Electronic Supplementary Material (ESI) for Analyst.
This journal is © The Royal Society of Chemistry 2023

Electronic supplementary information (ESI)

Raman Classification of Selected Subtypes of Acute Lymphoblastic Leukemia (ALL)

Adriana Adamczyk,[a] Anna M. Nowakowska,[a] Justyna Jakubowska,[b] Marta Zabczynska,[b] Maja Bartoszek,[a] Sviatlana Kashyrskaya,[a] Agnieszka Fatla,[a] Kacper Stawoski,[a] Kacper Siakala,[a] Agata Pastorczak,[b] Kinga Ostrowska,[b] Wojciech Mlynarski,[b] Katarzyna Majzner,[a*] and Malgorzata Baranska [a,d]

[a.] *Faculty of Chemistry, Jagiellonian University, 2 Gronostajowa St., 30-387 Krakow, Poland.*
[b.] *Department of Pediatric, Oncology and Hematology, Medical University of Lodz, Sporna St. 36/50 Lodz, Poland*
[c.] *Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, 14 Bobrzynskiego St., 30-348 Krakow, Poland*

*Correspondence: katarzyna.b.majzner@uj.edu.pl (K.M.)

Table S1 Band assignments based on the literature studies and measured Raman spectra of pure compounds.

Bands / cm^{-1}	Assignment	Reference
674	C=C-O-C deformation in lipids	1
680	Guanine ring breathing mode in nucleosides	2
702	Cholesterol	1
718, 725	$N^+(CH_3)_3$ group vibrations in phospho- and sphingolipids	1
730	A	
745	Tryptophane	3
756	Porphyrin ring breathing mode in haemoproteins	4,5
790	Ring breathing mode of U, T, C in the DNA and RNA	6
820	Phospho-L-serine	Fig. S1
825	o-phospho-L tyrosine	Fig. S1
828	O-P-O stretching in DNA, phosphodiesters, and tyrosine out-of-plane breathing mode	4,7
842	C-C stretching mode of glucose	8
833, 851	Tyrosine doublet	4, Fig. S1
856	L-serine	
833, 860	Tryptophane doublet	4
867, 915 and 970	RNA	6
860, 890	C-O-O skeletal modes	1

875	Asymmetric $N^+(CH_3)_3$ stretching in phosphatidylcholine	1
938	C-C stretching of α-helix in proteins	4
950, 955	N-C_α-C vibration of proteins	4
955	Tryptophane and valine	4
1006	Ring breathing of phenylalanine of proteins	7,9
1024	C-N stretching of proteins	6
1076	C-C stretching of fatty acids	1
1093	O-P-O stretching of nucleic acids and phospholipids	4,6
1130	C-C, C-O, and C-N stretching of proteins fatty acids and haemoproteins	1,4,7
1143	C-N stretch in proteins , glucose oxidase	4
1168	C-C stretching in fatty acids	1
1180	C, G	10
1242-1227	Amide III, β-sheet	2,4
1250-1240	Amide III, unordered secondary structure	
1272-1264	Amide III, α-helix	
1260	C=C-H deformation of unsaturated lipids	1,7
1304	Twisting mode of $(CH)_2$ in lipids	1,7,11
1312	Ring breathing mode of tryptophane	4
1336	A and G breathing mode of DNA	2,6,7
1345	A	2
1340, 1360	Tryptophan doublet	4
1368	G	2
1379	T, A, G	2,6
1455	CH_2 deformation of lipids and proteins	4,6,7,11
1492	Breathing mode of G and A	7
1580-1590	Pyrimidine ring; C=C bending in Phe, nucleic acids	2,4,7,11
1645-1657	Amide I, α-helix	2,4
1660-1665	Amide I, unordered secondary structure	
1660	Amide I, C=C stretching	1,7,11
1665-1680	Amide I, β-sheet	2,4

A-adenine, C-cytidine, G-guanine, T-thymidine, U-uridine

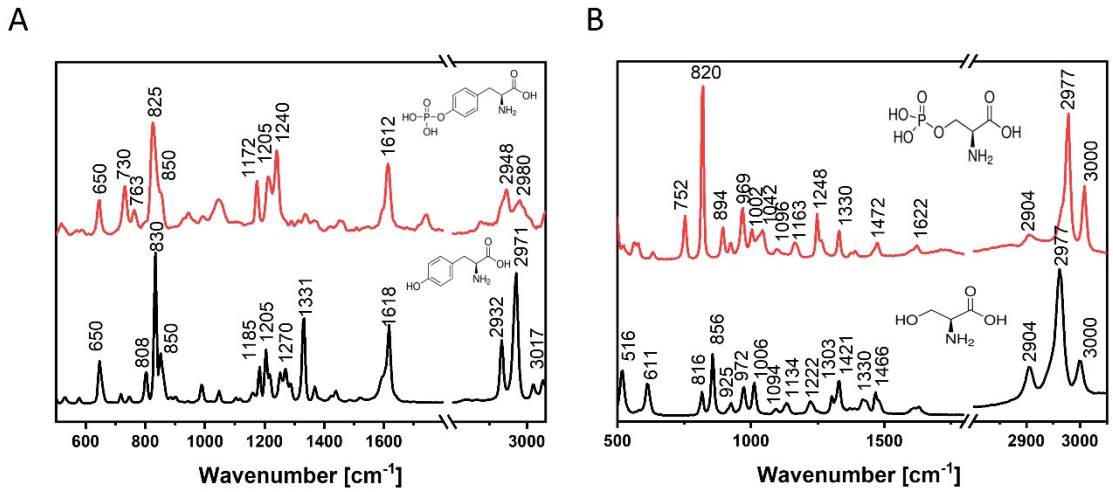

Figure S1. Raman spectra of solid (A) L-tyrosine and o-phospho-L-tyrosine and (B) L-serine and phospho-L-serine collected with a 532 nm excitation line with an integration time of 3 s and 10 accumulations. Compounds were purchased from Sigma-Aldrich (St. Louis, USA).

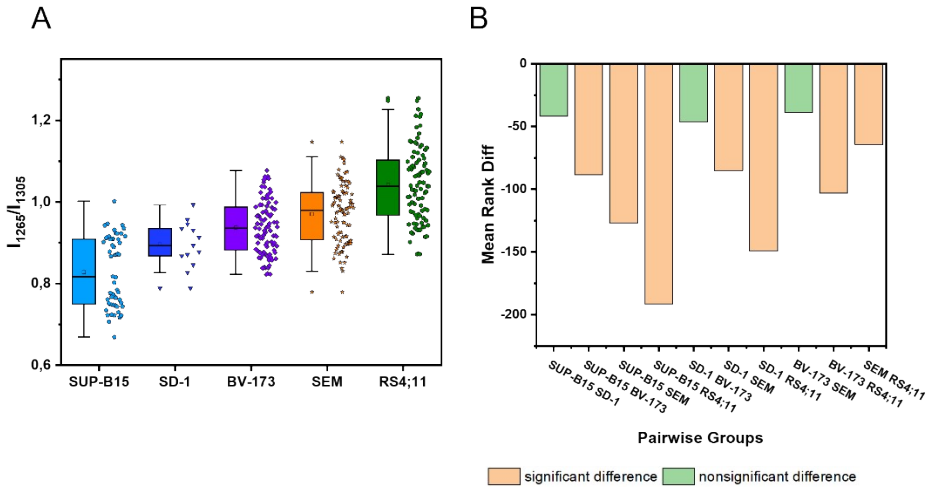

Figure S2. Band intensity ratio of 1260 cm-1 to 1306 cm-1 (A) for the average spectra of the lipid-rich classes. Values are given as median ± 1.5 SD (whiskers) and are shown in box plots: median (horizontal line), 25-75 percentile values (box), 1.5 SD (whiskers). (B) Parwise group difference significance plot based on the statistical Kruskal-Wallis ANOVA test.

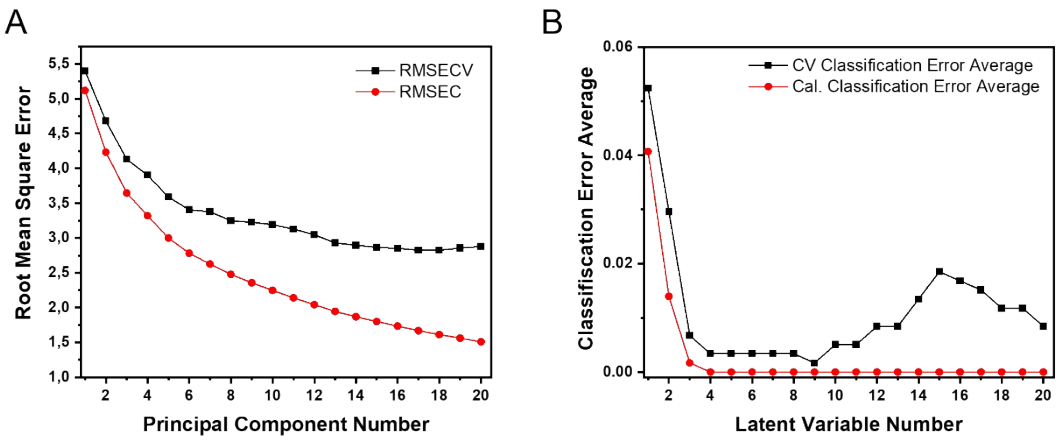

Figure S3. (A) Root mean square error of calibration (RMSEC) and cross-validation (RMSECV) in PCA model and (B) classification error of calibration and cross PLS-DA models, respectively, of spectra of B-ALL cell lines and B cells.

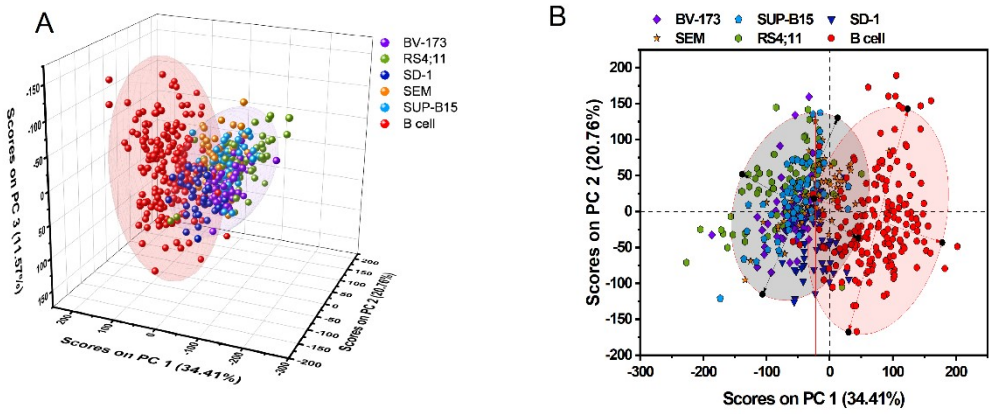

Figure S4. (A) 3D and (B) 2D score plots of PCA showing a homogenous distribution of different cell lines within the BCP-ALL cells.

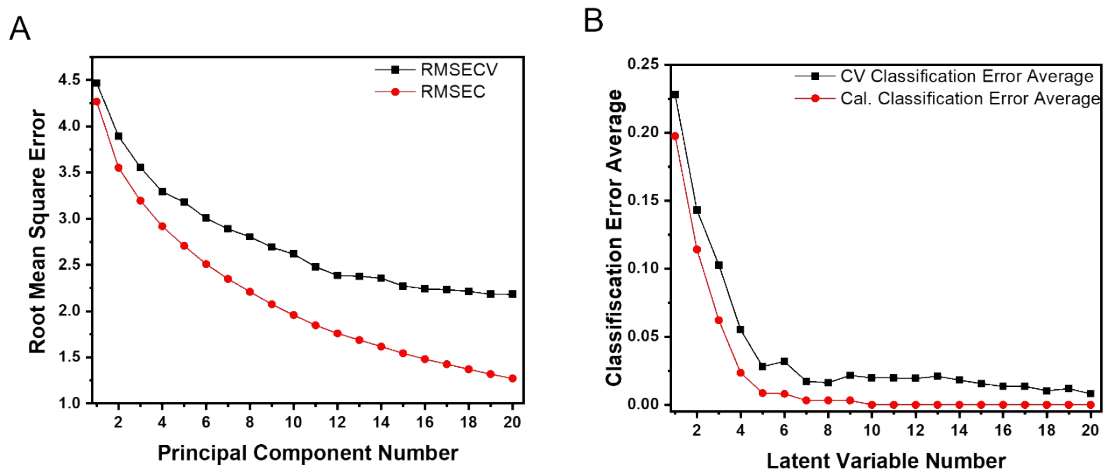

Figure S5. (A) Root mean square error of calibration (RMSEC) and cross-validation (RMSECV) in PCA model and (B) classification error of calibration and cross PLS-DA models, respectively, respectively, of *KMT2A* and *BCR-ABL1* - positive cell lines.

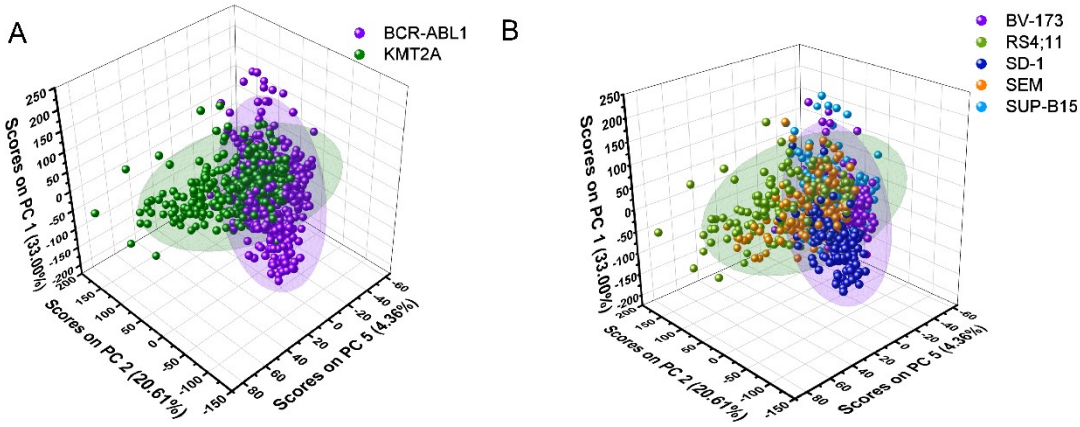

Figure S6. 3D PCA score plot for PC1, PC2 and PC5 with A) leukaemia subtypes marked B) leukaemia cell lines marked and 95% confidence ellipse.

References

1 K. Czamara, K. Majzner, M. Pacia, K. Kochan, A. Kaczor and M. Baranska, *J. Raman Spectrosc.*, 2014, **46**, 4–20.

2 C. Krafft, in *Encyclopedia of Analytical Chemistry*, 2018, pp. 1–15.

3 N. Huang, M. Short, J. Zhao, H. Wang, H. Lui, M. Korbelik and H. Zeng, *Opt. Express*, 2011, **19**, 22892–22909.

4 A. Rygula, K. Majzner, K. M. Marzec, A. Kaczor, M. Pilarczyk and M. Baranska, *J. Raman Spectrosc.*, 2013, **44**, 1061–1076.

5 H. R. Adams, C. Krewson, J. E. Vardanega, S. Fujii, T. Moreno, Chicano, Y. Sambongi, D. Svistunenko, J. Paps, C. R. Andrew and M. A. Hough, *Chem. Sci.*, 2019, **10**, 3031–3041.

6 J. W. Chan, D. S. Taylor, T. Zwerdling, S. M. Lane, K. Ihara and T. Huser, *Biophys. J.*, 2006, **90**, 648–656.

7 P. Leszczenko, A. Borek-Dorosz, A. M. Nowakowska, A. Adamczyk, S. Kashyrskaya, J. Jakubowska, M. Ząbczyńska, A. Pastorczak, K. Ostrowska, M. Baranska, K. M. Marzec and K. Majzner, *Cancers* , 2021, 13.

8 E. Wiercigroch, E. Szafraniec, K. Czamara, M. Z. Pacia, K. Majzner, K. Kochan, A. Kaczor, M. Baranska and K. Malek, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.*, 2017, **185**, 317–335.

9 A. Rygula, M. Z. Pacia, L. Mateuszuk, A. Kaczor, R. B. Kostogrys, S. Chlopicki and M. Baranska, *Analyst*, 2015, **140**, 2185–2189.

10 A. J. Ruiz-Chica, M. A. Medina, F. Sánchez-Jiménez and F. J. Ramírez, *J. Raman Spectrosc.*, 2004, **35**, 93–100.

11 A. Borek-Dorosz, A. M. Nowakowska, P. Leszczenko, A. Adamczyk, A. Pieczara, J. Jakubowska, A. Pastorczak, K. Ostrowska, M. Ząbczyńska, K. Sowinski, W. I. Gruszecki, M. Baranska, K. M. Marzec and K. Majzner, *J. Adv. Res.*, 2022, **41**, 191–203.