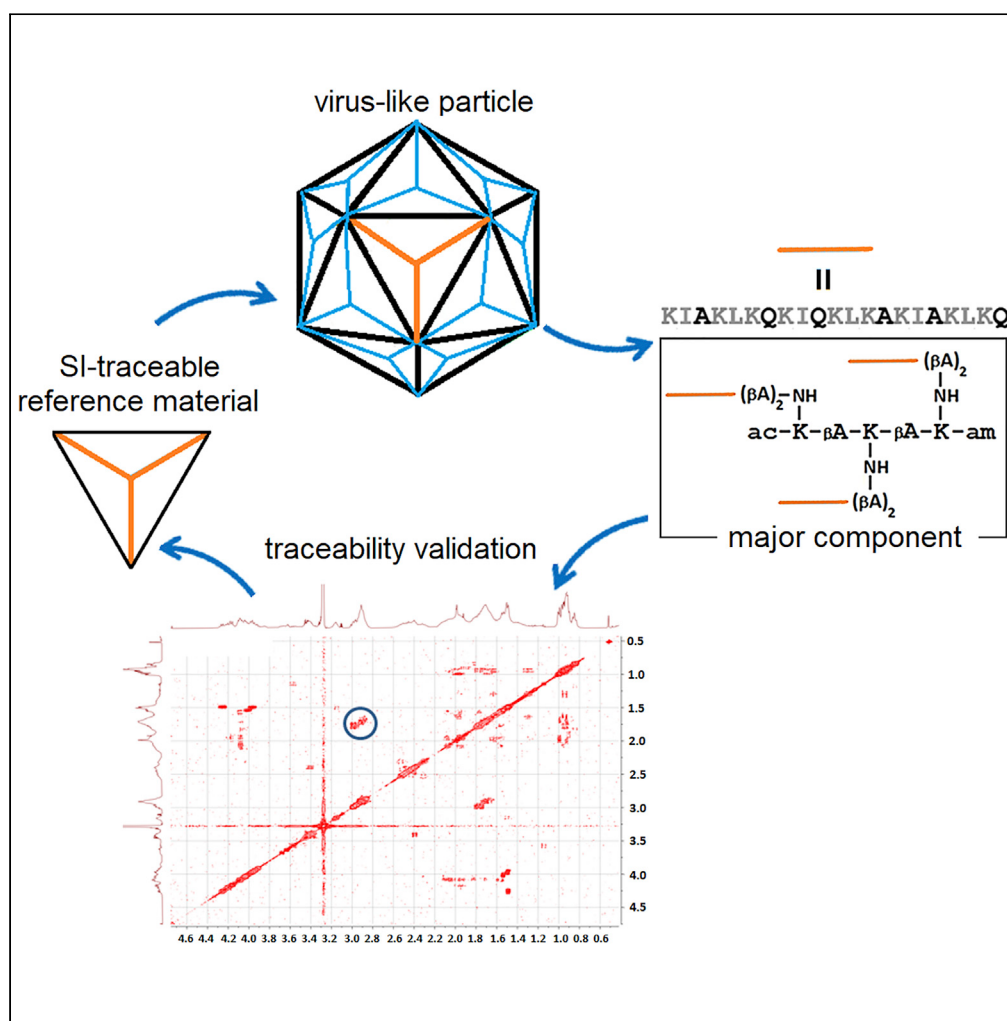


Article

An SI-traceable reference material for virus-like particles



Andrea Briones,
Gustavo Martos,
Magali Bedu,
Tiphaine Choteau,
Ralf D. Josephs,
Robert I.
Wielgosz, Maxim
G. Ryadnov

max.ryadnov@npl.co.uk

Highlights

A reference material for virus-like particles with traceability to the SI

The material is a major component of virus-like particles capable of gene delivery

Purity profile of the material is evaluated to the highest metrological order

The material allows comparability of physicochemical properties of virus-like systems

Briones et al., iScience 25,
104294
May 20, 2022 © 2022 The
Author(s).
<https://doi.org/10.1016/j.isci.2022.104294>

Article

An SI-traceable reference material for virus-like particles

Andrea Briones,^{1,2} Gustavo Martos,² Magali Bedu,² Tiphaine Choteau,² Ralf D. Josephs,² Robert I. Wielgosz,² and Maxim G. Ryadnov^{1,3,4,*}

SUMMARY

A reference material for virus-like particles traceable to the International System of Units (Système International d'Unités – the SI) is reported. The material addresses the need for developing reference standards to benchmark virus-like gene delivery systems and help harmonize measurement approaches for characterization and testing. The material is a major component of synthetic polypeptide virus-like particles produced by the state-of-the-art synthetic and analytical chemistry methods used to generate gene delivery systems. The purity profile of the material is evaluated to the highest metrological order demonstrating traceability to the SI. The material adds to the emerging toolkit of reference standards for quantitative biology.

INTRODUCTION

Continuous progress in nanotechnology and synthetic biology opens up new avenues for therapeutic and diagnostic applications (Mastrobattista et al., 2006; Tan et al., 2021). Advanced therapies emerge as particular beneficiaries, with a repertoire of biomimetic materials becoming available to access intracellular processes with a nanoscale precision (Glover et al., 2005; Mitchell et al., 2021; Zhang et al., 2018). Virus-like particles (VLPs) constitute a major type of such materials. They meet the need for non-toxic and biologically differential activities and provide effective solutions to structural uniformity by emulating the architecture of viral capsids (Fietze et al., 2016; Matsuura, 2018; Rohovie et al., 2017; Steinmetz et al., 2020). The physical attributes of VLPs find use in diagnostics as contrast agents for tissue imaging and internal standards for virus detection in biopsies as well as in medicines manufacturing as in-process calibrators to validate the efficacy of viral clearance in downstream processes (Chung et al., 2020; Johnson et al., 2017; Schwarz and Douglas, 2015). Traditional uses of VLPs such as vaccine platforms respond to new opportunities with the advent of RNA vaccines, exploring platforms capable of encapsulating RNA (Ho et al., 2021; Biddlecome et al., 2019) while emerging antimicrobial VLPs address the challenges of antimicrobial resistance including the targeting of intracellular pathogens (Kepiro et al., 2020; Rey et al., 2021).

Virus-like gene delivery systems present a rapidly evolving area of gene therapy, with the number of clinical trials steadily increasing (Ginn et al., 2018). The steady progress in the area, compounded by the lack of approved systems for systemic use, re-emphasizes the need for suitable standards, reference materials, and methods to benchmark the performance attributes of both synthetic and virus-derived gene delivery systems, and help harmonize approaches for characterization and testing (Iglesias-López et al., 2019; Faruqi et al., 2020). SI-traceable reference materials play a decisive role in this regard as they allow us to demonstrate reproducibility and traceability of the systems in correlation with their physicochemical and biological properties (Plant et al., 2014; Coxon et al., 2019).

One relevant type of reference materials relies on high-purity nucleic acids (Kalman et al., 2016). These materials are well characterized, provide an effective means to evaluate the accuracy of next-generation sequencing assays, and can support the quantification of the total nucleic acid content. However, these materials fall short of providing a quantitative insight into gene delivery systems themselves: their identity, purity, and amount. Virus serotypes can be used to evaluate the particle, genome, and infectious titres of viral vectors (Ayuso et al., 2014) but these materials are multi-component, not necessarily homogeneous and their constituents are not fully characterized, which can lead to variations in measurement results (Lock et al., 2010).

¹National Physical Laboratory, Hampton Road, Teddington TW11 0LW, UK

²Bureau International des Poids et Mesures, Pavillon de Breteuil, 92312 Sèvres, France

³Department of Physics, King's College London, London WC2R 2LS, UK

⁴Lead contact

*Correspondence: max.ryadnov@npl.co.uk
<https://doi.org/10.1016/j.isci.2022.104294>



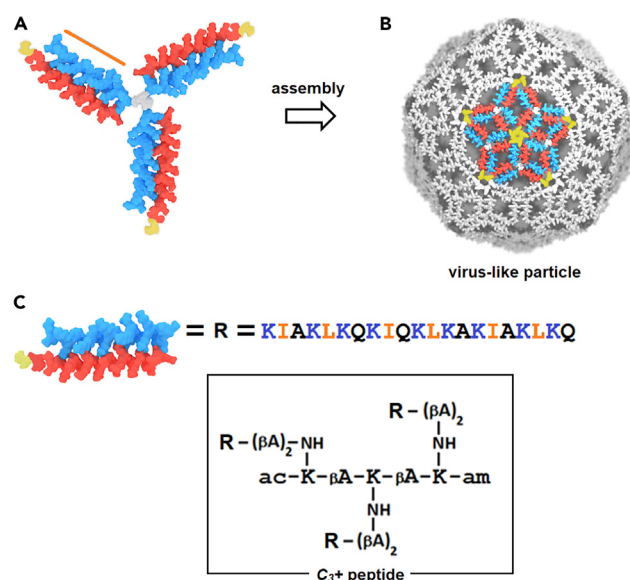


Figure 1. Virus-like particle material and its major component: C₃-triskelion peptide

(A) A molecular model of the C₃-triskelion peptide (blue) in combination with its assembly partner peptide (red) (PDB entry 4DMD rendered by Macromodel Schrödinger). Each subunit of the triskelion corresponds to the altitude of an equilateral triangle. The peptide adopts a 3-fold symmetry (C₃) matching that of triangular faces in an icosahedron, assembling into a capsid.

(B) A molecular model of T = 4 icosahedron assembled from the C₃-triskelions. For clarity, only one 5-fold axis is shown in color.

(C) A molecular model of one arm of the triskelion (blue), its corresponding sequence, and the chemical structure (box). In the sequence, lysine residues are highlighted in blue, leucines and isoleucines are in orange.

Each virus and virus-like particle assembles from multiple copies of one or more polypeptide monomers. The monomers pack around the surface of a sphere with a cubic symmetry thereby forming a regular nano-scale cage (Perlmutter and Hagan, 2015). Therefore, the quantification of a VLP or a virus comes down to the quantification of its monomer (Burlina et al., 2006; Rakowska et al., 2014).

Proteins and peptides with a therapeutic or diagnostic value are benchmarked against relevant reference polypeptide materials with an established traceability to the International System of Units (Système International d'Unités – the SI). The first and necessary step in the development of such materials is the validation of their purity with traceability established to the SI as a measure of their physicochemical properties (Josephs et al., 2019). These materials must contain a known mass fraction of the major component, which for VLPs is a reference monomer material. As per regulatory directives for advanced therapy medicinal products (e.g., 2009/120/EC), such as gene therapy products, a reference material should be relevant and specific to products and substances it is used to benchmark, while its physicochemical properties shall be characterized and documented. Therefore, a candidate for a reference material for VLPs must be synthetically accessible and able to predictably assemble into VLPs, encapsulate, and transfer nucleic acids into human cells and be devoid of the unwanted effects of replication-competent viruses. Here, we report such a material, with its purity established with traceability to the SI.

RESULTS

Virus-like particle material and reference measurement procedures

The material under study, which meets the outlined requirements for a candidate reference material, is based on the triskelion component of synthetic peptide capsids designed as virus-like icosahedra reported elsewhere (De Santis et al., 2017). This triskelion is synthetically accessible, predictably assembles into VLPs, which encapsulate and transfer genes into human cells and are devoid of the adverse effects typical of viruses (De Santis et al., 2017). The triskelion supports a C₃ rotational symmetry in the assembly of the capsids and is cationic. Therefore, it is referred to as C₃+ (Figure 1).

The C₃+ peptide was chemically synthesized, purified by ultra-performance liquid chromatography (UPLC), and lyophilized to give rise to the C₃+ material. The purity of the purified peptide estimated by LC-UV from the relative peak area response was above 98%. The purity of the peptide was then assessed against primary reference calibrants (amino acids and peptides with certified purity traceable to the SI with associated uncertainty) through an uninterrupted chain of calibrations including the associated uncertainty evaluations to establish the traceability of the mass fraction value.

The reference measurement procedure used to characterize the material has been validated for a range of peptide and protein analytes (Li et al., 2021; Little et al., 2017; Stoppacher et al., 2013; Burkitt et al., 2008). The purity assignment by the mass balance approach is achieved by employing a combination of analytical techniques to quantify the mass fraction content of all impurities present and estimate the mass fraction of the main peptide component (Westwood et al., 2013). Thus, the purity value calculated was cross-validated with two orthogonal methods based on amino acid analysis (AAA) and quantitative nuclear magnetic resonance (qNMR).

Mass balance approach to assign the purity of the material

Metrological traceability and measurement of uncertainty are prerequisites for ensuring that substances comply with international quality control regulations. In clinical chemistry, drug development, and diagnostics, purity is understood as a mass fraction of a major component that is present in the material of interest. The mass fraction for each of the impurities in the material is determined and the summation of these subtracted from the upper limit of 100% (1000 mg/g).

In this study, we utilized the mass balance method developed at the Bureau International des Poids et Mesures (BIPM) to assign the purity of the major (peptide) component in the C₃+ material (Westwood et al., 2013; Josephs et al., 2018). Under this approach, the material of interest is assessed and quantified in terms of the following:

- related impurities (w_{RS})

These are modified forms of the peptide component in the material, which can arise from side reactions during chemical synthesis, deletions, or insertions in the amino acid sequence or from the compromised stability of the final product. These are assessed by liquid chromatography-mass spectrometry (LC-MS).

- water (w_W)

An SI-traceable Karl Fisher (KF) coulometric titration method with associated water standards is a principal method to determine water content in the material. The results obtained by KF are cross-validated using microanalysis (% C, H, N content).

- residual organic solvent (total volatile organic compounds) (w_{OS})

Residual organic solvents can be determined by direct injection of a peptide solution into a gas chromatography-mass spectrometer (GC-MS) or by ¹H qNMR.

- non-volatiles (w_{NV}), with the potential contributions from inorganic impurities and non-volatile organics.

Ion chromatography (IC) is used to quantify these impurities in ionic forms including trifluoroacetate (TFA), other anions such as formate, nitrate, oxalate, chloride, acetate, phosphate, and sulfate and cations such as ammonium, magnesium, calcium, potassium, and sodium. The measurements are performed following calibration using relevant anion or cation reference materials and typically in combination with other approaches. For example, to measure the fraction of TFA as a by-product of chemical synthesis, one can apply an orthogonal approach by cross-correlating the data with that obtained by ¹⁹F-NMR, microanalysis (% N, C, H content), or TGA for the total content of non-volatile inorganic impurities.

Table 1. Mass fraction assignments of the C₃+ material components by the mass balance method

Impurity components	w	u (mg g ⁻¹)	U	Method	Traceability
Water	62.3	7.4	15	KFT	NIST SRM 2890
TFA ^a	267	1.1	2.2	¹⁹ F-qNMR	BTMSB-F ₄ internal standard
Cations	2.01	0.09	0.18	IC	
Peptide	669	7.5	15		

^aTFA was also confirmed by ion chromatography: $w = 253$, $u = 8$, $U = 16$ (mg g⁻¹); for the calculations of the combined impurities the value obtained by qNMR was used.

The identification and mass fraction assignments of impurities for the C₃+ material was performed using such an approach, with each method calibrated with standards of certified purity, ensuring that the results are traceable to the SI.

The assigned mass fraction values for all impurity types are presented in Equation 1, which allows to quantify the mass fraction of the peptide (w_P) in mg g⁻¹.

$$w_P = 1000 - (w_{RS} + w_W + w_{OS} + w_{NV}) \quad (\text{Equation 1})$$

with, w_{RS} , w_W , w_{OS} , and w_{NV} as mass fractions in mg g⁻¹ of related structural impurities, water, organic solvents, and non-volatiles, respectively.

In addition, the combined standard uncertainty of peptide purity is calculated by a quadratic combination of the standard uncertainties of each impurity assignment.

$$u(w_P) = \sqrt{u(w_{RS})^2 + u(w_W)^2 + u(w_{OS})^2 + u(w_{NV})^2} \quad (\text{Equation 2})$$

In the cases where there is no statistically significant difference between the blank control and sample, the assigned value is zero with an associated uncertainty corresponding to the limit of detection of the method.

The mass fraction of non-volatile impurities in the C₃+ material such as TFA was determined by IC and ¹⁹F-qNMR. For IC measurements, oxytocin was used as a peptide reference control, while external instrument calibrations were performed with appropriate reference materials (see STAR Methods) (Li et al., 2021). The results were obtained for two series of measurements performed on different days.

IC revealed TFA as a significant component in the C₃+ material, with lower signals (<0.1% of that for TFA) observed for acetate, formate, chloride, nitrate, and sulfate ions (Figure S1A). The TFA mass fraction value determined by IC was 253 ± 16 mg g⁻¹. This value agreed with the TFA mass fraction determined by ¹⁹F qNMR (267.0 ± 2.2 mg g⁻¹). The NMR signal used for TFA quantification was the singlet at -76.6 ppm, along with its four asymmetric satellite peaks due to the ¹⁹F coupling to ¹³C present either in the α (large J coupling) or β position (small J coupling) of the TFA molecule (Figure S1B). IC results also revealed small amounts of cations (mainly Na⁺, K⁺, and Ca²⁺) corresponding to 2.01 ± 0.18 mg g⁻¹.

Signals at chemical shift regions characteristic of residual solvents were searched for in the ¹H NMR spectra as well. Only some trace levels of ethanol and dimethyl sulfoxide could be ascertained (not shown).

Water content was measured by Karl Fischer titration (KFT), which indicated the mass fraction of water (w_W) in the material to be 62.3 ± 15 mg/g ($k = 2$) (Figure S2). Table 1 shows mass fraction assignments and associated expanded uncertainties for each of the contributing impurities and the pure peptide in the material.

After subtracting impurity contributions from the upper limit of 100% or 1000 mg g⁻¹, the peptide component mass fraction in the material was 669 ± 15 mg g⁻¹. The mass fractions of elemental carbon, hydrogen, and nitrogen determined by elemental analysis were compared to the expected mass fraction values as determined from the mass balance approach.

Table 2. Elemental analysis results for the C₃+ material

Elemental content (mol/mol)	C ₃ + peptide (C ₃₈₆ H ₇₂₀ N ₁₁₄ O ₈₄)	TFA (C ₂ O ₂ F ₃)	H ₂ O	Cation	Total (found, mol)	Total (found, g/mol)	Relative amount (%)	
							calc.	found.
N	114	0	0	0	114	1597	12.90	12.88
C	386	58	0	0	444	5333	43.08	43.39
H	720	0	86	0	806	812	6.56	6.67
O	84	58	43	0	185	2960	23.91	N.A.
F	0	87	0	0	87	1653	13.35	N.A.
cations	0	0	0	0.6	0.6	24	0.19	N.A.

Thus, the impurity mass fractions served to calculate each impurity molar ratio to the C₃+ peptide resulting in the material composition formula: C₃₈₆H₇₂₀N₁₁₄O₈₄ · (H₂O)₄₃ · (TFA)₂₉ · (cations)_{0.6}. Subsequently, the theoretical C, H, and N mass fractions were calculated and compared to the elemental analysis results (Table 2). The results from the two methods agreed to within <2%, providing an independent confirmation of the mass balance results.

Amino acid analysis isotope dilution mass spectrometry (AAA-IDMS)

The purity values for the peptide component by AAA-IDMS were calculated from mass fractions of leucine and isoleucine residues. The two amino acid residues are the constituent amino acids of the C₃⁺ material, which are stable under the conditions of complete hydrolysis and yield excellent recovery rates for quantification. This is supported by the results of key comparison studies (Josephs et al., 2017), under the Consultative Committee for Amount of Substance (CCQM). Importantly, for both amino acid residues calibrator, amino acid materials with assigned SI-traceable purity values are available, which allowed for accurate and traceable quantification of the material. The values for leucine were slightly higher (657 ± 14 mg/g) than those for isoleucine (614 ± 56 mg/g), averaging the final purity value to 636 ± 45 mg/g (Figure 2A).

Given the variations between the two residues, different contributions of uncertainty sources to the final purity value were assessed (Figure 2B). For leucine the major uncertainty source was the intermediate precision, *u*(IntPre), which accounts for variability in the efficacy of hydrolysis and the signal integrations of mass spectra for injection replicates. Other minor uncertainty contributors are from weighing and the purity of the amino acid standards used in the analysis, *u*(IDMS) (Figure 2B).

¹H quantitative nuclear magnetic resonance (¹H qNMR)

In accordance with the AAA-IDMS data, analyses by ¹H qNMR returned the value of 639 ± 16 mg g⁻¹ (*k* = 2) for the mass fraction of the peptide in the material. The calculated value and uncertainty were based on the resonance signals of the lysine ε-CH₂ and leucine and isoleucine methyl protons. Proton assignments in the peptide were performed using homonuclear correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY), and by comparing chemical shifts obtained to those of free amino acids in solution. Spectral regions for lysine ε-CH₂ protons (labeled D) and for leucine and isoleucine -CH₃ protons (labeled B and C) were integrated for comparison and to display relative proton stoichiometry (Table 3 and Figure 3). The peptide was measured against an internal standard.

Proton-proton correlations in COSY and TOCSY experiments are indicated in brackets (Figure 3A). BTMSB-F₄ was used as the internal standard and was chosen to perform ¹H and ¹⁹F qNMR measurements on the same samples (Westwood et al., 2019). The mass fraction of the peptide was calculated based on lysine ε-methylene and leucine and isoleucine methyl ¹H NMR signals, while the TFA mass fraction was calculated based on the trifluoromethyl ¹⁹F NMR signal.

The purity mass fraction values of 636 ± 45 mg g⁻¹ by AAA-IDMS and 639 ± 16 mg/g by ¹H qNMR were averaged to a total peptide mass fraction of 638 ± 24 mg g⁻¹. Related structure impurities were investigated by LC-hrMS in order to rule out the presence of major peptide components other than C₃+ peptide. Despite numerous efforts to separate chromatographically C₃+ peptide from minor co-eluting species of

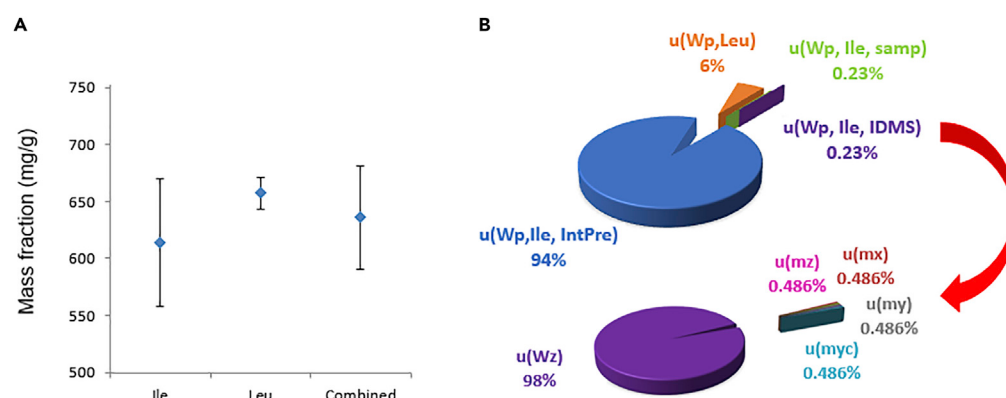


Figure 2. The purity of the C₃+ material established by amino acid analysis

(A) Purity values for peptide in C₃+ material calculated from the contributions of leucine and isoleucine residues, with the averaged value of 636 ± 45 mg/g. Error bars are expanded uncertainties ($k = 2$).

(B) Relative uncertainty contributions to the final assigned purity value of the C₃+ material, where $u(W_p)$ is the standard uncertainty of the peptide mass fraction, $u(\text{IntPre})$ is an uncertainty due to the variance in the hydrolysis and the MS signal between injections, $u(\text{IDMS})$ is the combined uncertainty of the standard amino acid solution mass fraction and weighing, my is the mass of labeled amino acid solution in sample, myc is the mass of labeled amino acid solution in the calibrant, mz is the mass of standard amino acid solution in calibrant, Wz is the mass fraction of the standard amino acid solution, mx is the mass of peptide solution in sample.

similar mass ($\pm 5\%$), only one impurity could be clearly differentiated from in-source or adduct ions (Figure S3). This impurity of mass 7918.31 Da was attributed to a modified C₃+ peptide that lacked 3 Lys residues and its mass fraction was estimated to be 18 mg g^{-1} with an asymmetric expanded uncertainty of $+21 \text{ mg g}^{-1}$ and -18 mg g^{-1} . In the absence of an appropriate standard, this value remained indicative only and was not used to correct the calculated total peptide mass fraction.

Taking into consideration the expanded uncertainty values for different measurements, it can be concluded that the total peptide mass fraction value calculated by the mass balance approach agreed with the combined value calculated through orthogonal methods based on isotope dilution amino acid analysis and ¹H qNMR (Figure 3C).

DISCUSSION

The field of VLPs is burgeoning with many systems available commercially. The applications range from research tool kits to therapeutics approved for clinic or undergoing clinical trials. VLPs developed as vaccines against different viruses, from hepatitis B virus to SARS-CoV-2, have been licensed and commercialized (Van Bömmel and Berg, 2021; Knoll and Wonodi, 2020), with at least over 100 vaccines available against influenza alone (Quan et al., 2016). Similarly, in gene therapy, virus-derived and VLP-based therapeutics continue to dominate clinical trials (Ginn et al., 2018). However, the growing number of commercial systems demands comparability assessments from early stages, i.e., R&D, to manufacturing. The need is further echoed by existing and emerging requirements regulatory policies set up for vendors to demonstrate the traceability of their products. Given that VLPs vary in composition, there remain difficulties in the comparability of their physicochemical properties against the highest point of reference, i.e., SI-traceable reference materials, which is lacking. Therefore, the first SI reference materials should be developed to benchmark the physicochemical properties of VLPs to help address variabilities in their composition.

Among the key drivers to develop synthetic VLPs for intracellular delivery, one is to overcome the inherent drawbacks of viruses including limited loading capacities and insertional mutagenesis. Designs from the bottom up allow better predictability for relating structure with activity and increasing the complexity of next generation designs to be endowed with specialist functions. Emulating the virus architecture appears to be an obvious route for designing VLP systems as it allows us to capitalize on the advantageous properties of viruses such as morphological uniformity and modularity. Top-down designs of virus-derived systems are equally valuable as they build upon bio-functional optimizations of the virus architecture. However, irrespective of apparent differences between bottom-up and top-down designs, quantification and traceability are the ultimate quality attributes of all gene delivery systems. SI-traceable reference materials

Table 3. Partial assignment of ^1H NMR signals in the C_3^+ material based on chemical shifts (δ) and observed 2D correlations in COSY and TOCSY spectra

Multiplet	H δ (ppm)	Id	#H/molec	Evidence
A	0.37	BTMSB- F_4 [$(-\text{CH}_3)_3$] ₂	18	δ
B	0.85	Ile (γCH_3 -)	27 (9 Ile x 3)	δ , COSY [1.82 (Ile βCH)], TOCSY [1.17 (Ile γH), 3.63]
C	0.93	Ile ($\gamma\text{-CH}_3$, $\delta\text{-CH}_3$) & Leu $\gamma\text{-CH}_3$) ₂	108 (18 Ile/Leu x 6)	δ , COSY [1.29, 1.71, 1.99], TOCSY [3.63, 4.07 (αH Leu and Ile)]
D	2.91	Lys (ϵCH_2 -)	54 (27 Lys x 2)	δ , COSY [1.72 (δCH_2)], TOCSY [1.46, 1.72, 1.95, 3.97]
E	8–8.6	Backbone NH protons (partial D exchanged)		δ

provide an optimal solution in this regard and help meet emerging requirements for commercial vendors to demonstrate traceability for their products and technologies. As a rule of thumb, the purity of such reference materials must be demonstrated to the highest metrological order. In this report, we described the characterization of a synthetic reference material demonstrating its traceability to the SI. The material was produced by the state-of-the-art synthetic and analytical approaches used by the developers and manufacturers of commercial systems, thus ensuring compatibility with their processes and final products. Impurities of the material reported are typical of synthetic materials and hence the obtained impurity profile can be used to benchmark those of commercial systems. Impurities were quantified and subtracted from the total sample mass to obtain the content of the major component, i.e. the C_3^+ peptide. The mass fraction of the total peptide content was $669 \pm 15 \text{ mg g}^{-1}$ based on the mass balance approach, which was cross-validated using an orthogonal approach consisting in the combination of qNMR and isotope dilution amino acid analysis that resulted in a confirmatory value of $638 \pm 24 \text{ mg g}^{-1}$. The reported reference material can be used for the characterization and validation of native and synthetic virus-like particles as well as other peptide- and protein-based drug delivery systems. The applicability of the material is multifold. For example, it can be used to confirm the purity of a commercial VLP, or virus-derived products as required by the manufacturer or a regulatory body or can help quantify the encapsulation efficacy of a designed gene delivery system. Often, recombinantly prepared VLPs contain nucleic acids, either via nonspecific encapsulation of host nucleic acids or via designed encapsulation. The challenge here is that the average amount of nucleic acids per VLP remains uncertain. With its proven ability to encapsulate nucleic acids upon assembly into a VLP (De Santis et al., 2017), the reported material can be adapted to establish the ratio of loaded versus empty VLP particles and probe the stoichiometry between protein and gene constituents. All such validations can be performed in different sample matrices, *in vitro* and in cell extracts, thus providing an SI-traceable evaluation route for the amount of a desired material in target media. The material, the first of the type to the best of our knowledge, adds to the emerging toolkit of reference standards for quantitative biology.

Limitations of the study

The study introduced an SI-traceable reference material for virus-like particles as gene delivery systems. It can also be used for related applications which require the quantification of an exogenous material in a given biological matrix. Biomaterials developed for immunoengineering and engineering biology are exemplar areas of material dissemination. Practical considerations for benchmarking experimental gene delivery systems or other systems, e.g. vaccines, are not discussed in the study. The study also does not address the promise of the material for correlative imaging applications, e.g., the monitoring of transfection in live cells, the measurement and trafficking of VLPs *in vivo*, or the quantification of nucleic acid content per VLP, which would require orthogonal measurements, e.g., PCR, and complementary reference materials, i.e., nucleic acids. These types of analysis merit a series of follow-up investigations focusing on the development of appropriate extraction and measurement protocols and associated uncertainty evaluations.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

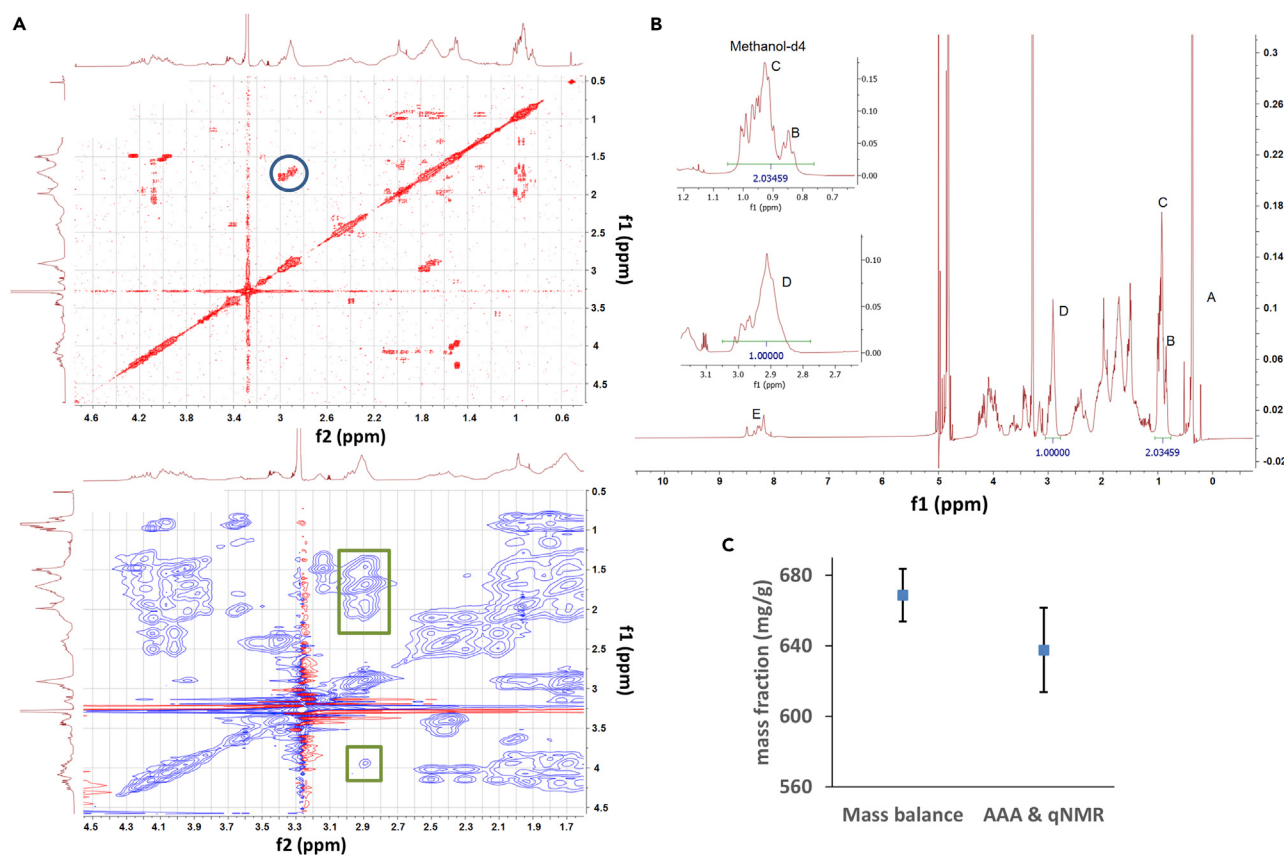


Figure 3. Quantitative analysis of the C₃⁺ material by qNMR

(A) 2D homonuclear COSY (top) and TOCSY (bottom) correlations for the C₃⁺ material. Cross signals highlighted in blue and green indicate proton-proton interactions within the same lysine spin systems.

(B) ¹H NMR spectrum of the peptide containing BTMSB-F₄ as internal standard. Partially assigned regions are indicated with letters (see Table 3) and are integrated to display the relative proton stoichiometry.

(C) Comparison of results of the mass balance method and AAA & qNMR. Where, the value obtained for the C₃⁺ peptide in the mass balance method was of 669 ± 15 mg/g (k = 2) and the averaged value obtained for the measurements with the AAA & qNMR was 638 ± 24 mg/g (k = 2). Error bars are expanded uncertainties (k = 2).

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead contact
 - Materials availability
 - Specialist reagents
 - Data and code availability
- **METHOD DETAILS**
 - General analytical information
 - Liquid chromatography high resolution tandem mass spectrometry (LC-hrMS/MS)
 - AAA coupled to IDMS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2022.104294>.

ACKNOWLEDGMENTS

We acknowledge funding from the UK's Industrial Strategy Challenge Fund and thank Prof Paul Brewer for the critical assessment of the work.

AUTHOR CONTRIBUTIONS

All authors designed and performed the experiments. A.B. and M.G.R. wrote the manuscript. All authors analysed the data and contributed to the writing of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 31, 2022

Revised: April 6, 2022

Accepted: April 20, 2022

Published: May 20, 2022

REFERENCES

- Ayuso, E., Blouin, V., Lock, M., McGorray, S., Leon, X., Alvira, M.R., Auricchio, A., Bucher, S., Chtarto, A., Clark, K.R., et al. (2014). Manufacturing and characterization of a recombinant adeno-associated virus type 8 reference standard material. *Hum. Gene Ther.* 25, 977–987. <https://doi.org/10.1089/hum.2014.057>.
- Biddlecome, A., Habte, H.H., McGrath, K.M., Sambanthamoorthy, S., Wurm, M., Sykora, M.M., Knobler, C.M., Lorenz, I.C., Lasaro, M., Elbers, K., and Gelbart, W.M. (2019). Delivery of self-amplifying RNA vaccines in vitro reconstituted virus-like particles. *PLoS One* 14, e0215031. <https://doi.org/10.1371/journal.pone.0215031>.
- Burkitt, W.I., Pritchard, C., Arsene, C., Henrion, A., Bunk, D., and O'Connor, G. (2008). Toward Système International d'Unité-traceable protein quantification: from amino acids to proteins. *Anal. Biochem.* 376, 242–251. <https://doi.org/10.1016/j.ab.2008.02.010>.
- Burlina, F., Sagan, S., Bolbach, G., and Chassaing, G.A. (2006). A direct approach to quantification of the cellular uptake of cell-penetrating peptides using MALDI-TOF mass spectrometry. *Nat. Protoc.* 1, 200–205. <https://doi.org/10.1038/nprot.2006.30>.
- Chung, Y.H., Cai, H., and Steinmetz, N.F. (2020). Viral nanoparticles for drug delivery, imaging, immunotherapy, and theranostic applications. *Adv. Drug Deliv. Rev.* 156, 214–235. <https://doi.org/10.1016/j.addr.2020.06.024>.
- Coxon, C.H., Longstaff, C., and Burns, C. (2019). Applying the science of measurement to biology: why bother? *PLoS Biol.* 17, e3000338. <https://doi.org/10.1371/journal.pbio.3000338>.
- De Santis, E., Alkassam, H., Lamarre, B., Faruqui, N., Bella, A., Noble, J.E., Micale, N., Ray, S., Burns, J.R., Yon, A.R., et al. (2017). Antimicrobial peptide capsids of de novo design. *Nat. Commun.* 8, 2263. <https://doi.org/10.1038/s41467-017-02475-3>.
- Faruqui, N., Kummrow, A., Fu, B., Divieto, C., Rojas, F., Kisulu, F., Cavalcante, J.J.V., Wang, J., Campbell, J., Martins, J.L., et al. (2020). Cellular metrology: scoping for a value proposition in extra- and intracellular measurements. *Front. Bioeng. Biotechnol.* 7, 456. <https://doi.org/10.3389/fbioe.2019.00456>.
- Frietze, K.M., Peabody, D.S., and Chackerian, B. (2016). Engineering virus-like particles as vaccine platforms. *Curr. Opin. Virol.* 18, 44–49. <https://doi.org/10.1016/j.coviro.2016.03.001>.
- Ginn, S.L., Amaya, A.K., Alexander, I.E., Edelstein, M., and Abedi, M.R. (2018). Gene therapy clinical trials worldwide to 2017: an update. *J. Gene Med.* 20, e3015. <https://doi.org/10.1002/jgm.3015>.
- Glover, D.J., Lipps, H.J., and Jans, D.A. (2005). Towards safe, non-viral therapeutic gene expression in humans. *Nat. Rev. Genet.* 6, 299–310. <https://doi.org/10.1038/nrg1577>.
- Ho, W., Gao, M., Li, F., Li, Z., Zhang, X.Q., and Xu, X. (2021). Next-generation vaccines: nanoparticle-mediated DNA and mRNA delivery. *Adv. Healthc. Mater.* 10, e2001812. <https://doi.org/10.1002/adhm.202001812>.
- Iglesias-López, C., Agustí, A., Obach, M., and Vallano, A. (2019). Regulatory framework for advanced therapy medicinal products in Europe and United States. *Front. Pharmacol.* 10, 921. <https://doi.org/10.3389/fphar.2019.00921>.
- Johnson, S., Brorson, K.A., Frey, D.D., Dhar, A.K., and Cetlin, D.A. (2017). Characterization of non-infectious virus-like particle surrogates for viral clearance applications. *Appl. Biochem. Biotechnol.* 183, 318–331. <https://doi.org/10.1007/s12010-017-2447-y>.
- Josephs, R.D., Li, M., Song, D., Westwood, S., Stoppacher, N., Dairea, A., Choteau, T., Wielgosz, R., Xiao, P., Liu, Y., et al. (2017). Key comparison study on peptide purity – synthetic human C-peptide. *Metrologia* 54, 08007. <https://doi.org/10.1088/0026-1394/54/1a/08007>.
- Josephs, R.D., Stoppacher, N., Dairea, A., Choteau, T., Lippa, K.A., Phinney, K.W., Westwood, S., and Wielgosz, R.I. (2018). State-of-the-art and trends for the SI traceable value assignment of the purity of peptides using the model compound angiotensin I. *Trends Anal. Chem.* 101, 108–119. <https://doi.org/10.1016/j.trac.2017.09.026>.
- Josephs, R.D., Martos, G., Li, M., Wu, L., Melanson, J.E., Quaglia, M., Beltrão, P.J., Prevoo-Franzen, D., Boeuf, A., Delatour, V., et al. (2019). Establishment of measurement traceability for peptide and protein quantification through rigorous purity assessment – a review. *Metrologia* 56, 044006. <https://doi.org/10.1088/1681-7575/ab27e5>.
- Kalman, L.V., Datta, V., Williams, M., Zook, J.M., Salit, M.L., and Han, J.Y. (2016). Development and characterization of reference materials for genetic testing: focus on public partnerships. *Ann. Lab. Med.* 36, 513–520. <https://doi.org/10.3343/alm.2016.36.6.513>.
- Kepiro, I.E., Marzuoli, I., Hammond, K., Ba, X., Lewis, H., Shaw, M., Gunnoo, S.B., De Santis, E., Łapińska, U., Pagliara, S., et al. (2020). Engineering chirally blind protein pseudocapsids into antibacterial persisters. *ACS Nano* 14, 1609–1622. <https://doi.org/10.1021/acsnano.9b06814>.
- Knoll, M.D., and Wonodi, C. (2020). Oxford-AstraZeneca COVID-19 vaccine efficacy. *Lancet* 397, 72–74. [https://doi.org/10.1016/s0140-6736\(20\)32623-4](https://doi.org/10.1016/s0140-6736(20)32623-4).
- Li, M., Josephs, R.D., Dairea, A., Choteau, T., Westwood, S., Martos, G., Wielgosz, R.I., and Li, H. (2021). Structurally related peptide impurity identification and accurate quantification for synthetic oxytocin by liquid chromatography-high-resolution mass spectrometry. *Anal. Bioanal. Chem.* 413, 1861–1870. <https://doi.org/10.1007/s00216-021-03154-5>.
- Little, R.R., Wielgosz, R.I., Josephs, R., Kinumi, T., Takatsu, A., Li, H., Stein, D., and Burns, C. (2017). Implementing a reference measurement system for C-peptide: successes and lessons learned. *Clin. Chem.* 63, 1447–1456. <https://doi.org/10.1373/clinchem.2016.269274>.
- Lock, M., McGorray, S., Auricchio, A., Ayuso, E., Beecham, E.J., Blouin-Tavel, V., Bosch, F., Bose, M., Byrne, B.J., Caton, T., et al. (2010). Characterization of a recombinant adeno-associated virus type 2 reference standard material. *Hum. Gene Ther.* 21, 1273–1285. <https://doi.org/10.1089/hum.2009.223>.
- Mastrobattista, E., van der Aa, M.A., Hennink, W.E., and Crommelin, D.J.A. (2006). Artificial viruses: a nanotechnological approach to gene delivery. *Nat. Rev. Drug Discov.* 5, 115–121. <https://doi.org/10.1038/nrd1960>.
- Matsuura, K. (2018). Synthetic approaches to construct viral capsid-like spherical nanomaterials. *Chem. Commun. (Camb.)* 54, 8944–8959. <https://doi.org/10.1039/c8cc03844a>.
- Mitchell, M.J., Billingsley, M.M., Haley, R.M., Wechsler, M.E., Peppas, N.A., and Langer, R. (2021). Engineering precision nanoparticles for drug delivery. *Nat. Rev. Drug Discov.* 20, 101–124. <https://doi.org/10.1038/s41573-020-0090-8>.

- Perlmutter, J.D., and Hagan, M.F. (2015). Mechanisms of virus assembly. *Annu. Rev. Phys. Chem.* 66, 217–239. <https://doi.org/10.1146/annurev-physchem-040214-121637>.
- Plant, A.L., Locascio, L.E., May, W.E., and Gallagher, P.D. (2014). Improved reproducibility by assuring confidence in measurements in biomedical research. *Nat. Methods* 11, 895–898. <https://doi.org/10.1038/nmeth.3076>.
- Quan, F.S., Lee, Y.T., Kim, K.H., Kim, M.C., and Kang, S.M. (2016). Progress in developing virus-like particle influenza vaccines. *Expert Rev. Vaccines* 15, 1281–1293. <https://doi.org/10.1080/14760584.2016.1175942>.
- Rakowska, P.D., Lamarre, B., and Ryadnov, M.G. (2014). Probing label-free intracellular quantification of free peptide by MALDI-ToF mass spectrometry. *Methods* 68, 331–337. <https://doi.org/10.1016/j.ymeth.2014.03.012>.
- Rey, S., Faruqi, N., Hoose, A., Dondi, C., and Ryadnov, M.G. (2021). Designer protein pseudo-capsids targeting intracellular bacteria. *Biomater. Sci.* 9, 6807–6812. <https://doi.org/10.1039/d1bm01235e>.
- Rohovie, M.J., Nagasawa, M., and Swartz, J.R. (2017). Virus-like particles: next-generation nanoparticles for targeted therapeutic delivery. *Bioeng. Transl. Med.* 2, 43–57. <https://doi.org/10.1002/btm2.10049>.
- Schwarz, B., and Douglas, T. (2015). Development of virus-like particles for diagnostic and prophylactic biomedical applications. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* 7, 722–735. <https://doi.org/10.1002/wnan.1336>.
- Steinmetz, N.F., Lim, S., and Sainsbury, F. (2020). Protein cages and virus-like particles: from fundamental insight to biomimetic therapeutics. *Biomater. Sci.* 8, 2771–2777. <https://doi.org/10.1039/d0bm00159g>.
- Stoppacher, N., Josephs, R.D., Dairea, A., Choteau, T., Westwood, S.W., and Wielgosz, R.I. (2013). Impurity identification and determination for the peptide hormone angiotensin I by liquid chromatography-high-resolution tandem mass spectrometry and the metrological impact on value assignments by amino acid analysis. *Anal. Bioanal. Chem.* 405, 8039–8051. <https://doi.org/10.1007/s00216-013-6953-7>.
- Tan, X., Letendre, J.H., Collins, J.J., and Wong, W.W. (2021). Synthetic biology in the clinic: engineering vaccines, diagnostics, and therapeutics. *Cell* 184, 881–898. <https://doi.org/10.1016/j.cell.2021.01.017>.
- Van Bömmel, F., and Berg, T. (2021). Three are better than one – increasing HBV seroprotection by a tri-antigenic vaccine. *Lancet Infect. Dis.* 21, 1197–1198. [https://doi.org/10.1016/s1473-3099\(20\)30845-8](https://doi.org/10.1016/s1473-3099(20)30845-8).
- Westwood, S., Choteau, T., Dairea, A., Josephs, R.D., and Wielgosz, R.I. (2013). Mass balance method for the SI value assignment of the purity of organic compounds. *Anal. Chem.* 85, 3118–3126. <https://doi.org/10.1021/ac303329k>.
- Westwood, S., Yamazaki, T., Huang, T., Garrido, Bruno, Ü, Zhang, W., Martos, G., Stoppacher, N., Saito, T., and Wielgosz, R. (2019). Development and validation of a suite of standards for the purity assignment of organic compounds by quantitative NMR spectroscopy. *Metrologia* 56, 064001. <https://doi.org/10.1088/1681-7575/ab45cb>.
- Zhang, W., Zhang, X.E., and Li, F. (2018). Virus-based nanoparticles of simian virus 40 in the field of nanobiotechnology. *Biotechnol. J.* 13, e1700619. <https://doi.org/10.1002/biot.201700619>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Acetonitrile $\geq 99.9\%$, HPLC grade	VWR	20060.320
Formic acid $\geq 99.9\%$, HPLC grade	VWR	84865.260P
Sodium trifluoroacetate	Merck	132101
Sodium bicarbonate/carbonate	Merck	62414
HYDRANAL TM – Coulomat CG-K reagent	VWR	34821
HYDRANAL TM – Coulomat AG-oven	VWR	34836
Lactose standard 5%	Merck	112939
NIST SRM 2890 standard	Merck	NIST2890
Acetanilide	Merck	NIST141E
Magnesium oxide, 99.99%	Merck	203718
2,3,5,6-tetrafluoro-1,4-bis(trimethylsilyl)-benzene (BTMSB-F ₄)	NMIJ	CRM 4602a
Isotopically labelled L-Ile- ¹³ C ₆	Cambridge Isotope laboratories	CLM-2248-H-PK
Isotopically labelled L-Leu- ¹³ C ₆ - ¹⁵ N	Cambridge Isotope laboratories	CNLM-615-PK
Naturally occurring L-Ile	Sigma-Aldrich	I2752
Naturally occurring L-Leu	Sigma-Aldrich	L8000
Software and algorithms		
MestReNova version 14	Mestrelab Research	https://mestrelab.com/download/mnova/
Tiamo version 2.3	Metrohm	https://www.metrohm.com/en_nl/products/6/6056/66056923.html
Analyst version 1.7.1	SCIEX	https://sciex.com/products/software/analyst-software
Other		
Jupiter LC column, 250 mm × 4.6 mm, 5 μ m, 300 Å	Phenomenex	00G-4053-E0
Primesep 100 column, 250 mm × 4.6 mm, 5 μ m, 100 Å	SIELC Technologies	100-46.250.0510

RESOURCE AVAILABILITY

Lead contact

Further information and requests for reagents may be directed to, and will be fulfilled by, the lead contact, Maxim G. Ryadnov (max.ryadnov@npl.co.uk).

Materials availability

This study generated an SI-traceable reference material for the physicochemical characterisation of viral-like particles.

Specialist reagents

Ultrapure water (18.2 M Ω ·cm⁻¹) was produced using a Milli-Q system from Millipore (Molsheim, France). The C₃₊ peptide was purchased from CEM (North Carolina, USA). Isotopically labelled amino acids L-Ile-¹³C₆ and L-Leu-¹³C₆-¹⁵N and naturally occurring L-Ile and L-Leu were characterised in-house for purity after the purchase as per the [key resources table](#).

Data and code availability

- All data supporting the measurements carried on the C₃+ peptide can be found within the manuscript and the [supplemental information](#) or can be shared by the [lead contact](#) upon request.
- The paper does not report an original code. All software and pipelines used in the study are listed and referenced in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request

METHOD DETAILS

General analytical information

An LC series 1200 instrument (Agilent, Waldbronn, Germany) consisting of a model G1379B degasser, G1312B binary pump SL, G1367C autosampler SL, and G1316B column thermostat SL was employed for LC-hrMS/MS analysis. The LC was coupled online to an LTQ-Orbitrap XL high-resolution tandem mass spectrometer (Thermo Electron, Bremen, Germany) fitted with a heated electrospray ionization (HESI) source. For AAA by IDMS, a microwave hydrolysis system (CEM, Saclay, France) and a rotational vacuum concentrator connected to a freeze-dryer system (Christ, Germany) were used to hydrolyse and concentrate samples, respectively. An Exion LC coupled online to a QTRAP 6500+ mass spectrometer (AB Sciex, Les Ulis, France) fitted to an electrospray ionisation (ESI) source was used for the detection of amino acids. A 930 compact IC flex and 863 compact autosampler (Metrohm AG, Switzerland) and a Metrohm Titrand 851 coupled to a Metrohm 874 Oven sample processor were used for IC and KF titration, respectively. Elemental analysis was performed using a Vario Micro Cube Elemental Analyser (Elementar, Lyon, France), whereas NMR measurements were performed on an ECS-400 MHz NMR instrument equipped with a direct type (Royal) automatic tuning probe (JEOL, Tokyo, Japan).

Liquid chromatography high resolution tandem mass spectrometry (LC-hrMS/MS)

LC conditions

Chromatographic separations were performed using a Jupiter column, 250 mm × 4.6 mm, 5 μm, 300 Å (Phenomenex, Le Pecq, France). After optimisation, the following conditions were used for all subsequent LC-hrMS/MS experiments with C₃+ peptide solutions: flow rate at 1 mL/min, the column thermostat temperature at 38°C, the autosampler rack temperature at 5°C, and the injection volume at 10 μL. Eluent A was water containing 0.1% formic acid (v/v) and eluent B was AcN containing 0.1% formic acid (v/v). The LC gradient programme used was as follows: 10% eluent B (5 min hold), gradient to 30% eluent B in 55.1 min, gradient to 90% B in 10 s (5 min hold), back to 30% B in 10 s, and gradient back to 10% B in 10 min, and hold for 10 min (total method duration 82 min).

MS parameters

The LTQ-Orbitrap-XL MS instrument was operated in a positive ionisation mode. The spray voltage was set at +4 kV. The flow rate sheath gas was 90 a.u. and the flow rate of the auxiliary gas was 5 a.u. Capillary temperature was 370°C and the tube lens was set at 105 V. The full scan detection mode at 30000 resolution and dependent MS/MS (HCD activation) scans at 15000 resolution targeting the most intense ions were used across the chromatographic elution.

AAA coupled to IDMS

Preparation of peptide and standard solutions

An internal standard (IS) was prepared in ultrapure water containing the ¹³C isotopically labelled amino acids. Similarly, a standard mixture prepared in ultrapure water contained the unlabelled amino acid reference materials of known purity. The IS was added to the standard mixture and to the C₃+ peptide solution to prepare the calibration blend and the sample blend, respectively. The calibration blend was prepared so that it contained the same amount of unlabelled and labelled amino acids as expected in the sample blend after acid hydrolysis (exact matching). Furthermore, the amount of IS was adjusted so as to produce the same signal intensity as the unlabelled amino acids in the mass spectrometer during LC-MS/MS analysis. The blends were prepared in glass microvials and volumetric additions were controlled gravimetrically.

Hydrolysis

All solutions prepared were transferred and dried in a rotational vacuum concentrator coupled to a freeze-dryer vacuum and left overnight. These were then placed in a microwave hydrolysis system to hydrolyze amino acids under the following settings: 150 W and 170°C for 3 h. Upon completion, microvials were transferred to a rotational vacuum concentrator coupled to a freeze dryer system and left to dry overnight. Samples were prepared for LC-MS/MS analyses by adding 100 µL of HPLC mobile phase (AcN/water at 35v/65v).

LC conditions

Chromatographic separations were performed on a Primesep 100 column, 250 mm × 4.6 mm, 5µm, 100 Å (SIELC Technologies, Wheeling, USA). After optimization, the following conditions were used for all LC-MS/MS experiments for AAA: flow rate at 1 mL/min, the column thermostat temperature at 25°C, the autosampler rack temperature at 5°C, and the injection volume at 10 µL. The eluent A consisted of AcN/water (35v/65v) and eluent B was water/AcN (35v/65v) with 1% formic acid. The LC gradient program was as follows: 60% eluent B (1 min hold), gradient to 100% eluent B in 31 min (3 min hold), back to 60% B in 2 min and hold for 4 min (total method duration 40 min).

MS parameters

The instrument was operated in a positive mode (ESI+). The ion spray voltage was set at +5.5 kV with ion source temperature at 500°C. Ion source gas and curtain gas were set up at 50 and 25, respectively. The instrument was optimized by direct infusion of single LC standards of Ile, Leu, Phe, Pro and Val until protonation was observed for all. Ions were detected and monitored by Multiple Reaction Monitoring (MRM) approach. Parameters for each amino acid and their internal standards are listed in Table S2. Data acquisition and processing was performed using Analyst software version 1.7.1 (SCIEX). Amino acid concentrations were calculated comparing original amino acid (¹²C-AA) peak area to isotope labelled internal standards (¹³C-AA) ratio from sample and calibrant blends as per Equation 3:

$$w_x = w_z \cdot \frac{m_z}{m_{yc}} \cdot \frac{m_y}{m_x} \cdot \frac{R'_B}{R_{B_c}} \quad (\text{Equation 3})$$

w_x and w_z are AA mass fraction in the sample solution (µg/g) and mass fraction of the standard AA solution added to the calibrant blend (µg/g). The terms m_z , m_{yc} , m_y and m_x indicate the mass of standard AA solution added to the calibrant blend (g), the mass of labelled AA solution added to the calibration blend (g), the mass of labelled AA solution to the sample blend (g) and the mass of sample solution added to the sample blend (g). R'_B is the analyte/IS ratio measured in the sample blend and R_{B_c} is the analyte/IS ratio measured in the calibration blend.

IC: standards and sample solutions

Two independent 1000 µg/g C₃+ peptide solutions were prepared and diluted ten times in mobile phase and analyzed by ion chromatography on two different days. Standard solutions were prepared from concentrated sodium trifluoroacetate (Na⁺TFA⁻) (>99%) supplied by Sigma Aldrich Chemicals. TFA and anions content were determined using ion chromatography system composed of 930 compact IC flex and 863 compact autosampler (Metrohm AG, Switzerland) coupled with suppressed conductivity detection performed with 0.05 M sulfuric acid regenerated suppressor module. Metrosep A Supp 5, 250 × 4.0 mm analytical anions exchange (Metrohm AG, Switzerland) with sodium carbonate/bicarbonate (3.2 mM/1mM) eluent flowing at 0.7 mL/min was used at 30°C.

KF titration

KF titration was performed using a Titando 851 titrator instrument (Metrohm, Herisau, Switzerland), coupled to a Metrohm 874 oven sample processor and operated by the Tiamo (version 2.3) software. The C₃+ peptide samples (between 2 and 6 mg) were placed in the oven and the released water was transferred to the titrator cell via a stream of dried, high purity nitrogen carrier gas. The coulometric determinations were carried out with the KF titration reagents HYDRANAL™ Coulomat CG-K and Coulomat AG-Oven (VWR, Fontenay-sous-Bois, France). Water content was calculated after correction of water determined in blank samples. A lactose standard 5% (Merck, France) was used as a control sample and the SI traceability of the results was established through calibration with NIST SRM 2890 standard (NIST, MD, USA).

Determination of water mass fraction content (w_w)

A temperature gradient experiment was run to define the optimal oven temperature for the sample. The release of water was observed up to 180°C and the sample completely degraded at 200°C. The Karl Fisher titrator was then prepared by purging with dried carrier gas to remove any water traces. After the determination of the blank value, the samples were analyzed using the optimized method (Figure S2). A minimum sample intake of 2.3 mg was used for each analysis. The average mass fraction of water was 62.3 ± 7.4 mg/g ($k = 2$).

Elemental analysis

The elemental analysis was performed using a Vario Micro elemental analyser (Elementar Analysensysteme GmbH, Langenselbold, Germany). The analyser was calibrated using standard samples of acetanilid (Merck, France) ranging from 1 to 10 mg. Subsequently, four C_3+ samples were mixed with magnesium oxide (Merck, France) in small tin crucibles, sealed and placed into the furnace for high-temperature oxygen combustion, followed by gas chromatography and thermal conductivity detection.

Quantitative 1H and ^{19}F NMR

The peptide material and the internal standard (IS) 2,3,5,6-tetrafluoro-1,4-bis(trimethylsilyl)-benzene (BTMSB- F_4 , NMIJ Certified Reference Material 4602a) were accurately weighed in an ultra-micro balance XP2U (Mettler Toledo), co-dissolved in deuterated methanol, transferred to an NMR tube and analysed first by 1H qNMR and then by ^{19}F qNMR. Four replicate spectra were acquired for each of the two prepared sample mixtures. The determined 90° pulse widths were 6.2 and 14.4 μs for 1H and ^{19}F , respectively. The longest T1 relaxation times were 2.17 and 2.8 s corresponding to the internal standards in 1H and ^{19}F NMR, respectively. The repetition time was set to $15 \times T1$ and the number of scans was 128. The transmitter offsets were set at the midpoint between the quantification and internal standard signals and ^{13}C decoupling and spinning were not used. Post-acquisition data processing was performed using MestReNova software version 14 (Mestrelab Research).

The purity value was calculated using Equation 4 and its uncertainty corresponded to the square root of the sum of the squared uncertainties associated to each of the intervening factors in the equation (Westwood et al., 2019).

$$w_x = \frac{I_x}{I_{Std}} \frac{N_{Std}}{N_x} \frac{M_x}{M_{Std}} \frac{m_{Std}}{m_x} w_{Std} \quad (\text{Equation 4})$$

where I_{Std} , N_{Std} , M_{Std} , m_{Std} and w_{Std} are the signal area, number of protons or fluorines, molecular weight, weighed mass and mass fraction of the IS, respectively. The terms I_x , N_x , M_x , m_x and w_x indicate the signal area, number of protons or fluorines, molecular weight, weighed mass and mass fraction of the analyte, respectively.