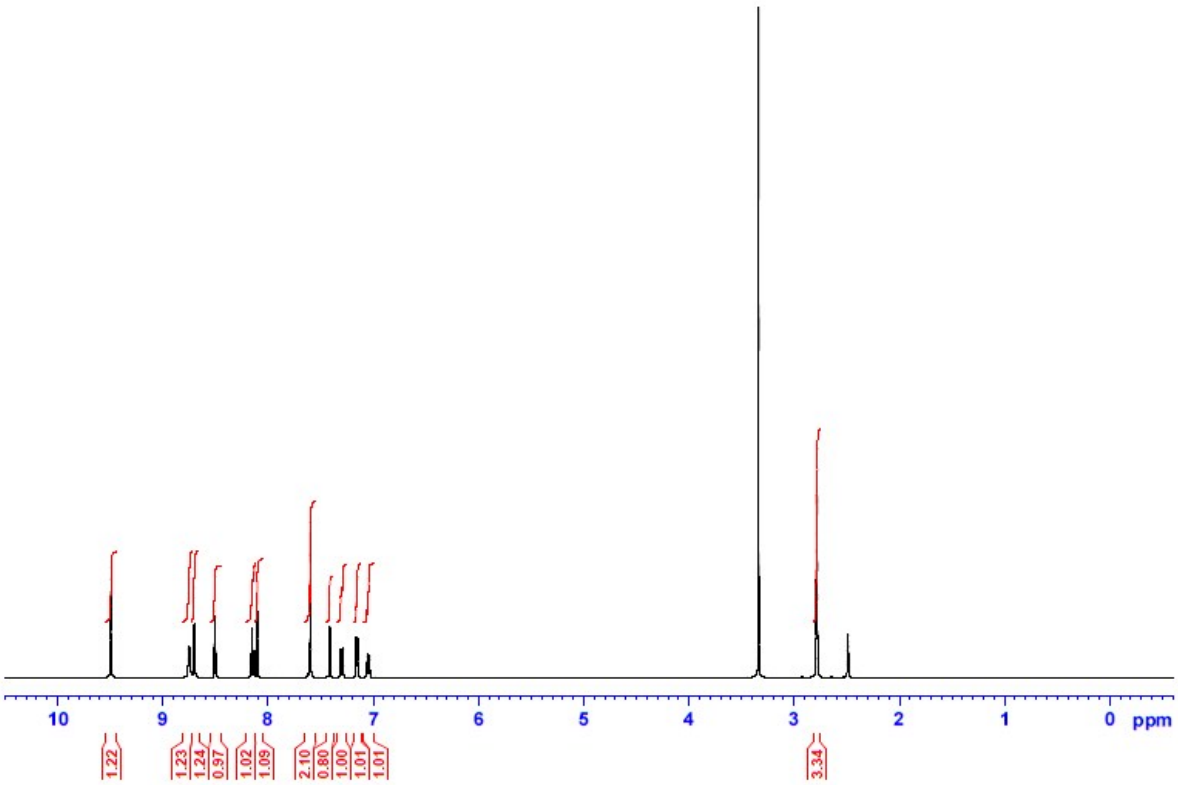


Fig. S1. ^{1}H NMR spectra of REG (500 MHz, DMSO-d_6)

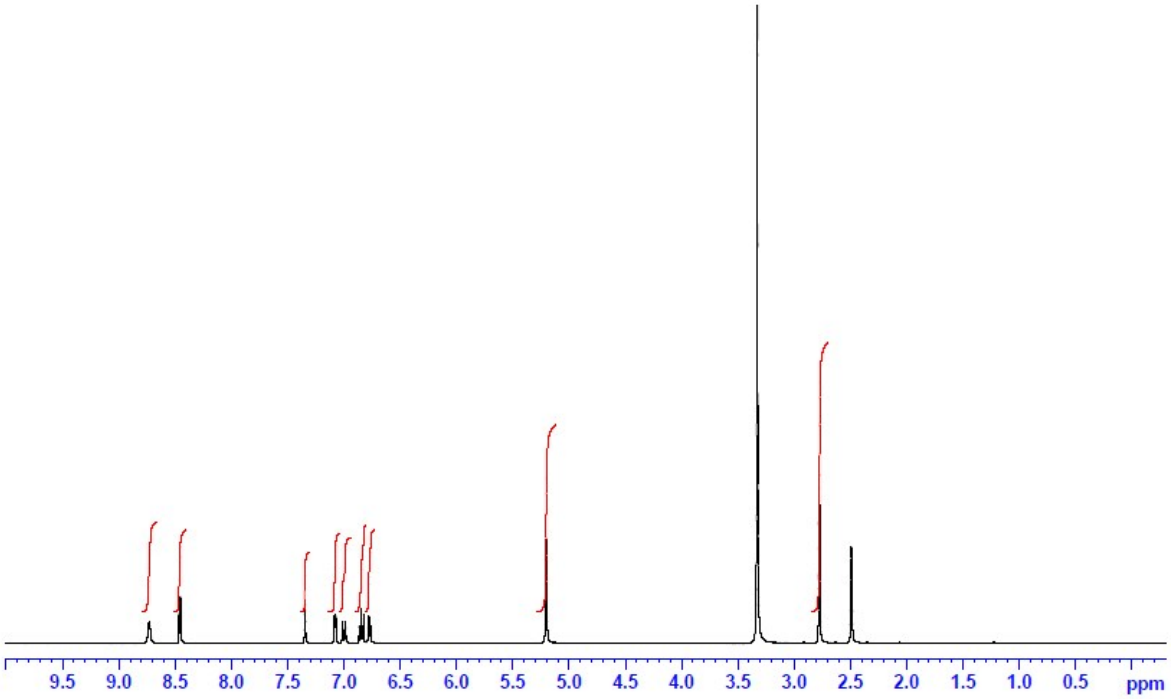


Fig. S2. ^{1}H NMR spectra of DP1 (500 MHz, DMSO-d_6)

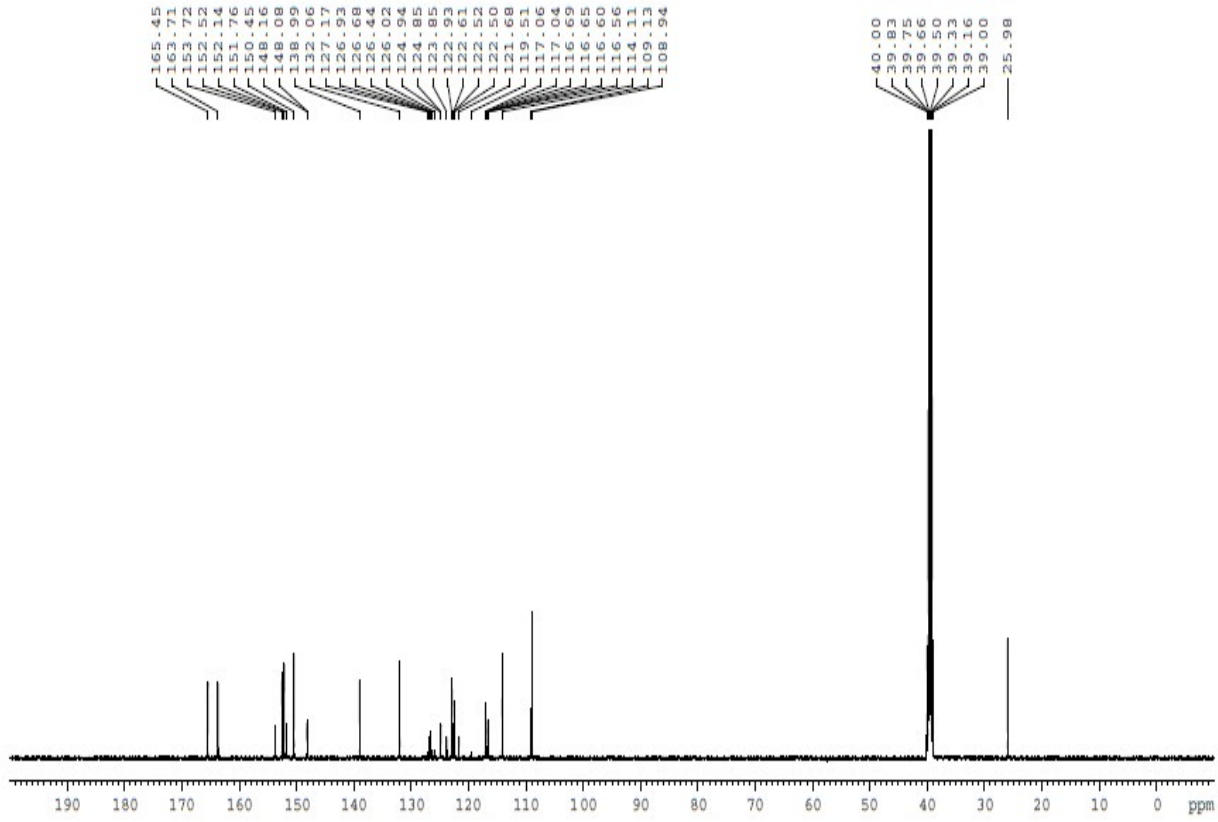


Fig. S3. ^{13}C NMR spectra of REG (500 MHz, DMSO-d_6)

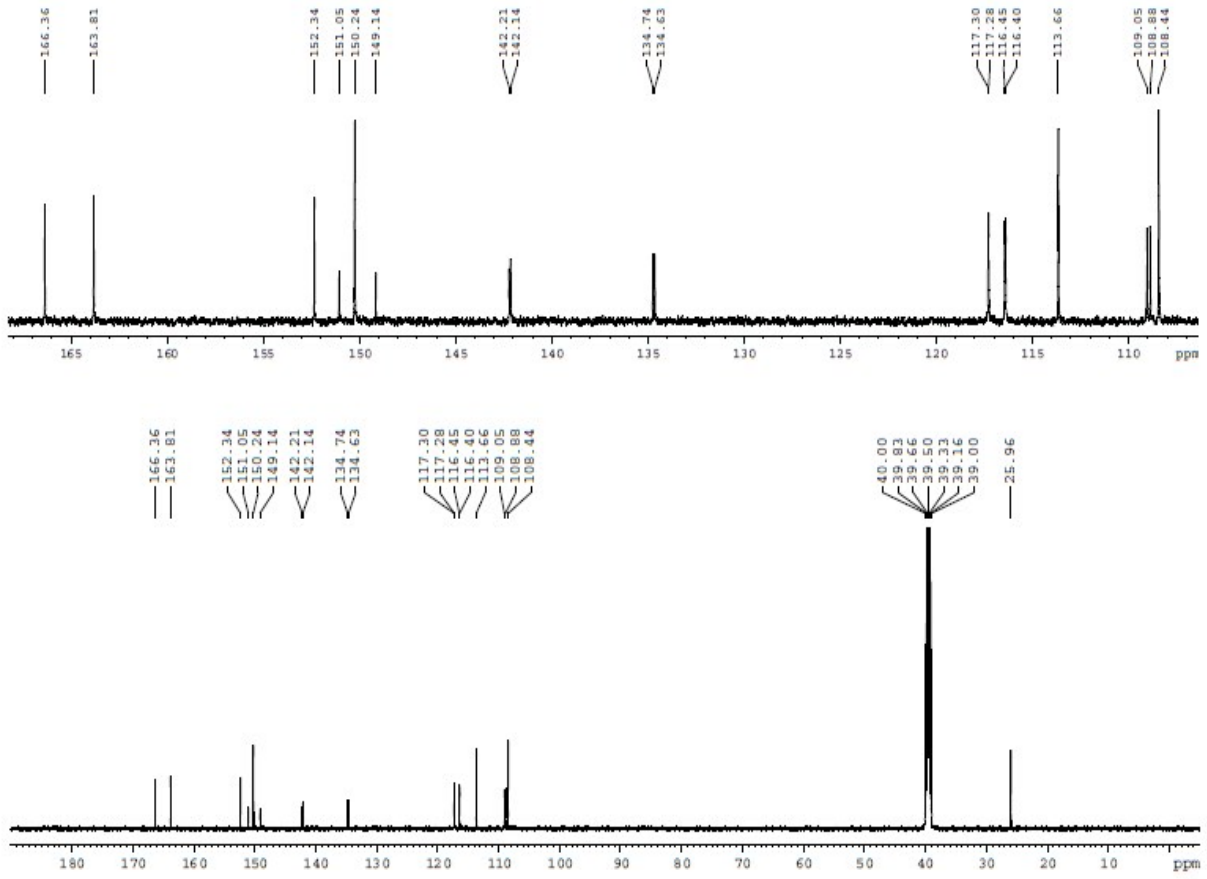


Fig. S4. ^{13}C NMR spectra of DP1 (500 MHz, DMSO-d_6)

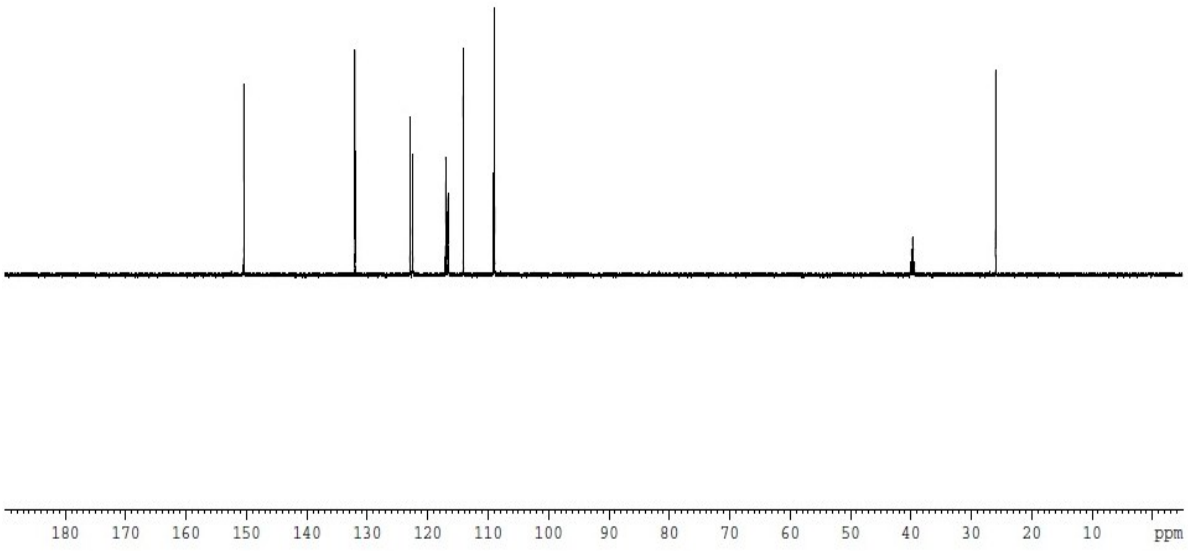


Fig. S5. DEPT135 NMR spectra of REG (500 MHz, DMSO-d_6)

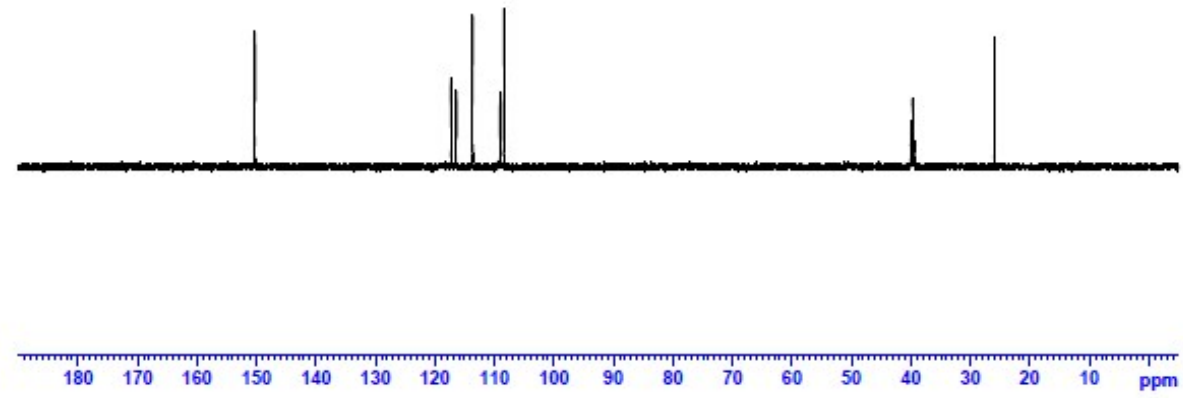


Fig. S6. DEPT135 NMR spectra of DP1 (500 MHz, DMSO-d_6)

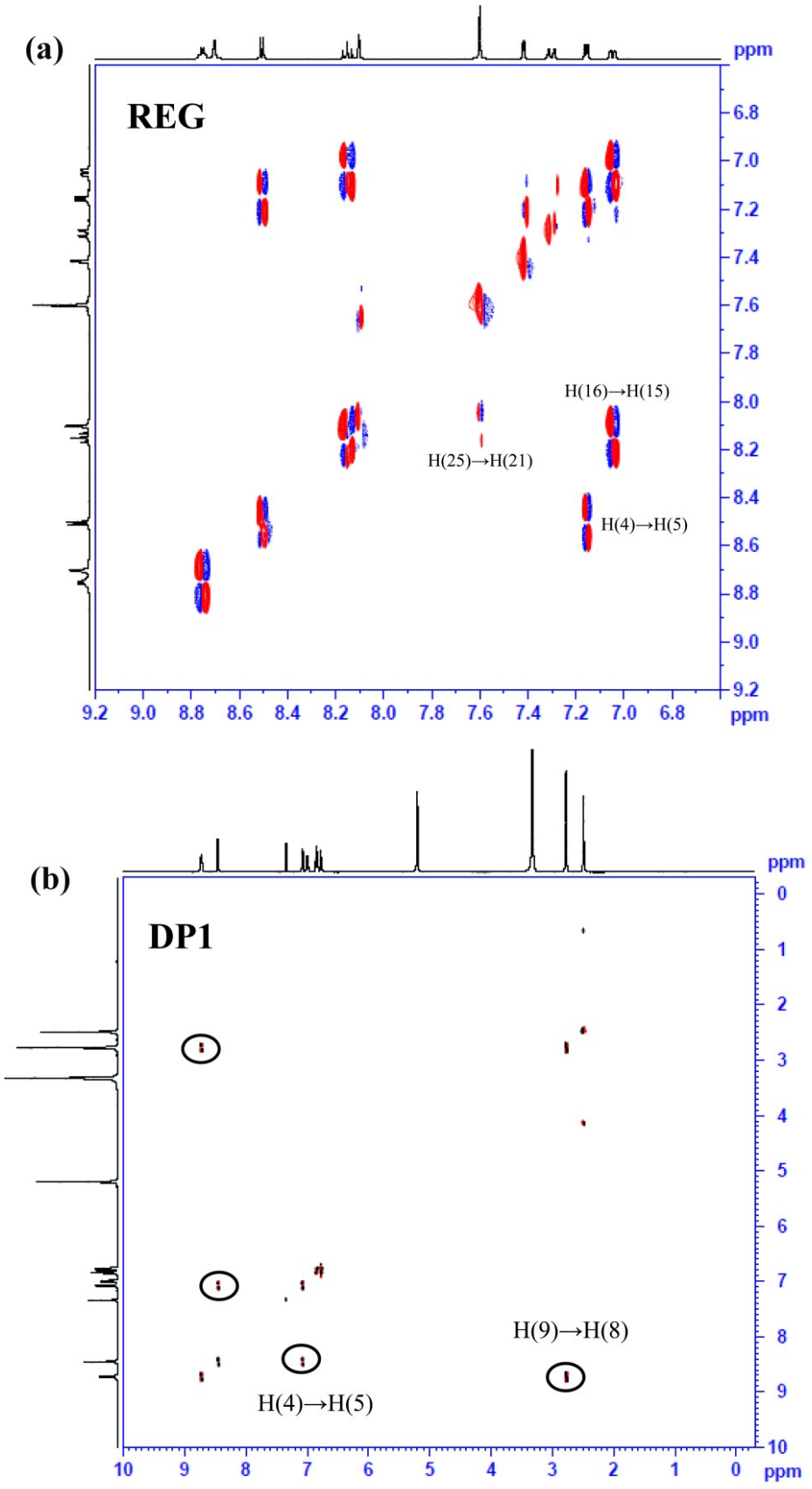


Fig. S7. COSY spectra of REG and DP1 (500 MHz, DMSO-d6)

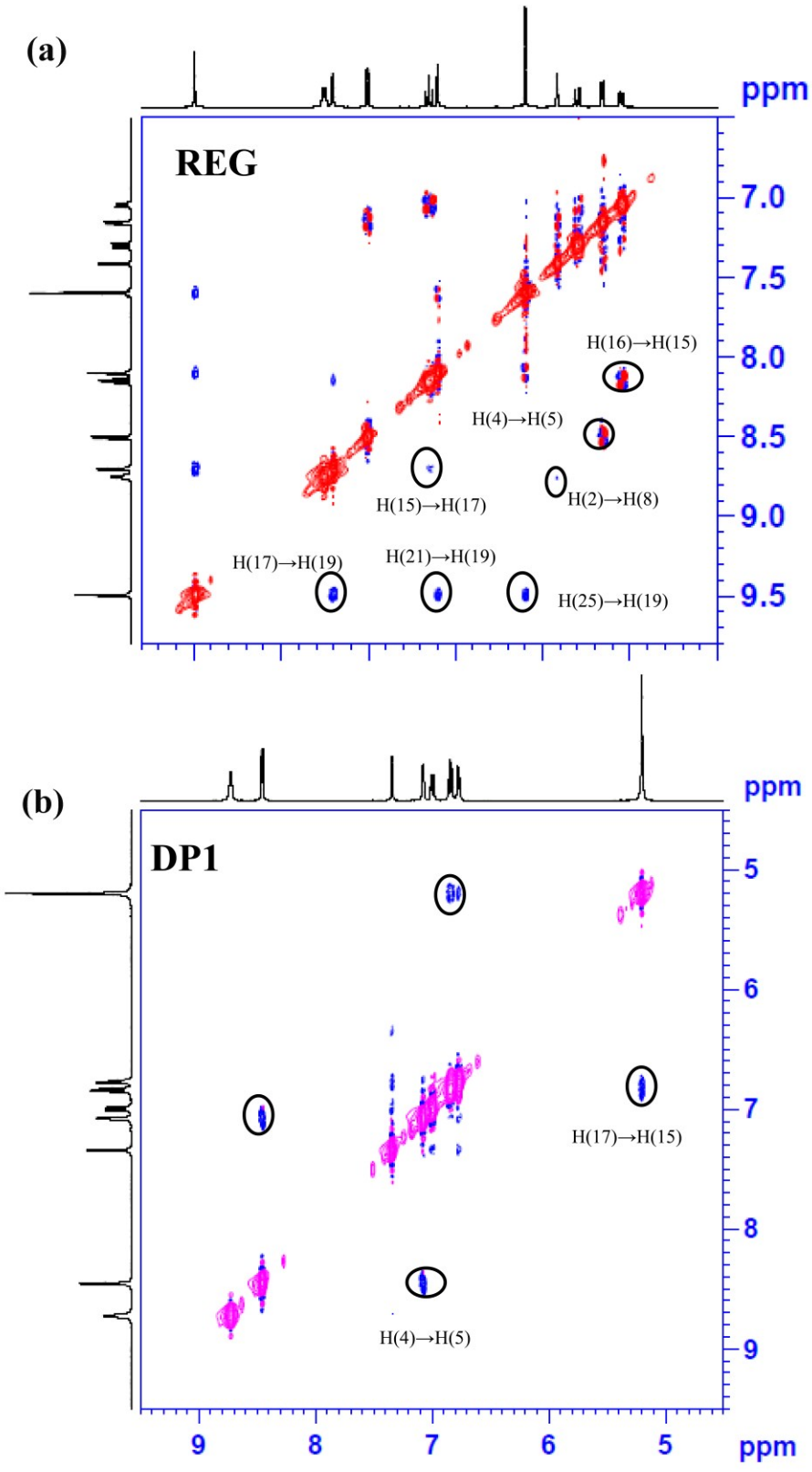


Fig. S8. ROESY spectra of REG and DP1 (500 MHz, DMSO-d6)

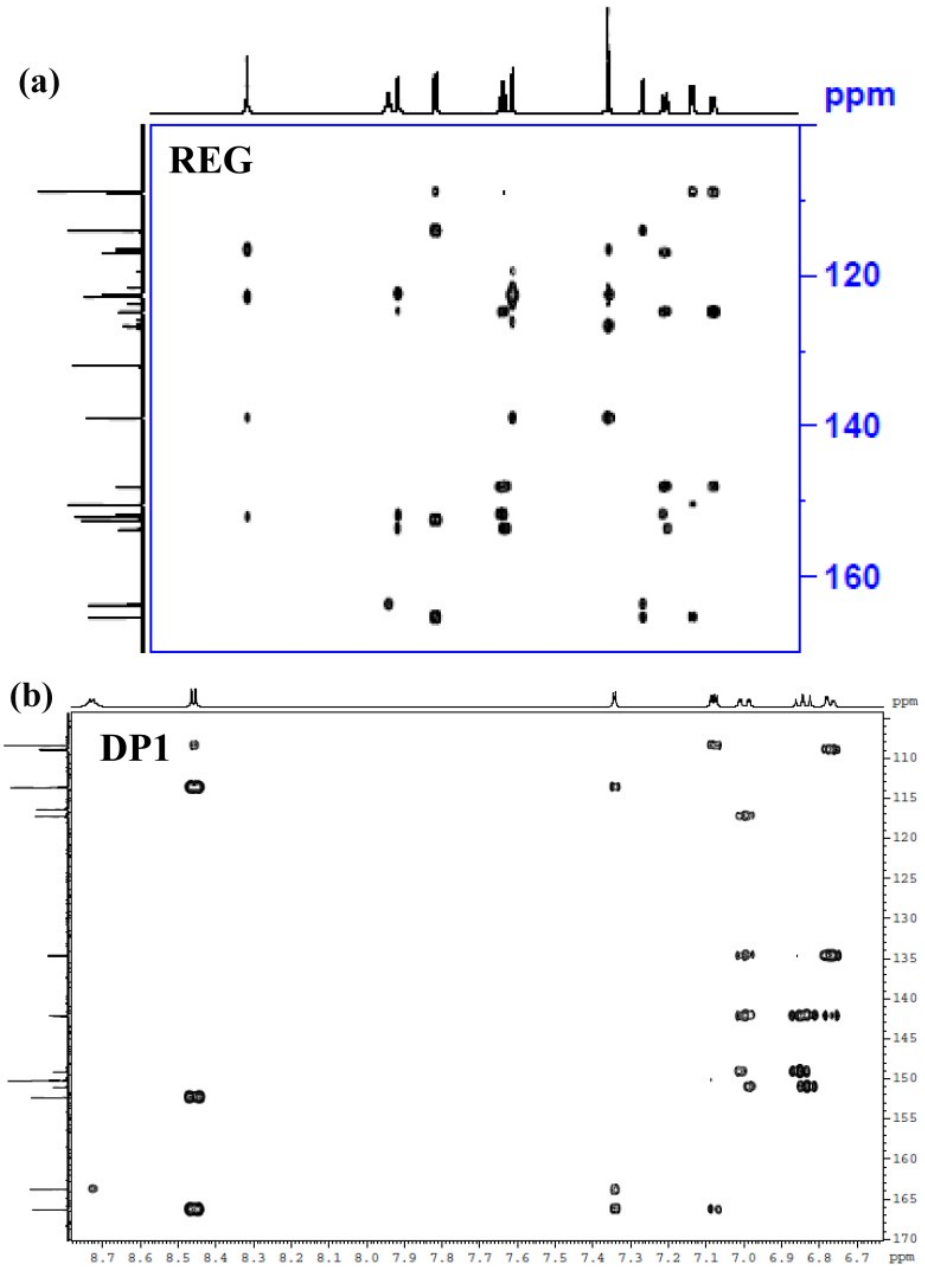


Fig. S9. HMBC spectra of REG and DP1 (500 MHz, DMSO-d6)

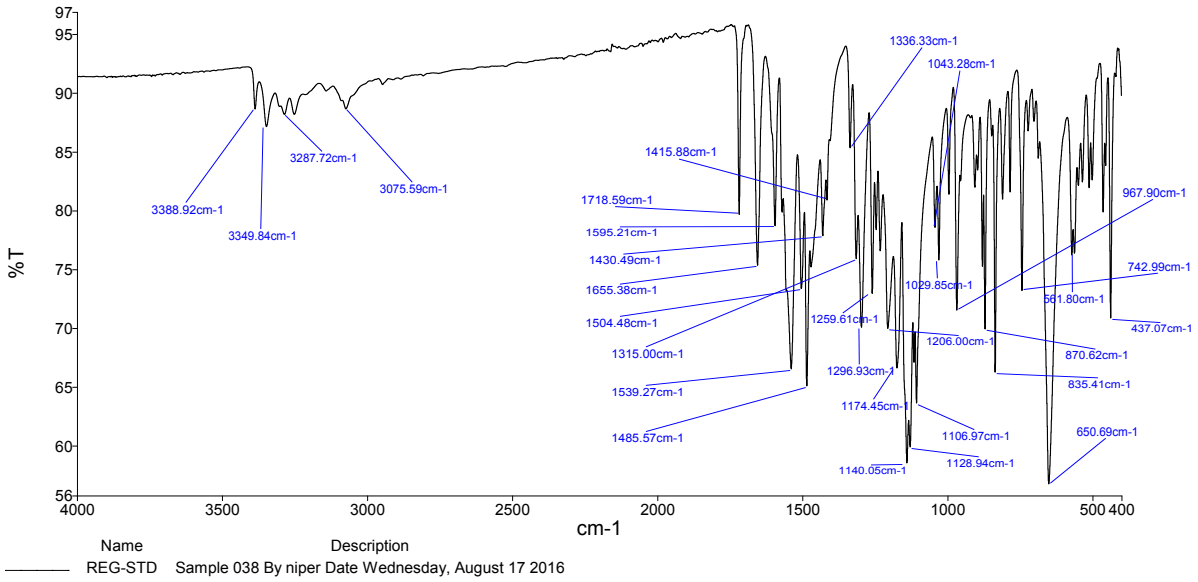


Fig. S10. IR spectra of REG (4000–400 cm^{-1})

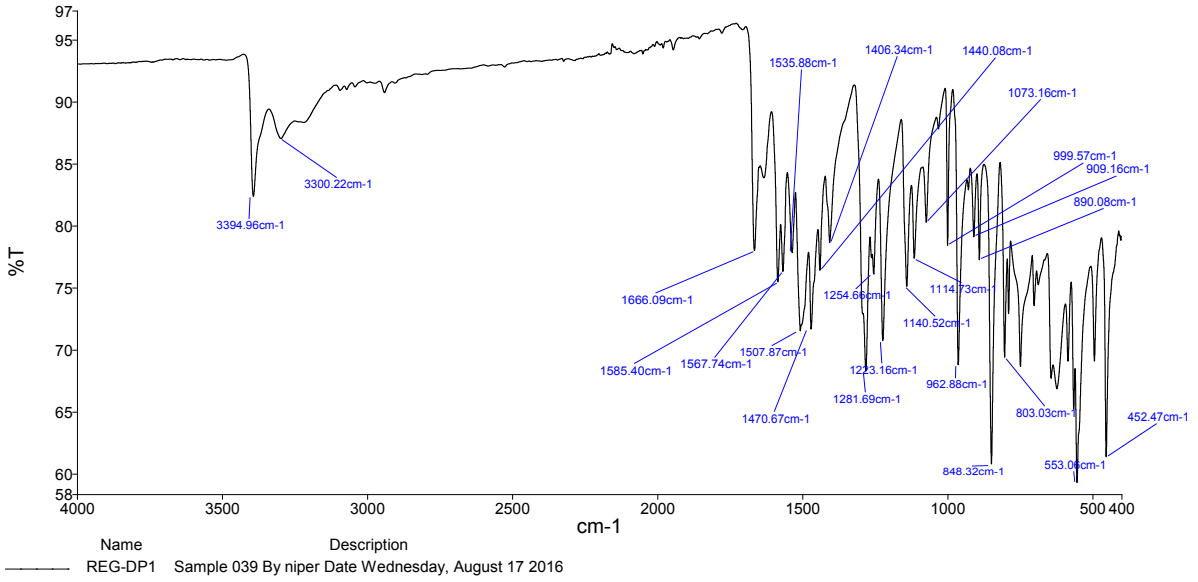


Fig. S11. IR spectra of REG (4000–400 cm^{-1})

Table S1 High resolution mass spectrometry (HRMS) data of product ions of protonated regorafenib and its degradation products

REG and DPs	Proposed formula	Calculated mass	Observed mass	Error (ppm)
REG-483	$C_{21}H_{16}ClF_4N_4O_3^+$	483.0842	483.0849	-1.45
	$C_{21}H_{14}ClF_4N_4O_2^+$	465.0736	465.0747	-2.37
	$C_{21}H_{14}ClF_2N_4O_3^+$	443.0717	443.0719	-0.45
	$C_{19}H_{11}ClF_4N_3O_2^+$	424.047	424.0479	-2.12
	$C_{19}H_{10}ClF_3N_3O_2^+$	404.0408	404.041	-0.49
	$C_{14}H_{11}FN_3O_3^+$	288.0779	288.0783	-1.39
	$C_{14}H_9FN_3O_2^+$	270.0673	270.0676	-1.11
	$C_{13}H_{13}FN_3O_2^+$	262.0986	262.0986	0.00
	$C_{13}H_{11}FN_3O^+$	244.0881	244.0881	0.00
	$C_{12}H_6FN_2O_2^+$	229.0408	229.0406	0.87
	$C_8H_3ClF_2NO^+$	201.9866	201.9865	0.50
	$C_9H_6ClN_2O^+$	193.0163	193.0162	0.52
DP1-262	$C_{13}H_{13}FN_3O_2^+$	262.0986	262.0979	2.67
	$C_{13}H_{11}FN_3O^+$	244.0881	244.0884	-1.23
	$C_{11}H_8FN_2O^+$	203.0615	203.0618	-1.48
	$C_{10}H_8FN_2^+$	175.0666	175.0666	0.00
	$C_{10}H_7N_2^+$	155.0604	155.0607	-1.93
	$C_9H_7FN^+$	148.0557	148.0559	-1.35
	$C_6H_7FNO^+$	128.0506	128.0498	6.25
	$C_6H_5FNO^+$	126.035	126.0353	-2.38
	$C_6H_6FN^{\bullet+}$	111.0479	111.0483	-3.60
	$C_8H_5^+$	101.0386	101.0389	-2.97
	$C_5H_5FN^+$	98.0401	98.0403	-2.04
	$C_6H_5^+$	77.0386	77.0388	-2.60
	$C_4H_3^+$	51.0229	51.0232	-5.88
DP2-196	$C_7H_6ClF_3N^+$	196.0135	196.0138	-1.53
	$C_7H_3ClF_3^+$	178.987	178.9867	1.68
	$C_7H_6F_3N^{\bullet+}$	161.0447	161.0448	-0.62
	$C_7H_6F_2N^+$	142.0463	142.0462	0.70
	$C_6H_6ClN^{\bullet+}$	127.0183	127.0183	0.00
	$C_7H_3F_2^+$	125.0197	125.0198	-0.80
	$C_6H_6FN^{\bullet+}$	111.0479	111.0479	0.00
	$C_5HF_2^+$	99.0041	99.0043	-2.02
	$C_6H_6N^+$	92.0495	92.0495	0.00
	$C_3HF_2^+$	75.0041	75.0041	0.00
	$C_6H_2^{\bullet+}$	74.0151	74.0152	-1.35
	CF_3^+	68.9947	68.9948	-1.45
	$C_5H_5^+$	65.0386	65.0388	-3.08
DP3-349	$C_{14}H_{10}ClF_4N_2O_2^+$	349.0361	349.0353	2.29

	$C_8H_4ClF_3NO^+$	221.9928	221.9936	-3.60
	$C_7H_6ClF_3N^+$	196.0135	196.014	-2.55
	$C_7H_5ClF_2N^+$	176.0073	176.0071	1.14
	$C_6H_3ClF_3^+$	166.987	166.9871	-0.60
	$C_7H_6F_3N^{\bullet+}$	161.0447	161.0449	-1.24
	$C_7H_5FNO_2^+$	154.0299	154.0303	-2.60
	$C_7H_5F_2N^{\bullet+}$	141.0385	141.0377	5.67
	$C_6H_7FNO^+$	128.0506	128.0504	1.56
	$C_6H_5FNO^+$	126.035	126.0353	-2.38
	$C_6H_6NO^+$	108.0444	108.0442	1.85
	$C_5H_4FO^+$	99.0241	99.0237	4.04
	$C_5H_4F^+$	83.0292	83.029	2.41
	$C_4H_5^+$	53.0386	53.0384	3.77
DP4-304	$C_{15}H_{15}FN_3O_3^+$	304.1092	304.1099	-2.30
	$C_{15}H_{13}FN_3O_2^+$	286.0986	286.0987	-0.35
	$C_{14}H_9FN_3O^+$	254.0724	254.0725	-0.39
	$C_{13}H_{10}FN_2O_2^+$	245.0721	245.0723	-0.82
	$C_{13}H_8FN_2O^+$	227.0615	227.0621	-2.64
	$C_{11}H_8FN_2O^+$	203.0615	203.0617	-0.98
	$C_{11}H_7N_2O^+$	183.0553	183.0553	0.00
	$C_{10}H_8FN_2^+$	175.0666	175.0665	0.57
	$C_{10}H_7N_2^+$	155.0604	155.0606	-1.29
	$C_9H_7FN^+$	148.0557	148.0556	0.68
	$C_6H_7FNO^+$	128.0506	128.0499	5.47
	$C_6H_5FNO^+$	126.035	126.035	0.00
	$C_6H_6FN^{\bullet+}$	111.0479	111.0481	-1.80
	$C_5H_5FN^+$	98.0401	98.0404	-3.06
DP5-499	$C_{21}H_{16}ClF_4N_4O_4^+$	499.0791	499.0798	-1.40
	$C_{21}H_{14}ClF_4N_4O_3^+$	481.0685	481.068	1.04
	$C_{21}H_{13}ClF_3N_4O_3^+$	461.0623	461.0638	-3.25
	$C_{19}H_{13}ClF_4N_3O_3^+$	442.0576	442.0565	2.49
	$C_{19}H_{11}ClF_4N_3O_2^+$	424.047	424.0474	-0.94
	$C_{19}H_{10}ClF_3N_3O_2^+$	404.0408	404.0407	0.25
	$C_{19}H_{10}F_4N_3O_2^+$	388.0704	388.0703	0.26
	$C_{19}H_9F_3N_3O_2^+$	368.0641	368.0644	-0.82
	$C_{14}H_{11}FN_3O_4^+$	304.0728	304.0731	-0.99
	$C_{13}H_{13}FN_3O_3^+$	278.0935	278.0939	-1.44
	$C_{12}H_8FN_2O_3^+$	247.0513	247.0513	0.00
	$C_{12}H_6FN_2O_2^+$	229.0408	229.0407	0.44
	$C_8H_3ClF_2NO^+$	201.9866	201.9861	2.48
	$C_{10}H_5FNO^+$	174.035	174.0352	-1.15
	$C_6H_8F_2NO^+$	148.0568	148.056	5.40
	$C_6H_5FNO^+$	126.035	126.0345	3.97
	$C_6H_6FN^{\bullet+}$	111.0479	111.0478	0.90

Table S2 ^{1}H NMR chemical shifts (δ in ppm) and ^{13}C NMR chemical shifts (δ in ppm) for DP 1 and REG

REG					
Atom number	**^{1}H**	**ppm/J**	**^{13}C**	**DEPT**	**HSQC**
REG					
1	-	-	152.5	-	
2	1H	7.41/*d* 2.6	108.9	CH	CH
3	-	-	165.4	-	
4	1H	7.15/*dd* 2.6, 5.6	114.1	CH	CH
5	1H	8.50/*d* 5.6	150.4	CH	CH
6	-	-	-	-	
7	-	-	163.7	-	
8	1H	8.75/*q* 5.0	-	-	
9	3H	2.78/*d* 5.0	25.9	CH3	CH
10	-	-	-	-	
11	-	-	148.2, 148.1/*d*	-	
12	1H	7.30/*dd* 2.7, 11.5	109.1, 108.9/*d*	CH	CH
13	-	-	153.7, 151.8/*d*	-	
14	-	-	124.9, 124.8/*d*	-	
15	1H	8.15/*t* 9.0	122.6, 122.5/*d*	CH	CH
16	1H	7.05/*dd* 1.3, 2.7, 9.0	117.1, 117.0/*d*	CH	CH
17	1H	8.70/*d* 2.0	-	-	
18	-	-	152.1	-	
19	1H	9.49/s	-	-	
20	-	-	138.9	-	
21	1H	8.10/*d* 1.6	116.7,116.6, 116.6, 116.6/*q*	CH	CH
22	-	-	127.2, 126.9, 126.9,126.4/*q*	-	
23	-	-	122.5	-	
24	1H	7.65-7.75/	122.9	CH	CH
25	1H	7.65-7.75/	132.1	CH	CH
26	-	-	126.0, 123.8, 121.7, 119.5/*q*	-	
DP1					
1	-	-	152.3	-	
2	1H	7.34	108.4	CH	CH
3	-	-	166.4	-	
4	1H	7.07	113.7	CH	CH
5	1H	8.45	150.2	CH	CH
6	-	-	-	-	
7	-	-	163.8	-	
8	1H	8.72	-	-	
9	3H	2.77	26.0	CH3	CH
10	-	-	-	-	
11	-	-	134.7, 134.6/*d*	-	
12	1H	6.99	109.0, 108.9/*d*	CH	CH
13	-	-	151.0, 149.1/*d*	-	
14	-	-	142.2, 142.1/*d*	-	
15	1H	6.84	116.4, 116.4/*d*	CH	CH
16	1H	6.77	117.3, 117.3/*d*	CH	CH
17	2H	5.20	-	-	

Table S3 Probability values of different toxicity models of regorafenib (REG) and its degradation products by TOPKAT analyses

MODEL	DP1	DP2	DP3	DP4	DP5	PD (REG)
Rat_Male_NTP	0.052	0.716	0.017	0.001	0.000	0.000
Ames_Mutagenicity	0.000	0.162	0.000	0.000	0.012	0.002
Rat_Female_NTP	0.998	0.000	0.000	0.001	0.000	0.123
Mouse_Male_NTP	1.000	0.001	1.000	1.000	1.000	1.000
Mouse_Female_NTP	1.000	0.000	0.835	1.000	1.000	1.000
Rat_Male_FDA_None_vs_ Carcinogen	0.000	0.152	1.000	0.000	1.000	1.000
Rat_Male_FDA_Single_vs_ Multiple_Prediction	0.999	1.000	1.000	0.001	1.000	1.000
Rat_Female_FDA_None_vs_ Carcinogen	0.002	0.000	0.003	0.002	0.001	0.001
Rat_Female_FDA_Single_vs _Multiple_Prediction	1.000	1.000	1.000	1.000	1.000	1.000
Mouse_Male_FDA_None_vs _Carcinogen	0.367	0.799	0.999	0.004	0.969	0.662
Mouse_Male_FDA_Single vs Mult	0.000	0.000	0.000	0.000	0.000	0.000
Mouse_Female_FDA_None_ vs_Carcinogen	0.990	0.000	0.000	0.508	0.000	0.000
Mouse_Female_FDA_Single vs Mult	0.000	0.000	0.000	0.000	0.000	0.000
Weight_of_Evidence_Rodent _Carcinogenicity	0.999	0.997	1.000	0.991	0.000	0.991
Developmental_Toxicity_Pot ential	1.000	0.429	0.004	1.000	0.997	1.000
Rat_Oral_LD50	1.5 g/kg	1.0 g/kg	2.5 g/kg	2.0 g/kg	178.5 mg/kg	112.5 mg/kg
Rat_Maximum_Tolerated_ Dose_Feed	623.5 mg/kg	58.2 mg/kg	1.6 g/kg	72.3 mg/kg	188.7 mg/kg	56.9 mg/kg
Rat_Maximum_Tolerated_ Dose_Gavage	1.7 g/kg	58.2 mg/kg	4.3 g/kg	200.2 mg/kg	522.4 mg/kg	157.6 mg/kg
Rat_Inhalational_LC50	10 g/m3/H	10 g/m3/H	10 g/m3/H	10 g/m3/H	10 g/m3/H	10 g/m3/H
Chronic_LOAEL	44.9 mg/kg	353.6 mg/kg	96.7 mg/kg	98.8 mg/kg	7.0 mg/kg	4.4 mg/kg
Skin_Irritancy_None_vs_ Irritant	0.003	0.278	0.999	0.000	1.000	1.000
Skin_Sensitization_None_vs _Sensitizer	0.000	1.000	0.000	0.000	0.000	0.000
Skin_Sensitization_Weak_vs _Strong	1.000	0.002	0.000	1.000	1.000	1.000
Ocular_Irritancy_Mild_vs_ Moderate_Severe	1.000	0.514	0.919	1.000	0.996	1.000
Ocular_Irritancy_Moderate_ vs_Severe	0.000	0.000	0.000	0.001	0.000	0.000
Ocular_Irritancy_None_vs_ Irritant	0.998	0.992	1.000	1.000	1.000	1.000
Aerobic_Biodegradability	0.000	0.000	0.000	0.000	0.000	0.000
Fathead_Minnow_LC50	497.7 ug/l	100.6 mg/l	437.5 ug/l	105.9 ug/l	1.7 ug/l	3.0 ug/l
Daphnia_EC50	436.3 ug/l	668.2 ug/l	16.7 ug/l	1.5 mg/l	3.5 ug/l	6.6 ug/l

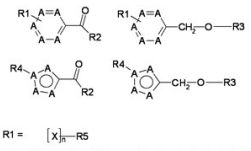
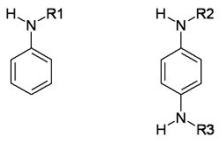
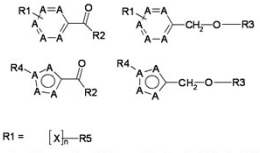
NTP: National Toxicological Program

FDA: Food and Drug Administration

LOAEL: Lowest Observed Adverse Effect Level

PD: Parent Drug

Table S4 Qualitative toxicity prediction of regorafenib (REG) and its degradation products by DEREK analysis

Regorafenib and its degradation products	Carcinogenicity	Skin Sensitization	Peroxisome Proliferation
Structural alert	Alkylaryl or bisaryl carboxylic acid or precursor  R1 = [X]n—R5 where X is any atom (but cannot be in a ring) attached at the meta or para position and that atom A at the point of attachment must be a carbon. R5 is an aromatic carbon in a 5 or 6 membered ring. n = 1, 2 or R1 = a chain of 7 atoms of the form A-A-C-C-C-C-C R2 = H, OR, NRR, where R= H, C R3 = H, C=O, P R4 = R1 but no meta or para position requirement	Aromatic primary or secondary amine  (I) (II) R1 = H, CH3, CH2CH3, CH2CH=CH2, CH2C≡CH R2, R3 = H, C Amino substituted coumarins are excluded	Alkylaryl or bisaryl carboxylic acid or precursor  R1 = [X]n—R5 where X is any atom (but cannot be in a ring) attached at the meta or para position and that atom A at the point of attachment must be a carbon. R5 is an aromatic carbon in a 5 or 6 membered ring. n = 1, 2 or R1 = a chain of 7 atoms of the form A-A-C-C-C-C-C R2 = H, OR, NRR, where R= H, C R3 = H, C=O, P R4 = R1 but no meta or para position requirement
Comments	The alert also covers ester, amide, aldehyde and alcohol precursors of the acid, and carbon and phosphorus ester derivatives of alcohol precursors. This alert covers peroxisome proliferators characterized by a carboxylic acid group attached to an aromatic ring with a lipophilic alkyl group or another aromatic ring arranged in a more or less linear arrangement which could be comparable in shape to a long chain fatty acid. Peroxisome metabolism of the carboxylic acid is blocked by the aryl ring.	This alert describes the skin sensitisation of aromatic amines and their N-protonated forms according to the toxicophores shown. In order to elicit a sensitisation response aromatic amines require transformation to a species capable of reacting with a skin protein nucleophilic group. Three key mechanisms through which this could be achieved have been postulated, all of which may have a role to play. The presence of a skin sensitisation structural alert within a molecule indicates the molecule has the potential to cause skin sensitisation. Whether or not the molecule will be a skin sensitiser will also depend upon its percutaneous absorption. Generally, small lipophilic molecules are more readily absorbed into the skin and are therefore more likely to cause sensitisation.	The alert also covers ester, amide, aldehyde and alcohol precursors of the acid, and carbon and phosphorus ester derivatives of alcohol precursors. This alert covers peroxisome proliferators characterized by a carboxylic acid group attached to an aromatic ring with a lipophilic alkyl group or another aromatic ring arranged in a more or less linear arrangement which could be comparable in shape to a long chain fatty acid. Peroxisome metabolism of the carboxylic acid is blocked by the aryl ring.
REG	*√	**NA**	*√
DP1	*√	√	*√
DP2	**NA**	√	**NA**
DP3	**NA**	**NA**	**NA**
DP4	*√	**NA**	*√
DP5	*√	**NA**	*√

***√:** Indicates carcinogenicity & peroxisome proliferation in Mouse and Rat.

√: Indicates skin sensitization in all species.