

2021

## Infrared based saliva screening test for COVID-19

Bayden R. Wood

*Monash University, Clayton Campus, 3800, Victoria, Australia*

Kamila Kochan

*Monash University, 3800 Victoria, Australia*

Diana E. Bedolla

*Monash University, Clayton Campus, 3800 Victoria Australia*

*See next page for additional authors*

Follow this and additional works at: <https://arrow.tudublin.ie/nanolart>

### Recommended Citation

Wood, B.R., Kochan, K., Bedolla, D.E. (2021). Infrared based saliva screening test for COVID-19. *Angewandte Chemie International Edition*, vol. 60, no. 31, pg. 17102–17107 doi:10.1002/anie.20210445360

This Article is brought to you for free and open access by the NanoLab at ARROW@TU Dublin. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@TU Dublin. For more information, please contact [arrow.admin@tudublin.ie](mailto:arrow.admin@tudublin.ie), [aisling.coyne@tudublin.ie](mailto:aisling.coyne@tudublin.ie), [gerard.connolly@tudublin.ie](mailto:gerard.connolly@tudublin.ie).



This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 4.0 License](https://creativecommons.org/licenses/by-nc-sa/4.0/)

---

**Authors**

Bayden R. Wood, Kamila Kochan, Diana E. Bedolla, Natalia Salazar-Quiroz, Samantha L. Grimley, David Perezguaita, Matthew J. Baker, Jitraporn Vongsvivut, Mark J. Tobin, Keith R. Bambery, Dale Christensen, Shivani Pasricha, Anthony K. Eden, and Aaron Mclean

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/351919867>

# Infrared based saliva screening test for COVID-19

Article in *Angewandte Chemie International Edition* · May 2021

DOI: 10.1002/anie.202104453

CITATIONS

14

READS

251

23 authors, including:



**Bayden R Wood**

Monash University (Australia)

275 PUBLICATIONS 8,565 CITATIONS

SEE PROFILE



**Kamila Kochan**

Monash University (Australia)

61 PUBLICATIONS 1,794 CITATIONS

SEE PROFILE



**Diana Bedolla**

Sincrotrone Trieste S.C.p.A.

44 PUBLICATIONS 383 CITATIONS

SEE PROFILE



**Natalia Salazar-Quiroz**

Emory University

6 PUBLICATIONS 80 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



International Society for Clinical Spectroscopy (<http://clirspec.org/>) [View project](#)



BTN2A1 & cancer immunity [View project](#)

A Journal of the Gesellschaft Deutscher Chemiker

# Angewandte Chemie

GDCh

International Edition

www.angewandte.org

## Accepted Article

**Title:** Infrared based saliva screening test for COVID-19

**Authors:** Bayden Robert Wood, Kamila Kochan, Diana E. Bedolla, Natalia Salazar-Quiroz, Samantha Grimley, David Perez-Guaita, Matthew J. Baker, Jitraporn Vongsvivut, Mark Tobin, Keith Bambery, Dale Christensen, Shivani Pasricha, Anthony K. Eden, Aaron Mclean, Supti Roy, Jason Roberts, Julian Druce, Deborah A. Williamson, Julie McAuley, Mike Catton, Damian Purcell, Dale Godfrey, and Philip Heruad

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

**To be cited as:** *Angew. Chem. Int. Ed.* 10.1002/anie.202104453

**Link to VoR:** <https://doi.org/10.1002/anie.202104453>

## RESEARCH ARTICLE

**Infrared based saliva screening test for COVID-19**

Bayden R. Wood<sup>[a]\*†</sup>, Kamila Kochan<sup>[a]†</sup>, Diana E. Bedolla<sup>[a,b,c]2,3†</sup>, Natalia Salazar-Quiroz<sup>[d]</sup>, Samantha Grimley<sup>[d]</sup>, David Perez-Guaita<sup>[a,e,f]</sup>, Matthew J. Baker<sup>[g]</sup>, Jitraporn Vongsvivut<sup>[h]</sup>, Mark Tobin<sup>[h]</sup>, Keith Bambery<sup>[h]</sup>, Dale Christensen<sup>[h]</sup>, Shivani Pasricha<sup>[d]</sup>, Anthony K. Eden<sup>[a]</sup>, Aaron Mclean<sup>[a]</sup>, Supti Roy<sup>[a]</sup>, Jason Roberts<sup>[i]</sup>, Julian Druce<sup>[i]</sup>, Deborah A. Williamson<sup>[d,j]</sup>, Julie McAuley<sup>[d]</sup>, Mike Catton<sup>[i]</sup>, Damian Purcell<sup>[d,k]\*</sup>, Dale Godfrey<sup>[d,k]\*</sup>, and Philip Heraud<sup>[a,l]\*</sup>

- [a] Prof. Bayden R. Wood, Dr. Kamila Kochan, Dr. Diana E. Bedolla, Dr. David Perez-Guaita, Mr. Anthony Eden, Mr. Aaron Mc Lean, Dr. Supti Roy, Dr. Philip Heraud  
Centre for Biospectroscopy and School of Chemistry,  
Monash University, Clayton Campus,  
3800, Victoria, Australia.  
E-mail: [bayden.wood@monash.edu](mailto:bayden.wood@monash.edu), [phil.heraud@monash.edu](mailto:phil.heraud@monash.edu)
- [b] Dr. Diana E. Bedolla  
Area Science Park,  
Padriciano 99, 34149, Trieste, Italy.
- [c] Dr. Diana E. Bedolla  
Elettra-Sincrotrone Trieste,  
S.S. 14 Km 163.5, 34149, Trieste, Italy.
- [d] Dr. Natalia Salazar Quiroz, Dr. Samantha Grimley, Dr. Shivani Pasricha, Prof. Deborah A. Williamson, Dr. Julie McAuley, Prof. Damian Purcell, Prof. Dale Godfrey  
Department of Microbiology and Immunology,  
University of Melbourne,  
Melbourne, Australia.  
E-mail: [dfjp@unimelb.edu.au](mailto:dfjp@unimelb.edu.au), [godfrey@unimelb.edu.au](mailto:godfrey@unimelb.edu.au)
- [e] Dr. David Perez-Guaita  
FOCAS Research Institute,  
Dublin Institute of Technology,  
D04, Dublin, Ireland.
- [f] Dr. David Perez-Guaita  
Department of Analytical Chemistry, University of Valencia,  
50 Dr. Moliner Street, Research Building,  
46100 Burjassot Valencia, Spain.  
Address 2
- [g] Dr. Matthew J. Baker  
Department of Pure and Applied Chemistry  
University of Strathclyde,  
Technology and Innovation Centre Level 7,  
99 George Street, Glasgow, G1 1RD, Scotland, UK.
- [h] Dr. Jitraporn Vongsvivut, Dr. Mark Tobin, Dr. Keith Bambery  
Australian Synchrotron,  
Victoria 3168, Australia.
- [i] Dr. Jason Roberts, Dr. Julian Druce, Mike Catton  
Victorian Infectious Diseases Reference Laboratory,  
The Peter Doherty Institute for Infection and Immunity,  
Melbourne, Vic, Australia.
- [j] Prof. Deborah A. Williamson  
Department of Microbiology,  
Royal Melbourne Hospital,  
Melbourne, Australia.
- [k] Prof. Dale Godfrey  
Australian Research Council Centre of Excellence in Advanced Molecular Imaging,  
University of Melbourne,  
Melbourne, Victoria 3010.
- [l] Dr. Philip Heraud  
Department of Microbiology and the Biomedicine Discovery Institute,  
Faculty of Medicine, Nursing and Health Sciences, Monash University,  
Victoria, Australia.
- [†] These authors contributed equally to this work.

Supporting information for this article is given via a link at the end of the document.

## RESEARCH ARTICLE

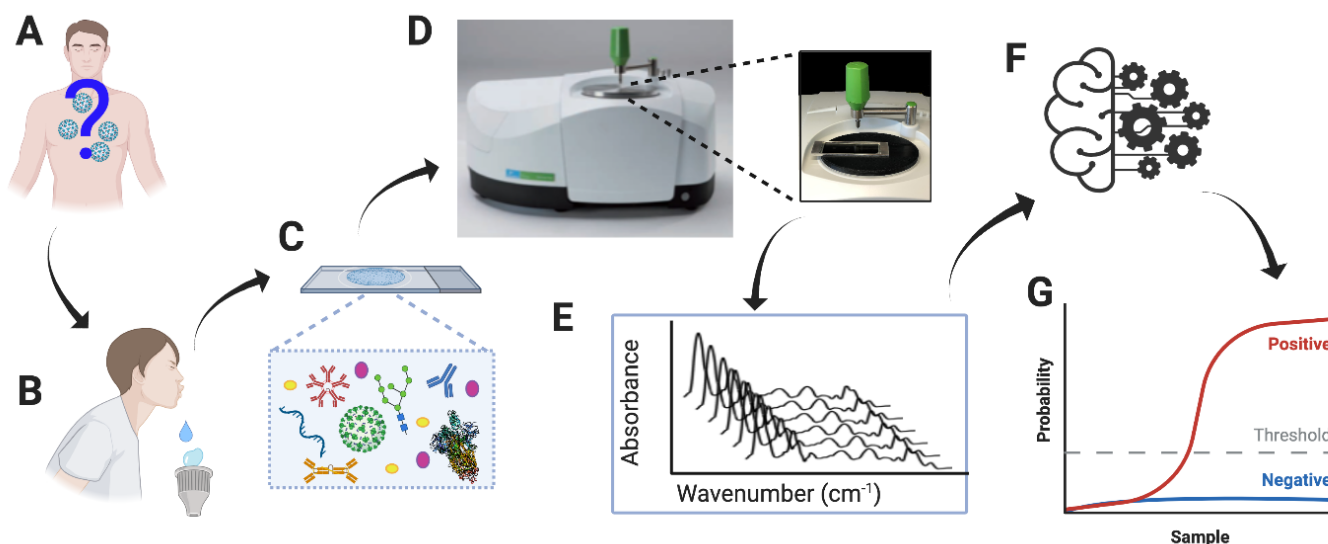
**Abstract:** Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in an unprecedented need for diagnostic testing that is critical in controlling the spread of COVID-19. We propose a portable infrared spectrometer with purpose-built transfection accessory for rapid point-of-care detection of COVID-19 markers in saliva. Initially, purified virion particles were characterized with Raman spectroscopy, synchrotron infrared (IR) and AFM-IR. A data set comprising 171 transfection infrared spectra from 29 patients testing positive for SARS-CoV-2 by RT-qPCR and 28 testing negative, was modeled using Monte Carlo Double Cross Validation with 50 randomized test and model sets. The testing sensitivity was 93 % (27/29) with a specificity of 82 % (23/28) that included positive samples on the limit of detection for RT-qPCR. Here, we demonstrate a proof-of-concept high throughput infrared COVID-19 test that is rapid, inexpensive, portable and utilizes sample self-collection thus minimizing the risk to healthcare workers and ideally suited to mass screening.

## Introduction

Coronavirus disease 2019 (COVID-19) is a highly transmissible respiratory disease caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). This virus has infected over 160 million people resulting in greater than 3 million deaths surpassing malaria infections as the world's most devastating infectious disease.[1] The high mortality rate and tremendous socio-economic cost of the virus has necessitated mass testing to detect infected individuals who can then be isolated to reduce transmission. The gold standard for detection of SARS-CoV-2 is reverse transcription Polymerize Chain Reaction (RT-PCR) assay using material collected from nasopharyngeal swabs.[2] Whilst the assay has been effective in reducing transmission rates, the

turnaround time for the result can be several days. The assay itself takes ~4 hours, requiring relatively expensive instrumentation and consumables including PCR primers, enzymes and other reagents, which can be difficult to implement in under-resourced health care systems. Collection of swab samples for RT-PCR requires skilled health workers and close face to face contact with patients; hence necessitating personal protective equipment (PPE), which is expensive and may be in short supply.[3] Nasal swab collection is uncomfortable, sometimes even causing bleeding and is likely to be unacceptable by some workers, where regular surveillance sampling is required.[4] While saliva testing with RT-qPCR is being explored[5], this does not overcome the limitations associated with expense and time. Considering these many drawbacks, the implementation of new diagnostic approaches that can be performed rapidly, affordably, and with minimal risk to patient and health care workers, is highly desirable.

Infrared spectroscopy is a routine analytical technique mainly used to identify molecular functional groups in various materials. Infrared light interacting with the intrinsic vibrational modes of molecules generates a spectrum that represents a unique chemical fingerprint of the sample. The development of portable Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) spectrometers has seen a plethora of applications in the field of biomedical diagnostics [6]. More recently, the technology has been applied to diagnose Plasmodium sp. in blood [7] and viruses including hepatitis B and hepatitis C in serum [8]. The technique relies on detecting both the molecular phenotype of the pathogen and the host immune response. Advanced machine learning techniques can then be applied to predict diagnostic outcomes based on the chemical differences between infected and uninfected samples.



**Figure 1. Schematic diagram of an infrared saliva-based test for COVID-19.** (A-E) (A) Patient presents with COVID-19 symptoms. (B) The patient dribbles into a container containing Viral Transport Medium (VTM). (C) The saliva, which contains a range of COVID-19 biomarkers including ACE2, adenosine deaminase, immunoglobulin G, immunoglobulin M, RNA and secretory immunoglobulin A is deposited onto an infrared transfection substrate and dried (10 mins). (D) The spectra were acquired in triplicate (5 minutes) using a PerkinElmer Spectrum Two spectrometer with a modified PerkinElmer® reflective accessory optimised for transfection slides (insert). (E) The spectra, which represent a chemical snapshot of the entire saliva chemistry, including COVID-19 markers. (F) A Monte Carlo double cross validation model is used to predict COVID-19 based on spectral markers. (G) The results are presented as PLS-DA prediction plots and Receiver Operating Curves.

## RESEARCH ARTICLE

Recently, Barauna et al.[9] applied ATR-FTIR to analyze saliva collected from COVID-19 diagnosed patients using pharyngeal swabs. In their measurements the swab was placed onto ATR spectrometer and the spectrum recorded of both the saliva and swab. Interestingly their spectra of purified virus did not show the characteristic amide modes of proteins that are expected given the dominant S-proteins that characterize SARS-CoV-2 virions, nor did they show the characteristic RNA bands that would also be expected from a large RNA virus. [10]

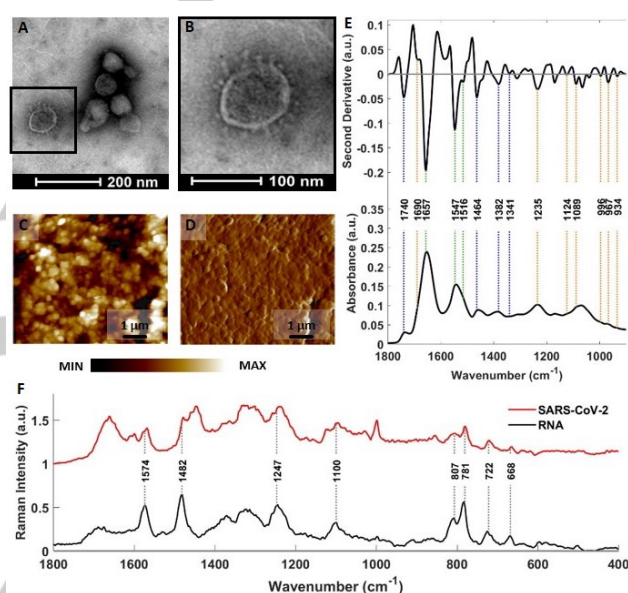
ATR spectroscopy has a number of disadvantages when it comes to high throughput screening. First, the sample has to be dried onto the internal reflection element (IRE) prior to the measurement to maximise the absorbance. Secondly, the residue must be cleaned off the IRE, which increases the potential of transmission through aerosols and via surfaces, posing a danger to the operator. More recently, ATR disposable silicon substrates have been developed and deployed to detect brain cancer in serum samples. [11] These substrates negate the drying and cleaning steps but are prohibitively expensive for routine COVID-19 testing at present. In pursuit of a cheaper approach to high throughput screening, we propose utilizing infrared reflective slides and a modified reflection accessory to obtain high quality spectra of saliva. Hitherto, most infrared diagnostics have focused on blood components and only few studies have investigated less invasively collected samples such as saliva [12]. Saliva is emerging as an attractive medium for point-of-care diagnosis of COVID-19 in the current pandemic. The SARS-CoV-2 virus has a preferential tropism to human airway epithelial cells that express the cellular receptor angiotensin-converting enzyme 2 (ACE2) [13]. ACE2 was found to be higher in salivary glands compared to the lungs, indicating salivary glands could be a potential target for SARS-CoV-2 [14]. Wylie et al. [15] reported that SARS-CoV-2 RNA copies were higher in saliva (5.58 mean log copies per millilitre) compared to nasopharyngeal swabs (4.93 mean log copies per millilitre) from 70 COVID-19 positive patients. Besides SARS-CoV-2 virions, saliva is known to have many other COVID-19 biomarkers, including ACE2, adenosine deaminase, immunoglobulin G, immunoglobulin M, RNA and secretory immunoglobulin A [16]. The number of potential biomarkers and complexity of the spectra requires a machine modelling approach, which has already been employed to data mine blood clinical parameters for COVID-19 detection [17] and detecting SARS-CoV-2 in nasal swabs using matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS). [18] From a practical perspective, saliva can be self-collected by dribbling into a sterile container reducing time and cost associated with the specimen collection and minimising nosocomial transmission of the disease [13b, 19]. The infrared technique is eminently suited to the analysis of saliva samples because it requires no additional reagents or consumables, is very rapid (less than 5 minutes to record spectra from three replicate samples and do the computation), the sample dries to a homogenous deposit and the data can be directly transferred to a machine learning model for diagnosis.

## Results and Discussion

A comprehensive schematic diagram of the COVID-19 diagnostic approach is presented in Figure 1(A-G). In developing a new infrared based test, the first aim was to record a high-quality

spectrum of the purified virus SARS-CoV-2 and assign the important marker bands. The SARS-CoV-2 virions were purified from the supernatant of infected Vero cells deactivated by fixation with 4% formalin. [20] Aliquots from the purified stock solution were deposited and dried onto a BaF<sub>2</sub> window.

The presence of virions was confirmed by means of Transmission Electron Microscopy (TEM). TEM images clearly show spherical particles approximately 120 nm in diameter with multiple spikes forming the solar crown structure, characteristic of coronaviruses (Figure 2 A, B) [21]. This is consistent with previously reported TEM images of SARS-CoV. [22] Atomic Force Microscopy (AFM) confirmed the presence of spherical particles of approximately 120 nm in diameter, notably aggregated together in large clusters (Figure 2 C, D). Synchrotron-FTIR spectra were collected from the individual clusters and the mean and second derivative spectrum calculated (Figure 2E).

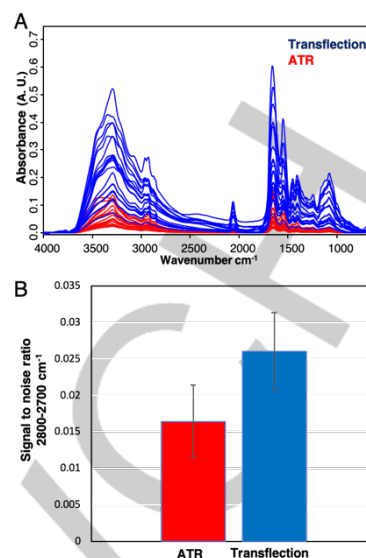


**Figure 2. AFM, TEM, synchrotron FTIR and Raman characterisation of SARS-CoV-2 virus. (A-E)** (A) TEM image of SARS-CoV-2 sample with single virion marked by black square. (B) Magnification of the area marked by the black square in (A), showing a single SARS-CoV-2 particle with its characteristic morphological appearance. (C) AFM height and (D) AFM deflection images of SARS-CoV-2 sample, demonstrating multiple round structures. The color scale for (C) and (D) is provided below. (E) Synchrotron FTIR spectrum and its 2<sup>nd</sup> derivative transform with the most prominent bands marked. (F) Raman spectrum (532 nm) of SARS-CoV-2 virions (red) compared to spectrum of purified RNA (black) with labelled bands. The bands are color-coded as follows: lipids (blue), proteins (green) and nucleic acids (orange).

Numerous bands are observed, reflecting the main chemical constituents of the virion particles (Table S1). These include protein bands at 1657, 1547, 1517 cm<sup>-1</sup> (from spike, envelope, membrane and nucleocapsid proteins) along with lipid bands at 1740, 1464, 1382 and 1341 cm<sup>-1</sup> (from lipid bilayer surrounding the nucleocapsid). In particular, multiple bands associated with RNA are present, including 1690, 1235, 1124, 1089, 996, 967 and 934 cm<sup>-1</sup>. [23] A similar spectral profile was observed using nanoscale IR spectroscopy (Figure S1), which enables single-virion spectra to be recorded at a vertical resolution of approximately 100 nm. It is important to note that the spectra will have minor contributions from fixative and residual media

## RESEARCH ARTICLE

attached to the virus particles. In particular, bands at  $1235\text{ cm}^{-1}$  and  $1092\text{ cm}^{-1}$  were assigned to residual paraformaldehyde, which polymerizes when 4 % formalin dries (Figure S2). Raman spectra of SARS-CoV-2 virions compared to purified RNA show nucleic acid markers at  $1242$ ,  $1110$ ,  $782$ ,  $723$  and  $670\text{ cm}^{-1}$  and an RNA-specific marker band at  $813\text{ cm}^{-1}$  (Figure 2F). Raman spectroscopy confirms the purity of the virion extraction and shows the very strong RNA bands that characterize this RNA rich virus. The bands between  $800\text{--}700\text{ cm}^{-1}$  are unique identifiers for RNA as evidenced by the corresponding RNA spectrum, which matches perfectly to many bands in the virion particle spectrum. The lack of bands from media or fixative confirms the purity of the virions. The relative contribution of RNA bands to the Raman spectrum is very high, likely reflecting the very high content of RNA in SARS-CoV-2, which contains the largest genome RNA-containing viruses. In pursuit of a high throughput tool for point-of-care diagnosis, we optimized a PerkinElmer Spectrum Two infrared spectrometer to perform transfection measurements using a modified reflection accessory for infrared reflective slides. A top plate was 3D printed to enable an aluminum slide holder to be slid along the mount so that the background spectrum and each sample deposit could be measured successively. To improve the throughput the slides were mounted as low as possible in the top plate and the mirrors of the reflection accessory adjusted so the 8 mm beam matched the size of the deposit to maximize absorbance. The larger infrared beam spot enables more sample and consequently more virions to be detected in the saliva compared to ATR-FTIR and Raman spectroscopy. Raman has the advantage of being able to measure aqueous samples. However, the diameter of laser beam ( $\sim 1\text{ }\mu\text{m}$ ) results in much less virions being detected compared to the infrared based approaches. Close up photographs of the accessory are found in the Supporting Information (Figure S3). The transfection approach resulted in greater absorbance compared to ATR and less noise (Figure 3) because more sample is being interrogated by the infrared beam and thus the ability to detect virions improved. Other advantages include the ability to batch dry samples; viral transmission is minimized by self-collection of saliva with no swab contribution, the instrument does not need to be cleaned between measurements, and the slides can be easily stored for future analysis making the transfection approach more conducive to point-of-site testing and mass screening compared to ATR-FTIR. To test directly the application of the infrared approach as a diagnostic for COVID-19 infection, we recorded 171 infrared spectra of triplicate dried saliva deposits from 57 human donors presenting to the Royal Melbourne Hospital with COVID-19 like symptoms (Table S2). Of these 29 patients tested positive with SARS-CoV-2 by RT-qPCR and 28 tested negative. Figure 4A shows the averaged spectra ( $4000\text{--}800\text{ cm}^{-1}$ ) for the positive and negative SARS-CoV-2 infected saliva samples along with the corresponding second derivative spectra. Some of the spectra show a small contribution from Viral Transport Medium (VTM), which has bands at  $1578\text{ cm}^{-1}$ ,  $1408\text{ cm}^{-1}$  and  $1078\text{ cm}^{-1}$  (Figure S4). In addition to signatures from infectious agents, infrared spectra acquired from saliva will contain information from components including mucins, proline-rich proteins, cysteins, histatins, statherins, amylases, carbonic anhydrases, salivary peroxidases, lipids, carbohydrates and inorganic compounds including nitrate and thiocyanate.[24] Saliva spectra have a distinctive band from thiocyanate at  $2059\text{ cm}^{-1}$  and strong bands from proteins at  $3288\text{ cm}^{-1}$ ,  $1658\text{ cm}^{-1}$  and  $1549\text{ cm}^{-1}$  assigned to

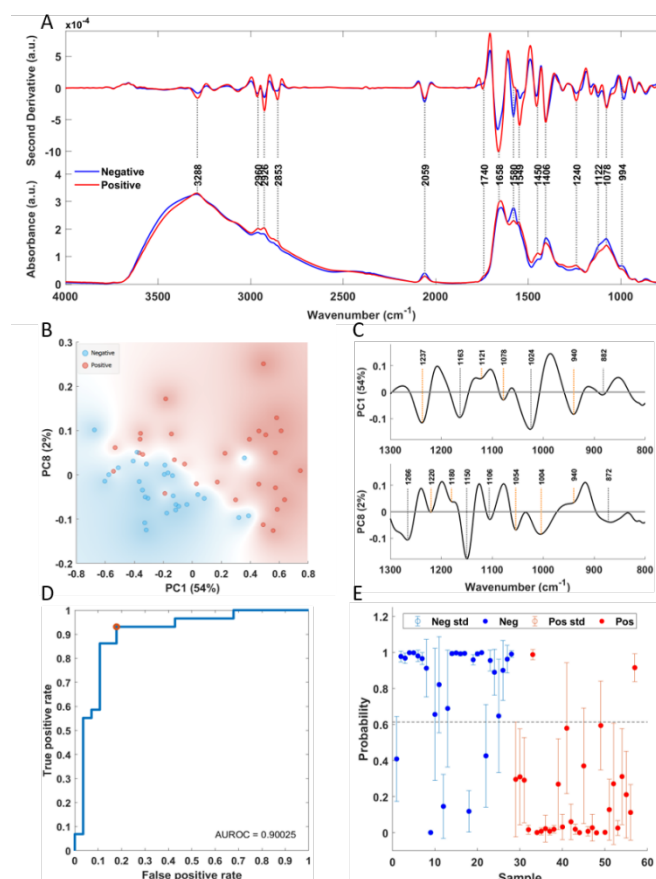


**Figure 3.** (A) Comparison of transfection spectra recorded using a Perkin Elmer Spectrum Two FT-IR spectrometer by depositing and drying the saliva samples (5 volunteers) directly onto the ATR internal reflection element (red) and using the new transfection accessory with infrared transfection slides. (B) Signal-to-noise comparison between the two techniques showing standard deviation error bars.

the amide A, amide I and II modes, respectively. The protein bands on average appear more intense in the positive samples compared to the negative samples. In general, the protein level of saliva is relatively small compared to other biofluids like serum. [24a] Hence, the significant difference in protein levels between infected and uninfected samples would likely be from viral proteins associated with the virus or antibodies produced by the host. On average, the thiocyanate band appears slightly more intense in the negative samples. Interestingly, thiocyanate is converted by salivary peroxidases to hypothiocyanite ( $\text{OSCN}^{-1}$ ) and is a potent antibacterial agent in the mouth, however, these levels will also change in response to  $\text{SCN}^{-1}$  intake in food and smoking [24b, 24c]. The bands at  $1240\text{ cm}^{-1}$  and  $1078\text{ cm}^{-1}$  are assigned to the phosphodiester groups associated with nucleic acids [23] and are more intense in the positive saliva samples indicating a contribution from viral RNA to this band. However, the  $1078\text{ cm}^{-1}$  band could also have minor contributions from VTM. We first applied Principal Component Analysis (PCA) to investigate general trends in the spectral data. The modelling was performed in the phosphodiester region ( $1300\text{--}800\text{ cm}^{-1}$ ) on vector normalized second derivative spectra where each spectrum is the average of the three replicates (Source Data 1 and External Modelling S1). The  $1300\text{--}800\text{ cm}^{-1}$  region was chosen because this is where many important RNA and glycoprotein marker bands are located and there is less interference from VTM. The PC1 vs. PC8 scores plot shows a general separation of infected and uninfected saliva sample spectra along PC1 with most negatives clustered in the lower left quadrant (Figure 4B). The corresponding PC1 vector indicated a number of RNA bands highlighted in orange ( $1237$ ,  $1121$ ,  $1078$  and  $940\text{ cm}^{-1}$ ) distinguishing the positive from the negative samples (Figure 4C). It should be noted that other bands including glycoproteins would also contribute to these bands in this region of the spectrum but the correlation with the averaged second derivative spectrum shown of virion particles in Figure 2E indicates that the major contribution to the infected saliva spectra is from RNA.

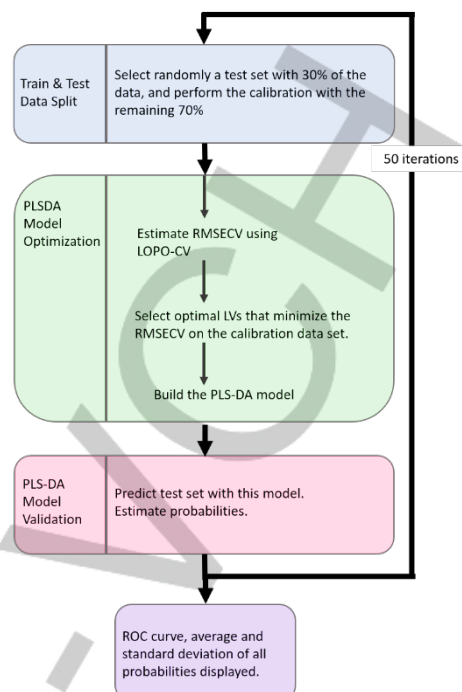


## RESEARCH ARTICLE



**Figure 4.** Patient modelling (A-E) (A) FTIR averaged spectra (bottom) and corresponding second derivative transformed spectra (top) from 87 spectra from 29 COVID-19 positive patients (red) and 84 spectra from 28 COVID-19 negative patients (blue) with important bands labelled. (B) Principal Component Analysis (PCA) scores plot showing PC1 (54 % explained variance) and PC8 (2 % explained variance) with each spectrum plotted as a single point in multidimensional space with red dots associated with positive and blue dots negative spectra. The red and blue shading shows the general separation of the positive from negative spectra, which occurs along the diagonal. (C) PC1 and PC8 loading plots showing the important bands discriminating the positive from negative sample spectra along each PC. The positive scores are associated with negative loadings and *vice versa* because the modelling was based on the second derivative spectra. (D) Receiver Operating Characteristic (ROC) Curve showing the false-positive rate versus the true positive rate for the MCDCV model. (E) Probability diagram showing the averaged sample spectra plotted as points in multi-dimensional space. The threshold value of 0.6 is shown by the dotted line, which is optimized to minimize the number of false negatives classified by the spectroscopic model. The error bars represent the standard deviation for the 50 randomized test and model sets.

Partial Least Squares-Discriminant Analysis (PLS-DA) was employed to create discrimination models to predict the positivity of the sample using first derivative spectra from the patient cohort (Source Data 2 and External Modelling S2). The methodology was evaluated using Monte Carlo Double Cross Validation (MCDCV) with 50 randomized test (30 %) and model sets (70 %) (Figure 5). The test spectra were averaged, and no test replicates were included in the model avoiding over optimistic modelling, the so-called technical replicate trap. By randomizing both the test and model set ensures the most robust and unbiased way of testing the model's performance. Furthermore, by keeping the numbers of variables small and using only the 1300-900  $\text{cm}^{-1}$  region further increases the robustness of the modelling because the more variables in the model, such as the variable amount of



**Figure 5.** MCDCV flow chart showing model architecture with inner cross validation and outer prediction loops.

viral transport media in the sample, the more chance of finding spurious correlations. Figure 4D shows the Receiver Operating area under the ROC of 0.90, which is excellent for a small sample cohort. The MCDCV modelling approach achieved a sensitivity of 93 % (27/29) and a specificity of 82 % (23/28) based on the selection of a threshold of 0.6, which was optimized to reduce the number of false-negatives (Figure 4E).

## Conclusion

Here we report on a new transfection infrared based saliva test for COVID-19. The results showed a 93 % sensitivity and 82 % specificity using the MCDCV modelling approach. Furthermore, we identified specific SARS-CoV-2 infrared and Raman marker bands from highly purified virions. To improve the sensitivity and specificity and determine if the technique has the specificity to distinguish SARS-CoV-2 from other respiratory viruses including influenza and other coronaviruses a larger patient cohort is required. RNA RT-qPCR remains the “gold standard” diagnostic technique for SARS-CoV-2 infection. However, there is an urgent need for a point-of-care screening technique that could potentially triage patients for specific RT-qPCR testing. Such a tool would be extremely useful in the current pandemic enabling onsite screening at airports, sporting venues, universities and schools. An infrared based saliva test is logistically easier to perform, rapid, and minimizes the risk of transmission to health workers. Furthermore, self-collection of saliva would reduce patient discomfort and improve community participation rates in testing.

## RESEARCH ARTICLE

## Acknowledgements

Mr. Finlay Shanks for instrumental support. Dr Emanuele Pedersoli for his assistance in graphic scripting. Funding: Part of this research was undertaken on the IRM beamline at Australian Synchrotron (Victoria, Australia), part of the Australian Nuclear Science and Technology Organisation (ANSTO). We acknowledge the support of the beamtime (Proposal ID. 16460 and 16476) provided by ANSTO, funded by the Australian Government. This work was in-part funded by Monash Green Chemical Futures Special Project fund. Part of this work was funded by the A2 milk foundation. Instruments were loaned and modified for transfection infrared slides by Dr Jerry Sellors and Dr Robert Packer courtesy of PerkinElmer®. N.S.Q, S.G. and J.M. were supported by funds from the Jack Ma Foundation through a grant to D.P. and D.I.G. D.P. and D.I.G. are supported by the National Health and Medical Research Council of Australia (NHMRC), including a Senior Principal Research Fellowship (1117766) and the Australian Research Council CE140100011 for D.I.G..

## Data and model availability

The datasets and model codes are available in the Zenodo repository (<https://zenodo.org/record/4156646>).

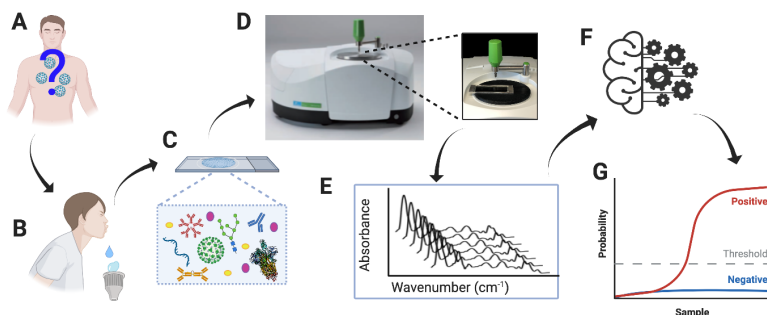
**Keywords:** COVID-19 diagnostic• Fourier transform infrared spectroscopy (FTIR)• Raman spectroscopy• SARS-CoV-2• Saliva

[1] World Health Organisation Coronavirus Disease (COVID-19) Dashboard, W.H.O., <https://covid19.who.int/> 2020.  
 [2] L. M. Czumbel, S. Kiss, N. Farkas, I. Mandel, A. Hegyi, A. Nagy, Z. Lohinai, Z. Szakacs, P. Hegyi, M. C. Steward, G. Varga, *Front. Med. (Lausanne)* **2020**, *7*, 465.  
 [3] W. Walters, T. Collins, A. Evans, *Br. J. Oral Maxillofac. Surg.* **2020**, *58*, e320-e322.  
 [4] F. Ali, D. A. Sweeney, *Clin. Infect. Dis.* **2020**, *72*, e357–e358.  
 [5] E. Williams, K. Bond, B. Zhang, M. Putland, D. A. Williamson, *J. Clin. Microbiol.* **2020**, *58*, e00776-20.  
 [6] M. J. Baker, J. Trevisan, P. Bassan, R. Bhargava, H. J. Butler, K. M. Dorling, P. R. Fielden, S. W. Fogarty, N. J. Fullwood, K. A. Heys, C. Hughes, P. Lasch, P. L. Martin-Hirsch, B. Obinaju, G. D. Sockalingum, J. Sule-Suso, R. J. Strong, M. J. Walsh, B. R. Wood, P. Gardner, F. L. Martin, *Nat. Protoc.* **2014**, *9*, 1771-1791.  
 [7] P. Heraud, P. Chatchawal, M. Wongwattanakul, P. Tippayawat, C. Doerig, P. Jearanaikoon, D. Perez-Guaita, B. R. Wood, *Malar. J.* **2019**, *18*, 348.  
 [8] S. Roy, D. Perez-Guaita, S. Bowden, P. Heraud, B. R. Wood, *Clin. Spectrosc.* **2019**, *1*, 100001.  
 [9] V. G. Barauna, M. N. Singh, L. L. Barbosa, W. D. Marcarini, P. F. Vassallo, J. G. Mill, R. Ribeiro-Rodrigues, L. C. G. Campos, P. H. Warnke, F. L. Martin, *Anal. Chem.* **2021**, *93*, 2950-2958.  
 [10] Z. Ke, J. Oton, K. Qu, M. Cortese, V. Zila, L. McKeane, T. Nakane, J. Zivanov, C. J. Neufeldt, B. Cerikan, J. M. Lu, J. Peukes, X. Xiong, H.-G. Kräusslich, S. H. W. Scheres, R. Bartenschlager, J. A. G. Briggs, *Nat.* **2020**, *588*, 498-502.  
 [11] H. J. Butler, P. M. Brennan, J. M. Cameron, D. Finlayson, M. G. Hegarty, M. D. Jenkinson, D. S. Palmer, B. R. Smith, M. J. Baker, *Nature (London)* **2019**, *10*, 1-9.  
 [12] a) D. C. Caixeta, E. M. G. Aguiar, L. Cardoso-Sousa, L. M. D. Coelho, S. W. Oliveira, F. S. Espindola, L. Raniero, K. T. B. Crosara, M. J. Baker, W. L. Siqueira, R. Sabino-Silva, *PLoS One* **2020**, *15*, e0223461; b) I. C. C. Ferreira, E. M. G. Aguiar, A. T. F. Silva, L. L. D. Santos, L. Cardoso-Sousa, T. G. Araujo, D. W. Santos, L. R. Goulart, R. Sabino-Silva, Y. C. P. Maia, *J. Oncol.* **2020**, *2020*, 4343590; c) I. Maitra, C. L. M. Morais, K. M. G. Lima, K. M. Ashton, R. S. Date, F. L. Martin, *Analyst (Cambridge, U.K.)* **2019**, *144*, 7447-7456; d) D. R. Paschotto, B. Pupin, T. T. Bhattacharjee, L. E. S. Soares, *Anal. Sci.* **2020**, *36*, 1059-1064.

[13] a) V. J. Munster, M. Koopmans, N. van Doremalen, D. van Riel, E. de Wit, *N. Engl. J. Med.* **2020**, *382*, 692-694; b) P. Hari Krishnan, *J. Craniofac. Surg.* **2020**, *31*, e653-e655.  
 [14] J. Xu, Y. Li, F. Gan, Y. Du, Y. Yao, *J. Dent. Res.* **2020**, *99*, 989.  
 [15] A. L. Wyllie, J. Fournier, A. Casanovas-Massana, M. Campbell, M. Tokuyama, P. Vijayakumar, J. L. Warren, B. Geng, M. C. Muenker, A. J. Moore, C. B. F. Vogels, M. E. Petrone, I. M. Ott, P. Lu, A. Venkataraman, A. Lu-Culligan, J. Klein, R. Earnest, M. Simonov, R. Datta, R. Handoko, N. Naushad, L. R. Sewanan, J. Valdez, E. B. White, S. Lapidus, C. C. Kalinich, X. Jiang, D. J. Kim, E. Kudo, M. Linehan, T. Mao, M. Moriyama, J. E. Oh, A. Park, J. Silva, E. Song, T. Takahashi, M. Taura, O.-E. Weizman, P. Wong, Y. Yang, S. Bermejo, C. D. Odio, S. B. Omer, C. S. Dela Cruz, S. Farhadian, R. A. Martinello, A. Iwasaki, N. D. Grubaugh, A. I. Ko, *N. Engl. J. Med.* **2020**, *383*, 1283-1286.  
 [16] L. L. Fernandes, V. B. Pacheco, L. Borges, H. K. Athwal, F. de Paula Eduardo, L. Bezinelli, L. Correa, M. Jimenez, N. Dame-Teixeira, I. M. A. Lombaert, D. Heller, *J. Dent. Res.* **2020**, 22034520960070.  
 [17] M. Peng, Y. Jie, Q. Shi, L. Ying, H. Zhu, G. Zhu, X. Ding, Z. He, J. Qin, J. Wang, H. Yan, X. Bi, B. Shen, D. Wang, L. Luo, H. Zhao, C. Zhang, Z. Lin, L. Hong, J. Li, *SSRN Electronic Journal* **2020**, 10.2139/ssrn.3541119.  
 [18] F. M. Nachtigall, A. Pereira, O. S. Trofymchuk, L. S. Santos, *Nat. Biotech.* **2020**, *38*, 1168-1173.  
 [19] F. Schwendicke, *Evid. Based Dent.* **2020**, *21*, 62.  
 [20] a) L. Caly, J. Druce, J. Roberts, K. Bond, T. Tran, R. Kostecki, Y. Yoga, W. Naughton, G. Taiaroa, T. Seemann, M. B. Schultz, B. P. Howden, T. M. Korman, S. R. Lewin, D. A. Williamson, M. G. Catton, *Med J Aust* **2020**, *212*, 459-462; b) J. Y. H. Lee, N. Best, J. McAuley, J. L. Porter, T. Seemann, M. B. Schultz, M. Sait, N. Orlando, K. Mercouliou, S. A. Ballard, J. Druce, T. Tran, M. G. Catton, M. J. Pryor, H. L. Cui, A. Luttick, S. McDonald, A. Greenhalgh, J. C. Kwong, N. L. Sherry, M. Graham, T. Hoang, M. Herisse, S. J. Pidot, D. A. Williamson, B. P. Howden, I. R. Monk, T. P. Stinear, *J. Med. Microbiol.* **2020**, *69*, 1169-1178.  
 [21] a) P. S. Masters, in *Adv. Virus Res.*, Vol. 66, Academic Press, **2006**, pp. 193-292; b) A. R. Fehr, S. Perlman, in *Coronaviruses: Methods and Protocols* (Eds.: H. J. Maier, E. Bickerton, P. Britton), Springer New York, New York, NY, **2015**, pp. 1-23.  
 [22] C. S. Goldsmith, K. M. Tatti, T. G. Ksiazek, P. E. Rollin, J. A. Comer, W. W. Lee, P. A. Rota, B. Bankamp, W. J. Bellini, S. R. Zaki, *Emerg. Infect. Dis.* **2004**, *10*, 320-326.  
 [23] B. R. Wood, *Chem. Soc. Rev.* **2016**, *45*, 1980-1998.  
 [24] a) C. P. Schultz, M. K. Ahmed, C. Dawes, H. H. Mantsch, *Anal. Biochem.* **1996**, *240*, 7-12; b) K. M. Pruitt, B. Mansson-Rahemtulla, D. C. Baldone, F. Rahemtulla, *Biochemistry* **1988**, *27*, 240-245; c) J. Tenovou, K. M. Pruitt, *J. Oral Pathol.* **1984**, *13*, 573-584.

## RESEARCH ARTICLE

## Entry for the Table of Contents



Using a modified infrared reflection accessory designed for high throughput saliva analysis we show proof of concept of a COVID-19 point-of-care screening device.

Institute and/or researcher Twitter usernames: @MonashCfb, @clirspec, @MonashUni