



**University of
Nottingham**

UK | CHINA | MALAYSIA

**Enabling *in situ* protein analysis for
pharmaceutical applications
using 3D OrbiSIMS**

Anna Maria Kotowska

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

September 2021

Abstract

In situ characterisation of proteins at surfaces is of interest to biological, biotechnological and pharmaceutical fields, specifically areas of research such as tissue imaging, biomaterials discovery, biosensor development and tissue engineering. Label-free protein characterization at surfaces is commonly achieved by matrix assisted laser desorption ionisation (MALDI) and requires digestion and/or matrix deposition prior to mass spectrometry, limiting the spatial resolution in two and three dimensions. This work presents the use of secondary ion mass spectrometry, 3D OrbiSIMS, to achieve the first matrix-, digestion- and label-free assignment of proteins in solid three-dimensional samples.

Primary ion beam-induced fragmentation of undigested proteins resulted in spectra containing characteristic fragment ions, which can be used for *de novo* sequencing protein identification. This approach, termed ballistic sequencing, enabled amino acid sequence assignment covering from 5 to 53% of the sequence in 16 analysed proteins. The similarities and differences from classical proteomic methods were established and features such as proton, ligand or cell-binding sites were detected in the analysed proteins.

The method proved to be sensitive enough to detect highly specific protein fragments in the 3D OrbiSIMS spectra from a protein monolayer sample, equating to 40 femtomoles of a protein per analysis. This demonstrates the unique capability of ballistic sequencing to characterise protein surfaces, when the amount of analyte is too small to be extracted from the surface for analysis via existing methods such as MALDI. To further test the method, three example proteins were mapped throughout the depth of a complex biological sample, human skin. Collagen, keratin and

corneodesmosin were detected in the dermis, epidermis and *stratum corneum*, respectively, which is consistent with known distribution of these proteins in skin.

A reproducible method of manufacturing monodisperse PLGA micro- and nanoparticles was successfully developed for subsequent incubation with biological media and investigation of protein adsorption on the surface. A reliable size tuning in a range between 120-250 nm and 5-20 μm was achieved by the use of droplet microfluidics and modifying parameters such as the polymer concentration and the flow rate of continuous and dispersed phases.

The polymeric particles were incubated with serum albumin and whole serum and analysed *in situ* using cryogenic capabilities of the 3D OrbiSIMS. The chemical image of particle surface was not achieved, however the preliminary results obtained in this work, such as detection of characteristic serum albumin peaks in the frozen particles sample, indicate that further method development could enable use of the 3D OrbiSIMS as an emerging technique for *in situ* analysis of particles' interaction with biological media.

Acknowledgements

Firstly, I would like to acknowledge the EPSRC Centre for Doctoral Training in Advanced Therapeutics & Nanomedicines (EP/L01646X/1) for funding my PhD. I am thankful to the CDT management for coordinating the programme and everyone in the CDT and in the School of Pharmacy for sharing knowledge, technical or moral support.

I am particularly grateful to my supervisors: David Scurr, Jon Aylott and Alex Shard, for their endless patience, mentorship and guidance. Specifically, David for every “you got 10 mins?” turning into hours of talk about science and beyond. Jon for always seeing things in perspective and tirelessly teaching me to do the same. Alex for always keeping an eye on the details and asking difficult questions. I would also like to thank all collaborators, especially Morgan Alexander, Phil Williams, Paula Mendes and Gustavo Trindade for invaluable support and frustratingly eye-opening conversations.

In no particular order, thank you to Paulina, Ola, Ola, Li Min, Sofya and Jean-Luc for more than they imagine. The biggest thanks to my family for the continuous encouragement, my Grandma for always believing in everything I do and my parents for always believing that whatever I do, I can do better. Thank you.

Nunc est bibendum.

Table of contents

1	General introduction.....	16
1.1	Classical methods of analys of proteins at surfaces	16
1.2	Secondary ion mass spectrometry	19
1.3	Biotechnological and biological applications of <i>in situ</i> protein analysis	25
1.4	Pharmaceutical applications of <i>in situ</i> protein analysis.....	28
1.4.1	Nanotechnology for controlled and targeted drug delivery	28
1.4.2	Methods and materials in particle manufacturing.....	29
1.4.3	Nanoparticle interaction with biological environment.....	31
1.4.4	Nanoparticle characterisation methods	36
1.5	The scope of this thesis	38
2	Materials and methods	39
2.1	Materials	39
2.2	Protein sample preparation	40
2.2.1	Protein films	40
2.2.2	Protein monolayer	41
2.3	Micro- and nanoparticle preparation	41
2.3.1	Microfluidics	41
2.3.2	Polymer particle functionalisation	42
2.4	Analytical methods.....	42
2.4.1	Dynamic light scattering (DLS).....	42

2.4.2	Laser diffraction (LD)	43
2.4.3	Scanning electron microscopy (SEM)	43
2.4.4	X-ray photoelectron spectroscopy (XPS).....	43
2.4.5	Time of flight secondary ion mass spectrometry (ToF-SIMS)	43
2.4.6	3D OrbiSIMS	44
2.4.7	Optical profilometry.....	46
2.4.8	Ellipsometry	46
2.4.9	Statistical analysis	47
3	Development of ballistic sequencing for assignment and identification of proteins using 3D OrbiSIMS.....	48
3.1	Chapter aims	48
3.2	Methods	49
3.2.1	Protein films	49
3.2.2	ToF-SIMS measurements	49
3.2.3	3D OrbiSIMS measurements	50
3.3	Results and discussion.....	51
3.3.1	Capabilities and limitations of ToF-SIMS for assignment of protein fragments.....	51
3.3.2	Protein identification using GCIB/Orbitrap using model samples	56
3.3.3	Interpretation of ballistic fragmentation spectra by comparison with classical mass spectrometry proteomics methods	59
3.3.4	Lateral resolution and sensitivity of the ballistic sequencing	69

3.4	Conclusions	71
4	Application of ballistic sequencing for <i>in situ</i> analysis of biological samples ..	72
4.1	Chapter aims	72
4.2	Methods	73
4.2.1	3D OrbiSIMS measurement of human skin	73
4.2.2	Preparation of the protein biochip.....	74
4.2.3	Measurement of layer thickness.....	74
4.3	Results and discussion.....	75
4.3.1	Detection of distinct protein fragments in a simple mixture.....	75
4.3.2	Characterisation of human serum by 3D OrbiSIMS	78
4.3.3	Detection of distinct protein fragments in human skin	84
4.3.4	Detection of protein fragments in a protein biochip	91
4.3.5	Detection of functional and active sites of proteins	93
4.4	Conclusions	96
5	Manufacturing of polymeric micro- and nanoparticles for the investigation of protein corona.....	98
5.1	Chapter aims	98
5.2	Methods	99
5.2.1	Micro- and nanoparticle production.....	99
5.2.2	Particle size measurements	100
5.3	Results and discussion.....	101
5.3.1	Particle preparation using microfluidics	101

5.3.2	PLGA microparticle size tuning	102
5.3.3	PLA microparticle size tuning	106
5.3.4	Nanoparticle size tuning.....	109
5.3.5	Impact of the surfactant on the nanoparticle size.....	111
5.4	Conclusions	113
6	Protein-functionalised polymer particles	114
6.1	Chapter aims.....	114
6.2	Methods	115
6.2.1	Particle incubation.....	115
6.2.2	ToF-SIMS measurement of freeze dried particles	115
6.2.3	3D OrbiSIMS measurement of frozen-hydrated particles	116
6.3	Results and discussion.....	117
6.3.1	Detection and assignment of diagnostic human serum albumin fragments using cryo 3D OrbiSIMS	119
6.3.2	Particle chemistry analysed <i>in situ</i> using cryo OrbiSIMS	122
6.4	Conclusions	128
7	Summary and future work.....	130
8	References	134

Figures

Figure 1.1. Schematic workflow of MALDI analysis.....	18
Figure 1.2. Protein samples can be distinguished by first and second principal component.....	20
Figure 1.3. Positive polarity ToF-SIMS spectrum of a 1.2 kDa peptide obtained with 20 kV Ar ₂₀₀₀ ⁺ primary beam.	22
Figure 1.4. Schematic of the 3D OrbiSIMS	23
Figure 1.5. 3D OrbiSIMS image overlay of algae membrane.	24
Figure 1.6. Temporal representation of the foreign body response	27
Figure 1.7. Schematic of the microfluidic setup.	30
Figure 1.8. Physico-chemical characteristics of nanoparticles mediating protein interaction.....	33
Figure 3.1. ToF-SIMS spectrum of lysozyme film masked by a gold grid	52
Figure 3.2. Comparison of ToF-SIMS and 3D OrbiSIMS spectra of lysozyme.....	53
Figure 3.3. Scores of PC1 and PC2. The amino acid fragment intensity patterns are consistent within the proteins and vary between different proteins.	54
Figure 3.4. Principal component analysis of 7 different proteins immobilised on TCD-terminated chips.....	55
Figure 3.5. Inverted overlay comparison of positive polarity ToF-SIMS (red) and 3D OrbiSIMS (blue) spectra of purified chicken egg lysozyme film.....	57
Figure 3.6. Positive polarity 3D OrbiSIMS spectrum of thin insulin film.....	58
Figure 3.7. Inverted overlay comparison of full 3D OrbiSIMS spectrum of insulin and 3D OrbiSIMS MS/MS of the m/z 620.29	59
Figure 3.8. 16 proteins of different biological functions were analysed.....	60

Figure 3.9. Proposed structures of observed ions dissimilar to CID MS of tryptic peptides, similar to HCD or MALDI-ISD fragmentation.	62
Figure 3.10. Structure and amino acid sequence of bovine α -chymotrypsin.	63
Figure 3.11. Peak intensity depending on ion type and sequence composition, N-terminus sequence in a) positive polarity and b) negative polarity.	66
Figure 3.12. Peak intensity depending on ion type and sequence composition, negative polarity.	66
Figure 3.13. Comparison of structures and amino acid sequences of bovine serum albumin and human serum albumin.	68
Figure 3.14. 3D OrbiSIMS image of a protein film under a TEM grid, obtained with a focussed 5 μm diameter Argon ₃₀₀₀ ⁺ primary beam.	69
Figure 4.1. Spectrum of a thin film of 1:1 lysozyme: insulin mixture.	77
Figure 4.2. GCIB Orbitrap images of the model mixture.	77
Figure 4.3. 3D OrbiSIMS spectrum of human serum.	79
Figure 4.4. 3D OrbiSIMS depth profile overlay of example ions of three proteins; keratin (stratum corneum and epidermis, blue), corneodesmosin (stratum corneum, red) and collagen (dermis, black) in a human skin sample.	86
Figure 4.5. 3D OrbiSIMS depth profile overlay of example ions representing collagen.	87
Figure 4.6. 3D OrbiSIMS depth profile overlay of example ions representing QHGSG sequence of corneodesmosin (CDSN).	87
Figure 4.7. 3D OrbiSIMS depth profile overlay of example ions representing keratin sequence SFGGGG.	88
Figure 4.8. Layer thickness measurements of the self assembled monolayer (SAM) and SAM with immobilised lysozyme.	92

Figure 4.9. Overlay of 3D OrbiSIMS spectra of catalase and cytochrome c.....	94
Figure 4.10. Inverted overlay of 3D OrbiSIMS spectra of catalase and cytochrome c with corresponding structures of ferric and ferrous heme.	95
Figure 4.11. Fragmentation of heme using Ar ₃₀₀₀ ⁺ primary beam..	95
Figure 5.1. Droplet junction chip (190µm etch depth) for production of a nanoparticles by flow focusing and b microparticles by droplet formation.	102
Figure 5.2. Comparison of three particle sizing techniques, SEM (a), light microscopy (b) and laser diffraction(c).....	105
Figure 5.3. Particle size tuning by manipulating polymer concentration and continuous phase flow rate, assessed by SEM.	105
Figure 5.4. Particle size tuning by manipulating polymer concentration and continuous phase flow rate, assessed by SEM.	107
Figure 5.5. The impact of polymer choice on the particle size.	108
Figure 5.6. Comparison of the SEM images of the surface of a PLGA particles and b PLA particles produced with the same flow parameters.....	108
Figure 5.7. Particle diameter measured by DLS, impact of PLGA concentration and a ratio of flow rates	110
Figure 5.8. Impact of surfactant and the total flow rate on nanoparticle size.	112
Figure 5.9. Size distribution of PLGA particles produced with 1% PLGA solution and high total flow rates, measured by dynamic light scattering.	112
Figure 6.1. Comparison of ToF-SIMS images of freeze-dried a-d) bare particles, b-h)particles incubated with PBS and c-l) particles incubated with bovine serum albumin. Total ion images of (a) bare particles, (e) particles incubated with PBS and (i) particles incubated with BSA. Example PLGA ion C ₃ H ₅ O ₂ ⁻ in (b) bare particles, (f) particles incubated with PBS and (j) particles incubated with BSA.....	118

Figure 6.2. Inverted overlay comparison of human serum albumin measured in cryogenic conditions and dried protein film in room temperature.....	120
Figure 6.3. Negative polarity spectrum of human serum albumin acquired in cryogenic conditions.	120
Figure 6.4. Comparison of cryo ToF-SIMS spectra (a) and images of particles b) incubated with bovine serum albumin and c) not incubated.....	123
Figure 6.5. Cryo ToF-SIMS images of PLGA particles high-pressure frozen and freeze-fractured after incubation with bovine serum albumin.	124
Figure 6.6. Experimental setup of the cryo sample holder and the comparison of the size of the freeze-fracture planchette and the 6 mm planchette.	125
Figure 6.7. Inverted overlay comparison of cryo OrbiSIMS spectra of pure PLGA particles, obtained from samples high-pressure frozen in a freeze-fracture cell and samples in the 6 mm planchettes, plunge-frozen by dipping in liquid nitrogen.	125
Figure 6.8. Overlay comparison of human serum albumin b ₈ ion, present in the positive polarity cryo OrbiSIMS spectrum of the particles after incubation with human serum albumin and full serum	126
Figure 6.9. Cryo ToF-SIMS total ion images of pure PLGA particles, particles incubated with human serum albumin and particles incubated with human serum.	127

Tables

Table 3.1. List of analysed proteins from the smallest (insulin, 51 amino acid sequence) to the largest (fibronectin, 2446 amino acid sequence).....	60
Table 3.2. Peak list, description of sequence, measured mass, assignment and deviation from calculated mass for ions from positive polarity spectrum of α -chymotrypsin from bovine liver, ions assigned as C-terminal sequence of the A chain GVP AIQPVLSGL and a full A chain CGVP AIQPVLSGL.	63
Table 3.3. Peak list exported from SurfaceLab negative mode spectrum of bovine α -chymotrypsin, consisting of ions detected in the spectrum and assigned fragments of C-terminal sequence TALVNWVQQTLAAN.	65
Table 4.1. Assignments of C-terminus and internal sequences of human serum albumin, assigned by molecular formula prediction and SIMSdenovo.	80
Table 4.2. Highest-scoring proteins detected automatically in human serum sample using SIMS-MFP and SIMSdenovo.	82
Table 4.3. Highest intensity lipid peaks assigned automatically in human serum 3D OrbiSIMS spectrum, using molecular formula prediction and LipidMaps database of bulk lipid assignments.....	83
Table 4.4. Example assignments of the most abundant salt peaks in the 3D OrbiSIMS spectrum of human serum.	84
Table 4.5. Peak list exported from SurfaceLab negative mode depth profile through human skin, consisting of ions detected in the spectrum and assigned as sequence PGE of collagen.	89
Table 4.6. Peak list exported from SurfaceLab negative mode depth profile through human skin, consisting of ions detected in the spectrum and assigned as sequence QHGSG of corneodesmosin.....	89

Table 4.7. Peak list exported from SurfaceLab negative mode depth profile through human skin, consisting of ions detected in the spectrum and assigned as sequence SFGGGG of keratin.	90
Table 4.8. Lysozyme fragments visible in the spectra obtained from a protein monolayer sample.	92
Table 4.9. Examples of functional sites detected in the 16 analysed proteins. The location of the functional sites is based on UniProt database.	93
Table 5.1. Particle size tuning by changing parameters such as PLGA concentration and continuous phase flow rate.	104
Table 5.2. Particle size tuning by changing parameters such as PLA concentration and continuous phase flow rate.	106
Table 5.3. Parameters used in PLGA nanoparticle size tuning. Concentrations of 0.25%, 0.5%, 0.75% and 1% were used and continuous phase flow rates were modified to achieve flow ratios of 4-16.	109
Table 6.1. Assignments of N-terminus ions detected in the positive polarity cryo OrbiSIMS spectrum of human serum albumin.	121
Table 6.2. Assignments of C-terminus ions detected in negative polarity cryo OrbiSIMS spectrum of human serum albumin.	121
Table 6.3. Assignments of lipids detected in the the positive polarity cryo OrbiSIMS spectrum of PLGA particles incubated with serum, acquired from samples in the 6 mm planchettes.	128

Publications

1. "Direct nanobody immobilization on gold sensors" Simões B., Guedens W., Keene C., Kubiak K., Mulheran P., **Kotowska A.M.**, Scurr D.J., Alexander M.R., Broisat A., Johnson S., Muyldermans S., Devoogdt N., Adriaensens P., Mendes P.M., ACS Applied Materials and Interfaces, 2021
2. "*In situ* protein identification and mapping using secondary ion mass spectrometry" **Anna M. Kotowska**, Gustavo F. Trindade, Philip M. Williams, Jonathan W. Aylott, Alexander G. Shard, Morgan R. Alexander and David J. Scurr, Nature Communications, 2020
3. "Reversible, High-Affinity Surface Capturing of Proteins Directed by Supramolecular Assembly" Di Palma G., **Kotowska A.M.**, Hart L.R., Scurr D.J., Rawson F.J., Tommasone S., Mendes P.M., ACS Applied Materials and Interfaces, 2019

Conference presentations

1. "*In situ* imaging and depth profiling of proteins using 3D OrbiSIMS" 41st Annual Meeting of the British Mass Spectrometry Society, Sheffield, 9th September 2021
2. "Ballistic sequencing – analysis of intact proteins in the 3D OrbiSIMS" 85th UK Surface Analysis Forum, Manchester Zoom, 8-9th July 2020
3. "Acquisition of primary structure information from proteins using 3D OrbiSIMS" 22nd International Conference on Secondary Ion Mass Spectrometry, Kyoto, 20-25 October 2019

1 General introduction

1.1 Classical methods of analysis of proteins at surfaces

Characterisation of proteins on a surface (as opposed to analysis in liquid or gas phase) is of interest to areas of research such as tissue imaging, biomaterials discovery, biosensor development and tissue engineering. Particularly, *in situ* analysis is invaluable in providing information about protein distribution and conformation on the surface, which affects cellular adhesion, spreading and differentiation¹. Such spatial characterisation *in situ* is crucial to investigating the structure of the interface between a material and native tissue, which makes it key to developing medical devices. Characterising proteins in tissues is used to gain improved understanding of molecular processes underlying diseases such as Alzheimer's disease², cancers³⁻⁵ or lung diseases⁶. *In situ* analysis of tissues enables simultaneous detection and mapping of different proteins and other endogenous or exogenous compounds, creating a more comprehensive image of the disease. Assessment of the spatial distribution and orientation of proteins at surfaces is crucial for characterising biosensors used in medical and biotechnological fields, to ensure availability of protein active sites, region selectivity on the surface and the functionality of the device⁷. These biological and pharmaceutical applications of *in situ* protein analysis will be discussed in detail in sections 1.3 and 1.4.

The methods of analysing proteins at surfaces can be separated into two main groups, the labelled and label-free methods. The labelled methods include immune staining or the use of recombinant proteins with attached fluorescent tags (affinity tagging)³. Immunofluorescence enables locating proteins within the cell organelles in

a targeted way, but proteome-wide studies are not possible with this method⁸. This review will focus on label-free methods of imaging proteins. These methods allow analysis of a wide range of molecules simultaneously, enabling discovery based research and are based on mass spectrometry, with the most widely employed being matrix assisted laser desorption/ionisation (MALDI)⁹.

The schematic of MALDI imaging workflow is presented in Figure 1.1. Firstly, the sample must be coated with a matrix, specific to type of analyte (Figure 1.1a)¹⁰.

Next, mass spectra are acquired by moving the stage under the laser beam, as shown in Figure 1.1b. Each analysis forms a pixel in a map of the sample area (Figure 1.1c).

MALDI data is often supplemented by LC-MS/MS analysis of sample extracted from the surface, which can improve identification of species (Figure 1.1d).

Due to soft ionisation producing singly charged ions and no intense fragmentation, the analysis of large proteins with MALDI requires enzymatic digestion which may modify the structure of the sample surface. The process also requires matrix deposition, limiting the spatial distribution to 50 μm ^{3,11}. Recent non-commercial instrument developments pushed the limits of this technique allowing MALDI imaging of lipids and small peptides with lateral resolution of 1.4 μm by the use of laser focusing objective and specialised matrix deposition procedures¹².

Other methods used for protein characterisation at surfaces are desorption electrospray ionisation (DESI)¹³ and liquid extraction surface analysis mass spectrometry (LESA-MS)¹⁴, which differ from MALDI by operating in ambient conditions¹⁴. DESI and LESA require extraction of the analyte from the surface, therefore these techniques have limited lateral resolution of several hundred micrometres to millimetre scale¹⁵.

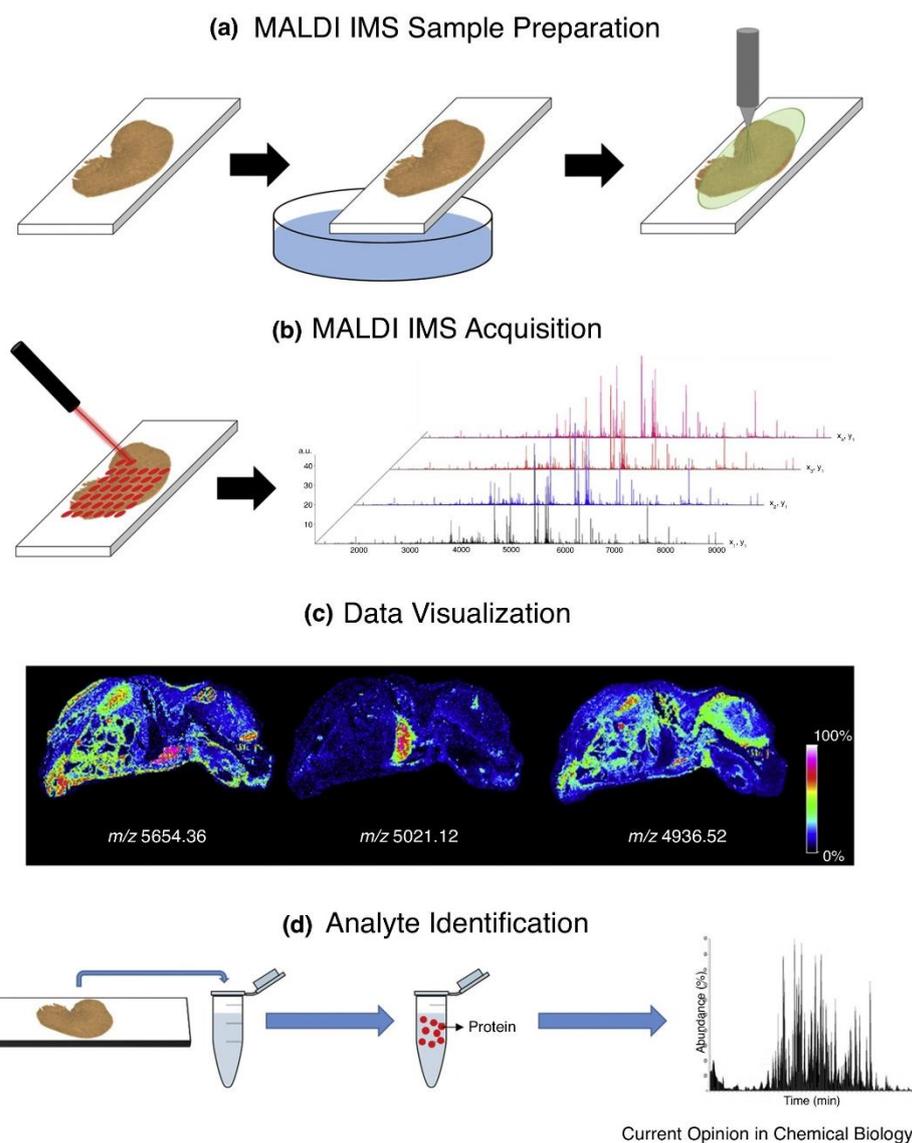


Figure 1.1. Schematic workflow of MALDI analysis. Copied from Ryan et.al *Current Opinion in Chemical Biology*, 2018.

Recently this limitation was challenged by the development of Nano-DESI, enabling lateral resolution of 200 μm for proteins up to 40 kDa¹⁶. Both techniques can enable direct analysis of intact proteins and complexes with a capability to detect molecules up to 35 kDa in DESI¹⁷ and 70 kDa in LESA-MS¹⁸. While allowing *in situ* analysis of proteins, MALDI, DESI and LESA-MS are limited to two-dimension analysis. Some 3D information can be achieved from reconstruction of multiple section images, however the depth resolution of such analysis is restricted by the capability to create thin tissue sections¹⁹.

1.2 Secondary ion mass spectrometry

Classical methods of label-free analysis of proteins at surfaces described in Section 1.1 do not enable high spatial resolution mapping of proteins in three dimensional samples. Time of flight secondary ion mass spectrometry (ToF-SIMS) is a surface analysis technique in which the analyte is ejected from the surface and ionised by the primary ion beam²⁰. Compared to MALDI or DESI, SIMS allows higher lateral resolution (<200 nm) and 3D analysis of the sample²¹. It can achieve subcellular imaging^{22,23} and analysis of samples with topography²⁴. However, in SIMS, energetic primary ion beams, most commonly Bi₃⁺, cause intense fragmentation of the molecules ejected from the surface, which is particularly problematic in the analysis of large biomolecules.

These characteristics make the technique more suitable for analysis of small molecules rather than proteins. Large biomolecules are heavily fragmented by the primary ion beams resulting in small ($m/z < 200$), unspecific ions being detected. Conclusions regarding protein distribution can be drawn from imaging and profiling of amino acid ions²⁵⁻²⁷ in positive polarity and CN⁻ and CNO⁻ peaks in negative polarity²⁸.

Despite typical SIMS spectra of proteins containing only single amino acid ions, relative intensities of the peaks create patterns, encoding information about protein identity^{29,30}, conformation³¹ or orientation³²⁻³⁵. These patterns can be identified and compared using multivariate analysis^{34,36-38}. From various multivariate statistics methods, principal component analysis (PCA) is the most suitable for distinguishing different protein samples³⁶. In this type of analysis no prior information about the

samples is given into the model. The PCA algorithm detects the biggest differences between the samples. The first principal component (PC1) explains the most variation in the dataset. Subsequent principal components (PC2, PC3 etc.) explain a smaller fraction of variation within the dataset, until no variation is detected. Figure 1.2 demonstrates the possibility of distinguishing protein samples using PCA.

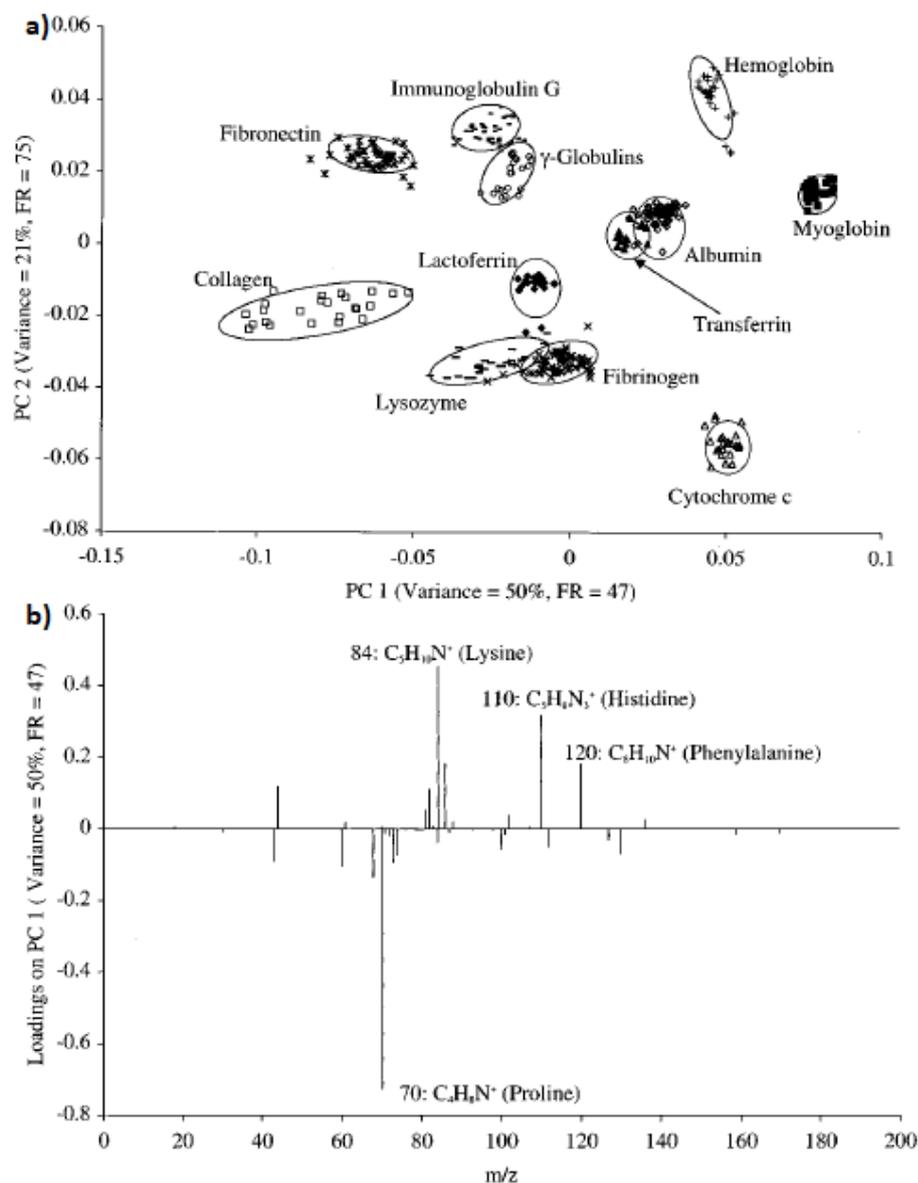


Figure 1.2. Protein samples can be distinguished by first and second principal component, which together explain 71% of differences between the samples. a) Scores plot of PCA on single protein samples. b) Loadings plot shows amino acid intensity patterns. Copied from Wagner M. et al. *Anal.Chem.* 2002.

Recently machine learning was employed to use the patterns in amino acid ion intensities to identify peptides³⁹. These approaches can provide insight into simple protein systems, however they cannot be used to differentiate proteins in more complex samples, for example protein mixtures, as the ion intensities will be affected by the matrix effect and the presence of salts^{40,41}.

In a few reports, multi amino acid fragments and molecular ions of peptides up to 1.6 kDa have also been detected by using Bi_3^+ and Au_3^+ small polyatomic metal primary ion beams^{42,43}. However, a breakthrough in the SIMS analysis of biological samples came with the development of polyatomic primary ion beams such as C_{60} and gas cluster ion beams (GCIB) such as Ar_n ⁴⁴. Large cluster primary ion source analysis beams carry less energy per atom resulting in less intense fragmentation of the analyte and enhancing the ion yield of the higher mass ions⁴⁵. Additionally, the use of continuous or quasicontinuous primary beam enables 3D profiling of biological samples without compromising mass resolution^{46,47}. This technological improvement enabled lipid profiling at cellular level⁴⁸ and detection of multi amino acid fragments from peptides up to 3 kDa⁴⁹. Figure 1.3 demonstrates fragmentation of bovine angiotensin I in a study comparing various argon cluster sizes, which found the 20kV Ar_{2000}^+ primary beam to enable assignment of multiple distinct peptide peaks⁵⁰.

Recently, the detection of molecular ions of whole proteins up to 12 kDa using Ar_{2000}^+ primary beam has been reported⁵¹. Nonetheless, large proteins have yet to be identified without pre-analysis digestion to constituent peptides in SIMS⁵². Detected ions cannot always be unambiguously assigned, especially in the case of complex biological samples due to limited mass resolving power of the ToF analyser. The development of hybrid instruments such as the PHI *nanoTOF* II⁵³ and the 3D

OrbiSIMS⁵⁴, employing alternative types of detectors could allow acquisition of more detailed macromolecular information.

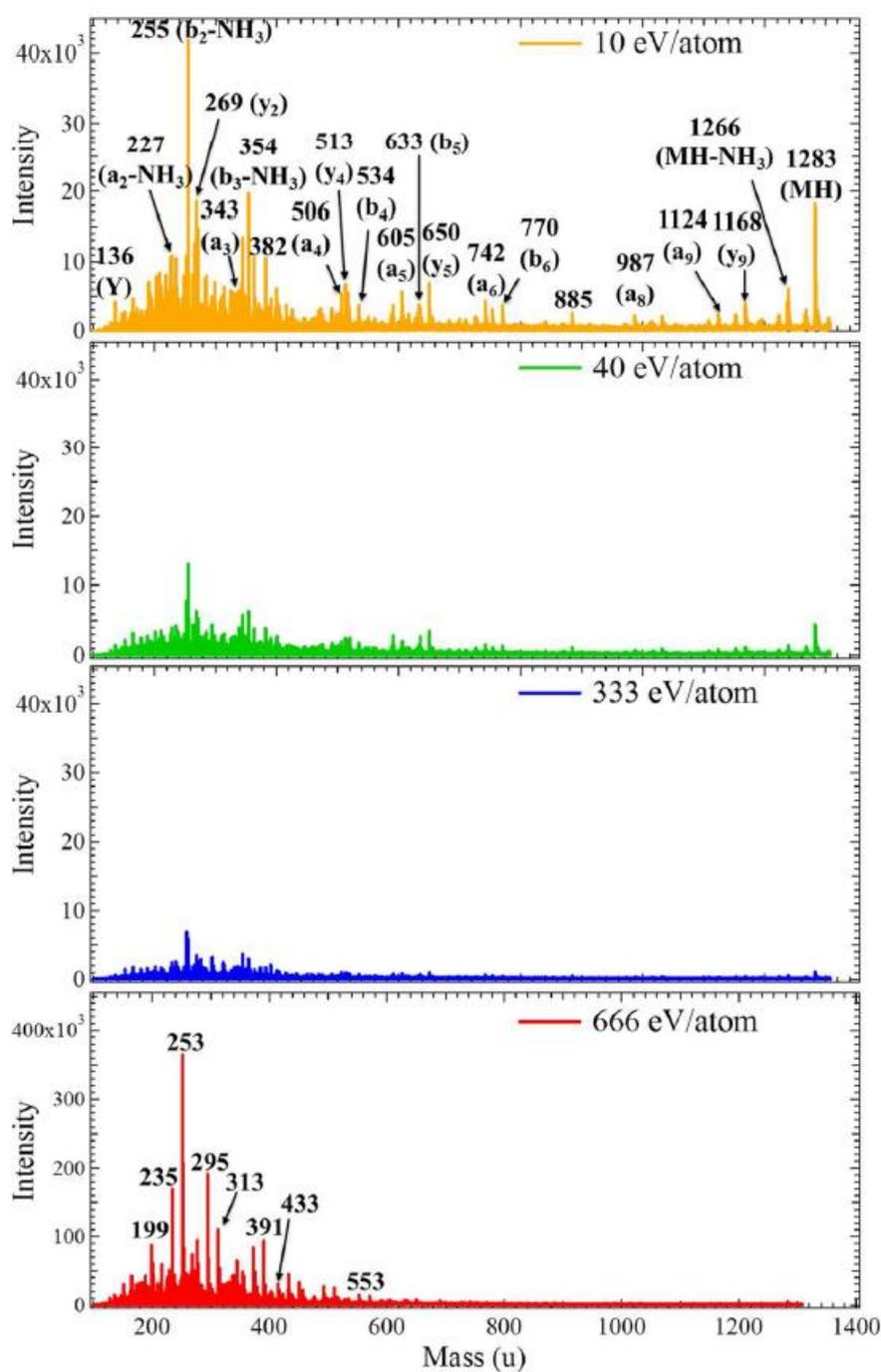


Figure 1.3. Positive polarity ToF-SIMS spectrum of a 1.2 kDa peptide obtained with 20 kV Ar_{2000}^+ primary beam. The use of a GCIB enables detection of multi amino acid peptide fragments and the molecular ion. Copied from Yokoyama Y. et al. *Anal. Chem.* 2016.

The 3D OrbiSIMS comprises of a ToF-SIMS 5 (IONTOF GmbH, Germany) and a Q Exactive HF Orbitrap™ (Thermo Fisher Scientific, Germany) analyser, used commonly in proteomic and metabolomic research⁵⁵. The instrument combines the high mass resolving power ($> 240,000$ at $200 m/z$ and high mass accuracy (<1 ppm) of an Orbitrap™ and the 3D imaging capability of time of flight secondary ion mass spectrometry (ToF-SIMS) and is presented schematically in Figure 1.4.

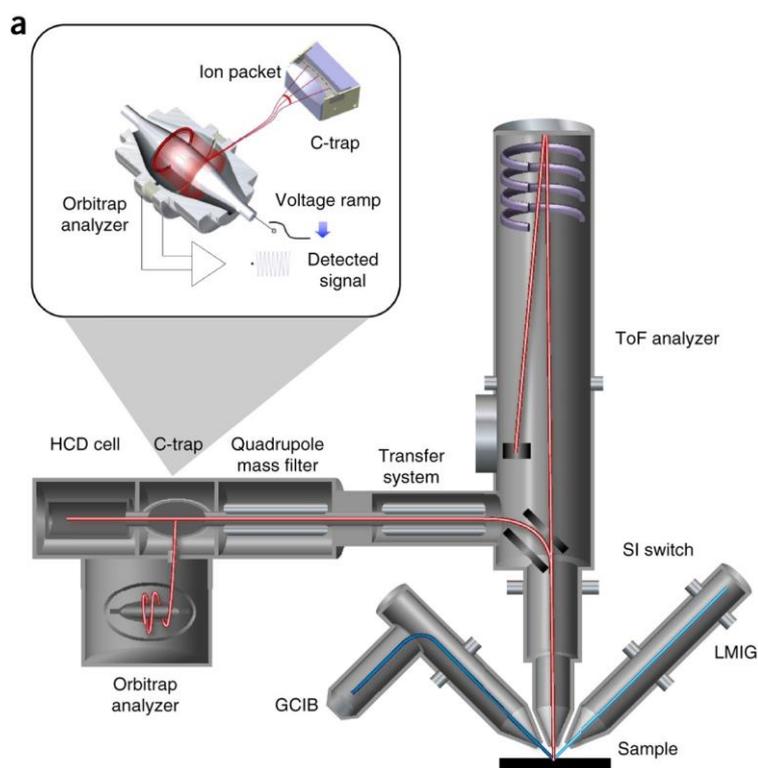


Figure 1.4. Schematic of the 3D OrbiSIMS. The instrument has a dual ion beam (LMIG and GCIB) and dual analyser (ToF and Orbitrap) capability. Copied from Passarelli M. et al. *Nature Methods* 2017.

The 3D OrbiSIMS can be used in 10 different modes, which include single beam (Bi_n^+ or Ar_n^+), single analyser (ToF or Orbitrap), as well as single beam double analyser and 3D imaging employing both beams and both analysers working together in an alternate pattern. For surface analysis, which does not require time consuming imaging, 3D OrbiSIMS can provide high-throughput metabolomic analysis of microarrays, providing complementary information to classical

metabolomic techniques such as gas chromatography mass spectrometry or liquid chromatography mass spectrometry⁵⁶.

The 3D OrbiSIMS enables relatively high lateral resolution high chemical specificity imaging of cell components too fragile to be isolated from the sample. In the example presented in Figure 1.5, different biomolecule groups, phosphocholine (red), chlorophyll (green) and polysaccharide (blue) were detected simultaneously during *in situ* analysis of algae membrane⁵⁷. Crucially, this analysis enabled for the first time the assessment of spatial distribution of oligosaccharide, which is expected to play a role in intercellular communication.

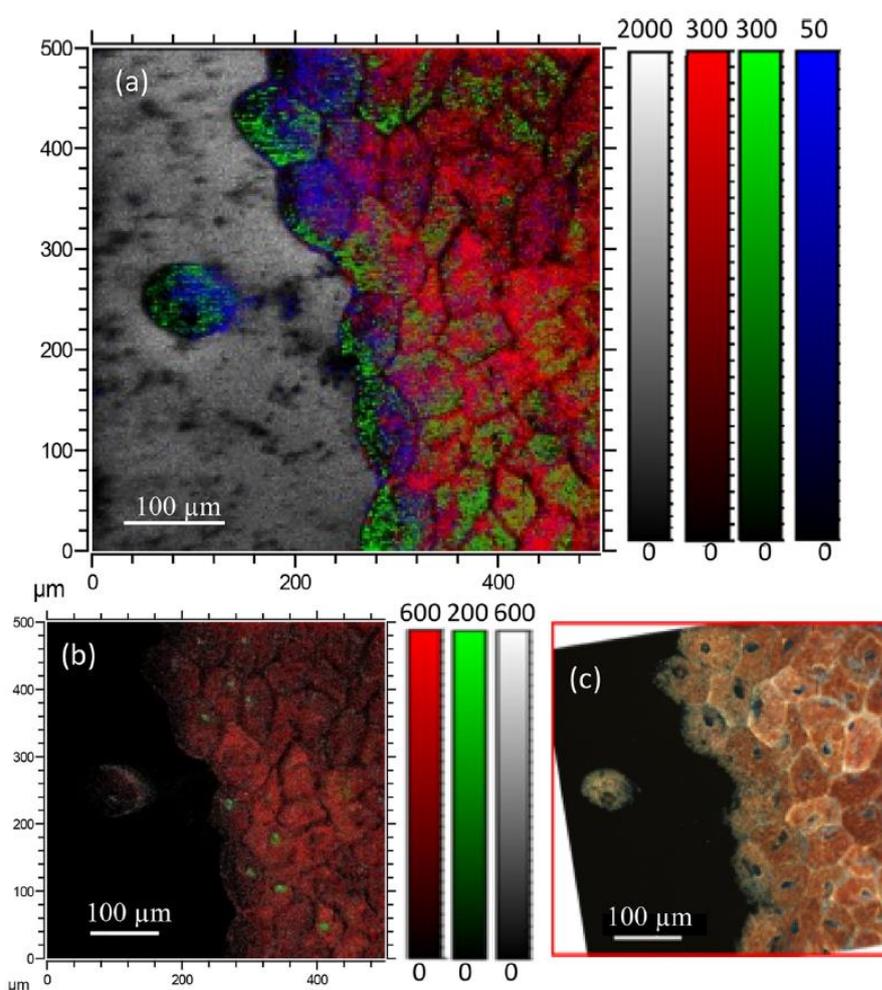


Figure 1.5. 3D OrbiSIMS image overlay of algae membrane, phosphocholine headgroup ($C_5H_{15}NO_4P^+$, red), chlorophyll A ($C_{30}H_{25}N_4OMg^+$, green) and polysaccharide ($C_{18}H_{30}O_{15}Na^+$, blue). Copied from Aoyagi S. *et.al Anal. Chem* 2019.

Importantly, the 3D OrbiSIMS has the capability to operate in cryogenic conditions, which not only enables to preserve the native state of the biological material, but also has been found to improve the detection of organic species, including peptides⁵⁸⁵⁹⁶⁰. In addition to improving ion yield of high mass ions through facilitating protonation, cryogenic OrbiSIMS also allows *in situ* detection of volatile compounds such as fatty acids, aldehydes and aromatic hydrocarbons⁶¹ and 3D mapping of the chemistry of highly hydrated samples such as bacterial biofilm⁶².

1.3 Biotechnological and biological applications of *in situ* protein analysis

A particularly challenging and growing field in biotechnology is the area of protein biochips. Biochips are platforms for high-throughput analysis of protein-protein, protein-drug interactions and protein phosphorylation. Due to being surface bound, they are useful for characterisation of highly acidic, highly alkaline or hydrophobic proteins, which are challenging in liquid-based methods⁶³. The paramount challenges in the biochip development are maintaining protein activity, conformation and adoption of a suitable orientation on the surface to ensure availability of the active site. In addition to complicated manufacturing techniques, quality control methods assessing the robustness and effectiveness are a bottleneck in the biochip development^{64,65}. In-situ protein analysis methods are necessary to rapidly evaluate the biochips during manufacturing process in order to rationally design efficient platforms⁶⁶.

Imaging distribution of proteins is valuable in manufacturing patterned protein surfaces used in tissue engineering research. It has been shown that spatial organisation of proteins guides cell attachment, spreading and differentiation¹. *In situ* characterisation of protein patterns is crucial for addressing the main challenges in the field of developmental biology which are closely mimicking the native environment⁶⁷ and linking localisation to cellular behaviour⁶⁸.

In addition to biomolecule patterning in 2D, proteins are used in engineering whole tissues. Incorporating proteins in networks and scaffolds for cartilage engineering helps mimic native tissue⁶⁹. Creating polymer-protein complexes improves the compatibility of cells with the material and improves transport of oxygen and signalling and nutrition compounds⁷⁰. Overall understanding protein role at the interface between engineered and biological material helps integrate the material with the native tissue.

In situ analysis of proteins at implantable materials helps predict the performance of the device in the body. Medical devices such as catheters, stents or bone replacements provoke an immune reaction after being implanted in the host. This response is known as foreign body reaction (FBR) and is initiated by adsorption of extracellular matrix proteins on the material surface, as presented in Figure 1.6⁷¹. This event activates immune cells, which produce collagens, forming a thick fibrous layer around the material. As a result, the material is separated from the host tissue, causing slow healing, impairing device integration and increasing the risk of infection⁷². Understanding the mechanism of protein adsorption and activation of the immunological cascade is necessary for manufacturing of successful medical devices by altering the chemistry of biomaterials⁷³.

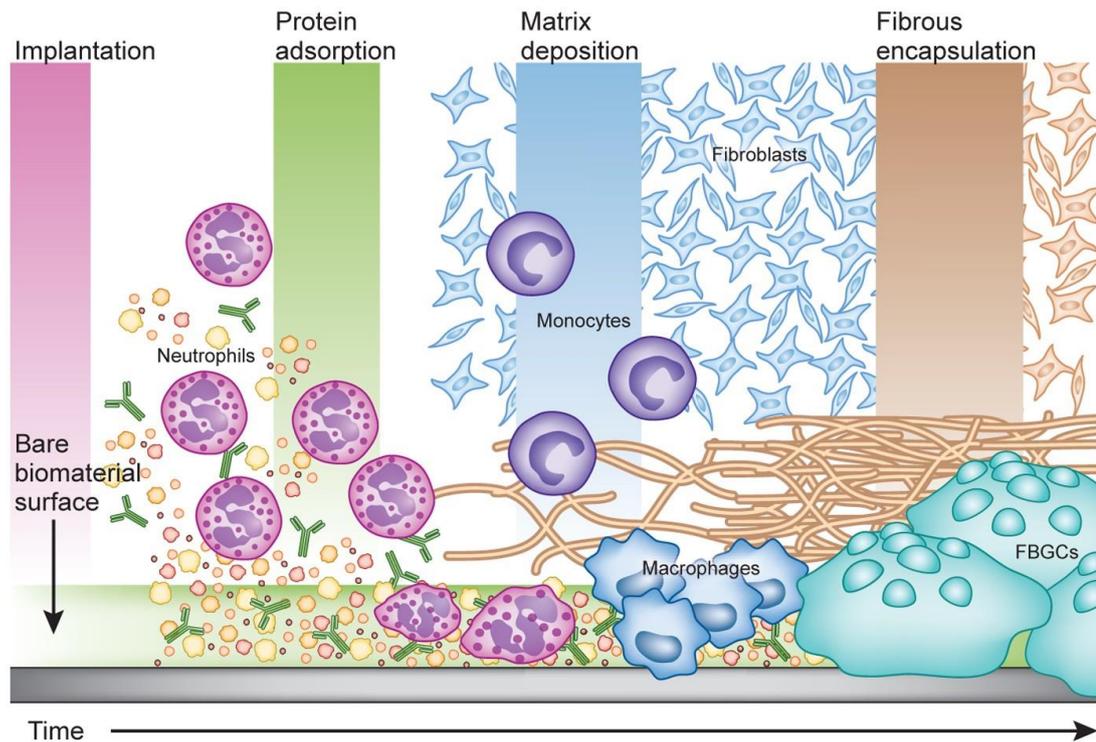


Figure 1.6. Temporal representation of the foreign body response. Protein adsorption is the first step of the immunological reaction to an implanted material. Copied from Granger D.W., *Nature Biotechnology*, 2013.

Design of biomaterials promoting the right cell adhesion and reducing bacterial adhesion directly links to the development of successful bone replacement implants or catheters and drug delivery implants such as insulin pumps. For the same reason knowledge of protein-material interaction contributes to the advancement of disease and medicine testing. Extracellular matrix proteins play a role in the pathophysiology of diseases affecting various organs, including heart, liver, kidneys or lungs⁷⁴. More than 50 diseases are caused by accumulation and misfolding of proteins into aggregates called amyloids. However, due to the heterogeneity of these fibrous deposits, the cause or treatment of related amyloidosis or neurodegenerative disorders remains unknown⁷⁵. Investigating proteins in the diseased tissues may provide insight into molecular background of the disease leading to the development of effective treatments.

1.4 Pharmaceutical applications of *in situ* protein analysis

1.4.1 Nanotechnology for controlled and targeted drug delivery

Nanomedicines aim to overcome challenges in drug development such as poor solubility, availability, pharmacokinetics, efficacy, target specificity or toxicity *in vivo*⁷⁶. Nanoparticle formulations can be approved to market if they are confirmed to reduce drug toxicity or improve on-target efficacy⁷⁷ and until 2016, 51 nanoparticle products (including the same formulation registered for different indication), had been approved by the FDA⁷⁸. Nanoparticles can increase efficacy by targeting specific cells in a passive and/or active way.

Passive targeting is based on size of the particle. For example in case of cancer therapy, nanoparticles pass through tumour blood vessels, which are leaky due to growing faster than healthy cells⁷⁹. This effect is called enhanced permeability and retention (EPR) and is usually observed for particles of diameter smaller than 100 nm⁸⁰. Active targeting of specific cells can, in theory, be achieved by immobilising ligands⁸¹, chemical⁸² or biological functionalities^{83,84} on nanoparticle surfaces. These ligands are devised to bind with receptors overexpressed on target cells.

Multiple materials are being studied for suitability as nanoparticle formulations, the most prevalent being polymer, lipid, nanocrystal, micelle or protein particles. This work will focus on polymeric particles as the most widely studied nanomaterial.

Many features of polymeric nanoparticles make them particularly advantageous for overcoming drug delivery challenges. These are their ability to achieve sustained drug release, encapsulation of poorly soluble drugs, multi-drug co-delivery and the ease of surface functionalisation^{85,86}.

1.4.2 Methods and materials in particle manufacturing

Established methods of polymeric particle production include different types of polymerisations such as dispersion, precipitation, emulsion or suspension polymerisation and emulsification methods. Solvent evaporation is one of the most commonly used emulsification methods of producing polymeric particles. In this method, a polymer solution is mixed with surfactant solution and either homogenised (microparticles) or sonicated (nanoparticles), followed by evaporation of the solvent^{87,88}. Droplets initially formed at the interface of the two solvents shrink as the solvent evaporates resulting in solid polymeric particles. A modification of these methods in which the interface between the solvents can be accurately controlled is called microfluidics^{89,90}.

Conventional methods and microfluidics have different capabilities and limitations. Advantages of microfluidics over microemulsion and nanoemulsion methods include narrow size distribution with polydispersity below 5%, capability to produce particles of non-conventional, non-spherical shapes, control over particles composition and internal structure. Particles produced using microfluidics present narrower particle size distribution, the production is continuous (rather than batch) and can be scaled-up in production process⁹¹.

In microfluidic systems two or more solvents are pushed through micro-sized channels by pumps enabling precise control over flow speed^{92,93}. The solvent containing the polymer (and the active pharmaceutical ingredient (API) in the case of drug-loaded particles) is called the dispersed phase. The solvent containing surfactant is called the continuous phase. Depending on solvent miscibility, droplets

or nanoemulsions are created on the intersection of two channels⁹⁴. There is a variety of shapes and sizes of microfluidic chips, each offering benefits for different applications. Microfluidic chips are typically produced of one of two materials, namely polydimethylsiloxane (PDMS) or glass. PDMS chips are low-cost and can easily be produced in various shapes using lithography methods. Glass chips on the other hand have better solvent compatibility and durability. The most popular chip geometries used in producing polymeric microspheres and nanospheres are T-junction and flow-focusing x-junction. An example of x-junction type chip installed in microfluidic setup is shown on Figure 1.7.

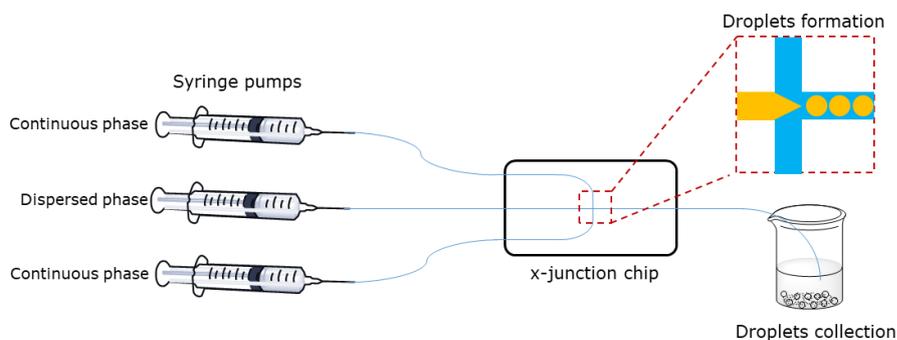


Figure 1.7. Schematic of the microfluidic setup. The magnified region shows droplets formation at the interface of two liquids.

Droplets are transformed into particles by solvent evaporation. Microfluidics enables the production of various materials such as polymeric, protein or inorganic particles and liposomes^{95,96}. The size of the droplets and resulting particles is tuned by changing flow rates of the two solvents and manipulating the composition of the liquids used, specifically surfactant type and concentration (1-3%) as well as polymer concentration (0.25%-2.5%)^{97,98}. The proportion between continuous phase flow rate and dispersed phase flow rate is known as the flow ratio and drives the character of the flow. As flow rate decreases, flow regime changes from dripping through jetting to laminar flow. Dripping results in largest droplets, jetting produces

smaller particles and in laminar flow two solvents are flowing next to each other and no particles are produced⁹⁹⁻¹⁰¹.

1.4.3 Nanoparticle interaction with biological environment

The use of polymeric nanoparticles could be particularly advantageous for formulation of biologics such as proteins, peptides, nucleotides growing in popularity as researched therapeutic agents. By acting at the source of the disease, they give promise of not only treating, but curing the condition and reducing the side effects. However, the development of such medicines is more labour- and cost-consuming than that of small molecule drugs.

Recent success of messenger ribonucleic acid (mRNA) nanoparticle-based vaccines demonstrates a potential of this technology to overcome the special stability and efficacy restrictions of fragile biological drugs¹⁰². These products were successful as they are designed to work in the circulation and interact with the immune system. In the case of medicines, whose target is outside of the circulation, the nanoparticles face a range of physiological barriers that limit their delivery to the site of action. As a result, only a small percentage of nanomedicines make it past proof-of-concept stage^{76,103-105}.

In 2016, Wilhelm *et al.* gathered results from 117 reports across last 10 years of investigating nanoparticles delivery to tumours, which revealed that a median of only 0.7% of the dose injected to mice reached the target¹⁰⁵. Even in cases of formulations being successful in animal studies, insufficient evidence of efficacy and safety often cause many nanoparticle products to fail in phase III clinical trials due to

inconclusive results and patient-to-patient variability¹⁰⁶. There have also been reports of formulations which have made it to the market but were subsequently withdrawn due to safety concerns (eg. Feruglose®, Resovist®).

In the laboratory, medicines are studied in cell cultures maintained in buffers, however in animal studies, nanoparticle formulations are being injected into the bloodstream, a more complex environment. The initial environment the particle interacts with is blood, consisting of blood cells suspended in plasma. Plasma is 90% water, with 1% ions (K, Na, Ca, Mg, Cl), dissolved gases, microelements and vitamins, and 9% organic compounds such as protein, glucose and lipids. There are 2000 different plasma proteins, of which albumin is the most abundant (about 60%). Plasma proteins have many roles, from maintaining osmotic pressure, transport of lipids, hormones, vitamins and minerals, to being a part of immune system and binding foreign bodies in the circulation.

Plasma protein adsorption on the surface of nanoparticles, which happens rapidly once they encounter the bloodstream, is one of the major reasons for largely unpredictable behaviour of nanoparticles *in vivo*¹⁰⁷⁻¹⁰⁹. This layer of biomolecules, termed the protein corona or biocorona, becomes a defining property of a nanoparticle thereby impairing cellular uptake or increasing it^{107,110,111}. Despite years of inventive designing of particles, research is still often thwarted by biology, which leads to limited clinical translation of targeted therapeutics¹¹². To overcome this, we have to account for the fact that nanoparticles behave differently in simple media than in tissues and take a more systematic approach of predicting and increasing *in vivo* performance by investigating the basics of biological mechanisms in nanoparticle delivery process.

The effect of proteins adsorbing on flat surfaces of different hydrophobicity or hydrophilicity was first described in the 1960s¹¹³. It was then found that this effect is especially pronounced on nanoparticles and since then notable progress has been made in both nanoparticle development and nanoparticle-biological environment studies¹¹¹. Research in past decades has contributed to a fundamental understanding in the mechanisms of the protein adsorption on nanoparticles: it has been established that size, surface charge, material and possible targeting molecules affect protein corona composition^{114,115}. Illustration of nanomaterial properties affecting protein adsorption on the surface of the particle is presented in Figure 1.8.

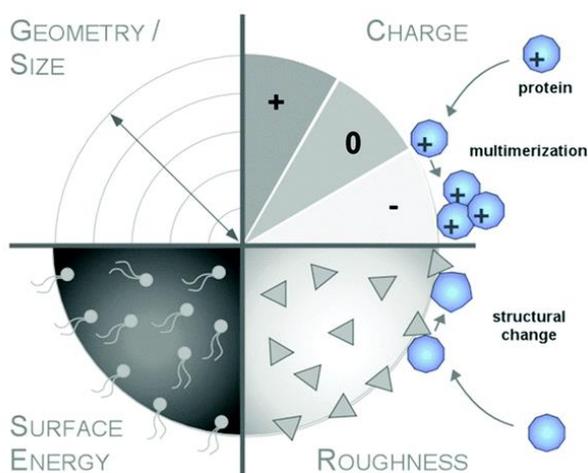


Figure 1.8. Physico-chemical characteristics of nanoparticles mediating protein interaction are particle size, charge, surface energy and roughness. Illustration copied from Docter, D. et al. *Chem. Soc. Rev.* 2015.

Beyond establishing how particle properties affect the protein corona, it has been suggested that corona composition may vary depending on nanoparticle entry site¹¹⁶. Moreover, recently it has been found that the corona contains more than just one type of organic molecules and this complex layer has been given the name of ‘biocorona’ or ‘biomolecule corona’¹¹⁷. Proteins and other biomolecules such as lipids, sugars and metabolites adsorbed on nanoparticles provide a new biological

identity to the particle. The adsorbed layer affects both passive and active targeting by altering particle size (additional layer or creating aggregates), surface energy, solubility, charge and accessibility of targeting moieties originally attached on the surface¹¹⁸.

The change in these physical properties of the particle in turn affects its interaction with the body. The first line effect is the rapid detection of the particle as a foreign body and removal from the system by the immune cells, specifically the macrophages. This recognition as a threat is guided by blood proteins, which cover unwanted substances in a process of opsonisation. Immunoglobulins and complement proteins improve uptake to specialised immune system cells and activate phagocytosis¹¹⁹. If the particle is not removed by the macrophages, it faces further challenges such as delivery to the site of action, which may be affected if the targeting moieties are hidden on the surface, and the release of the active pharmaceutical ingredient (API), which is then affected by a layer of macromolecules. In multiple studies, the protein coat impairs particle uptake to the cell, therefore reducing its efficacy and toxicity^{120–123}.

The mechanism of biocorona formation has been linked to the material hydrophilicity, as the proteins adsorb on the surface to reduce the surface tension. Due to this, nanoparticles with hydrophilic surface modification, such as poly (ethylene glycol) (PEG) have been shown to overcome negative consequences of protein adsorption, such as macrophagial uptake and rapid removal from the circulation^{124,125}. This so called stealth effect was introduced in 1995, when the first successful nanoparticle formulation was approved by the Food and Drug Administration (FDA)¹²⁶. This first nanoparticle drug, Doxil was a liposome formulation with a PEG layer. This effect is powerful to the extent that all

nanoparticle products currently available on the market are using the stealth PEG technology^{78,127,128}.

However, PEG is not biodegradable¹²⁹, may trigger immune responses¹³⁰ and its interaction with blood components is not clear¹³¹. For example, several recent studies have shown that PEGylation of a surface does not decrease protein adsorption in general, but reduces the non-specific protein attachment on the surface of PEGylated particles^{132,133} and promotes adsorption of 'don't eat me' proteins¹³⁴. These findings highlighted one of the key unanswered questions in the field, namely whether or not the physiological response is influenced by the whole corona or by specific proteins. Other not fully understood aspects include the dynamics of corona formation¹³⁵, weakly and strongly bound proteins¹³⁶ or presence of non-protein components in the biocorona¹³⁷. These questions remain unanswered due to lack of technology capable of *in situ* characterisation of the particle chemistry. A fundamental understanding of how nanoparticles interact with biological environment is needed to make these nanomedicine-biology interactions provide a desired response such as using the protein adsorption to increase selective cellular uptake^{118,138} or diagnostic purposes¹³⁹.

As found in many studies, size, surface charge, material and possible targeting molecules affect protein corona composition^{140,141}. Some reports are conflicting, pointing to relevance or irrelevance of protein concentration to their adsorption on the surface of nanoparticles^{142,143}. In several studies, proteins were shown to be binding to particles in a competitive manner, which may be a reason for some study-to-study disparities^{121,122}. Finally, the cell type and environment is also believed to play a role in the process of nanoparticle recognition and in the outcome of the encounter¹⁴⁴.

1.4.4 Nanoparticle characterisation methods

Protein corona research is driven by an observation of disproportion between studied and approved nanomedicines¹⁴⁵⁻¹⁴⁷. There is a similar imbalance between the number of protein corona studies and their application in nanomedicine development. The area lacks well defined protocols necessary for reproducible and reliable results^{115,148,149}. There is an evident variability in experimental setup: sample preparation conditions, sample manipulation and analysis. There is also irregularity in reporting all experimental detail, with several studies not mentioning crucial information about sample incubation such as temperature or length of time.

A clear factor affecting protein corona composition is the incubation medium, whether it is a single protein, a mixture of few chosen proteins, serum, plasma or whole blood. Obtained bio-nano conjugates differ qualitatively (different proteins adsorbed in different mixtures due to competitive character of the interactions) and quantitatively (different amount of proteins adsorbed due to dynamic adsorbing-desorbing on the surface). Concentration of proteins, type of buffer used in dilution and pH of the solution also greatly affect both particle stability in solution and the composition of the biocorona¹⁵⁰.

One of the key parameters determining experimental repeatability and reproducibility is the nanoparticle:protein ratio. This should be assured to remain constant and is measured as particle surface area to number of protein molecules¹⁵¹. Additional factors affecting results are temperature, length of incubation time or agitation of the particles during incubation, which vary in reported protocols from 5 minutes do 24 hours and from 4 °C to 37 °C^{152,153}.

In previously reported studies, there is a multitude of incubation and washing methods. After incubation, the unbound and weakly bound proteins are removed by centrifugation (with or without sucrose cushion, at various speed rates, different temperatures and for different length of time) or size exclusion chromatography, which is less disruptive and can preserve the weaker bound proteins on the surface of the particles. Magnetic particles can be separated from incubation media using magnetic separation¹⁵⁴.

With respect to characterisation of the particle coating, traditionally used methods can be divided into direct and indirect approaches. The direct methods provide information about protein identity in the corona. These methods include sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), often followed by electrospray mass spectrometry (ESI-MS)^{155,156}. These methods allow to identify large number of proteins from plasma incubated particles. This type of analysis requires extensive sample manipulation, separation of proteins from the particles, protein denaturation and digestion prior to the analysis, therefore it loses information about physical structure of the particle. Moreover, SDS-PAGE and MS focus only on protein while it has been shown recently that the corona may contain non-protein elements, such as lipids, sugars, nucleic acids and metabolites^{157,158}.

Indirect methods are used to supplement information about coat thickness and uniformity. This includes techniques such as dynamic light scattering (DLS), differential centrifugal sedimentation (DCS) or transmission electron microscopy (TEM) imaging and X-ray photoelectron spectroscopy (XPS)¹⁵⁹. Of these methods, only DCS can provide information about loosely bound corona morphology as the particle suspension can be analysed without prior washing¹⁶⁰. XPS and TEM can

operate in high vacuum therefore analysed protein is denatured and has lost its native structure.

Methods of analysing the biocorona described in section 1.1.4 provide important information about the number and types of proteins adsorbed on the surface of the particles, however, they require extensive sample manipulation and do not provide information about structure and distribution of the proteins and other biomolecules on the surface of the particles. A chemical imaging method is therefore needed to correlate spatial and chemical information regarding the biocorona¹⁶¹.

1.5 The scope of this thesis

The aim of this project is to develop a method of 3D *in situ* protein analysis without the need for matrix deposition or protein digestion. The recently developed 3D OrbiSIMS will be used for this goal. In contrast to classically used MALDI, LESA and DESI, SIMS enables three-dimensional analysis with minimal sample preparation. SIMS allows analysis of samples in three dimensions with lateral resolution < 200 nm, however previously only non-specific amino acid fragments of proteins and multi amino acid fragments from small peptides have been assigned in SIMS and protein fragments derived from undigested proteins have not been assigned to date. As described in Section 1.2, the 3D OrbiSIMS combines an argon GCIB analysis beam and OrbitrapTM analyser, which are hypothesised to enable more detailed analysis of proteins due to reduced fragmentation and high accuracy of the assignments.

Further aims include testing the suitability of the method for biotechnological and pharmaceutical analysis, specifically for studying the interaction of synthetic materials with biology. Drug delivery particles in biological media, which currently evade *in situ* characterisation by conventional analysis methods, will be manufactured using microfluidics and studied using the 3D OrbiSIMS as an example of a challenging pharmaceutical application.

2 Materials and methods

2.1 Materials

Proteins: lysozyme from chicken egg white, α -chymotrypsin from bovine pancreas, insulin solution human, recombinant bovine serum albumin, horse skeletal muscle myoglobin, L-lactate dehydrogenase from rabbit muscle, human holo-transferrin, concanavalin A from Jack Bean, bovine plasma fibronectin, alcohol dehydrogenase from *Saccharomyces cerevisiae*, porcine lipase, bovine liver catalase, human serum albumin and cytochrome c from equine heart were purchased from Sigma Aldrich.

Porcine pepsin and porcine trypsin were purchased from Promega.

Amicon Ultra 0.5 mL Centrifugal Filters were purchased from Merck Milipore.

Human skin was purchased from Tissue Solutions (<https://www.tissue-solutions.com/>). female > 40 years of age, ex vivo skin from cosmetic surgery.

Heptakis-(6-deoxy-6-thio)- β -cyclodextrin was purchased from Cyclodextrin-Shop (Tilburg, Netherlands) to >97% purity. Hydrogen peroxide (H₂O₂) 30% w/w

solution, sulfuric acid (H₂SO₄) reagent grade 95-95% and N,N-dimethylformamide for molecular biology ≥99% were purchased from Sigma Aldrich.

Gold slides for protein immobilisation, polycrystalline gold with 30nm thickness on special flat glass (Nexterion B from Schott AG) were purchased from Georg Albert PVD – Beschichtungen.

Poly (D,L-lactide-co-glycolide) (PLGA), lactide:glycolide (50:50), molecular weight (MW) 30,000-60,000 and poly (vinyl alcohol) (PVA) MW 9,000-10,000, 80% hydrolysed were purchased from Sigma Aldrich.

Microfluidic 100 µm flow focusing chip and 190 µm droplet junction chip were purchased from Dolomite. All chip fittings and chip interface were purchased from Dolomite.

2.2 Protein sample preparation

2.2.1 Protein films

100 µM protein solutions were prepared by dissolving appropriate amount of protein in 1 mL in MiliQ water (18 MΩ). Gold slides were cleaned by UV and sonication in isopropanol for 10 minutes, subsequently rinsed with MiliQ water and dried under Argon flow. Purified proteins were spotted and dried onto a gold slide 3 times to obtain a thick protein film. Protein:peptide mixture was prepared by 1:1 volumetric mixing of 100 µM insulin and 100 µM lysozyme concentrations. The mixture of purified lysozyme and insulin was spotted and dried onto a gold slide 3 times to obtain a thick protein film. Gold grids for electron transmission microscopy (TEM) of 200 mesh × 125 µm pitch were purchased from Sigma Aldrich. For the imaging experiment, the bare gold grid was placed on top of the protein film sample. A bare gold slide and a bare gold grid have been analysed as control samples.

2.2.2 Protein monolayer

For the preparation of protein monolayer biochips, the gold substrates were cleaned by immersion in piranha solution (70% H₂SO₄, 30% H₂O₂) at room temperature for 8 minutes, rinsed with MiliQ water and dried with an argon flow. (Caution: Piranha solution reacts violently with organic solvents and should be handled with care). The clean gold substrates were immersed for 24 h in 50 μM DMF solutions of heptakis-(6-deoxy-6-thio)-β-cyclodextrin (TCD) for preparation of a self-assembled monolayer (SAM). Subsequently, the gold substrates were rinsed with DMF and MiliQ water and dried under an argon flow. The TCD SAMs were immersed in a 0.05 mM lysozyme in PBS solution for 2 h. Following protein immobilization, the samples were washed with PBS buffer followed by submersion in MiliQ water for 1 min. The samples were then dried under argon and placed in the instrument for the analysis immediately.

2.3 Micro- and nanoparticle preparation

2.3.1 Microfluidics

10% w/v PLGA solution in DCM was prepared by dissolving 1 g PLGA in 10 mL DCM. For the tuning of the microparticle size, dilutions of 7.5%, 5%, 2.5%, 1%, 0.75%, 0.5% and 0.25% were used. For the production of the nanoparticles, 1%, 0.75%, 0.5% and 0.25% solutions of PLGA in acetone were prepared. 1% w/v PVA solution was prepared by dissolving 2 g of PVA in 200 mL MiliQ water at 90°C for 40 minutes under constant stirring. The polymer solution was placed in a 10 mL glass syringe purchased from Scientific Glass Engineering. Two plastic 10 mL syringes were filled with water (nanoparticle production) or surfactant solution

(microparticle production). The syringes were placed on the Harvard Apparatus pumps and attached to microfluidic setup via luer lock fittings. X-junction of the Dolomite 100 μm flow focusing chip was used for the microparticle production. X-junction of the Dolomite 190 μm droplet chip was used for the nanoparticle production. The flow ratio was calculated as ratio between 2 times the continuous phase flow divided by the dispersed phase flow, as the continuous phase is flowing through two channels compared to one channel of the dispersed phase. In case of microparticle production, the droplets were collected into glass vials containing 1 mL 1%w/v PVA solution. The receiving vial was placed on a magnetic plate and a stirring bar was placed inside the vial to enhance solvent evaporation during droplet collection. In case of the nanoparticle production, the nanoemulsion was collected into an empty glass vial.

2.3.2 Polymer particle functionalisation

Washed particles (1 mL) were incubated with 0.5 mL of BSA, HSA or full human serum solution. The incubation was carried out in glass vials under constant agitation for 2 hours at room temperature. After the incubation the particles were washed 3 times by centrifugation at 9000 g.

2.4 Analytical methods

2.4.1 Dynamic light scattering (DLS)

DLS measurements were taken using Zetasizer Nano. Refractive index (RI) of 50:50 PLGA in light scattering measurements calculations was applied as found in literature, with sources reporting $\text{RI}=1.47\text{-}1.61$ ¹⁶² or $\text{RI}=1.51\text{-}1.67$ ¹⁶³.

2.4.2 Laser diffraction (LD)

Laser diffraction measurements were taken with Coulter LS230. Refractive index (RI) of 50:50 PLGA in light scattering measurements calculations was applied as found in literature, with sources reporting RI=1.47-1.61¹⁶² or RI=1.51-1.67¹⁶³.

2.4.3 Scanning electron microscopy (SEM)

For SEM measurements, the particles suspended in water were dropped onto carbon tape stuck to a SEM stub. The drops were allowed to dry overnight in the fumehood. After drying, the samples were coated with platinum for 120 s. SEM measurements were taken using Jeol-6490. The energy was set to 10 keV. Images were processed in ImageJ in order to measure the average diameter of all particles in the field of view. The intensity threshold was adjusted for each image separately depending on contrast, the pixel to length was set based on the scale bar and the Particle analysis function was used to obtain the diameter of all particles within the threshold.

2.4.4 X-ray photoelectron spectroscopy (XPS)

The self-assembled monolayer and the self-assembled monolayer with immobilised lysozyme were measured using an Axis-Ultra XPS instrument (Kratos Analytical, UK) with monochromated Al K α X-ray source, based at the National Physical Laboratory (NPL), Teddington, UK. Wide scans were acquired for analysis of organic layer thickness on the gold slide with step size 1000 meV, pass energy 160 eV, dwell time 300 ms, in a range 1300 to -10 eV.

2.4.5 Time of flight secondary ion mass spectrometry (ToF-SIMS)

For the acquisition of LMIG ToF-SIMS image, a 30 keV Bi₃⁺ primary beam was used. LMIG current was 0.05 pA. The 256 \times 256 pixel ToF image was acquired over an area of 200 \times 200 μ m using random raster. The cycle time was set to 250 ms.

Optimal target potential was set to +58 V. Three separate areas were analysed on each sample and each measurement lasted 15 scans, the total ion dose per measurement was 9.44×10^{10} ions/cm².

The results were analysed in SurfaceLab 7. Calibration of positive mode spectra was made using CH₃⁺, C₂H₃⁺, C₃H₅⁺, C₇H₇⁺ and Au₃⁺ ions. Negative mode spectra were calibrated to CN⁻, CNO⁻, Au⁻, Au₃⁻ and Au₅⁻.

2.4.6 3D OrbiSIMS

3D OrbiSIMS spectra

A 20 keV Ar₃₀₀₀⁺ analysis beam of 20 μm diameter, was used as the primary ion beam. The GCIB duty cycle was set to 4.4%, with current 218 pA. The depth profile was run on the area of 200 × 200 μm using random raster mode with crater size 280 × 280 μm. The cycle time was set to 400 μs. Optimal target potential was approximately +68 V. Argon gas flooding was in operation in order to aid charge compensation, pressure in the main chamber was maintained at 9.0×10^{-7} mbar. The spectra were collected in positive polarity, in the *m/z* range 150 - 2250. The injection time was set to 500 ms. Three separate areas were analysed on each sample and each measurement lasted 30 scans, the total ion dose per measurement was 1.63×10^{11} ions/cm². Mass resolving power was set to 240,000 at *m/z* = 200.

3D OrbiSIMS MS/MS

A 20 keV Ar₃₀₀₀⁺ analysis beam of 20 μm diameter, was used as primary ion beam. Ar₃₀₀₀⁺ with duty cycle set to 4.4%, GCIB current was 218 pA. The depth profile was run on the area of 200 × 200 μm using random raster mode with crater size 280 ×

280 μm . The cycle time was set to 400 μs . The target potential was set to +68 V. Argon gas flooding was in operation in order to aid charge compensation, pressure in the main chamber was maintained at 9.0×10^{-7} mbar. The spectra were collected in positive polarity, in m/z range 75 - 1125. The injection time was set to 500 ms. Three separate areas were analysed on the sample and each measurement lasted 20 scans, the total ion dose per measurement was 1.63×10^{11} ions/ cm^2 . The mass resolving power was set to 240,000 at m/z 200. The normalised collision energy (NCE) was set to 35 and the isolation width was $\Delta m/z$ 10. The normalised collision energy is the collision energy normalised to the collision of a singly charged precursor ion at mass $m/z = 500$ and it increases the energy linearly with increasing mass of the precursor ion, therefore the collision energy in this experiment ($m/z = 620$) was 43.4 eV.

3D OrbiSIMS image

20 keV Ar_{3000}^+ analysis beam of 20 μm diameter (imaging of the protein mixture) or a 20 keV Ar_{3000}^+ imaging beam of 5 μm diameter (imaging of lysozyme film masked with a gold grid) was used as primary ion beam. The 20 μm analysis beam was configured as described in the spectra acquisition section. The 5 μm imaging beam duty cycle set to 37.7%, GCIB current was 18 pA. The Q Exactive images were run on the area of $200 \times 200 \mu\text{m}$ using random raster mode. The cycle time was set to 400 μs . Optimal target potential was set to +58 V. Argon gas flooding was in operation in order to aid charge compensation, pressure in the main chamber was maintained at 9.0×10^{-7} mbar. The images were collected in positive polarity, in m/z range 150-2250. The injection time was set to 500 ms. Three separate areas were analysed on each sample and each measurement lasted 1 scan, the total ion dose per measurement was 1.11×10^{12} ions/ cm^2 . Mass resolving power was set to 240000 at $m/z = 200$.

3D OrbiSIMS depth profile of human skin

20 keV Ar_{3000}^+ analysis beam of 20 μm diameter, was used as primary ion beam. Ar_{3000}^+ with duty cycle set to 15%, GCIB current was 230 pA. The Q Exactive depth profile was run on the area of 200 \times 200 μm using random raster mode with crater size 280 \times 280 μm . The cycle time was set to 200 μs . Optimal target potential was approximately -384.4 V. The acquisition was run in temperature -170°C. The depth profile was collected in negative polarity, in m/z range 75 - 1125. The injection time was set to 500 ms. The measurement lasted 18307 scans, the total ion dose per measurement was 2.65×10^{13} ions/ cm^2 . Mass resolving power was set to 240,000 at $m/z = 200$.

2.4.7 Optical profilometry

The depth of the sputtered material was estimated by the SurfaceLab software as 10 nm per scan and confirmed by profilometry as 300 nm after acquisition of 30 scans. Optical profilometry scans were obtained using a Zeta-20 optical microscope (Zeta Instruments, CA, USA). The scans were acquired in a Z range of 4.6 μm . The number of steps was set to 328, allowing for step size (Z resolution) of 0.014 μm .

2.4.8 Ellipsometry

The thickness of the self-assembled monolayers and the self-assembled monolayer with immobilised lysozyme was measured using J.A. Woollam Co. Inc. alpha-SETM spectroscopic ellipsometer. The CompleteEASE software was employed to determine the thickness values and the calculations were based on a two-phase organic/Au model, in which the organic layer was assumed to be isotropic and assigned a refractive index of 1.50. The thickness reported is the average of three

different measurements on SAM or SAM and protein samples, with the errors reported as standard deviation.

2.4.9 Statistical analysis

Multivariate analysis of ToF-SIMS results was done in Matlab GUI, simsMVA117, using Matlab R2017b. The peak list from table A1 in appendix was applied to 4 regions of interest on all samples. Peak areas normalised to total ion count were loaded into simsMVA. The data was preprocessed by Poisson scaling and mean centring. PCA was run in algorithm mode, retaining all principal components.

3 Development of ballistic sequencing for assignment and identification of proteins using 3D OrbiSIMS

3.1 Chapter aims

The use of gas cluster ion beams such as Ar_n^+ in SIMS allows assignments on peptide fragments up to 1200 Da and detection of molecular ions of peptides up to 6000 Da, as detailed in Section 1.2. However, protein fragments derived from larger undigested proteins have not been assigned in SIMS. This chapter aims to assess the capability of the 3D OrbiSIMS to acquire chemically specific information from protein samples by the use of Ar_{3000}^+ primary beam and OrbitrapTM analyser. The additional MS/MS capability is expected to help in the understanding of the complex datasets. Limit of detection and lateral resolution achievable with the method will also be assessed.

Further aims of this chapter are to analyse a benchmark of 16 proteins to validate the method and demonstrate the capabilities and limitations of this method as well as place this emerging technology in the context of established mass spectrometry proteomic methods.

3.2 Methods

3.2.1 Protein films

100 μM protein solutions were prepared by dissolving appropriate amount of protein in 1 mL in MiliQ water (18 M Ω). Gold slides were cleaned by UV and sonication in isopropanol for 10 minutes, subsequently rinsed with MiliQ water and dried under Argon flow. Purified proteins were spotted and dried onto a gold slide 3 times to obtain a thick protein film. Protein:peptide mixture was prepared by 1:1 volumetric mixing of 100 μM insulin and 100 μM lysozyme concentrations. The mixture of purified lysozyme and insulin was spotted and dried onto a gold slide 3 times to obtain a thick protein film. Gold grids for electron transmission microscopy (TEM) of 200 mesh \times 125 μm pitch were purchased from Sigma Aldrich. For the imaging experiment, the bare gold grid was placed on top of the protein film sample. A bare gold slide and a bare gold grid have been analysed as control samples.

3.2.2 ToF-SIMS measurements

For the acquisition of LMIG ToF-SIMS image, a 30 keV Bi_3^+ primary beam was used. LMIG current was 0.05 pA. The 256 \times 256 pixel ToF image was acquired over an area of 200 \times 200 μm using random raster. The cycle time was set to 250 ms. Optimal target potential was set to +58 V. Three separate areas were analysed on each sample and each measurement lasted 15 scans, the total ion dose per measurement was 9.44×10^{10} ions/cm². The results were analysed in SurfaceLab 7. Positive mode spectra were calibrated using CH_3^+ , C_2H_3^+ , C_3H_5^+ , C_7H_7^+ and Au_3^+ . Negative mode spectra were calibrated to CN^- , CNO^- , Au^- , Au_3^- and Au_5^- .

Multivariate analysis of ToF-SIMS results was done in simsMVA, using Matlab R2017b. The peak list normalised to total ion count was applied to 4 regions of interest on all samples. The data was preprocessed by Poisson scaling and mean centring. PCA was run in algorithm mode, retaining all principal components.

3.2.3 3D OrbiSIMS measurements

A 20 keV Ar_{3000}^+ analysis beam of 20 μm diameter, was used as the primary ion beam. The GCIB duty cycle was set to 4.4%, with current 218 pA. The depth profile was run on the area of $200 \times 200 \mu\text{m}$ using random raster mode with crater size $280 \times 280 \mu\text{m}$. The cycle time was set to 400 μs . Optimal target potential was approximately +68 V. Argon gas flooding was in operation in order to aid charge compensation, pressure in the main chamber was maintained at 9.0×10^{-7} mbar. The spectra were collected in positive polarity, in the m/z range 150 - 2250. The injection time was set to 500 ms. Three separate areas were analysed on each sample and each measurement lasted 30 scans, the total ion dose per measurement was 1.63×10^{11} ions/ cm^2 . Mass resolving power was set to 240,000 at $m/z = 200$.

For 3D OrbiSIMS MS/MS, normalised collision energy (NCE) was set to 35 and the isolation width was $\Delta m/z$ 10. The normalised collision energy is the collision energy normalised to the collision of a singly charged precursor ion at mass $m/z = 500$ and it increases the energy linearly with increasing mass of the precursor ion, therefore the collision energy in this experiment ($m/z = 620$) was 43.4 eV.

For the 3D OrbiSIMS imaging of lysozyme film masked with a gold grid, 20 keV Ar_{3000}^+ of 5 μm diameter was used as primary ion beam. The 5 μm imaging beam duty cycle was set to 37.7% and the GCIB current was 18 pA.

3.3 Results and discussion

3.3.1 Capabilities and limitations of ToF-SIMS for assignment of protein fragments

Lateral resolution of and sensitivity of SIMS is advantageous in comparison to classical methods of analysing proteins on the surface such as MALDI and DESI as described in Section 1.3. Here, a lysozyme film, prepared by drop-drying a solution of lysozyme on a gold slide, was masked by the bare gold TEM grid. Relatively intense fragmentation by Bi_3^+ primary beam results in the detection of unspecific amino acid fragment ions as shown in Figure 3.1. The peaks below $m/z = 200$ are amino acid fragments assigned according to Lhoest *et. al*²⁹ and are present in all analysed proteins. Peaks at the higher ($m/z > 200$) mass region were found to be related to pure gold and gold clusters, Au-S bond or other non-specific organic fragments as shown in Figure 3.1.

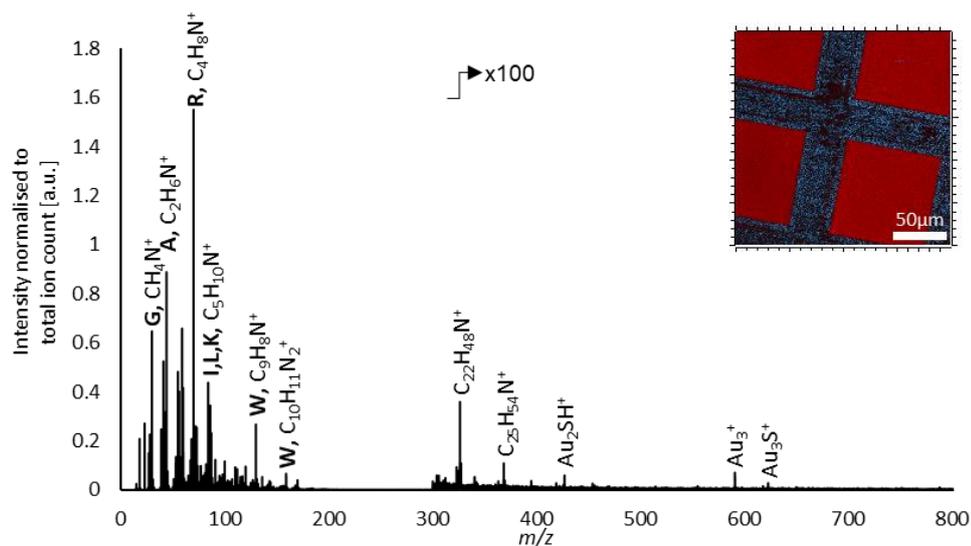


Figure 3.1. ToF-SIMS spectrum of lysozyme film masked by a gold grid and corresponding image (inset). The spectrum represents a region of interest consisting of the areas under the grid (red in the image), however several gold-related peaks are still visible in the spectrum (blue in the image).

In addition to the intense fragmentation induced by the LMIG, the mass resolving power of the ToF analyser does not always allow for confident assignment of the protein fragments. Figure 3.2 presents a comparison of LMIG/ToF and GCIB/OrbitrapTM analyses of lysozyme film dried on a gold slide, focused on the area of the spectrum, assigned previously as a tryptophan fragment ion ($C_{10}H_{11}N_2^+$). The OrbiSIMS spectrum reveals that the ToF peak includes four lower intensity non-specific organic fragments in addition to the previously assigned amino acid fragment.

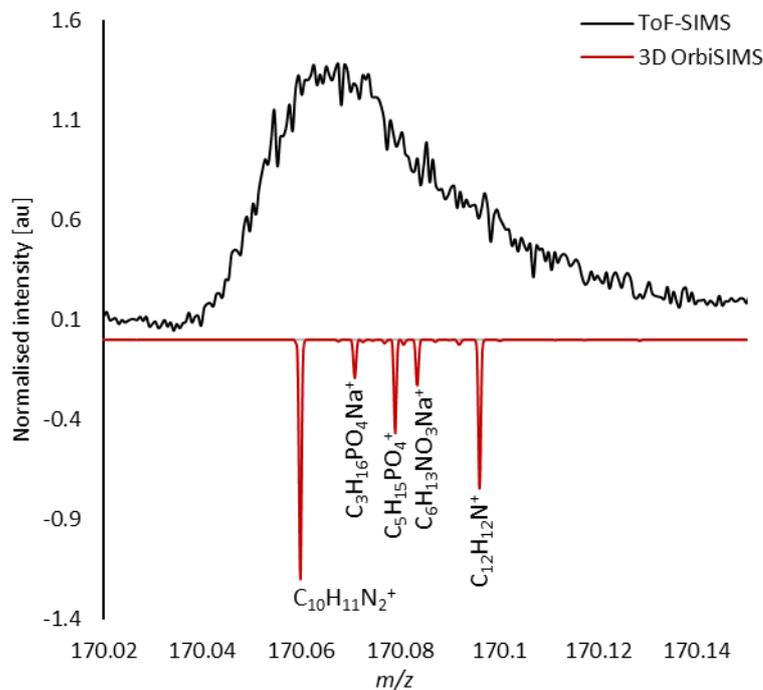


Figure 3.2. Comparison of ToF-SIMS and 3D OrbiSIMS spectra of chicken egg lysozyme, the 170.0616 m/z peak area.

Despite all proteins being composed of the same units, a set of 20 amino acids, the relative intensities of the secondary ion peaks detected by the time of flight analyser create patterns, encoding information about protein identity³⁷, conformation¹⁶⁴ or orientation³¹. These patterns can be identified and compared using multivariate statistics. Principal component analysis (PCA) was chosen as the most suitable method for distinguishing different protein samples using the LMIG ToF spectra.

PCA was performed on 7 different proteins immobilised on heptakis-(6-deoxy-6-thio)- β -cyclodextrin (TCD) layers to assess whether different protein samples can be distinguished. The amino acid peak list assigned by Lhoest *et al*⁵ was used as the series of variables. Figure 3.3 presents the scores of first principal component (PC1) and second principal component (PC2).

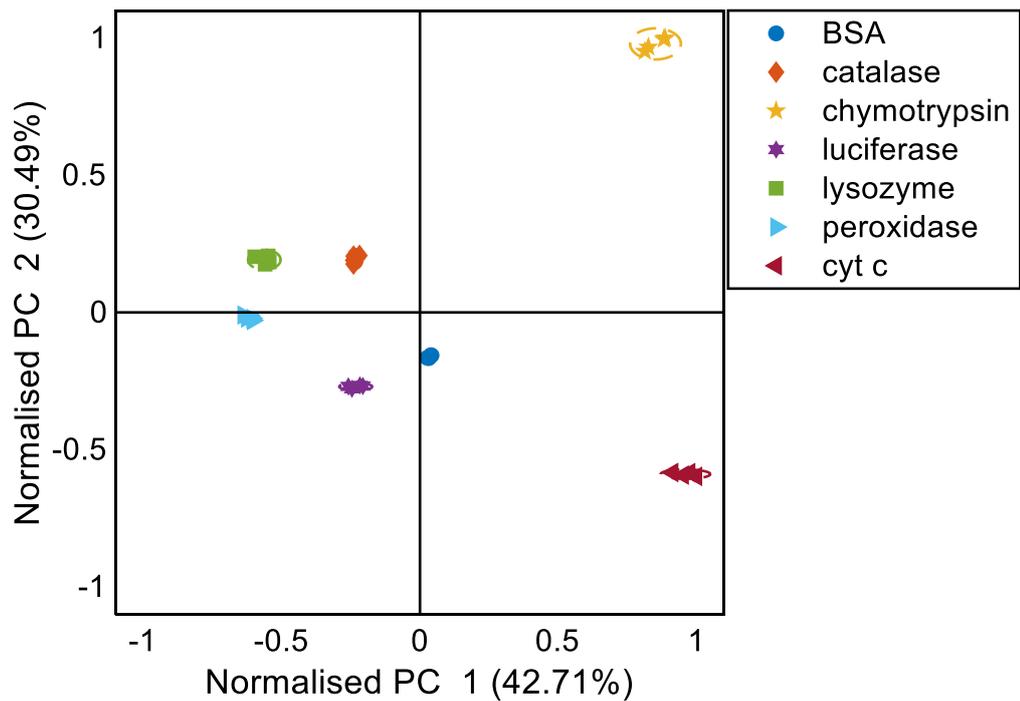


Figure 3.3. Scores of PC1 and PC2. The amino acid fragment intensity patterns are consistent within the proteins and vary between different proteins.

The scores plot in Figure 3.3 suggests that sample chemistry varies for different proteins, therefore the samples can be distinguished between each other by using this statistical method. Four regions of interest were analysed on each sample and 95% confidence intervals were added to all sample types. PC1 and PC2 together explain 73.2% variation in the dataset. The scores plot does not provide sufficient information on its own and has to be followed by a loadings graph (Figure 3.4).

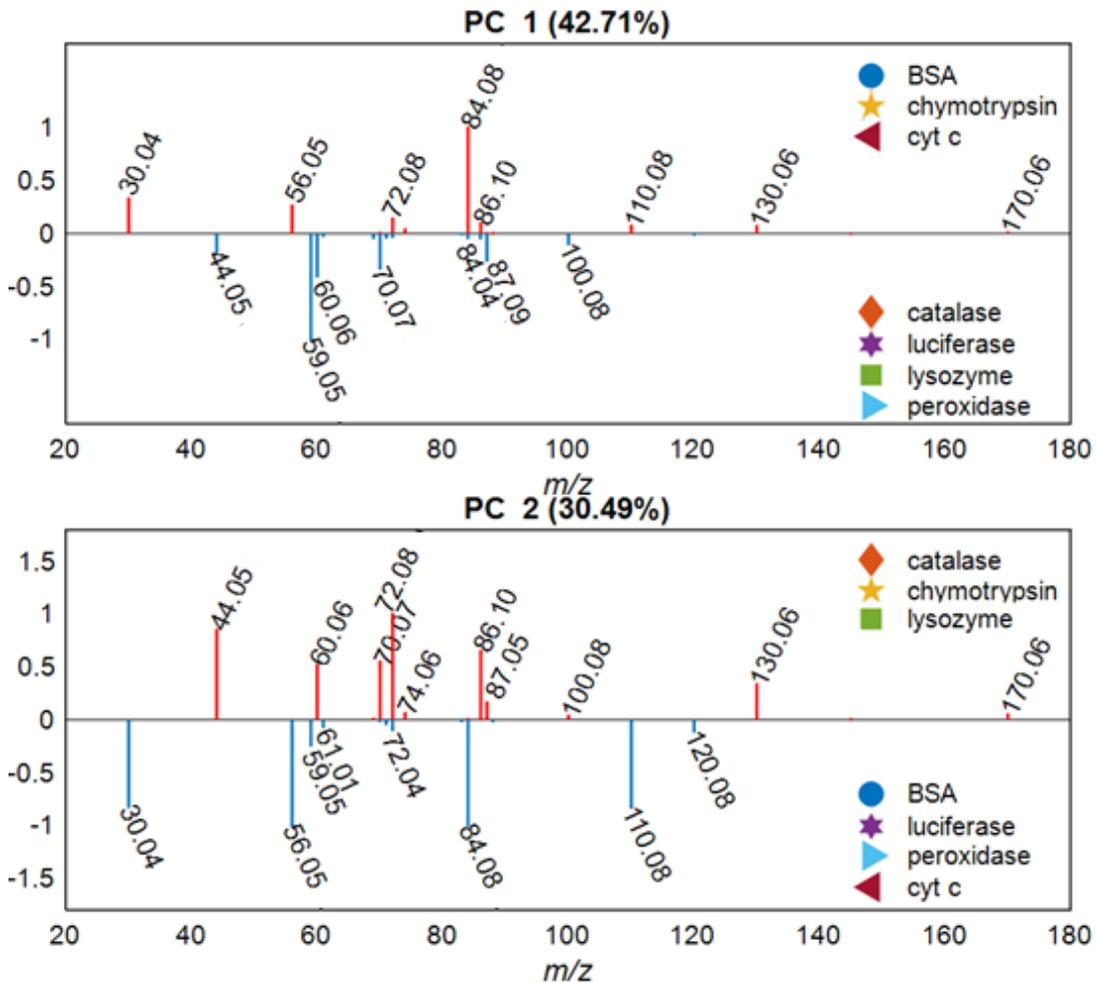


Figure 3.4. Principal component analysis of 7 different proteins immobilised on TCD-terminated chips. Loadings plot of first (PC1) and second (PC2) principal component. Detailed explanation in the main text.

The loadings plot demonstrates the variables, in this case amino acid fragments peak intensities, responsible for variation in the dataset. The main differences between the samples are caused by peaks at m/z 30.04 (glycine), m/z 58.05 (arginine) and m/z 84.08 (leucine/isoleucine/lysine). Samples with positive scores on PC1 are represented by positive loadings on PC1, which means that signals at m/z 30 and m/z 84 are more prevalent in lysozyme, catalase and chymotrypsin than in the other proteins (Figure 3.3).

As previously shown by Lhoest *et al.* and other studies, and as illustrated using a series of proteins in this work (Figure 3.3), single protein systems can be

distinguished between each other by using statistical analysis. With the use of appropriate control samples, PCA can also provide information about protein orientation on a biochip¹⁶⁵. This approach can provide important information and is useful in comparing simple samples, however it does not provide sufficient detail for protein identification in complex systems, especially biological samples, in which amino acid peaks may overlap with other organic molecules.

3.3.2 Protein identification using GCIB/Orbitrap using model samples

Multivariate analysis of amino acid ion intensities acquired with LMIG and ToF analyser provides information related to protein identity (Section 3.2.1), however it can only be applied to single protein samples or simple (binary) mixtures. Large cluster ion beams such as Ar_{3000}^+ have been demonstrated to produce protein fragments larger than single amino acids¹⁶⁶. There are however, issues relating to the ability of the ToF analyser to assign relatively high mass secondary ions and resolve similar mass species produced by such systems. The relatively high mass resolving power of the OrbitrapTM analyser in combination with the GCIB was anticipated to enable assignment of observed peaks with high confidence, achieving more detailed characterization of proteins at surfaces.

For this chapter, 100 μM protein solutions were spotted and dried onto a gold slide to obtain protein films. The spectra acquired from the protein films using Ar_{3000}^+ primary ion beam and the OrbiTrapTM analyser contain large ($m/z > 300$) protein fragments compared to LMIG/ToF as shown in Figure 3.5. However, the spectra do not match any existing protein databases such as Mascot or the National Institute of

Standards and Technology (NIST) database. These resources are created for protein digests and the peptidic secondary ion fragments observed in GCIB/Orbi analysis are derived from ballistic fragmentation of the intact proteins in contrast to enzymatically digested proteins analysed in classical LC-MS, therefore they cannot be readily assigned using ‘peptide fingerprint’ identification. Nevertheless the spectra contain multiple repeating neutral losses between the peaks, as highlighted in Figure 3.5. The neutral losses belong to different amino acids and can be assigned with accuracy due to the high mass resolving power of the Orbitrap analyser. The presence of amino acid neutral losses allows the *de novo* sequencing method of protein identification.

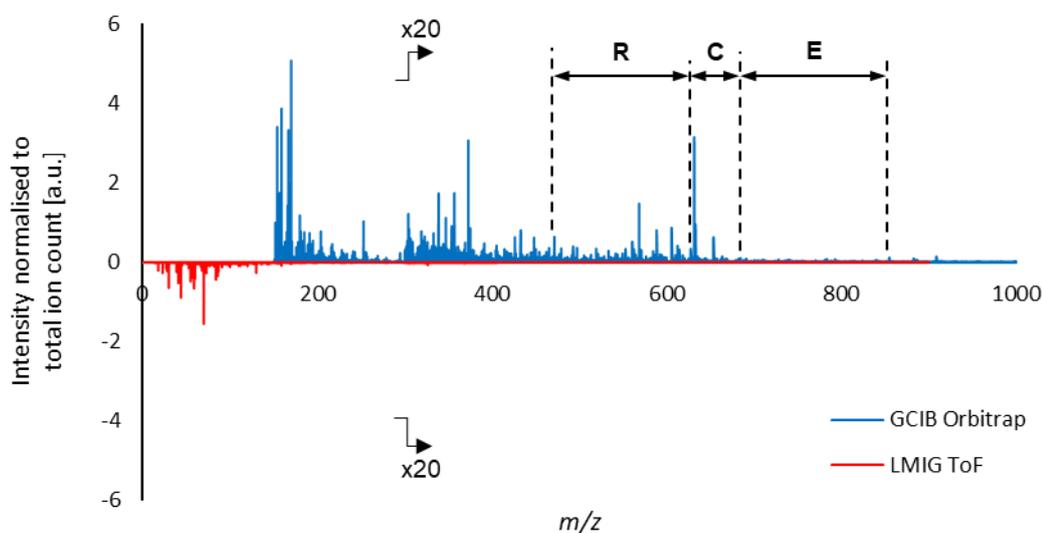


Figure 3.5. Inverted overlay comparison of positive polarity ToF-SIMS (red) and 3D OrbiSIMS (blue) spectra of purified chicken egg lysozyme film. The total ion dose per LMIG ToF measurement was 9.44×10^{10} ions/cm², the total ion dose per GCIB Orbitrap measurement was 1.63×10^{11} ions/cm².

In addition to the main 20 amino acid residues, neutral losses of 27.9949 u (CO) and 17.0265 u (NH₃) are frequently observed in the spectra. These residues are the effect of the different sites of the amide bond cleavage during the ballistic fragmentation,

producing fragments known in proteomics as a, b and c ions (Figure 3.6b)¹⁶⁷. The most common types of ions in traditional mass spectrometry of proteins are the b ions and the corresponding C-terminus y ions. These ions are achieved by the lowest collision energy¹⁶⁸. Higher collision energy, or use of electrons in the fragmentation process, leads to a and c ions¹⁶⁹.

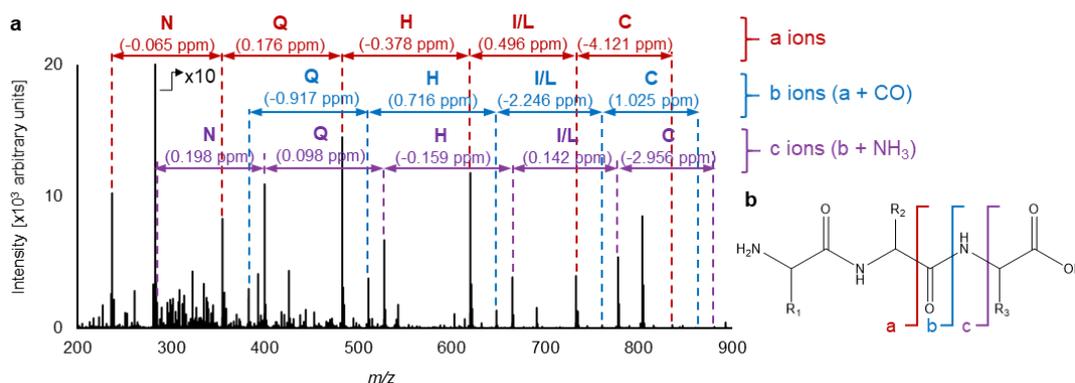


Figure 3.6. Positive polarity 3D OrbiSIMS spectrum of thin (300 nm) insulin film. The first 8 amino acids of the insulin B chain (FVNQHLC) dominate the spectrum. Each peptide fragment (FVN, FVNQ, FVNQH etc.) is observed as a-, b- and c-type ion. (b) N-terminus ions a, b and c are formed when different bonds related to the peptide bond are broken.

The initial hypothesis of acquiring primary structure information from spectra by *de novo* sequencing was confirmed by an MS/MS experiment carried out during 3D OrbiSIMS analysis of human insulin sample. The m/z 620.29 peak, assigned as the FVNQHLC sequence of insulin was further fragmented with collision energy of 43 eV, using the MS/MS capability of the instrument. The resulting spectrum (Figure 3.7, red), contained the ions assigned as the FVNQHLC sequence in the full spectrum (Figure 3.7, black). Additionally, the MS/MS experiment enabled confirmation of the identity of the smallest fragment in the sequence (FVN) by detection of additional neutral losses (Figure 3.7, red), which were not clearly visible in the full survey spectrum.

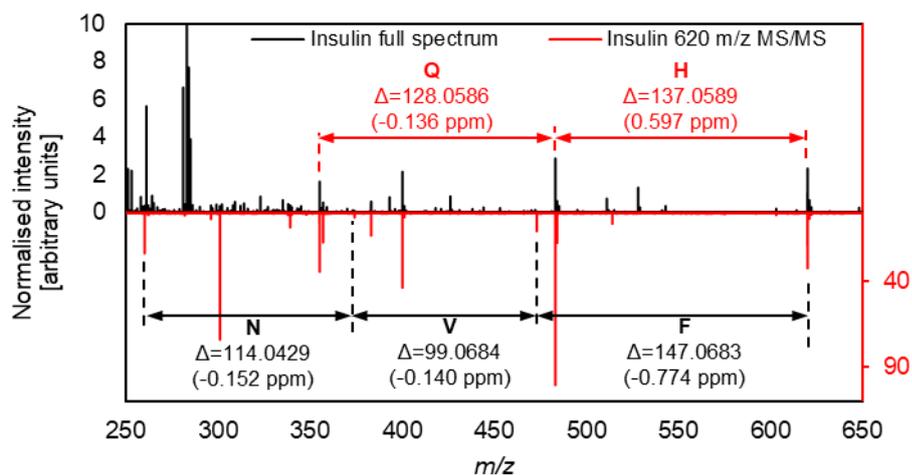


Figure 3.7. Inverted overlay comparison of full 3D OrbiSIMS spectrum of insulin (black) and 3D OrbiSIMS MS/MS of the m/z 620.29 (red) ion, assigned as the first 5 amino acids in the insulin chain *b* sequence: FVNQH. All labelled amino acid neutral losses are present in both the full spectrum and the MS/MS spectrum, which confirms the suggested fragmentation. The peak at m/z 300.05 in MS/MS spectrum (red) is an instrument artefact and should be ignored.¹⁷⁰

3.3.3 Interpretation of ballistic fragmentation spectra by comparison with classical mass spectrometry proteomics methods

In order to validate the ballistic sequencing method, 16 model protein films covering a range of biological functions and sizes from insulin (6 kDa) to fibronectin (272 kDa) were analysed using the 3D OrbiSIMS (Figure 3.8). Amino acid sequence coverages from 5% (33 of 679 amino acids for transferrin) to 53% (68 of 129 amino acids for lysozyme) can be readily assigned in the positive polarity spectra (Table 3.1).

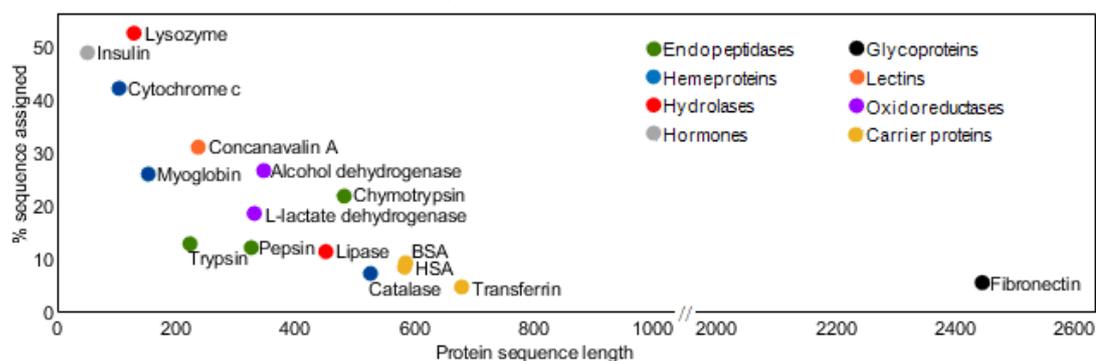


Figure 3.8. 16 proteins of different biological functions were analysed. From 5% (transferrin) to 53% (lysozyme) sequence coverage was achieved.

Table 3.1. List of analysed proteins from the smallest (insulin, 51 amino acid sequence) to the largest (fibronectin, 2446 amino acid sequence). Sequence coverage achieved with the 3D OrbiSIMS is presented as number of assigned amino acids and fraction of the whole protein sequence.

Name	UniProt database code	Length of the protein sequence (amino acids)	Number of assigned amino acids	% of the sequence assigned
Insulin	INS_HUMAN	51	25	49.02
Cytochrome c	CYC_HORSE	104	44	42.31
Lysozyme	LYSC_CHICK	129	68	52.71
Myoglobin	MYG_HORSE	153	40	26.14
Trypsin	TRYP_PIG	223	29	13.00
Concanavalin A	CONA_CANEN	237	74	31.22
Pepsin	PEPA_PIG	326	40	12.27
L-lactate dehydrogenase	LDHA_RABIT	331	62	18.73
Alcohol dehydrogenase	ADH1_YEAST	347	93	26.80
Lipase	LIPL_PIG	451	52	11.53
Chymotrypsin	CTRA_BOVIN	482	47	9.75
Catalase	CATA_BOVIN	526	39	7.41
BSA	ALBU_BOVIN	583	55	9.43
HSA	ALBU_HUMAN	585	50	8.58
Transferrin	TRFE_HUMAN	679	33	4.86
Fibronectin	FINC_BOVIN	2446	139	5.68

The 3D OrbiSIMS spectra of proteins contain multiple types of ions, which complicates the interpretation of the data. Additionally, apart from N- and C-terminal ions, a large proportion of the peaks observed in the spectrum are internal fragments of the sequence. Comparison of the ballistic sequencing spectra and the methods classically used in proteomics was carried out to understand the rules by

which the ions are formed and facilitate assignment of multiple amino acid sequences in different proteins.

Similarities between the peptidic fragments observed in the GCIB/OrbitrapTM spectra and spectra obtained by low energy collision induced ionisation (CID) MS/MS of tryptic peptides include the presence of a, b and a-NH₃ ions^{167,168}. In the GCIB/OrbitrapTM spectra, ions representing the N-terminus (a, b, c, a-NH₃ ions) of the protein sequence are detected as well as the C-terminus fragments (y ions). Similarly to CID fragmentation, cysteine residues participating in disulphide bonds have been found to form R-S, R-SH₂ ions in the GCIB/OrbitrapTM spectra, resulting from partial side chain cleavage^{171,172}. This has been shown on examples of insulin, lipase, lysozyme, trypsin, bovine serum albumin, human serum albumin and fibronectin and the assignments are available in [Supplementary Table 3](#)¹⁷³.

In contrast to CID fragmentation, c and z fragments, previously seen in electron capture dissociation (ECD)¹⁷⁴ are commonly detected in the GCIB/OrbitrapTM spectra ([Supplementary Tables 10-147](#)¹⁷³). The abundance of internal ions is a similarity to high energy collision dissociation (HCD)¹⁶⁹. Internal fragments are formed when a y ion undergoes further fragmentation into ya, yb, yc and ya-NH₃ ions. The proposed structure of the internal fragment ya for the NAWVA sequence is presented in Figure 3.9a. At C-termini, in addition to y ions, z, z+1 and z+2 ions, characteristic of MALDI in-source decay (ISD)¹⁷⁵⁻¹⁷⁷ are observed in the GCIB spectra. The proposed structure of a z+1 ion for the RGCRL sequence is given in Figure 3.9b.

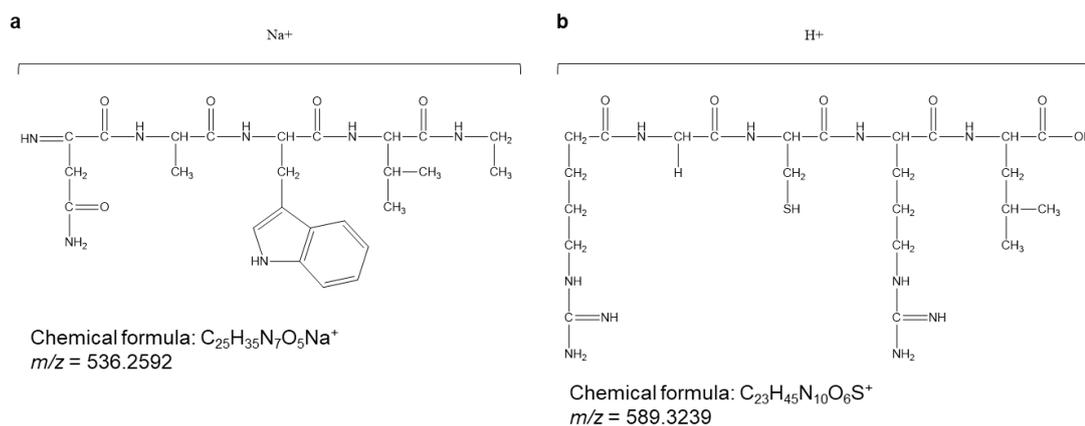


Figure 3.9. Proposed structures of observed ions dissimilar to CID MS of tryptic peptides, similar to HCD or MALDI-ISD fragmentation. (a) Internal ion *ya* is presented on an example NAWVA sequence observed in the lysozyme spectrum. (b) C-terminal ion *z+1* is presented on an example RGCRL sequence observed in the lysozyme spectrum.

As the charge usually resides at the N- or C-terminus of the protein¹⁷⁸ and only one main chain bond needs to be broken in the formation of N- and C-terminus ions, these types of fragments form the longest sequences in the obtained GCIB/OrbitrapTM spectra of 16 model proteins. For proteins composed of multiple chains, terminus sequences were observed for one or more chains, as presented on the example of bovine chymotrypsin (Figure 3.10). In the case of α -chymotrypsin, chain A, which consists of only 13 amino acids, was also detected as a whole peptide (Figure 3.10, Table 3.2), which suggests an ability of complete sequence coverage and therefore identification of peptides composed of short ($m/z < 1300$) chains.

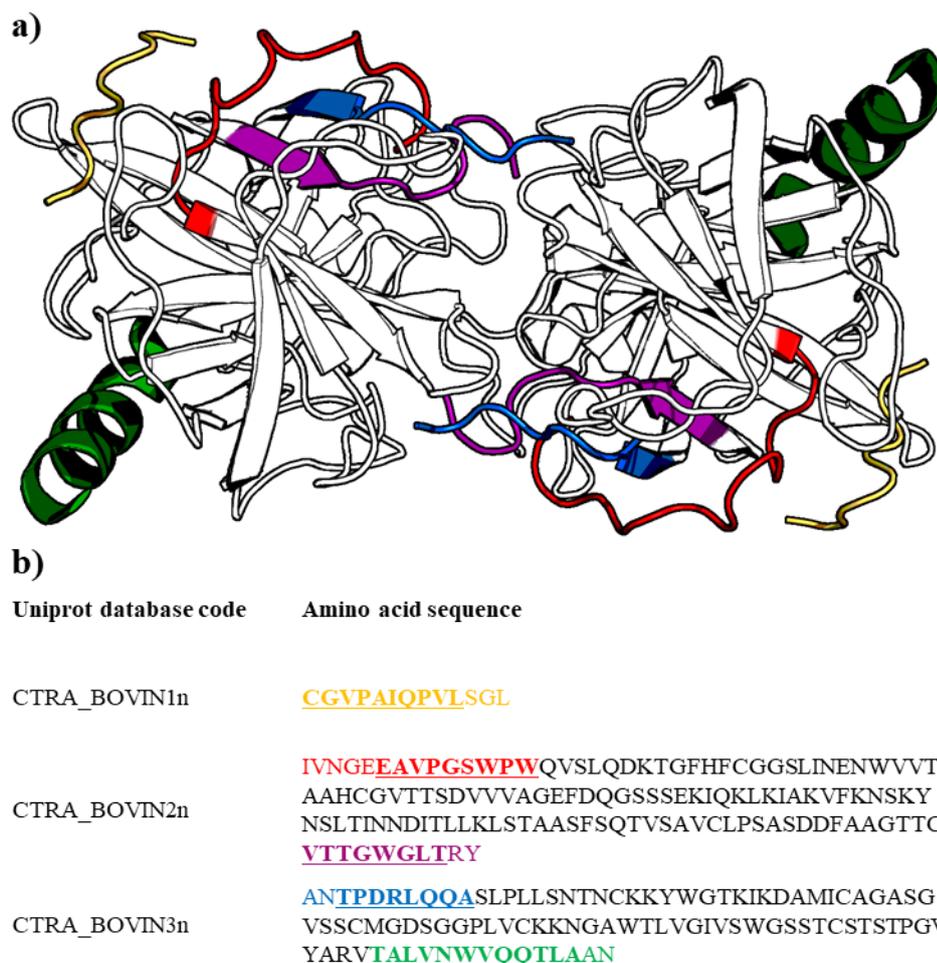


Figure 3.10. Structure and amino acid sequence of bovine α -chymotrypsin. The colours in the structure cartoon correspond to the highlighted parts of the amino acid sequence and in the colours of ion assignments in [Supplementary Table 148-152¹⁷³](#).

Table 3.2. Peak list, description of sequence, measured mass, assignment and deviation from calculated mass for ions from positive polarity spectrum of α -chymotrypsin from bovine liver, ions assigned as C-terminal sequence of the A chain GVP AIQPVLSGL and a full A chain CGVPAIQPVLSGL. The colour corresponds to the presence of the observed sequence in α -chymotrypsin presented in Figure 3.10.

Description	m/z	Assignment	Dev. (ppm)
SGL	298.1374	C ₁₁ H ₂₁ N ₃ O ₅ Na ⁺	0.1814
LSGL	411.2216	C ₁₇ H ₃₂ N ₄ O ₆ Na ⁺	0.3605
PVLSGL	607.3429	C ₂₇ H ₄₈ N ₆ O ₈ Na ⁺	0.4458
IQPVLSGL	848.4858	C ₃₈ H ₆₇ N ₉ O ₁₁ Na ⁺	0.6353
PAIQPVLSGL	1016.5762	C ₄₆ H ₇₉ N ₁₁ O ₁₃ Na ⁺	1.1239
GVP AIQPVLSGL	1172.6668	C ₅₃ H ₉₁ N ₁₃ O ₁₅ Na ⁺	1.5263
CGVPAIQPVLSGL	1275.6733	C ₅₆ H ₉₆ N ₁₄ O ₁₆ SNa ⁺	-0.6437

Chymotrypsin was selected for detailed analysis of negative polarity spectrum as it is a large protein, composed of multiple chains. The negative polarity spectra in proteomics can provide additional information to complement the positive mode, such as details of phosphorylation sites^{179,180}. This work focuses on the analysis of positive mode data in order to place this emerging technique in the context of the widely used established CID and MALDI methods. Fragmentation patterns analogous to positive polarity spectra were observed in the negative polarity spectra acquired with Ar₃₀₀₀⁺ GCIB and OrbitrapTM (Table 3.3).

The C-termini of proteins contain a free carboxyl group, therefore it was hypothesised that the C-terminus sequences will dominate the negative polarity spectra. Indeed, deprotonated C-terminus ions: y, z-H and x ions are observed in negative polarity spectra ([Supplementary Tables 148,150,152](#)¹⁷³). Deprotonated N-terminus fragments: a, b and c are also present ([Supplementary Tables 149, 151](#)¹⁷³). These types of ions are consistent with negative polarity ions reported in previous studies for methods such as CID¹⁸¹ or ECD.¹⁷⁴

Further analysis of the sequences assigned in the chymotrypsin sample aimed to improve understanding of which amino acids guide formation of different types of ions. In agreement with previous work, due to the unique structure of proline (P), specifically its side chain being bound to the peptide backbone, N-terminus sequences fragmented at proline result in low intensity or lack of a, b or c ions (Figure 3.11), at the same time promoting high intensity C-terminus peaks, such as y ions (Figure 3.12)^{182,183}. This effect is observed in both positive and negative polarity 3D OrbiSIMS spectra.

Table 3.3. Peak list exported from SurfaceLab negative mode spectrum of bovine α -chymotrypsin, consisting of ions detected in the spectrum and assigned fragments of C-terminal sequence TALVNWVQQTLAN. The colour corresponds to the presence of the observed sequence in human insulin presented in 3.10.

Description	y			z-H			x		
	m/z	Assignment	Dev. (ppm)	m/z	Assignment	Dev. (ppm)	m/z	Assignment	Dev. (ppm)
AN	x			187.0726	C ₇ H ₁₁ N ₂ O ₄ ⁻	0.7264			
AAN	x			258.1098	C ₁₀ H ₁₆ N ₃ O ₅ ⁻	1.0986			
LAAN	x			371.1935	C ₁₆ H ₂₇ N ₄ O ₆ ⁻	-0.2595			
TLAAN	x			472.2414	C ₂₀ H ₃₄ N ₅ O ₈ ⁻	0.2003			
QTLAAN	x			600.3005	C ₂₅ H ₄₂ N ₇ O ₁₀ ⁻	1.1395			
QQTLAN	743.3710	C ₃₀ H ₅₁ N ₁₀ O ₁₂ ⁻	2.2270	728.3595	C ₃₀ H ₅₀ N ₉ O ₁₂ ⁻	1.4595	771.3654	C ₃₁ H ₅₁ N ₁₀ O ₁₃ ⁻	1.5380
VQQTLAN	842.4389	C ₃₅ H ₆₀ N ₁₁ O ₁₃ ⁻	1.3952	827.4279	C ₃₅ H ₅₉ N ₁₀ O ₁₃ ⁻	1.2157	870.4337	C ₃₆ H ₆₀ N ₁₁ O ₁₄ ⁻	1.1739
WVQQTLAN	1028.5179	C ₄₆ H ₇₀ N ₁₃ O ₁₄ ⁻	0.7758	1013.5070	C ₄₆ H ₆₉ N ₁₂ O ₁₄ ⁻	0.7728	1056.5129	C ₄₇ H ₇₀ N ₁₃ O ₁₅ ⁻	0.8869
NWVQQTLAN	x			1127.5506	C ₅₀ H ₇₅ N ₁₄ O ₁₆ ⁻	1.3082	1170.5565	C ₅₁ H ₇₆ N ₁₅ O ₁₇ ⁻	1.3377
VNVVQQTLAN	1241.6290	C ₅₅ H ₈₅ N ₁₆ O ₁₇ ⁻	0.4496	1226.6187	C ₅₅ H ₈₄ N ₁₅ O ₁₇ ⁻	0.9778	1269.6252	C ₅₆ H ₈₅ N ₁₆ O ₁₈ ⁻	1.4429
LVNVVQQTLAN	1354.7145	C ₆₁ H ₉₆ N ₁₇ O ₁₈ ⁻	1.5284	1339.7028	C ₆₁ H ₉₅ N ₁₆ O ₁₈ ⁻	0.8937	1382.7075	C ₆₂ H ₉₆ N ₁₇ O ₁₉ ⁻	0.0865
ALNVVQQTLAN	1425.7515	C ₆₄ H ₁₀₁ N ₁₈ O ₁₉ ⁻	1.3644	1410.7400	C ₆₄ H ₁₀₀ N ₁₇ O ₁₉ ⁻	0.9025	1453.7453	C ₆₅ H ₁₀₁ N ₁₈ O ₂₀ ⁻	0.5211
TALNVVQQTLAN	1526.7994	C ₆₈ H ₁₀₈ N ₁₉ O ₂₁ ⁻	1.3900	1511.7875	C ₆₈ H ₁₀₇ N ₁₈ O ₂₁ ⁻	0.7344			

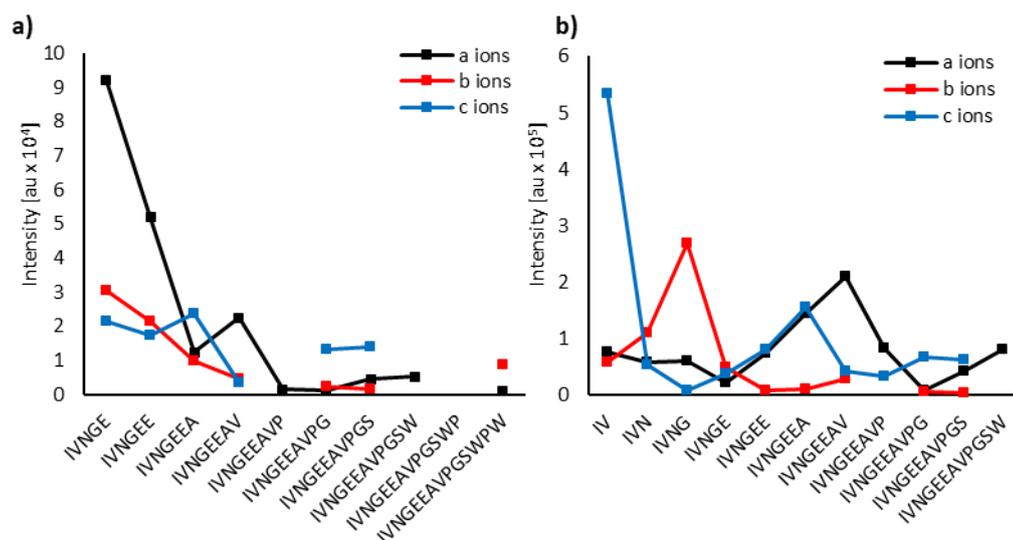


Figure 3.11. Peak intensity depending on ion type and sequence composition, N-terminus sequence in a) positive polarity and b) negative polarity. Sequences terminated by proline are of lower intensity or absent in both positive and negative polarities.

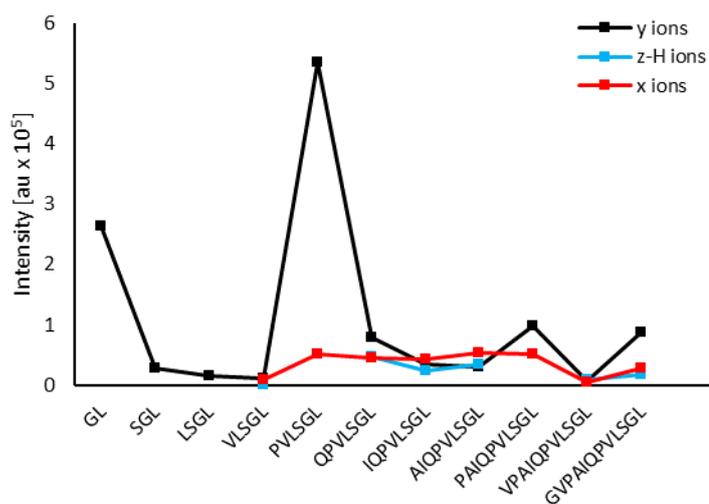


Figure 3.12. Peak intensity depending on ion type and sequence composition, negative polarity, C-terminus sequence. y ions of sequences containing proline have higher intensity than sequences not terminated with proline.

Due to the complexity of the 3D OrbiSIMS spectra of proteins, specifically the existence of different types of ions, presence of sodium adducts and sequences derived from chain termini as well as the middle of the chains, no other useful rules

on peak intensity in relation to different amino acid sequences have been found so far in this work. Statistical analysis of large dataset may be needed to uncover the mechanism behind which sequences preferentially appear in the spectra.

Analysis of the spectra of similar structures, such as the same proteins from different species, revealed that both common and distinct fragments can be identified in the samples. Bovine serum albumin (BSA) and human serum albumin (HSA) have 76% identical sequences (441 of 584 amino acids) calculated by protein BLAST¹⁸⁴. Their spectra contain 32% (1178 of 3709 peaks) common peaks. Common sequences and species-specific sequences were detected in the GCIB/OrbitrapTM spectra and are presented in Figure 3.13. This finding demonstrates the ability of the 3D OrbiSIMS to provide sufficiently detailed information to distinguish two spectra coming from similar proteins, which was not possible in SIMS analysis previously, as described in Section 1.2.

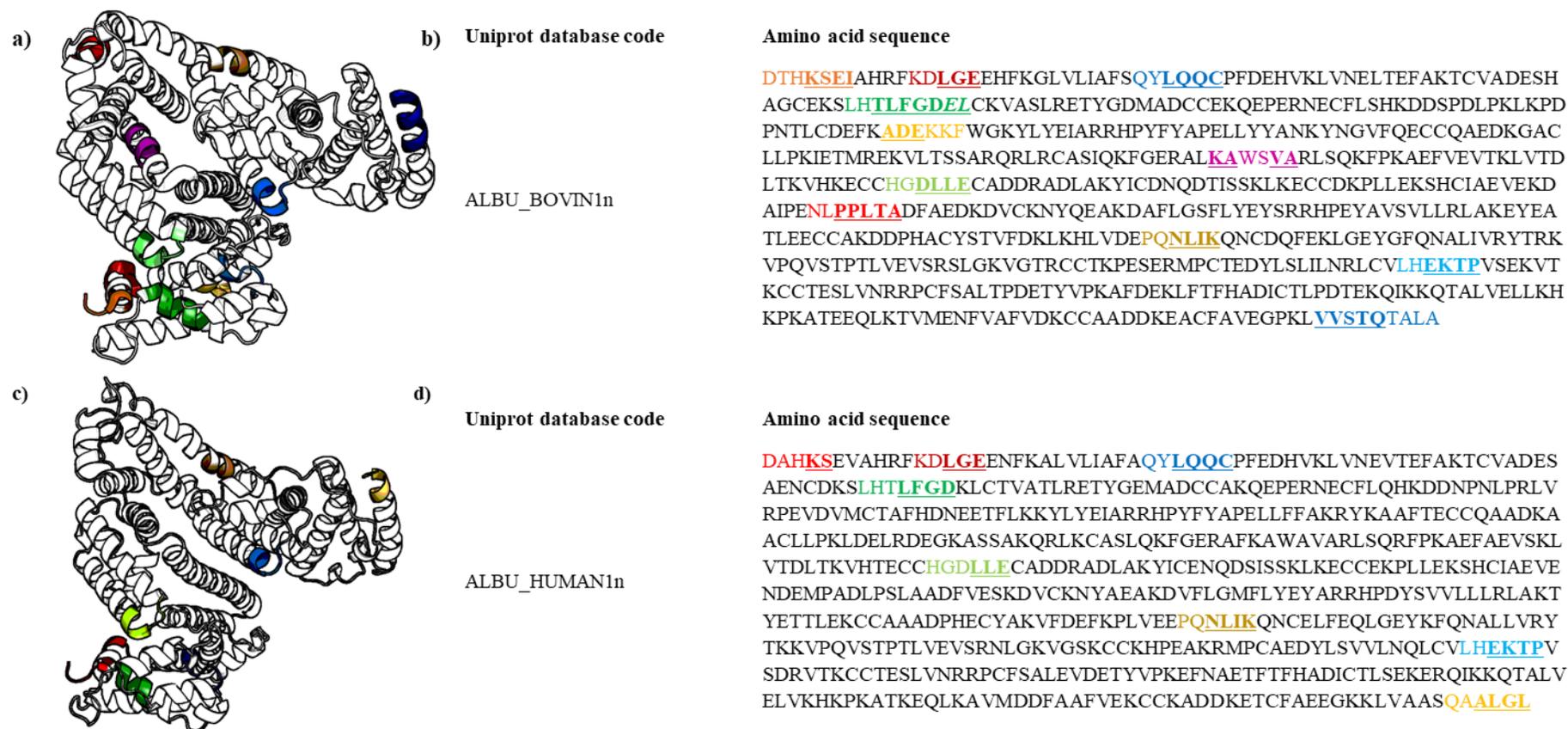


Figure 3.13. Comparison of structures and amino acid sequences of bovine serum albumin and human serum albumin.

3.3.4 Lateral resolution and sensitivity of the ballistic sequencing

Previously imaging of proteins in SIMS was limited to mapping amino acid fragments in positive polarity²⁶ or CN^- and CNO^- peaks in negative polarity^{27,185}. The ability to image the proteins with high chemical specificity and relatively high lateral resolution using the GCIB/OrbitrapTM ballistic sequencing was studied here. A patterned protein surface was created by masking a thin (300 nm) lysozyme film using a transmission electron microscopy grid. The images obtained with a focussed 5 μm diameter Argon₃₀₀₀⁺ primary beam and OrbitrapTM analyser achieved lateral resolution of 10 μm using the sum of 27 of the characteristic peptidic fragments of lysozyme as illustrated in Figure 3.14.

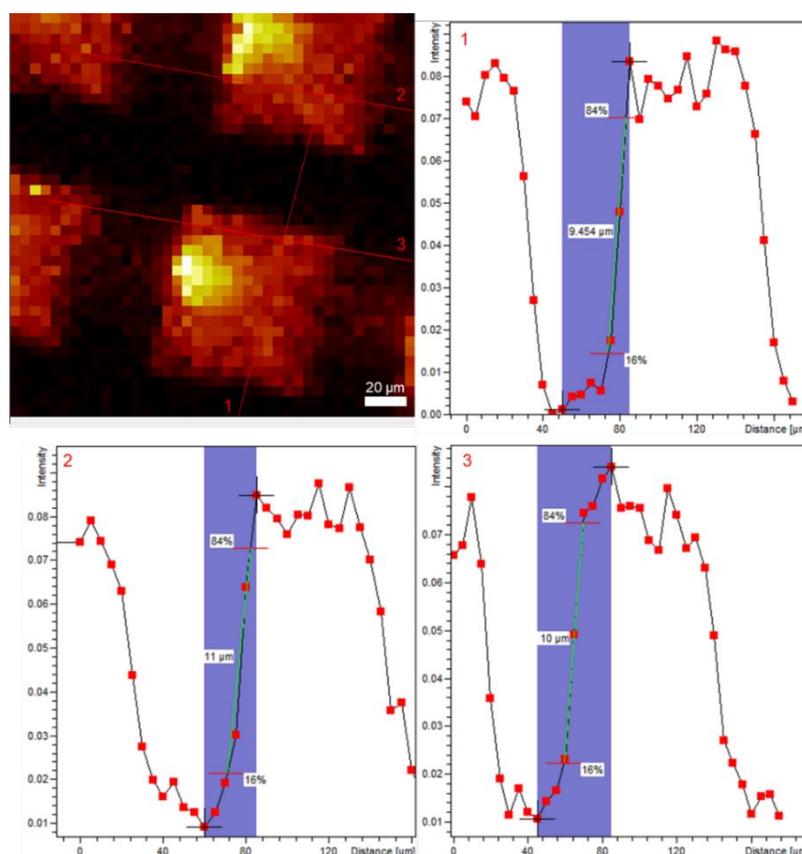


Figure 3.14. 3D OrbiSIMS image of a protein film under a transmission electron microscopy grid, obtained with a focussed 5 μm diameter Argon₃₀₀₀⁺ primary beam.

The image shows a sum of lysozyme peaks KVFG c, KVFG+Na c, KVFGRC a5, KVFGRC a5-NH3, KVFGRC b5, KVFGRC b5-NH3, KVFGRC b-CN3H4, KVFGRC c, KVFGRC a, KVFGRC b, KVFGRC-S a, KVFGRC-SH2 a, KVFGRC a, KVFGRC b, RL y, RL z+1, RL z-1, RG yb, IRG yb, WIRG yb, WIRG ya-NH3, AWIRG yb, AWIRGCRL z-1, VQAWIRG yc, DVQAWIRG yb, NAWV+Na yc, FNTQ+Na a, normalised to total ion count. Line scans 1, 2 and 3 demonstrate the ability to obtain a high chemical specificity image with lateral resolution of approximately 10 μm . Mean of three (n=3) measurements was 10.15 μm and the standard deviation was $\text{SD} = 0.78 \mu\text{m}$.

In the reference lysozyme film samples, the amount of material removed during sputtering is assumed to consist of pure protein. The amount of protein molecules can be calculated by equation $n = \frac{\rho A d N_A}{M_W}$, where ρ is the density of the protein (1.37 g/mL)¹⁸⁶, A is the analysed area, $200 \times 200 \mu\text{m}$ ($40000 \mu\text{m}^2$), d is the depth of material consumed during the analysis, estimated by the SurfaceLab software and confirmed by profilometry (Table 2.1). As calculated, the number of lysozyme molecules that enabled the protein fragment assignment in the analysed sample was 6.9×10^{11} (1 picomole), which demonstrates high sensitivity of the method.

Profilometry measurement of a crater formed during GCIB/OrbitrapTM analysis of lysozyme film allows for calculation of the depth resolution of the instrument with the chosen settings. In model protein samples, with 300 nm total depth after 30 scans, depth per scan was calculated as 10 nm. This demonstrates unique capability of the 3D OrbiSIMS to map proteins in three dimensional samples with relatively high spatial resolution.

3.4 Conclusions

This work demonstrates the capability of the 3D OrbiSIMS to achieve the first matrix- and label-free *in situ* assignment of undigested proteins (up to 272 kDa), which can move forward a range of research areas investigating protein surfaces. A bottom-up *de novo* sequencing approach allowed sequence coverage from 5 to 53% in 16 analysed proteins. High chemical specificity of the method was demonstrated by assignment of common and distinct peaks in the spectra of similar proteins, bovine serum albumin and human serum albumin.

Similarities and differences between the 3D OrbiSIMS analysis and most common proteomic analysis methods were assessed in order to aid the protein identification by GCIB/OrbitrapTM ballistic sequencing. The GCIB induced ballistic sequencing produced a, b, c and a-NH₃ ions at the N-terminus of the amino acid sequence, y, z and x ions at the C-terminus and ya, yb, yc and ya-NH₃ ions from sequences in the middle of the protein chain. Analogous peak types were present in both positive and negative spectra. The general types of ions have been described, however the mechanism of ion formation mechanisms to enable high-throughput identification is subject for extended work.

Chemical structures of selected fragments were suggested and the assignments were confirmed using the MS/MS capability of the instrument. This method allows characterisation of undigested proteins with unprecedented lateral resolution of 10 μm and high sensitivity, as 1 picomole of lysozyme was sufficient for assignment of peptidic sequences.

4 Application of ballistic sequencing for *in situ* analysis of biological samples

4.1 Chapter aims

The results presented in Chapter 3 described the development of ballistic sequencing using the 3D OrbiSIMS on 16 example model proteins. This method allows, for the first time, matrix- label- and digestion-free analysis of proteins at surfaces with 10 μm resolution. It is proposed that this method can be beneficial to a range of biological applications in medicine and medical device development. The work described in this chapter aims to extend the ballistic sequencing beyond model samples and assess the method's capability to investigate challenging biological systems.

The ballistic sequencing approach will be applied to simple mixtures (two-component) and complex human serum which contains different chemical components (proteins, lipids, salts). The ability to characterise these systems using the 3D OrbiSIMS data will help to confirm the extent to which this approach can be applied beyond model systems.

In order to assess the depth profiling and imaging capability of this approach as applied to real samples, human skin and a biochip will be analysed. These represent extreme examples of highly challenging samples owing to the highly biologically complex and varied nature of skin tissue and the monolayer nature of the skin and biochip respectively. Characterisation of skin protein profile is relevant to

understanding skin diseases and their treatment, while detection of characteristic protein fragments in a biochip sample could aid the development of selective biosensors and medicines (Section 1.3).

Investigation of active sites of proteins is crucial for proteomic analyses and may provide insight into protein identity and action. Active sites are responsible for the protein's function by forming chemical bonds with external substrates and their highly specific chemical structure could be diagnostic and aid identification. One of the aims of this chapter is to investigate whether sites such as ligand, proton or cell binding sites can be detected in protein spectra by ballistic sequencing. Additionally, assignment of small molecule characteristic features within the protein structure such as heme will be explored.

4.2 Methods

4.2.1 3D OrbiSIMS measurement of human skin

20 keV Ar_{3000}^+ analysis beam of 20 μm diameter, was used as primary ion beam. Ar_{3000}^+ with duty cycle set to 15%, GCIB current was 230 pA. The Q Exactive depth profile was run on the area of $200 \times 200 \mu\text{m}$ using random raster mode with crater size $280 \times 280 \mu\text{m}$. The cycle time was set to 200 μs . Optimal target potential was approximately -384.4 V. The acquisition was run in temperature -170°C. The depth profile was collected in negative polarity, in m/z range 75 - 1125. The injection time was set to 500 ms. The measurement lasted 18307 scans, the total ion dose per measurement was 2.65×10^{13} ions/cm². Mass resolving power was set to 240,000 at $m/z = 200$.

4.2.2 Preparation of the protein biochip

For the preparation of protein monolayer biochips, the gold substrates were cleaned by immersion in piranha solution (70% H₂SO₄, 30% H₂O₂) at room temperature for 8 minutes, rinsed with MiliQ water and dried with an argon flow. (Caution: Piranha solution reacts violently with organic solvents and should be handled with care). The clean gold substrates were immersed for 24 h in 50 μM DMF solutions of heptakis-(6-deoxy-6-thio)-β-cyclodextrin (TCD) for preparation of a self-assembled monolayer (SAM). Subsequently, the gold substrates were rinsed with DMF and MiliQ water and dried under an argon flow. The TCD SAMs were immersed in a 0.05 mM lysozyme in PBS solution for 2 h. Following protein immobilization, the samples were washed with PBS buffer followed by submersion in MiliQ water for 1 min. The samples were then dried under argon and placed in the instrument for the analysis immediately.

4.2.3 Measurement of layer thickness

The self-assembled monolayer and the self-assembled monolayer with immobilised lysozyme were measured using an Axis-Ultra XPS instrument (Kratos Analytical, UK) with monochromated Al K α X-ray source, based at the National Physical Laboratory (NPL), Teddington, UK. Wide scans were acquired for analysis of organic layer thickness on the gold slide with step size 1000 meV, pass energy 160 eV, dwell time 300 ms, in a range 1300 to -10 eV.

The thickness of the self-assembled monolayers and the self-assembled monolayer with immobilised lysozyme was measured using J.A. Woollam Co. Inc. alpha-SETM

spectroscopic ellipsometer. The CompleteEASE software was employed to determine the thickness values and the calculations were based on a two-phase organic/Au model, in which the organic layer was assumed to be isotropic and assigned a refractive index of 1.50. The thickness reported is the average of three different measurements on SAM or SAM and protein samples, with the errors reported as standard deviation.

4.3 Results and discussion

4.3.1 Detection of distinct protein fragments in a simple mixture

3D OrbiSIMS can enable *in situ* analysis of proteins at surfaces by ballistic sequencing and technical aspects of this method such as ion annotation, sensitivity and lateral resolution are described using 16 model protein samples in Chapter 3. To test the limits of practical applications of the 3D OrbiSIMS analysis of proteins, an equimolar mixture of lysozyme and insulin was imaged using the Ar_{3000}^+ primary beam focused to 20 μm . As presented in Figure 4.1, sequences from both lysozyme and insulin can be assigned in the 3D OrbiSIMS spectrum, allowing the unambiguous identification of both proteins in the mixture.

Segments of amino acids originating from insulin (light and dark blue) and lysozyme (light and dark red) are presented in the spectrum (Figure 4.1a). Segments of lysozyme and insulin observed in the spectrum are highlighted blue and red in the respective protein structure schematics (Figure 4.1b-c). An overlay of insulin, mixture and lysozyme samples was magnified to demonstrate lysozyme (d) and insulin (e) specific fragments. Peaks assigned as lysozyme sequence NAWVAW

(dark red) are detected in the mixture spectrum (black) and are absent in the insulin spectrum (blue) (d). Peaks assigned as insulin sequence FVNQHLC (blue) are detected in the mixture spectrum (black) and are absent in the lysozyme spectrum (red) (e).

In order to confirm that the lysozyme : insulin sample is a homogenous mixture and both proteins are detected simultaneously in the same area, images of assigned ions were generated. Fragment ions originating from both lysozyme (FN, FNT, FNTQ, FNTQA, NTQA, GI, GIL, GILQ, KVF, KVFG, KVFGRC, KVFGRC,NA, NAW, NAWV, NAWVA, NAWVAWR, RG) and insulin (FVN, FVNQ, FVNQH, FVNQHL, FVNQHLC, FVNQHLCG, FVNQHLCGS, FVNQHLCGSH, GI, GIV, GIVEQ, GIVEQC) are distributed across the whole analysis area and are detected simultaneously from a mixture, as demonstrated in Figure 4.2.

The intensity of the sum of insulin ions is 10 times higher than the intensity of the sum of lysozyme ions (Figure 4.2). This result may be due to the fact that smaller proteins (insulin at 6 kDa) and peptides offer better availability of the chain ends for fragmentation. Additionally, during sample preparation, lysozyme is likely to be localised closer to the bottom of the sample while insulin is closer to sample surface due to the larger size of lysozyme and increased number of possible electrostatic interactions and hydrogen bonds with the surface^{187,188}. To account for this effect, the analysis should be carried out until the whole sample depth is sputtered.

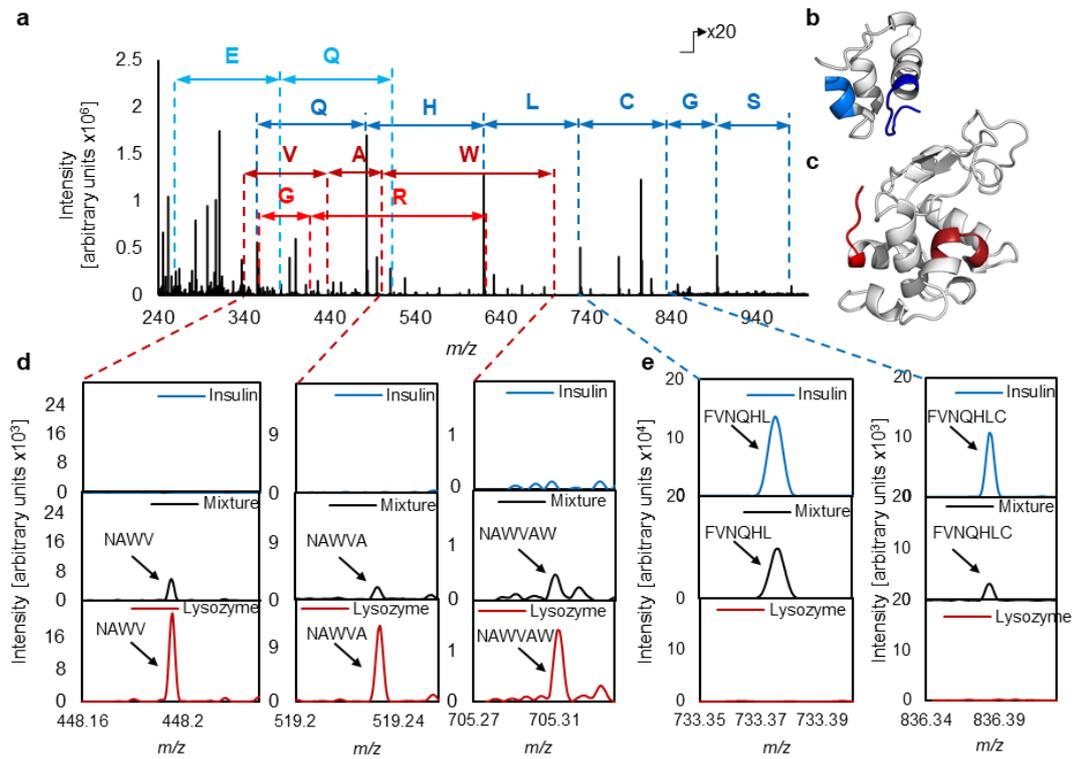


Figure 4.1. Spectrum of a thin film (300 nm) of 1:1 lysozyme: insulin mixture. Segments of amino acids originating from insulin (blue) and lysozyme (red) are presented in the spectrum (a) and highlighted blue in the insulin cartoon (b) and red in the lysozyme cartoon (c). An overlay of insulin, mixture and lysozyme samples was magnified to demonstrate lysozyme (d) and insulin (e) related fragments.

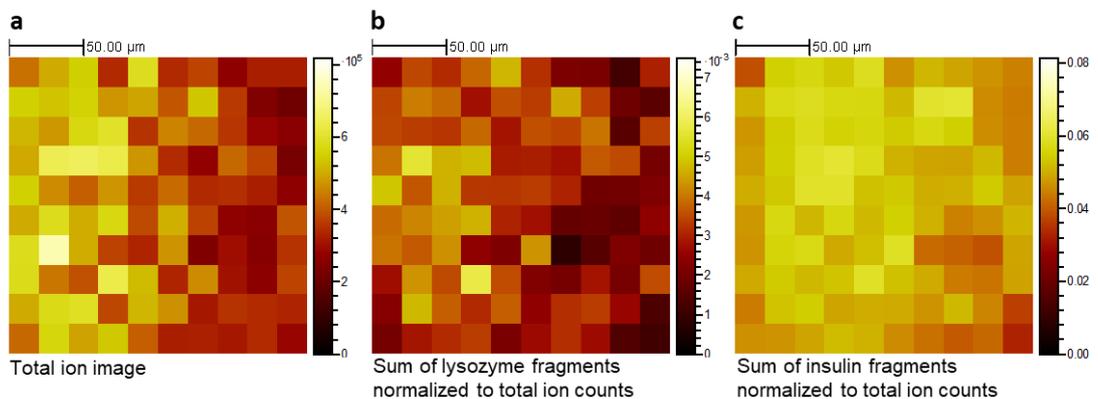


Figure 4.2. GCIB Orbitrap images of the model mixture. The maps represent the total ion image (a), sum of lysozyme fragments, normalized to total (b) and sum of insulin fragments, normalized to total (c).

Detection and simultaneous assignment of fragment ions originating from both lysozyme and insulin in this binary mixture suggests that the ballistic sequencing can

enable analysis of proteins in complex samples with high chemical specificity, which was not possible in SIMS previously. This finding suggests the possibility of applying the 3D OrbiSIMS approach to mixed systems such as biological samples such as materials interacting with cells or tissue sections.

4.3.2 Characterisation of human serum by 3D OrbiSIMS

This section aims to apply the 3D OrbiSIMS for the first time to analysis of proteins in a complex, but well characterised biological sample, human serum. Serum is a fraction of blood remaining after removal of blood cells and clotting factors. It is composed of proteins, lipoprotein particles, electrolytes, antibodies, antigens and hormones. Most abundant serum proteins and their concentrations are available in the Protein Atlas¹⁸⁹. Three abundant serum proteins, albumin, transferrin and insulin, were manually assigned in single protein model spectra in Section 3.2. These assignments were applied to the 3D OrbiSIMS spectrum of human serum sample, enabling identification of characteristic sequences from these proteins, as presented in Figure 4.3. Intense high mass ($m/z > 600$) ions have been assigned as other serum components, lipids, by searching for sodiated or protonated ions in the LipidMaps® In-Silico Structure Database (link: <https://www.lipidmaps.org/resources/databases/lmissd/search.php>). Example di- and triglycerols have been highlighted in the 3D OrbiSIMS spectrum of human serum shown in Figure 4.3.

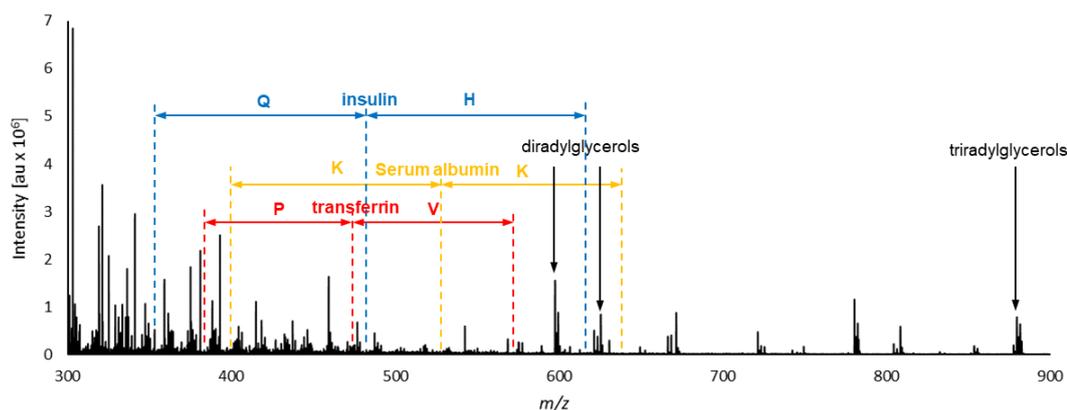


Figure 4.3. 3D OrbiSIMS spectrum of human serum. Sequences of abundant serum proteins, insulin, albumin and transferrin and di- and triradylglycerols were manually assigned based on previous model protein samples and LipidMaps®.

Manual assignment of sequences in the 3D OrbiSIMS spectra allows only targeted search for known species. In order to detect proteins in an untargeted way, SIMS Molecular Formula Prediction tool (SIMS-MFP) and SIMSdenovo¹⁷³ were employed. Work described in this chapter contributed to the development of SIMS-MFP software, which will be available on Github and described by Edney *et al.* (manuscript in preparation). Briefly, the software generates possible elemental composition assignments based on accurate masses of selected elements. The elemental composition can be confirmed by double bond equivalent, a value describing the level of unsaturation of a molecule, calculated from the amounts of hydrogen, nitrogen and carbon in the molecule. SIMSdenovo was developed as a part of the work described in Chapter 3¹⁷³ and is available on Github (<https://github.com/guferraz/simsdenovo/>). This software detects amino acid residues and sequences in the spectrum, followed by database matching of found sequences and protein identification.

Combined use of both SIMS-MFP and SIMSdenovo enables automatic elemental and amino acid assignment of characteristic fragments of known proteins in the 3D OrbiSIMS spectrum of human serum. Example assignments of a and a-NH₃ type

ions of amino acid sequences of albumin detected in human serum are presented in

Table 4.1.

Table 4.1. Assignments of C-terminus and internal sequences of human serum albumin, assigned automatically by molecular formula prediction and SIMSdenovo.

Description	a				a-NH3			
	m/z	Assignment	Dev. (ppm)	Intensity	m/z	Assignment	Dev. (ppm)	Intensity
QA	194.0902	C ₇ H ₁₃ N ₃ O ₂ Na ⁺	1.05	385964	177.0635	C ₇ H ₁₀ N ₂ O ₂ Na ⁺	0.29	285218
QAA	265.127	C ₁₀ H ₁₈ N ₄ O ₃ Na ⁺	-0.42	436399	248.1005	C ₁₀ H ₁₅ N ₃ O ₃ Na ⁺	-0.25	1190002
QAAL	378.2111	C ₁₆ H ₂₉ N ₅ O ₄ Na ⁺	-0.20	396449	361.1846	C ₁₆ H ₂₆ N ₄ O ₄ Na ⁺	-0.07	827740
QAALG	435.2324	C ₁₈ H ₃₂ N ₆ O ₅ Na ⁺	-0.55	191754	418.2059	C ₁₈ H ₂₉ N ₅ O ₅ Na ⁺	-0.45	670771
QAALGL	548.3166	C ₂₄ H ₄₃ N ₇ O ₆ Na ⁺	-0.19	51031	531.2899	C ₂₄ H ₄₀ N ₆ O ₆ Na ⁺	-0.48	96204
AAS	224.1005	C ₈ H ₁₅ N ₃ O ₃ Na ⁺	-0.28	31795	207.0741	C ₈ H ₁₂ N ₂ O ₃ Na ⁺	0.42	45786
AASQ	352.1591	C ₁₃ H ₂₃ N ₅ O ₅ Na ⁺	-0.11	26119	335.1326	C ₁₃ H ₂₀ N ₄ O ₅ Na ⁺	0.03	61195
AASQA	423.1962	C ₁₆ H ₂₈ N ₆ O ₆ Na ⁺	-0.13	12343	406.1695	C ₁₆ H ₂₅ N ₅ O ₆ Na ⁺	-0.50	37858
					477.2068	C ₁₉ H ₃₀ N ₆ O ₇ Na ⁺	-0.04	19870
					590.2909	C ₂₅ H ₄₁ N ₇ O ₈ Na ⁺	0.03	15154
SQA	281.122	C ₁₀ H ₁₈ N ₄ O ₄ Na ⁺	-0.09	25662	264.0954	C ₁₀ H ₁₅ N ₃ O ₄ Na ⁺	-0.29	81403
SQAA	352.1591	C ₁₃ H ₂₃ N ₅ O ₅ Na ⁺	-0.11	26119	335.1326	C ₁₃ H ₂₀ N ₄ O ₅ Na ⁺	0.03	61195
SQAAL	465.2429	C ₁₉ H ₃₄ N ₆ O ₆ Na ⁺	-0.65	26477	448.2165	C ₁₉ H ₃₁ N ₅ O ₆ Na ⁺	-0.35	87638
SQAALG					505.238	C ₂₁ H ₃₄ N ₆ O ₇ Na ⁺	-0.24	42735
SQAALGL	635.3499	C ₂₇ H ₄₈ N ₈ O ₈ Na ⁺	1.84	8793	618.3218	C ₂₇ H ₄₅ N ₇ O ₈ Na ⁺	-0.62	13877
NLG	279.1427	C ₁₁ H ₂₀ N ₄ O ₃ Na ⁺	-0.22	341594	262.1161	C ₁₁ H ₁₇ N ₃ O ₃ Na ⁺	-0.43	1169208
NLGK	407.2376	C ₁₇ H ₃₂ N ₆ O ₄ Na ⁺	-0.30	13181	390.2111	C ₁₇ H ₂₉ N ₅ O ₄ Na ⁺	-0.19	525061
NLGKV	506.3059	C ₂₂ H ₄₁ N ₇ O ₅ Na ⁺	-0.47	7986	489.2794	C ₂₂ H ₃₈ N ₆ O ₅ Na ⁺	-0.39	136384
NLGKVG	563.3278	C ₂₄ H ₄₄ N ₈ O ₆ Na ⁺	0.35	2937	546.3009	C ₂₄ H ₄₁ N ₇ O ₆ Na ⁺	-0.28	63813
NLGKVGS					633.3332	C ₂₇ H ₄₆ N ₈ O ₈ Na ⁺	0.19	18536
VAA	236.1369	C ₁₀ H ₁₉ N ₃ O ₂ Na ⁺	-0.20	433040	219.1103	C ₁₀ H ₁₆ N ₂ O ₂ Na ⁺	-0.45	403335
VAAS	323.169	C ₁₃ H ₂₄ N ₄ O ₄ Na ⁺	0.07	139369	306.1423	C ₁₃ H ₂₁ N ₃ O ₄ Na ⁺	-0.42	122913
VAASQ	451.2274	C ₁₈ H ₃₂ N ₆ O ₆ Na ⁺	-0.34	36487	434.2008	C ₁₈ H ₂₉ N ₅ O ₆ Na ⁺	-0.47	83855
VAASQA	522.2677	C ₂₁ H ₃₇ N ₇ O ₇ Na ⁺	-0.23	27082	505.238	C ₂₁ H ₃₄ N ₆ O ₇ Na ⁺	-0.24	42735.
VAASQAA	593.3047	C ₂₄ H ₄₂ N ₈ O ₈ Na ⁺	0.87	15197	576.2752	C ₂₄ H ₃₉ N ₇ O ₈ Na ⁺	-0.06	17364
DLP	320.1581	C ₁₄ H ₂₃ N ₃ O ₄ Na ⁺	0.07	143126				
DLPS	407.1897	C ₁₇ H ₂₈ N ₄ O ₆ Na ⁺	-1.00	13172				
DLPSL	520.2736	C ₂₃ H ₃₉ N ₅ O ₇ Na ⁺	-1.10	4999				
DLPSLA	591.3109	C ₂₆ H ₄₄ N ₆ O ₈ Na ⁺	-0.65	5614				
DLPSLAA	662.3491	C ₂₉ H ₄₉ N ₇ O ₉ Na ⁺	-1.23	4557				

Importantly, three of the assigned sequences form 10-membered amino acid sequence at the C-terminus of albumin, as highlighted in the amino acid sequence of human serum albumin:

DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCFEDHVKLVNEVTEFAKTCVADESAEN
CDKSLHTLFGDKLCTVATLRETYGEMADCCAQEPERNECFHQHDDNPNLPRLVRPEVDV
MCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAACLLPKLDE
LRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTEC
CHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA**DLPSLAA**
DFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHE
CYAKVFDEFKPLVEEPQNLIKQNCLEFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSR**NL**
GKVGSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCFSAL
EVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA
FVEKCKADDKETCFAEEGKKL**VAASQAALGL**

As described in Section 3.2, amino acid sequences generated by ballistic sequencing may originate from N- or C-terminus of the protein chain or from the middle of the protein chain. The fact that automatic assignment results in terminus sequences is valuable, because protein termini often determine protein functions and contain important information about the protein's identity. It was shown that over 85% of all proteins contain a unique N-terminus sequence^{173,190,191}. Several protein identification strategies are focusing on the assignment of this part of the sequence^{178,192} and implementing this fact in the SIMSdenovo software could improve the confidence in untargeted identification of proteins in the 3D OrbiSIMS. Additionally, since the protein termini contain motifs responsible for protein action, they are often a target of engineered modifications for biological studies or drug development¹⁹³ and observing these motifs in the spectra could help assess the results of structure manipulation.

SIMS-MFP and SIMSdenovo were used to achieve untargeted assignment and identification of proteins in human serum, the top 47 highest scoring proteins are

listed in Table 4.2. Proteins marked green are present in the human serum according to Protein Atlas, however the first 6 highest scoring proteins (white) are false-positive results, which should not be detected in human serum. This demonstrates that untargeted protein identification by our dedicated software is not yet achieved and implementation of full elemental assignment as well as merging of overlapping sequences based on the premise of Basic Local Alignment Search Tool (BLAST)¹⁹⁴ are currently being developed to improve the reliability of the identifying function of SIMSdenovo.

Table 4.2. Highest-scoring proteins detected automatically in human serum sample using SIMS-MFP and SIMSdenovo.

Number	Protein	Uniprot code	Score	Number	Protein	Uniprot code	Score
1	MUC19_HUMAN	1n	9990	24	TENX_HUMAN	1n	2890
2	OBSCN_HUMAN	1n	5126	25	SSPO_HUMAN	1n	2873
3	KMT2D_HUMAN	1n	4874	26	PRR12_HUMAN	1n	2861
4	CO7A1_HUMAN	1n	4452	27	CO5A3_HUMAN	1n	2859
5	AHMK2_HUMAN	1n	4363	28	CO1A1_HUMAN	1n	2846
6	MUC12_HUMAN	1n	4102	29	COBA2_HUMAN	1n	2836
7	HMCN2_HUMAN	1n	3988	30	CO3A1_HUMAN	1n	2815
8	ZN469_HUMAN	1n	3936	31	CO4A6_HUMAN	1n	2801
9	BSN_HUMAN	1n	3737	32	CO4A4_HUMAN	1n	2799
10	MUC5B_HUMAN	1n	3734	33	TNC18_HUMAN	1n	2796
11	NACAM_HUMAN	1n	3706	34	SYNE1_HUMAN	1n	2762
12	EPIPL_HUMAN	1n	3665	35	CO5A1_HUMAN	1n	2733
13	PKD1_HUMAN	1n	3505	36	CO4A3_HUMAN	1n	2721
14	AHMK_HUMAN	1n	3466	37	AGRV1_HUMAN	2n	2716
15	MUC5A_HUMAN	1n	3391	38	TACC2_HUMAN	1n	2711
16	FCGBP_HUMAN	1n	3333	39	CO4A1_HUMAN	1n	2701
17	PCD16_HUMAN	1n	3325	40	HERC2_HUMAN	1n	2689
18	APOA_HUMAN	1n	3297	41	SHAN1_HUMAN	1n	2672
19	PCLO_HUMAN	1n	3265	42	PGCA_HUMAN	1n	2653
20	CO4A5_HUMAN	1n	3257	43	COGA1_HUMAN	1n	2649
21	SRCAP_HUMAN	1n	3201	44	CO1A2_HUMAN	1n	2644
22	PGBM_HUMAN	1n	2928	45	EP400_HUMAN	1n	2644
23	AGRV1_HUMAN	1n	2912	46	CO2A1_HUMAN	1n	2620
				47	COMA1_HUMAN	1n	2617

SIMS-MFP search with elemental composition C₄₋₁₀₀, H₄₋₂₂₀, N₀₋₂, O₀₋₅₀, S₀₋₁, P₀₋₅, Na₁₋₁ was used to screen peaks with possible lipid composition in the human serum. Then the LIPID MAPS® database of bulk lipid species¹⁹⁵ was coded to automatically match formulae in the list from MFP. The database groups major classes of lipids based on the molecular formula and indicates general classification of a lipid, but not specific chain positions or double bond regiochemistry and geometry. The formulae suggested by the script were matched with the database, resulting in 226 assigned and classified lipid peaks and the most abundant of these are listed in Table 4.3. Additionally, phosphate salts were assigned in the serum spectrum by searching compounds of composition: C₀, H₁₋₁₀, N₀₋₀, O₀₋₁₀, S₀₋₁, P₀₋₅, Na₀₋₅, Ca₀₋₅, Cl₀₋₅ in the list of peaks unassigned as either lipids or proteins (Table 4.4). These results demonstrate that protein and lipid fragments can be readily assigned in the 3D OrbiSIMS spectra of a biological sample such as human serum using the SIMS-MFP.

Table 4.3. Highest intensity lipid peaks assigned automatically in human serum 3D OrbiSIMS spectrum, using molecular formula prediction and LipidMaps database of bulk lipid assignments.

<i>m/z</i>	Assignment	dev. (ppm)	Lipid group
303.2294	C ₁₈ H ₃₂ O ₂ Na ⁺	-0.17	FA 18:2
459.2481	C ₂₁ H ₄₁ O ₇ PNa ⁺	-0.24	LPA 18:1/LPA O-18:2;O
478.3290	C ₂₉ H ₄₅ NO ₃ Na ⁺	-0.34	NAE 27:7;O
487.2793	C ₂₃ H ₄₅ O ₇ PNa ⁺	-0.43	LPA 20:1/LPA O-20:2;O/PA O-20:1
518.3214	C ₂₄ H ₅₀ NO ₇ PNa ⁺	-0.60	LPC 16:0/LPC O-16:1;O/LPE 19:0/LPE O-19:1;O
542.4906	C ₃₄ H ₆₅ NO ₂ Na ⁺	-0.27	Cer 34:2;O/NAE 32:2
597.4852	C ₃₇ H ₆₆ O ₄ Na ⁺	-0.22	DG O-34:4/MG 34:4/MG O-34:5;O
599.5009	C ₃₇ H ₆₈ O ₄ Na ⁺	-0.13	DG O-34:3/MG 34:3/MG O-34:4;O
621.4850	C ₃₉ H ₆₆ O ₄ Na ⁺	-0.53	CE 12:1;O ₂ /DG O-36:6
623.5007	C ₃₉ H ₆₈ O ₄ Na ⁺	-0.45	CE 12:0;O ₂ /DG O-36:5
625.5163	C ₃₉ H ₇₀ O ₄ Na ⁺	-0.53	DG O-36:4
649.5164	C ₄₁ H ₇₀ O ₄ Na ⁺	-0.35	CE 14:1;O ₂ /DG O-38:6
652.6001	C ₄₂ H ₇₉ NO ₂ Na ⁺	-0.31	Cer 42:3;O

666.4832	C ₃₆ H ₇₀ NO ₆ PNa ⁺	-0.14	CerP 36:2;O ₂ /LPC O-28:3/LPE O-31:3
671.5737	C ₄₅ H ₇₆ O ₂ Na ⁺	-0.07	CE 18:2
721.4777	C ₃₉ H ₇₁ N ₀ O ₈ PNa ⁺	-0.24	LPG O-33:5/PA 36:3/PA O-36:4;O
722.4811	C ₃₈ H ₆₉ NO ₁₀ Na ⁺	-0.37	HexCer 32:3;O ₄
723.4933	C ₃₉ H ₇₃ O ₈ PNa ⁺	-0.31	LPG O-33:4/PA 36:2/PA O-36:3;O
725.5567	C ₃₉ H ₇₉ N ₂ O ₆ PNa ⁺	-0.13	PE-Cer 37:1;O ₂ /SM 34:1;O ₂
749.5087	C ₄₁ H ₇₅ O ₈ PNa ⁺	-0.64	PA 38:3/PA O-38:4;O
780.5511	C ₄₂ H ₈₀ NO ₈ PNa ⁺	-0.35	CerP 42:3;O ₄ /LPC 34:3;O/PC 34:2/PC O-34:3;O/PE 37:2/PE O-37:3;O
782.5669	C ₄₂ H ₈₂ NO ₈ PNa ⁺	-0.16	CerP 42:2;O ₄ /LPC 34:2;O/PC 34:1/PC O-34:2;O/PE 37:1/PE O-37:2;O
804.5517	C ₄₄ H ₈₀ NO ₈ PNa ⁺	0.40	CerP 44:5;O ₄ /PC 36:4/PC O-36:5;O/PE 39:4/PE O-39:5;O
806.5670	C ₄₄ H ₈₂ NO ₈ PNa ⁺	-0.03	CerP 44:4;O ₄ /PC 36:3/PC O-36:4;O/PE 39:3/PE O-39:4;O
808.5826	C ₄₄ H ₈₄ NO ₈ PNa ⁺	-0.09	CerP 44:3;O ₄ /PC 36:2/PC O-36:3;O/PE 39:2/PE O-39:3;O
853.7250	C ₅₃ H ₉₈ O ₆ Na ⁺	-0.66	DG 50:3;O/DG O-50:4;O ₂ /TG 50:2/TG O-50:3;O
855.7407	C ₅₃ H ₁₀₀ O ₆ Na ⁺	-0.60	DG 50:2;O/DG O-50:3;O ₂ /TG 50:1/TG O-50:2;O
877.7251	C ₅₅ H ₉₈ O ₆ Na ⁺	-0.52	CE 28:1;O ₄ /DG 52:5;O/DG O-52:6;O ₂ /TG 52:4/TG O-52:5;O
879.7408	C ₅₅ H ₁₀₀ O ₆ Na ⁺	-0.47	CE 28:0;O ₄ /DG 52:4;O/DG O-52:5;O ₂ /TG 52:3/TG O-52:4;O
881.7565	C ₅₅ H ₁₀₂ O ₆ Na ⁺	-0.41	DG 52:3;O/DG O-52:4;O ₂ /TG 52:2/TG O-52:3;O

Table 4.4. Example assignments of the most abundant salt peaks in the 3D OrbiSIMS spectrum of human serum.

<i>m/z</i>	Dev. (ppm)	Intensity	Assignment
158.9131	0.27	17374	PO ₄ HNaCa ⁺
163.9222	0.13	536719	PO ₄ Na ₃ ⁺
164.9301	0.58	35374622	PO ₄ HNa ₃ ⁺
174.8780	-0.31	15440	PO ₄ Ca ₂ ⁺
186.9121	0.81	29309137	PO ₄ Na ₄ ⁺
250.8916	-0.03	6897698	PO ₈ Na ₄ ⁺

4.3.3 Detection of distinct protein fragments in human skin

Assignment of proteins in the 3D OrbiSIMS spectra of human serum sample, described in Section 4.2.2, required knowledge of serum proteome gathered in the Protein Atlas. This sample was a challenging biological and homogenous system

without any characteristic features. In this section, the ballistic sequencing is applied to detect proteins in a depth profile through human skin composed of multiple layers.

Detection of proteins and qualitative and quantitative changes in the protein composition of skin is necessary for fundamental understanding of healthy and pathological function of skin and is of particular interest in the context of topically applied medicines, care products and exposure to chemicals^{196,197}. Conventional analysis of the skin proteome is challenging due to the tissue being rich in lipids, insoluble proteins and proteins cross-linked with sugars, lipids or other proteins. These factors contribute to complicated multi-step isolation and digestion of proteins for conventional mass spectrometry identification¹⁹⁸.

The capability of the 3D OrbiSIMS to identify and locate proteins within complex biological tissue, human skin, *in situ* without complicated sample preparation was assessed by depth profiling in cryogenic conditions. Human skin was placed on an ITO slide without any additional sample preparation and analysed in -170°C. Three examples; corneodesmosin, keratin and collagen, known to be abundant in the *stratum corneum*, epidermal and dermal layers respectively, were detected in the depth profile. Sequences characteristic to these proteins were assigned *in situ* in the skin layers where they are known to be present exclusively or predominantly. The layers of the skin in the depth profile, specifically the *stratum corneum*, underlying epidermis and dermis, were assigned based on the abundance of a phospholipid marker (PO_3^-), as described in previous analysis of skin in SIMS by Starr *et al.*¹⁹⁹ and are presented in Figure 4.4. Protein locations in the skin layers were based on the immunochemically stained tissue sections gathered in the Protein Atlas.¹⁸⁹

Hydroxyproline is an indicator of collagen, as it forms 13.5% of the collagen sequence and is rarely found in other proteins²⁰⁰. The intensity of hydroxyproline, methoxyproline and fragments of hydroxyproline, as well as the PGE sequence all increased with the depth of analysis, reaching a maximum in the dermis region (Figure 4.5, Table 4.5). Similarly, the characteristic sequence QHGSG was assigned to corneodesmosin, present in the *stratum corneum* (Figure 4.6, Table 4.6) and SFGGGG sequence assigned to keratin, abundant throughout the *stratum corneum* and the epidermis (Figure 4.7, Table 4.7). The method, together with the knowledge of skin composition, allowed targeted assignment of corneodesmosin, keratin and collagen *in situ*, which indicates the capability to profile different proteins in complex biological systems.

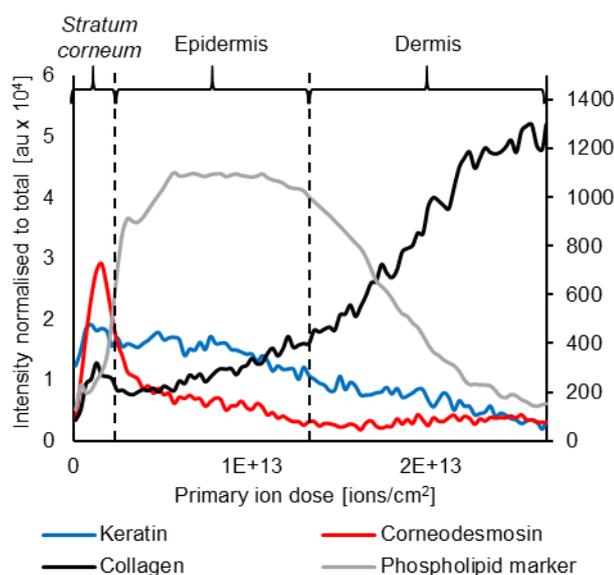


Figure 4.4. 3D OrbiSIMS depth profile overlay of example ions of three proteins; keratin (*stratum corneum* and *epidermis*, blue), corneodesmosin (*stratum corneum*, red) and collagen (*dermis*, black) in a human skin sample. The dashed lines indicate borders between the skin layers, assigned based on the profile of phospholipid marker (PO_3^-)¹⁹⁹.

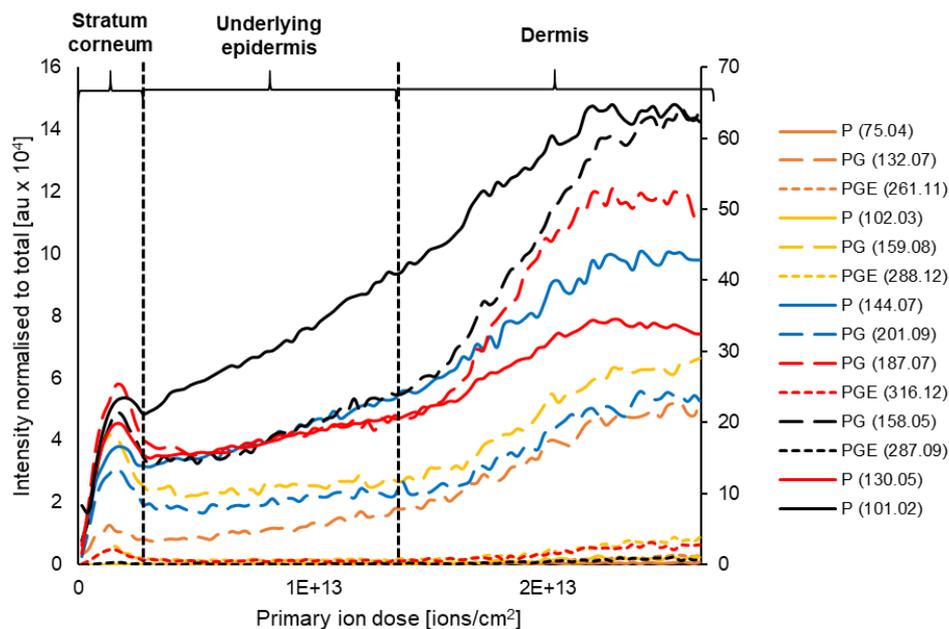


Figure 4.5. 3D OrbiSIMS depth profile overlay of example ions representing collagen. The sequence PGE, frequently occurring in the collagen sequence, is profiled through the skin with hydroxyproline (130.05 m/z), metoxyproline (144.07 m/z) or hydroxyproline fragments (75.04, 101.02, 102.03 m/z) as starting points. Each colour represents one sequence and all ions are assigned in Table 4.5.

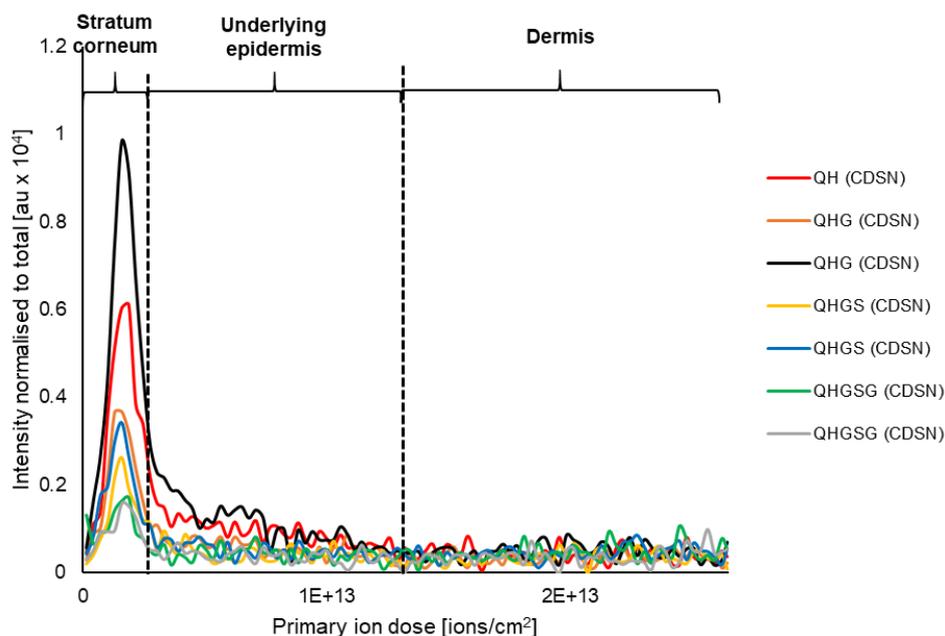


Figure 4.6. 3D OrbiSIMS depth profile overlay of example ions representing QHGS sequence of cormenodesmosin (CDSN), abundant predominantly in the stratum corneum. All presented ions are assigned in Table 4.5.

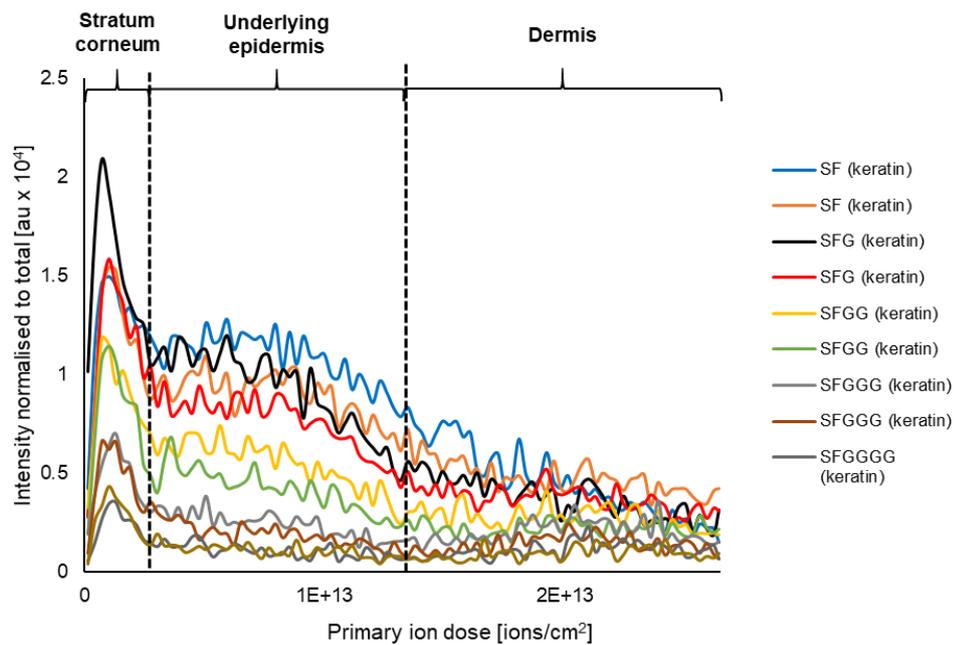


Figure 4.7. 3D OrbiSIMS depth profile overlay of example ions representing keratin sequence SFGGGG, detected throughout the stratum corneum and the underlying epidermis. All ions are assigned in Table 4.7.

Table 4.5. Peak list exported from SurfaceLab negative mode depth profile through human skin, consisting of ions detected in the spectrum and assigned as sequence PGE of collagen. The *m/z* values represent the experimentally observed center mass of each peak. The deviation (*dev.*) represents the parts per million (*ppm*) accuracy of the assignment. The colours represent the labelling in the depth distribution of the assigned peaks in Figure 4.5.

Description	P			PG			PGE		
	<i>m/z</i>	Assignment	Dev. (ppm)	<i>m/z</i>	Assignment	Dev. (ppm)	<i>m/z</i>	Assignment	Dev. (ppm)
Metoxyproline	144.0665	C ₆ H ₁₀ NO ₃ ⁻	-1.0039	201.0880	C ₈ H ₁₃ N ₂ O ₄ ⁻	-0.3301	330.1306	C ₁₃ H ₂₀ N ₃ O ₇ ⁻	-0.1054
Hydroxyproline fragment	75.0448	C ₃ H ₇ O ₂ ⁻	-5.2346	132.0665	C ₅ H ₁₀ NO ₃ ⁻	-0.8348	261.1094	C ₁₀ H ₁₇ N ₂ O ₆ ⁻	0.8199
Hydroxyproline fragment	101.0243	C ₄ H ₅ O ₃ ⁻	-0.8169	158.0457	C ₆ H ₈ NO ₄ ⁻	-0.9927	287.0886	C ₁₁ H ₁₅ N ₂ O ₇ ⁻	0.4788
Hydroxyproline fragment	102.0322	C ₄ H ₆ O ₃ ⁻	-0.5861	159.0774	C ₆ H ₁₁ N ₂ O ₃ ⁻	-0.9402	288.1202	C ₁₁ H ₁₈ N ₃ O ₆ ⁻	0.4506
Hydroxyproline	130.0508	C ₅ H ₈ NO ₃ ⁻	-1.0131	187.0723	C ₇ H ₁₁ N ₂ O ₄ ⁻	-0.6061	316.1148	C ₁₂ H ₁₈ N ₃ O ₇ ⁻	-0.5758

Table 4.6. Peak list exported from SurfaceLab negative mode depth profile through human skin, consisting of ions detected in the spectrum and assigned as sequence QHGS of corneodesmosin. The *m/z* values represent the experimentally observed center mass of each peak. The deviation (*dev.*) represents the parts per million (*ppm*) accuracy of the assignment. The depth distribution of the assigned peaks is presented in Figure 4.6.

Description	a			b			c			a-NH3		
	<i>m/z</i>	Assignment	Dev. (ppm)									
QH	236.1153	C ₁₀ H ₁₄ N ₅ O ₂ ⁻	0.0475	264.1104	C ₁₁ H ₁₄ N ₅ O ₃ ⁻	0.5526	281.1369	C ₁₁ H ₁₇ N ₆ O ₃ ⁻	0.4540	219.0886	C ₁₀ H ₁₁ N ₄ O ₂ ⁻	-0.5614
QHG	293.1370	C ₁₂ H ₁₇ N ₆ O ₃ ⁻	0.9411	321.1315	C ₁₃ H ₁₇ N ₆ O ₄ ⁻	-0.5284	338.1585	C ₁₃ H ₂₀ N ₇ O ₄ ⁻	0.8865	276.1102	C ₁₂ H ₁₄ N ₅ O ₃ ⁻	0.0404
QHGS	380.1689	C ₁₅ H ₂₂ N ₇ O ₅ ⁻	0.3898	408.1636	C ₁₆ H ₂₂ N ₇ O ₆ ⁻	-0.2528				363.1422	C ₁₅ H ₁₉ N ₆ O ₅ ⁻	-0.0194
QHGS	437.1901	C ₁₇ H ₂₅ N ₈ O ₆ ⁻	-0.4396	465.1849	C ₁₈ H ₂₅ N ₈ O ₇ ⁻	-0.6714				420.1639	C ₁₇ H ₂₂ N ₇ O ₆ ⁻	0.3438

Table 4.7. Peak list exported from SurfaceLab negative mode depth profile through human skin, consisting of ions detected in the spectrum and assigned as sequence SFGGGG of keratin. The *m/z* values represent the experimentally observed center mass of each peak. The deviation (*dev.*) represents the parts per million (ppm) accuracy of the assignment. The depth distribution of the assigned peaks is presented in Figure 4.7.

Description	a			b			c			a-NH3		
	<i>m/z</i>	Assignment	Dev. (ppm)									
SF	207.1138	C ₁₁ H ₁₅ N ₂ O ₂ ⁻	-0.5626	235.1088	C ₁₂ H ₁₅ N ₂ O ₃ ⁻	-0.0630	252.1354	C ₁₂ H ₁₈ N ₃ O ₃ ⁻	0.3139	190.0873	C ₁₁ H ₁₂ NO ₂ ⁻	-0.4510
SFG	264.1355	C ₁₃ H ₁₈ N ₃ O ₃ ⁻	0.6372	292.1304	C ₁₄ H ₁₈ N ₃ O ₄ ⁻	0.5157	309.1568	C ₁₄ H ₂₁ N ₄ O ₄ ⁻	-0.0093	247.1088	C ₁₃ H ₁₅ N ₂ O ₃ ⁻	0.1317
SFGG	321.1566	C ₁₅ H ₂₁ N ₄ O ₄ ⁻	-0.6429	349.1515	C ₁₆ H ₂₁ N ₄ O ₅ ⁻	-0.6948	366.1780	C ₁₆ H ₂₄ N ₅ O ₅ ⁻	-0.7797	304.1302	C ₁₅ H ₁₈ N ₃ O ₄ ⁻	-0.1081
SFGGG	378.1781	C ₁₇ H ₂₄ N ₅ O ₅ ⁻	-0.5307	406.1728	C ₁₈ H ₂₄ N ₅ O ₆ ⁻	-0.8831	423.1997	C ₁₈ H ₂₇ N ₆ O ₆ ⁻	-0.0624	361.1517	C ₁₇ H ₂₁ N ₄ O ₅ ⁻	-0.2364
SFGGGG	435.1995	C ₁₉ H ₂₇ N ₆ O ₆ ⁻	-0.5562	463.1944	C ₂₀ H ₂₇ N ₆ O ₇ ⁻	-0.4827	480.2210	C ₂₀ H ₃₀ N ₇ O ₇ ⁻	-0.4193	418.1728	C ₁₉ H ₂₄ N ₅ O ₆ ⁻	-0.9422

4.3.4 Detection of protein fragments in a protein biochip

The ballistic sequencing method developed in Section 3.2 allows the analysis of proteins at surfaces without any pre-treatment which is particularly advantageous for investigating protein distribution on materials when the amount of the protein on the surface is in pico- or femtomolar range. The efficacy of the 3D OrbiSIMS in these challenging biological applications was assessed by analysing a biochip produced using a method developed by Di Palma *et al.*¹⁶⁵ consisting of a lysozyme monolayer immobilised on a self-assembled monolayer (SAM). For the preparation of protein biochips, the gold substrates were cleaned by immersion in piranha solution, rinsed with MiliQ water and dried with an argon flow. The clean gold substrates were functionalised with heptakis-(6-deoxy-6-thio)- β -cyclodextrin (TCD). Subsequently, the gold substrates were rinsed with DMF and MiliQ water and dried under an argon flow. The TCD SAMs were immersed in a lysozyme in PBS solution and following protein immobilization, the samples were washed with PBS buffer followed by submersion in MiliQ water for 1 min. The samples were then dried under argon and placed in the instrument for the analysis immediately. The samples were confirmed to be a monolayer of proteins by ellipsometry and X-ray photoelectron spectroscopy (XPS) using calculations developed by Shard²⁰¹, as illustrated in Figure 4.8.

The amount of protein molecules in the analysed volume is calculated by equation =

$\frac{A}{A_{TCD}}$, where A is the analysed area, $200 \times 200 \mu\text{m}$ ($40000 \mu\text{m}^2$). A_{TCD} is the area of one molecule of thiol β -cyclodextrin (TCD) forming the self assembled monolayer. A_{TCD} is 1.83 nm^2 , as the diameter of one TCD molecule is 1.53 nm .

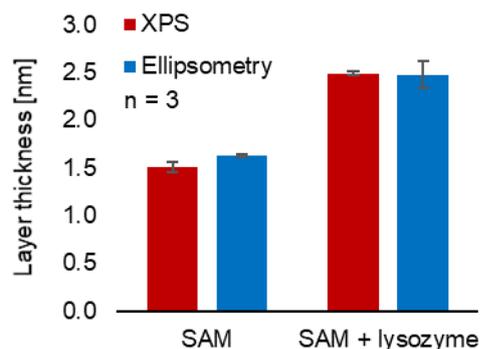


Figure 4.8. Layer thickness measurements of the self assembled monolayer (SAM) and SAM with immobilised lysozyme, derived from XPS results (red) and measured by ellipsometry (blue). The bars represent mean of three measurements on one sample. The error bars represent the standard deviation between three different areas at the surface of one sample.

The actual amount of protein molecules is smaller than the amount of cyclodextrin molecules and is determined by the size of the protein. Therefore, the maximum amount of lysozyme molecules in the analysed area is 2.18×10^{10} (40 femtomoles). Seven diagnostic fragment ions from the N- and C- terminus of lysozyme were detected in the samples confirming lysozyme immobilisation on the biochip, as presented in Table 4.8. This illustrates that the sensitivity of the analytical technique is sufficient to detect and assign monolayer protein coverage.

Table 4.8. Lysozyme fragments visible in the spectra obtained from a protein monolayer sample. Protein monolayer was obtained by immobilisation of the protein on a self assembled monolayer (SAM) on a gold slide. Three areas on a sample were analysed and a SAM sample without the protein was analysed as a control. The Supplementary Table presents average intensity (arbitrary units) of selected ions across the three measurements per sample.

<i>m/z</i>	Assignment	Description	SAM + lysozyme		SAM only		Bare gold	
			Run 1	Run 2	Run 1	Run 2	Run 1	Run 2
214.1295	C ₈ H ₁₆ N ₅ O ₂ ⁺	RG	46277	59307	217	0	0	0
271.1765	C ₁₂ H ₂₃ N ₄ O ₃ ⁺	RL	3421	1859	0	0	0	0
449.2864	C ₂₂ H ₃₇ N ₆ O ₄ ⁺	KVFG	2058	3253	0	24	0	0
560.3665	C ₂₇ H ₄₆ N ₉ O ₄ ⁺	KVFGRC	652	1236	0	0	0	0
588.3612	C ₂₈ H ₄₆ N ₉ O ₅ ⁺	KVFGRC	9752	6227	0	0	0	0
605.3875	C ₂₈ H ₄₉ N ₁₀ O ₅ ⁺	KVFGRC	19300	14136	0	33	0	0
631.4030	C ₃₀ H ₅₁ N ₁₀ O ₅ ⁺	KVFGRC-S	6798	3267	0	0	0	0

4.3.5 Detection of functional and active sites of proteins

Section 4.2.2 discussed the importance of amino acid sequences coming from protein chain termini for identification of proteins due to these termini coding information about protein function. However, active sites can be found in various locations in the protein sequence and fragments originating from the middle of the amino acid sequence can also carry information specific to certain proteins. The sequences identified in 16 model protein samples described in Section 3.2 include a number of biologically important binding and active sites located at various locations within the protein chains. Active sites such as heme-, substrate- or metal binding sites were identified based on the details in their UniProt entries and are listed in Table 4.9.

Table 4.9. *Examples of functional sites detected in the 16 analysed proteins. The location of the functional sites is based on UniProt database.*

Functional site	Location in the amino acid sequence	Protein
heme-binding site	H18 and M80	Cytochrome c
NAD-binding site	R99	L-lactate dehydrogenase
	R341	Alcohol dehydrogenase
substrate binding site	T248	L-lactate dehydrogenase
carbonate binding site	T139, T471	Transferrin
	H66	Alcohol dehydrogenase
metal binding site	H604	Transferrin
	H3, E6, D13, H67, D248	Bovine serum albumin
	E8, D10	Concanavalin A
active site	H74	Catalase
fibrin- and heparin binding site	C53-R273, T1813-T2083, C2298-G2429	Fibronectin
collagen binding site	C309-T609	Fibronectin
cell attachment site	V1359-T1632	Fibronectin
charge relay system	C194	Trypsin
	H57, D102, S195	Chymotrypsin

The analysis of the model protein samples resulted in identification of heme, a small molecule (< 700 Da) within the protein, characteristic to a group of enzymes. The heme consists of a porphyrin ring with covalently attached iron. This complex is

present in a range of proteins: haemoglobin, myoglobin, cytochromes, peroxidases and catalases. Figure 4.9 shows an overlay of positive polarity 3D OrbiSIMS spectra of two proteins containing heme: cytochrome c (red) and catalase (blue).

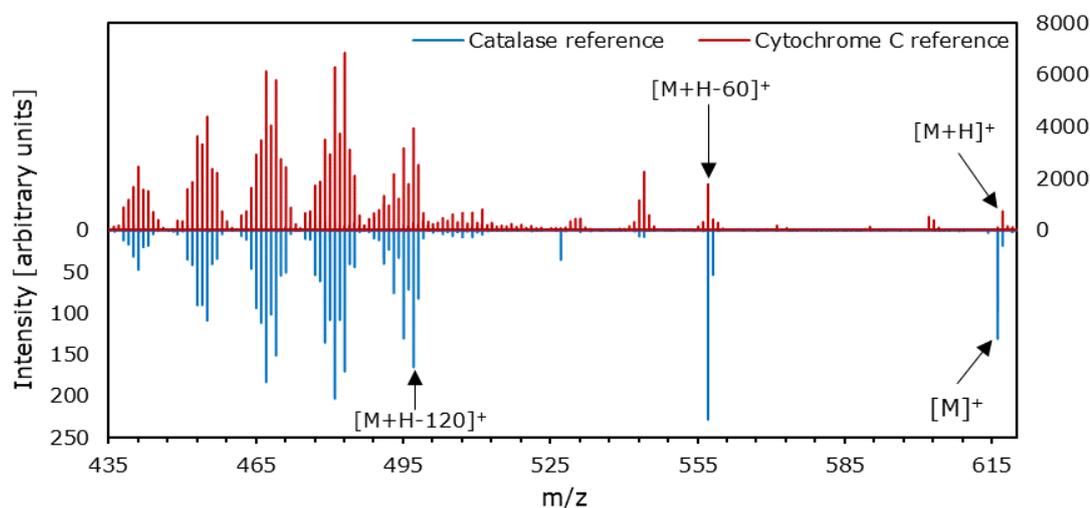


Figure 4.9. Overlay of 3D OrbiSIMS spectra of catalase (red) and cytochrome c (blue).

The spectrum of cytochrome c (Figure 4.9, red) and catalase (Figure 4.9, blue) both show similar peak patterns between m/z 430 - 530. A difference between the spectra is the presence of the heme peak at m/z 616.1 in catalase and at m/z 617.1 in the cytochrome c spectra. In heme, iron is bound to two pyrrole rings via a coordination bond with nitrogen. As a result, the heme containing Fe(II) (ferrous heme) is uncharged and requires protonation in order to be detected. Therefore a $[M+H]^+$ ion is observed at m/z 617.1 in the cytochrome c spectrum (Figure 4.10, blue). The heme containing Fe(III) (ferric heme) carries positive charge without protonation and appears at m/z 616.1 (Figure 4.10, red). This finding is in agreement with the previous studies^{202–204}, catalase heme resting state is Fe(III)²⁰⁵, while heme in cytochrome c prefers to be in reduced state: Fe(II)²⁰⁶. The heme fragments (series of peaks between m/z 430 - 530) can be observed in both spectra. The proposed heme fragmentation is presented in Figure 4.11.

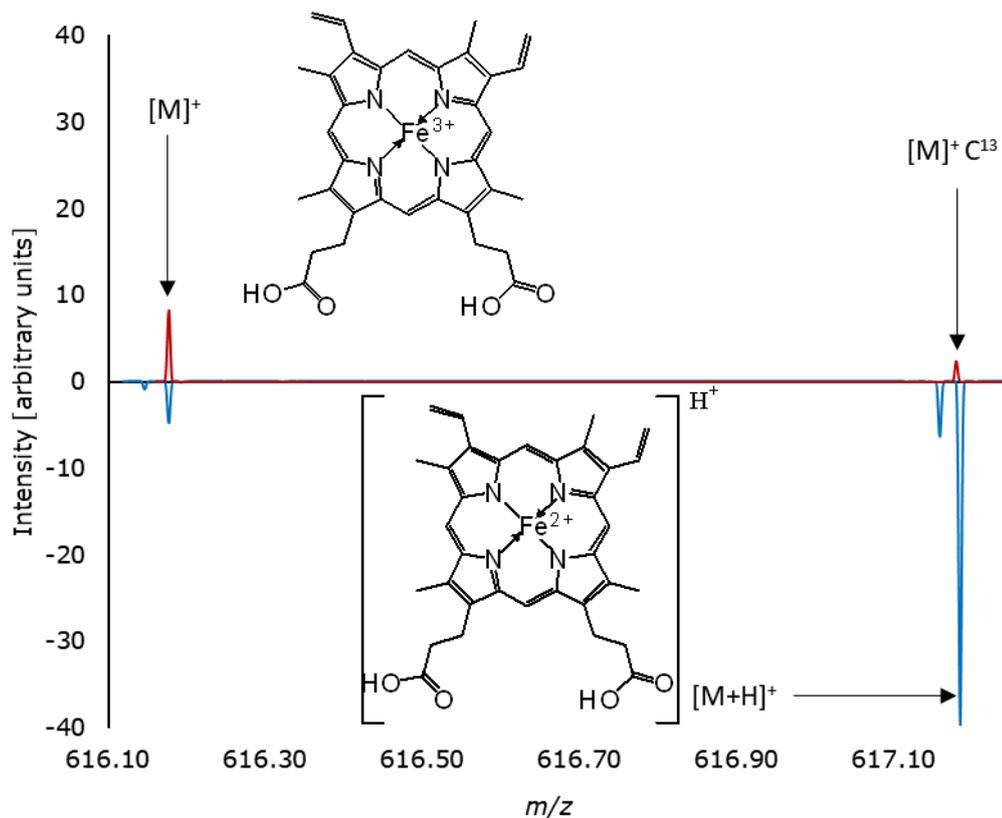


Figure 4.10. Inverted overlay of 3D OrbiSIMS spectra of catalase (red) and cytochrome c (blue) with corresponding structures of ferric (top) and ferrous (bottom) heme.

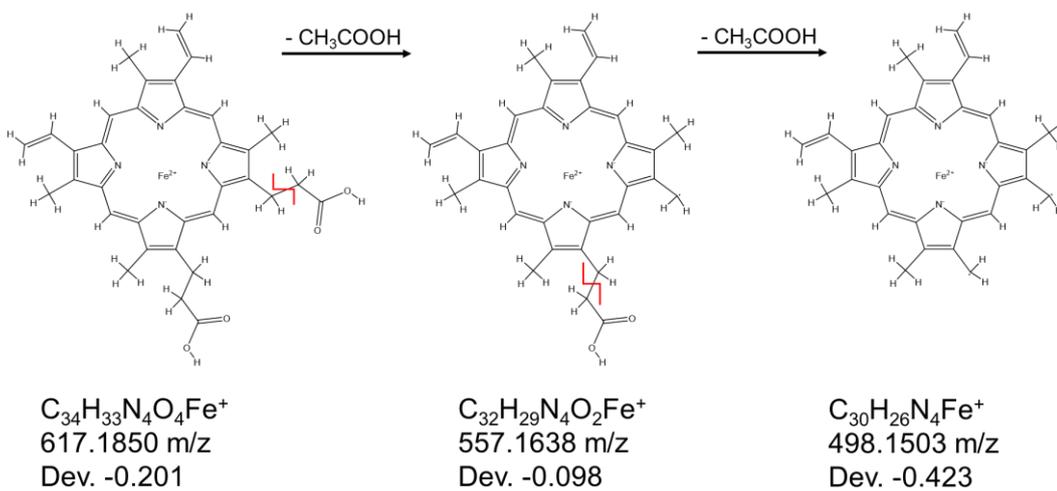


Figure 4.11. Fragmentation of heme using Ar_{3000}^+ primary beam. The original structure loses acetic acid fragment (CH_3COOH , $\Delta = 60$ u) or two acetic acid fragments ($\Delta = 120$ u).

The heme $[M+H]^+$ ion loses acetic acid fragment to yield a peak at m/z 557.1. In further fragmentation second acetic acid is lost. These fragmentation assignments are confirmed by previously reported MALDI measurements followed by tandem MS on heme²⁰⁷⁻²⁰⁹. Detection of the pattern of heme fragment peaks between m/z 430 - 530 has also been reported using ToF-SIMS^{210,211}. This pattern is visible in cytochrome c spectra obtained with Bi_3^+ primary ion beam and ToF analyser, however the peaks are very broad and of low intensity, therefore cannot be identified with confidence.

4.4 Conclusions

This work successfully extends the ballistic sequencing of proteins using the 3D OrbiSIMS to broad real life applications. Abundant serum proteins, albumin, transferrin, fibronectin and insulin were assigned in the 3D OrbiSIMS spectra by comparison with previously acquired model protein spectra. Data collected in this work contributed to the development of dedicated data analysis software, SIMS-MFP and SIMSdenovo, and sequences of the most abundant serum protein, albumin, were automatically assigned in the 3D OrbiSIMS spectrum of human serum by combined use of this approach. The annotated sequences included C-terminus of the protein, which is informative of protein identity and function. This finding is expected to improve untargeted protein identification by SIMSdenovo, a function which currently remains in progress. Other components of serum such as lipids and salts were successfully automatically assigned in the 3D OrbiSIMS spectrum. Importantly, characteristic features such as proton, ligand or cell-binding sites were detected in the analysed proteins, which could improve the confidence in untargeted

protein identification and could be applied to studying protein-ligand interactions. Identification of small molecule (< 700 Da) ligands characteristic to a class of proteins, e.g. heme, can be used as a confirmation of protein identity, as the heme structure generates multiple high intensity peaks in the 3D OrbiSIMS.

The analysis of a biochip revealed that highly specific (5-membered amino acid sequence) protein fragments can be assigned even in the 3D OrbiSIMS spectra a protein monolayer sample, equating to 40 femtomoles of a protein per analysis. This demonstrates the unique capability of ballistic sequencing to characterise protein surfaces, when the amount of analyte is too small to be extracted from the surface for classical analysis such as ESI-MS.

Human skin was analysed to demonstrate detection of proteins in a complex biological system. Proteins known to be abundant within specific layers of human skin were assigned in a targeted way and mapped in the tissue by 3D OrbiSIMS depth profiling. Specifically, collagen was detected predominantly in the dermis, keratin in the stratum corneum and throughout the epidermis and corneodesmosin only in the *stratum corneum*.

The identification of protein active sites and method's sensitivity allowing to detect characteristic fragments from a protein monolayer shows the potential to use ballistic sequencing for biological samples. By the analysis of human skin and a protein biochip, this chapter demonstrated broad applications of the ballistic sequencing in areas crucial to health, medicine and medical devices development.

5 Manufacturing of polymeric micro- and nanoparticles for the investigation of protein corona

5.1 Chapter aims

This chapter aims to obtain monodisperse polymeric particles in a size range from 100 nm to 100 μm for subsequent study of protein adsorption on the particle surface. A combination of three polymers will be used to manipulate the hydrophilic or hydrophobic character of the particles as it has been found that the size and hydrophobicity of nano- and microparticles play an important role in the protein corona formation (Section 1.4.3).

Poly lactic co-glycolic acid (PLGA) with equal proportions of lactic and glycolic acid will be used as a core polymer. Poly vinyl alcohol will be used as a surfactant and a factor increasing surface hydrophilicity. Hydrophobic particles will be produced by exchanging 50:50 PLGA with homopolymeric poly lactic acid (PVA).

The particles will be prepared by microfluidics due to its relatively easy and reproducible manipulation of particle size by altering the flow rates and the capability to produce particles with polydispersity lower than 5%. Additionally, the process can be scaled up and allows the control of the amount of particles produced, which is an important factor in protein corona studies (Section 1.4.2).

5.2 Methods

5.2.1 Micro- and nanoparticle production

10% w/v PLGA solution in DCM was prepared by dissolving 1 g PLGA in 10 mL DCM. For the tuning of the microparticle size, dilutions of 7.5%, 5%, 2.5%, 1%, 0.75%, 0.5% and 0.25% were used. For the production of the nanoparticles, 1%, 0.75%, 0.5% and 0.25% solutions of PLGA in acetone were prepared. 1% w/v PVA solution was prepared by dissolving 2 g of PVA in 200 mL MiliQ water at 90°C for 40 minutes under constant stirring. The polymer solution was placed in a 10 mL glass syringe purchased from Scientific Glass Engineering. Two plastic 10 mL syringes were filled with water (nanoparticle production) or surfactant solution (microparticle production). The syringes were placed on the Harvard Apparatus pumps and attached to microfluidic setup via luer lock fittings. X-junction of the Dolomite 100 μm flow focusing chip was used for the microparticle production. X-junction of the Dolomite 190 μm droplet chip was used for the nanoparticle production. The flow ratio was calculated as ratio between 2 times the continuous phase flow divided by the dispersed phase flow, as the continuous phase is flowing through two channels compared to one channel of the dispersed phase. In case of microparticle production, the droplets were collected into glass vials containing 1 mL 1% w/v PVA solution. The receiving vial was placed on a magnetic plate and a stirring bar was placed inside the vial to enhance solvent evaporation during droplet collection. In case of the nanoparticle production, the nanoemulsion was collected into an empty glass vial.

5.2.2 Particle size measurements

DLS measurements were taken using Zetasizer Nano. Refractive index (RI) of 50:50 PLGA in light scattering measurements calculations was applied as found in literature, with sources reporting RI=1.47-1.61¹⁶² or RI=1.51-1.67¹⁶³.

Laser diffraction measurements were taken with Coulter LS230. Refractive index (RI) of 50:50 PLGA in light scattering measurements calculations was applied as found in literature, with sources reporting RI=1.47-1.61¹⁶² or RI=1.51-1.67¹⁶³.

For scanning electron microscopy (SEM) measurements, the particles suspended in water were dropped onto carbon tape stuck to a SEM stub. The drops were allowed to dry overnight in the fumehood. After drying, the samples were coated with platinum for 120 s. SEM measurements were taken using Jeol-6490. The energy was set to 10 keV. Images were processed in ImageJ in order to measure the average diameter of all particles in the field of view. The intensity threshold was adjusted for each image separately depending on contrast, the pixel to length was set based on the scale bar and the Particle analysis function was used to obtain the diameter of all particles within the threshold. Particle analysis function provides particle area as a result and particle radius was calculated as $r = \sqrt{\frac{A}{\pi}}$.

5.3 Results and discussion

5.3.1 Particle preparation using microfluidics

For the production of micro- and nanoparticles, the polymer solution was placed in a glass syringe and two plastic syringes were filled with water (nanoparticle production) or surfactant solution (microparticle production). The syringes were placed on the syringe pumps and attached to microfluidic setup. X-junction of the 100 μm flow focusing chip was used for the microparticle production. X-junction of the 190 μm droplet chip was used for the nanoparticle production. The droplets were collected into glass vials containing PVA solution (microparticles) or into an empty glass vial (nanoparticles).

Both nano- and microparticles can be prepared using the droplet junction chip (190 μm channel diameter) and the flow focusing chip (100 μm channel diameter) can be used for production of microparticles. The flow focusing chip was not successfully used for production of nanoparticles due to extensive PLGA aggregation in the 100 μm channels. The solubility of PLGA in acetone is lower than in dichloromethane (DCM), therefore the aggregation frequently occurs in the flow focusing chip²¹².

During the production of nanoparticles, the dispersed phase, PLGA dissolved in dichloromethane, becomes focused by the continuous phase, DI water, flowing in the outer two channels of the X-junction, as shown in Figure 5.1.a. During the production of microparticles, droplet formation can be observed at the X-junction, as demonstrated in Figure 5.1b.

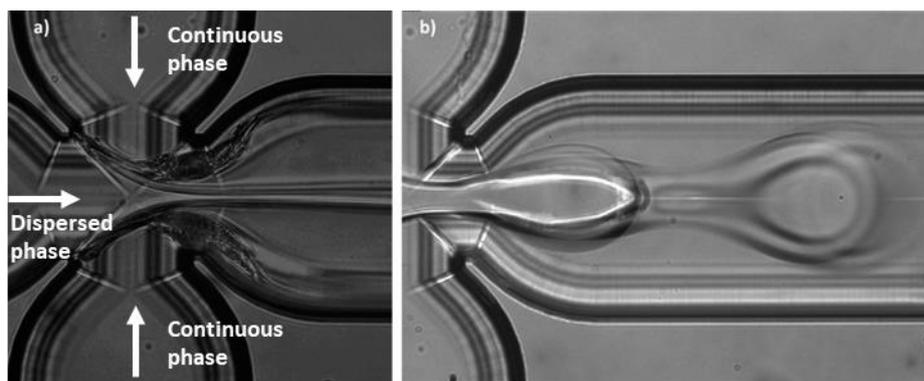


Figure 5.1. Droplet junction chip (190 μm etch depth) for production of **a** nanoparticles by flow focusing and **b** microparticles by droplet formation. The continuous phase (water or water with surfactant) is flowing through two outer channels, focusing the dispersed phase (organic solvent) at the X junction. The ratio of the summed continuous phase flow and the dispersed phase flow is altered to achieve different size particles.

5.3.2 PLGA microparticle size tuning

The diameter of obtained particles was measured by laser diffraction, using refractive index of 50:50 PLGA RI = 1.57, a mean value of RI found in previous studies, with sources reporting RI=1.47-1.61¹⁶² or RI=1.51-1.67¹⁶³. Additionally, scanning electron microscopy (SEM) was used for size and morphology assessment. SEM images were processed in ImageJ in order to measure the average diameter of all particles in the field of view using the particle analysis function.

Monodisperse microparticles in a range of sizes from 5 μm to 20 μm were produced using the flow focusing chip and the size tuning parameters are presented in Table 5.1. In this work the dispersed phase was constant and the continuous phase was altered to investigate the impact of the flow ratio on the particle size. The flow ratio is the ratio between 2 times the continuous phase flow divided by the dispersed phase flow, as the continuous phase is flowing through two channels compared to one channel of the dispersed phase. It has been found that the continuous phase flow rate is the key factor affecting the particle size²¹³. Solutions from 0.25% to 1%

PLGA were used to produce particles in the 100 μm flow focusing chip and the flow ratios $F_{\text{cont}}/F_{\text{disp}}$ 4 - 12 resulted in monodisperse particles of size 4 - 20 μm .

As shown in Table 5.1, the particles produced with higher flow ratios present lower standard deviation than the particles produced with lower flow ratios. This means that the microparticles tend to be more uniform when a higher flow rate of the continuous phase is used, which is with agreement to previous studies by Kovalchuk *et al.*⁹⁷ and Seo *et al.*¹⁰⁰. At lower flow ratios, droplets are formed in a squeezing regime, in which the droplet is larger than the channel width at the junction, due to which the droplets may vary in size. At the optimal flow ratios of 10 $\mu\text{l}/\text{min}$: 1 $\mu\text{l}/\text{min}$ or 12 $\mu\text{l}/\text{min}$: 1 $\mu\text{l}/\text{min}$, the droplets are formed in a dripping or jetting regime and are smaller than the narrowest part of the channel, therefore nothing obstructs droplet formation leading to more uniform particles.

Above the ratio $F_{\text{cont}}/F_{\text{disp}} = 12$, the flow is unstable and tends to break and reverse, leading to a large degree of polydispersity in obtained particle batch, unreliability and lack of reproducibility. Below a flow ratio of 4, a laminar flow is observed and no droplets are produced^{101,214}. According to Donno *et al.*, further modification of particle size can be achieved by using different molecular weight polymer^{215,216}, however the diversity of particles obtained with the 30,000 - 60,000 Da polymer is sufficient for the further work in this project.

The particle size results obtained with laser diffraction (LD) are larger than the diameters observed in the SEM images. This difference in size is believed to be caused by swelling due to hydration when analysing particles in water (LD) as opposed to dried on the surface (SEM). Additionally, the calculation of particle size

from LD results may be inaccurate due to uncertainty regarding the refractive index of the particles.

Table 5.1. Particle size tuning by changing parameters such as PLGA concentration and continuous phase flow rate. Scanning electron microscopy (SEM) and laser diffraction (LD) particle diameter measurement results are presented for each batch. Average diameter and standard deviation were calculated for a minimum of 50 particles in case of SEM and for 3 bulk measurements in case of LD. Flow ratio is the ratio between 2 times the continuous phase flow divided by the dispersed phase flow.

Sample name	PLGA conc. %w/w	Dispersed phase flow rate [$\mu\text{l}/\text{min}$]	Continuous phase flow rate [$\mu\text{l}/\text{min}$]	Flow ratio	Mean diameter [μm]			
					SEM	LD	SEM	LD
AM055	1.00	10	20	4	17.73	± 0.5	23.6	± 1.8
AM056	1.00	10	35	7	14.73	± 0.6	16.4	± 1.2
AM057	1.00	10	50	10	12.04	± 0.4	13.4	± 1.7
AM058	1.00	10	60	12	11.39	± 0.3	8.6	± 1.4
AM059	0.75	10	20	4	13.83	± 0.4	15.3	± 1.9
AM060	0.75	10	35	7	11.13	± 0.5	13.9	± 3.0
AM061	0.75	10	50	10	8.03	± 0.4	10.2	± 0.8
AM062	0.75	10	60	12	7.10	± 0.2	9.3	± 1.7
AM063	0.50	10	20	4	10.95	± 0.7	14.1	± 2.5
AM064	0.50	10	35	7	9.51	± 0.6	10.9	± 0.5
AM065	0.50	10	50	10	7.66	± 0.4	8.5	± 1.4
AM066	0.50	10	60	12	6.65	± 0.2	8.2	± 1.2
AM067	0.25	10	20	4	9.33	± 0.6	9.3	± 2.4
AM068	0.25	10	35	7	8.66	± 0.5	8.8	± 1.4
AM069	0.25	10	50	10	6.91	± 0.4	7.3	± 3.0
AM070	0.25	10	60	12	5.34	± 0.5	5.9	± 0.9

To compare size and morphology of the particles in a hydrated state, additional measurement was carried out using light microscopy as showed in Figure 5.2.. Mean diameter of particles in batch AM057, calculated from light microscopy images was $12 \mu\text{m} \pm 0.5$, which is consistent with SEM results. The particles are smaller than the laser diffraction results suggest and the observed difference between the results is attributed to the inaccuracy of the refractive index value (RI=1.47-1.61¹⁶² or RI=1.51-1.67¹⁶³) and the fact that the particles analysed using SEM are dried. The

average particle diameter between the three techniques for batch AM057 is 12.5 μm with standard deviation of 0.87 μm .

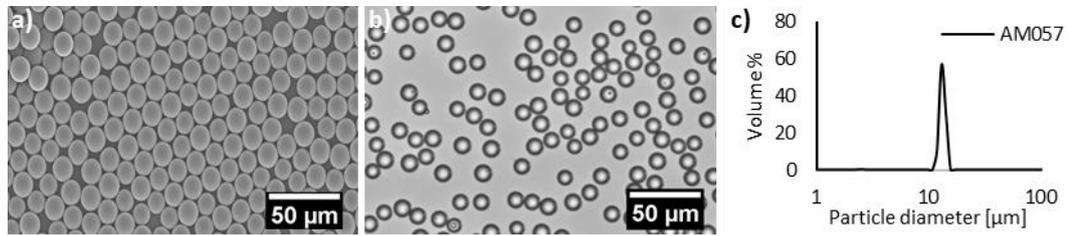


Figure 5.2. Comparison of three particle sizing techniques, SEM (a), light microscopy (b) and laser diffraction(c). Batch AM057 was chosen as an example. The particles have average diameter of $12.5 \mu\text{m} \pm 0.87 \mu\text{m}$ between the three techniques. The circularity of the particles was confirmed by light microscopy and high resolution SEM images

PLGA solutions were produced resulting in particles of 15 - 20, 10 - 15, 6 - 11 and 5 - 9 μm diameter for 1, 0.75, 0.5 and 0.25% PLGA solutions respectively. The relationship between the flow ratio and the particle diameter is linear between flow ratio values of 4 - 12 and can be described by equations presented in Figure 5.3.

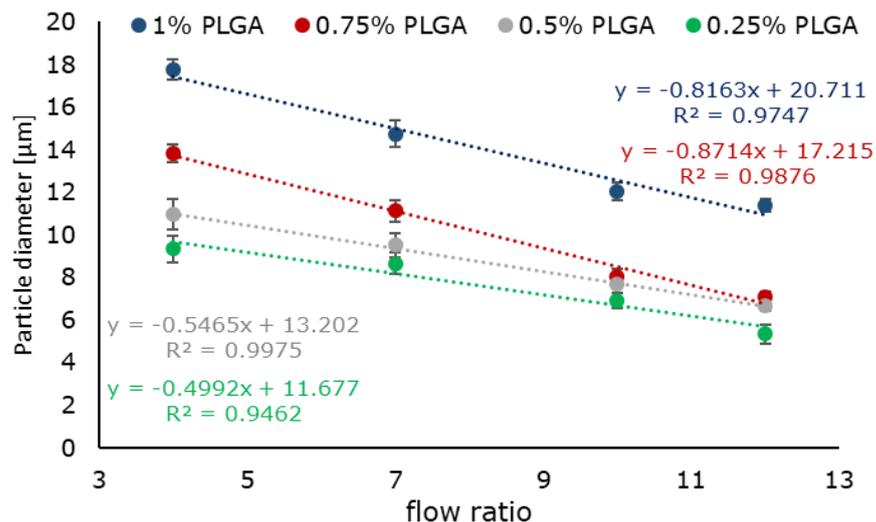


Figure 5.3. Particle size tuning by manipulating polymer concentration and continuous phase flow rate, assessed by SEM. Every data point is an average of 50 measurements in one.

Further changes to the production parameters were tested in order to obtain submicron particles with size increasing in regular increments. The particles of size

approximately 50, 75, 100 and 200 μm have been produced using PLGA concentrations 2 - 5%, however these particles are highly polydisperse (data not shown). Microfluidics allows for production of monodisperse particles, however it requires parameter optimisation (appropriate flow rates, chip size and type) in order to achieve monodispersity. Due to increased density of more concentrated solutions, optimal settings for the production of the 50 - 200 μm particles have not been achieved.

5.3.3 PLA microparticle size tuning

Microparticles made with poly lactic acid (PLA) were produced as an example of particles with a hydrophobic surface. The same chip, 1% PVA surfactant and the same flow parameters were applied as for the PLGA particles discussed in section 5.2.2 and are shown in Table 5.2.

Table 5.2. Particle size tuning by changing parameters such as PLA concentration and continuous phase flow rate. Average size and standard deviation were calculated for a minimum of 50 particles in SEM images.

Sample name	PLA conc. %w/w	Dispersed phase flow rate [$\mu\text{l}/\text{min}$]	Continuous phase flow rate [$\mu\text{l}/\text{min}$]	Flow ratio	Mean diameter [μm]	Standard deviation [μm]
AM089	1.00	10	20	4	21.86	3.3
AM090	1.00	10	35	7	16.85	4.6
AM091	1.00	10	50	10	13.01	3.5
AM092	1.00	10	60	12	12.70	4.1
AM093	0.75	10	20	4	14.40	4.3
AM094	0.75	10	35	7	11.56	3.4
AM095	0.75	10	50	10	9.00	2.9
AM096	0.75	10	60	12	7.95	0.9
AM097	0.50	10	20	4	12.11	4.4
AM098	0.50	10	35	7	9.90	2.5
AM099	0.50	10	50	10	7.77	0.8
AM100	0.50	10	60	12	7.34	1.7

PLA particles from 7 to 22 μm were obtained using the optimised parameters presented in Table 5.2. The polydispersity, measured as a standard deviation from the mean size, was larger than in the case of PLGA particles. However, similarly to PLGA particles, the polydispersity was lower in the particles produced with flow ratios $F_{\text{cont}}/F_{\text{disp}}$ 10 - 12. Similarly to PLGA particles, the relationship between the flow ratio and the particle diameter is linear between flow ratio values of 4 - 12, as shown on Figure 5.4.

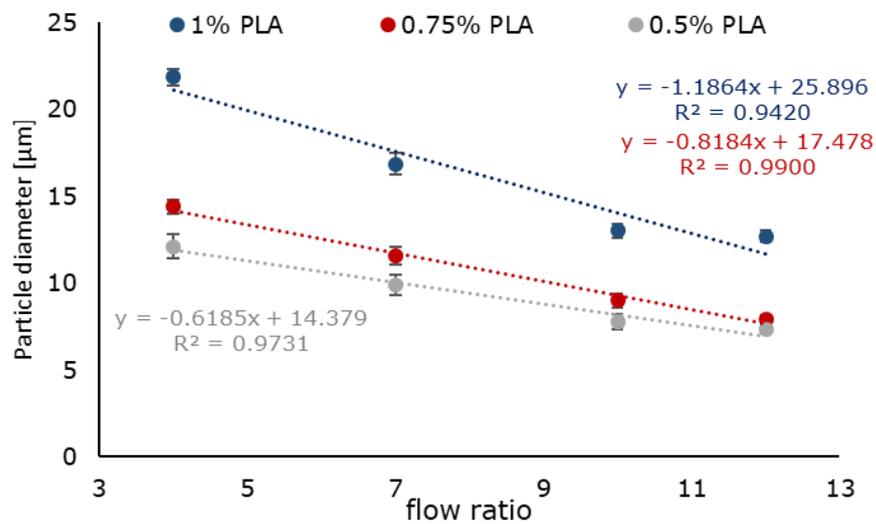


Figure 5.4. Particle size tuning by manipulating polymer concentration and continuous phase flow rate, assessed by SEM. Every data point is an average of 50 measurements in one.

The impact of polymer choice on the particle size on the example of 50:50 PLGA and PLA is presented in Figure 5.5. The average diameter of PLA particles is larger than that of the PLGA particles at each tested flow rate, however the size difference is overall not significant ($P > 0.05$). The polydispersity measured as standard deviation from the mean particle diameter is larger in the PLA samples than in PLGA samples as shown in Figure 5.5. The increased polydispersity may negatively

affect observation of the impact of particle size on the protein adsorption in the further studies.

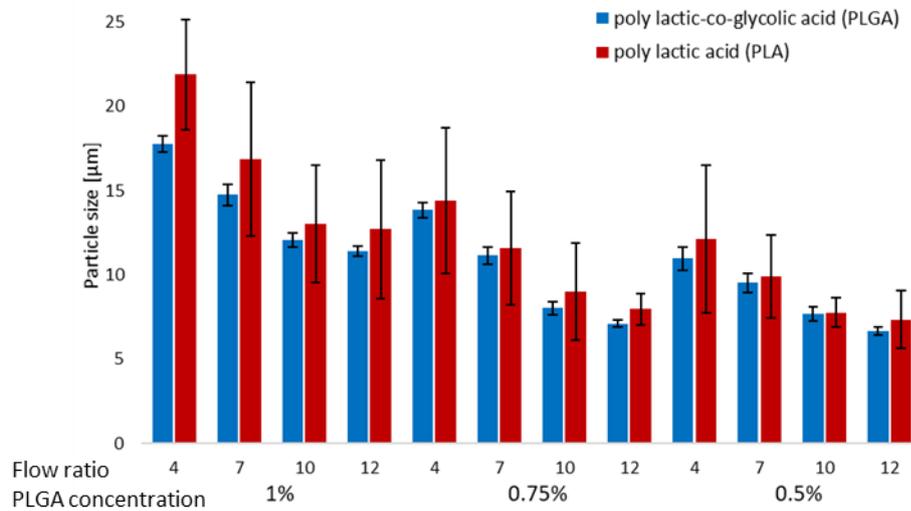


Figure 5.5. The impact of polymer choice on the particle size. Average size and standard deviation were calculated for a minimum of 50 particles in SEM images.

Particles produced using PLA are larger and have a rougher, more irregular surface when compared to the comparatively smooth PLGA particles, as shown on Figure 5.6. The difference in surface roughness and particle porosity has been reported previously²¹⁷ and is attributed to differences in polymer-solvent interactions.

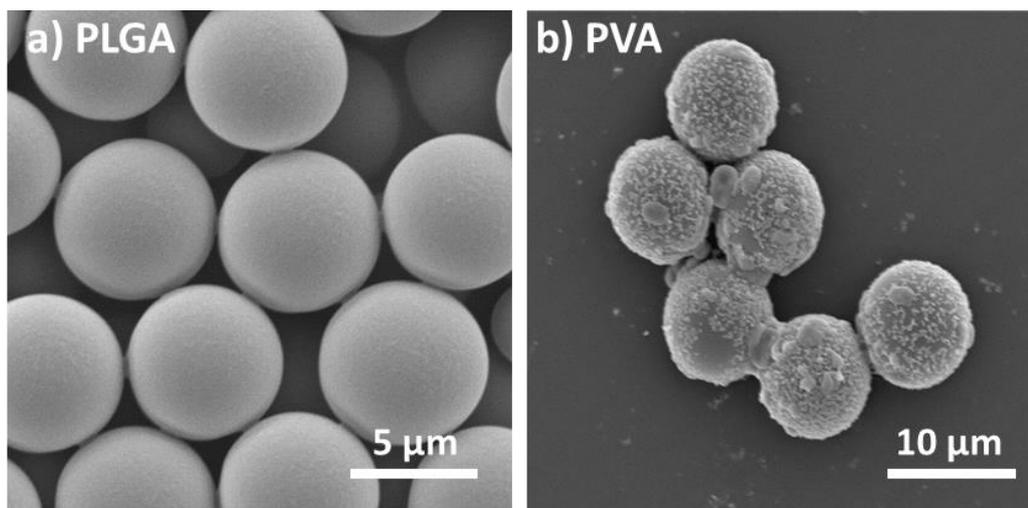


Figure 5.6. Comparison of the SEM images of the surface of a PLGA particles of 6.6 µm diameter (AM066) and b PLA particles produced with the same flow parameters (AM100), 8.4 µm diameter.

A higher concentration of PVA (2% or 5%) has been shown to result in smoother surface of the PLA particles ^{218,219}, however in this work, to obtain a more hydrophobic surface of the particle, the PVA concentration was kept at 1%.

5.3.4 Nanoparticle size tuning

The nanoparticles were prepared by flow focusing of the acetone PLGA solution in the 190 μm etch depth microfluidic chip. The parameters were optimised to obtain a stable flow and monodisperse particles and the settings used for particle size tuning are detailed in Table 5.3.

Table 5.3. Parameters used in PLGA nanoparticle size tuning. Concentrations of 0.25%, 0.5%, 0.75% and 1% were used and continuous phase flow rates were modified to achieve flow ratios of 4-16.

Sample name	PLGA concentration [%w/w]	Dispersed phase flow rate [$\mu\text{l}/\text{min}$]	Continuous phase flow rate [$\mu\text{l}/\text{min}$]	Flow ratio
AN112	1.00	25	50	4
AN113	1.00	25	100	8
AN114	1.00	25	150	12
AN115	1.00	25	200	16
AN108	0.75	25	50	4
AN109	0.75	25	100	8
AN110	0.75	25	150	12
AN111	0.75	25	200	16
AN104	0.50	25	50	4
AN105	0.50	25	100	8
AN106	0.50	25	150	12
AN107	0.50	25	200	16
AN100	0.25	25	50	4
AN101	0.25	25	100	8
AN102	0.25	25	150	12
AN103	0.25	25	200	16

Figure 5.7 presents the results of nanoparticle size tuning by altering the polymer concentration and the continuous phase flow rates. The changes to continuous or dispersed phase flow rates are statistically relevant for lower concentration PLGA solutions, but not for 0.75% and 1% PLGA solutions (Figure 5.7), which may be caused by higher density of the 0.75% and 1% solutions. The particle size was tuned between 120 - 180 nm. Each incremental increase in PLGA concentration resulted in statistically significant ($P \leq 0.05$) increase in the particle size (Figure 5.7). All produced nanoparticles were homogenous with the polydispersity index (PDI) below 0.09 and no correlation was found between the flow ratio and the PDI value.

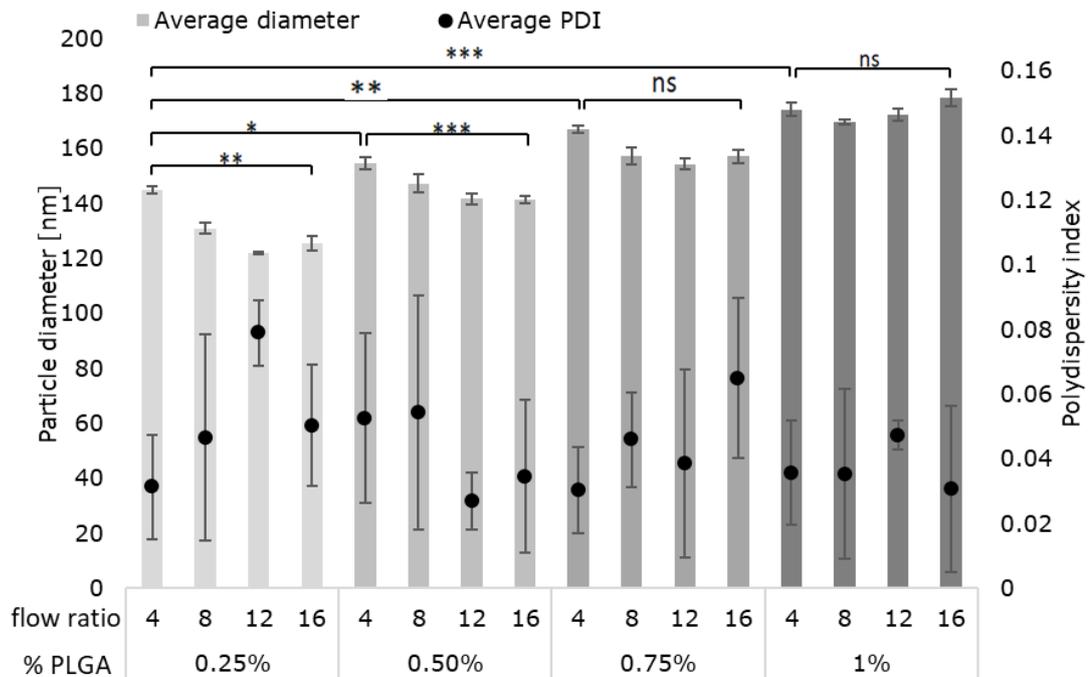


Figure 5.7. Particle diameter measured by DLS, impact of PLGA concentration (0.25%-1%) and a ratio of flow rates (4-16). Average diameter and standard deviation is calculated for 4 replicates. T-test was performed to assess significance of manufacturing parameters on particle size. ns $P > 0.05$, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$.

Further increase of PLGA concentration (higher than 1% w/w) causes blockages in the microfluidic chip, limiting the capability to produce nanoparticles larger than 180

nm. The use of concentrations lower than 0.25% results in wide particle size distribution. This behaviour of PLGA solutions with concentration outside 0.25-1% range determines the limits of the size tuning possibilities, as the concentration was found to be the determining factor in obtaining a desired particle size.

5.3.5 Impact of the surfactant on the nanoparticle size

The use of surfactant is not necessary in the production of PLGA nanoparticles from an acetone solution, however, a surfactant can be added in order to increase particle hydrophilicity or modify the particle size. PVA was chosen for this study, as it was the surfactant used in the preparation of the microparticles. When PVA is used as a surfactant, it becomes incorporated in the particle core, therefore expanding its diameter^{220,221}. In agreement with previous studies²²², PVA has been found to increase the size of PLGA nanoparticles while the impact of the total flow rate was not statistically significant ($P > 0.05$) as shown in Figure 5.8.

The impact of the total flow rate was also investigated, as previous studies reported that the total flow rate in some cases has²²³ or does not have²²⁴ an impact on the nanoparticle size. The hypothesised impact was that increased total flow rate would decrease the particle size due to increased pressure at the junction. The total flow rate was increased by a factor of 2 while maintaining the flow ratios by increasing both dispersed and continuous flow rates. In analysed particles produced with 1% PLGA and 1% PLA solutions, the increased total flow rate did not result in significant ($P > 0.05$) decrease in the particle diameter, as shown on Figure 5.8. Further increase of total flow rate resulted in PLGA blocking the junction.

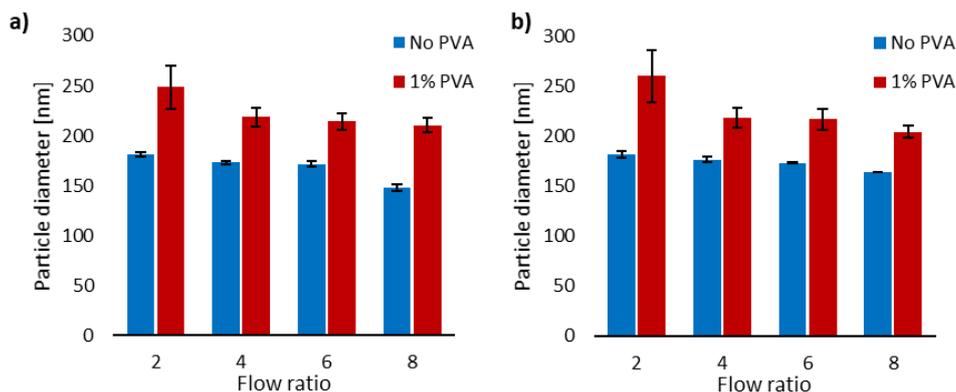


Figure 5.8. Impact of surfactant and the total flow rate on nanoparticle size. Particle diameter was measured by DLS, each bar represents average of 4 measurements and the error bars are standard deviation of 4 measurements. PLGA particles were produced with (red) and without (blue) 1% PVA in the water phase with the continuous phase flow rate of **a)** 25µm/ml and **b)** 50µm/ml.

In order to test the impact of total flow rate without blocking the chip, PLGA dissolved in DCM was investigated, specifically for settings used in batch AM055. The flow ratio of 4 was maintained while increasing the total flow rate 50-fold and 100-fold. Particles approximately 1 µm in diameter were produced using PLGA dissolved in DCM at high flow rates of dispersed : continuous phases 500 µl/min : 1000 µl/min and 1 ml/min : 2 ml/min (Figure 5.9). The particles produced in these conditions have diameters on the border between the nano- and microparticle formulations, however, they have a high polydispersity index (0.6 - 0.8).

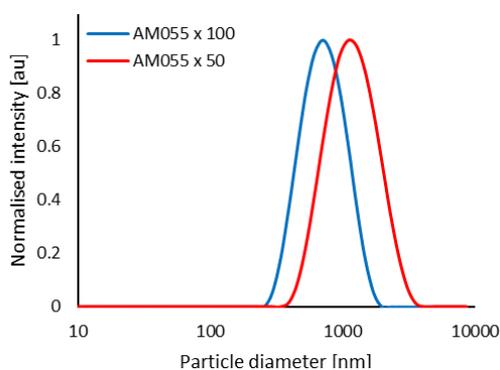


Figure 5.9. Size distribution of PLGA particles produced with 1% PLGA solution and high total flow rates, measured by dynamic light scattering. Particles produced with flow rates 500 µl/min : 1000 µl/min result in particles of approximately 1300 nm (red) and with flow rates 1 ml/min : 2 ml/min, approximately 730 nm (blue).

5.4 Conclusions

Particles produced in this work will be used to determine the impact of two key factors, namely size and hydrophobicity, on the adsorption of proteins on the surface of the particle. A reproducible method of manufacturing monodisperse PLGA micro- and nanoparticles using microfluidics was successfully developed. A reliable size tuning in a range between 120 - 250 nm and 5 - 20 μm was achieved by modifying parameters such as the polymer concentration and the flow rate of continuous and dispersed phases. Larger particles of diameter 50 - 200 μm as well as particles on the border between the micro and nanoscale have been produced, however they do not have sufficient monodispersity to be used in a study of particle size impact on protein adsorption.

The impact of the use of surfactant, 1% PVA, on the nanoparticle size was assessed in order to produce comparable particle sizes of different hydrophilicity. As the PVA becomes incorporated in the particle core, it enlarges the particle. A range of nanoparticle sizes were successfully produced to provide comparable particle diameters.

A range of microparticle sizes composed of poly lactic acid (PLA) was prepared to test the impact of higher hydrophobicity on the protein corona formation. The PLA particles were not significantly larger ($P > 0.05$) than the PLGA particles and were more polydisperse. SEM imaging revealed that the surface of PLA particles was less smooth than that of corresponding PLGA particles, which is caused by different interaction of the polymer with the solvent and may affect how the particles interact with proteins in solution.

6 Protein-functionalised polymer particles

6.1 Chapter aims

As described in the introduction (Section 1.4.4), current methods of analysing the chemistry of the biocorona, forming on the surface of the particles upon contact with biological media, require extensive sample preparation and therefore do not reflect the state of the particle in solution. This chapter aims to overcome this limitation by employing the ballistic sequencing to achieve the first *in situ* characterisation of protein-particle complexes.

Chapters 1 and 2 described the development and application of ballistic sequencing using the 3D OrbiSIMS for label-, matrix- and digestion-free analysis of proteins at surfaces. Chapter 5 focused on manufacturing and characterisation of monodisperse polymeric micro- and nanoparticles in a range of sizes. This chapter aims to combine the methods developed in sections 1 to 3 in order to characterise the microparticles after incubation with biological media with minimal sample preparation.

Microparticles will be incubated in media of increasing complexity: phosphate buffer saline (PBS), bovine serum albumin (BSA), human serum albumin (HSA) and full human serum prior to analysis using the 3D OrbiSIMS. The capabilities of the GCIB OrbitrapTM and LMIG ToF-SIMS analytical modes to provide chemical information of the particle surface in room temperature and cryogenic conditions will be assessed.

Experimental parameters such as duration of incubation, temperature, particle concentration and washing protocols affect the qualitative and quantitative outcome of the biocorona studies, as discussed in Section 1.4.4. Moreover, the proposed

method of observing particle interactions with the biological environment is the first attempt to characterise such complex systems *in situ*, different from any approach previously employed to this field of research, therefore this chapter will focus on method development.

6.2 Methods

6.2.1 Particle incubation

Particles produced in chapter 5 were washed 3 times by centrifugation at 9000 g, replacing supernatant with MiliQ water after each centrifugation. Washed particles (1 mL) were incubated with 0.5 mL of BSA, HSA or full human serum solution. The incubation was carried out in glass vials under constant agitation for 2 hours at room temperature. After the incubation the particles were washed 3 times by centrifugation at 9000 g.

6.2.2 ToF-SIMS measurement of freeze dried particles

After incubation, washed particles were freeze-dried using Lablyo Freeze Drier. For the acquisition of LMIG ToF-SIMS image, a 30 keV Bi₃⁺ primary beam was used. LMIG current was 0.05 pA. The 256 × 256 pixel ToF image was acquired over an area of 200 × 200 μm using random raster. The cycle time was set to 250 ms. Optimal target potential was set to +58 V. Three separate areas were analysed on each sample and each measurement lasted 15 scans, the total ion dose per measurement was 9.44×10^{10} ions/cm².

The results were analysed in SurfaceLab 7. Calibration of positive mode spectra was made using CH_3^+ , C_2H_3^+ , C_3H_5^+ , C_7H_7^+ and Au_3^+ ions. Negative mode spectra were calibrated to CN^- , CNO^- , Au^- , Au_3^- and Au_5^- .

6.2.3 3D OrbiSIMS measurement of frozen-hydrated particles

Particles were high-pressure frozen in 2mm freeze-fracture planchettes using Leica EM ICE High Pressure Freezer or plunge-frozen by dipping in liquid nitrogen. Leica transfer system was used to place the frozen samples in the instrument to prevent collecting moisture from the air.

20 keV Ar_{3000}^+ analysis beam of 20 μm diameter, was used as primary ion beam. Ar_{3000}^+ with duty cycle set to 4.4%, GCIB current was 190 pA. The Q Exactive depth profile was run on the area of $200 \times 200 \mu\text{m}$ using random raster mode with crater size $280 \times 280 \mu\text{m}$. The cycle time was set to 200 μs . Optimal target potential was approximately -293 V. The acquisition was run in temperature -170°C . The spectra were collected in positive polarity, in m/z range 150 - 2250. The injection time was set to 500 ms. Mass resolving power was set to 240,000 at $m/z = 200$.

6.3 Results and discussion

The aim of this chapter was to observe the biomolecule corona on the drug delivery particles without removing the coat from these conjugates, as a particularly challenging pharmaceutical application of the ballistic sequencing. The PLGA microparticles, produced by microfluidics in the work described in chapter 5 were washed and incubated with BSA, HSA or full human serum solution. The incubation was carried out under constant agitation for 2 hours at room temperature and after the incubation the particles were washed 3 times by centrifugation at 9000 g.

The first experiment aimed to assess whether the protein coat can be investigated by ToF-SIMS analysis of freeze-dried particles. Negative polarity imaging of the bare particles, particles incubated with PBS and particles incubated with bovine serum albumin (BSA) enabled detection of poly-lactic-co-glycolic acid (PLGA) and polyvinyl alcohol (PVA) peaks and protein related peaks, as presented in Figure 6.1. The overlay image of the PLGA ($C_3H_5O_2^-$ fragment) and backbone of the protein chain (CN^- fragment) diagnostic ions are showed in Figure 6.1. The protein fragment (green) is visible in the entire field of view and is anticorrelated with the particle ion (red). This analysis revealed that freeze-drying is not a suitable sample preparation method, as the protein detaches from the particles in the drying process.

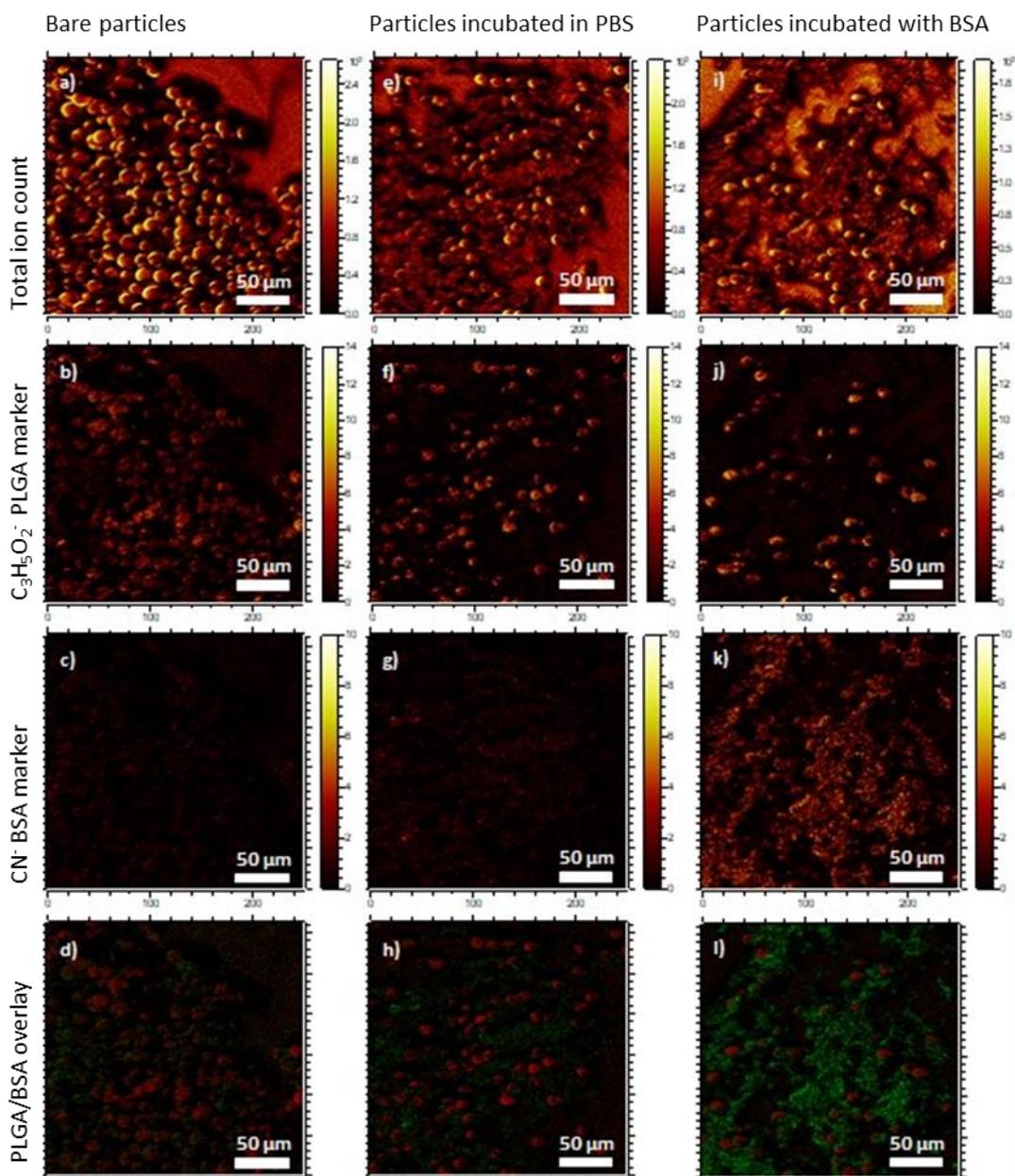


Figure 6.1. Comparison of ToF-SIMS images of freeze-dried a-d) bare particles, b-h) particles incubated with PBS and c-l) particles incubated with bovine serum albumin. Total ion images of (a) bare particles, (e) particles incubated with PBS and (i) particles incubated with BSA. Example PLGA ion $C_3H_5O_2^-$ in (b) bare particles, (f) particles incubated with PBS and (j) particles incubated with BSA. CN^- protein marker ion in (c) bare particles, (g) particles incubated with PBS and (k) particles incubated with BSA. Overlay of the $C_3H_5O_2^-$ and CN^- markers.

6.3.1 Detection and assignment of diagnostic human serum albumin fragments using cryo 3D OrbiSIMS

The aim of this chapter is to utilise the cryogenic capabilities of the 3D OrbiSIMS to capture the particles *in situ* in the solution in the form in which they may be presented to cells. However, another aspect of SIMS experiments in the presence of frozen water is increased ionisation of the sample components. This effect is particularly pronounced with higher molecular weight compounds and is associated with the presence of H_3O^+ ions capable of proton donation and reduced sample damage during sputtering of frozen material²²⁵. Signal enhancement has been studied in frozen samples containing water⁶², spraying water above the sample⁶⁰ and by using water clusters as primary ion beams^{59,226}. These methods have been found to increase the $[\text{M}+\text{H}]^+$ signals and also the salt adducts such as $[\text{M}+\text{Na}]^+$ and $[\text{M}+\text{K}]^+$. However, the amount of protein in the sputtered volume is smaller than in the case of dry protein films due to the dilution in water or buffer. Human serum albumin at blood concentration, 45 mg/mL, was measured in the cryogenic conditions to assess whether highly specific protein ions are detectable using this method.

The positive polarity cryo OrbiSIMS spectrum of human serum albumin dissolved in water contains up to 13-membered sequences originating from the N-terminus, which stand out above the average intensity peaks as presented in Figure 6.2 and Table 6.1. 8-membered sequences originating from the C-terminus were also present in the spectrum, however they are left out of the figure for clarity. The sequence assignments are listed in Table 6.1. and they are present as a, b and c ions, following the fragmentation patterns described in Section 3.2.3. These high molecular weight

fragments ($m/z > 600$) were not previously detected in the room temperature measurement of dried human serum albumin films, as shown in Figure 6.2.

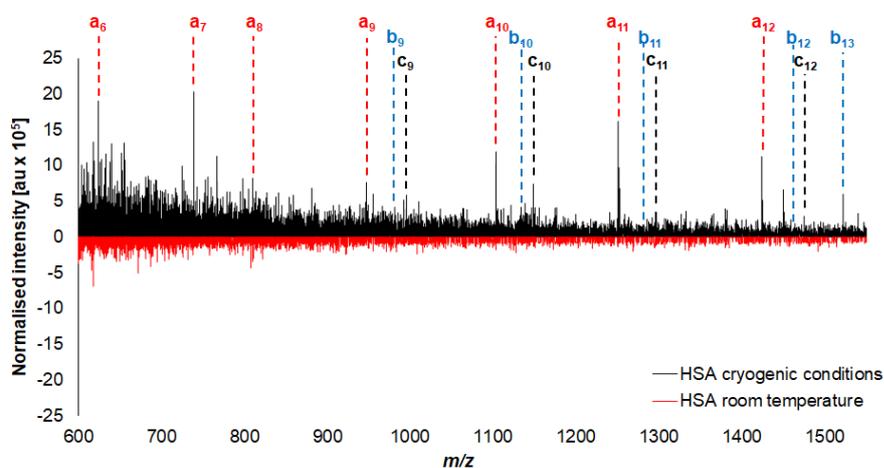


Figure 6.2. Inverted overlay comparison of human serum albumin measured in cryogenic conditions (black, top) and dried protein film in room temperature (red, bottom).

As described in Section 3.2.3, the negative polarity spectra can provide additional information about the analysed protein and in the 3D OrbiSIMS negative polarity spectra contain y, x and z ions. Similarly to the positive polarity frozen-hydrated 3D OrbiSIMS spectrum of HSA, the terminus ions have higher intensities than average peaks in the spectrum. Up to 10-membered sequences from the C-terminus are present in the negative polarity, as demonstrated in Figure 6.3. and Table 6.2.

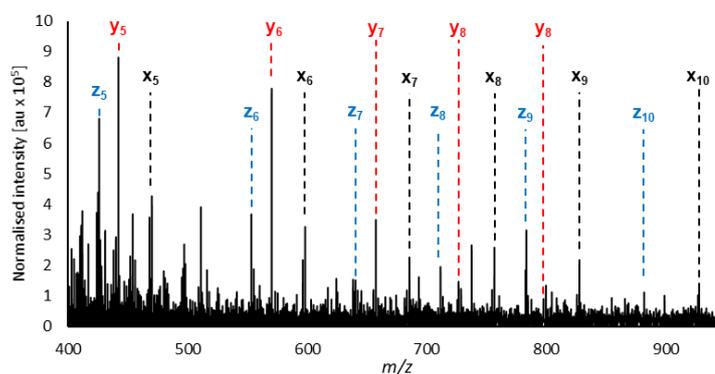


Figure 6.3. Negative polarity spectrum of human serum albumin acquired in cryogenic conditions.

Table 6.1. Assignments of N-terminus ions detected in the positive polarity cryo OrbiSIMS spectrum of human serum albumin.

Description	a			b			c		
	<i>m/z</i>	Assignment	Dev. (ppm)	<i>m/z</i>	Assignment	Dev. (ppm)	<i>m/z</i>	Assignment	Dev. (ppm)
a4	424.2306	C ₁₈ H ₃₀ N ₇ O ₅ ⁺	0.65	452.2255	C ₁₉ H ₃₀ N ₇ O ₆ ⁺	0.72	469.2523	C ₁₉ H ₃₃ N ₈ O ₆ ⁺	1.18
a5	511.2626	C ₂₁ H ₃₅ N ₈ O ₇ ⁺	0.61	539.2580	C ₂₂ H ₃₅ N ₈ O ₈ ⁺	1.41	556.2844	C ₂₂ H ₃₈ N ₉ O ₈ ⁺	1.06
a6	640.3057	C ₂₆ H ₄₂ N ₉ O ₁₀ ⁺	1.22	668.3007	C ₂₇ H ₄₂ N ₉ O ₁₁ ⁺	1.32	685.3269	C ₂₇ H ₄₅ N ₁₀ O ₁₁ ⁺	0.71
a7	739.3742	C ₃₁ H ₅₁ N ₁₀ O ₁₁ ⁺	1.18	767.3694	C ₃₂ H ₅₁ N ₁₀ O ₁₂ ⁺	1.45	784.3957	C ₃₂ H ₅₄ N ₁₁ O ₁₂ ⁺	1.16
a8	810.4111	C ₃₄ H ₅₆ N ₁₁ O ₁₂ ⁺	0.79	838.4064	C ₃₅ H ₅₆ N ₁₁ O ₁₃ ⁺	1.26			
a9	947.4702	C ₄₀ H ₆₃ N ₁₄ O ₁₃ ⁺	0.94	975.4639	C ₄₁ H ₆₃ N ₁₄ O ₁₄ ⁺	-0.39	992.4922	C ₄₁ H ₆₆ N ₁₅ O ₁₄ ⁺	1.44
a10	1103.5708	C ₄₆ H ₇₅ N ₁₈ O ₁₄ ⁺	0.33	1131.5654	C ₄₇ H ₇₅ N ₁₈ O ₁₅ ⁺	0.00	1148.5922	C ₄₇ H ₇₈ N ₁₉ O ₁₅ ⁺	0.21
a11	1250.6391	C ₅₅ H ₈₄ N ₁₉ O ₁₅ ⁺	0.15	1406.7290	C ₆₂ H ₉₆ N ₂₁ O ₁₇ ⁺	0.18	1295.6613	C ₅₆ H ₈₇ N ₂₀ O ₁₆ ⁺	0.77
a12	1378.7326	C ₆₁ H ₉₆ N ₂₁ O ₁₆ ⁺	-0.90	1521.7571	C ₆₆ H ₁₀₁ N ₂₂ O ₂₀ ⁺	0.92	1423.7564	C ₆₂ H ₉₉ N ₂₂ O ₁₇ ⁺	0.78

Table 6.2. Assignments of C-terminus ions detected in negative polarity cryo OrbiSIMS spectrum of human serum albumin.

Description	y			x			z		
	<i>m/z</i>	Assignment	Dev. (ppm)	<i>m/z</i>	Assignment	Dev. (ppm)	<i>m/z</i>	Assignment	Dev. (ppm)
y 1				158.082	C ₇ H ₁₂ NO ₃ ⁻	-1.54			
y 2	187.1086	C ₈ H ₁₅ N ₂ O ₃ ⁻	-1.23	215.1036	C ₉ H ₁₅ N ₂ O ₄ ⁻	-0.83	172.0977	C ₈ H ₁₄ NO ₃ ⁻	-1.22
y 3	300.1924	C ₁₄ H ₂₆ N ₃ O ₄ ⁻	-1.45	328.1873	C ₁₅ H ₂₆ N ₃ O ₅ ⁻	-1.45	285.1817	C ₁₄ H ₂₅ N ₂ O ₄ ⁻	-1.05
y 4	371.2293	C ₁₇ H ₃₁ N ₄ O ₅ ⁻	-1.88	399.2241	C ₁₈ H ₃₁ N ₄ O ₆ ⁻	-1.91	356.2188	C ₁₇ H ₃₀ N ₃ O ₅ ⁻	-0.92
y 5	442.2663	C ₂₀ H ₃₆ N ₅ O ₆ ⁻	-1.84	470.261	C ₂₁ H ₃₆ N ₅ O ₇ ⁻	-2.13	427.2553	C ₂₀ H ₃₅ N ₄ O ₆ ⁻	-2.09
y 6	570.3245	C ₂₅ H ₄₄ N ₇ O ₈ ⁻	-2.07	598.3198	C ₂₆ H ₄₄ N ₇ O ₉ ⁻	-1.36	555.3137	C ₂₅ H ₄₃ N ₆ O ₈ ⁻	-1.91
y 7	657.3569	C ₂₈ H ₄₉ N ₈ O ₁₀ ⁻	-1.30	685.3514	C ₂₉ H ₄₉ N ₈ O ₁₁ ⁻	-1.81	642.3457	C ₂₈ H ₄₈ N ₇ O ₁₀ ⁻	-1.79
y 8	728.3939	C ₃₁ H ₅₄ N ₉ O ₁₁ ⁻	-1.27	756.3882	C ₃₂ H ₅₄ N ₉ O ₁₂ ⁻	-2.06	713.3829	C ₃₁ H ₅₃ N ₈ O ₁₁ ⁻	-1.50
y 9	799.4303	C ₃₄ H ₅₉ N ₁₀ O ₁₂ ⁻	-2.03	827.4254	C ₃₅ H ₅₉ N ₁₀ O ₁₃ ⁻	-1.71	784.4193	C ₃₄ H ₅₈ N ₉ O ₁₂ ⁻	-2.25
y 10	898.4991	C ₃₉ H ₆₈ N ₁₁ O ₁₃ ⁻	-1.40	926.4933	C ₄₀ H ₆₈ N ₁₁ O ₁₄ ⁻	-2.15	883.4872	C ₃₉ H ₆₇ N ₁₀ O ₁₃ ⁻	-2.56

6.3.2 Particle chemistry analysed *in situ* using cryo OrbiSIMS

In order to avoid protein detachment from the surface observed during freeze-drying and obtain samples, in which particles are partially suspended in the solution, the samples were prepared using a freeze-fracture cell and Leica high pressure freezing system.

The use of Bi_3^+ primary beam and the ToF analyser enabled fast imaging of the particles suspended in buffer, as shown on Figure 6.4. Example amino acid fragment ions representing glycine ($m/z = 30$, CH_4N^+), alanine ($m/z = 44$, $\text{C}_2\text{H}_6\text{N}^+$) and methionine ($m/z = 61$, $\text{C}_2\text{H}_5\text{S}^+$) were localised on the surface of the particles incubated with BSA and absent in the bare particle spectra and images.

Phosphate salt ions with distribution contrasting to the distribution of amino acid peaks in the image have been assigned, as presented in Figure 6.5a. The red colour represents a sum of amino acid fragments and the blue colour represents a sum of salt ions: $\text{Na}_2\text{HPO}_4^+$, PO_3H^+ . This suggests that the proteins are only present on the surface of the particles, not in the surrounding solution, however does not provide chemically specific information about the protein identity, therefore GCIB Orbitrap measurement was run in the same sample area. ToF-SIMS images were acquired before and after the GCIB Orbitrap measurement and are presented in Figure 6.5(a) and Figure 6.5(b) respectively. The GCIB sputtering removed the salts (Figure 6.5, blue) from the solution surrounding the particles and sodium (Figure 6.5d) from the surface of the particles. In the ToF-SIMS image acquired after the GCIB Orbitrap analysis, the distribution of amino acid peaks follows the distribution of total ion

image and there are no contrasting ions in the image, therefore it cannot be assured that the protein is present only on the surface of the particles.

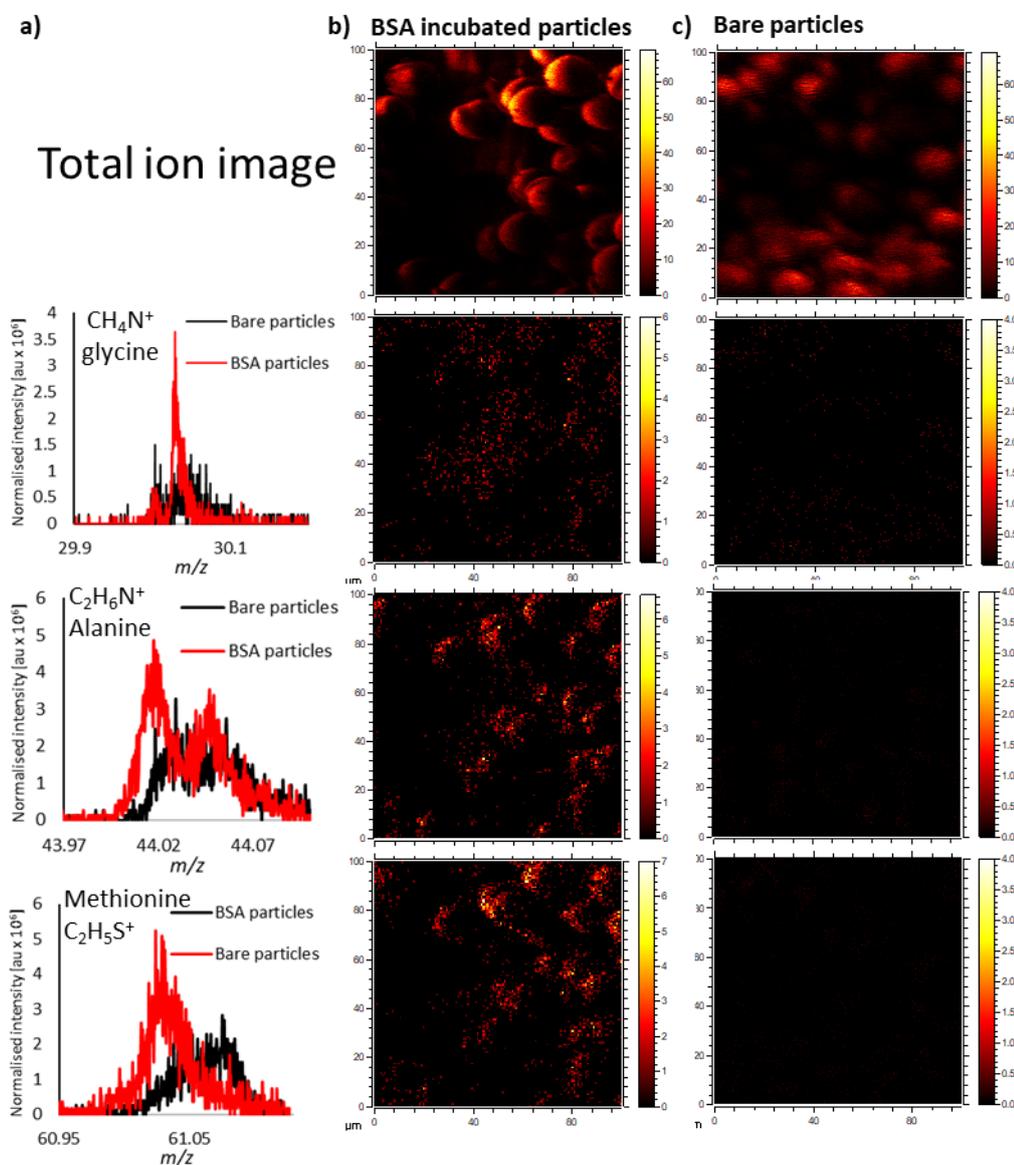


Figure 6.4. Comparison of cryo ToF-SIMS spectra (a) and images of particles (b) incubated with bovine serum albumin and (c) not incubated. The samples were high-pressure frozen and freeze-fractured prior to the analysis. Amino acid fragment ions representing glycine ($m/z = 30$, CH_4N^+), alanine ($m/z = 44$, $\text{C}_2\text{H}_6\text{N}^+$) and methionine ($m/z = 61$, $\text{C}_2\text{H}_5\text{S}^+$) are detected on the surface of the particles incubated with BSA and absent in the bare particle spectra and images.

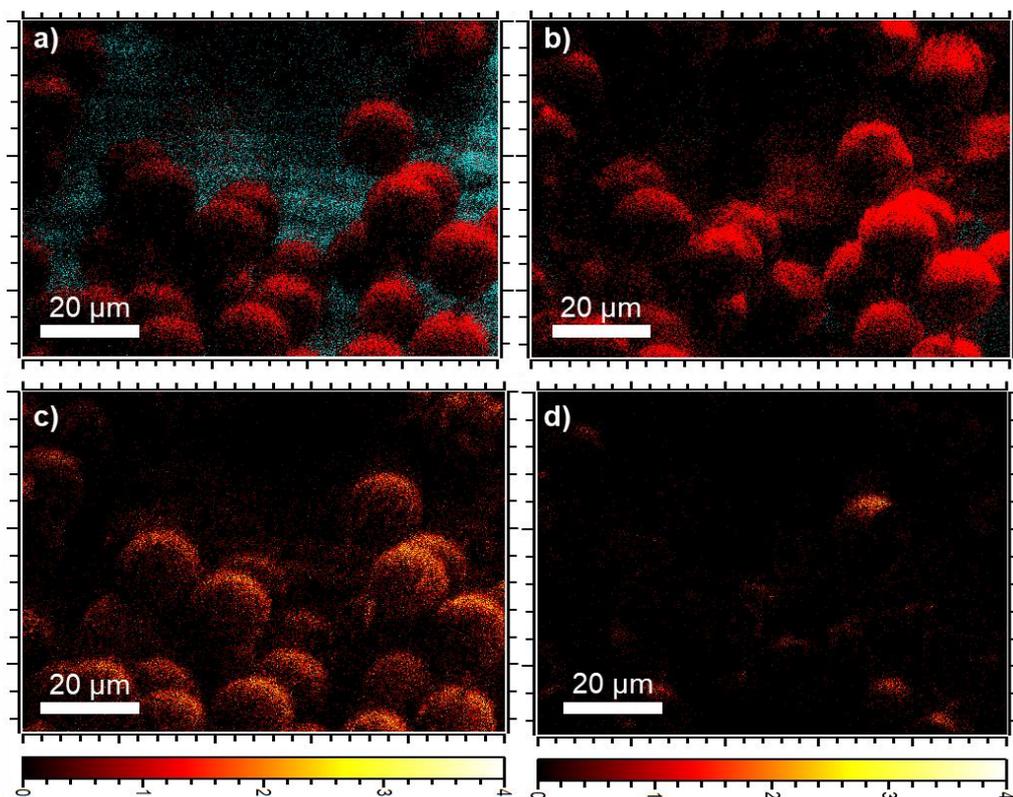


Figure 6.5. Cryo ToF-SIMS images of PLGA particles high-pressure frozen and freeze-fractured after incubation with bovine serum albumin. (a) overlay of amino acid fragments (red) and salt ions (blue) before acquisition of GCIB Orbitrap spectra. (b) Overlay of the same amino acid fragments (red) and salt ions (blue) after the GCIB sputtering. (c) sodium ion distribution before GCIB sputtering and (d) after sputtering. The images have been rotated by 90° clockwise and reduced to 71% of height to counteract the image distortion due to the beam coming from 45° to the sample.

The relatively small (2 mm diameter) area of the freeze-fracture cells, topography of the samples and challenging charge compensation of the frozen water-based samples resulted in poor quality (low, 10^4 total ion count/s, presence of characteristic noise) of the GCIB Orbitrap spectra of all particle samples, in which characteristic ions from neither PLGA or protein can be detected. Therefore, the presence of the protein on the surface of the particles or in the surrounding solution cannot be confirmed.

Figure 6.6 demonstrates different available substrates and sample holders. The increased working area of 6 mm planchettes allowed optimisation of charge compensation and improved the quality of the acquired spectrum, as presented in

Figure 6.7. The total ion count was increased by a factor of 100, the noise was reduced and distinct fragments of PLGA were detected in the spectrum (Figure 6.7).

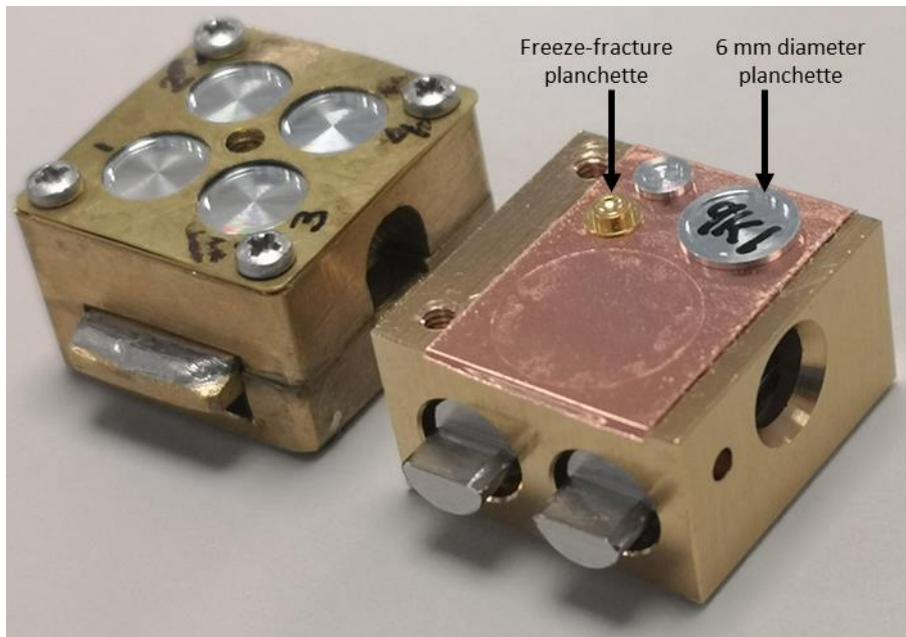


Figure 6.6. Experimental setup of the cryo sample holder (left) and the comparison of the size of the freeze-fracture planchette and the 6 mm planchette (right).

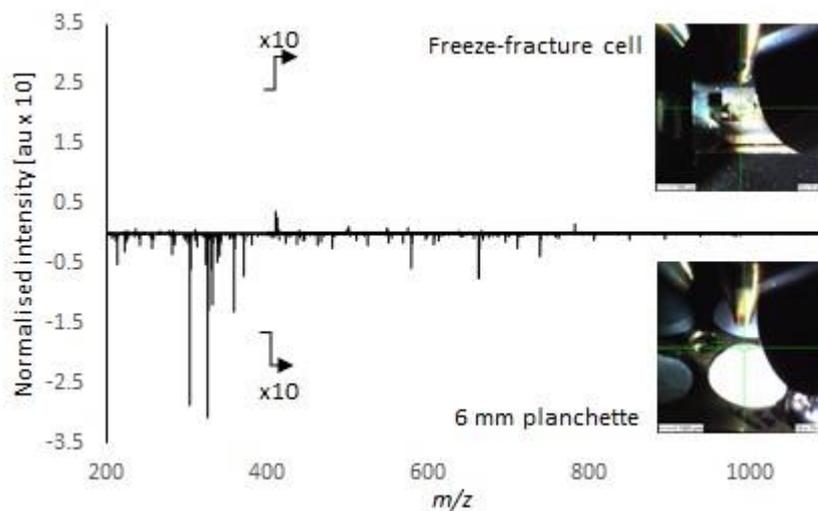


Figure 6.7. Inverted overlay comparison of the cryo OrbiSIMS spectra of pure PLGA particles, obtained from samples high-pressure frozen in a freeze-fracture cell (top) and the samples in the 6 mm planchettes, plunge-frozen by dipping in the liquid nitrogen (bottom).

Despite improved overall ion count, only one distinct human serum albumin ion was detected in the cryo OrbiSIMS spectra of particles incubated with HSA and particles incubated with serum, acquired from samples in the 6 mm planchettes. The detected ion, b₈ fragment of HSA sequence (C₃₅H₅₆N₁₁O₁₃⁺) was detected with similar intensity in HSA-incubated particles and serum-incubated particles as shown in Figure 6.8. The characteristic ion detected in the cryo OrbiSIMS spectra may indicate the presence of the protein on the surface of the particles, however the ion intensity was not sufficient to detect it in 3D OrbiSIMS imaging, therefore the presence of the protein in the surrounding solution cannot be excluded. The particles after incubation were washed by centrifugation, which may have caused protein removal from the particle surface. To circumvent this effect, large cut-off (150 kDa) dialysis should be used in the future to wash the particles after the incubation.

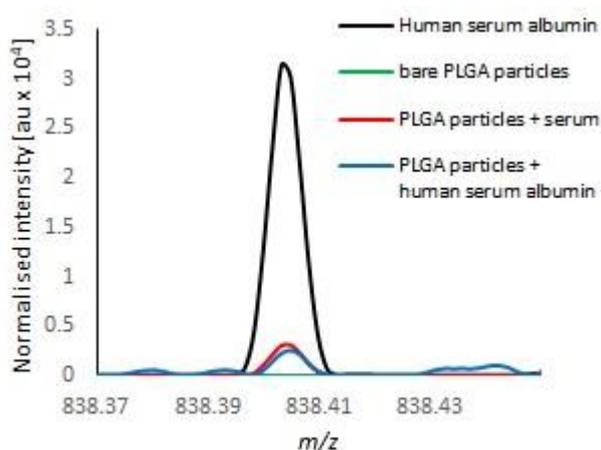


Figure 6.8. Overlay comparison of human serum albumin b₈ ion, present in the positive polarity cryo OrbiSIMS spectrum of the particles after incubation with human serum albumin (blue) and full serum (red), acquired from samples in the 6 mm planchettes.

Cryo ToF-SIMS imaging of the pure particles, particles incubated with HSA and particles incubated with serum was carried out to visualise the location of the

microparticles in the sample. The microparticles are not visible in the image of the pure particles and particles incubated with serum (Figure 6.9 a and c) and are only indistinctly outlined in the image of particles incubated with HSA (Figure 6.9b). These results suggest that during sample preparation the particles sank to the bottom of the 6mm planchette and are buried in the depth of the sample, not exposed for the analysis in a way they were available while using the freeze-fracture cell. Further method development is necessary, specifically depositing a thinner layer of the solutions on the planchettes or letting excess buffer evaporate before freezing.

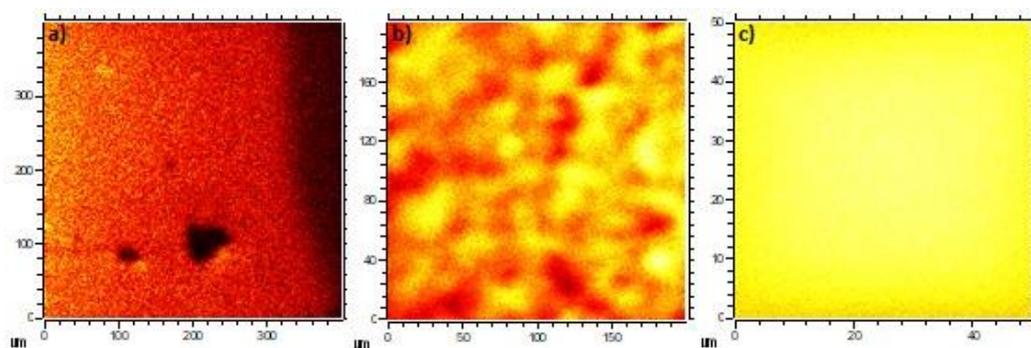


Figure 6.9. Cryo ToF-SIMS total ion images of a) pure PLGA particles, b) particles incubated with human serum albumin and c) particles incubated with human serum. The images were acquired from samples in the 6 mm planchettes.

Due to the particles not being visible in the the acquired cryo ToF-SIMS images, it is unclear whether the lack of protein fragments detected in the spectra is caused by not sampling the surface of the particles, the protein being below the limit of OrbiSIMS detection, or the protein is being removed from the surface of the particles in the process of washing by centrifugation.

In the positive cryo OrbiSIMS spectrum of particles incubated with the human serum acquired from samples in the 6 mm planchettes, multiple serum lipids: diglycerides, triglycerides and ceramides were detected, as shown in Table 6.3. However, they

may form a residue remaining in the supernatant after washing, not bound to the particle surface. Washing the particles by dialysis could reduce this effect and preparation of thin layer samples could provide more opportunity to image the lipid ions in relation to the particles.

Table 6.3. Assignments of lipids detected in the the positive polarity cryo OrbiSIMS spectrum of PLGA particles incubated with serum, acquired from samples in the 6 mm planchettes.

<i>m/z</i>	Assignment	Dev. (ppm)	Description
184.0732	C ₃ H ₁₃ PNO ₄ ⁺	-0.50	Phosphocholine
597.4853	C ₃₇ H ₆₆ O ₄ Na ⁺	-0.10	DG O-34:4/MG 34:4/MG O-34:5;O
599.5011	C ₃₇ H ₆₈ O ₄ Na ⁺	0.25	DG O-34:4/MG 34:4/MG O-34:5;O
621.4854	C ₃₉ H ₆₆ O ₄ Na ⁺	0.15	CE 12:1;O2/DG O-36:6
623.5009	C ₃₉ H ₆₈ O ₄ Na ⁺	-0.11	CE 12:1;O2/DG O-36:6
666.4833	C ₃₆ H ₇₀ NO ₆ PNa ⁺	0.01	CerP 36:2;O2/LPC O-28:3/LPE O-31:3
671.5737	C ₄₅ H ₇₆ O ₂ Na ⁺	-0.04	CE 18:2

6.4 Conclusions

This chapter aimed to visualise the chemistry of the interaction between drug delivery particles and the biological environment *in situ*. Such an approach has never been achieved before and current biocorona analysis methods require separation of the biomolecules from the surface of the particles. To enable *in situ* study of such bioconjugates, a method of observing proteins at surfaces without any sample manipulation was developed in the previous parts of this project (Chapter 3 and 4). With the use of this ballistic sequencing method and the monodisperse polymeric particles described in Chapter 5, this work aimed to provide new insight into biomolecule adsorption on the surface of particles to guide design and production of drug delivery or diagnostic vehicles.

Analysis of human serum albumin solution in cryogenic conditions resulted in highly specific up to 13-membered amino acid sequences from the N-terminus of the protein in positive polarity and up to 10-membered sequences from the C-terminus in negative polarity. This enhancement of the informative signals by the use of cryo OrbiSIMS demonstrates the capability to analyse proteins in native state and a potential to observe proteins on the surface of the particles.

One characteristic human serum albumin fragment and multiple serum lipid ions were detected in the cryo OrbiSIMS spectra of the PLGA particles incubated with human serum. However, no particles were visible in the image of the surface of the sample, therefore the detected fragments may originate from the surrounding solution rather than from the surface of the microparticles.

The results obtained in this work highlight the challenges in the sample handling and characterisation of the protein corona, specifically the reversible binding of the proteins to the particles. Although the structure-performance relationships between the particle size and material and the biocorona composition were not achieved, the preliminary results obtained in this work could form a foundation for future *in situ* study of material-biology interaction. Further development is needed to capture the true *in situ* composition of the particle coat.

7 Summary and future work

This work developed the first matrix- and label-free *in situ* assignment of undigested proteins using 3D OrbiSIMS for applications focusing on the interface between biology and synthesised materials. The findings represent a step change in the capabilities for biological analysis, providing the opportunity to move forward a range of scientific research areas including understanding the molecular mechanism of diseases, the development of medicines and the organism reaction to implanted medical devices. Fragmentation induced by Ar_{3000}^+ primary beam resulted in spectra containing large ($m/z > 500$) ions, which can be used for *de novo* sequencing protein identification and this approach was termed ballistic sequencing. From 5 to 53% of the sequence was assigned in 16 analysed proteins up to 272 kDa in size.

The chemical specificity of ions assigned by this approach allowed identification of common and distinct peaks in the spectra of similar proteins, bovine serum albumin and human serum albumin. The results were compared with classical proteomic methods and 3D OrbiSIMS analysis showed similarities and differences from the most commonly used low energy collision induced ionisation and analogous peak types were present in both positive and negative spectra. Specifically, the ballistic sequencing resulted in a, b, c and a-NH₃ ions at the N-terminus of the amino acid sequence, y, z and x ions at the C-terminus as well as ions from sequences in the middle of the protein chain. The assignments were confirmed by the MS/MS capability of the 3D OrbiSIMS. The lateral resolution of the ballistic sequencing was assessed by masking the protein film with a TEM grid, achieving 10 μm distance between the 16% intensity to 84% intensity of the sum of protein fragment ions.

Characteristic features such as proton, ligand or cell-binding sites were detected in the analysed proteins, which could improve the confidence in untargeted protein identification and could be applied to studying protein-ligand interactions.

Fragments of abundant proteins were assigned in complex biological mixtures such as human serum and corneodesmosin, keratin and collagen were mapped across human skin. Additionally, the analysis of a biochip revealed that highly specific lysozyme fragments can be assigned in the 3D OrbiSIMS spectra a protein monolayer sample. This finding demonstrates broad applications of the ballistic sequencing in areas crucial to health, medicine and medical devices development, particularly samples with pico- or femtomolar amount of protein on the surface.

A reproducible method of manufacturing monodisperse PLGA micro- and nanoparticles using microfluidics was successfully developed for subsequent *in situ* investigation of protein adsorption on the surface of the particles using 3D OrbiSIMS. A reliable size tuning was achieved in a range between 120 - 250 nm and 5 - 20 μm and the impact of the surfactant and hydrophobic polymer on the size and surface characteristics of the particles was assessed. One characteristic human serum albumin fragment and multiple serum lipid ions were detected in the cryo OrbiSIMS spectra of the PLGA particles incubated with human serum. The detected fragments may originate from the surrounding solution rather than from the surface of the microparticles and further method development is needed to capture the true *in situ* composition of the particle coat. However, 3D OrbiSIMS analysis of human serum albumin solution in cryogenic conditions resulted in enhanced detection of highly specific protein fragments, compared to room temperature analysis. This finding demonstrates the capability to analyse proteins in native state and a potential to observe proteins on the surface of the particles in the future.

Further development of the ballistic sequencing method should include investigation of the limits of detection, lateral and spatial resolution for model samples and the matrix effect causing suppression of protein fragments in complex mixtures.

Extension of the benchmark of proteins studied in this work together with computational data analysis of the 3D OrbiSIMS datasets can help understand protein fragmentation under Ar_{3000}^+ , improve confidence in the automated sequence assignment described in Section 4.2.2 and enable high throughput protein analysis. 3D OrbiSIMS analysis of peptides and employment of different argon cluster sizes is also expected to improve understanding of ballistic fragmentation. The potential to apply this method to other aspects of proteomics, such as analysis of insoluble proteins, detection of post translational modification or protein-protein interactions and biopolymer analysis (such as DNA, RNA or polysaccharides) should also be explored.

Moreover, future studies should include considerations of the mechanism of the increased detection of large C- and N-terminus sequences in the cryo OrbiSIMS spectra. The mechanism may be related to the presence of water, which is known to enhance ionisation of molecules in SIMS (Section 1.2), reduced fragmentation due to the frozen state of the sample or the native state of the protein as opposed to dried denatured protein films. Investigating these mechanisms would involve preparation of a series of diluted samples as well as dry protein film and denatured protein in cryogenic conditions.

Further work towards *in situ* analysis of biocorona should focus on the following aspects: firstly, thin samples of few layers of particles should be deposited into the 6 mm planchettes and frozen by plunge-freezing in liquid nitrogen. Such preparation would benefit from large sample area while maintaining the particles exposed close

to the sample surface. Secondly, optimisation of particle incubation and washing. Parameters such as temperature, time of incubation and particle concentration should be controlled and altered to assess the impact of these changes. Post-incubation washing procedure should be optimised to avoid harsh removal of the proteins from the surface of the particles. Large cut-off dialysis is suggested for this purpose. Further, the impact of particle size and chemistry, such as hydrophobic, hydrophilic character, presence of targeting molecules or protein pre-coating on the composition of the biocorona should be assessed. For this the different particles developed in chapter 3 could be used.

8 References

1. Ekblad, T. & Liedberg, B. Protein adsorption and surface patterning. *Curr. Opin. Colloid Interface Sci.* **15**, 499–509 (2010).
2. Michno, W., Wehrli, P. M., Zetterberg, H. & Blennow, K. GM1 locates to mature amyloid structures implicating a prominent role for glycolipid-protein interactions in Alzheimer pathology. *BBA - Proteins Proteomics* 1–10 (2018).
3. Hardesty, W. M. & Caprioli, R. M. In situ molecular imaging of proteins in tissues using mass spectrometry. *Anal. Bioanal. Chem.* **391**, 899–903 (2008).
4. Stauber, J. *et al.* On-Tissue Protein Identification and Imaging by MALDI-Ion Mobility Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **21**, 338–347 (2010).
5. Chaurand, P., Schwartz, S. A. & Caprioli, R. M. Assessing Protein Patterns in Disease Using Imaging Mass Spectrometry. *J. Proteome Res.* **3**, 245–252 (2004).
6. Saleem, M. & Galla, H. J. Surface view of the lateral organization of lipids and proteins in lung surfactant model systems-A ToF-SIMS approach. *Biochim. Biophys. Acta - Biomembr.* **1798**, 730–740 (2010).
7. Veisheh, M., Zhang, Y., Hinkley, K. & Zhang, M. Two-dimensional protein micropatterning for sensor applications through chemical selectivity technique. *Biomed. Microdevices* **3**, 45–51 (2001).
8. Lundberg, E. & Borner, G. H. H. Spatial proteomics: a powerful discovery tool for cell biology. *Nat. Rev. Mol. Cell Biol.* **20**, 285–302 (2019).
9. Kompauer, M., Heiles, S. & Spengler, B. Autofocusing MALDI mass

- spectrometry imaging of tissue sections and 3D chemical topography of nonflat surfaces. *Nat. Methods* **14**, 1156–1158 (2017).
10. Ryan, D. J., Spraggins, J. M. & Caprioli, R. M. Protein identification strategies in MALDI imaging mass spectrometry: a brief review. *Curr. Opin. Chem. Biol.* **48**, 64–72 (2019).
 11. Ryan, D. J., Spraggins, J. M. & Caprioli, R. M. Protein identification strategies in MALDI imaging mass spectrometry: a brief review. *Curr. Opin. Chem. Biol.* **48**, 64–72 (2019).
 12. Kompauer, M., Heiles, S. & Spengler, B. Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4- μm lateral resolution. *Nat. Methods* **14**, 90–96 (2016).
 13. Rao, W., Celiz, A. D., Scurr, D. J., Alexander, M. R. & Barrett, D. A. Ambient DESI and LESA-MS analysis of proteins adsorbed to a biomaterial surface using in-situ surface tryptic digestion. *J. Am. Soc. Mass Spectrom.* **24**, 1927–1936 (2013).
 14. Kocurek, K. I., Griffiths, R. L. & Cooper, H. J. Ambient ionisation mass spectrometry for in situ analysis of intact proteins. *J. Mass Spectrom.* **53**, 565–578 (2018).
 15. Han, J. *et al.* Imaging of protein distribution in tissues using mass spectrometry: An interdisciplinary challenge. *TrAC - Trends Anal. Chem.* **112**, 13–28 (2019).
 16. Cooper, H. J. & Hale, O. J. Native mass spectrometry imaging of proteins and protein complexes by nano-desi. *Anal. Chem.* **93**, 4619–4627 (2021).

17. Griffiths, R. L. & Cooper, H. J. Direct Tissue Profiling of Protein Complexes: Toward Native Mass Spectrometry Imaging. *Anal. Chem.* **88**, 606–609 (2016).
18. Griffiths, R. L., Konijnenberg, A., Viner, R. & Cooper, H. J. Direct Mass Spectrometry Analysis of Protein Complexes and Intact Proteins up to >70 kDa from Tissue. *Anal. Chem.* **91**, 6962–6966 (2019).
19. Crecelius, A. C. *et al.* Three-dimensional visualization of protein expression in mouse brain structures using imaging mass spectrometry. *J. Am. Soc. Mass Spectrom.* **16**, 1093–1099 (2005).
20. Benninghoven, A. Surface analysis by Secondary Ion Mass Spectrometry (SIMS). *Surf. Sci.* **299–300**, 246–260 (1994).
21. Gilmore, I. S. SIMS of organics—Advances in 2D and 3D imaging and future outlook. *J. Vac. Sci. Technol. A Vacuum, Surfaces, Film.* **31**, 050819 (2013).
22. Lockyer, N. P. & Vickerman, J. C. Progress in cellular analysis using ToF-SIMS. *Appl. Surf. Sci.* **231–232**, 377–384 (2004).
23. Henss, A. *et al.* High resolution imaging and 3D analysis of Ag nanoparticles in cells with ToF-SIMS and delayed extraction. *Biointerphases* **13**, 03B410 (2018).
24. Rangarajan, S. & Tyler, B. J. Interpretation of static time-of-flight ion mass spectral images of adsorbed protein films on topographically complex surfaces. *Appl. Surf. Sci.* **231–232**, 406–410 (2004).
25. Kingshott, P., McArthur, S., Thissen, H., Castner, D. G. & Griesser, H. J. Ultrasensitive probing of the protein resistance of PEG surfaces by secondary

- ion mass spectrometry. *Biomaterials* **23**, 4775–4785 (2002).
26. Zheng, L. *et al.* Investigating lipid-lipid and lipid-protein interactions in model membranes by ToF-SIMS. *Appl. Surf. Sci.* **255**, 1190–1192 (2008).
 27. Hasegawa, U. *et al.* Estimation of protein adsorption on dialysis membrane by means of TOF-SIMS imaging. *J. Memb. Sci.* **236**, 91–99 (2004).
 28. Brison, J. *et al.* ToF-SIMS imaging and depth profiling of HeLa cells treated with bromodeoxyuridine. *Surf. Interface Anal.* **43**, 354–357 (2011).
 29. Lhoest, J. B., Wagner, M. S., Tidwell, C. D. & Castner, D. G. Characterization of adsorbed protein films by time of flight secondary ion mass spectrometry. *J. Biomed. Mater. Res.* **57**, 432–440 (2001).
 30. Wagner, M. S. & Castner, D. G. Analysis of adsorbed proteins by static time-of-flight secondary ion mass spectrometry. **232**, 366–376 (2004).
 31. Xia, N. & Castner, D. G. Preserving the structure of adsorbed protein films for time-of-flight secondary ion mass spectrometry analysis. *J. Biomed. Mater. Res. - Part A* **67**, 179–190 (2003).
 32. Okada, K. *et al.* Evaluation of immobilized-lysozyme by means of TOF-SIMS. *Appl. Surf. Sci.* **255**, 1104–1106 (2008).
 33. Aoyagi, S. *et al.* Structure analysis of immobilized-bovine serum albumin by means of TOF-SIMS. *e-Journal Surf. Sci. Nanotechnol.* **4**, 614–618 (2006).
 34. Baugh, L. *et al.* Probing the orientation of surface-immobilized protein G B1 using ToF-SIMS, sum frequency generation, and NEXAFS spectroscopy. *Langmuir* **26**, 16434–16441 (2010).

35. Park, J. W., Cho, I. H., Moon, D. W., Paek, S. H. & Lee, T. G. ToF-SIMS and PCA of surface-immobilized antibodies with different orientations. *Surf. Interface Anal.* **43**, 285–289 (2011).
36. Graham, D. J. & Castner, D. G. Multivariate analysis of ToF-SIMS data from multicomponent systems: The why, when, and how. *Biointerphases* **7**, 1–12 (2012).
37. Wagner, M. S. & Castner, D. G. Characterization of Adsorbed Protein Films by Time-of-Flight Secondary Ion Mass Spectrometry with Principal Component Analysis. *Langmuir* **17**, 4649–4660 (2001).
38. Wagner, M. S., Tyler, B. J. & Castner, D. G. Interpretation of static time-of-flight secondary ion mass spectra of adsorbed protein films by multivariate pattern recognition. *Anal. Chem.* **74**, 1824–1835 (2002).
39. Aoyagi, S. *et al.* Evaluation of Time-of-Flight Secondary Ion Mass Spectrometry Spectra of Peptides by Random Forest with Amino Acid Labels: Results from a Versailles Project on Advanced Materials and Standards Interlaboratory Study. *Anal. Chem.* **93**, 4191–4197 (2021).
40. Jones, E. A., Lockyer, N. P. & Vickerman, J. C. Depth profiling brain tissue sections with a 40 keV C60+ primary ion beam. *Anal. Chem.* **80**, 2125–2132 (2008).
41. Shard, A. G., Spencer, S. J., Smith, S. A., Havelund, R. & Gilmore, I. S. The matrix effect in organic secondary ion mass spectrometry. *Int. J. Mass Spectrom.* **377**, 599–609 (2015).
42. Solé-Domènech, S., Johansson, B., Schalling, M., Malm, J. & Sjövall, P.

- Analysis of opioid and amyloid peptides using time-of-flight secondary ion mass spectrometry. *Anal. Chem.* **82**, 1964–1974 (2010).
43. Schneider, P. *et al.* Direct Analysis of Ion-Induced Peptide Fragmentation in Secondary-Ion Mass Spectrometry. *Anal. Chem.* **92**, 15604–15610 (2020).
 44. Fletcher, J. S. Latest applications of 3D ToF-SIMS bio-imaging. *Biointerphases* **10**, 018902 (2015).
 45. Fletcher, J. S. & Vickerman, J. C. A new SIMS paradigm for 2D and 3D molecular imaging of bio-systems. *Anal. Bioanal. Chem.* **396**, 85–104 (2010).
 46. Fletcher, J. S. *et al.* A New Dynamic in Mass Spectral Imaging of Single Biological Cells. **80**, 9058–9064 (2008).
 47. Fletcher, J. S. Evaluation of biomolecular distributions in rat brain tissues by means of ToF-SIMS using a continuous beam of Ar clusters. **307**, 1–6 (2016).
 48. Tian, H. *et al.* Spatiotemporal lipid profiling during early embryo development of *Xenopus laevis* using dynamic ToF-SIMS imaging. (2016).
 49. Yokoyama, Y. *et al.* Peptide Fragmentation and Surface Structural Analysis by Means of ToF-SIMS Using Large Cluster Ion Sources. *Anal. Chem.* **88**, 3592–3597 (2016).
 50. Aoyagi, S. *et al.* Peptide structural analysis using continuous Ar cluster and C60ion beams. *Anal. Bioanal. Chem.* **405**, 6621–6628 (2013).
 51. Son, J. G. *et al.* Ar-gas cluster ion beam in ToF-SIMS for peptide and protein analysis. *Biointerphases* **15**, 021011 (2020).
 52. Rabbani, S., Barber, A. M., Fletcher, J. S., Lockyer, N. P. & Vickerman, J. C.

- TOF-SIMS with argon gas cluster ion beams: A comparison with C 60+.
Anal. Chem. **83**, 3793–3800 (2011).
53. Ogrinc Potočnik, N., Fisher, G. L., Prop, A. & Heeren, R. M. A. Sequencing and Identification of Endogenous Neuropeptides with Matrix-Enhanced Secondary Ion Mass Spectrometry Tandem Mass Spectrometry. *Anal. Chem.* **89**, 8223–8227 (2017).
54. Havelund, R. *et al.* Supplementary Protocol: Label-free Imaging of Biomolecules in Murine Brain Sections Using the 3D OrbiSIMS. *Nat. Methods* **14**, 1175–1183 (2017).
55. Zubarev, R. A. & Makarov, A. Orbitrap Mass Spectrometry. *Anal. Chem.* **85**, 5288–5296 (2013).
56. Meurs, J. *et al.* Sequential Orbitrap Secondary Ion Mass Spectrometry and Liquid Extraction Surface Analysis-Tandem Mass Spectrometry-Based Metabolomics for Prediction of Brain Tumor Relapse from Sample-Limited Primary Tissue Archives. *Anal. Chem.* **93**, 6947–6954 (2021).
57. Aoyagi, S. *et al.* OrbiSIMS Imaging Identifies Molecular Constituents of the Perialgal Vacuole Membrane of *Paramecium bursaria* with Symbiotic *Chlorella variabilis*. *Anal. Chem.* (2019).
58. Cliff, B., Lockyer, N. P., Corlett, C. & Vickerman, J. C. Development of instrumentation for routine ToF-SIMS imaging analysis of biological material. *Appl. Surf. Sci.* **203–204**, 730–733 (2003).
59. Sheraz Née Rabbani, S., Razo, I. B., Kohn, T., Lockyer, N. P. & Vickerman, J. C. Enhancing ion yields in time-of-flight-secondary ion mass spectrometry:

- A comparative study of argon and water cluster primary beams. *Anal. Chem.* **87**, 2367–2374 (2015).
60. Piwowar, A. M. *et al.* Effects of cryogenic sample analysis on molecular depth profiles with TOF-secondary ion mass spectrometry. *Anal. Chem.* **82**, 8291–8299 (2010).
61. Newell, C. L., Vorng, J. L., MacRae, J. I., Gilmore, I. S. & Gould, A. P. Cryogenic OrbiSIMS Localizes Semi-Volatile Molecules in Biological Tissues. *Angew. Chemie - Int. Ed.* **59**, 18194–18200 (2020).
62. Zhang, J. *et al.* Cryo-OrbiSIMS for 3D Molecular Imaging of a Bacterial Biofilm in Its Native State. *Anal. Chem.* **92**, 9008–9015 (2020).
63. Mitchell, P. In the pursuit of industrial proteomics Robust , reproducible ways to identify and quantify the protein complement of cells are still some way off , as are the means of turning these technologies into viable businesses . *Nat. Biotechnol.* **21**, 233–237 (2003).
64. Jonkheijm, P., Weinrich, D., Schröder, H., Niemeyer, C. M. & Waldmann, H. Chemical strategies for generating protein biochips. *Angew. Chemie - Int. Ed.* **47**, 9618–9647 (2008).
65. Sheridan, C. Protein chip companies turn to biomarkers. *Nat. Biotechnol.* **23**, 3–4 (2005).
66. Rusmini, F., Zhong, Z. & Feijen, J. Protein immobilization strategies for protein biochips. *Biomacromolecules* **8**, 1775–1789 (2007).
67. Wang, S. *et al.* Gradient lithography of engineered proteins to fabricate 2D and 3D cell culture microenvironments. *Biomed. Microdevices* **11**, 1127–1134

- (2009).
68. Menon, G. & Krishnan, J. Spatial localisation meets biomolecular networks. *Nat. Commun.* **12**, (2021).
 69. Singh, A. *et al.* Enhanced lubrication on tissue and biomaterial surfaces through peptide-mediated binding of hyaluronic acid. *Nat. Mater.* **13**, 988–995 (2014).
 70. Armstrong, J. P. K. *et al.* Artificial membrane-binding proteins stimulate oxygenation of stem cells during engineering of large cartilage tissue. *Nat. Commun.* **6**, (2015).
 71. Grainger, D. W. All charged up about implanted biomaterials. *Nat. Biotechnol.* **31**, 507–509 (2013).
 72. Doloff, J. C. *et al.* Colony stimulating factor-1 receptor is a central component of the foreign body response to biomaterial implants in rodents and non-human primates. *Nat. Mater.* **16**, 671–680 (2017).
 73. Bank, R. A. Limiting biomaterial fibrosis. *Nat. Mater.* **18**, 781 (2019).
 74. Kuivaniemi, H. & Tromp, G. Type III collagen (COL3A1): Gene and protein structure, tissue distribution, and associated diseases. *Gene* **707**, 151–171 (2019).
 75. Chiti, F. & Dobson, C. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu. Rev. Biochem.* **86**, 1–42 (2017).
 76. Ventola, C. L. Progress in Nanomedicine: Approved and Investigational Nanodrugs. *P T* **42**, 742–755 (2017).

77. Ehmann, F. *et al.* Next-generation nanomedicines and nanosimilars: EU regulators' initiatives relating to the development and evaluation of nanomedicines. *Nanomedicine* **8**, 849–856 (2013).
78. Bobo, D., Robinson, K. J., Islam, J., Thurecht, K. J. & Corrie, S. R. Nanoparticle-Based Medicines: A Review of FDA-Approved Materials and Clinical Trials to Date. *Pharm. Res.* **33**, 2373–2387 (2016).
79. Björnmalm, M., Thurecht, K. J., Michael, M., Scott, A. M. & Caruso, F. Bridging Bio-Nano Science and Cancer Nanomedicine. *ACS Nano* **11**, 9594–9613 (2017).
80. Nichols, J. W. & Bae, Y. H. Odyssey of a cancer nanoparticle: From injection site to site of action. *Nano Today* **7**, 606–618 (2012).
81. Bloom, A. N., Tian, H., Schoen, C. & Winograd, N. Label-free visualization of nilotinib-functionalized gold nanoparticles within single mammalian cells by C60- SIMS imaging. *Anal. Bioanal. Chem.* **409**, 3067–3076 (2017).
82. Herda, L. M., Hristov, D. R., Lo Giudice, M. C., Polo, E. & Dawson, K. A. Mapping of Molecular Structure of the Nanoscale Surface in Bionanoparticles. *J. Am. Chem. Soc.* jacs.6b12297 (2016).
83. Pietersz, G. A., Wang, X., Yap, M. L., Lim, B. & Peter, K. Therapeutic targeting in nanomedicine: The future lies in recombinant antibodies. *Nanomedicine* **12**, 1873–1889 (2017).
84. Chen, H. *et al.* Peptide functionalized gold nanoparticles : the influence of pH on binding efficiency. *Nanotechnology* **28**, 11pp (2017).
85. Hu, C. M. J., Fang, R. H., Luk, B. T. & Zhang, L. Polymeric

nanotherapeutics: Clinical development and advances in stealth functionalization strategies. *Nanoscale* **6**, 65–75 (2014).

86. Karlsson, J., Vaughan, H. J. & Green, J. J. Biodegradable Polymeric Nanoparticles for Therapeutic Cancer Treatments. *Annu. Rev. Chem. Biomol. Eng.* **9**, 105–127 (2018).
87. Perez, A., Hernández, R., Velasco, D., Voicu, D. & Mijangos, C. Poly (lactic-co-glycolic acid) particles prepared by microfluidics and conventional methods. Modulated particle size and rheology. *J. Colloid Interface Sci.* **441**, 90–97 (2015).
88. Ankrum, J. A. *et al.* Engineering cells with intracellular agent-loaded microparticles to control cell phenotype. *Nat. Protoc.* **9**, 233–245 (2014).
89. Teh, S.-Y., Lin, R., Hung, L.-H. & Lee, A. P. Droplet microfluidics. *Lab Chip* **8**, 198 (2008).
90. Abalde-Cela, S., Taladriz-Blanco, P., De Oliveira, M. G. & Abell, C. Droplet microfluidics for the highly controlled synthesis of branched gold nanoparticles. *Sci. Rep.* **8**, 1–6 (2018).
91. Holtze, C. Large-scale droplet production in microfluidic devices - An industrial perspective. *J. Phys. D. Appl. Phys.* **46**, (2013).
92. Shang, L., Cheng, Y. & Zhao, Y. Emerging Droplet Microfluidics. *Chem. Rev.* **117**, 7964–8040 (2017).
93. Valencia, P. M., Farokhzad, O. C., Karnik, R. & Langer, R. Microfluidic technologies for accelerating the clinical translation of nanoparticles. *Nat. Nanotechnol.* **7**, 623–629 (2012).

94. Lim, J. M. *et al.* Parallel microfluidic synthesis of size-tunable polymeric nanoparticles using 3D flow focusing towards in vivo study. *Nanomedicine Nanotechnology, Biol. Med.* **10**, 401–409 (2014).
95. Kong, T. *et al.* Droplet based microfluidic fabrication of designer microparticles for encapsulation applications. *Biomicrofluidics* **6**, (2012).
96. Xu, Q. *et al.* Preparation of monodisperse biodegradable polymer microparticles using a microfluidic flow-focusing device for controlled drug delivery. *Small* **5**, 1575–1581 (2009).
97. Kovalchuk, N. M. *et al.* Effect of surfactant on emulsification in microchannels. *Chem. Eng. Sci.* **176**, 139–152 (2018).
98. Wojnicki, M. *et al.* Micro droplet formation towards continuous nanoparticles synthesis. *Micromachines* **9**, 1–13 (2018).
99. Gañán-Calvo, A. M., Montanero, J. M., Martín-Banderas, L. & Flores-Mosquera, M. Building functional materials for health care and pharmacy from microfluidic principles and Flow Focusing. *Adv. Drug Deliv. Rev.* **65**, 1447–1469 (2013).
100. Seo, M. *et al.* Continuous Microfluidic Reactors for Polymer Particles. *Langmuir* **21**, 11614–11622 (2005).
101. Wang, J. *et al.* Droplet microfluidics for the production of microparticles and nanoparticles. *Micromachines* **8**, 1–23 (2017).
102. Dolgin, B. E. The tangled history of mRNA vaccines. *Nature* **597**, 318–324 (2021).
103. Min, Y., Caster, J. M., Eblan, M. J. & Wang, A. Z. Clinical Translation of

- Nanomedicine. *Chem. Rev.* **115**, 11147–11190 (2015).
104. Coty, J. B. & Vauthier, C. Characterization of nanomedicines: A reflection on a field under construction needed for clinical translation success. *J. Control. Release* **275**, 254–268 (2018).
105. Wilhelm, S. *et al.* Analysis of Nanoparticle Delivery to Tumours. *Nat. Rev. Mater.* **1**, 1–138 (2016).
106. Bobo, D., Robinson, K. J., Islam, J., Thurecht, K. J. & Corrie, S. R. Nanoparticle-Based Medicines: A Review of FDA-Approved Materials and Clinical Trials to Date. *Pharm. Res.* **33**, 2373–2387 (2016).
107. Jain, P. *et al.* In-vitro in-vivo correlation (IVIVC) in nanomedicine: Is protein corona the missing link? *Biotechnol. Adv.* (2017).
108. Anselmo, A. C., Prabhakarandian, B., Pant, K. & Mitragotri, S. Clinical and commercial translation of advanced polymeric nanoparticle systems: opportunities and material challenges. *Transl. Mater. Res.* **4**, 014001 (2017).
109. Lazarovits, J., Chen, Y. Y., Sykes, E. A. & Chan, W. C. W. Nanoparticle–blood interactions: the implications on solid tumour targeting. *Chem. Commun.* (2015).
110. Schäffler, M. *et al.* Blood protein coating of gold nanoparticles as potential tool for organ targeting. *Biomaterials* **35**, 3455–3466 (2014).
111. Ke, P. C., Lin, S., Parak, W. J., Davis, T. P. & Caruso, F. A Decade of the Protein Corona. *ACS Nano* **11**, 11773–11776 (2017).
112. Walkey, C. D. & Chan, W. C. W. Understanding and controlling the interaction of nanomaterials with proteins in a physiological environment.

- Chem. Soc. Rev.* **41**, 2780–2799 (2012).
113. Vroman, L. Effect of adsorbed proteins on the wettability of hydrophilic and hydrophobic solids. *Nature* **196**, 476–477 (1962).
114. Walczyk, D., Baldelli Bombelli, F., Monopoli, M. P., Lynch, I. & Dawson, K. A. What the cell ‘sees’ in bionanoscience. *J. Am. Chem. Soc.* **132**, 5761–5768 (2010).
115. Docter, D. *et al.* The nanoparticle biomolecule corona: lessons learned – challenge accepted? *Chem. Soc. Rev.* **44**, 6094–6121 (2015).
116. Boraschi, D. *et al.* The bio-nano-interface in predicting nanoparticle fate and behaviour in living organisms: Towards grouping and categorising nanomaterials and ensuring nanosafety by design. *BioNanoMaterials* **14**, 195–216 (2013).
117. Hadjidemetriou, M. & Kostarelos, K. Nanomedicine: Evolution of the nanoparticle corona. *Nat. Nanotechnol.* **12**, 288–290 (2017).
118. Dai, Q. *et al.* Particle Targeting in Complex Biological Media. *Adv. Healthc. Mater.* **7**, 1–32 (2018).
119. Lesniak, A. *et al.* Effects of the presence or absence of a protein corona on silica nanoparticle uptake and impact on cells. *ACS Nano* **6**, 5845–5857 (2012).
120. Laurent, S., Burtea, C., Thirifays, C., Rezaee, F. & Mahmoudi, M. Significance of cell ‘observer’ and protein source in nanobiosciences. *J. Colloid Interface Sci.* **392**, 431–445 (2013).
121. Westmeier, D., Chen, C., Stauber, R. H. & Docter, D. The bio-corona and its

- impact on nanomaterial toxicity. *Eur. J. Nanomedicine* **7**, 153–168 (2015).
122. Lin, S. *et al.* NanoEHS beyond toxicity – focusing on biocorona. *Environ. Sci. Nano* **4**, 1433–1454 (2017).
123. Ritz, S. *et al.* Protein Corona of Nanoparticles: Distinct Proteins Regulate the Cellular Uptake. *Biomacromolecules* **16**, 1311–1321 (2015).
124. Schöttler, S., Landfester, K. & Mailänder, V. Controlling the Stealth Effect of Nanocarriers through Understanding the Protein Corona. *Angew. Chemie - Int. Ed.* **55**, 8806–8815 (2016).
125. Dai, Q., Walkey, C. & Chan, W. C. W. Polyethylene glycol backfilling mitigates the negative impact of the protein corona on nanoparticle cell targeting. *Angew. Chemie - Int. Ed.* **53**, 5093–5096 (2014).
126. Barenholz, Y. Doxil® - The first FDA-approved nano-drug: Lessons learned. *J. Control. Release* **160**, 117–134 (2012).
127. Wang, R., Billone, P. S. & Mullett, W. M. Nanomedicine in Action : An Overview of Cancer Nanomedicine on the Market and in Clinical Trials. **2013**, (2013).
128. Mello, S. R. D. *et al.* The evolving landscape of drug products containing nanomaterials in the United States. *Nat. Nanotechnol.* **12**, 523–529 (2017).
129. Gulati, N. M., Stewart, P. L. & Steinmetz, N. F. Bioinspired Shielding Strategies for Nanoparticle Drug Delivery Applications. *Mol. Pharm.* **15**, 2900–2909 (2018).
130. Wibroe, P. P., Ahmadvand, D., Ali, M., Yagmur, A. & Moghimi, S. M. An integrated assessment of morphology , size , and complement activation of the

- PEGylated liposomal doxorubicin products Doxil[®], *J. Control. Release* **221**, 1–8 (2016).
131. García, K. P. *et al.* Zwitterionic-coated ‘stealth’ nanoparticles for biomedical applications: Recent advances in countering biomolecular corona formation and uptake by the mononuclear phagocyte system. *Small* **10**, 2516–2529 (2014).
132. Naidu, P. S. R. *et al.* The Protein Corona of PEGylated PGMA-Based Nanoparticles is Preferentially Enriched with Specific Serum Proteins of Varied Biological Function. *Langmuir* **33**, 12926–12933 (2017).
133. Papi, M. *et al.* Clinically approved PEGylated nanoparticles are covered by a protein corona that boosts the uptake by cancer cells. *Nanoscale* **9**, 10327–10334 (2017).
134. Schöttler, S. *et al.* Protein adsorption is required for stealth effect of poly(ethylene glycol)- and poly(phosphoester)-coated nanocarriers. *Nat. Nanotechnol.* **11**, 372–377 (2016).
135. Pulido-Reyes, G., Leganes, F., Fernández-Piñas, F. & Rosal, R. Bio-nano interface and environment: A critical review. *Environ. Toxicol. Chem.* **36**, 3181–3193 (2017).
136. Shannahan, J. The biocorona: A challenge for the biomedical application of nanoparticles. *Nanotechnol. Rev.* **6**, 345–353 (2017).
137. Müller, J. *et al.* Beyond the protein corona – lipids matter for biological response of nanocarriers. *Acta Biomater.* **71**, 420–431 (2018).
138. Maiolo, D., Del Pino, P., Metrangolo, P., Parak, W. J. & Baldelli Bombelli, F.

- Nanomedicine delivery: Does protein corona route to the target or off road?
Nanomedicine **10**, 3231–3247 (2015).
139. Rampado, R., Crotti, S., Caliceti, P., Pucciarelli, S. & Agostini, M. Recent Advances in Understanding the Protein Corona of Nanoparticles and in the Formulation of “Stealthy” Nanomaterials. *Front. Bioeng. Biotechnol.* **8**, 1–19 (2020).
140. Sakulkhu, U., Mahmoudi, M., Maurizi, L., Salaklang, J. & Hofmann, H. Protein corona composition of superparamagnetic iron oxide nanoparticles with various physico-Chemical properties and coatings. *Sci. Rep.* **4**, 1–9 (2014).
141. Stolnik, S. *et al.* Surface Modification of Poly(lactide-co-glycolide) Nanospheres by Biodegradable Poly(lactide)-Poly(ethylene glycol) Copolymers. *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists* **11**, 1800–1808 (1994).
142. Ding, H. ming & Ma, Y. qiang. Computer simulation of the role of protein corona in cellular delivery of nanoparticles. *Biomaterials* **35**, 8703–8710 (2014).
143. Yan, Y. *et al.* Differential roles of the protein corona in the cellular uptake of nanoporous polymer particles by monocyte and macrophage cell lines. *ACS Nano* **7**, 10960–10970 (2013).
144. Dawson, K. A. & Yan, Y. Current understanding of biological identity at the nanoscale and future prospects. *Nat. Nanotechnol.* **16**, 229–242 (2021).
145. Venditto, V. J. & Szoka, F. C. Cancer nanomedicines: So many papers and so

- few drugs! *Adv. Drug Deliv. Rev.* **65**, 80–88 (2013).
146. Doyle, S. Promise of nano revolution hasn't materialized. *CMAJ* **184**, 133–134 (2012).
147. Maeda, H. & Khatami, M. Analyses of repeated failures in cancer therapy for solid tumors : poor tumor - selective drug delivery , low therapeutic efficacy and unsustainable costs. *Clin. Transl. Med.* 1–20 (2018).
148. Faria, M. *et al.* Minimum information reporting in bio–nano experimental literature. *Nat. Nanotechnol.* **13**, 777–785 (2018).
149. Winzen, S. *et al.* Complementary analysis of the hard and soft protein corona: sample preparation critically effects corona composition. *Nanoscale* (2015).
150. Moore, T. L. *et al.* Nanoparticle colloidal stability in cell culture media and impact on cellular interactions. *Chem. Soc. Rev.* **44**, 6287–6305 (2015).
151. Docter, D. *et al.* Quantitative profiling of the protein coronas that form around nanoparticles. *Nat. Protoc.* **9**, 2030–2044 (2014).
152. Mahmoudi, M., Shokrgozar, M. A. & Behzadi, S. Slight temperature changes affect protein affinity and cellular uptake/toxicity of nanoparticles. *Nanoscale* **5**, 3240 (2013).
153. Shang, L. & Nienhaus, G. U. In Situ Characterization of Protein Adsorption onto Nanoparticles by Fluorescence Correlation Spectroscopy. *Acc. Chem. Res.* (2017).
154. Bonvin, D., Chiappe, D., Moniatte, M., Hofmann, H. & Mionić Ebersold, M. Methods of protein corona isolation for magnetic nanoparticles. *Analyst* **142**, 3805–3815 (2017).

155. Pederzoli, F. *et al.* Protein corona and nanoparticles: how can we investigate on? *Wiley Interdiscip. Rev. Nanomedicine Nanobiotechnology* **9**, 1–23 (2017).
156. Carrillo-Carrion, C., Carril, M. & Parak, W. J. Techniques for the experimental investigation of the protein corona. *Curr. Opin. Biotechnol.* (2017).
157. Pink, M., Verma, N., Kersch, C. & Schmitz-Spanke, S. Identification and characterization of small organic compounds within the corona formed around engineered nanoparticles. *Environ. Sci. Nano* **5**, 1420–1427 (2018).
158. Hellstrand, E. *et al.* Complete high-density lipoproteins in nanoparticle corona. *FEBS J.* **276**, 3372–3381 (2009).
159. Baer, D. R., Gaspar, D. J., Nachimuthu, P., Techane, S. D. & Castner, D. G. Application of surface chemical analysis tools for characterization of nanoparticles. *Anal. Bioanal. Chem.* **396**, 983–1002 (2010).
160. Bekdemir, A. & Stellacci, F. A centrifugation-based physicochemical characterization method for the interaction between proteins and nanoparticles. *Nat. Commun.* **7**, 1–8 (2016).
161. Nel, A. E. *et al.* Understanding biophysicochemical interactions at the nano-bio interface. *Nat. Mater.* **8**, 543–557 (2009).
162. Dhandayuthapani, B. & Sakthi, D. kumar. Biomaterials for Biomedical Applications. *Biomed. Appl. Polym. Mater. Compos.* **1706635**, 1–20 (2016).
163. Patterson, J., Stayton, P. S. & Li, X. In situ characterization of the degradation of PLGA microspheres in hyaluronic acid hydrogels by optical coherence tomography. *IEEE Trans. Med. Imaging* **28**, 74–81 (2009).

164. Okada, K. *et al.* Evaluation of immobilized-lysozyme by means of TOF-SIMS. *Appl. Surf. Sci.* **255**, 1104–1106 (2008).
165. Di Palma, G. *et al.* Reversible, high-affinity surface capturing of proteins directed by supramolecular assembly. *ACS Appl. Mater. Interfaces* **11**, 8937–8944 (2019).
166. Gnaser, H. *et al.* Peptide dissociation patterns in secondary ion mass spectrometry under large argon cluster ion bombardment. *Rapid Commun. Mass Spectrom.* **27**, 1490–1496 (2013).
167. Chu, I. K. *et al.* Proposed nomenclature for peptide ion fragmentation. *Int. J. Mass Spectrom.* **390**, 24–27 (2015).
168. Steen, H. & Mann, M. The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* **5**, 699–711 (2004).
169. Samgina, T. Y. *et al.* Mass spectrometric de novo sequencing of natural non-tryptic peptides : comparing peculiarities of collision-induced dissociation (CID) and high energy collision dissociation (HCD). 2595–2604 (2014).
170. Scigelova, M. & Makarov, A. Orbitrap mass analyzer - Overview and applications in proteomics. *Proteomics* **1**, 16–21 (2006).
171. Na, S. *et al.* Characterization of disulfide bonds by planned digestion and tandem mass spectrometry. *Mol. Biosyst.* **11**, 1156–1164 (2015).
172. Tsai, P. L., Chen, S. F. & Huang, S. Y. Mass spectrometry-based strategies for protein disulfide bond identification. *Rev. Anal. Chem.* **32**, 257–268 (2013).
173. Kotowska, A. M. *et al.* Protein identification by 3D OrbiSIMS to facilitate in situ imaging and depth profiling. *Nat. Commun.* **11**, 1–8 (2020).

174. Zhurov, K. O., Fornelli, L., Wodrich, M. D., Laskay, Ü. A. & Tsybin, Y. O. Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis. *Chem. Soc. Rev.* **42**, 5014–5030 (2013).
175. Sakakura, M. & Takayama, M. In-source decay and fragmentation characteristics of peptides using 5-aminosalicylic acid as a matrix in matrix-assisted laser desorption/ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **21**, 979–988 (2010).
176. Asakawa, D., Calligaris, D., Zimmerman, T. A. & Pauw, E. De. In-source decay during matrix-assisted laser desorption/ionization combined with the collisional process in an FTICR mass spectrometer. *Anal. Chem.* **85**, 7809–7817 (2013).
177. Asakawa, D., Takahashi, H., Iwamoto, S. & Tanaka, K. Fundamental study of hydrogen-attachment-induced peptide fragmentation occurring in the gas phase and during the matrix-assisted laser desorption/ionization process. *Phys. Chem. Chem. Phys.* **20**, 13057–13067 (2018).
178. Jacob, E. & Unger, R. A tale of two tails: Why are terminal residues of proteins exposed? *Bioinformatics* **23**, 225–230 (2007).
179. Riley, N. M. *et al.* The negative mode proteome with activated ion negative electron transfer dissociation (AI-NETD). *Mol. Cell. Proteomics* **14**, 2644–2660 (2015).
180. Tran, T. T. N., Wang, T., Hack, S., Hoffmann, P. & Bowie, J. H. Can collision-induced negative-ion fragmentations of [M-H]⁻ anions be used to identify phosphorylation sites in peptides? *Rapid Commun. Mass Spectrom.*

- 25**, 3537–3548 (2011).
181. Lee, J. *et al.* One-Step Peptide Backbone Dissociations in Negative-Ion Free Radical Initiated Peptide Sequencing Mass Spectrometry. *Anal. Chem* **85**, 7044–7051 (2013).
182. Dong, N. P., Zhang, L. X. & Liang, Y. Z. A comprehensive investigation of proline fragmentation behavior in low-energy collision-induced dissociation peptide mass spectra. *Int. J. Mass Spectrom.* **308**, 89–97 (2011).
183. Harper, B., Miladi, M. & Solouki, T. Loss of internal backbone carbonyls: Additional evidence for sequence-scrambling in collision-induced dissociation of γ -type ions. *J. Am. Soc. Mass Spectrom.* **25**, 1716–1729 (2014).
184. Bateman, A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).
185. Tyler, B. J., Bruening, C., Rangaranjan, S. & Arlinghaus, H. F. TOF-SIMS imaging of adsorbed proteins on topographically complex surfaces with Bi_3^+ primary ions. *Biointerphases* **6**, 135–141 (2011).
186. Ray, S. *et al.* Neutralized chimeric avidin binding at a reference biosensor surface. *Langmuir* **31**, 1921–1930 (2015).
187. Vasilev, K. *et al.* Creating gradients of two proteins by differential passive adsorption onto a PEG-density gradient. *Biomaterials* **31**, 392–397 (2010).
188. Mathé, C. *et al.* Structural determinants for protein adsorption/non-adsorption to silica surface. *PLoS One* **8**, 1–13 (2013).
189. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* (80-.). **347**, (2015).

190. McDonald, L., Robertson, D. H. L., Hurst, J. L. & Beynon, R. J. Positional proteomics: Selective recovery and analysis of N-terminal proteolytic peptides. *Nat. Methods* **2**, 955–957 (2005).
191. Sharma, S. & Schiller, M. The carboxy-terminus, a key regulator of protein function. *Crit Rev Biochem Mol Biol.* **54**, 85–102 (2019).
192. Bahir, I. & Linial, M. ProTeus: Identifying signatures in protein termini. *Nucleic Acids Res.* **33**, 277–280 (2005).
193. Deng, J. R. *et al.* N-Terminal selective modification of peptides and proteins using 2-ethynylbenzaldehydes. *Commun. Chem.* **3**, 1–9 (2020).
194. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
195. Liebisch, G. *et al.* Update on LIPID MAPS classification, nomenclature, and shorthand notation for MS-derived lipid structures. *J. Lipid Res.* **61**, 1539–1555 (2020).
196. Parkinson, E. *et al.* Proteomic analysis of the human skin proteome after In Vivo treatment with sodium dodecyl sulphate. *PLoS One* **9**, (2014).
197. Dengjel, J., Bruckner-Tuderman, L. & Nyström, A. Skin proteomics—analysis of the extracellular matrix in health and disease. *Expert Rev. Proteomics* **17**, 377–391 (2020).
198. Bliss, E., Heywood, W. E., Benatti, M., Sebire, N. J. & Mills, K. An optimised method for the proteomic profiling of full thickness human skin. *Biol. Proced. Online* **18**, 1–7 (2016).
199. Starr, N. J. *et al.* Enhanced vitamin C skin permeation from supramolecular

- hydrogels, illustrated using in situ ToF-SIMS 3D chemical profiling. *Int. J. Pharm.* **563**, 21–29 (2019).
200. Hofman, K., Hall, B., Cleaver, H. & Marshall, S. High-throughput quantification of hydroxyproline for determination of collagen. *Anal. Biochem.* **417**, 289–291 (2011).
201. Shard, A. G. A straightforward method for interpreting XPS data from core-shell nanoparticles. *J. Phys. Chem. C* **116**, 16806–16813 (2012).
202. Ganem, B. Studies on Heme Binding in Myoglobin, Hemoglobin, and Cytochrome c by Ion Spray Mass Spectrometry. (1993).
203. Schmidt, A. & Karas, M. The Influence of Electrostatic Interactions on the Detection of Heme-Globin Complexes in ESI-MS. *J. Am. Soc. Mass Spectrom.* **0305**, 0–6 (2001).
204. Gledhill, M. The detection of iron protoporphyrin (heme b) in phytoplankton and marine particulate material by electrospray ionisation mass spectrometry - comparison with diode array detection. *Anal. Chim. Acta* **841**, 33–43 (2014).
205. Shi, Y. & Zhang, Y. Bioinorganic Chemistry Mechanisms of HNO Reactions with Ferric Heme Proteins. 16654–16658 (2018).
206. Cianetti, S., Ne, M., Vos, M. H., Martin, J. & Kruglik, S. G. Photodissociation of Heme Distal Methionine in Ferrous Cytochrome c Revealed by Subpicosecond Time-Resolved Resonance Raman Spectroscopy. *J. Am. Chem. Soc.* **126**, 13932–13933 (2004).
207. SHIMMA, S. & SETOU, M. Mass Microscopy to Reveal Distinct Localization of Heme B (m/z 616) in Colon Cancer Liver Metastasis. *J. Mass*

- Spectrom. Soc. Jpn.* **55**, 145–148 (2011).
208. Gemperline, E. & Li, L. MALDI-Mass Spectrometric Imaging for the Investigation of Metabolites in *Medicago truncatula* Root Nodules. 1–11 (2014).
209. Pashynska, V. A., Van Den Heuvel, H., Claeys, M. & Kosevich, M. V. Characterization of noncovalent complexes of antimalarial agents of the artemisinin-type and Fe(III)-heme by electrospray mass spectrometry and collisional activation tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* **15**, 1181–1190 (2004).
210. Mazel, V. *et al.* Identification of ritual blood in African artifacts using TOF-SIMS and synchrotron radiation microspectroscopies. *Anal. Chem.* **79**, 9253–9260 (2007).
211. Suo, Z., Avci, R., Schweitzer, M. H. & Deliorman, M. Porphyrin as an Ideal Biomarker in the Search for Extraterrestrial Life. *Astrobiology* **7**, 605–615 (2007).
212. Makadia, H. K. & Siegel, S. J. Poly Lactic-co-Glycolic Acid (PLGA) as biodegradable controlled drug delivery carrier. *Polymers (Basel)*. **3**, 1377–1397 (2011).
213. Vladisavljević, G. T., Duncanson, W. J., Shum, H. C. & Weitz, D. A. Emulsion templating of poly(lactic acid) particles: Droplet formation behavior. *Langmuir* **28**, 12948–12954 (2012).
214. Mark, D., Haeberle, S., Roth, G., von Stetten, F. & Zengerle, R. From microfluidic application to nanofluidic phenomena issue. *Chem. Soc. Rev.* **39**,

- 1153–82 (2010).
215. Bally, F. *et al.* Improved size-tunable preparation of polymeric nanoparticles by microfluidic nanoprecipitation. *Polym. (United Kingdom)* **53**, 5045–5051 (2012).
216. Donno, R. *et al.* Nanomanufacturing through microfluidic-assisted nanoprecipitation: Advanced analytics and structure-activity relationships. *Int. J. Pharm.* **534**, 97–107 (2017).
217. Anderson, J. M. & Shive, M. S. Biodegradation and biocompatibility of PLA and PLGA microspheres. *Adv. Drug Deliv. Rev.* **64**, 72–82 (2012).
218. Blasi, P. Poly(lactic acid)/poly(lactic-co-glycolic acid)-based microparticles: an overview. *J. Pharm. Investig.* **49**, 337–346 (2019).
219. Ocal, H., Arca-Yegin, B., Vural, I., Goracinova, K. & Cals, S. 5-Fluorouracil-loaded PLA/PLGA PEG–PPG–PEG polymeric nanoparticles: formulation, in vitro characterization and cell culture studies. **9045**, 560–567 (2014).
220. Halayqa, M. & Domańska, U. PLGA biodegradable nanoparticles containing perphenazine or chlorpromazine hydrochloride: Effect of formulation and release. *Int. J. Mol. Sci.* **15**, 23909–23923 (2014).
221. Scholes, P. D. *et al.* Detection and determination of surface levels of poloxamer and PVA surfactant on biodegradable nanospheres using SSIMS and XPS. *J. Control. Release* **59**, 261–278 (1999).
222. Saadati, R. & Dadashzadeh, S. Marked effects of combined TPGS and PVA emulsifiers in the fabrication of etoposide-loaded PLGA-PEG nanoparticles : In vitro and in vivo evaluation. *Int. J. Pharm.* **464**, 135–144 (2014).

223. Anna, S. L., Bontoux, N. & Stone, H. A. Formation of dispersions using ““ flow focusing ”” in microchannels. **82**, 364–366 (2003).
224. Zhao, C. X. Multiphase flow microfluidics for the production of single or multiple emulsions for drug delivery. *Adv. Drug Deliv. Rev.* **65**, 1420–1446 (2013).
225. Robinson, M. A. & Castner, D. G. Characterization of sample preparation methods of NIH/3T3 fibroblasts for ToF-SIMS analysis. *Biointerphases* **8**, 1–12 (2013).
226. Dimovska Nilsson, K., Karagianni, A., Kaya, I., Henricsson, M. & Fletcher, J. S. $(\text{CO}_2)_n^+$, $(\text{H}_2\text{O})_n^+$, and $(\text{H}_2\text{O})_n^+(\text{CO}_2)$ gas cluster ion beam secondary ion mass spectrometry: analysis of lipid extracts, cells, and Alzheimer’s model mouse brain tissue. *Anal. Bioanal. Chem.* 4181–4194 (2021).