Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/4/eaay5206/DC1

Supplementary Materials for

Direct observation of glucose fingerprint using in vivo Raman spectroscopy

Jeon Woong Kang, Yun Sang Park, Hojun Chang, Woochang Lee, Surya Pratap Singh, Wonjun Choi, Luis H. Galindo, Ramachandra R. Dasari, Sung Hyun Nam*, Jongae Park, Peter T. C. So*

*Corresponding author. Email: sh303.nam@samsung.com (S.H.N.); ptso@mit.edu (P.T.C.S.)

Published 24 January 2020, *Sci. Adv.* **6**, eaay5206 (2020)
DOI: 10.1126/sciadv.aay5206

This PDF file includes:

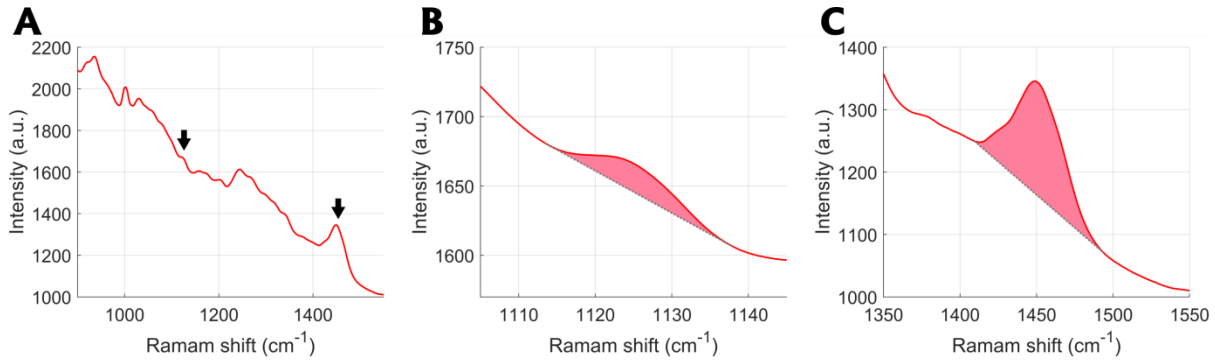

Fig. S1. Band area feature. (A) An acquired Raman spectrum in red. The two black arrows indicate two bands used as an example to show how to extract the band area feature, as illustrated in panels B and C. **(B)** A band area feature extracted over a glucose fingerprint peak at 1125 cm^{-1}. It is intensity integration (area in pink under the red curve) of the background-subtracted spectrum over the band when a tangential line (gray dotted line) over the band is used for the background subtraction. **(C)** Another illustration about a band area feature extracted from a band at 1450 cm^{-1} for protein and lipid of the skin. The *y*-axis labeled as "area integral of two bands" in Fig. 2B and Fig. 3D represents the ratio between band area features calculated as done in panel B and C.

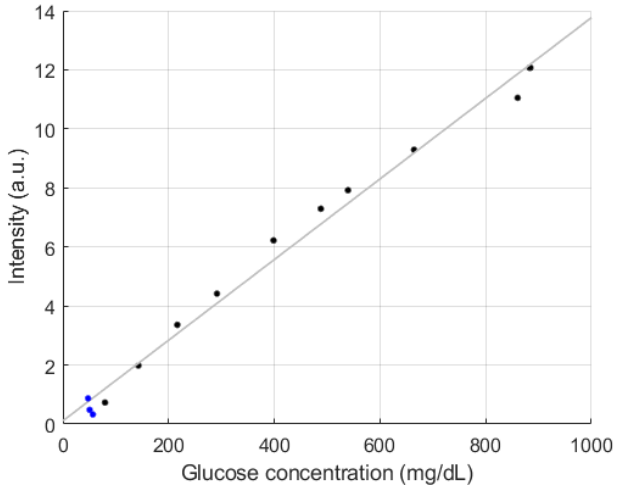

Fig. S2. Estimation of the LoD using a linear regression. If the instrument response *y* is linearly related to the concentration *x* as $y = a + bx$, LoD is defined as $3SD_a / b$, where SD_a is the standard deviation of *y*-residuals, and *b* is the slope of the linear curve (*37*). It illustrates spectrum intensities vs corresponding glucose concentrations during the time period when fluorescence stayed relatively flat (see Fig. 5C). Using the above LoD definition with minimal glucose concentrations around 52 mg/dL, the LoD of our measurements was calculated as ~75 mg/dL. It is noted that the standard deviation was calculated with the datasets nearest to zero concentration.

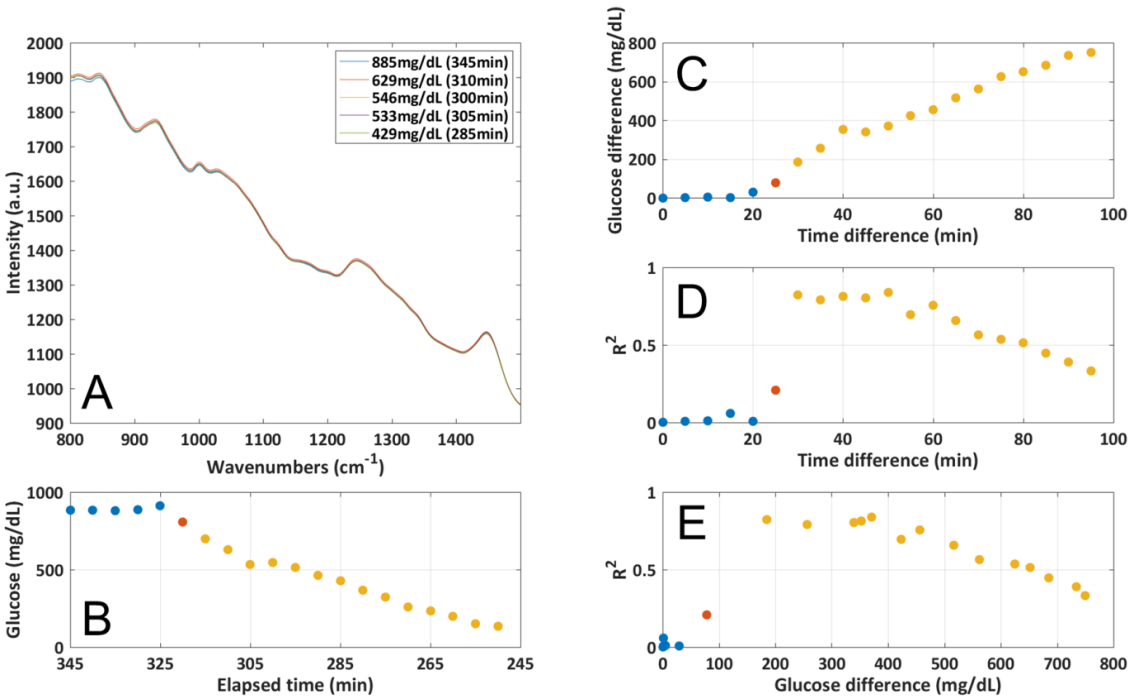

Fig. S3. Analysis of the validity of calibration using the background subtraction method used in Fig. 1A and the LoD measurement. (A) Raw spectra during the period when fluorescence stayed relatively flat**. (B)** Glucose concentrations from the elapsed time of 345 min to that of 250 min in time reversal manner. **(C)** Glucose concentration differences depending on the time difference from the reference at 345 min. The time of subtraction reference at the elapsed time of 345 min is located at the time difference of zero in panels C and D. The abscissa of the time difference in C and D also goes in time reversal manner. **(D)** Change in the squared Pearson correlation coefficient between the subtraction spectra and the spectrum of pure glucose in solution. This demonstrates high correlation coefficients for ~ 50 min when enough differences started to appear in the corresponding glucose concentrations between subtraction spectra. The initial low correlations were due to the small glucose change from the approximately flat glucose level. This indicates that the calibration using the subtraction reference spectrum can stay valid and reliable for ~ 50 min. **(E)** The squared Pearson correlation coefficients as a function of glucose concentration difference, combining the results in panels C and D. For glucose concentration differences smaller than ~ 30 mg/dL (marked as blue dots), glucose signal differences in the corresponding subtraction spectra were observed to be buried below the noise level. At the glucose concentration difference of 78 mg/dL (a red dot), a distinguishing correlation coefficient from the prior data started to appear (the same color codes were applied in panels B to E.). The unfilled gap with data between 30 mg/dL and 78 mg/dL was due to the limited number of datasets from our *in vivo* experiment.

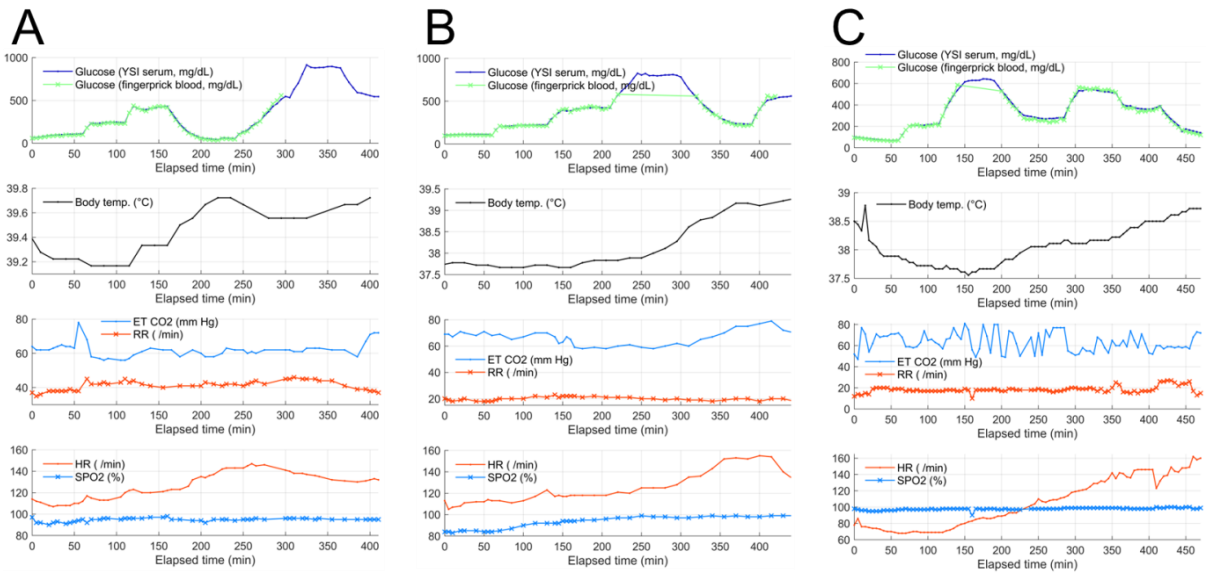

Fig. S4. Change in glucose concentration measured by a YSI glucose analyzer and Accu-Chek finger-prickers (top panels) and vital signs from the subject. Trials 1 to 3 in Panels A to C, respectively. The vital signs include the subject's body temperature, end-tidal CO_2 (ET CO2), respiration rate (RR), heart rate (HR), and blood oxygen saturation (SPO2).

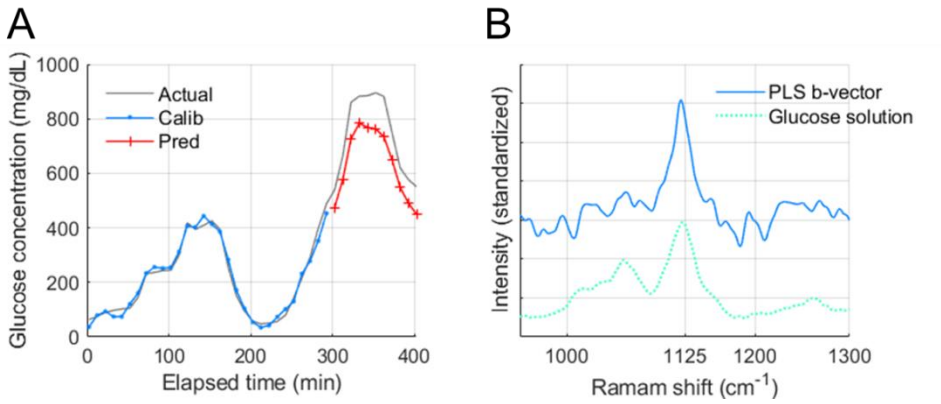

Fig. S5. Linear regression with full-range spectra. (A) Glucose concentration prediction in partial least squares regression analysis using full-range background-subtracted spectra in Trial 1. It results in R=0.99 and MARD=14.1 % in prediction. **(B)** The PLSR b-vector used in (A) is comparable to glucose solution spectrum, and this represents the Raman spectra capture glucose signals (*34*).

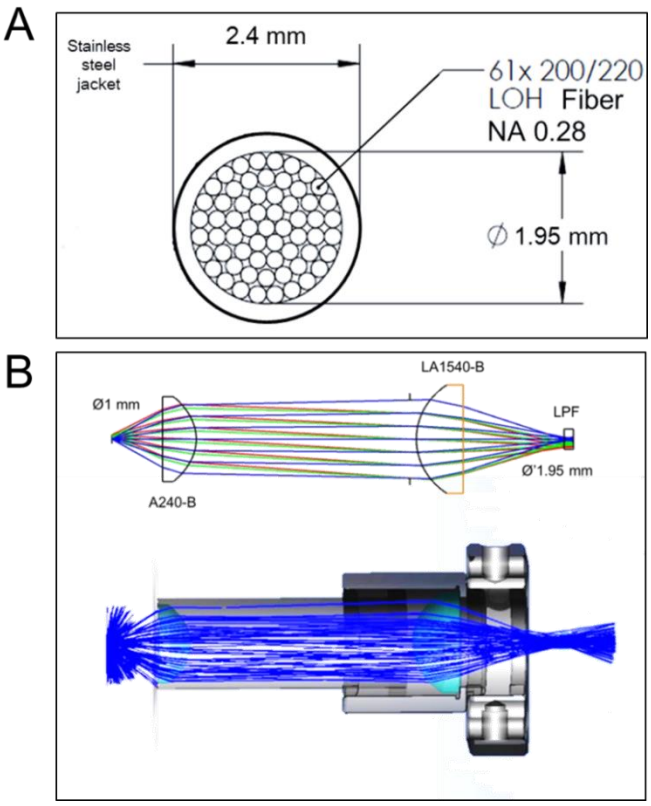

Fig. S6. Raman probe design. (A) Raman probe design based on sequential ray-tracing simulation. **(B)** Non-sequential ray-tracing simulation to confirm the performance of the designed Raman probe. It should be noted that a different design for Raman probe was used for the Trial 2. Instead of using 2 mm-diameter fiber-bundle directly over the sampling volume for collection of Raman photons, a simple imaging type Raman probe with lenses were used for higher NA collection from skin. The magnification of the probe is set to match the diameter of the 1.95 mm-diameter input aperture of the fiber-bundle.

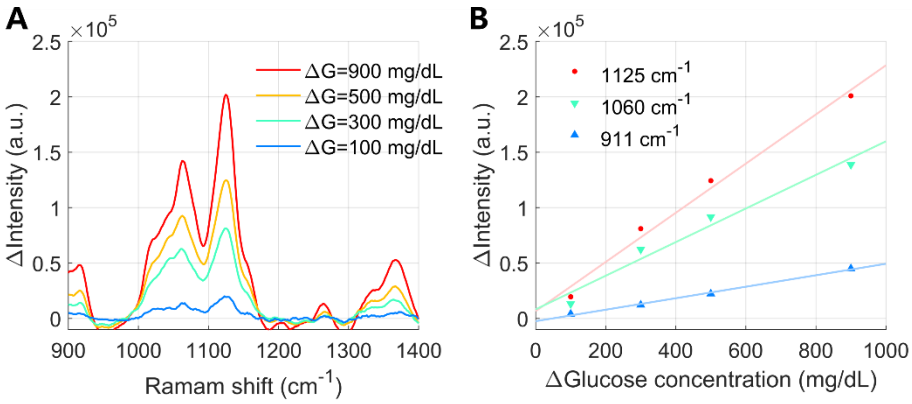

Fig. S7. Glucose signal change corresponding to several Δ*G* values in glucose solution. (A) Difference spectra with several ΔG (glucose concentration difference) values in glucose solution. The intensity of the signal, considered to be glucose, increases as ΔG increases. **(B)** Changes in the intensity at the three glucose characteristic Raman peaks (911 cm^{-1}, 1060 cm^{-1}, and 1125 cm^{-1}) in panel A. The Raman peak at 1125 cm^{-1} responds in the most prominent way to glucose concentration change among the three fingerprint peaks.

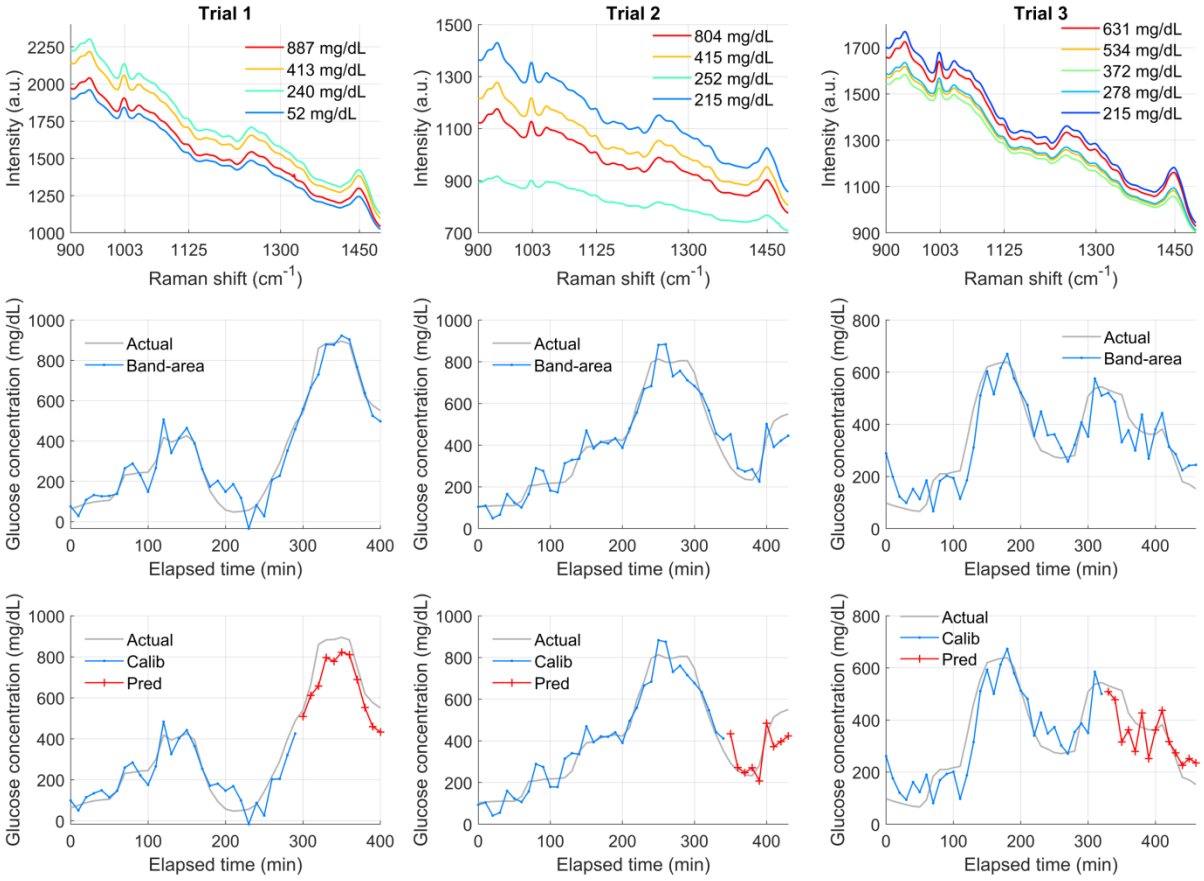

Fig. S8. Spectra and time course results from the band area features in MLR analysis from all the three trials. Raman spectra and their corresponding glucose concentrations in the *in vivo* experiments are shown in the top row. The middle row represents time course results from the four band-area features in MLR analysis with all the data. When the data are split into two parts for calibration and prospective prediction, the time course results from the band-area features in MLR analysis are illustrated in the bottom row.

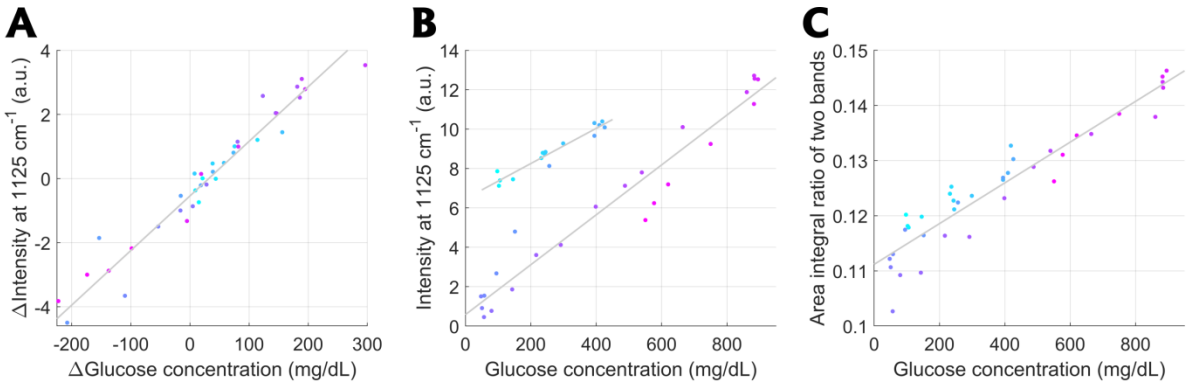

Fig. S9. Raman intensity changes with different glucose values. (A) Re-plotted Fig. 1B (ΔG vs. ΔI). Data points colored in cyan to magenta are arranged in chronological time order, and colored data points on each plot correspond to those in the same color on the other two plots. **(B)** Linearity between G vs. I. Linearity of the Raman peak to glucose concentration becomes poor for the entire recordings (R=0.74), though two partial time-continuous recordings show great linearity (R=0.96 with data points both in cyan color in the first one third recordings and R=0.97 with those in magenta color in the last two thirds.). The broken linearity is considered to occur due to the subject's movement. **(C)** The corresponding band-area ratio vs. glucose concentration. Panel C shows the improved linearity in the entire recordings as R=0.94. This confirms the effectiveness of the suggested band-area feature for blood-glucose prediction.