Supporting Information

Mitigating the impact of gelatin capsule variability on detection of substandard and falsified pharmaceuticals with near-IR spectroscopy

Olatunde Awotunde [a], Jiaqi Lu [a], Jin Rui [a], Nicholas Roseboom[a], Ornella Joseph [a], Sarah Honegger[a], Alyssa Wicks[a], Kathleen Hayes[a], Marya Lieberman [a]*

[a] Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN 46556 USA

Corresponding Author

Marya Lieberman, PhD, - Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN USA, mlieberm@nd.edu (corresponding author) ORCID: 0000-0003-3968-8044

Authors

Olatunde Awotunde - Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN USA, aawotund@nd.edu ORCID: 0000-0002-5037-7549
Jiaqi Lu - Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN USA, jlu22@nd.edu ORCID: 0000-0003-2399-6245
Jin Rui-- Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN USA, jcai@nd.edu ORCID: 0000-0003-0708-5257
Nicholas Roseboom - Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN USA, nroseboo@nd.edu ORCID: 0000-0001-9611-1191
Sarah Honegger - Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN USA, shonegge@nd.edu
Ornella Joseph - Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN USA, ojoseph2@nd.edu ORCID: 0000-0001-7460-2555
Alyssa Wicks - Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN USA, awicks@nd.edu ORCID: 0000-0003-4584-0532
Kathleen Hayes - Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame IN USA, khayes5@nd.edu ORCID: 0000-0003-1217-0050

Table of Contents

Table S1: Contents of lab-made isoniazid formulations.

Isoniazid (mg)	Microcrystalline Cellulose (mg)	% isoniazid w/w
100	600	14
200	500	29
300	400	43
500	200	71

Table S2: Contents of lab-made doxycycline formulations.

Doxycycline (mg)	Lactose (mg)	% Doxycycline w/w
100	600	14
200	500	29
300	400	43
500	200	71

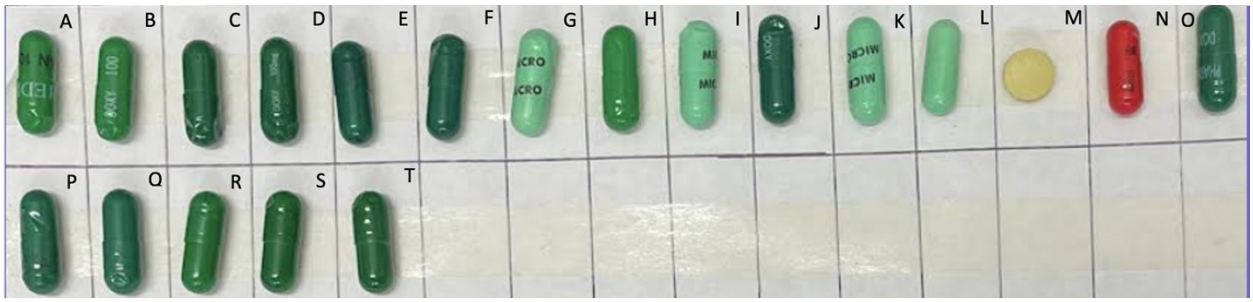

Figure S1i: Additional photographs of doxycycline capsules from products purchased in Kenya and Liberia in 2020.

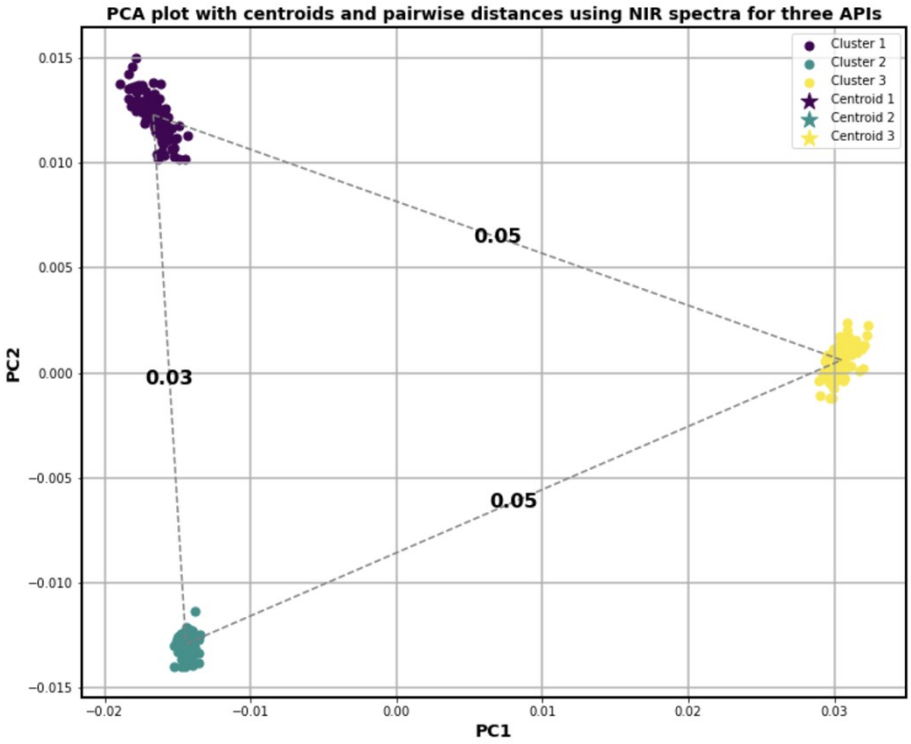

Figure 1ii: Histograms comparing all the inter-cluster distances for clusters of 17 capsule types/colors housing the same amount API (INH), versus the inter-cluster distances for the three different APIs/formulations (isoniazid, acetaminophen, and vitamin C) all in identical clear capsule type.

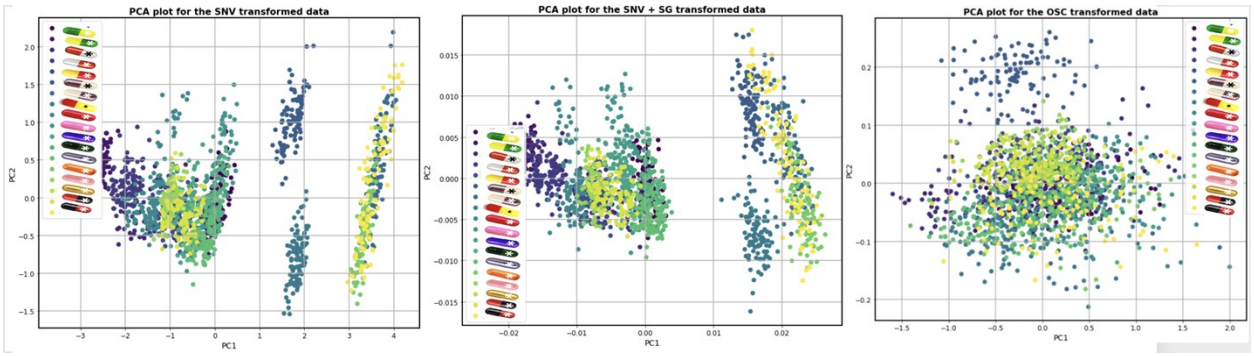

Figure S1iii . PCA scatter plot of the treated spectra of the same sample from which the dot plot pairwise Euclidean distances were measured (NB the scale of SNV-SG pretreatment is different due second polynomial and derivative treatment).

Table S3: Brands and theoretical % w/w API estimates for some of the field collected doxycycline capsules from Liberia and Kenya.

Sample ID	Collection date/location	Capsule letter	theoretical % w/w API range	Actual % w/w Doxycycline content in the Pill	Batch number	Actual API content (mg)	Pass/Fail
21L-004	2021/Liberia	1	29 - 37	31	DAOC-003	102.2	Pass
21L-007	2021/Liberia	2	35 - 41	34	DAOC-003	95.9	Pass
21L-010	2021/Liberia	3	28 - 52	43	DAOC-003	101.6	Pass
21L-012	2021/Liberia	4	31 - 44	36	DAOC-003	96.1	Pass
21L-001	2021/Liberia	5	39 - 41	34	K40901	99.7	Pass
21L-003	2021/Liberia	6	33 - 42	35	K40901	105.3	Pass
21L-008	2021/Liberia	7	31 - 61	51	K40901	91.5	Pass
21L-011	2021/Liberia	8	29 - 39	32	K40901	92.5	Pass
21L-013	2021/Liberia	9	31 - 41	34	K40901	96.2	Pass
21L-002	2021/Liberia	10	30 - 39	33	SC-027	95.4	Pass
21L-006	2021/Liberia	11	28 - 37	31	SC-027	90.3	Pass
16_0775	2016/Kenya	A	43 - 57	26	150502	54.89	Fail
16_0788	2016/Kenya	B	46 - 62	27	150502	52.52	Fail
16_0612	2016/Kenya	C	45 - 60	24	150502	48.00	Fail
16_0269	2016/Kenya	D	46 - 62	27	150502	53.00	Fail
16_0267	2016/Kenya	E	46 - 62	26	150502	50.00	Fail
19_0558bw	2019/Kenya	12	38 - 51	45	FBW4K807	106.00	Pass
19_0558cg	2019/Kenya	13	38 - 51	45	FBW4K807	106.00	Pass
19_0559bw	2019/Kenya	14	38 - 50	42	FBW4K806	101.00	Pass
19_0559cg	2019/Kenya	15	38 - 50	42	FBW4K806	101.00	Pass
19_0555bw	2019/Kenya	16	39 - 52	47	FBW4K806	109.00	Pass
19_0555cg	2019/Kenya	17	39 - 52	47	FBW4K806	109.00	Pass
16-0787	2019/Kenya	18	39 - 52	47	483	109.00	Pass

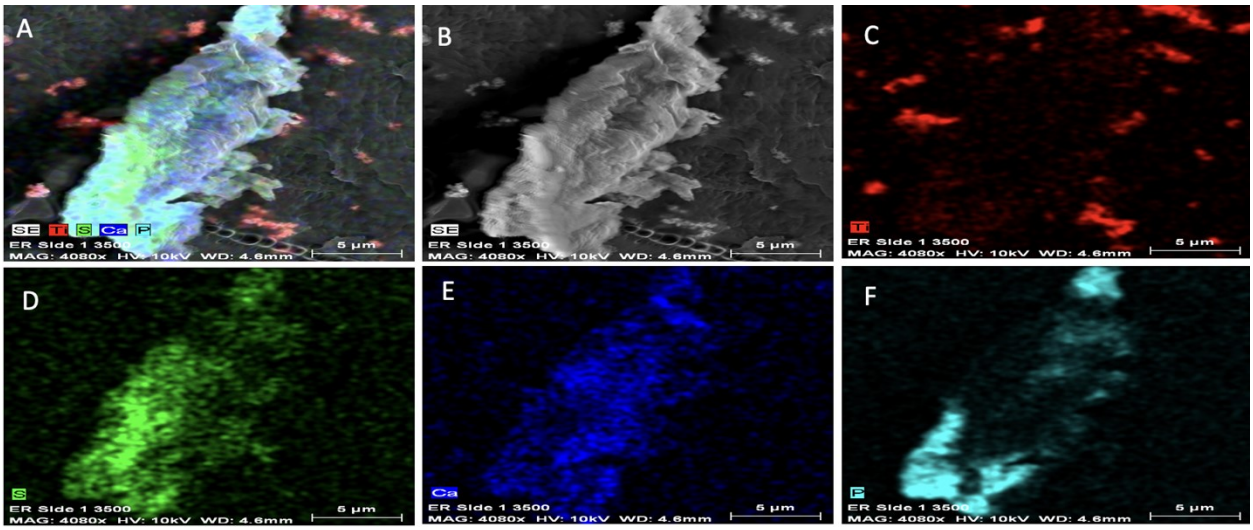

Figure S2: SEM EDX image shows surface roughness and uneven distribution of some elements.

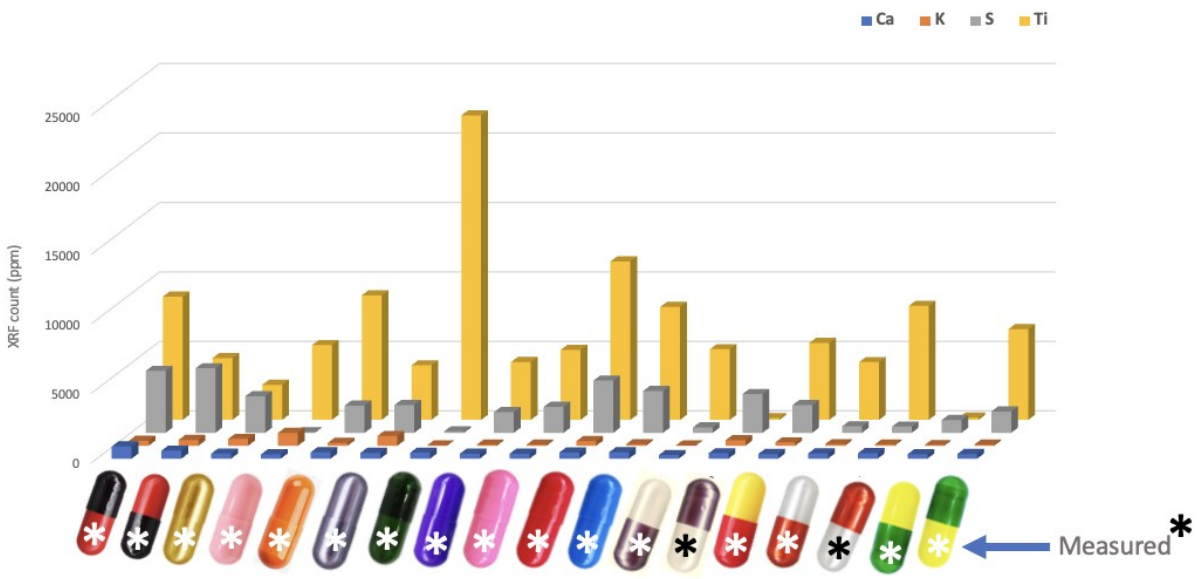

Figure S3: Approximate content of Ca (blue), K (orange), S (grey) and Ti (yellow) in 00 sized capsules, as measured by X-ray fluorimetry.
Scale is in units of parts per million (ppm) using the soil calibration of the XRF instrument. Ti is likely from titanium dioxide, a common opaquing agent.

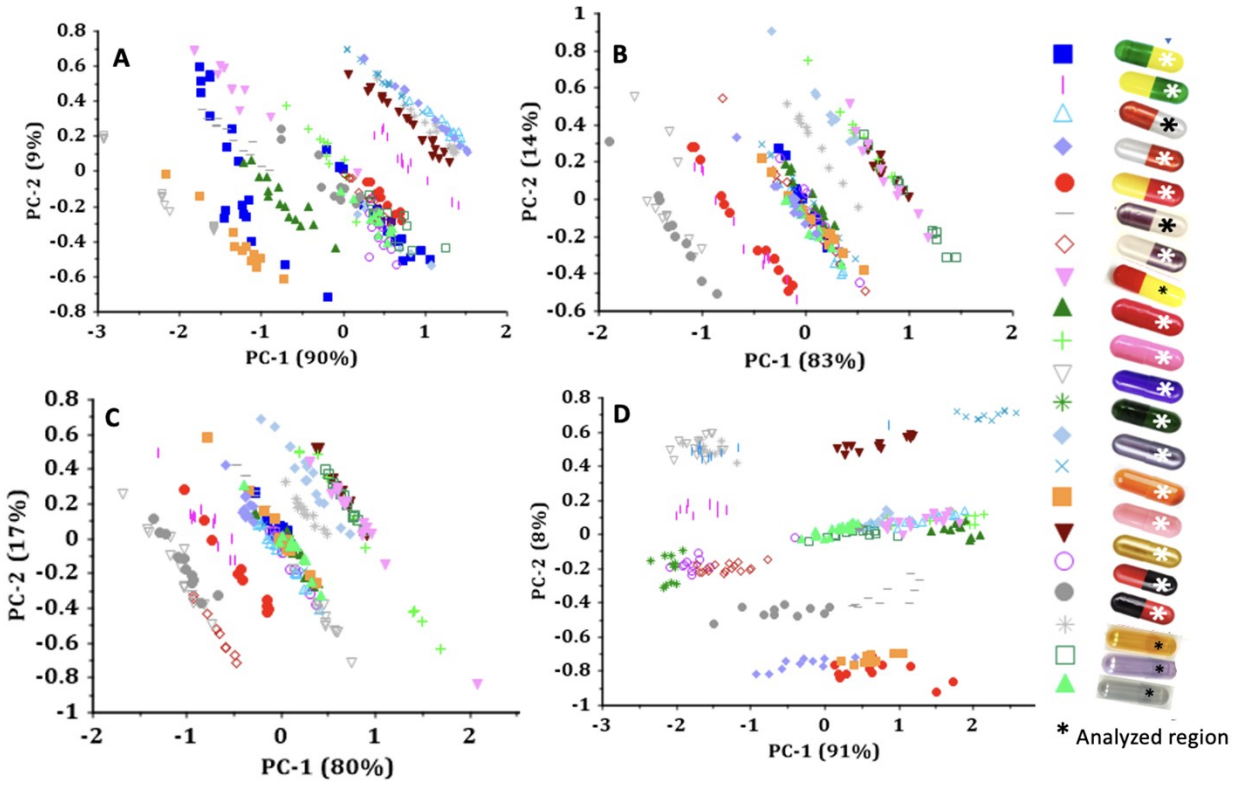

Figure S4: Principal components of the raw NIR spectra of 13 capsules of different colors (in addition to 3 semi-transparent capsules) filled with A) 14% B) 29%, C) 43% and D) 71% isoniazid (no data smoothing or pre-treatment).

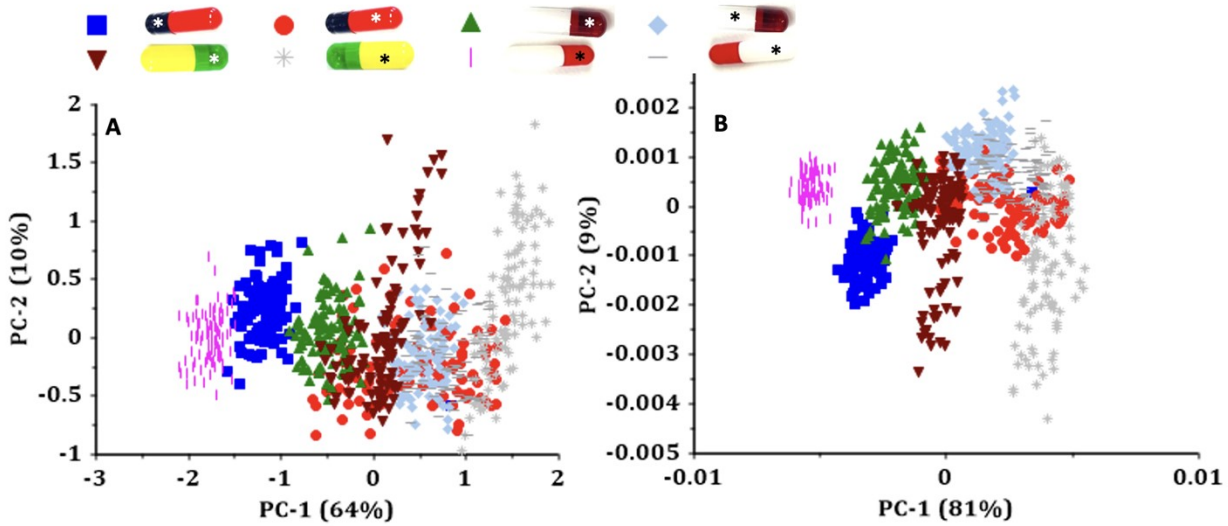

Figure S5: SNV (A) and SNV-SG (B) transformed NIR spectra of isoniazid acquired in the body and cap of four two-colored capsules.

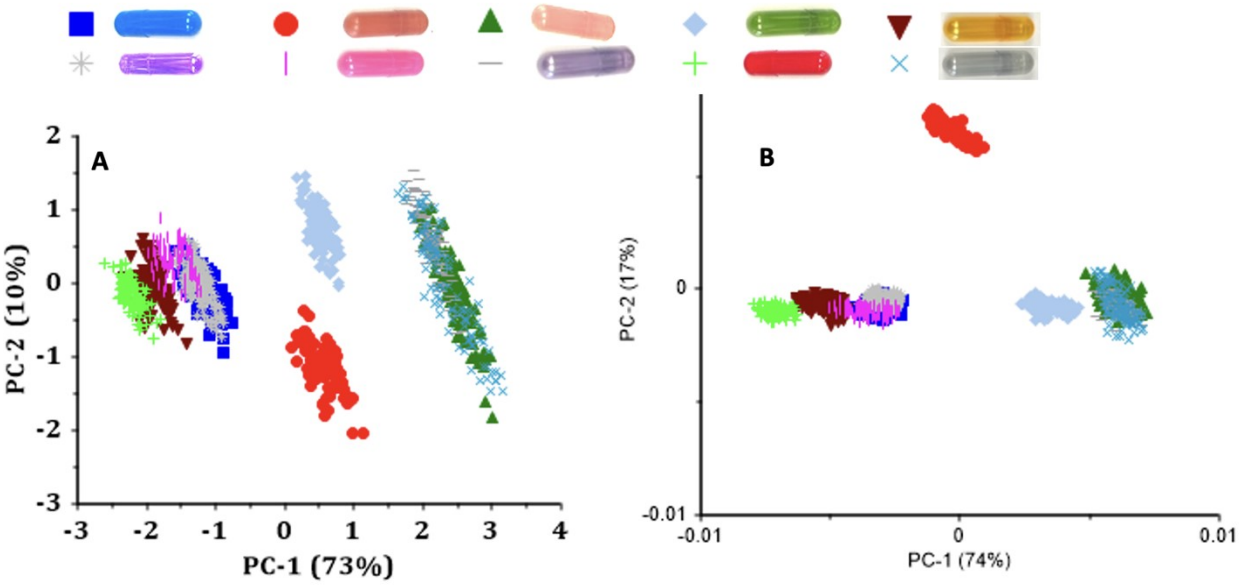

Figure S6: SNV (A) and SNV-SG (B) transformed NIR spectra of doxycycline acquired in ten single-colored capsules.

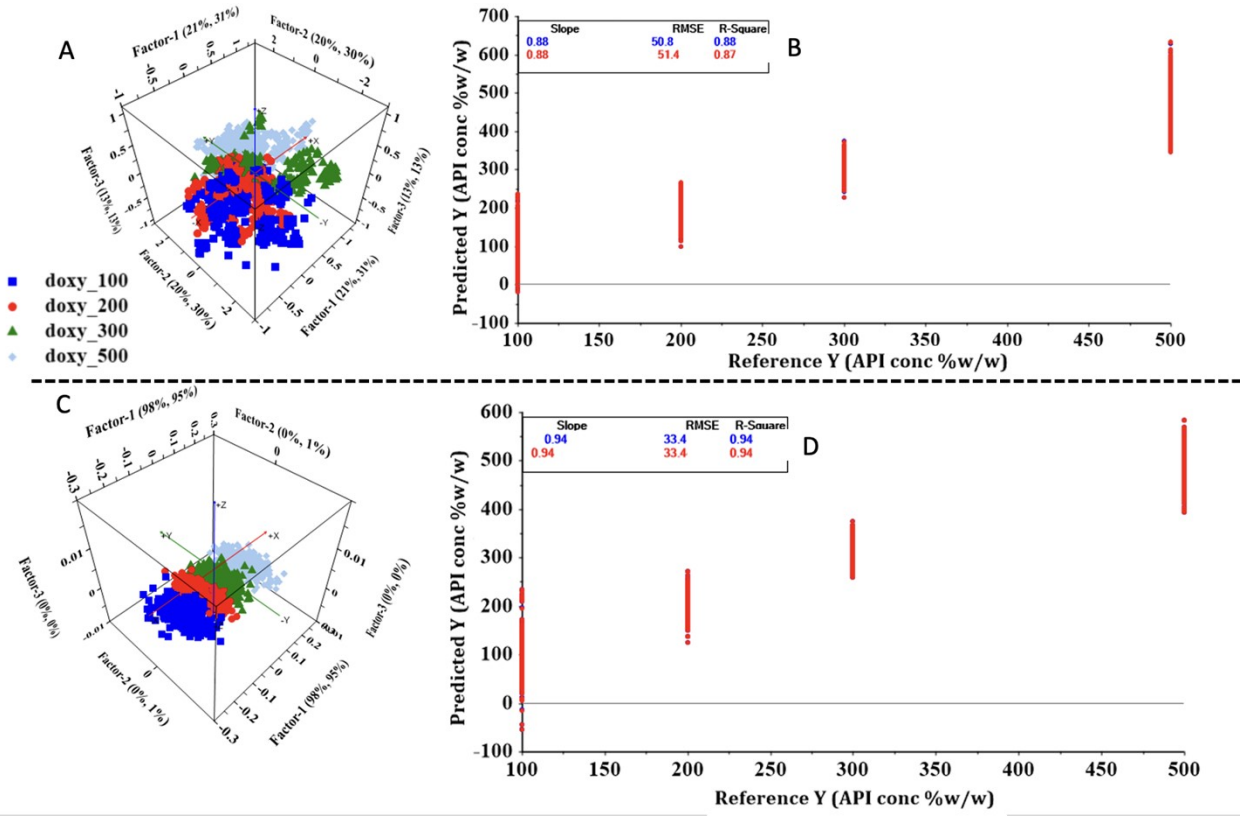

Figure S7: PLS-R performance lab formulated doxycycline and lactose binary mixtures predictions using capsules of different colors of SNV-SG (A & B) and SG-OSC (C & D). doxy_100, doxy_200, doxy_300 & doxy_500 is 14%, 28%, 42% and 71% isoniazid respectively. Panels A and C show principal component projections of the NIR data, panels B and D show both the calibration (blue text) and validation (red text) performance.

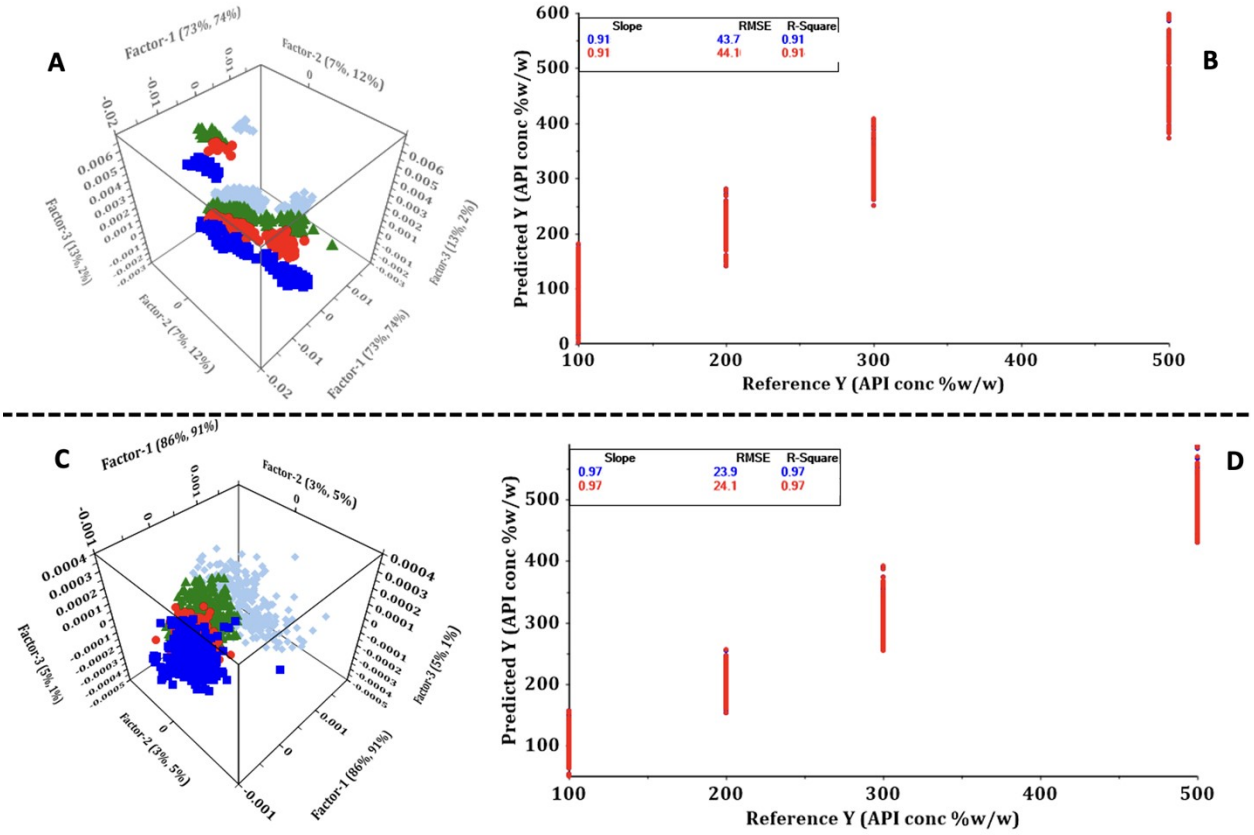

Figure S8: PLS-R performance lab formulated isoniazid and cellulose binary mixtures predictions using capsules of different colors of SNV-SG (A & B) and SG-OSC (C & D). INH_100, INH_200, INH_300 & INH_500 is 14%, 28%, 42% and 71% isoniazid respectively. Panels A and C show principal component projections of the NIR data, panels B and D show the calibration (blue text) and validation (red data points and text) performance.

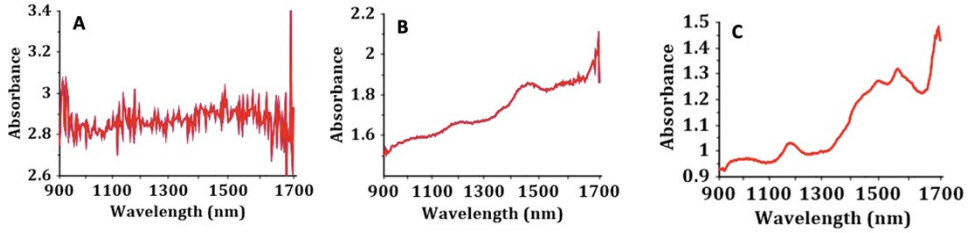

Figure S9: A. Blank B. empty capsule C. capsule housing API (Ethambutol) showed the role of CAD sample holder in ensuring consistency in sample presentation to NIR light with minimal to no interference.