Proteomic Investigation of Salivary Biomarkers in Periodontal Diseases

Mulli, Tonnie Kituku

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to:
  • Share: to copy, distribute and transmit the work

Under the following conditions:
  • Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
  • Non Commercial: You may not use this work for commercial purposes.
  • No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waivered if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Title: Proteomic Investigation of Salivary Biomarkers in Periodontal Diseases

Author: Tonnie Mulli

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

**END USER LICENSE AGREEMENT**

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. http://creativecommons.org/licenses/by-nc-nd/3.0/

You are free to:
- Share: to copy, distribute and transmit the work

Under the following conditions:
- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

**Take down policy**

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
PROTEOMIC INVESTIGATION OF
SALIVARY BIOMARKERS IN
PERIODONTAL DISEASES

Tonnie Kituku Mulli
(BDS, M Clin Dent Periodontology)

Thesis Submitted to King’s College London for the Degree of
Doctor of Philosophy (PhD)

Unit of Periodontology, Department of Restorative Dentistry
Dental Institute
King’s College London

London

2012
ABSTRACT

Introduction: The aim of the study was to investigate novel biomarkers of periodontal disease present in GCF and saliva.

Methods: The identities of specific putative biomarkers identified by SELDI-TOF MS were confirmed using 1-D PAGE coupled with LC-MS/MS. These biomarkers were then tested using ELISA in the samples that had been initially screened using SELDI-TOF MS as well as in saliva samples in an independent cohort selected according to diabetic status rather than periodontal status. The effect of non-surgical periodontal therapy, diabetes, saliva sample collection method, eating, time of day, day of the week and storage conditions on the salivary concentration of human neutrophil peptide 1-3 (HNP 1-3), cathelicidin LL-37 and protein S100A8 were tested. Copy number variation for HNP 1-3 gene, DEFA1A3 was tested in an independent cohort and the data analysed for association with presence and degree of periodontal disease.

Results: Antimicrobial peptides, HNP 1-3, LL-37 and S100A8 were identified. GCF and salivary HNP 1-3 and LL-37 concentrations were significantly elevated in periodontitis than in gingivitis or health. LL-37 was significantly elevated in aggressive periodontitis (AgP) compared with chronic periodontitis (CP) and gingivitis/health. Periodontal therapy reduced the salivary levels of HNP 1-3. Diabetes, eating and day of week had no significant effect on AMP levels. Saliva collection using spitting method yielded significantly more AMPs than using Salivettes®. Sample collection at 16:00 hours yielded significantly higher amounts of AMPs than samples collected at other times. Sample storage at room temperature (RT) for up to 48 hours had no significant impact on the salivary HNP 1-3 and LL-37 concentration while a proteinase inhibitor did not did improve the recovery of the AMPs after 72 hours. ≤ 6 copies of DEFA1A3 copy numbers were partly associated with periodontal disease.

Conclusion: GCF and salivary HNP 1-3 and LL-37 are potential novel biomarkers capable of differentiating between periodontal disease from health with high specificity and sensitivity. They may be useful for point-of-care diagnostics, screening for and monitoring periodontal health status as well as in epidemiologic studies. They may also be useful in understanding the pathogenesis of periodontal disease.
ACKNOWLEDGEMENTS

Firstly I would like to sincerely thank my supervisor, Prof Francis Hughes who has been like a true caring patriarch to me throughout my studies in the UK. His kindness and open door attitude has made my studies and work such an enjoyable exercise. The sacrifices he has made will never be forgotten.

Many thanks to the Head of Department and my second supervisor Dr. Mark Ide, for his valued guidance and to Dr. Ian McKay and Dr. Fiona M Clarke for steadying my hands in my initial days in the laboratory world.

I would also like to acknowledge Dr. Penny Hodge and her colleagues for generously donating their sample collection of saliva.

I would also like to appreciate the friendship and encouragements from fellow PhD students both at QMUL and KCL.

If I had any strength, if I had any encouragement, it is thanks to those hopeful eyes of my lovely children Joseph and Joy and the backbone of my life, my dear wife Ruth. I will never be able to thank you enough.

Finally and most important, is my gratefulness to God. Without Him, all would be in vain.

This work has been made possible by support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust, from King’s College London, Queen Mary University of London and Glaxo SmithKline Consumer Health Care plc. and the Oral and Dental Research Trust.
# TABLE OF CONTENTS

| LIST OF FIGURES | 8 |
| LIST OF TABLES | 15 |
| LIST OF SUPPLIERS | 18 |
| ABBREVIATIONS | 19 |

## CHAPTER 1 - LITERATURE REVIEW

1.1 INTRODUCTION .................................................. 24
   1.1.1 DISEASES OF THE PERIODONTIUM ....................... 25
   1.1.2 AETIOLOGY OF PERIODONTAL DISEASES .................. 26
   1.1.3 DIAGNOSIS AND CLASSIFICATION OF PERIODONTITIS .... 27
   1.1.3.1 GINGIVITIS ........................................... 28
   1.1.3.2 CHRONIC PERIODONTITIS ............................ 29
   1.1.3.3 AGGRESSIVE PERIODONTITIS ......................... 30
   1.1.3.4 CASE DEFINITIONS FOR PERIODONTITIS ............. 31
   1.1.3.5 USE OF CURRENT CLINICAL PARAMETERS FOR PERIODONTAL DIAGNOSIS 32
   1.1.4 DIAGNOSTICS AND BIOMARKERS ......................... 33
      1.1.4.1 WHAT ARE BIOMARKERS? ............................ 34
      1.1.4.2 THE PROCESS OF BIOMARKER DISCOVERY .......... 35
      1.1.4.3 ORAL FLUIDS AS SOURCES OF BIOMARKERS ....... 36
      1.1.4.4 SALIVA ............................................ 37
      1.1.4.5 GCF ............................................... 38
      1.1.4.6 RECENT STUDIES ON BIOMARKER DISCOVERY .......... 39
   1.1.5 PROTEINChip® TECHNOLOGY ............................. 40
   1.1.6 ANTIMICROBIAL PEPTIDES .............................. 41
1.2 SUMMARY AND AIMS ........................................... 42
   1.2.1 SPECIFIC OBJECTIVES .................................. 43

## CHAPTER 2 - PROTEIN IDENTIFICATION

2.1 INTRODUCTION .................................................. 44
   2.1.1 BOTTOM UP APPROACH .................................. 45
   2.1.2 TOP-DOWN APPROACH ................................... 46
   2.1.3 POTENTIAL CHALLENGES OF PROTEIN IDENTIFICATION .. 47
   2.1.3.1 ION EXCHANGE CHROMATOGRAPHY (IEC) ............. 48
   2.1.3.2 IMMobilized METAL AFFINITY CHROMATOGRAPHY (IMAC) 49
   2.2 AIMS ..................................................... 50
2.3 METHODS AND MATERIALS ..................................... 51
   2.3.1 PATIENT COHORT ....................................... 52
   2.3.2 SAMPLES ................................................ 53
   2.3.3 SALIVA AND GCF CLARIFICATION AND ENRICHMENT ...... 54
   2.3.3.1 FRACTIONATION .................................... 55
   2.3.3.2 CONFIRMATION OF PEAKS USING SELDI-TOF MS .... 56
   2.3.4 PROTEIN IDENTIFICATION USING LC-MS/MS ........... 57

4
# TABLE OF CONTENTS

2.3.4.1 **Determining the extent and quality of protein detection by LC-MS/MS**  
1D Gel SDS-page  
In-gel digestion  
LC-MS/MS  
Data analysis  
2.3.4.2 Protein detection in specific molecular weight bands of CM10, PH6 fractions.  

2.4 Results  
2.4.1 Saliva and GCF clarification and enrichment  
2.4.2 Protein identification using LC-MS/MS  
2.4.2.1 The extent and quality of protein detection by LC-MS/MS.  
2.4.2.2 Identified salivary proteins in gel band regions of 3-6kDa and 10-17kDa  

2.5 Discussion  

CHAPTER 3 – PRELIMINARY BIOMARKER VALIDATION  
3.1 Introduction  
3.1.1 Aims and objectives  
3.2 Methodology and materials  
3.2.1 Ethics approval  
3.2.2 Patient cohort and sample acquisition.  
3.2.2.1 Samples  
3.2.2.2 Diagnostic cohort  
3.2.2.3 Prognostic cohort  
3.2.3 Total protein quantification  
3.2.4 ELISAs  
3.2.5 Data analysis  
3.3 Results  
3.3.1 Diagnostic cohort  
3.3.1.1 Total protein  
3.3.1.2 HNP 1-3  
3.3.1.3 LL-37  
3.3.1.4 HBD-3  
3.3.1.5 IL-1α  
3.3.2 Prognostic cohort  
3.3.2.1 Total protein quantification  
3.3.2.2 HNP 1-3  
3.3.2.3 LL-37  
3.3.2.4 HBD-3  
3.3.2.5 IL-1α  
3.4 Discussion  
3.4.1 Conclusion  

CHAPTER 4 VALIDATION IN AN INDEPENDENT COHORT  
4.1 Introduction
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.1</td>
<td>AIMS AND OBJECTIVES</td>
<td>138</td>
</tr>
<tr>
<td>4.2</td>
<td>MATERIALS &amp; METHODS</td>
<td>139</td>
</tr>
<tr>
<td>4.2.1</td>
<td>EXCLUSION CRITERIA</td>
<td>139</td>
</tr>
<tr>
<td>4.2.2</td>
<td>SAMPLE ACQUISITION</td>
<td>139</td>
</tr>
<tr>
<td>4.2.3</td>
<td>TOTAL PROTEIN QUANTIFICATION</td>
<td>140</td>
</tr>
<tr>
<td>4.2.4</td>
<td>ELISAS</td>
<td>140</td>
</tr>
<tr>
<td>4.2.5</td>
<td>DATA ANALYSIS</td>
<td>141</td>
</tr>
<tr>
<td>4.2.5.1</td>
<td>CASE DEFINITIONS</td>
<td>141</td>
</tr>
<tr>
<td>4.2.5.2</td>
<td>DATA ANALYSIS</td>
<td>141</td>
</tr>
<tr>
<td>4.3</td>
<td>RESULTS</td>
<td>143</td>
</tr>
<tr>
<td>4.3.1</td>
<td>DEMOGRAPHIC AND CLINICAL SUMMARY OF SAMPLE POPULATION</td>
<td>143</td>
</tr>
<tr>
<td>4.3.2</td>
<td>THE APPLICATION OF PREVIOUSLY DETERMINED HNP 1-3 AND LL-37 CUT-OFF POINTS TO DIAGNOSE PERIODONTAL DISEASE AMONG THE NEW COHORT.</td>
<td>144</td>
</tr>
<tr>
<td>4.3.3</td>
<td>INFLUENCE OF DIABETIC STATUS ON ANTIMICROBIAL PEPTIDE LEVELS.</td>
<td>145</td>
</tr>
<tr>
<td>4.3.4</td>
<td>SALIVARY ANTIMICROBIAL PEPTIDE LEVELS BASED ON PERIODONTAL DIAGNOSIS</td>
<td>148</td>
</tr>
<tr>
<td>4.3.4.1</td>
<td>HNP 1-3</td>
<td>148</td>
</tr>
<tr>
<td>4.3.4.2</td>
<td>LL-37</td>
<td>151</td>
</tr>
<tr>
<td>4.3.4.3</td>
<td>S100A8</td>
<td>154</td>
</tr>
<tr>
<td>4.3.4.4</td>
<td>TOTAL PROTEIN</td>
<td>157</td>
</tr>
<tr>
<td>4.3.5</td>
<td>CORRELATION BETWEEN AMPS AND TOTAL PROTEIN CONCENTRATIONS IN SALIVA.</td>
<td>159</td>
</tr>
<tr>
<td>4.3.6</td>
<td>TESTING THE DIAGNOSTIC UTILITY OF COMBINED MARKER</td>
<td>163</td>
</tr>
<tr>
<td>4.4</td>
<td>DISCUSSION</td>
<td>170</td>
</tr>
<tr>
<td>4.4.3</td>
<td>CONCLUSION</td>
<td>174</td>
</tr>
<tr>
<td>CHAPTER 5 - TRANSLATIONAL STUDIES</td>
<td></td>
<td>175</td>
</tr>
<tr>
<td>5.1</td>
<td>INTRODUCTION</td>
<td>176</td>
</tr>
<tr>
<td>5.2</td>
<td>AIMS</td>
<td>177</td>
</tr>
<tr>
<td>5.2.1</td>
<td>OBJECTIVES</td>
<td>177</td>
</tr>
<tr>
<td>5.3</td>
<td>MATERIALS AND METHODS</td>
<td>178</td>
</tr>
<tr>
<td>5.3.1</td>
<td>ETHICS APPROVAL</td>
<td>178</td>
</tr>
<tr>
<td>5.3.2</td>
<td>PATIENT COHORT AND SAMPLE ACQUISITION</td>
<td>178</td>
</tr>
<tr>
<td>5.3.3</td>
<td>SAMPLING AND SAMPLE COLLECTION</td>
<td>178</td>
</tr>
<tr>
<td>5.3.3.1</td>
<td>EFFECT OF COLLECTION METHOD ON SALIVARY AMPS RECOVERY</td>
<td>179</td>
</tr>
<tr>
<td>5.3.3.2</td>
<td>EFFECT OF DIURNAL VARIATION AND EATING ON SALIVARY AMPS CONCENTRATION</td>
<td>179</td>
</tr>
<tr>
<td>5.3.3.3</td>
<td>SALIVARY AMP STABILITY UNDER DIFFERENT STORAGE TEMPERATURE AND PROTEINASE INHIBITOR.</td>
<td>179</td>
</tr>
<tr>
<td>5.3.4</td>
<td>TOTAL PROTEIN QUANTIFICATION</td>
<td>180</td>
</tr>
<tr>
<td>5.3.5</td>
<td>ELISAS</td>
<td>180</td>
</tr>
<tr>
<td>5.3.6</td>
<td>STATISTICAL ANALYSIS</td>
<td>180</td>
</tr>
<tr>
<td>5.4</td>
<td>RESULTS</td>
<td>182</td>
</tr>
<tr>
<td>5.4.1</td>
<td>EFFECT OF COLLECTION METHOD ON THE RECOVERY OF AMPS IN WHOLE SALIVA.</td>
<td>182</td>
</tr>
<tr>
<td>5.4.1.1</td>
<td>HNP 1-3</td>
<td>182</td>
</tr>
<tr>
<td>5.4.1.2</td>
<td>LL-37</td>
<td>182</td>
</tr>
<tr>
<td>5.4.2</td>
<td>THE EFFECT OF SAMPLE STORAGE TEMPERATURE AND PROTEINASE INHIBITOR ON SALIVARY AMPS.</td>
<td>186</td>
</tr>
<tr>
<td>5.4.2.1</td>
<td>THE EFFECT OF SAMPLE STORAGE TEMPERATURE ON SALIVARY AMPS.</td>
<td>186</td>
</tr>
</tbody>
</table>
# 5.4.3 The Effect of Diurnal Variations on Salivary AMP Concentrations

## 5.4.3.1 Total Protein

## 5.4.3.2 HNP 1-3

## 5.4.3.3 LL-37

## 5.4.3.4 S100A8

# 5.4.4 The Effect of Eating on Salivary AMP Concentrations

# 5.5 Discussion

## 5.5.1 Conclusion

---

# Chapter 6 - Genomic Studies on DEFA1A3

## 6.1 Introduction

## 6.2 Aims

## 6.2.1 Objectives

## 6.3 Methodology

### 6.3.1 Sample Description

### 6.3.2 Genomic DNA Extraction and Quality Control

### 6.3.3 qPCR

## 6.4 Results

## 6.5 Discussion

### 6.5.1 Conclusion

# Chapter 7 - Final Discussion, Future Studies and Conclusions

## 7.1 Final Discussion

## 7.1.1 Future Studies

## 7.1.2 Conclusions

# Chapter 8 – Appendices

## 8.1 Chapter 2: Protein Identification

### 8.1.1 Biomarker Discovery Using SELDI-TOF MS

### 8.1.2 Prefractionation Spin Column Protocols

### 8.1.2.1 IMAC 30 Spin Column

### 8.1.2.2 CM10 Spin Column

### 8.1.3 Confirmation of Peaks Using SELDI-TOF MS

### 8.1.3.1 ProteinChip Array Treatment and Buffers

### 8.1.4 Stacking Gel Preparation for Membrane Proteomics

### 8.1.5 Protein Identification Using LC-MS/MS

## 8.2 Chapter 3: Preliminary Biomarker Validation

### 8.2.1 ELISA Protocols

### 8.2.1.1 HNP 1-3

### 8.2.1.2 HBD3

### 8.2.1.3 IL-1A

### 8.2.1.4 LL-37

## 8.3 Chapter 5: Translational Studies

# Chapter 9 - References
Chapter 1

Figure 1-1: The sources of whole saliva (WS) constituents. Adapted from (Kaufman and Lamster, 2000b). .................................................................................................................................................. 41
Figure 1-2: Flow chart of the biomarker discovery process. Adapted from (Ciphergen, 2007). 42
Figure 1-3: Venn diagram showing the overlapping protein identification among plasma, whole saliva and parotid/SM/SL. Adapted from (Loo et al., 2010). ........................................................................................................................................ 44
Figure 1-4: A list of the potential biomarkers for periodontal diagnosis. ............................................................................................................................................... 48
Figure 1-5: The SELDI process using a ProteinChip array and TOF-MS (Bio-Rad Laboratories, 2007) 51
Figure 1-6 ProteinChip CM10 array surface chemistry with protein. Adapted from (Bio-Rad Laboratories, 2007). .................................................................................................................................................. 52
Figure 1-7: ProteinChip IMAC30 array metal-binding surface with metal and protein. Adapted from (Bio-Rad Laboratories, 2007) ........................................................................................................................................ 53
Figure 1-8: Comparative IMAC30 ProteinChip array tracings for periodontitis and gingivitis in GCF showing higher relative peak intensity of peptides corresponding to HNP 2, 1 and 3 respectively in periodontitis. ........................................................................................................................................ 55

Chapter 2

Figure 2-1: Gold (Au) ProteinChip array showing the map for the different samples based on CM10 fractions (Above) and IMAC30 fractions (Below). From these representative tracings, it can be seen that the IMAC-imidazole 100mM and CM10 at pH6 aliquots gave the best resolution and peak intensities for the peaks of interest (in this case, HNP 1-3 as indicated by small red arrow below the spectra). ........................................................................................................................................ 68
Figure 2-2: A schematic view of the fractionation process showing the final number of fractions collected per each starting sample. ........................................................................................................................................ 69
Figure 2-3: A schematic representation of the protein identification workflow for pooled saliva samples based on a 1D-gel electrophoresis, trypsin digestion and LC-MS/MS. ........................................................................................................................................ 73
Figure 2-4: A representative GoldChip® array spectra for the initial exploratory settings based on SPA matrix showing poor quality for most of the spectra. The samples used here had been prepared by fractionation using IMAC columns and eluted using 100mM imidazole elution buffer. ........................................................................................................................................ 78
Figure 2-5: Representative GoldChip® array spectra based on CHCA matrix and the optimised settings in Table 2-2 showing good quality for most of the spectra. The samples used here had been prepared by fractionation using CM10 columns elution buffer of pH 8. ........................................................................................................................................ 79
Figure 2-6: Representative spectra for CM10 fractions at different pH in both GCF and saliva performed on the same GoldChip® array. Notice the high peak intensities in the pH 6 fractions for both GCF and saliva compared to other fractions of the same media. In these spectra, the peaks for HNP 2, 1 and 3 respectively can be seen (red arrows between the regions m/z 3300 to 3500). ........................................................................................................................................ 80
Figure 2-7: SDS-PAGE preparation of salivary protein samples (CM10-purified pH6 fraction and IMAC-purified 100mM imidazole fraction). The interface between the 4% and 20%w/v gel sections is marked. MW corresponds to unresolved molecular weight markers. Volumes of 15µl of sample and 5µl of size markers were loaded. The region excised is indicated by red boxes. ........................................................................................................................................ 81
Figure 2-8: SDS-PAGE separation of protein samples in GCF (CM10 pH 6 and IMAC 100mM imidazole fractions). The interface between the 4% and 20%w/v gel sections is marked. MW corresponds to unresolved molecular weight markers. Volumes of 15µl of sample and 5µl of size markers were loaded. The red boxes indicate the regions excised. ........................................................................................................................................ 82
Figure 2-9: Mascot match for cathelicidin in the 10-19KDa sample. .................................................. 83

Figure 2-10: Scaffold list of identified proteins and peptides in both CM10 pH6 and IMAC 100mM imidazole fractions of saliva. Green highlighting indicates protein identification with high probability. A total of 15 proteins were identified in both fractions................................................................. 84

Figure 2-11: Scaffold list of identified proteins and peptides in both CM10 pH6 fractions of GCF. The IMAC portion is not shown, as it did not contain any IDs. Green highlighting indicates protein identification with high probability. A total of 8 proteins were identified in this fraction. HNP 1 is boxed in red....................................................................................................................................... 85

Figure 2-12: SDS-PAGE preparation of salivary protein samples (CM10 pH6). MW corresponds to unresolved molecular weight markers. Volumes of 30µl of 1:2 sample and 5µl of size markers were loaded. The region excised is highlighted in red with the bromophenol blue dye front being visible at the bottom. The image shown is at high contrast.................................................................................................................. 86

Figure 2-13: Mass spectrometric coverage of neutrophil granule peptide HP1 and matching proteins in saliva. Other proteins assigned to the same two peptides are stated in the text. .......................................................... 87

Figure 2-14: Scatter graph of the LC-MS/MS peaks for HNP 1 and Probably HNP 2 and 3 in saliva..... 87

Figure 2-15: Scatter graph of the LC-MS/MS peaks for HNP 1 and Probably HNP 2 and 3 in saliva..... 87

Figure 2-16: SDS-PAGE separation of saliva sample. Volumes of 40µl of 1:2 sample (X) and 5µl of size markers (M) were loaded. The region excised for in-gel digestion is indicated by the red box. ........ 88

Figure 2-17: Four scatter graphs for calcium-binding protein S100A8 specific peptide sequences from the LC-MS/MS peaks of CM10 fraction (pH 6) of saliva.................................................................................... 90

Figure 2-18: Scaffold 3 display of aligned samples. Percentage values are scores indicating confidence of protein identification. Identifications are from Uniprot database human entries. Results are shown using minimal stringency settings for peptide and protein identification in Scaffold........................................ 96

Chapter 3

Figure 3-1: ROC construct showing the dimensions that determine the classification of a test as either perfect, better or hopeless. FPR (False positive rate) and TPR (true positive rate). (Adapted from online Wikipedia· http://en.wikipedia.org/wiki/File:ROC_space-2.png) ......................................................... 109

Figure 3-2: Scatter graph for salivary total protein concentration in the diagnostic samples from subjects with gingivitis (G), chronic periodontitis (CP) and aggressive periodontitis (AgP). .......... 116

Figure 3-3: HNP 1-3 levels measured using ELISA in GCF showing significantly greater levels in both chronic periodontitis (CP) and aggressive periodontitis (AgP) than in gingivitis (G). .................... 117

Figure 3-4: Receiver characteristic curve (ROC) for HNP 1-3 differentiating between (A) gingivitis and chronic periodontitis and (B) gingivitis and aggressive periodontitis........................................ 117

Figure 3-5: (A) Scatter plots of salivary HNP 1-3 in different periodontal health status showing significantly higher expression in CP than in gingivitis. (B) Box and whisker (Tukey) plot showing higher expression of salivary HNP 1-3 in periodontitis than in gingivitis................................. 119

Figure 3-6: ROC construct for salivary HNP 1-3 with 75.3% AUC (p=0.0066) a sensitivity of 72.73% and specificity of 61.11% to differentiate gingivitis from CP at a cut off of >0.2087 µg/ml ......................... 119

Figure 3-7: A composite graph of HNP 1-3 levels in saliva and GCF showing higher expression in GCF than in saliva. ........................................................................................................ 120

Figure 3-8: Linear regression plot for HNP 1-3 levels in saliva against that in GCF paired within subjects showing lack of significant correlation within subjects................................................................. 121
Chapter 4

Figure 4-1: Scatter plot for salivary concentrations of HNP 1-3 in different diabetic status regardless of periodontal diagnosis showing no significant difference by one way ANOVA with Bonferroni's correction....................................................................................................................... 146

Figure 4-2: Scatter plot for salivary concentrations of LL-37 in different diabetic status regardless of periodontal diagnosis showing no significant difference by one way ANOVA with Bonferroni's correction....................................................................................................................... 146

Figure 4-3: Scatter plot for salivary concentrations of S100A8 in different diabetic status regardless of periodontal diagnosis showing no significant difference by one way ANOVA with Bonferroni's correction....................................................................................................................... 147

Figure 4-4: Scatter plot for salivary concentrations of total protein in different diabetic status regardless of periodontal diagnosis showing no significant difference by one way ANOVA with Bonferroni's correction....................................................................................................................... 147

Figure 4-5: Scatter graph showing differential expression of salivary HNP 1-3 in different periodontal disease severity regardless of the diabetic status. P values are based on 1-way ANOVA with Bonferroni’s post test correction. *** refers to < 0.0001....................................................................................................................... 148

Figure 4-6: Scatter plots for mean salivary HNP 1-3 concentration in non-diabetics according to periodontal status. Means are significantly different by 1-way ANOVA (p = 0.03) but not with Bonferroni’s post test correction....................................................................................................................... 149
Disease severities regardless of the diabetic status. There was weak but statistically positive

Figure 4-7: Scatter plots for mean salivary HNP 1-3 concentration in well controlled diabetics according to periodontal status. P- values are based on 1-way ANOVA with Bonferroni’s post test correction. * refers to p < 0.05...

Figure 4-8: Scatter plots for mean salivary HNP 1-3 concentration in poorly controlled diabetics according to periodontal status. P- values are based on 1-way ANOVA with Bonferroni’s post test correction.

Figure 4-9: A representative ROC construct for salivary HNP 1-3 ability to distinguish between gingivitis/mild periodontitis from severe periodontitis in all cases regardless of their diabetic status.

Figure 4-10: Scatter graph showing no statistically significant differential expression of salivary LL-37 by 1-way ANOVA with Bonferroni’s post-test correction in different periodontal status regardless of the diabetic status.

Figure 4-11: Scatter plot showing no statistically significant differential expression of salivary LL-37 by 1-way ANOVA with Bonferroni’s post-test correction in the non-diabetics.

Figure 4-12: Scatter plot showing no statistically significant differential expression of salivary LL-37 by 1-way ANOVA with Bonferroni’s post-test correction in the well controlled diabetics.

Figure 4-13: Scatter plot showing no statistically significant differential expression of salivary LL-37 by 1-way ANOVA with Bonferroni’s post-test correction in the poorly controlled diabetics.

Figure 4-14: Scatter graph showing no statistically significant differential expression of salivary S100A8 by 1-way ANOVA with Bonferroni’s post-test correction in different periodontal status regardless of the diabetic status.

Figure 4-15: Scatter plots for mean salivary S100A8 concentration in non-diabetics according to periodontal status. P- values are based on 1-way ANOVA with Bonferroni’s post test correction.

Figure 4-16: Scatter plots for mean salivary S100A8 concentration in well-controlled diabetics according to periodontal status. P- values are based on 1-way ANOVA with Bonferroni’s post test correction showing no statistically significant different.

Figure 4-17: Scatter plots for mean salivary S100A8 concentration in poorly controlled diabetics according to periodontal status. P- values are based on 1-way ANOVA with Bonferroni’s post-test correction showing no statistically significant different.

Figure 4-18: ROC construct for salivary S100A8 ability to distinguish between gingivitis/mild periodontitis from severe periodontitis in the non-diabetic group.

Figure 4-19: Salivary protein content in different periodontal severity showing statistically significant higher salivary concentration of total proteins in the moderate periodontitis group than the mild/healthy group based on one-way ANOVA with Bonferroni’s comparison test, p = 0.01. The cases are not grouped according to diabetic status.

Figure 4-20: Salivary concentrations of total protein in different periodontal disease severities among the non-diabetic group showing no statistical difference.

Figure 4-21: Salivary concentrations of total protein in different periodontal disease severities among the well controlled diabetic group showing no statistical difference.

Figure 4-22: Salivary concentrations of total protein in different periodontal disease severities among the poorly controlled diabetic group showing no statistical difference.

Figure 4-23: Correlations coefficient between salivary HNP 1-3 and S100A8 in different periodontal disease severities regardless of the diabetic status. There was weak but statistically positive

LIST OF FIGURES
correlation in the mild and moderate periodontitis groups that was lost in the severe periodontitis group. .......................................................................................................................................................... 162

Figure 4-24: This is a representative scatter plot for normalised values to 2.9 µg/ml and 255.4 ng/ml cut-offs for HNP 1-3 & S100A8 respectively................................................................................................................................................... 164

Figure 4-25: Representative graph showing an arbitrary choice of x and y-intercepts to be used as the cut-offs for normalizing the two markers (HNP 1-3 on the x-axis and S100A8 on the y-axis). In this case, 12.5 µg/ml for HNP 1-3 and 550 ng/ml for S100A8 were chosen.................................................................................................................. 165

Figure 4-26: ROC curve for the combined normalised values for HNP 1-3 and S100A8 using the cut-offs 12.5 µg/ml for HNP 1-3 and 550 ng/ml for S100A8. .................................................................................................................. 166

Figure 4-27: Representative graph showing an arbitrary choice of x and y-intercepts to be used as the cut-offs for normalizing the two markers (HNP 1-3 on the x-axis and S100A8 on the y-axis). In this case, 5 µg/ml for HNP 1-3 and 500 ng/ml for S100A8 were chosen.................................................................................................................. 166

Figure 4-28: Scatter plot of summed value of normalised concentration of HNP 1-3 and S100A8 (normalization cut-offs are 5 µg/ml and 500 ng/ml respectively) showing statistically significant higher mean values in severe periodontitis. P = 0.0003 by 1-way ANOVA with Bonferroni’s correction................. 167

Figure 4-29: ROC curve for the combined normalised values for HNP 1-3 and S100A8 using the cut-offs 5 µg/ml for HNP 1-3 and 500 ng/ml for S100A8.................................................................................................................. 167

Figure 4-30: Representative graph showing choice of x and y-intercepts to be used as the thresholds for normalizing the two markers (HNP 1-3 on the x-axis and S100A8 on the y-axis) based on optimized cut-off points in the individual marker ROC curves (Figure 4-9 and Figure 4-18). In this case, 2.9 µg/ml for HNP 1-3 and 255.4 ng/ml for S100A8 were chosen.................................................................................................................. 168

Figure 4-31: Scatter plot of summed value of normalised concentration of HNP 1-3 and S100A8 (normalization cut-offs are 2.9 µg/ml and 255.4 ng/ml respectively) showing statistically significant higher mean values in severe periodontitis as compare to the other two. P = 0.0007 by 1-way ANOVA with Bonferroni’s correction........................................................................................................................................... 168

Figure 4-32: ROC curve for the combined normalised values for HNP 1-3 and S100A8 using the cut-offs 2.9 µg/ml for HNP 1-3 and 255.4 ng/ml for S100A8.................................................................................................................. 169

Chapter 5

Figure 5-1: A graph of salivary HNP 1-3 concentration showing a 20.8 fold higher mean concentration in saliva collected by spitting as compared to Salivette®. T-test p = 0.0001.................................................................................................................. 183

Figure 5-2: A plot showing the relationship between salivary HNP 1-3 concentration in saliva collected by spitting as compared to Salivette®. The concentrations were higher in the spitting sample than in the Salivette sample for each paired sample. .................................................................................................................................................. 183

Figure 5-3: Regression analysis between salivary concentration of HNP 1-3 collected using Salivette® versus spitting method showing a moderate but significant positive correlation between the two. r² = 0.42 (p = 0.0002).................................................................................................................................................. 184

Figure 5-4: A graph of salivary LL-37 concentration showing a 22.9 fold higher mean concentration in saliva collected by spitting as compared to Salivette®. T-test p = 0.0002.................................................................................................................. 184

Figure 5-5: A plot showing the relationship between salivary LL-37 concentration in saliva collected by spitting as compared to Salivette®. The concentrations were higher in the spitting sample than in the Salivette sample for each paired sample .................................................................................................................................................. 185
Figure 5-6: Regression analysis between salivary concentration of LL-37 collected using salivette® versus spitting method showing a moderate but significant positive correlation between the two. \( r^2 = 0.4 \) \((p = 0.0005)\).................................................................................................................................185

Figure 5-7: Scatter graphs for (A) HNP 1-3 and (B) LL-37 showing Salivette®/spitting ratios with no significant difference regardless of the order of sample collection method. T-test \( p \) value = 0.45 and 0.72 for HNP 1-3 and LL-37 respectively.................................................................187

Figure 5-8: Scatter graph of salivary HNP 1-3 concentrations for 24 samples left at room temperature for 0, 6, 24, 48 and 72 hours. Notice one outlier at time zero which corresponds to sample ID number 2118 that had mucin residuals that probably affected HNP 1-3 measurement. There was no statistical difference in the concentrations at different time points (1-way ANOVA with Bonferroni’s post test comparisons, \( p \) value = 0.16) ................................................................................................................................188

Figure 5-9: Bar graph showing proportion change of salivary HNP 1-3 from baseline levels (Time zero). Error bars are SD. There is significant reduction in HNP 1-3 at ‘72Hrs+PI’ compared to all other time points \((p < 0.0001, \text{one-way ANOVA with Bonferroni’s post test correction})\). ........189

Figure 5-10: Scatter graph of salivary LL-37 concentrations for 24 samples left at room temperature for 0, 6, 24, 48 and 72 hours. There was a significant reduction in the concentration between time zero and both 48 and 72 hours, and also between 6 hours and both 48 and 72 hours, 1-way ANOVA with Bonferroni’s post test comparison, \( p \) value = 0.001. .............................................................................................................................190

Figure 5-11: Bar graph showing proportion change of salivary LL-37 from baseline levels (Time zero). Error bars are SD. There is significant reduction in LL-37 at ‘72Hrs & ‘72Hrs+PI’ compared to baseline and ‘6Hrs’ \((p = 0.001, \text{one-way ANOVA with Bonferroni’s post test correction})\). ........191

Figure 5-12: Scatter plot of salivary total protein concentration normalised to the mean value for the total samples collected. No significant difference in different time points \((\text{ANOVA with Bonferroni’s post test comparison, } p = 0.144)\).........................................................................................................................192

Figure 5-13: Bar graph showing variations in mean total salivary protein concentrations at different time points. However, variation did not reach statistical significance by 1-way ANOVA with Bonferroni’s post-test comparison.................................................................................................................................193

Figure 5-14: Scatter plot of salivary HNP 1-3 concentration normalised to the intra-subject mean value. Notice peaks at 12:00 and 16:00 hours but difference is not statistically significant \((\text{ANOVA with Bonferroni’s post test comparison, } p = 0.282)\). .................................................................................................................................194

Figure 5-15: Bar graph showing variations in mean salivary HNP 1-3 concentrations at different time points. However, variation did not reach statistical significance using 1-way ANOVA with Bonferroni’s post test comparison................................................................................................................................195

Figure 5-16: Scatter plot of salivary LL-37 concentration normalised to the intra-subject mean value. No significant difference in different time points is noted \((\text{ANOVA with Bonferroni’s post test comparison, } p = 0.07)\).................................................................................................................................195

Figure 5-17: Bar graph showing variations in actual mean salivary LL-37 concentrations at different time points. However, variation did not reach statistical significance by 1-way ANOVA with Bonferroni’s post test comparison................................................................................................................................196

Figure 5-18: Scatter plot of salivary S100A8 concentration normalised to the intra-subject mean value. Notice significant peak at 16:00 hours compared to all other time points \((\text{ANOVA with Bonferroni’s post test comparison, } p = 0.0001)\).................................................................................................................................197

Figure 5-19: Bar graph showing variations in mean salivary S100A8 concentrations at different time points. Note the statistically significant peak at 16:00 hours compared to 10:00, 14:00 hours and days 2 and 3 \((\text{One-way ANOVA with Bonferroni’s post test comparison, } p = 0.001)\).................................................................................................................................198
Chapter 6

Figure 6-1: Schemata of Chromosome 8 showing the arrangement of the DEFA1A3 and its neighbouring defensin genes. Adapted from (Nuytten et al., 2009) .............................................................. 208

Figure 6-2: Flow through schemata for DNA extraction using DNeasy extraction kit (Qiagen Group) ................................................................................................................................. 213

Figure 6-3: NCBI Primer blast for DEFA1 PCR template input >ref|NT_023736.17|:c6827602-6825171 Homo sapiens chromosome 8 genomic contig, GRCh37 reference primary assembly (Range 6827602-6825171) Green = Forward primer, Pink = Probe and Blue = Reverse primer all in 5’ to 3’ .......... 216

Figure 6-4: Frequency histogram showing the frequency (%) of the various copies of the DEFA1A3 CNV. The most frequent CN was 6 (31.9%). .................................................................................. 218

Figure 6-5: Bar graph showing the proportion of cases with copy numbers ≤6 versus > 6 within each periodontal diagnosis. .................................................................................................................... 219

Figure 6-6: Bar graph of the absolute numbers of cases with copy numbers ≤6 versus > 6 within each periodontal diagnosis showing statistically significant difference between the UK reference population and the disease cases but not the periodontally healthy group. .................................................................................. 219

Figure 6-7: A representative print-out for the RT-PCR amplification plot showing excellent amplification plots for samples and no amplification for NTC samples. .......................................................... 220

Figure 6-8: A schematic model for a probable relationship between DEFA1A3 gene copy numbers with corresponding baseline salivary HNP 1-3 and how this may influence the quantitative and qualitative development of the microbial niche in the gingival crevice either promoting periodontal health or development of periodontal disease. .................................................................................. 224

Chapter 8

Figure 8-1: Apparent molecular weights of SeeBlue Plus2 Pre-stained standard on a NuPAGE Novex 4-12% Bis-Tris Gel w/MES .................................................................................. 241
Chapter 1

Table 1-1: Abbreviated Version of the 1999 Classification of Periodontal Diseases and Conditions (Armitage, 1999) ................................................................. 31
Table 1-2: The specific features of the two different forms of AgP are stated here. Adapted (Lang, 1999; Tonetti and Mombelli, 1999). .................................................................................................................. 34
Table 1-3: A summary of data on the prevalence of periodontitis according to different case definitions as described in a study looking at the impact of using different case definitions for periodontal research (Costa et al., 2009). Def = definition. .............................................................................................................. 36
Table 1-4: Highlight of the limitations of current clinical diagnostic parameters ............................................. 39
Table 1-5: A sample list of some of the host derived enzymes and their inhibitors in GCF. Adapted (Armitage, 2004a) ........................................................................................................................................ 46
Table 1-6: A summary of patient-based clinical data for the GCF study (Av.) = Average, CP = Chronic periodontitis and AgP = Aggressive periodontitis. ..................................................................................... 54

Chapter 2

Table 2-1: List of buffers used in IMAC fractionation of both saliva and GCF samples ..................................... 70
Table 2-2: SELDI-TOF MS parameters for running the GoldChip® arrays in the different optimization phases ........................................................................................................................................ 73
Table 2-3: A summary list of proteins and peptides identified with high confidence. In both CM10 pH 6 and IMAC 100mM imidazole fractions of saliva. A total of 16 proteins were identified in both fractions. Notice the overlap in the two fractions. All the proteins identified in the IMAC fraction were also present in the CM10 fraction except for the two highlighted in blue. .................................................................................. 91
Table 2-4: Final proteomic analysis of SDS-PAGE gel region 3-6kDa for saliva CM10 pH6 sample.............. 92
Table 2-5: Proteomic analysis of SDS-PAGE gel bands in GCF Sample CM10 pH6 & IMAC 100mM imidazole fractions. Gene index (protein accession) numbers are provided for similar proteins containing the same assigned peptides. Identities of similar proteins can be found at http://www.ncbi.nlm.nih.gov/ or BLAST searched at http://www.uniprot.org/. Neutrophil granule peptide HP1 is highlighted in the CM10 fraction in yellow colour. All proteins with a Scaffold probability of greater than 0% are included rather than only those with a probability of 99% or more. .................................................................................................................................. 95
Table 2-6: Proteomic analysis of SDS-PAGE gel bands 3-6KDa and 10-19KDa for CM10 pH6 saliva sample. Uniprot protein accession numbers (human entries) are provided for each assigned protein and where appropriate for database entries containing the same set of assigned peptides. Identities of similar proteins can be found or BLAST-searched at http://www.ncbi.nlm.nih.gov/ and http://www.uniprot.org/.................................................................................................................. 103

Chapter 3

Table 3-1: The identity of the potential biomarker peaks candidate for validation studies .................... 110
Table 3-2: Summary of the clinical criteria for the diagnosis of the periodontal condition .................... 113
Chapter 4

Table 4-1: Table showing the criteria used for case definition (CAL = clinical attachment loss; BOP = bleeding on probing)............................................................141

Table 4-2: Table showing the distribution of the cohort according to diabetic status and periodontal diagnosis according to the case definition. Numbers in parenthesis represent the proportion in percentage of periodontal diagnosis in each diabetic status followed by proportion of diabetic status in each periodontal severity group ................................................................................................................................................................................144

Table 4-3: A table showing summary demographic and basic periodontal clinical parameters according to different diabetic status (M/F = male/female ratio; BOP = Bleeding on probing; PPD = Probing pocket depth; CAL = Clinical attachment loss; HbA1c = glycosylated haemoglobin). P-values (*) are based on one-way ANOVA with Bonferroni’s post test correction.................................................................144

Table 4-4: Table showing the cut-off point concentrations of salivary HNP 1-3 that differentiate gingivitis from chronic periodontitis and LL-37 that differentiate gingivitis from periodontitis............145

Table 4-5: Matrix tables for the correlation between salivary HNP 1-3, LL-37, S100A8 and total protein concentration in saliva showing very weak but significant correlation in all pairs of analytes (data are from all cases regardless of their periodontal or diabetic status). (R = correlation coefficient; ‘p’ is the p value)..........................................................................................................................................................................................160

Table 4-6: Matrix tables for the correlation between salivary HNP 1-3, LL-37, S100A8 and total protein concentration in saliva showing (A) very weak but significant correlation in all pairs of analytes among the gingivitis/mild periodontitis group (B) weak but significant correlation between the AMPs but not with total protein concentration among the moderate periodontitis group and (C) weak correlation between HNP 1-3 and LL-37 and no significant correlation between all the other analytes. (R = correlation coefficient; ‘p’ is the p value)..................................................................................................................................................161

Chapter 5

Table 5-1: Two-way ANOVA with Bonferroni’s multiple comparison summary table for salivary HNP 1-3 showing no significant contribution by RT storage time and inter-individual variation to the total variation observed..................................................................................................................................................................................188

Table 5-2: Two-way ANOVA with Bonferroni’s multiple comparison summary table for salivary LL-37 showing significant effect of both RT storage time and inter-individual variation to the total variation observed........................................................................................................................................................................190

Table 5-3: Two-way ANOVA summary table for total salivary protein showing lack of significant effect of time and day of sample collection as well as inter-individual variation.................................................192

Table 5-4: Two-way ANOVA summary table for salivary HNP 1-3 showing a small but significant contribution by inter-individual variation to the total variation observed. There was lack of significant effect of time and day of sample collection. ..............................................................................................................................................194

Table 5-5: Two-way ANOVA (Bonferroni’s multiple comparisons) summary table for salivary LL-37 showing a moderate but significant contribution of inter-individual variation to the total variation observed. There was lack of significant effect of time and day of sample collection.........................................................196

Table 5-6: Two-way ANOVA with Bonferroni’s multiple comparison summary table for salivary S100A8 showing a moderate but significant contribution of inter-individual variation to the total variation observed as well as a significant effect of time and day of sample collection........................................197

Table 5-7: Summary table showing the mean levels (SD) of different salivary analytes at different time points. There was no statistical difference between the different time points in total protein, HNP 1-3
and LL-37 concentrations. S100A8 was significantly higher at 16:00 hrs (day 1) than 10:00 hrs (days 1, 2 and 4) and 14:00 hrs (day 1). ANOVA with Bonferroni’s post test comparison, p value = 0.001....199

Table 5-8: Summary table showing mean (SD) of the different analytes normalised as the percentage proportion of the individual mean over the 7 time points. .................................................................200

Chapter 6

Table 6-1: Table showing the frequency of the different periodontal status. .................................................211
Table 6-2: This table shows the composition of the RT-PCR reaction mix (volume in microliters) .........213
Table 6-3: RT-PCR run parameters for DEFA1A3 CNV assay using ABI 7900HT Fast system on Absolute quantification standard setting. ........................................................................................................214
Table 6-4: Table showing Custom TaqMan® Copy Number Assays details used to perform the RT-PCR experiment. The primers are shown in 5’ to 3’ direction except for the reverse primer which is in 3’ to 5’. The assay ID is DEFA1_CCWR1LZ from Applied Biosystems.................................................215
Table 6-5: Total counts of subjects with DEFA1A3 copy numbers ≤ 6 or > 6 in different periodontal diagnostic categories with no significant difference between the groups by Chi square distribution (p = 0.20, df = 3). The UK and Belgian populations’ DEFA1A3 CNV data are shown for comparison, adapted from (Aldred et al., 2005; Nuytten et al., 2009) respectively.................................................................218

Chapter 8

Table 8-1: Partial list of putative GCF & Salivary markers for periodontal disease as flagged by SELDI-TOF MS. The markers with tentatively known identity are in bold font. Different suspected fragments of partially processed hCAP38/LL-37 are shown with an asterisk (*). .........................................................235
Table 8-2: Protocol for GCF Profiling Using a Copper-Enriched Array On-Spot ........................................238
Table 8-3: Protocol for GCF Profiling On-Spot with CM10 ......................................................................239
Table 8-4: Table showing the composition of 20% resolving gel used in 1D-PAGE before LC-S/MS ....239
Table 8-5: Table showing the composition of 4% stacking gel using 10 well, 1mm comb used before LC-S/MS. ........................................................................................................................................240
Table 8-6: Complete list of all the proteins and peptides identified using LC-MS/MS in both GCF and saliva. ........................................................................................................................................248
Table 8-7: Randomization list for collecting saliva samples using either salivette or by spitting first. ........................................................................................................................................251
LIST OF SUPPLIERS

Abcam PLC, Cambridge, UK
Alpha Diagnostics International Inc., Texas, USA
Applied Biosystems Inc., Warrington, UK
Bio-Rad Laboratories Ltd., Hemel Hempstead, UK
BMG LABTECH Ltd., Aylesbury, UK
Caltag-Medsystems Ltd., Buckingham, UK
Cambridge Bioscience, Cambridge, UK
Eppendorf UK Ltd., Cambridge, UK
Florida Probe Corp., Gainesville, USA
GraphPad Software Inc., California, USA
Invitrogen Ltd., Paisley, UK
MBL International Corporation, Massachusetts, USA
Nunc, Loughborough, UK
Pro Flow Inc., New York, USA
Qiagen Ltd., Crawley, UK
R&D Systems Europe Ltd., Abingdon, UK
Sarstedt Ltd., Leicester, UK
Sigma-Aldrich Company Ltd., Dorset, UK
Thermo Fisher Scientific Inc., Hemel Hempstead, UK
ABBREVIATIONS

%  Percentage
®  Registered trade mark
°C  Degrees Celsius
2-D  Two-dimensional
A.a  *Aggregatibacter actinomycetemcomitans*
AAP  American academy of periodontology
ACN  Acetonitrile
AgP  Aggressive periodontitis
AL  Attachment Loss
AMP  Antimicrobial peptide
ANOVA  Analysis of variance
AUC  Area under the curve
*B.f*  *Bacteroides forsythus*
BB  Binding buffer
BCA  Bicinchoninic Acid Protein
BMT  Bone marrow transplant
BOP  Bleeding on probing
BSA  Bovine serum albumin
CAL  Clinical attachment loss
CDC  Centres for disease control and prevention
CHCA  Alpha-cyano-4-hydroxy cinnamic acid
CM10  Weak cation exchanger
CN  Copy number
CNV  Copy number variation
CP  Chronic periodontitis
Cl  cycle threshold
Da  Dalton
DEFA1  Alpha defensin-1 gene
DEFA1A3  Alpha defensin-1 and -3 gene
DEFA3  Alpha defensin-3 gene
DF  Degrees of freedom
DI  Deionized water
DM  Diabetes mellitus
DNA  Deoxyribonucleic acid
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotide</td>
</tr>
<tr>
<td>EAM</td>
<td>Energy absorbing matrix/molecules</td>
</tr>
<tr>
<td>EB</td>
<td>Elution buffer</td>
</tr>
<tr>
<td>EDM</td>
<td>Expression difference mapping</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethyaledmine-tetra-acetic-acid electrophoresis</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme linked immunosorbent assay</td>
</tr>
<tr>
<td>EOP</td>
<td>Early onset periodontitis</td>
</tr>
<tr>
<td>FPR</td>
<td>False positive values or rate</td>
</tr>
<tr>
<td>FT-ICR</td>
<td>Fourier transform ion cyclotron resonance</td>
</tr>
<tr>
<td>G</td>
<td>Gingivitis</td>
</tr>
<tr>
<td>G-CSF</td>
<td>Recombinant granulocyte colony-stimulating factor</td>
</tr>
<tr>
<td>GAP</td>
<td>Generalized Aggressive Periodontitis</td>
</tr>
<tr>
<td>GCF</td>
<td>Gingival crevicular fluid</td>
</tr>
<tr>
<td>gDNA</td>
<td>Genomic DNA</td>
</tr>
<tr>
<td>GeLC-MS/MS</td>
<td>Gel LC-MS/MS</td>
</tr>
<tr>
<td>GI</td>
<td>Gingival index</td>
</tr>
<tr>
<td>HbA1c</td>
<td>Glycosylated haemoglobin</td>
</tr>
<tr>
<td>hBD</td>
<td>Human β-defensins</td>
</tr>
<tr>
<td>HBD-3</td>
<td>Human beta defensin-3</td>
</tr>
<tr>
<td>hCAP38</td>
<td>Human cationic antimicrobial peptide of 38 amino acids.</td>
</tr>
<tr>
<td>HNP</td>
<td>Human neutrophil peptide</td>
</tr>
<tr>
<td>Hrs</td>
<td>Hours</td>
</tr>
<tr>
<td>ICTP</td>
<td>Pyridinoline cross-linked carboxyterminal telopeptide of type I collagen</td>
</tr>
<tr>
<td>ID</td>
<td>Identification</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilized metal affinity capture</td>
</tr>
<tr>
<td>Inc.</td>
<td>Incorporated</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>iTRAQ</td>
<td>Isobaric tags for relative and absolute quantification</td>
</tr>
<tr>
<td>JP</td>
<td>Juvenile periodontitis</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilo Dalton</td>
</tr>
<tr>
<td>LAP</td>
<td>Localized Aggressive Periodontitis</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>Liquid Chromatography quadruple Mass Spectrometry</td>
</tr>
<tr>
<td>LL-37</td>
<td>Cathelicidin LL-37</td>
</tr>
<tr>
<td>Ltd</td>
<td>Limited</td>
</tr>
<tr>
<td>M/F</td>
<td>Male to female ratio</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass charge ratio</td>
</tr>
<tr>
<td>MALDI-TOF</td>
<td>Matrix assisted laser desorption and ionization time of flight</td>
</tr>
<tr>
<td>mg</td>
<td>Milligram</td>
</tr>
<tr>
<td>min</td>
<td>Minute</td>
</tr>
<tr>
<td>ml</td>
<td>Milliliter</td>
</tr>
<tr>
<td>mm</td>
<td>Millimeter</td>
</tr>
<tr>
<td>MRC</td>
<td>Medical Research Council</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MRP-8</td>
<td>Myeloid-related protein 8</td>
</tr>
<tr>
<td>mw</td>
<td>molecular weight</td>
</tr>
<tr>
<td>ng</td>
<td>Nano gram</td>
</tr>
<tr>
<td>NHS</td>
<td>National health service</td>
</tr>
<tr>
<td>Nm</td>
<td>Nanometers</td>
</tr>
<tr>
<td>NO</td>
<td>Nitric oxide</td>
</tr>
<tr>
<td>NP</td>
<td>Normal phase</td>
</tr>
<tr>
<td>NR</td>
<td>Non responders</td>
</tr>
<tr>
<td>NTA</td>
<td>Nitroacetic acid</td>
</tr>
<tr>
<td>NTC</td>
<td>No template control</td>
</tr>
<tr>
<td>OHI</td>
<td>Oral hygiene instruction</td>
</tr>
<tr>
<td><em>P. intermedia</em></td>
<td><em>Prevotella intermedia</em></td>
</tr>
<tr>
<td><em>P.g</em></td>
<td><em>Porphyromonas gingivalis</em></td>
</tr>
<tr>
<td>pg</td>
<td>Pico gram</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PGE$_2$</td>
<td>Prostaglandin E 2</td>
</tr>
<tr>
<td>PI</td>
<td>Proteinase Inhibitor</td>
</tr>
<tr>
<td>PMN</td>
<td>Polymorphonuclear cells</td>
</tr>
<tr>
<td>PPD</td>
<td>Probing pocket depth</td>
</tr>
<tr>
<td>ppm</td>
<td>Parts per million</td>
</tr>
<tr>
<td>PTM</td>
<td>Post-translational modification</td>
</tr>
<tr>
<td>qPCR</td>
<td>quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>QTOF</td>
<td>Quadruple time of Flight</td>
</tr>
<tr>
<td>RLH</td>
<td>Royal London Hospital</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver operating characteristic</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Real time polymerase chain reaction</td>
</tr>
<tr>
<td>S. mutans</td>
<td>Streptococcus mutans</td>
</tr>
<tr>
<td>S100A8</td>
<td>Protein S100A8</td>
</tr>
<tr>
<td>SCX</td>
<td>Strong cationic exchanger</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulphate polyacrylamide gel</td>
</tr>
<tr>
<td>Sec</td>
<td>Second</td>
</tr>
<tr>
<td>SELDI</td>
<td>Surface enhanced laser desorption ionization</td>
</tr>
<tr>
<td>SPA</td>
<td>Sinapinic acid</td>
</tr>
<tr>
<td>T. Forsythus</td>
<td>Tannerella Forsythus</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>TIC</td>
<td>Total ion current</td>
</tr>
<tr>
<td>TM</td>
<td>Trade mark</td>
</tr>
<tr>
<td>TMT</td>
<td>Tandem mass tags</td>
</tr>
<tr>
<td>TOF-MS</td>
<td>Time of flight mass spectrometry</td>
</tr>
<tr>
<td>TPR</td>
<td>True positive values</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>UP</td>
<td>Ultra pure</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>WS</td>
<td>Whole saliva</td>
</tr>
<tr>
<td>µl</td>
<td>Microlitre</td>
</tr>
</tbody>
</table>
Chapter 1 - Literature Review
1.1 Introduction

For over a century the diagnosis and classification of periodontal disease has been based on full clinical examination. This process is time consuming, often inaccurate, laced with subjectivity and provides little etiological information to aid treatment planning and estimates of prognosis. The medical world has made great strides in advancing diagnostic capabilities such as by the use of biochemical analytes, which may provide specific information about disease status including aetiologic pathways. The potential utility of biomarkers in periodontology may include the ability to improve diagnosis and treatment planning of patients with periodontal disease (for example by allowing stratification of patients who would benefit from different treatments), the potential to identify novel biological pathways which may result in development of new periodontal treatments, and as an epidemiological tool to allow the rapid screening of patients with periodontal disease in an objective and cost effective way (i.e. quick and inexpensive diagnostics that do not require highly skilled staff to perform). It is envisaged that such biomarkers may render themselves amenable to be packaged for-at-home use and would be ideal for the development of point of care (POC) diagnostics.

In most biological systems, there is a delicate balance that keeps the biochemical constitution of the organism stably regulated. A change in the levels of these substances may indicate pathology either locally or further afield. Many of these biochemical substances are proteins and peptides that play a significant role in maintaining health and therefore changes in proteins and peptides and their expression levels would reflect changes in health status. The current concept of the pathogenesis of periodontitis is based on the understanding that there is interplay between the oral microbiota and the host response that is maintained in a delicate balance to maintain health. Therefore, a tip in this balance may result to periodontal destruction. In the process, the host and the microbiota release a lot of proteins and or peptides whether as enzymes by the bacteria for nutrition purpose or to evade detection by the host, or other host factors to defend the host (Arnim, 1966; Giannobile, 2008; Hendler et al., 2010; Taubman et al., 2007). These proteins and peptides might act as biomarkers for periodontal disease. Common sources of such biomarkers in...
the oral cavity include saliva and gingival crevicular fluid, (GCF) which are relatively easily accessible and acceptable to patients compared to blood or urine products (Koka et al., 2008). In addition, the proximity of these fluids to the diseased sites makes them very attractive for the search for biomarkers associated with periodontal disease or health.

Ideally, diseases would be classified on the basis of their aetiology. In the absence of such a classification then it is necessary to classify diseases according to their clinical features. Periodontal disease can be viewed as a syndromic condition requiring the description of several signs and symptoms before arriving at a diagnosis. Even then what mostly is described, as a periodontal diagnosis is empirical rather than definitive in nature. As in many complex diseases, the causes of periodontal disease are multifactorial. For this reason, a model for periodontal disease identification based on a combination of markers may yield better results for diagnosis than looking for a single marker (Hanioka et al., 2005; Ramseier et al., 2009). Potential markers of disease in saliva and GCF include specific bacteria, bacterial products, host cell products, host cell derived enzymes, indicators of humoral/cell mediated immunity and products of host connective tissue breakdown. This thesis describes studies, which attempt to identify some of these biomarkers that may demonstrate diagnostic and prognostic value in the management of destructive periodontal disease as well as reveal new insights into the pathogenic pathways of the condition.

1.1.1 Diseases Of The Periodontium

Periodontal disease is the most common of all oral diseases (and probably of any disease) in the human kind and as such stakes its claim as a major public health issue (Marshall-Day, 1951; Pihlstrom et al., 2005). It also is the second major cause of tooth loss (Albander et al., 1997; Albander et al., 1999; Baelum and Scheutz, 2002; Brown and Löe, 1993; Russell, 1957). Periodontal disease may be defined as an inflammatory condition initiated by oral microorganisms affecting the tissue structures surrounding and supporting the tooth also known as the periodontium. Broadly speaking, there are many other diseases or conditions of the periodontium including those of genetic, inflammatory,
developmental, neoplastic, traumatic, atrophic or metabolic origin (Armitage, 2004b). The inflammatory component forms the bulk of periodontal conditions affecting a staggering 90% of the human population in its various forms (Pihlstrom et al., 2005; Scherp, 1964). The mildest and apparently the commonest of these is gingivitis, which reversibly affects the gingiva without involving the deeper tooth supporting structures. The destructive form of inflammatory periodontal disease is periodontitis. Periodontitis is an irreversible condition consisting of an inflammatory process in the tooth investing tissues including the gingiva, alveolar bone, cementum and the periodontal ligament. It is defined as “an inflammatory disease of the supporting tissues of the teeth caused by specific microorganisms, resulting in progressive destruction of the periodontal ligament, and alveolar bone with pocket formation, recession, or both” (Guerini, 1909; Highfield, 2009).

The pathognomonic features of periodontitis are the breakdown of periodontal connective tissue with infiltration of the sub-epithelial area adjacent to the deepened pathological sulcus (periodontal pocket) by an inflammatory infiltrate. This loss of connective tissue coupled with the apical migration of the epithelial attachment and marginal alveolar bone loss leads to loss of attachment that can culminate in tooth morbidity and ultimately tooth loss. In fact, periodontitis is one of the top two risk factors leading to tooth loss (Aida et al., 2006; Aida et al., 2009; Akhter et al., 2008; Angelillo et al., 1996; Chrysanthakopoulos, 2011; McCaul et al., 2001; Morita et al., 1994; Oral health in America: a report of the Surgeon General, 2000; Shigli et al., 2009).

Since the dawn of human history, there has been consistent evidence of affliction by periodontal disease (Guerini, 1909). Epidemiologic studies have however shown variability in the degree of destruction between populations, subgroups of the same population and between and within individuals (Albander et al., 1999; Brown et al., 1990). Evidently even within the same individual, different sites may be affected to various degrees (Baelum et al., 1996; Papapanou et al., 1988).
1.1.2 Aetiology Of Periodontal Diseases

Our understanding of the aetiology of inflammatory periodontal diseases has advanced considerably in the past 50 years. Historically the aetiology of periodontal disease had been variously explained by such things as nutritional deficiencies, worms and poor vasculature (Guerini, 1909). As well, the understanding in the 1950s was that when left untreated periodontal disease followed a natural progression from gingivitis to periodontitis and finally to tooth loss (Arnim, 1966; Heitz-Mayfield et al., 2003; Russell, 1957; Schatzle et al., 2003). The periodontal disease index (PDI) in Russell’s classic epidemiological studies demonstrates clearly the understanding at the time of the continuum of periodontal disease progression from health, to gingivitis then to periodontitis. In this index, the numerical values are used in an increasing order from periodontal health, gingivitis to the higher scores representing severe periodontitis and ultimately tooth loss (Russell, 1957). However from the mid 1960s and early 1970s, studies began to show the role of bacteria in the etiology of the disease (Keyes, 1964; Lindhe, 1973; Loe et al., 1965; Löe, 1986; Silness and Loe, 1964; Theilade et al., 1966). This led to the simplistic view that periodontal disease had one cause (plaque bacteria) and as such removal of plaque and its retentive factors is all that was needed to control the disease.

However in the 80’s and early 90’s, this concept of periodontal disease as one disease with one cause was challenged by the findings of longitudinal epidemiological studies. These revealed that despite the high prevalence and incidence of developing signs of periodontal disease, only about 10% of the population faced a risk of severe destructive disease and that even in the same individual, there were variations in the rate and severity of the disease at different sites of the dentition (Baelum et al., 1996; Brown and Löe, 1993; Hugoson and Rylander, 1982; Löe, 1986). Recent papers have described this phenomenon by explaining the role of the host in destructive periodontal disease. Greenstein suggested that it is the host and not the bacteria that primarily determined the variation in disease severity (Greenstein, 2005). Moreover, studies revealed that it is not all the bacteria in dental plaque mass that are pathognomonic to periodontal disease but specific species that possess
unique characteristics that enabled them to both colonise and evade the host immune defense as well as induce disease (Newman et al., 1976; Socransky et al., 1984; Socransky and Haffajee, 1991).

The current thinking is that causation of periodontal disease is multifactorial with bacterial challenge and the consequent host response to this challenge playing a primary role. It is now generally acknowledged that in the natural history of periodontitis, plaque bacteria are necessary but not sufficient to cause periodontitis (Carvel and Carr, 1982; Gemmell and Seymour, 1994; Gemmell et al., 2002; Gemmell and Seymour, 2004; Kinane, 2001; Miyasaki, 1991; Schenkein, 2006). As such, different risk factors and risk determinants together play a critical role in the aetiology of disease. Risk factors are environmental or behavioral characteristics of the individual that increase the probability of developing disease. These are biologically related to the presence of the disease but not necessarily in a causative manner. They are usually modifiable e.g. plaque control and smoking. Risk determinants on the other hand are usually non-modifiable and include factors such as genetic makeup, gender and age of the subject.

From the forgoing, it appears that plaque bacteria endowed with specific virulence factors are responsible for the initiation of periodontal disease in susceptible hosts whose immune response is capable of perpetuating the progression of the disease. This host response is typically effected via production of active molecules that may lead to further amplification of the response in a positive feed back loop resulting in to overexpression of potentially destructive biomolecules such as cytokines and immunoglobulins. It has been anticipated that this orchestration between the bacterial virulence factors and the host response is capable of generating molecular trails (proteins and peptide biomarkers) in the periodontal tissues that leak out in to the oral fluids specifically the GCF and saliva (Giannobile et al., 2009).

1.1.3 Diagnosis and Classification of Periodontitis

Diagnosis suggests that a disease entity can be comprehended based on the key features of the disease and implicitly that this information may have an
impact on its management. Medical diagnosis has been defined as the conclusion reached through the process of identifying a disease by its signs and symptoms, and the results of various biological assessments including the treatment outcome and other risk factors. Diagnostic procedures may be used to identify people at risk of developing disease, detect early stage disease in clinically asymptomatic individuals, classify disease categories, predict likely responders to specific treatments monitor treatment efficacy and detect disease recurrence (Mombelli, 2005).

As in other disciplines of medicine, classification of periodontal diseases has been a way of trying to understand these diseases. Therefore in the past periodontitis has been classified in diverse ways reflecting the knowledge and to some extent the consensus view at the time of aetiology, diagnosis and prognosis.

Several classifications in the past have attempted this uphill task including the 1989 classification (Catton, 1989), the 1993 European classification that recognized early onset periodontitis as a distinct entity from adult periodontitis (Attstrom, 1993) and most recently the 1999 Classification by the International Workshop for Classification of Periodontal Diseases (Armitage, 1999). This current classification rejected the various classifications of early onset periodontitis (EOP) and in particular separated chronic periodontitis from aggressive periodontitis, with the latter encompassing most of what was previously known as EOP. Table 1-1 below shows a summary of the current classification of periodontal diseases.

Periodontitis is a syndrome, the clinical manifestations of which may vary in distribution and severity. Therefore, in spite of all the effort to diagnose and classify periodontal diseases, there is no natural basis for a sharp distinction between health and disease or between different forms of periodontitis. The impact of this is evident considering that there has been over 10 different classifications within a period of 20 years and none of these has been conclusive or without criticism (Baelum and Lopez, 2003). It has been recently commented convincingly that ‘periodontal diagnosis, which, despite rapid advances in knowledge, still lingers in the 20th century, and is based upon
subjective clinical examination procedures, which are time-consuming and poorly implemented in general dental practice’ (Chapple, 2009). The consequence of this is under-diagnosis of periodontal disease within general practice leading to significant amounts of untreated disease (Chapple, 2009; Morgan, 2001).

<table>
<thead>
<tr>
<th>Clinical parameter</th>
<th>Other descriptive characteristics</th>
</tr>
</thead>
</table>
| **Gingival diseases** | A. Dental plaque induced gingival diseases  
B. Non-plaque induced gingival lesions |
| **Chronic periodontitis** | A. Localised  
B. Generalised  
- Slight: 1-2 mm CAL  
- Moderate: 3-4 mm CAL  
- Severe: ≥ 5 mm CAL |
| **Aggressive periodontitis** | A. Localised  
B. Generalised  
- Slight: 1-2 mm CAL  
- Moderate: 3-4 mm CAL  
- Severe: ≥ 5 mm CAL  
Patients usually below 35 years of age |
| **Periodontitis as a manifestation of systemic diseases** | A. Associated with hematological disorders  
B. Associated with genetic disorders  
C. Not otherwise specified |
| **Necrotising periodontal diseases** | A. Necrotizing ulcerative gingivitis  
B. Necrotizing ulcerative periodontitis |
| **Abscesses of the periodontium** | A. Gingival abscess  
B. Periodontal abscess  
C. Pericoronal abscess |
| **Periodontitis associated with endodontic lesions** | A. Combined periodontic-endodontic lesions |
| **Developmental or Acquired deformities and conditions** | A. Localized tooth-related factors that modify or predispose plaque-induced gingival diseases/periodontitis  
B. Mucogingival deformities and conditions around the tooth  
C. Mucogingival deformities and conditions on edentulous ridges |

Table 1-1: Abbreviated Version of the 1999 Classification of Periodontal Diseases and Conditions (Armitage, 1999)

In this thesis, the main clinical conditions that will be discussed are gingivitis, chronic periodontitis and aggressive periodontitis, which are the main periodontal conditions of the human. These are briefly described in the following paragraphs.
1.1.3.1 Gingivitis

Gingivitis is an inflammatory condition of the gingival tissues without evidence of attachment loss or alveolar bone destruction. It results from the inflammatory response to bacterial plaque (Mariotti, 1999). It is reversible and can also occur in cases of reduced periodontium due to previous periodontitis. It has been reported that gingivitis has a varied prevalence with reports ranging from less than 20% to almost 100% of the population (Bimstein et al., 1989; Brown and Löe, 1993; Spencer et al., 1983) and this prevalence peaks among the adolescent and child population (Brown and Löe, 1993). Classic experimental gingivitis studies have demonstrated that clinical gingivitis develops within 21 days of abstinence from oral hygiene practices and that gingival health is restored once these oral hygiene measures are reinstituted (Jensen et al., 1968; Löe et al., 1965; Löe et al., 1967; Theilade et al., 1966). Marginal inflammation accompanies periodontitis but the latter is not a natural progression from gingivitis as there are those who develop gingivitis throughout life but never advance into periodontitis.

1.1.3.2 Chronic Periodontitis

Chronic periodontitis is an inflammatory condition of the periodontium involving inflammation of the gingiva, apical migration of epithelial attachment together with destruction of the alveolar bone and loss of marginal periodontal ligament. It progresses in cycles of bursts of activity and quiescence all of which may vary in duration (Socransky et al., 1984).

According to the 1999 classification of periodontal conditions (Armitage, 2004b) the main clinical features and characteristics of chronic periodontitis are:

- Most prevalent in adults but can occur in children and adolescents
- Amount of destruction is consistent with the presence of local factors
- Subgingival calculus is a frequent finding
- Associated with a variable microbiological pattern
- Slow to moderate rate of progression but may have periods of rapid progression
- Can be associated with local predisposing factors (e.g. tooth-related or iatrogenic factors)
• May be modified by and/or associated with systemic diseases (e.g. diabetes mellitus)
• Can be modified by factors other than systemic disease such as cigarette smoking and emotional stress

1.1.3.3 Aggressive Periodontitis

There are major difficulties when one tries to classify a disease whose aetiology is not fully known or understood, such as aggressive periodontitis. As it has been noted; ‘Many diseases remain of unknown cause; known causes are of diverse types; causation may be complex, with interplay of several factors’ (Scadding, 1996).

Aggressive periodontitis (AgP) can only be diagnosed syndromically by defining a set of signs and symptoms that together form a distinct clinical entity (Tonetti and Mombelli, 1999). The cardinal features of inflammation of the marginal tissues, attachment loss and alveolar bone destruction characterize AgP just as much as in CP. The main differences between CP and AgP are the rate of progression (as the name suggests, ‘aggressive’ is more rapid compared to that seen in CP) the pattern of destruction, clinical signs of inflammation and the relative amount of local factors including plaque and that AgP tends to occur in a younger population than is observed in CP (Armitage, 1999; Armitage et al., 2010; Lang, 1999).

Aggressive periodontitis (AgP) can be further sub-classified as localized or generalized depending on several clinical presentations. The 1999 consensus group for the World Workshop on classification of periodontal diseases and conditions agreed that the term ‘Localised AgP’ (LAP) should replace ‘localised juvenile periodontitis’ (LJP) while ‘Generalised AgP’ (GAP) replace ‘generalised juvenile periodontitis’ (GJP). They also noted that “AgP is a specific type of periodontitis with clearly identifiable clinical and laboratory findings which make it sufficiently different from chronic periodontitis” (Tonetti and Mombelli, 1999).

This claim could be challenged for as of now there are no known diagnostic tests available for the definitive identification of periodontitis or that will reliably distinguish between chronic and aggressive periodontitis (Buduneli and Kinane,
2011). Besides, consensus is lacking in determining the case definition for AgP (Armitage et al., 2010). The only way to separate AgP from CP is by alluding to extra features that are not commonly found in CP.

<table>
<thead>
<tr>
<th>Localized AgP:</th>
<th>Generalized AgP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circumpubertal onset</td>
<td>Usually affecting people under 30 years of age, but patients may be older</td>
</tr>
<tr>
<td>Robust serum antibody response to infecting agents</td>
<td>Poor serum antibody response to infecting agents</td>
</tr>
<tr>
<td>Localized first molar/incisor presentation with interproximal attachment loss on at least two permanent molar teeth, one of which is a first molar, and involving no more than two other teeth other than the first molars and the incisors.</td>
<td>Pronounced episodic nature of the destruction of attachment loss and alveolar bone.</td>
</tr>
<tr>
<td>Generalized interproximal attachment loss affecting at least three permanent teeth other than first molars and incisors</td>
<td></td>
</tr>
</tbody>
</table>

Table 1-2: The specific features of the two different forms of AgP are stated here. Adapted (Lang, 1999; Tonetti and Mombelli, 1999).

The common features of both localized and generalized AgP are:
• Otherwise clinically healthy patients.
• Rapid attachment loss.
• Familial aggregation.

Other secondary features that are generally but not universally present are:
• Amount of microbial deposits are not consistent with the severity of the periodontal disease destruction.
• Elevated proportions of *Aggregatibacter actinomycetemcomitans* (*A.a*) and in some populations, *Porphyromonas gingivalis* (*P.g*), the latter especially in generalised AgP.
• Phagocyte abnormalities.
• Hyper-responsive macrophage phenotype, including elevated levels of PGE2 and IL-1β.
• Progression of attachment loss and bone loss may be self-limiting.

Despite the effort to place periodontal diseases in classes, it is evident that cross-sectional classifications are not predictive of the rate of progression of periodontal disease (Albander et al., 1997). On the other hand, ‘it can be argued that, at present, regardless of the enormous increase in knowledge of periodontal diseases, we still know too little to diagnose and classify the periodontal disease of a patient on an etiologic basis’, (Van Der Velden, 2005) while others believe that the mechanism of etiology of periodontitis is so complex and diverse that it precludes a classification based on etiology (Baelum and Lopez, 2003). It is then not surprising that definitive tests to determine the diagnosis of periodontitis are lacking.

1.1.3.4 Case Definitions for Periodontitis.

In addition to the obvious difficulties of classification apparent from the discussion above, there is no universal consensus as to what clinically constitutes a “case” of chronic or aggressive periodontitis. The availability of universally accepted case definitions of periodontitis may not have great impact in normal clinical practice, but they become of considerable importance for research purposes. For example, epidemiological descriptions of disease prevalence require a clear disease definition, which can be readily applied, preferably dichotomously, to subjects within a population cohort. Similarly, in studies such as described in this thesis comparing putative biomarkers with clinical diagnosis it is essential to have clear applicable diagnostic criteria. Because the syndromic features used to identify periodontal disease tend to be continuous variables (e.g. amount of attachment loss) and also that the condition is site-specific, a single case definition of disease is not intuitively obvious.

Studies have shown that the use of different case definitions has a great impact on the description of the prevalence, severity and extent of periodontal disease (Beck and Löe, 1993; Costa et al., 2009; Page and Eke, 2007). The commonest parameters that are currently utilised to measure the presence,
distribution and severity of periodontal disease both in the clinical situation and in the field of research are probing pocket depth (PPD) and clinical attachment level (CAL). These parameters being continuous variables beg the need for clear thresholds, either side of which the periodontal site or patient is classified as having disease or no disease and its severity and distribution can be established.

A study looking at the impact of using different case definitions based on different thresholds for CAL and PPD on the prevalence of periodontitis demonstrated that as the threshold increases; the prevalence of periodontitis reduces and vice versa (Table 1-3).

<table>
<thead>
<tr>
<th>Case def 1</th>
<th>Case def 2</th>
<th>Case def 3</th>
<th>Case def 4</th>
<th>Case def 5</th>
<th>Case def 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 1 site with PPD ≥ 4mm</td>
<td>≥ 4 sites with CAL ≥ 5mm and ≥ 1 site with PPD ≥ 4mm</td>
<td>≥ 2 teeth with CAL ≥ 6mm and 1 site with PPD ≥ 5mm</td>
<td>≥ 4 teeth with ≥ 1 site with CAL ≥ 3mm and ≥ 4 teeth with ≥ 1 site with PPD ≥ 4mm</td>
<td>≥ 2 sites with interproximal CAL ≥ 4mm not on same tooth or ≥ 2 site with interproximal PPD ≥ 4mm</td>
<td>≥ 2 sites with interproximal CAL ≥ 6mm not on same tooth or ≥ 1 site with interproximal PPD ≥ 6mm</td>
</tr>
</tbody>
</table>

| 65.3% | 23.8% | 16.5% | 38.6% | 41.2% | 13.8% |

Table 1-3: A summary of data on the prevalence of periodontitis according to different case definitions as described in a study looking at the impact of using different case definitions for periodontal research (Costa et al., 2009). Def = definition.

The risk of using a lower threshold to define a case of periodontitis would be the potential to overestimate the prevalence, severity or distribution of the disease in the same patient or population. Another problem is encountered when age of the patient is considered, as loss of periodontal attachment tends to be more severe in the elderly both as a result of cumulative recession from disease and other non-inflammatory reasons.
CAL is a more robust measure of periodontal disease severity, however, loss of periodontal attachment can also occur without the presence of periodontitis as occurs in cases with subgingival restorative overhangs or in the distal surface of the second molars after extraction of the third molar (Page and Eke, 2007). A useful case definition would have to be able to capture a true case of periodontitis. The more informative way would be to use both PPD and CAL (Costa et al., 2009) although this by itself would not entirely exclude the shortfalls described above.

The Centres for Disease Control and Prevention (CDC) and the American Academy of Periodontology (AAP) has proposed such a definition (≥ 2 sites with interproximal CAL ≥ 4mm not on same tooth or ≥ 2 sites with interproximal PPD ≥ 5mm) as the gold standard for periodontitis representing moderate periodontitis. Any cases below this threshold are classified as mild or healthy while severe periodontitis is defined as greater than 2 sites with interproximal CAL ≥ 6mm not on same tooth and ≥ 1 site with interproximal PPD ≥ 5mm (Page and Eke, 2007). Including the PPD ≥ 5mm in the threshold reduces the chances of assigning a case as “severe” based only on the CAL, which may overestimate periodontitis severity in the elderly or successfully treated cases due to accumulative loss of periodontal attachment (Page and Eke, 2007).

Overall, it seems unlikely that a consensus will be reached in the near future on what cut offs to use. It then lies on the clinician/researcher to decide what suits their needs best. However, the CDC/AAP definition seems to address many of the issues raised and will form the basis of case definition for the cases described in this thesis.

1.1.3.5 Use of Current Clinical Parameters for Periodontal Diagnosis
Currently the diagnosis of periodontal diseases relies on the assessment of a range of clinical parameters. The features of all periodontal diseases include change of colour and texture of the gingiva, e.g. reddening, swelling and tendency to bleed on probing. In addition destructive periodontal conditions such as aggressive and chronic periodontitis exhibit attachment loss, typically associated with increased pocket depth, recession and alveolar bone loss.
Other signs and symptoms include furcation involvement and mobility. Radiographic findings also shed light on extent and severity of the disease.

However, although clinical assessment is the gold standard for periodontal diagnosis today, there are considerable limitations to these assessments, some of which are outlined in Table 1-4.

Furthermore clinical diagnostic parameters provide only limited information on treatment options, prognosis, and future progression and require the expertise of a highly trained dental professional in appropriate clinical settings. Periodontal disease has an episodic nature of progression impeding the process of accurate monitoring of the disease. Measurement of molecular biomarkers of periodontal disease would allow rapid accurate diagnoses and dynamic monitoring of disease activity (Herr et al., 2007). Consequently there is considerable interest in identifying new diagnostic markers which would provide additional clinical information over and above those obtained by a clinical examination, which might have both diagnostic and prognostic utility, and would allow simple point of care or even home use diagnosis.
<table>
<thead>
<tr>
<th>Clinical Parameter</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| Gingival colour change, Swelling   | • Subjective  
• Difficult to detect in darkly pigmented gingiva                                                                                                                                                   |
| Bleeding on probing                | • Can't differentiate gingivitis, chronic periodontitis and aggressive periodontitis  
• Uncomfortable to patient  
• May be masked in smoking  
• Not a good prognostic marker   |
| Probing depth                      | • Not frequently reproducible between and within examiners due to variables such as degree of inflammation, probing force (Gabathuler, 1971), type of probe, angulation, root or tooth morphology, local factors e.g. calculus. (Listgarten et al., 1976; Listgarten, 1980; Magnusson, 1980).  
• Over or underestimation of the probing depth even when using pressure sensitive probe (Polson et al., 1980).  
• Measures past disease experience  
• Uncomfortable to patient  
• Does not necessarily measure attachment loss or gain (Listgarten, 1980).  
• Difficult to interpret the significance of probing measurements |
| Attachment level                   | • Similar challenges as in probing depth  
• Difficult to interpret the significance of attachment level measurements                                                                                                                                 |
| Tooth hypermobility                | • Not a good measure (there are other causes of hypermobility)  
• Subjective to a certain degree                                                                                                                                                                             |
| Furcation involvement             | • Difficult if soft tissue is covering entry  
• Uncomfortable to patient                                                                                                                                                                                                 |
| Radiographic findings             | • A two dimensional view  
• Covering structures such as bone tissue and teeth make it difficult to see the lingual and buccal bone crest (Lindhe et al., 2006).  
• Reproducibility is a challenge                                                                                                                                                                           |
| Sounding                          | • Uncomfortable to patient                                                                                                                                                                                                 |

Table 1-4: Highlight of the limitations of current clinical diagnostic parameters
1.1.4 Diagnostics And Biomarkers

The characteristics of an ideal diagnostic test have been proposed. The test should be (Chapple et al., 1993; Chapple, 1997):

- Quantitative
- Highly sensitive and capable of analyzing a single periodontal site in health or disease
- Highly specific
- Reproducible
- Simple to perform
- Rapid
- Non-invasive
- Versatile in-terms of sample handling, storage and transport
- Amenable to chair-side use
- Economical
- Dependent on simple and robust instrumentation.

Most of these qualities cannot be demonstrated exclusively in any existing current single test. There is need therefore for the continued search for diagnostic tests especially those based on multiple biomarkers. It is hoped that multiple biomarkers would improve the specificity and sensitivity of the test. A brief description of the biomarker discovery process is provided below.

1.1.4.1 What are biomarkers?

A biomarker is "a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" (Atkinson et al., 2001).

In all biological systems, it is the proteins that carry out the molecular functions of the human body such as the generation of energy, production of cellular components, degradation of waste products, regulation of cellular processes and fighting disease (Albeck et al., 2005) hence the focus on protein markers in many areas of medical research as potential biomarkers.
Potential markers to study the presence and characteristics of periodontal disease have been classified as follows (Chapple, 1997): Markers indicating the presence of putative periodontal pathogens, markers of periodontal or gingival inflammation, markers of the host’s immune response to certain pathogens and markers of tissue destruction. Regarding periodontal diseases, the potential sources for such disease markers are oral fluids (including saliva and GCF), plasma/serum, subgingival plaque and tissue biopsies. The focus of this thesis is biomarker discovery for periodontal disease based on the oral fluids especially saliva.

1.1.4.2 The process of biomarker discovery

In general, the biomarker discovery process involves a series of well-defined stages from clinical diagnostic need and study design right through to clinical implementation. These stages are illustrated in the flow chart in Figure 1-2. This flow chart has informed the entire experimental approach to the studies described in this thesis.
Figure 1-2: Flow chart of the biomarker discovery process. Adapted from (Ciphergen, 2007)
The final stage (Clinical assay development) is represented in a different color to highlight the final goal of this work but also to reveal that the process could not be accomplished during the time stipulated for this PhD program. It will however be an obvious starting point for those coming behind to carry the work forward towards the ultimate goal of realizing clinical diagnostic utility.

1.1.4.3 Oral fluids as sources of biomarkers

In dealing with a site-specific disease like periodontitis, potentially GCF has the most relevant markers that are specific to the site besides being probably one of the earliest to reflect changes in the metabolic status of the surrounding periodontal tissues. On the other hand, salivary markers would be more representative of the overall periodontal status of the patient. This makes salivary biomarkers most desirable because most therapeutic interventions are patient-based rather than targeted to specific sites. It is with this view in mind that the search for salivary biomarkers for periodontal disease will form the bulk of this thesis with GCF analysis providing the initial clues as to what the potential biomarkers could be in saliva.

1.1.4.4 Saliva

Saliva is derived from three paired major salivary glands and a host of other submucosal glands in the oral cavity. The constituents of the saliva from each of the major glands are unique to the gland. The product in the oral cavity is referred to as whole saliva (WS): It consists of the secretions from the salivary glands, serum and blood products such as neutrophils and their products, GCF and other products from the oral mucosa including epithelial cells, electrolytes, microorganisms and their products (Edgar, 1992) Figure 1-1.

There are reportedly over 1100 proteins that are unique to WS and not found in glandular saliva nor in plasma (Loo et al., 2010). These proteins in WS may be derived from several sources including the host defense system, oral epithelium, the oral microflora and diet. This results in a very complex mixture that reflects the metabolic status of the tissues involved in its production and its immediate environment. It is not surprising then that the discovery of salivary-based diagnostics has been a very attractive approach in the past decade for most disease conditions. For instance saliva has been used successfully to diagnose HIV-1 and -2, oral cancer and viral hepatitis A, B and C. It has also
been very useful in monitoring some drugs such as marijuana, cocaine and alcohol (Mandel, 1993).

Figure 1.3: Venn diagram showing the overlapping protein identification among plasma, whole saliva and parotid/SM/SL. Adapted from (Loo et al., 2010).

Saliva collection is simple and in most cases does not require the addition of preservatives or other chemicals to collect or store, as is the case with serum or plasma. It is also a painless process that can easily be done at the chair side or at home by unskilled persons. Minor challenges in collection may be experienced among special groups such as infants, those with mental disabilities and the elderly. However most of these challenges can be easily circumvented by available techniques for saliva collection such as use of absorbent media for saliva collection in infants. Saliva-based diagnostics are therefore less expensive and present less risk to both the patient and the provider than current methodologies (Segal and Wong, 2008). This makes it easily tolerable for many patients (Koka et al., 2008).

1.1.4.5 GCF
Gingival crevicular fluid (GCF) is an inflammatory exudate that seeps into the gingival sulcus or periodontal pocket around teeth with gingival inflammation. Its main constituents are serum and its products and a battery of locally
generated tissue breakdown products, inflammatory mediators and antibodies against dental biofilm and bacteria and their products such as bacterial enzymes (Armitage, 2004a; Cimasoni, 1983).

Since the 1960s when GCF analysis was first advocated for the detection of inflammation in the gingival and periodontal tissues (Brill, 1960), there has been a steady increase in the number of studies interested in the diagnostic value of GCF. Quantitative analysis of GCF has shown that the volume of GCF is consistently increased in inflammation (Smith et al., 1992; Uitto, 2003) and therefore may be a good indicator of disease activity though this has been questioned and suggested that GCF flow rate rather than volume is more reliable in reflecting disease activity (Griffiths et al., 1992).

Destructive periodontal diseases are site specific and as such there are times when a site-specific diagnosis and monitoring is desirable for better site-specific management. The specific markers in each site would be directly related to the disease process i.e. etiology and pathogenesis. Site-specific approach to diagnosis would be important in treatment planning of individual teeth as well as in the management of localized lesions for instance in diagnosing, treatment planning and evaluating of treatment outcome in sites not responding to previous therapy, or those needing guided tissue regeneration (GTR) (Persson, 2005; Tonetti and Cortellini, 1997) and dental implant therapy.

Cross-sectional studies have attempted to identify disease markers for periodontal disease using GCF analysis. However these have often been unable to distinguish gingivitis sites from those with periodontitis because the main markers in cross-sectional studies are enzymes or inflammatory markers, which apparently are increased in both conditions (Armitage, 2004a). As such, inflammatory markers are not reliable as diagnostic markers (Behle et al., 2009). Experimental studies on the other hand have been able to demonstrate significant associations between periodontal breakdown and increases in the GCF content of a large number of markers (Chambers et al., 1984; Giannobile et al., 1995; Nelson et al., 1992; Offenbacher et al., 1986; Shibutani et al., 1993). Table 1-5 is a summary of a few of the markers targeted in GCF. One
of the main limitations of previous GCF studies has been the inability to identify and measure the levels of all the possible potential biomarkers.

<table>
<thead>
<tr>
<th>Cathepsins</th>
<th>Immunoglobulin-degrading enzymes</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Cysteine proteinase (B, H, L)</td>
<td></td>
</tr>
<tr>
<td>• Serine proteinase (G)</td>
<td></td>
</tr>
<tr>
<td>• Cathepsin D</td>
<td></td>
</tr>
<tr>
<td>Alkaline phosphatase (ALP)</td>
<td>Glycosidases</td>
</tr>
<tr>
<td>Acid phosphatase (ACP)</td>
<td>Dipeptidyl peptidases</td>
</tr>
<tr>
<td>β-glucuronidase</td>
<td>Nonspecific neutral proteinases</td>
</tr>
<tr>
<td>Elastase</td>
<td>Collagenases</td>
</tr>
<tr>
<td></td>
<td>• MMP-1, 3, 8 and 13</td>
</tr>
<tr>
<td>Elastase inhibitors</td>
<td>Gelatinases</td>
</tr>
<tr>
<td></td>
<td>• MMP-2 and 9</td>
</tr>
<tr>
<td>• α2-macroglobulin</td>
<td></td>
</tr>
<tr>
<td>• α1-proteinase inhibitor</td>
<td></td>
</tr>
<tr>
<td>Aspartate aminotransferases</td>
<td>Tissue inhibitor of MMP-1 (TIMP-1)</td>
</tr>
<tr>
<td>Trypsin-like enzymes</td>
<td>Stromelysins</td>
</tr>
<tr>
<td>Myeloperoxidases</td>
<td>Lactate dehydrogenase</td>
</tr>
<tr>
<td>Arylsulfatase</td>
<td>Creatine kinase</td>
</tr>
<tr>
<td>β-N-acetyl-hexosaminidase</td>
<td></td>
</tr>
</tbody>
</table>

Table 1-5: A sample list of some of the host derived enzymes and their inhibitors in GCF. Adapted (Armitage, 2004a)

A recent systematic review concluded that, “No factor in GCF can distinguish between aggressive or chronic periodontitis” (Loos and Tjoa, 2005). The demonstration of a diagnostic and or prognostic biomarker in GCF would therefore be a great breakthrough in the diagnosis, classification and management of periodontal diseases.

In summary, the main challenge has been how to get all the diagnostic information from these oral fluids. As in many other diseases, periodontitis is heterogeneous in its etiology and thus the use of a single biomarker to offer insight in to diagnosis is arguably not adequate. Pattern-based diagnostics with probabilistic scoring schemes can adapt and model the complexity present in a heterogeneous disease (Johann et al., 2004). For this reason, new technologies such as SELDI-TOF mass spectrometry (MS) coupled with
bioinformatics could provide quick multiple biomarker identification and isolation of patterns that predict or identify the presence of periodontal disease. Protein identification of these biomarkers followed by validation studies would establish the utility of the biomarkers. It may be possible also to use the differential expression of these proteins or peptides to differentiate the various forms of periodontal disease, most usefully to differentiate periodontitis from gingivitis or between chronic periodontitis and aggressive periodontitis.

1.1.4.6 Recent Studies on Biomarker Discovery

There has been a lot of interest in the past decade for the discovery of diagnostic biomarkers and prognostic biomarker for periodontal disease in both saliva and GCF using different techniques. A lot of these have been based on identification of molecular biomarkers in GCF and far much less in saliva. These have included host inflammatory markers (e.g. C-reactive protein, IL-1beta (IL-1β), IL-6, and tumour necrosis factor –alpha (TNF- α)), connective tissue destruction markers (e.g. α2-macroglobulin, matrix metalloproteinases, tissue inhibitors of metalloproteinases (TIMPs), aminotransferases, Cathepsin G) and bone remodeling markers (e.g. alkaline phosphatase, osteocalcin, RANKL, OPG and C-terminal type 1 collagen degradation products such as pyridinoline cross-linked carboxyterminal telopeptide of type I collagen (ICTP)) (Buduneli and Kinane, 2011; Christodoulides et al., 2005; Fox, 1993; Gemmell et al., 1997; Kaufman and Lamster, 2000a; Lamster and Grbic, 1995; Miller et al., 2010; Pederson et al., 1995; Seymour and Gemmell, 2001) Figure 1-4. According to a recent review of the hundreds of the GCF-based biomarkers so far described, only few have shown some ability to define a case of periodontitis thus differentiating it from health or gingivitis, as well as being informative regarding the response to treatment (Buduneli and Kinane, 2011; Loos and Tjoa, 2005).
Of the reported oral fluid based biomarkers currently, matrix metalloproteinase 8 (MMP-8) seems the most promising as a potential biomarker for both diagnosing and possibly prognosticating periodontitis (Giannobile et al., 2009; Loos and Tjoa, 2005). MMP-8 (collagenase-2) is a neutrophil collagenase. Its active form, aMMP-8, is known to be the main host cell-derived collagenase leading to periodontal tissue destruction as a result of the degradation and remodeling of the extracellular matrix in particular the gingival and periodontal ligament collagen (Birkedal-Hansen, 1993; Buduneli et al., 2007; Lee et al., 1995; Reynolds et al., 1994; Sorsa et al., 2006). Levels of active MMP-8 have been shown to correlate with clinical and radiographic parameters of periodontitis (Buduneli et al., 2007; Costa et al., 2010).

However, it has been suggested that detection of both latent (precursor) and aMMP-8 in oral fluids may not necessarily correlate with periodontal disease progression, as the active form is associated with periodontitis, while the latent form is associated with gingivitis (Sorsa et al., 1988; Sorsa et al., 2006; Sorsa et al., 2010; Uitto et al., 1990). Besides, a broad range of intra-individual and inter-individual aMMP-8 values has been demonstrated in cases with and
without disease (Kraft-Neumarker et al., 2011). Also, in a study comparing MMP-8 activity levels between periodontitis and gingivitis, no differences were found implying that MMP-8 is clearly elevated due to the inflammatory process, and therefore not discriminatory for periodontitis (Loos and Tjoa, 2005; Nomura et al., 1998). The reliability of MMP-8 is further brought to question when a systemic condition such as diabetes commonly associated with periodontitis is present since MMP-8 levels are independently elevated in both conditions (Costa et al., 2010) making it a useless diagnostic biomarker in diabetic patients. For a biomarker to be useful, it should ideally be able to differentiate between gingivitis and periodontitis, both inflammatory conditions as well as be reliable in detecting periodontitis in the presence of co-morbid conditions such as diabetes and osteoarthritis.

Recent studies in our research group as part of a PhD thesis (Clarke, 2007) and work that I carried out as part of my masters dissertation (Mulli, 2008) have utilised SEDLI-TOF MS to investigate the presence of periodontal biomarkers in Saliva and GCF respectively. They are described in some detail below as they form the foundation work to the studies described in this thesis, and they have not yet been published in the literature.

1.1.5 ProteinChip® Technology

Proteomics can be described as a method of characterizing the entire complement of proteins within a biologic specimen. “Clinical proteomics” is the application of available proteomics technologies to current areas of clinical investigation (Clarke et al., 2003). This technology provides an opportunity to simultaneously and comprehensively study changes in large numbers of proteins in relationship to health and disease.

Currently, proteomic technology combines high throughput analytical methodologies such as Surface Enhanced Laser Desorption/Ionization and Time of Flight mass spectrometry (SELDI-TOF MS) methods with complex bioinformatics to study systems biology (Clarke et al., 2003). Proteomic analyses are already widely in use for clinical studies ranging from cancer (Petricoin III et al., 2002) cardiovascular disease, organ transplant, pharmacodynamic studies, (Clarke et al., 2003; Hong et al., 2006) and
degenerative diseases such as Alzheimer's and diabetes (Qiu et al., 2005; Simonsen et al., 2008).

An important characteristic of proteins is that their functional activity is often modulated by changes in their structure. Conventional approaches used to assay proteins have variable ability to detect these changes, and may depend on the specificity of the antibody to the original or altered forms of the proteins. Additionally, a conventional assay may inadvertently measure only one form of a protein while many may exist. SELDI-TOF MS thus has an advantage over traditional assays in characterizing and quantifying all disease markers (Ciphergen, 2007).

ProteinChip arrays selectively bind and retain whole classes of proteins from complex samples for detection by the ProteinChip SELDI reader. Each array has either a chemically treated surface (such as cationic, anionic, metal affinity, hydrophobic, or hydrophilic) or a reactive surface that can be coupled with biologically relevant molecules (such as antibodies or receptors) for specific interaction with proteins of interest. Selected washes create protein maps on the ProteinChip array surface, and the mass profile of the proteins bound to each surface is detected quantitatively in minutes by the ProteinChip SELDI reader Table 1-5.

Each ProteinChip array has eight chemically active sites, or spots, where sample binding occurs. The array format resembles one column of a standard microplate and facilitates simultaneous analysis of multiple samples. A hydrophobic coating retains samples on the spots, and a bar code allows quick identification of array type. Because molecules bind through specific chemical interaction with the array surface, it is often possible to learn about a protein's chemical properties simply by using ProteinChip arrays.
Typically, 5 microliters (µl) of sample applied on the ProteinChip array can yield sufficient protein for analysis. This is an advantage especially in cases where tissue or fluid samples are not in abundance as occurs in GCF collection.

**CM 10 Array**

The weak cation exchange ProteinChip CM10 array can be used to analyze molecules with a positive charge on the surface. The active spots contain weak anionic carboxylate groups that interact with the positive charges on the surface of target proteins, e.g., containing lysine, arginine or histidine residues. The surface binds proteins that are positively charged at a given pH. To generate selectivity, the pH of the binding buffer is increased or decreased, depending on the need. By decreasing the pH of the binding/wash buffer, an overall net positive charge is imparted on a greater number of proteins within the sample and the result is more binding. By increasing the pH of the binding/wash buffer, an overall net negative charge is imparted on the proteins, resulting in less binding (i.e., more specificity). Binding of proteins to ProteinChip CM10 arrays can also be affected by changing the ionic strength of the buffer. By increasing the ionic strength, competition is generated between the charged protein on the surface and the buffer ions, causing weakly bound proteins to elute from the array surface (i.e., more specificity).
CM10 arrays have a hydrophobic barrier for sample containment and are the array of choice for weak cation exchange applications as demonstrated in Figure 1-6 below (Bio-Rad Laboratories, 2007).

![Image of CM10 array with protein binding](image)

**Figure 1-6** ProteinChip CM10 array surface chemistry with protein. Adapted from (Bio-Rad Laboratories, 2007).

**Immobilized Metal Affinity Capture (IMAC) Array**

The ProteinChip IMAC30 array can be used to capture molecules that bind polyvalent cationic metals such as nickel, gallium, copper, iron, and zinc. The active spots contain NTA groups on the surface that chelate the metal ions. Proteins applied to the array surface may bind to the chelated metal ion through histidine, tryptophan, cysteine, and phosphorylated amino acids. To generate selectivity, the binding and/or wash buffers may contain increasing concentrations of competitors (e.g., imidazole), which compete with the metal on the NTA group for binding to the protein or peptide. ProteinChip IMAC30 (Figure 1-7) array has a hydrophobic barrier for sample containment and is the array of choice for metal affinity applications (Bio-Rad Laboratories, 2007).
Exploratory work done as a proof of principle on the use of SELDI-TOF MS as a tool for biomarker discovery in periodontal disease done as part of a PhD thesis (Clarke, 2007) and a masters dissertation (Mulli, 2008) in our department paved the way to this current work. These studies involved assaying both saliva and GCF for diagnostic markers using SELDI-TOF MS and are briefly described below. Both studies utilised CM10 and IMAC30 ProteinChips® for both peptide and protein arrays. In the GCF study, the mean age of the subjects was 44yrs (Range 19yrs – 75yrs) while the male to female ratio was 0.85. Table 1-6 shows a summary of the clinical parameters from the 24 subjects according to their diagnosis, highest probing pocket depth (PPD), bleeding on probing (BoP), plaque and percentage of teeth with mobility in a dichotomous manner.

Univariate analysis using the expression difference mapping (EDM) tool in the Ciphergen ProteinChip software revealed a total of 159 peaks. These peaks represent the corresponding spots or bands in a 2D-PAGE gel. The anionic (CM10) arrays had 36 and 48 peaks for the protein and peptide arrays respectively while the metal affinity (IMAC) arrays had 35 and 40 peaks for the protein and peptide arrays respectively. When all the arrays were considered, the EDM analysis revealed 76 peaks that were significantly different between
gingivitis and periodontitis samples. 77.6% of these had higher relative peak intensities in periodontitis than in gingivitis (Mulli, 2008).

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>n</th>
<th>Highest PPD (Av.) in mm</th>
<th>BoP (%)</th>
<th>Plaque (%)</th>
<th>Mobility (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gingivitis</td>
<td>8</td>
<td>5 (3.6)</td>
<td>34</td>
<td>53</td>
<td>0</td>
</tr>
<tr>
<td>CP</td>
<td>8</td>
<td>9 (6.6)</td>
<td>69</td>
<td>73</td>
<td>57</td>
</tr>
<tr>
<td>AgP</td>
<td>8</td>
<td>10 (7)</td>
<td>88</td>
<td>66</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 1-6: A summary of patient-based clinical data for the GCF study (Av.) = Average, CP = Chronic periodontitis and AgP = Aggressive periodontitis.

In the saliva study, 90 patients were recruited with the diagnosis of gingivitis (n=31), CP (n=29) or AgP (n=30). Data obtained for saliva samples using the same method as for GCF (above) yielded a total of 270 peaks of which 41 were differently expressed between gingivitis and periodontitis with peaks ranging from 3158 to 28595 Da. Twenty five (61%) of these peaks were overexpressed in periodontitis than in gingivitis (Clarke, 2007).

Overall, these two studies demonstrated that SELDI-TOF mass spectrometry could profile protein and peptides in WS and GCF samples in subjects with or without periodontal disease and identify those that may have diagnostic discrimination (Clarke, 2007; Mulli, 2008). Of all the peaks detected in both WS and GCF, at least fourteen showed significant ability to distinguish between gingivitis and chronic periodontitis with high specificity and were thus scheduled for protein identification in this thesis. Most of these proteins and or peptides that are differentially expressed between gingivitis samples and disease cases using SELDI-TOF MS are of unknown identity.

However from published mass spectrometry (MS) data, five of these peaks were identified as belonging to the cationic antimicrobial peptides human β-defensin 1 (hBD1) with a molecular weight 5060 Da, α-defensins human neutrophil peptide (HNP) – HNP 1 (mw 3443 Da), HNP 2 (mw 3372 Da) and
HNP3 (mw 3487 Da) (Figure 1-8) and human cationic antimicrobial peptide 18 (hCAP-18) also known as cathelicidin (LL-37) of mw 4493 Da (Dommisch et al., 2009; Lundy et al., 2005). These will be briefly discussed next.

**Figure 1-8: Comparative IMAC30 ProteinChip array tracings for periodontitis and gingivitis in GCF showing higher relative peak intensity of peptides corresponding to HNP 2, 1 and 3 respectively in periodontitis.**

### 1.1.6 Antimicrobial Peptides

Antimicrobial peptides (AMPs) are a group of basic cationic peptides rich in cysteine and arginine usually composed of about 29-40 amino acids found in both vertebrates and plants. They were initially identified in pulmonary macrophages of rabbits (Patterson-Delafield et al., 1980) and later in human neutrophils (Ganz et al., 1985). The 2 major antimicrobial peptides in human are the alpha and beta defensins. Their key differences are in their cysteine pairing and disulphide pairing. The alpha defensins have disulphide links on 1-6, 2-4 and 3-5 while beta defensins are linked at 1-5, 2-4 and 3-6 (Ganz and Lehrer, 1995; Liu et al., 1997; Tang and Selsted, 1993).

The alpha defensins are further classified as human neutrophil peptides 1-4 (HNP) and human defensins 5-6 (HD 5-6). HNPs are found mainly in azurophilic granules of neutrophils (Gabay et al., 1989; Ganz et al., 1985; Ganz and Lehrer, 1995) while HD 5-6 are found in paneth cells of the intestinal tract (Jones and Bevins, 1992; Jones and Bevins, 1993). Human beta defensins (HBD) are of 3 types i.e. 1, 2 and 3 and are all found in epithelium (Zhao et al., 1996).
The genes for both HNP1 and HBD-3 are closely located at region p23.1 of chromosome 8 (8p.23.1) and the two are separated by a distance of about 100-150 kb (Liu et al., 1997). This may reflect their structural and or functional similarities and may suggest synchronized regulation and expression. The impact of genetic influence on expression and clinical sequel of antimicrobial peptides has been demonstrated in studies on gene polymorphisms (Aldred et al., 2005; Hollox et al., 2003; Hu et al., 2004; Jurevic et al., 2002; Jurevic et al., 2003; Levy et al., 2005; Linzmeier and Ganz, 2005). For instance, the amount of HNP-1 and HNP-3 peptides expressed in neutrophils is reported to be proportional to the combined copy number of DEFA1 and DEFA3 (Linzmeier and Ganz, 2005). It would be interesting to demonstrate the relationship of these polymorphisms and periodontal disease.

The source of derivation of HNP 1-3 in saliva is not fully understood. It has been demonstrated that HNPs are not present in any of the major salivary glands (Mizukawa and Nobuyoshi, 1998). However, many neutrophils migrate into the gingival sulcus and to the oral cavity (Ashkenazi and Dennison, 1989). It is therefore plausible that a significant amount of HNP 1-3 may originate from the neutrophils found in the gingival crevice (Hosokawa et al., 2006) although evidently neutrophils are not the sole source, but that other cells such as lymphocytes, monocytes and other dendritic cells are capable of producing them (Agerberth et al., 2000; Puklo et al., 2008; Rodriguez-Garcia et al., 2010).

The production of HNPs is usually constitutive, probably for the regulation of oral commensals, to keep them at levels that can be accommodated by the host. However stimulatory production occurs during pathologic states such as trauma, infection by bacteria, fungi or viruses and neoplasia (Albrethsen et al., 2005; Boehncke, 2004; Cunliffe et al., 2002; Diamond and Bevins, 1998; Mizukawa and Nobuyoshi, 1998; Mizukawa et al., 1999a; Mizukawa et al., 1999b). It is this stimulatory production, which accounts for the increased levels of HNP 1-3 reported in disease conditions (Hosokawa et al., 2006; Mizukawa et al., 1999b; Puklo et al., 2008). HNP1-3 production in GCF has been shown to be correlated with the load of periodontopathogenic bacteria especially
Porphyromonas gingivalis (P. gingivalis), Tannerella forsythia (T. forsythia) and Treponema denticola (T. denticola) (Puklo et al., 2008).

Studies using ELISA kits for HNP 1-3 report salivary concentrations of 0.8µg/ml. This however increases with oral infection, neoplasia or inflammation (Mizukawa et al., 1999a). This pattern of expression is similar to that in plasma and is associated with influx of PMNs in the sites of pathology (Panyutich et al., 1993). As expected HNP 1-3 levels in GCF are higher than in saliva. According to a recent study, the GCF HNP 1-3 levels from patients with CP were 73.92 µg/µl (range 0.40–116.17 µg/µl), AgP (median 17.64 µg/µl, range 0.36–112.52 µg/µl), and healthy controls (median 1.20 µg/µl, range 0.01–29.18 µg/µl) (Puklo et al., 2008).

HNP 1-3 play different vital roles including antimicrobial activity against most of the oral pathogens, regulation of the adaptive immunity, stimulation of epithelial proliferation and possibly biosynthetic and tissue remodeling (Ichinose et al., 1996; Mizukawa et al., 1999a; Proha´szka et al., 1997; van den Berg et al., 1998; Yang et al., 2000; Yang et al., 2002). HNPs are cytotoxic to tumour cells (Lichtenstein et al., 1986) and to human fibroblasts in concentrations above the physiologic levels. This dual nature would suggest that they can be involved both in the defense of the periodontal tissues from destruction as in inflammatory periodontal disease as well as participating in the destruction of the periodontium when their production is hyper-regulated. The interaction between HNPs and other immune cells plays an important role in maintaining periodontal health. For instance, HNP 1-3 have been shown to be chemotactic to immature dendritic cells and human T lymphocytes thereby recruiting the adaptive immunity to protect the periodontium and modulate dendritic cell generation and maturation (Chertov et al., 1996; Rodriguez-Garcia et al., 2010; Yang et al., 2000). These peptides further lead to the production of pro-inflammatory cytokines such as interleukin-1 alpha (IL-1α) that play a key role in the orchestration of the periodontal immune response. HNPs 1-3 work synergistically with LL-37 (Nagaoka et al., 2000). However, to the best of my knowledge, there are no data reporting the differential expression of HNPs 1-3 in saliva in health, gingivitis and periodontitis.
1.2 Summary And Aims

There is a well recognized need for the identification and validation of biomarkers in periodontology for use in diagnosis, prognostic assessments, as a research tool and possibly to identify novel targets for therapeutic intervention. Both saliva and GCF have been tested for this purpose but to date no highly reliable markers have been identified. Our recent work with SELDI-TOF MS biomarker discovery studies suggests that a range of peptides and proteins have considerable potential as biomarkers, although the identity of these markers has not been shown. Not withstanding this, some of the peptide markers putative identified by the SELDI have peak values suggestive of them being HNP 1-3 and cathelicidin LL-37. To the best of my knowledge, there are no data reporting the differential expression of HNPs 1-3, LL-37 and Protein S100A8 in saliva in health, gingivitis and periodontitis.

Therefore the overall aim of the studies in this thesis are to test the hypothesis that profiles of protein and peptide markers of saliva can be identified which are able to distinguish between gingivitis, chronic and aggressive periodontitis with reliability.

1.2.1 Specific objectives

1. To confirm the identity of putative antimicrobial peptides in GCF and Saliva.
2. To test the ability of salivary HNP and LL-37 concentrations to distinguish between different periodontal diagnoses with high reliability
3. To establish the identity of other potential biomarkers identified in initial biomarker discovery studies.
4. To test the diagnostic utility of these identified biomarkers
5. To validate these pilot observations by testing the assays on a newly collected independent patient population.
6. To determine the inherent variation of these biomarkers for the range of parameters including time, influence of eating, sample collection and storage conditions and response to treatment.
7. Determine the relationship of copy number polymorphisms of HNP 1-3 gene (DEFA1A3) and the corresponding periodontal status.
Chapter 2 - Protein Identification
2.1 Introduction

As previously highlighted, the studies leading to this thesis were based on the identification of diagnostic and prognostic biomarkers in both saliva and GCF using the SELDI-TOF MS platform. One of the challenges of using SELDI-TOF MS for biomarker discovery has been the lack of direct identification of the discriminatory peaks discovered (Clarke, 2007; Maeland Nilsen et al., 2011; Mulli, 2008).

Once the potential biomarkers mentioned in the preceding chapter were detected using SELDI-TOF MS, (Table 8-1 in Appendices) the next logical step was to characterize their identity. It was envisaged that identifying these potential biomarkers would facilitate the development of assays to detect them in clinical samples for validation purposes and potentially for the development of future diagnostic assays.

There are a number of techniques available that can be used for protein identification including one- or two-dimensional polyacrylamide gel electrophoresis (1D- or 2D-PAGE), Western blotting and mass spectrometry-based methods. Traditionally, gel based techniques have been used. However with the advent of mass spectrometry techniques, the process of protein or peptide identification has greatly improved. Each of these methods has its pros and cons. However, mass spectrometry (MS) is the most comprehensive method for extensive proteomics studies (Yates et al., 2009). For instance 1D- and 2D-PAGE are cheap to run when compared to establishing a mass spectrometry based system. However, gel-based techniques lack the dynamic range to detect a large number of analytes. In addition, they require a large volume of sample to analyze. Their resolution is also inadequate in being able to detect the difference between two peptides of similar molecular weight as well as being prone to masking by the more abundant or larger proteins.

Mass spectrometry on the other hand has a wider dynamic range. There is the added advantage that the analytes can be digested with specific enzymes into smaller segments of amino acid chains that are unique to specific proteins as well as identifying specific modifications of the peptides/proteins such as phosphorylations and methylations. These post-translational modifications
(PTMs) are important in that they may be associated with pathology or specific metabolic processes and as such may be biomarkers by themselves.

The study of proteins and peptides using mass spectrometry (MS) has been further advanced by the introduction of sample ionization techniques including electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI). These can be coupled with mass analyzers including quadruple/linear ion trap, time of flight (TOF), quadruple TOF (QTOF), Orbitrap, and the Fourier transform ion cyclotron resonance (FT-ICR) among others. Furthermore, multiple mass analyzers can be used in tandem to increase the resolution, sensitivity, accuracy and throughput for protein identification (Aebersold and Mann, 2003; Binz et al., 2003; Calligaris et al., 2011; Compton et al., 2011; Cristoni and Bernardi, 2003; Dreger, 2003; Fang et al., 2007; Hu et al., 2006b; Kelleher, 2004; Seraglia et al., 2003). Due to the high degree of sensitivity, only very minute volumes of sample are needed for analysis using mass spectrometry, which is unlike the gel-based techniques (Wong, 2008).

There are broadly speaking two major principles in the proteomics systems, namely the “bottom-up” and the “top-down” approach.

2.1.1 Bottom up approach

This is the more common of the approaches and involves breaking down the complex protein structure into smaller fragments that are then sequenced by the mass analyzer. Most of the fragmentation or digestion is done by proteolytic enzymes especially trypsin. It is possible to incorporate 1D- or 2D-PAGE prior to the trypsin digestion in order to isolate protein complexes of a particular molecular weight (mw) and abundance therefore clarifying the product further before MS/MS sequencing of the peptides. Pre-fractionation of samples using ion exchange chromatography, gel filtration chromatography and capillary isoelectric focusing among other techniques can also be very useful in clarifying the sample before further processing using gel-based techniques coupled to MS/MS (Hu et al., 2006b).
CHAPTER 2 – PROTEIN IDENTIFICATION

2.1.2 Top-down approach

Sometimes, it may be useful to measure proteins in their native form before being digested by proteolytic enzymes. This process is referred to as the “top-down” approach to proteomics (Calligaris et al., 2011; Kelleher, 2004). The early top-down proteomics were originally based on 2-D gels that separated a complex mixture of sample into protein spots visualized by staining. This “top-down” technique focused largely on intact protein albeit at low molecular resolution as highlighted above. Current techniques utilise MS platforms of different types including the ESI and MALDI, FT-ICR and LCMS/MS (Kelleher, 2004; Yates et al., 2009). With these techniques, it is possible to interrogate the whole protein including the important characteristics such as post-translational modifications (PTMs). This is useful especially when the goal is biomarker discovery since the intact protein and possible PTMs may be reflective of disease.

The top-down approach is not as universally used as the bottom-up approach for various reasons. FTICR instrumentation has had limited availability, while intact proteins present challenges with sensitivity and throughput (Kelleher, 2004). However, as the techniques used are further refined, more and more laboratories are adopting it. A good example is the advent of SELDI-TOF MS for biomarker discovery using a chip-based platform (Al-Tarawneh and Bencharit, 2009; Bouamrani et al., 2006; Caputo et al., 2003; Cheng Wu, 2008; de Seny et al., 2008; Dommisch et al., 2009; Frankfort et al., 2008; Hong et al., 2005; Issaq et al., 2002; Kiehntopf et al., 2007; Lundquist et al., 2005; Mulli et al., 2008; Navaglia et al., 2009; Papale et al., 2008; Poon, 2007; Purohit et al., 2006; Schipper et al., 2007b; Schmid et al., 2005; Seibold et al., 2007).

2.1.3 Potential challenges of protein Identification

Whatever proteomic method used, the assumption is that the proteins or peptides that one is attempting to identify are already well characterized and present in the databases that one crosschecks with the data obtained from mass spectrometry or gel-based systems. Unfortunately, not all possible millions of peptides/proteins have been described or identified and especially in eukaryotic organisms not to mention in the human species. This limits the
number of proteins and peptides that can be positively identified. Another challenge is the enormous amount of data generated in the form of peaks that need to be characterized for identification (Kicman et al., 2007; Kumar and Mann, 2009; Mo and Karger, 2002). The advent of computer bioinformatics and algorithms has done a lot to improve data analysis and refine the process of protein identification using mass spectrometry (Dakna et al., 2009; Kelleher, 2004; Kumar and Mann, 2009; Yates et al., 2009).

Another problem in protein identification is the presence of unwanted noise from the most abundant proteins that may mask the less abundant ones or large peptides or proteins masking the smaller ones that are of interest. One way to mitigate this problem is to pre-fractionate the samples in order to reduce the presence of abundant species, increase the number of less abundant species, and enrich samples in elements of interest.

When fractionating samples in search for specific peptides or proteins identified in SELDI-TOF MS, it is advisable to use methods that mimic the sample processing and ProteinChip® array chemistry used in the SELDI-TOF MS experiments in order not to introduce new variables while interpreting the peaks discovered using SELDI-TOF MS.

The previous SELDI-TOF MS experiments had utilised both Ion exchange chromatography (CM10) and immobilized metal affinity chromatography (IMAC 30) ProteinChip® arrays.

2.1.3.1 Ion Exchange Chromatography (IEC)

This is the most frequently used chromatography technique for the separation of proteins, peptides, and other charged biomolecules. IEC achieves separation by taking advantage of the difference in net surface charge of these elements in a complex biological mixture.

Surface charge is determined by the presence of weak acidic and basic groups on a protein and as such is highly pH dependent. Typically, at pH values below a protein’s isoelectric point (pI), the protein acquires a net positive surface charge and will adsorb to cation exchange media. At pH values above the protein’s pI, the protein acquires a net negative surface charge and will adsorb
to anion exchange media. Variations in mobile phase pH are most often used to selectively bind a molecule of interest to an IEC support. The ideal pH value is one that creates a large net charge difference among the different sample components.

Desorption of bound proteins from IEC media is achieved by either altering the pH or increasing the salt concentration (introducing competing ions) of the mobile phase. The higher the surface charge of the protein, the higher the ionic strength that is needed to bring about desorption.

An example of an IEC media is the ProteinChip® CM (carboxymethyl) spin columns (Bio-Rad Laboratories, Inc.). These contain CM ceramic HyperD F sorbent, a cation exchange chromatography support that mimics the binding and elution characteristics of the ProteinChip® CM10 array.

2.1.3.2 Immobilized Metal Affinity Chromatography (IMAC)
This exploits the affinity of proteins for chelated metal ions. This affinity is due to coordination bonds formed between metal ions and certain exposed side chains of protein amino acids. IMAC is orthogonal to IEC separation and therefore combining the two techniques can greatly improve the range of proteins or peptides identified.

An example of a pre-fractionation platform based on IMAC is the ProteinChip® IMAC spin columns (Bio-Rad Laboratories, Inc.). Metal ions are firstly immobilized on the IMAC support, which uses polymer-bound iminodiacetic acid (IDA) to chelate the metal ions. The adsorbent then binds to nitrogen-containing functional groups (imidazole in histidine groups) on proteins, as well as phosphorylated proteins and peptides.

Elution can then be achieved by increasing concentrations of imidazole, by reducing the pH, or a combination of both. Typically, concentrations of 5–100 mM imidazole are used.

IMAC support mimics binding and elution properties of ProteinChip IMAC30 arrays.
2.2 Aims

The aim of this part of the study was to identify the proteins and peptides that were flagged up by the SELDI-TOF MS as potential biomarkers of periodontal disease in both saliva and GCF. To achieve this both GCF and saliva samples needed to be clarified by fractionation and tested to confirm that the desired peaks were present before protein identification using LCMS/MS.

Specific objectives

1. Saliva and GCF sample clarification and enrichment
2. Confirmation of desired biomarker peaks in the enriched samples
3. Identification of putative biomarkers flagged by SELDI-TOF MS in both saliva and GCF.
2.3 Methods And Materials

2.3.1 Patient cohort

Patients had been recruited from the new referrals clinic at The Dental Teaching Hospital, Bart’s and The London NHS Trust following assessment by two consultant periodontists. The cohort composed of 90 patients. Thirty patients with a clinical diagnosis of aggressive periodontitis with attachment loss of greater than 6 mm affecting a minimum of 6 teeth and under the age of 40 years old were recruited. Also thirty-one clinically diagnosed chronic periodontitis cases were included. Twenty-nine patients with gingivitis but no radiographic evidence of bone loss were recruited from the dental emergency clinic (Table 3-2). Exclusion criteria included history of periodontal treatment or antimicrobial therapy within the previous 6 months, or presence of any systemic condition or being on medication likely to affect the periodontal tissues.

Following written informed consent, patients received full clinical assessment and details of smoking status were recorded and supplemented by measurement of expired carbon monoxide using piCO Smokerlyzer® (Bedfont Scientific Ltd., Kent, UK). Patients were subsequently characterized as current smokers, previous smokers or never smokers. The protocols for this study were approved by the City and East London Health Authority Local Research Ethics Committee (05/Q0601/80) (Clarke, 2007; Mulli et al., 2008).

2.3.2 Samples

Saliva samples were collected using Salivettes® (Sarstedt Ltd) by asking the patient to chew gently on the cotton roll for 30 seconds. The samples were spun from the cotton roll by centrifugation at 3000rpm for 5 minutes. GCF samples were collected using Periopaper® (Pro Flow Inc., Connecticut, USA) inserted into the periodontal crevice at a depth of about 1mm for 30 seconds. The supragingival plaque was gently removed with cotton roll before sample acquisition. Both saliva and GCF samples were immediately stored at minus 80°C until time for analysis.
2.3.3 Saliva and GCF Clarification and Enrichment

The aim of these experiments was to enrich the samples prior to protein identification using LC-MS/MS to ensure that they contained the desired peaks (putative biomarkers) for identification. This was achieved through the pre-fractionation of the GCF and saliva samples. Sample fractions that contained high peaks of the diagnostically useful differential peaks identified by SELDI-TOF MS were then selected for further mass spectrometry analysis. In order to optimize the samples for protein identification, the samples had to be clarified and enriched as stated in preceding paragraphs. In general, this was accomplished using chromatographic columns. The enriched samples were then tested to see whether they had the desired peaks for identification using SELDI-TOF MS. The final step was protein identification using LC-MS/MS on the optimized samples. Even then, the protein identification process also had to be optimized and the experiments are described below.

Clarification of the samples was done by fractionation using chromatographic spin columns with similar chemistry to the ProteinChip® arrays used in the initial SELDI-TOF MS biomarker discovery phase. The newly acquired sample fractions were then checked using SELDI-TOF MS to confirm the presence of the desired peaks (Figure 2-1). Finally, the optimal samples were used for protein identification using LC-MS/MS.

To confirm the presence of the desired peaks in the samples using SELDI-TOF MS, several parameters including the ProteinChip® array matrix and SELDI-TOF MS settings had to firstly be adjusted empirically and then fine tuned to give the best readings (Table 2-2).
2.3.3.1 Fractionation

In order to optimize sample clarification and enrichment, a set of one saliva and one posterior site GCF sample from a gingivitis patient and another set from a chronic periodontitis patient were used assuming that all the markers with high expression in both conditions would be present in these samples. In order to increase the chance of finding all or most of the desired peaks corresponding to the putative biomarkers, it was necessary to pool patient samples for the final fractionation to cover the entire range of possible peaks. Therefore, once the process was optimized, pooled saliva and pooled GCF samples were obtained by pooling samples. Due to sample depletion in previous studies, the available samples were from 15 gingivitis, 23 CP and 16 AgP patients. All the GCF samples were selected from any of the posterior sites regardless of the pocket depth.

Chromatographic columns with similar chemistries to the SELDI-TOF ProteinChips® used in the biomarker discovery study were used to fractionate the samples. Each pooled GCF and saliva sample was passed through the metal-binding IMAC spin columns (Ciphergen) and eluted into four fractions based on the molarity of the elution buffer (imidazole at 100mM, 70mM, 40mM...
and 10mM) or through a weak cationic exchanger CM10 spin columns and eluted with sodium acetate buffer at pH 6, 7, 8 and 9 to make four new fractions besides the unbound fraction per original sample. Each original sample therefore ended up with ten new fractions (samples) as Figure 2-2 demonstrates.

![Diagram of fractionation process](image)

**Figure 2-2**: A schematic view of the fractionation process showing the final number of fractions collected per each starting sample.

**ProteinChip® CM Spin Column Fractionation**

For this column, sorbent equilibration was done using the same buffer as the sample binding buffer (i.e. 0.1M sodium acetate, pH 4.5) to ensure that most proteins bound to the column. The spin columns were tapped lightly to settle the sorbent to the bottom. The caps on the top and bottom were then removed and column set upright in a 1.5 ml microcentrifuge tube. The column was then centrifuged at 1,000 rpm for 30 seconds to remove the buffer then the bottom cap of the spin column was replaced.

200µl of binding buffer (0.1M sodium acetate, pH 4.5) was added to the column, then the top cap replaced on the column before vortexing the column to mix the sorbent with the buffer. Then the top and bottom caps of the spin column were removed; the column set upright in a new 1.5 ml microcentrifuge tube and centrifuged for 30 seconds to remove the buffer. This was repeated twice for a
total of three buffer washes. Then the bottom cap was replaced on the spin column.

The samples were added to the column and left to bind for 20 minutes in an orbital shaker. Then the column was centrifuged for 30 seconds and the unbound fraction collected in a 1.5 ml microcentrifuge tube. Elution buffer (EB), sodium acetate buffer at pH 6 was added to the tube and again centrifuged as before to collect the pH 6 fraction in a new 1.5 ml microcentrifuge tube. The same was done with elution buffer at pH 7, 8 and 9 to collect those fractions.

The new samples including the unbound fraction were then precipitated using cold acetone. 1.5 ml of 100% cold acetone was added to 500µl of each new CM10 fraction (sample) and left overnight in minus 20°C. The contents were then centrifuged for 10 minutes to get a white pellet. The excess acetone was then decanted carefully and the tubes left open for a few minutes to dry.

The samples were then stored in minus 80°C freezer until the time for profiling using the SELDI-TOF when they were re-suspended with 50µl of 50mM ammonium bicarbonate buffer (ambic buffer), pH 7.8 by vortexing.

**IMAC Spin Column Fractionation**

The buffers used in this process are indicated in Table 2-1:

<table>
<thead>
<tr>
<th>Charging Solutions</th>
<th>100mM copper sulphate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binding Buffer (BB) and</td>
<td>Cu^2+</td>
</tr>
<tr>
<td>Equilibration Buffers</td>
<td></td>
</tr>
<tr>
<td>Starting Buffer</td>
<td>Sodium phosphate</td>
</tr>
<tr>
<td>Elution Buffer</td>
<td>10mM, 40mM, 70mM and 100mM imidazole</td>
</tr>
</tbody>
</table>

Table 2-1: List of buffers used in IMAC fractionation of both saliva and GCF samples.

The column support was washed three times to eliminate ethanol and sodium chloride preservatives in the storage solutions. The spin columns were tapped
lightly to settle the preservative to the bottom. The top and bottom caps were removed from the column, then the column was set upright in a 1.5 ml microcentrifuge tube before spinning at 1,000 rpm for 30 seconds to remove the wash out. 500 µl of DI water was added to the column. The column was centrifuged to remove the water. This was repeated two more times and the bottom cap replaced on the column.

200µl of 100 mM copper sulfate was added to the column, and then the top cap replaced before incubation with mixing in an orbital shaker for 20 minutes at room temperature (RT). The top and bottom caps were then removed and the spin column set upright in a new 1.5 ml microcentrifuge tube before spinning for 30 seconds to remove the charging solution. This process of charging the column was repeated once more.

200 µl of DI water was added to the spin column and then centrifuged to remove the water. This was repeated two more times and the bottom cap replaced. Then 200µl of binding buffer (BB) was added to the column and the top cap replaced. Mixing was done in an orbital shaker for 20 minutes. The column was then centrifuged to remove the buffer and the bottom cap replaced. This was repeated two more times. Then the support was re-suspended with 90µl of BB.

For GCF, 150µl of pooled sample was diluted with 450µl of BB whereas saliva samples were used undiluted. The samples were added to the spin column, and the top cap replaced followed by mixing in an orbital shaker for 20 minutes.

Afterwards, the top and bottom caps were removed; the column set upright in a new 1.5 ml microcentrifuge tube and then centrifuged at 1000 rpm for 30 seconds and the flow through (unbound fraction) collected. The bottom cap was then replaced. 200µl of BB was added to column to wash unbound fraction and added to the first flow through.

100µl of 10mM EB was added to the spin column, the top cap replaced and column incubated with mixing for 5 minutes at RT. The top and bottom caps were then removed; the column set upright in a new 1.5 ml microcentrifuge tube and then centrifuged for 30 seconds and the fraction collected. The bottom cap
was then replaced and the elution step repeated once. The two fractions were then pooled and the tube labeled with the buffer used i.e. 10mM EB. This step was repeated with the other EB strengths in an increasing order (40mM, 70mM & 100mM) and the duplicate fractions pooled and labeled accordingly.

The samples were then concentrated by acetone precipitation and reconstituted with 50µl of 50mM ambic buffer, pH 7.8 as described earlier for the CM10 assay. All the samples were then stored in minus 80°C freezer until time for profiling using the SELDI-TOF MS.

### 2.3.3.2 Confirmation Of Peaks Using SELDI-TOF MS

The samples used can be broadly categorized in to two; firstly saliva and GCF samples from two different patients for the purpose of optimization of experiment settings and secondly, saliva and GCF samples pooled from all the remaining patients in the original SELDI-TOF biomarker discovery study.

In order to optimize SELDI-TOF MS spectra acquisition, it was necessary to run multiple experiments altering the different parameters indicated in Table 2-2. This was achieved using the fractionated saliva and GCF samples from the two patients mentioned previously. Once the optimal settings were achieved using these two samples, the pooled saliva and GCF samples were then analyzed using the optimized conditions and settings Table 2-2. 2µL of sample was applied on each of the 8 array spots on the Goldchip® array and left to dry for 5 minutes. 1µL of matrix (SPA or CHCA depending on stage of optimization) was applied on each array spot and left to dry for 5 minutes. The array chip was then analyzed using the SELDI-TOF MS (model PBS IIc, Ciphergen) using the parameters in Table 2-2.

The Goldchip® array spots were cleaned between each sample application by rubbing gently with a piece of cotton wool containing methanol and then left to dry before further sample application.

The peaks for specific putative biomarkers flagged by the initial SELDI-TOF MS study were then manually searched for in all the spectra from the pooled samples and the spectra with the clearest peaks of interest chosen for protein identification using LC-MS/MS.
2.3.4 Protein Identification using LC-MS/MS

Once the samples with the appropriate peaks were chosen, it was then necessary to optimize the protein identification (ID) procedure. Protein ID was carried out in both GCF and saliva using 1-D gel electrophoresis coupled with Liquid Chromatography MS/MS (LC-MS/MS) in collaboration with the Institute of Psychiatry Proteomics Unit, Kings College London (Figure 2-3).

Table 2-2: SELDI-TOF MS parameters for running the GoldChip® arrays in the different optimization phases.

<table>
<thead>
<tr>
<th>Optimization parameters:</th>
<th>Initial settings</th>
<th>Optimized settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firing</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>Laser Intensity</td>
<td>195</td>
<td>171</td>
</tr>
<tr>
<td>GoldChip® array Matrix</td>
<td>SPA</td>
<td>CHCA</td>
</tr>
</tbody>
</table>

Of particular interest were neutrophil-derived proteins, with the originator cells being the main cellular effectors of the host immune response during periodontal disease. The expected size of such proteins is between 2-16KDa. The extent of protein detection would determine whether further fractionation
was desirable and the nature of proteins would indicate targets of particular interest for future quantitation.

**Samples**

Two saliva and two GCF samples were chosen for analysis: CM10-purified pH6 fraction (CM10) and IMAC purified 100mM imidazole fraction (IMAC). Samples had been concentrated using CM10 weak cation exchange columns using acidity to elute and IMAC columns with incremental imidazole concentration as elution buffer as described previously. These were then analyzed in stages; firstly to determine the extent and quality of protein detection by LC-MS/MS in all fractions of both saliva and GCF followed by analysis of selected molecular weight bands of the fractions that had shown best detection of proteins (CM10 fractions).

### 2.3.4.1 Determining The Extent And Quality Of Protein Detection By LC-MS/MS

**1D gel SDS-PAGE**

The 2 fractions of both saliva and GCF were used. Approximately 15µl of each sample was added to an equal volume of Laemmli buffer (Laemmli, 1970) and boiled for 10mins. The samples were then loaded onto two 4-20% (w/v) SDS-PAGE gels for stacking alongside Novex® SeeBlue® Plus2 pre-stained standards (Invitrogen, Paisley, UK) (Figure 2-7 and Figure 2-8). The gels were fixed, Coomassie stained, destained and visualized with ImageQuant (GE Healthcare, Buckinghamshire, UK). Bands corresponding to the stacked proteins were then excised and stored in water until required for digestion.

**In-gel digestion**

Gel bands were cut into pieces approximately 2mm² in area. In-gel reduction, alkylation and digestion with trypsin was performed prior to subsequent analysis by mass spectrometry (Shevchenko et al., 1996). Briefly, cysteine residues were reduced with dithiothreitol and alkylated with iodoacetamide to form stable carbamidomethyl derivatives. Trypsin digestion was carried out overnight at 37°C in ammonium carbonate buffer. Peptides were extracted from the gel pieces by two washes with 50mM ammonium bicarbonate and acetonitrile. The extract was pooled with the initial supernatant and lyophilized. Each sample
was then re-suspended in 23µL of 50mM ammonium bicarbonate before LC-MS/MS analysis of 23µl sample per run.

**LC-MS/MS**

Chromatographic separations were performed using an Ultimate LC system (Dionex, UK). Peptides were resolved by reversed phase chromatography on a 75 µm C18 PepMap column. A 60-minute gradient of acetonitrile in 0.05% formic acid was delivered to elute the peptides at a flow rate of 200nL/min. Peptides were ionized by electrospray ionization using a Z-spray source fitted to a QTof-micro (Waters Corp., Milford, MA, USA). The instrument was set to run in automated switching mode, selecting the two most intense precursor ions for sequencing by collision-induced fragmentation. Precursor ions were surveyed across a mass-to-charge ratio (m/z) range of 400-1800m/z. The MS/MS analyses were conducted across a range of 100-2000m/z using collision energy profiles that were chosen based on the m/z and the charge state of the peptide. Each precursor mass was excluded from further selection from the mass list for 90 seconds following its fragmentation.

**Data analysis**

Tandem mass spectra were extracted by Masslynx version 4.0 and Protein Lynx Global Server 2.2.5. Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.2.03). Saliva samples were then analyzed first and afterwards GCF was also analyzed with minimal modification of the parameters as described below. For the saliva samples, Mascot was set up to search the sprot_56_8_090210.fasta_fix database (410518 entries) while for GCF, Mascot was modified to search human entries of the NCBI non-redundant database (70320 entries, July 2007). In all cases, Mascot was set assuming the digestion enzyme trypsin and up to 3 missed cleavages. For both saliva and GCF, Mascot was searched with a fragment ion mass tolerance of 0.60 Da and a parent ion tolerance of 1.2 Da. Oxidation of methionine, carbamidomethyl (iodoacetamide derivative) cysteine, deamidation of glutamine and asparagine, N-terminal acetylation and phosphorylation addition on serine, threonine and lysine were specified in Mascot as variable modifications. For the
GCF, raw data were recalibrated against internal tryptic peptides where necessary to produce a mass accuracy of less than 50ppm for most peptides.

Scaffold (version Scaffold_2_02_03, Proteome Software Inc., Portland, Oregon, USA) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm (Keller et al., 2002). Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al., 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

### 2.3.4.2 Protein Detection In Specific Molecular Weight Bands Of CM10, pH6 Fractions.

After the initial experiments to determine the extent and quality of protein detection using LC-MS/MS in both saliva and GCF, it was determined that the CM10, pH6 saliva fractions yielded the most IDs. This saliva sample was then prepared and analyzed in very similar ways as the previous samples except for four main differences:

Firstly, one CM10, pH6 saliva sample was loaded onto a 10% (w/v) instead of 4-20% (w/v) polyacrylamide NuPAGE® Novex® Bis-Tris Mini Gel SDS-PAGE gel for stacking alongside Novex® SeeBlue® Plus2 pre-stained standards (Invitrogen, Paisley, UK). This was to facilitate better resolution of different bands of proteins in the gel as is clear by comparing Figure 2-12 and Figure 2-7. The electrophoresis was stopped when the leading band had migrated three quarters of the gel height.

Secondly, another CM10, pH6 saliva sample was lyophilised to completion and 40µl of Laemmli buffer and 40µl of 50mM TEAB buffer added and boiled for 10mins. The sample was then loaded across 2 gel lanes onto a 4-12% (w/v) SDS-PAGE gel alongside Novex® SeeBlue® Plus2 pre-stained standards (Invitrogen, Paisley, UK) before electrophoresis. Both gels were then fixed,
Coomassie stained, destained and visualised with ImageQuant (GE Healthcare, Buckinghamshire, UK).

Thirdly, the regions corresponding to the targeted molecular weight range of proteins were then excised (highlighted in red boxes in Figure 2-12 and Figure 2-16) and then stored in water until required for proteolytic digestion. For the first gel, the region was that corresponding to 3-6kDa while the second gel had two regions excised corresponding to 3-6kDa and 10-19kDa. This latter band was included in order to include regions that may contain the 11kDa cathelicidin peptide.

Lastly, after LC-MS/MS the Mascot was set to search the mammalian section of the NCBI non-redundant database (154181 entries, July 2007) for the first gel digest and entries of the uniprot_sprot_100401 database (selected for humans, 20280 entries) for the second set of gel digests.
2.4 Results

2.4.1 Saliva and GCF Clarification and Enrichment

The fractionation process yielded 10 samples (fractions) per initial sample Figure 2-2. The initial samples from the 2 patients had been analyzed in the SELDI-TOF MS using SPA as matrix and the mass spectrometry settings as shown under ‘initial settings’ in Table 2-2. The MS spectra produced for this group of experiments were characterized by spectra noise and lacked distinctive peaks (Figure 2-4).

![Figure 2-4: A representative GoldChip array spectra for the initial exploratory settings based on SPA matrix showing poor quality for most of the spectra. The samples used here had been prepared by fractionation using IMAC columns and eluted using 100mM imidazole elution buffer.](image)

After changing the matrix to CHCA and the mass spectrometer settings, the spectra reading greatly improved for the samples from the same 2 patients (Figure 2-5). These settings were then used for the pooled samples yielding overall good quality spectra (Figure 2-6).
Further visual analysis of the spectra revealed optimal spectra peaks in IMAC 100mM fractions and CM10, pH 6 fractions for both GCF and saliva as can be seen in the representative peaks in Figure 2-6. Notice the peaks corresponding to HNP1, 2 and 3. The fractions at pH 6 and IMAC 100 were therefore chosen for further protein identification using LC-MS/MS.
Figure 2-6: Representative spectra for CM10 fractions at different pH in both GCF and saliva performed on the same GoldChip® array. Notice the high peak intensities in the pH 6 fractions for both GCF and saliva compared to other fractions of the same media. In these spectra, the peaks for HNP 2, 1 and 3 respectively can be seen (red arrows between the regions m/z 3300 to 3500).
2.4.2 Protein Identification Using LC-MS/MS

2.4.2.1 The extent and quality of protein detection by LC-MS/MS.

As previously described, initial experiments were done to determine the extent and quality of protein detection by LC-MS/MS in both saliva and GCF CM10 pH6 and IMAC 100mM imidazole fractions.

These fractions were run on an SDS-PAGE gel as previously described to desalt and to stack the proteins. The protein mixtures migrated as unresolved bands through the 4% w/v gel and electrophoresis was terminated when the bands reached the interface of the 20% w/v gel. Monitoring of the pre-stained molecular weight markers ensured that the proteins did not enter into the 20% gel (Figure 2-7 and Figure 2-8).

Figure 2-7: SDS-PAGE preparation of salivary protein samples (CM10-purified pH6 fraction and IMAC-purified 100mM imidazole fraction). The interface between the 4% and 20%w/v gel sections is marked. MW corresponds to unresolved molecular weight markers. Volumes of 15µl of sample and 5µl of size markers were loaded. The region excised is indicated by red boxes.
CHAPTER 2 – PROTEIN IDENTIFICATION

Mass spectral data were searched against the non-redundant SwissProt database according to the criteria described in ‘Methods’ and restricting to human sequences. The results of the analysis and validation by Scaffold are given in Table 2-3 and Table 2-5.

All peptide assignments are provided in supplemental data (Table 8-6 in Appendices). A search with an increased number of variable modifications (oxidation of methionine, carbamidomethyl cysteine, deamidation of glutamine and asparagine, N-terminal acetylation and phosphorylation addition on serine, threonine and tyrosine) did not increase the number of validated protein assignments (data not shown).

In order to assess false discovery rate and increase the confidence of protein assignments to the mass spectral results, the data for each sample was searched against a SwissProt decoy database using Mascot. This contains randomised sequences and therefore any peptides matching would be highlighted as a false positive. In saliva samples, 15 proteins were identified. No false positives were assigned within the dataset above Mascot’s identity threshold indicating a zero false discovery rate for peptides and thus high confidence of assignments.

Four (4) proteins were validated between the ranges of 10-16KDa. These were Cystatin-SN, Ig kappa chain C region, Ig lambda chain C regions and Cystatin-B. Alpha-amylase and serum albumin dominated the spectra in both runs as
seen by the number and percentage of total spectra for these proteins Table 2-6. Further proteins identified by Scaffold below the specified criteria can be found in Table 2-3. A greater number of proteins were identified from the CM10 sample and correspondingly the amount of protein visible from the SDS-PAGE gel was higher. Protein coverage for the same proteins was also lower in the IMAC sample indicating better purification of salivary proteins by the CM10 weak cation exchange method.

In GCF, no assigned protein was validated below 16KDa. Eight (8) proteins in total were validated according to the above criteria in the CM10 sample and none in the IMAC fraction. Four proteins were identified with a probability of 77% including neutrophil granule peptide HP1 whose assigned sequence is shared with four other proteins, Figure 2-11. The single peptide assigned to HP1 was however validated with 95% probability. An NCBI search for gi228797 provides the HP1 full sequence (used for Scaffold validation), which can be BLAST-searched using Uniprot giving a set of matches from the Uniprot/Swissprot database.
Serum albumin and keratin isoforms dominated the spectra in the CM10 fraction as seen by the number and percentage of total spectra for these proteins (Table 8-6 in Appendices). A lower number of proteins were identified at high confidence (8) in the CM10 sample compared to the equivalent fraction of saliva (13). Repeating the Mascot search to include variable asparagine and glutamine deamidation and N-terminal acetylation did not increase the number of validated protein assignments in Scaffold. However no false positives were assigned within the dataset above Mascot’s identity threshold indicating a zero false discovery rate for peptides and thus high confidence of assignments.

Figure 2-10: Scaffold list of identified proteins and peptides in both CM10 pH6 and IMAC 100mM imidazole fractions of saliva. Green highlighting indicates protein identification with high probability. A total of 15 proteins were identified in both fractions.
2.4.2.2 Identified salivary proteins in gel band regions of 3-6kDa and 10-17kDa

From the results above, it was obvious that the saliva CM10, pH6 fractions gave the best chance for protein ID and henceforth were used for further protein ID work. This fraction of saliva sample was run on different gels and different gel-bands corresponding to specific molecular weights trypsin-digested for analysis with LC-MS/MS as previously described in the methods section above.

The first saliva sample loaded onto a 10% (w/v) gel (Figure 2-12) as described in the methods section had only one protein validated between the range 3KDa to 6KDa (highlighted in yellow in Table 2-4). This was assigned as neutrophil granule peptide HP1. The same two peptides for this protein were also assigned to Chain A, Defensin HNP-3 (gi229858); defensin, alpha 3 preproprotein (gi4885179); neutrophil peptide 3 precursor (gi7702778) and
defensin, alpha 1 preproprotein (gi4758146). The coverage of this target protein is provided mapped to the sequence in Figure 2-13 and the spectra view for different sequences in Figure 2-14 and Figure 2-15.

Figure 2-12: SDS-PAGE preparation of salivary protein samples (CM10 pH6). MW corresponds to unresolved molecular weight markers. Volumes of 30µl of 1:2 sample and 5µl of size markers were loaded. The region excised is highlighted in red with the bromophenol blue dye front being visible at the bottom. The image shown is at high contrast.

The two peptides assigned to neutrophil granule peptide HP1 were found in the Human Salivary Proteome Central Repository SM/SL protein list (http://hspp.dent.ucla.edu/cgi-bin/spkbcgi-bin/main.cgi) as IPI00005721 (International Protein Index), which is indexed as Q6EZF6, defensin, alpha 1 in Uniprot. In total, 6 proteins were identified with high probability including protein S100A8 (highlighted in yellow in Table 2-4).
Figure 2.13: Mass spectrometric coverage of neutrophil granule peptide HP1 and matching proteins in saliva. Other proteins assigned to the same two peptides are stated in the text.

Figure 2.14: Scatter graph of the LC-MS/MS peaks for HNP 1 and Probably HNP 2 and 3 in saliva.

Figure 2.15: Scatter graph of the LC-MS/MS peaks for HNP 1 and Probably HNP 2 and 3 in saliva.

One false positive was assigned within this dataset above Mascot’s identity threshold but this protein was not validated by Scaffold resulting in a 4.17% false discovery rate for peptides using Mascot alone and therefore high confidence of assignments.

The second gel of CM10 pH 6 saliva sample had been lyophilised to completion and 40 µl of Laemmli buffer (Laemmli, 1970) and 40 µl of 50 mM TEAB buffer added before boiling for 10 mins and loading onto a 4-12% (w/v) SDS-PAGE
The single sample was run on an SDS-PAGE gel to desalt and to resolve proteins. The protein mixture migrated as shown in Figure 2-16.

Electrophoresis was terminated when the 3KDa marker migrated approximately three quarters the length of the gel. The portions corresponding to molecular weight 3-6KDa and 10-19KDa had been excised (Figure 2-16) trypsin digested and analysed by LC-MS/MS as previously described in methods. Mascot was set up to search entries of the uniprot_sprot_100401 database (selected for humans, 20280 entries).

Proteins were found at high confidence in both gel-bands with 31 proteins identified from the 10-19KDa region and 11 proteins from the 3-6KDa regions, which is in accordance with the visual difference in protein levels from the gel (Figure 2-16)

![SDS-PAGE separation of saliva sample. Volumes of 40µl of 1:2 sample (X) and 5µl of size markers (M) were loaded. The region excised for in-gel digestion is indicated by the red box.](image)

All proteins assigned with high confidence in the 3-6KDa region apart from 4 were also identified in the 10-19KDa gel region. None of these 4 proteins have an intact mass in the 3-6KDa region and so represent processed or degraded forms of the proteins. For example the Defensin (P59665) peptides are in the
C-terminus of the sequence and correspond to known processed forms of the protein.

The Scaffold assignment algorithms did not assign Cathelicidin. However, a search using Mascot conducted as stated in Methods resulted in one assignment as shown Figure 2-9. The query (i.e. spectrum) is also matched to a higher scoring entry in the list of proteins from the search but has higher mass accuracy for cathelicidin. As can be seen from the MS/MS fragmentation, the spectrum is a weak assignment with only 3 y-ions matched and no signal-to-noise ratio. However, the major peaks are assigned.

False discovery rate (FDR) testing using a Mascot decoy database was performed on all samples as before (See methods). Only the 3-6KDa sample generated decoy database matches above Mascot’s significance threshold for peptide identity. Two peptides were matched to the decoy database compared to 46 peptides matched to the Uniprot database giving a FDR of 4.35% and therefore high confidence of identification.

In all the gels that were run, a total of 115 proteins were identified both in saliva and GCF. 67% of these were identified with ≥99% probability.
Figure 2.17: Four scatter graphs for calcium-binding protein S100A8 specific peptide sequences from the LC-MS/MS peaks of CM10 fraction (pH 6) of saliva.
<table>
<thead>
<tr>
<th>Sample name</th>
<th>Protein name</th>
<th>Protein accession numbers</th>
<th>Protein molecular weight (Da)</th>
<th>Protein identification probability</th>
<th>Number of unique peptides</th>
<th>Number of unique spectra</th>
<th>Number of total spectra</th>
<th>Percentage of total spectra</th>
<th>Percentage sequence coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saliva CM10 pH6</td>
<td>Cystatin-B</td>
<td>CYTB_HUMAN</td>
<td>11,121.30</td>
<td>100.00%</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>1.15%</td>
<td>58.20%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Ig alpha-1 chain C region</td>
<td>IGHA1_HUMAN</td>
<td>37,635.80</td>
<td>100.00%</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>0.92%</td>
<td>17.30%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Alpha-amylose 1</td>
<td>AMY1_HUMAN</td>
<td>57,750.40</td>
<td>100.00%</td>
<td>27</td>
<td>40</td>
<td>77</td>
<td>17.70%</td>
<td>64.80%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Zinc-alpha-2-glycoprotein</td>
<td>ZA2G_HUMAN</td>
<td>33,854.10</td>
<td>99.80%</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0.69%</td>
<td>10.50%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Serum albumin</td>
<td>ALBU_HUMAN</td>
<td>69,348.90</td>
<td>100.00%</td>
<td>39</td>
<td>51</td>
<td>61</td>
<td>14.00%</td>
<td>72.40%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Cystatin-SN</td>
<td>CYTN_HUMAN</td>
<td>16,343.80</td>
<td>100.00%</td>
<td>6</td>
<td>7</td>
<td>9</td>
<td>2.06%</td>
<td>58.90%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Ig kappa chain C region</td>
<td>IGKC_HUMAN</td>
<td>11,590.50</td>
<td>100.00%</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>1.61%</td>
<td>72.60%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Alpha-1-antitrypsin</td>
<td>A1AT_HUMAN</td>
<td>46,719.90</td>
<td>100.00%</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.46%</td>
<td>5.74%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Prolactin-inducible protein</td>
<td>PIP_HUMAN</td>
<td>16,555.10</td>
<td>100.00%</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.46%</td>
<td>23.30%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Ig lambda chain C regions</td>
<td>IAC_HUMAN</td>
<td>11,218.10</td>
<td>100.00%</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>1.15%</td>
<td>65.70%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Leukocyte elastase inhibitor</td>
<td>ILEU_HUMAN</td>
<td>42,725.80</td>
<td>99.90%</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.46%</td>
<td>7.92%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Ig gamma-2 chain C region</td>
<td>IGHE2_HUMAN</td>
<td>35,881.20</td>
<td>100.00%</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0.69%</td>
<td>10.10%</td>
</tr>
<tr>
<td>Saliva CM10 pH6</td>
<td>Polymeric immunoglobulin receptor</td>
<td>PIGR_HUMAN</td>
<td>83,265.40</td>
<td>100.00%</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>1.38%</td>
<td>16.00%</td>
</tr>
<tr>
<td>Saliva IMAC 100mM</td>
<td>Keratin, type I cytoskeletal 10</td>
<td>K1C10_HUMAN</td>
<td>59,494.30</td>
<td>100.00%</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1.76%</td>
<td>6.41%</td>
</tr>
<tr>
<td>Saliva IMAC 100mM</td>
<td>Keratin, type II cytoskeletal 10</td>
<td>K2C1_HUMAN</td>
<td>66,001.20</td>
<td>100.00%</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>2.94%</td>
<td>10.10%</td>
</tr>
</tbody>
</table>

Table 2-3: A summary list of proteins and peptides identified with high confidence. In both CM10 pH 6 and IMAC 100mM imidazole fractions of saliva. A total of 16 proteins were identified in both fractions. Notice the overlap in the two fractions. All the proteins identified in the IMAC fraction were also present in the CM10 fraction except for the two highlighted in blue.
<table>
<thead>
<tr>
<th>Biological sample name</th>
<th>Protein name</th>
<th>Protein accession numbers</th>
<th>Protein molecular weight (Da)</th>
<th>Protein identification probability</th>
<th>Number of unique peptides</th>
<th>Number of unique spectra</th>
<th>Number of total spectra</th>
<th>Percentage of total spectra</th>
<th>Percentage sequence coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM10 pH6 3-6kDa humans</td>
<td>Keratin 1</td>
<td>gi</td>
<td>11935049</td>
<td>66,050.30</td>
<td>100.00%</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>3.79%</td>
</tr>
<tr>
<td>CM10 pH6 3-6kDa humans</td>
<td>Cystatin SA-III=potential precursor of acquired enamel pellicle</td>
<td>gi[235948,gi</td>
<td>352334,gi</td>
<td>4503109</td>
<td>14,171.70</td>
<td>99.80%</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>CM10 pH6 3-6kDa humans</td>
<td>Histone H4</td>
<td>gi</td>
<td>223582,gi</td>
<td>32097,gi</td>
<td>4504301</td>
<td>11,219.50</td>
<td>99.80%</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>CM10 pH6 3-6kDa humans</td>
<td>Neutrophil granule peptide HP1</td>
<td>gi</td>
<td>228797,gi</td>
<td>229858,gi</td>
<td>4758146,gi</td>
<td>4885179,gi</td>
<td>5702778</td>
<td>3430.10</td>
<td>99.80%</td>
</tr>
<tr>
<td>CM10 pH6 3-6kDa humans</td>
<td>Polyubiquitin</td>
<td>gi</td>
<td>11024714,gi</td>
<td>1304128,gi</td>
<td>13569612,gi</td>
<td>26,856.20</td>
<td>100.00%</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>CM10 pH6 3-6kDa humans</td>
<td>S100 calcium-binding protein A8</td>
<td>gi</td>
<td>21614544,gi</td>
<td>225541,gi</td>
<td>29888</td>
<td>10,934.50</td>
<td>100.00%</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2-4: Final proteomic analysis of SDS-PAGE gel region 3-6kDa for saliva CM10 pH6 sample.
<table>
<thead>
<tr>
<th>Biological sample name</th>
<th>Protein name</th>
<th>Protein accession numbers</th>
<th>Protein molecular weight (Da)</th>
<th>Protein identification probability</th>
<th>Number of unique peptides</th>
<th>Number of unique spectra</th>
<th>Number of total spectra</th>
<th>Percentage of total spectra</th>
<th>Percentage sequence coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM10</td>
<td>Chain A, Crystal Structure Of A Serpin:protease Complex</td>
<td>gi</td>
<td>11514321,gi</td>
<td>13787109,gi</td>
<td>15080499,gi</td>
<td>15990507,gi</td>
<td>177827,gi</td>
<td>177836,gi</td>
<td>1946269,gi</td>
</tr>
<tr>
<td>CM10</td>
<td>Keratin 1</td>
<td>gi</td>
<td>11935049</td>
<td>66,050.30</td>
<td>100.00%</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3.60%</td>
</tr>
<tr>
<td>CM10</td>
<td>Keratin, type I cytoskeletal 9 (Cytokeratin-9) (CK-9) (Keratin-9) (K9),</td>
<td>gi</td>
<td>81175178</td>
<td>62,113.00</td>
<td>77.50%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.90%</td>
</tr>
<tr>
<td>CM10</td>
<td>Keratin, type II cytoskeletal 2 epidermal (Cytokeratin-2e) (K2e) (CK2e)</td>
<td>gi</td>
<td>181402,gi</td>
<td>AAC83410.</td>
<td>gi</td>
<td>435476,gi</td>
<td>emb</td>
<td>CA82315.1</td>
<td>cytokeratin 9</td>
</tr>
<tr>
<td>CM10</td>
<td>Neutrophil granule peptide HP1</td>
<td>gi</td>
<td>228797,gi</td>
<td>229858,gi</td>
<td>4758</td>
<td>4758</td>
<td>146,gi</td>
<td>4885179,gi</td>
<td>7702778</td>
</tr>
<tr>
<td>CM10</td>
<td>Keratin, type II cytoskeletal 2 epidermal (Cytokeratin-2e) (K2e) (CK2e)</td>
<td>gi</td>
<td>547754</td>
<td>65,848.40</td>
<td>100.00%</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2.70%</td>
</tr>
<tr>
<td>CM10</td>
<td>Amylase, pancreatic, alpha-2B precursor</td>
<td>gi</td>
<td>10280622,gi</td>
<td>1421331,gi</td>
<td>15988375,gi</td>
<td>15988376,gi</td>
<td>163311,gi</td>
<td>178585,gi</td>
<td>18655992,gi</td>
</tr>
<tr>
<td>Protein Description</td>
<td>Accessions</td>
<td>Intensity</td>
<td>% Identity</td>
<td>Peptide Count</td>
<td>% Coverage</td>
<td>% Confidence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------</td>
<td>------------</td>
<td>---------------</td>
<td>------------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alloalbumin Venezia</td>
<td>gi</td>
<td>178345,gi</td>
<td>28592,gi</td>
<td>32124 56,gi</td>
<td>4389275,gi</td>
<td>4502027,gi</td>
<td>6013427</td>
<td>66,017.70</td>
<td>100.00%</td>
</tr>
<tr>
<td>Predicted: keratin 6A</td>
<td>gi</td>
<td>114644568,gi</td>
<td>1346345,gi</td>
<td>1346346,gi</td>
<td>1346347,gi</td>
<td>3406 9,gi</td>
<td>386849,gi</td>
<td>5031839</td>
<td>42,451.20</td>
</tr>
<tr>
<td>Chain A, Trypsin (E.C.3.4.21.4) Complexed With The Inhibitor Disopropyl-Fluorophosphofluoridate (DFP)</td>
<td>gi</td>
<td>1064991</td>
<td>pdb</td>
<td>1TRN</td>
<td>B Chain B, Trypsin (E.C.3.4.21.4) Complexed With The Inhibitor Disopropyl-Fluorophosphofluoridate (DFP)</td>
<td>gi</td>
<td>1064990,gi</td>
<td>11120626,gi</td>
<td>1120628,gi</td>
</tr>
<tr>
<td>Immunoglobulin heavy chain constant region</td>
<td>gi</td>
<td>10799664,gi</td>
<td>12054074,gi</td>
<td>12054078,gi</td>
<td>14030849,gi</td>
<td>184759,gi</td>
<td>25987831,gi</td>
<td>9857759</td>
<td>35,941.30</td>
</tr>
<tr>
<td>Serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 1</td>
<td>gi</td>
<td>13489087</td>
<td>42,725.80</td>
<td>100.00%</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3.60%</td>
<td>11.10%</td>
</tr>
<tr>
<td>Tumor factor receptor (TFNR)</td>
<td>gi</td>
<td>12232589,gi</td>
<td>9886705</td>
<td>245,708.10</td>
<td>53.60%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.12%</td>
</tr>
<tr>
<td>C2orf18 isoform B protein</td>
<td>gi</td>
<td>17046385,gi</td>
<td>55953063,gi</td>
<td>8393013</td>
<td>50,399.40</td>
<td>53.50%</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Unnamed protein product</td>
<td>gi</td>
<td>10434496</td>
<td>40,794.40</td>
<td>52.60%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.12%</td>
<td>3.01%</td>
</tr>
<tr>
<td>IMAC 100mM GCF</td>
<td>Calcyphosphine 2</td>
<td>gi</td>
<td>14211947</td>
<td>44,033.30</td>
<td>50.60%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.12%</td>
</tr>
<tr>
<td>IMAC 100mM GCF</td>
<td>KIAA1607 protein</td>
<td>gi</td>
<td>10047289,gi</td>
<td>18676518</td>
<td>210,598.50</td>
<td>52.80%</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2-5: Proteomic analysis of SDS-PAGE gel bands in GCF Sample CM10 pH6 & IMAC 100mM Imidazole fractions. Gene index (protein accession) numbers are provided for similar proteins containing the same assigned peptides. Identities of similar proteins can be found at http://www.ncbi.nlm.nih.gov/ or BLAST searched at http://www.uniprot.org/. Neutrophil granule peptide HP1 is highlighted in the CM10 fraction in yellow colour. All proteins with a Scaffold probability of greater than 0% are included rather than only those with a probability of 99% or more.
Figure 2-18: Scaffold 3 display of aligned samples. Percentage values are scores indicating confidence of protein identification. Identifications are from Uniprot database human entries. Results are shown using minimal stringency settings for peptide and protein identification in Scaffold.
<table>
<thead>
<tr>
<th>Biological sample category</th>
<th>Protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Protein accession numbers</td>
</tr>
<tr>
<td>3-6KDa THIO_HUMAN Thioredoxin OS=Homo sapiens GN=TXN PE=1 SV=3</td>
<td>P10599</td>
</tr>
<tr>
<td>3-6KDa H2B1K_HUMAN Histone H2B type 1-K OS=Homo sapiens GN=HIST1H2BK PE=1 SV=3</td>
<td>O60814,P57053, P58876,P62807, Q5QN6,W,Q9307 9,Q99877,Q99879</td>
</tr>
<tr>
<td>3-6KDa PIP_HUMAN Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1 SV=1</td>
<td>P12273</td>
</tr>
<tr>
<td>3-6KDa TRFE_HUMAN Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=2</td>
<td>P02787</td>
</tr>
<tr>
<td>3-6KDa A1AT_HUMAN Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3</td>
<td>P01009</td>
</tr>
<tr>
<td>3-6KDa SPLC2_HUMAN Short palate, lung and nasal epithelium carcinoma-associated protein 2 OS=Homo sapiens GN=SPLUNC2 PE=1 SV=2</td>
<td>Q96DR5</td>
</tr>
<tr>
<td>3-6KDa RYR2_HUMAN Ryanodine receptor 2 OS=Homo sapiens GN=RYR2 PE=1 SV=2</td>
<td>Q92736</td>
</tr>
<tr>
<td>3-6KDa AMY1A_HUMAN Alpha-amylase 1 OS=Homo sapiens GN=AMY1A PE=1 SV=2</td>
<td>P04745</td>
</tr>
<tr>
<td>3-6KDa RETN_HUMAN Resistin OS=Homo sapiens GN=RET1 PE=1 SV=2</td>
<td>Q9H89</td>
</tr>
<tr>
<td>3-6KDa CYTN_HUMAN Cystatin-SN OS=Homo sapiens GN=CST1 PE=1 SV=2</td>
<td>P01037</td>
</tr>
<tr>
<td>3-6KDa CYTS_HUMAN Cystatin-N OS=Homo sapiens GN=CST4 PE=1 SV=3</td>
<td>P01036</td>
</tr>
<tr>
<td>3-6KDa ILEU_HUMAN Leukocyte elastase inhibitor OS=Homo sapiens GN=SERPINB1 PE=1 SV=1</td>
<td>P30740</td>
</tr>
<tr>
<td>3-6KDa ACTB_HUMAN Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1</td>
<td>P60709,P63261</td>
</tr>
<tr>
<td>3-6KDa ENOA_HUMAN Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2</td>
<td>P06733</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CYTD_HUMAN</td>
<td>Cystatin-D OS=Homo sapiens GN=CST5 PE=1 SV=1</td>
</tr>
<tr>
<td>UBIQ_HUMAN</td>
<td>Ubiquitin OS=Homo sapiens GN=RPS2A PE=1 SV=1</td>
</tr>
<tr>
<td>QSOX1_HUMAN</td>
<td>Sulfhydryl oxidase 1 OS=Homo sapiens GN=QSOX1 PE=1 SV=3</td>
</tr>
<tr>
<td>LRQ1_HUMAN</td>
<td>Leucine-rich repeat and IQ domain-containing protein 1 OS=Homo sapiens GN=LRRIQ1 PE=2 SV=2</td>
</tr>
<tr>
<td>H31_HUMAN</td>
<td>Histone H3.1 OS=Homo sapiens GN=HIST1H3A PE=1 SV=2</td>
</tr>
<tr>
<td>FIBA_HUMAN</td>
<td>Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2</td>
</tr>
<tr>
<td>S100P_HUMAN</td>
<td>Protein S100 OS=Homo sapiens GN=S100P PE=1 SV=2</td>
</tr>
<tr>
<td>H4_HUMAN</td>
<td>Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2</td>
</tr>
<tr>
<td>NC2A_HUMAN</td>
<td>Drl1-associated corepressor OS=Homo sapiens GN=DRAP1 PE=1 SV=3</td>
</tr>
<tr>
<td>SPB3_HUMAN</td>
<td>Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2</td>
</tr>
<tr>
<td>LDHA_HUMAN</td>
<td>L-lactate dehydrogenase A chain OS=Homo sapiens GN=LDHA PE=1 SV=2</td>
</tr>
<tr>
<td>HSP71_HUMAN</td>
<td>Heat shock 70 kDa protein 1A/1B OS=Homo sapiens GN=HSPA1A PE=1 SV=5</td>
</tr>
<tr>
<td>ZG16B_HUMAN</td>
<td>Zymogen granule protein 16 homolog B OS=Homo sapiens GN=ZG16B PE=1 SV=3</td>
</tr>
<tr>
<td>DEF1_HUMAN</td>
<td>Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1</td>
</tr>
<tr>
<td>CYTT_HUMAN</td>
<td>Cystatin-SA OS=Homo sapiens GN=CST2 PE=1 SV=1</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>P01276</td>
<td>Human Protoonin 1</td>
</tr>
<tr>
<td>P02778</td>
<td>Human Albumin Translational</td>
</tr>
<tr>
<td>P33981,Q9UEW8</td>
<td>Human Tryptophan</td>
</tr>
<tr>
<td>P0C0S8,P20671,Q16777,Q6FI13,Q96KK5,Q99878,Q9BTM1</td>
<td>Human Histone H2A2</td>
</tr>
<tr>
<td>P68871</td>
<td>Human Hemoglobin A2</td>
</tr>
<tr>
<td>Q9UGM3</td>
<td>Human Transketolase</td>
</tr>
<tr>
<td>Q9UGM3</td>
<td>Human Dual Specificity Protein Kinase 1</td>
</tr>
<tr>
<td>P06733</td>
<td>Human Alpha 2 Enolase</td>
</tr>
<tr>
<td>P02787</td>
<td>Human Myeloperoxidase</td>
</tr>
<tr>
<td>P05164</td>
<td>Human Myeloperoxidase</td>
</tr>
<tr>
<td>P01040</td>
<td>Human Cystatin A</td>
</tr>
<tr>
<td>P10599</td>
<td>Human Thioredoxin</td>
</tr>
<tr>
<td>P08211</td>
<td>Human Guanidine S-Transferase</td>
</tr>
<tr>
<td>P60709,P63261</td>
<td>Human Actin alpha, cytoplasmic 1</td>
</tr>
<tr>
<td>P61769</td>
<td>Human Beta-2-Microglobulin</td>
</tr>
<tr>
<td>P02767</td>
<td>Human Transferrin</td>
</tr>
<tr>
<td>P01276</td>
<td>Human Trypsinogen</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>IGKC_HUMAN Ig kappa chain C region</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>PIGR_HUMAN Polymeric immunoglobulin receptor</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>CYTN_HUMAN Cystatin-SN OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>CYTS_HUMAN Cystatin-S OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>CYTC_HUMAN Cystatin-C OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>AMY1_HUMAN Alpha-amylase 1 OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>HPT_HUMAN Haptoglobin OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>ALBU_HUMAN Serum albumin OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>PUS7L_HUMAN Pseudouridylate synthase 7 homolog-like protein OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>HSP71_HUMAN Heat shock 70 kDa protein 1A/1B OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>TTHY_HUMAN Transthyretin OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>ILEU_HUMAN Leukocyte elastase inhibitor OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>S10A8_HUMAN Protein S100-A8 OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>SH3L3_HUMAN SH3 domain-binding glutamic acid-rich-like protein 3 OS=Homo sapiens</td>
</tr>
<tr>
<td>Accession</td>
<td>Name</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Q9UBG3</td>
<td>CRNN_HUMAN Cornulin OS=Homo sapiens</td>
</tr>
<tr>
<td>P06702</td>
<td>S10A9_HUMAN Protein S100-A9 OS=Homo sapiens</td>
</tr>
<tr>
<td>P69905</td>
<td>HBA_HUMAN Hemoglobin subunit alpha OS=Homo sapiens</td>
</tr>
<tr>
<td>P68871</td>
<td>HBB_HUMAN Hemoglobin subunit beta OS=Homo sapiens</td>
</tr>
<tr>
<td>P24158</td>
<td>PR10N3_HUMAN Myeloblastin OS=Homo sapiens</td>
</tr>
<tr>
<td>P04406</td>
<td>G3P_HUMAN Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens</td>
</tr>
<tr>
<td>P06744</td>
<td>G6P1_HUMAN Glucose-6-phosphate isomerase OS=Homo sapiens</td>
</tr>
<tr>
<td>P06396</td>
<td>GELS_HUMAN Gelsolin OS=Homo sapiens</td>
</tr>
<tr>
<td>P62937</td>
<td>PPIA_HUMAN Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens</td>
</tr>
<tr>
<td>P61626</td>
<td>LYSC_HUMAN Lysozyme C OS=Homo sapiens</td>
</tr>
<tr>
<td>P80188</td>
<td>NGAL_HUMAN Neutrophil gelatinase-associated lipocalin OS=Homo sapiens</td>
</tr>
<tr>
<td>Q01469</td>
<td>FABP5_HUMAN Fatty acid-binding protein, epidermal OS=Homo sapiens</td>
</tr>
<tr>
<td>P31025</td>
<td>LCN1_HUMAN Lipocalin-1 OS=Homo sapiens</td>
</tr>
<tr>
<td>P57071</td>
<td>PRD15_HUMAN PR domain zinc finger protein 15 OS=Homo sapiens</td>
</tr>
<tr>
<td>Q9HC84</td>
<td>MUC5B_HUMAN Mucin-5B OS=Homo sapiens</td>
</tr>
<tr>
<td>10-19KDa</td>
<td>PROTEIN IDENTIFICATION</td>
</tr>
<tr>
<td>----------</td>
<td>------------------------</td>
</tr>
<tr>
<td>A1AT_HUMAN</td>
<td>Alpha-1-antitrypsin</td>
</tr>
<tr>
<td>IL1RA_HUMAN</td>
<td>Interleukin-1 receptor antagonist protein</td>
</tr>
<tr>
<td>H2B1K_HUMAN</td>
<td>Histone H2B type 1-K</td>
</tr>
<tr>
<td>PIP_HUMAN</td>
<td>Prolactin-inducible protein</td>
</tr>
<tr>
<td>PROF1_HUMAN</td>
<td>Profilin-1</td>
</tr>
<tr>
<td>PLSL_HUMAN</td>
<td>Plastin-2</td>
</tr>
<tr>
<td>ACBP_HUMAN</td>
<td>Acyl-CoA-binding protein</td>
</tr>
<tr>
<td>H2A1_HUMAN</td>
<td>Histone H2A type 1</td>
</tr>
<tr>
<td>CYTB_HUMAN</td>
<td>Cystatin-B</td>
</tr>
<tr>
<td>LDHA_HUMAN</td>
<td>L-lactate dehydrogenase A chain</td>
</tr>
<tr>
<td>LAC_HUMAN</td>
<td>Ig lambda chain C regions</td>
</tr>
<tr>
<td>IGH1_HUMAN</td>
<td>Ig alpha-1 chain C region</td>
</tr>
</tbody>
</table>
Table 2-6: Proteomic analysis of SDS-PAGE gel bands 3-6KDa and 10-19KDa for CM10 pH6 saliva sample. Uniprot protein accession numbers (human entries) are provided for each assigned protein and where appropriate for database entries containing the same set of assigned peptides. Identities of similar proteins can be found or BLAST-searched at http://www.ncbi.nlm.nih.gov/ and http://www.uniprot.org/.
2.5 Discussion

The total number of proteins identified in all saliva and GCF samples was 115. However, only 67 (58.3%) of these were identified with ≥99% probability. More proteins were identified in the CM 10 pH6 fractions than in the IMAC fractions suggesting that CM10 weak cation exchange method was better at purification of salivary proteins.

Of the putative biomarkers flagged previously using SELDI-TOF MS, three peptides were identified conclusively using LC-MS/MS. These were HNP1 (sharing peptide sequence with HNP2 & 3), cathelicidin LL-37 and Protein S100A8.

Even after employing the use of 1D SDS-PAGE for the saliva samples and focusing on the lower molecular weight regions, only six proteins were identified with high confidence in this region. Further, few low molecular weight proteins were assigned and no others apart from the likely neutrophil protein were assigned with high confidence below 10KDa. There was the presence of fragments from large abundant proteins such as polyubiquitin and keratin1 dominating the assigned spectra indicating that further fractionation may be required to identify other proteins with a molecular weight of 3-6KDa.

A small number of proteins were identified with high confidence from the CM10 samples of purified GCF. Analysis of the IMAC fractions indicated lower quantity of proteins. No low molecular weight proteins were assigned and none assigned with high confidence below 10KDa but a 3KDa peptide was identified with 95% probability and assigned to neutrophil granule peptide HP1 and its isoforms.

As described in the previous section, the results show that with each proteomics strategy, partial overlapping subsets of saliva proteins are identified even as reported in other studies (Esser et al., 2008). It would be recommended to use different proteomic approaches in order to produce a more comprehensive identification of the saliva proteome.

The reason for the low detection of proteins in GCF is not clear although ultrafiltration to desalt using a 3KDa molecular weight cut-off filter followed by
in-solution digestion of retentate might improve upon any poor recovery by the in-gel digestion method. The disproportionate number of spectra from a small number of the proteins identified indicates the need for further fractionation of the samples to achieve more extensive coverage of the GCF proteome, as was the case with previous saliva analyses. The use of different strategies for the pre-fractionation of saliva and GCF has been suggested as a useful way of improving the spectrum of identifiable salivary proteome (Hu et al., 2006a). Proposed fractionation strategies include the use of molecular weight cut off filters prior to protein digestion and size exclusion chromatography using solid phase extraction. For the more focused approach required in screening for low molecular proteins, molecular weight cut-off filters or size exclusion could remove the higher molecular weight proteins particularly albumin and alpha-amylase but binding to the larger molecules might reduce the yield of proteins in the filtrate/eluate.

Alternatively, for more comprehensive discovery of the low molecular weight salivary proteome, the 2D-LC/MS/MS approach could be followed and sample proteins targeted for analysis i.e. separation of the low molecular weight fraction by SCX gradient elution and analysis of multiple sub-fractions. Also, the use of isotopically labeled chemical tags e.g. isobaric tags for relative and absolute quantification (iTRAQ) or tandem mass tags (TMT) would further enable 4-8 samples to be processed at the same time (with or without further fractionation) and semi-quantitative analysis performed on the assigned proteins discovered. The use of in-house Scaffold Q+ software would enable relative quantitation of detected proteins in a relatively short time.

One of the greatest challenges that still persisted was the inability to directly link the identified proteins with the SELDI-TOF MS peaks described before. This was due to the way the two different MS generate and handle the data produced. One possible way to circumvent this is to link the SELDI-TOF MS in tandem with the LC-MS/MS so that as soon as the ions representing a particular peak are detected by the SELDI-TOF MS, they are then passed to the LC-MS/MS for further analysis. In this way, it would be possible to analyse each of the peaks that SELDI-TOF MS predicts as a potential biomarker and ID it concurrently. Unfortunately, at the time of doing this study, such technology
was not available. However now, there are emerging claims of such techniques working as evidenced by the technique developed by collaboration between Bio-Rad (the manufacturers of SELDI-TOF MS) and Bruker Daltonic. These have attempted to successfully do on-chip identification, without proteolytic digestion using the Lucid Proteomics System™ (Bio-Rad and Bruker Daltonic).

Many of the peaks detected by SELDI-TOF MS as potential biomarkers were not identified using the LC-MS/MS technique used here. It is likely that masking by the larger and more abundant proteins occurred. At this point, the options for further study were either to continue to try to identify the other SELDI-TOF MS peaks with biomarker potential by methods discussed above or to validate the ones that had been identified in this chapter. Having obtained reliable data on new potential biomarkers, however, it was worth validating them at this point. The next two chapters will discuss the validation of these identified peptides including HNP1-3, cathelicidin LL-37 and Protein S100A8.
Chapter 3 – Preliminary Biomarker Validation
3.1 Introduction

Validation stems from the word validity, which means degree of closeness of the reported results to the true correct results. In other words, ‘how true are the results in reality?’ The process of validating biomarkers therefore must involve establishing a range of characteristics including the intrinsic qualities of the biomarker, its determinants, and the analytic procedure (Bonassi et al., 2001).

The aim of a validation phase is to assess the validity of the described putative biomarkers against a larger and more heterogeneous population of patients. During this phase, the robustness of the candidate markers would be tested against a level of biological variability that more accurately reflects the variability in the target population (De Bock et al., 2010). Therefore, in order to position salivary diagnostics to be a novel, accurate and feasible technology to help us attain definitive point-of-care assessment of individuals’ periodontal health status, validation studies are crucial (Li et al., 2005).

Validation can be done both in retrospective as well as prospective studies. However the latter is resource hungry in terms of time and cost. It can also be done in the same population or a different set of population depending on the purpose. A validation exercise aims to establish the sensitivity and specificity of a test, which in turn is used to determine the test’s threshold of disease detection.

In order to determine the most appropriate cut-off point for the biomarker levels obtained, receiver operating characteristic (ROC) values and scatter plots are useful to evaluate their potential as biomarkers. A ROC curve is generated by plotting the true positive values or rate (TPR or sensitivity) on the y-axis against false positive values or rate (FPR or 1-specificity) on the x-axis (Figure 3-1). Sensitivity is the ability of the test to detect the disease when the disease is present i.e. true positive rate (TPR), while specificity is the probability that the test will be negative when the disease is not present i.e. true negative rate. Subtracting the true negative rate from 1 gives the false positive rate (FPR). The ROC plot is very informative in that it provides insight into the accuracy of the values. The values of ROC lie between 0 and 1 and the closer the value is to 1 the more the chance that a random selection from the positive group will
have a score greater than a random selection from the negative cases. It follows that a straight line on the plot (red dotted line in Figure 3-1) would mean that the scores for the two groups do not differ and therefore the test is not useful.

![ROC Space Diagram](http://en.wikipedia.org/wiki/File:ROC_space-2.png)

**Figure 3-1**: ROC construct showing the dimensions that determine the classification of a test as either perfect, better or hopeless. FPR (False positive rate) and TPR (true positive rate). (Adapted from online Wikipedia- http://en.wikipedia.org/wiki/File:ROC_space-2.png)

As outlined in the previous chapter, the process of biomarker discovery involves several steps that have to be carried out in order to identify and establish the utility of the biomarkers so discovered. So after the initial screening for potential diagnostic and prognostic biomarkers for periodontal disease in GCF and saliva described in chapter one using SELDI-TOF MS, and protein identification in chapter two, there was need to validate those markers or peaks of known identity (Clarke, 2007; Mulli, 2008).

Based on published data of peaks from SELDI-TOF MS data, the identities of five of the putative biomarkers were tentatively established (Table 3-1). It was felt that measurement of these peptides using a different technique in the same
saliva and GCF samples previously analyzed using SELDI-TOF MS would serve to validate the identified putative biomarkers. If successful, further validation studies would be carried out in an independent cohort. Other new biomarkers identified by LC-MS/MS would also be tested in the new cohort.

<table>
<thead>
<tr>
<th>MASS FOR CHARGE (M/Z)</th>
<th>PEPTIDE IDENTITY</th>
<th>AMINO ACID SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3443</td>
<td>HNP 1</td>
<td>ACYCRIPACIAGERRYGTCIYQGRLWAFCCC</td>
</tr>
<tr>
<td>3372</td>
<td>HNP 2</td>
<td>CYCRIPACIAGERRYGTCIYQGRLWAFCCC</td>
</tr>
<tr>
<td>3487</td>
<td>HNP 3</td>
<td>DCYCRIPACIAGERRYGTCIYQGRLWAFCCC</td>
</tr>
<tr>
<td>4492</td>
<td>LL-37</td>
<td>LLGDFFRKSKEKIGKEFKRIVQIRIKDFLRNLVPRTES</td>
</tr>
<tr>
<td>5060</td>
<td>HBD 3</td>
<td>GIINTLQKYYCRVGRGRCAVLSCLPKEEQIGKCSTRGRKCCRRKK</td>
</tr>
</tbody>
</table>

Table 3-1: The identity of the potential biomarker peaks candidate for validation studies

There were two phases of validation. Firstly was preliminary validation and secondly, validation in an independent cohort. In order to confirm that the putative biomarkers discovered using SELDI-TOF MS were actually present and that in reality they were capable of differentiating between gingivitis and periodontitis, it was necessary to do preliminary validation using a different technique to measure the putative biomarkers in the same samples that had been used during the SELDI-TOF MS biomarker discovery phase. This will be discussed in the present chapter while the validation in an independent cohort will be discussed in Chapter 4.
3.1.1 Aims and Objectives

The aim of this study was to determine if the putative diagnostic biomarkers (AMPs) flagged by the SELDI-TOF MS biomarker discovery process (Table 3-1) retained their diagnostic utility when measured using a different technique.

The specific objectives were:

1. To determine if salivary AMP profiles of patients with periodontitis differ from those with gingivitis.
2. To determine if salivary AMP profiles of patients with chronic periodontitis differ from those with aggressive periodontitis.
3. To determine if salivary AMP profiles of patients differ between pre-treatment and post-treatment in subjects with AgP.
4. To determine if GCF AMP profiles of patients with periodontitis differ from those with gingivitis.
5. To determine if GCF AMP profiles of patients with chronic periodontitis differ from those with aggressive periodontitis.
6. To determine if GCF AMPs profiles of patients differ between pre-treatment and post-treatment in subjects with AgP.
3.2 Methodology and materials

3.2.1 Ethics approval

The protocol for this study was approved by the City and East London Health Authority Local Research Ethics Committee (05/Q0601/80). All participants signed an informed written consent form.

3.2.2 Patient Cohort and Sample Acquisition.

The current study utilized banked saliva and GCF samples from two cohorts of patients participating in two ongoing studies. The first cohort (Diagnostic cohort) had been recruited for a study on periodontal diagnostic biomarker discovery using SELDI-TOF MS. The second cohort (Prognostic cohort) was set to investigate prognostic factors for initial response to non-surgical periodontal treatment for generalized aggressive periodontitis. In both cases saliva samples were collected using Salivettes® (Sarstedt Ltd) while GCF was collected using Periopaper® (Pro Flow Inc., Connecticut, USA) inserted into the periodontal crevice at a depth of about 1mm for 30 seconds. The supragingival plaque was gently removed with cotton roll before sample acquisition. All samples were immediately stored at minus 80°C until time for analysis.

Inclusion criteria

For both cohorts, participants had to give informed written consent to take part in the study before full clinical assessment was carried out and samples collected. Diagnosis was assigned using the following clinical criteria as seen in Table 3-2.

Exclusion criteria

Patients were excluded if they had a history of periodontal treatment or antimicrobial therapy within the previous 6 months, or were affected by any systemic condition(s) or taking any medication known to affect the periodontal tissues.
Clinical Diagnosis | Clinical Criteria for Diagnosis
---|---
**Gingivitis (G)** | • Signs of inflammation, bleeding on probing.  
  • No clinical or radiographic signs of attachment loss.

**Chronic Periodontitis (CP)** | • Diagnosis of CP with attachment loss of greater than 6 mm affecting a minimum of 6 teeth.  
  • Over the age of 40 years old.

**Aggressive Periodontitis (AgP)** | • Diagnosis of AgP with attachment loss of greater than 6 mm affecting a minimum of 6 teeth  
  • Under the age of 40 years old.

Table 3-2: Summary of the clinical criteria for the diagnosis of the periodontal condition

### 3.2.2.1 Samples
In both cohorts, saliva samples were collected using Salivettes® (Sarstedt Ltd) by asking the patient to chew gently on the cotton roll for 60 seconds. The samples were spun from the cotton roll by centrifugation at 3000rpm for 5 minutes. GCF samples were collected using Periopaper® (Pro Flow Inc., Connecticut, USA) inserted into the periodontal crevice at a depth of about 1mm for 30 seconds. The supragingival plaque was gently removed with cotton roll before sample acquisition. All samples were immediately stored at minus 80 degrees Celsius until time for analysis.

### 3.2.2.2 Diagnostic Cohort
Patients had been recruited from the new referrals clinic at The Dental Teaching Hospital, Bart’s and The London NHS Trust following assessment by two consultant periodontists. The cohort composed of ninety patients. Thirty patients with a clinical diagnosis of aggressive periodontitis with attachment loss of greater than 6 mm affecting a minimum of 6 teeth and under the age of 40 years old were recruited. Also thirty-one clinically diagnosed chronic periodontitis cases were included. Twenty-nine patients with gingivitis but no radiographic evidence of bone loss were recruited from the dental emergency
clinic (Table 3-2). The inclusion and exclusion criteria described above were applied.

Following written informed consent, patients received full clinical assessment and details of smoking status were recorded and supplemented by measurement of expired carbon monoxide using piCO Smokerlyzer® (Bedfont Scientific Ltd., Kent, UK). Patients were subsequently characterized as current smokers, previous smokers or never smokers.

3.2.2.3 Prognostic Cohort

Details of the patient cohort have previously been described (Hughes et al., 2006). In summary, patients were recruited from the new referrals clinic, The Dental Teaching Hospital, Bart’s and The London NHS Trust following assessment by a consultant periodontist. Patients with a clinical diagnosis of aggressive periodontitis with attachment loss of greater than 6 mm affecting a minimum of six teeth between the ages of 18 and 40 years old were eligible for inclusion in this study. All patients included in the study had sites exhibiting 6 mm clinical attachment loss (CAL) affecting more than two teeth in addition to incisors and first molar. Samples of gingival crevicular fluid from six affected sites were taken using Periopaper® (Pro Flow Inc., Connecticut, USA).

Seventy-nine (79) patients with generalized aggressive periodontitis were included in this prospective follow-up intervention study. Patients’ clinical and demographic parameters were collected at baseline and 10 weeks following a standard course of treatment (four visits of non-surgical root surface debridement together with OHI as required). At the baseline and 10 weeks post-treatment, GCF and saliva samples were collected and stored immediately at minus 80 degrees Celsius before analysis.

3.2.3 Total Protein Quantification

The total protein content for saliva samples was assayed using a commercial kit (Bicinchoninic Acid Protein (BCA) Assay Kit, Sigma-Aldrich, Inc) according to the manufacturers instructions. In brief, samples were diluted by a factor of x10. 25µl of the standards, negative control and samples were applied in duplicates on a 96 well plate, then 200µl of the BCA working solution added to
the wells followed by a 30 minutes incubation at 37 °C. Serial dilution of bovine serum albumin (BSA) preserved in 0.05% sodium azide was used for standards. The absorbance of the developed colour was measured using a colorimetric reader (Dynex Revelation® 4.24) at 470nm. The standard curve was plotted using Numbers® spread sheet software for Apple Macintosh.

3.2.4 ELISAs

ELISA tests for HNP 1-3, LL-37, HBD-3 and IL-1α were performed as a means of validating the ProteinChip experiments (Clarke, 2007; Mulli et al., 2008). The IL-1α was measured as a surrogate marker for periodontal inflammation. ELISAs were done using commercial kits for HNP 1-3 and LL-37 (Hycult biotechnology b.v. The Netherlands) and HBD-3 (Alpha Diagnostic International, USA) while the IL-1α DuoSet ELISA development kit (RandD Systems, inc. USA) was used to determine the levels of IL-1α run according to the manufacturer’s instructions (See Appendix 8.2.1 for detailed protocol). In brief, samples and ELISA reagents were thawed at room temperature. Both the test and control samples for saliva were diluted at an optimized dilution factor of x5,000 for HNP 1-3, x5 for LL-37, x15 for HBD-3 and x10 (with a repeat at a dilution of x20 for those with a concentration above the detection limit) for IL-1α. HNP 1-3 ELISA for GCF was done with a dilution factor of x500000. Reading of optical density was done at 450nm using either Optima® or Dynex Revelation® 4.24. All samples were run in duplicates with an internal control of standards and standard curves plotted using computer spreadsheet software (Microsoft Excel® or Numbers® spread sheet software for Apple Macintosh).

3.2.5 Data analysis

Statistical analysis was carried out using Prism 5 for Mac OS X (GraphPad Software, Inc). One-way analysis of variance (ANOVA) was used to analyse the difference in expression of the different analytes in the various periodontal conditions. Where only two variables were compared, two tailed t-test was performed with significance set at p = 0.05. Scatter plots were then generated with standard deviation indicated. ROC constructs were created for those markers showing a significant difference in expression and cut off points were set for optimal sensitivity and specificity.
3.3 Results

3.3.1 Diagnostic Cohort

3.3.1.1 Total protein

Total protein concentration for saliva was determined in 57 samples out of the original 90. The missing samples were due to progressive sample depletion during previous SELDI-TOF MS experiments. There was no statistical difference ($p = 0.2$) between gingivitis (1.65 mg/ml, SD 0.56 mg/ml; $n = 20$), CP (2.18 mg/ml, SD 1.41 mg/ml; $n = 19$) and AgP (2.03 mg/ml, SD 0.73 mg/ml; $n = 18$) based on total protein content.

![Figure 3-2: Scatter graph for salivary total protein concentration in the diagnostic samples from subjects with gingivitis (G), chronic periodontitis (CP) and aggressive periodontitis (AgP).](image)

3.3.1.2 HNP 1-3

**GCF**

For GCF studies, there were 47 samples each from a single posterior site per patient comprising 19 gingivitis, 14 chronic periodontitis and 14 aggressive periodontitis cases. The mean concentration of HNP 1-3 in gingivitis was 29.34 µg/ml (SD 6.65 µg/ml); chronic periodontitis was 36.00 µg/ml (SD 1.89 µg/ml) and aggressive periodontitis 33.88 µg/ml (SD 4.34 µg/ml). One way Analysis of variance (ANOVA) for the three diagnostic variables revealed significant
difference in the levels of CP and AgP compared to gingivitis (p value = 0.0012) (Figure 3-3).

Figure 3-3: HNP 1-3 levels measured using ELISA in GCF showing significantly greater levels in both chronic periodontitis (CP) and aggressive periodontitis (AgP) than in gingivitis (G).

Figure 3-4: Receiver characteristic curve (ROC) for HNP 1-3 differentiating between (A) gingivitis and chronic periodontitis and (B) gingivitis and aggressive periodontitis.
The area under the curve (AUC) from the receiver operated curve (ROC) constructs was 87% (p value = 0.0003) for gingivitis versus CP and 74% (p value = 0.02) for gingivitis versus AgP (Figure 3-4). A cut off > 31.66 µg/ml could differentiate periodontitis from gingivitis with a sensitivity of 100% and a specificity of 57.89%. Increasing this cut off to >35.06 µg/ml retained a high sensitivity (78.59%) and specificity (84.21%). A cut of >31.68 µg/ml to differentiate between gingivitis and AgP had a High sensitivity (85.71%) and moderate specificity (57.89%). Adjusting this cut off to >34.96 µg/ml reduced the sensitivity to 57.14% while the specificity was increased to 84.21%.

**SALIVA**

For salivary HNP 1-3 levels, there were a total of 60 samples comprising of gingivitis (n = 18), CP (n = 22) and AgP (n = 20). There was a significantly higher level of salivary HNP 1-3 in CP (0.70 µg/ml; SD = 0.75 µg/ml) as compared to gingivitis (0.22 µg/ml SD 0.23 µg/ml) by t-test (p = 0.014). The levels in AgP were 0.48 µg/ml and this was not significantly different from either gingivitis or CP (Figure 3-5A). However, when both periodontitis cases (CP and AgP) were combined, salivary HNP 1-3 remained significantly higher in periodontitis (0.59 µg/ml; SD = 0.75 µg/ml) as compared to gingivitis by t-test (p = 0.046). See Figure 3-5B.

A cut off > 0.1762 µg/ml could differentiate CP from gingivitis with a sensitivity of 86.36% and a specificity of 55.56%. Increasing this cut off to >0.2942 µg/ml increased the specificity (83.33%) while moderately reducing sensitivity (63.64%).
Figure 3-5: (A) Scatter plots of salivary HNP 1-3 in different periodontal health status showing significantly higher expression in CP than in gingivitis. (B) Box and whisker (Tukey) plot showing higher expression of salivary HNP 1-3 in periodontitis than in gingivitis.

Figure 3-6: ROC construct for salivary HNP 1-3 with 75.3% AUC (p=0.0066) a sensitivity of 72.73% and specificity of 61.11% to differentiate gingivitis from CP at a cut off of >0.2087 µg/ml.
The levels of HNP 1-3 were higher in CP than AgP and gingivitis both in saliva and GCF. However, the levels in GCF were higher than in saliva by 132, 52 and 71 fold in gingivitis, CP and AgP respectively (Figure 3-7). However, using regression analysis on 35 subject paired GCF and saliva samples, there was no significant correlation between HNP 1-3 levels in saliva and those in GCF within subjects, $r^2 = 0.08$, p value = 0.10 (Figure 3-8).
Figure 3-8: Linear regression plot for HNP 1-3 levels in saliva against that in GCF paired within subjects showing lack of significant correlation within subjects.

3.3.1.3 LL-37

Due to the high volume of sample needed for running the LL-37 ELISA, only 30 samples (gingivitis n = 11, CP n = 11 and AgP, n = 8) were available for analysis. However for these samples (Figure 3-9), salivary LL-37 levels were significantly higher in AgP (53.9 ± 55.0ng/ml; mean ± SD) than in both G (p=0.007; 2.98 ± 3.8ng/ml, mean ± SD) and CP (p=0.013; 7.49 ± 10.1ng/ml; mean ± SD). Also when gingivitis was compared to periodontitis (CP + AgP), there was higher expression of LL-37 in the periodontitis group (29.04 ng/ml; SD = 47.0 ng/ml. n = 19) than in the gingivitis group (2.93 ng/ml; SD = 3.15 ng/ml. n = 11) though this was not statistically significant. ROC constructs for gingivitis versus AgP produced 100% specificity and sensitivity at a cut off point of salivary LL-37 levels >11.89 ng/ml (AUC of 100%, p = 0.0001) while that between CP and AgP were 100% sensitivity and 81.8% specificity at a cut off point of > 10.16 ng/ml (AUC of 93%, p = 0.002) Figure 3-10.
3.3.1.4 HBD-3

A total of 54 samples were analyzed for hBD-3 using a commercial ELISA kit. Although the mean levels of salivary hBD-3 were higher in both CP (571.4 pg/ml; SD = 268.5 pg/ml; n = 20) and AgP (468.6 pg/ml; SD = 653.7 pg/ml; n =
18) than in gingivitis (432.6 pg/ml; SD = 259.6 pg/ml; n = 16), this difference did not reach statistical significance Figure 3-11.

Figure 3-11: A graph of salivary hBD-3 expression in gingivitis, Chronic periodontitis (CP), Aggressive Periodontitis (AgP) and Periodontitis (CP + AgP) showing lack of significant difference in the different conditions.

3.3.1.5 IL-1α

There was no difference in the mean levels of salivary IL-1α in gingivitis (2.66 ng/ml; SD 1.35 ng/ml), CP (2.66 ng/ml; SD 1.13 ng/ml), and AgP (2.83 ng/ml; SD 1.67 ng/ml). When gingivitis was compared with periodontitis (CP and AgP combined) no statistically significant differences were found (Figure 3-12).
3.3.2 Prognostic Cohort

Analysis of the saliva samples for the prognostic cohort was done among aggressive periodontitis patients before and after treatment as described before. The number of pre-treatment samples available was not sufficient to match the number of post-treatment samples. Analysis was therefore done in two ways. Firstly, unmatched pre-treatment levels were compared with post-treatment levels without pairing. Secondly, where available, the paired samples were compared.

3.3.2.1 Total protein quantification

The average total protein for unmatched samples of saliva revealed significantly higher (t-test; p = 0.025) pre-treatment levels (1.74 mg/ml; SD = 0.46 mg/ml; n = 14) as compared to post-treatment levels (1.32 mg/ml; SD = 0.64 mg/ml; n = 48) Figure 3-13.
3.3.2.2 HNP 1-3

GCF

HNP 1-3 analysis (Figure 3-15A) in 29 paired GCF samples from AgP patients before treatment was significantly higher ($p = 0.024$) ($91.41 \mu g/ml$, SD 128.0 \mu g/ml) than after treatment ($35.63 \mu g/ml$, SD 35.87 \mu g/ml) (Figure 3-15A). ROC integral (Figure 3-15B) for HNP 1-3 in before and after treatment showed an AUC of 71% ($p = 0.005$). At a cut of point of < 25.44 \mu g/ml of HNP 1-3 in GCF in these patients, it was possible to differentiate treated from untreated cases with a sensitivity of 58.6% and a specificity of 75.9%. If the cut off was adjusted to < 63.63 \mu g/ml, the sensitivity increased to 86.2% while the specificity was 41.4%.

There was no difference (t-test, p value = 0.09) in the pre-treatment levels of HNP 1-3 in GCF of both responders (those with <30% non responding sites) as compared to non-responders (NR, those with >30% non-responding sites). Similarly, plaque score, smoking status and age did not show any correlation with the levels of HNP 1-3.
Figure 3-14: Scatter graph for GCF HNP 1-3 concentrations showing lack of statistical difference between responders and non-responders in AgP patients after non-surgical treatment.

Figure 3-15: (A) Scatter graph of HNP 1-3 in paired GCF samples before and after treatment in subjects with AgP showing statistically significant reduction after treatment (p = 0.024). (B) A ROC integral for HNP 1-3 showing AUC of 71% (p = 0.005).
SALIVA

All salivary levels of HNP 1-3 in the same cohort could not be paired due to insufficient sample volume. There was no significant difference in HNP 1-3 levels before and after treatment as well as between responders and non-responders to treatment. Even when only the few paired pre and post-treatment samples were compared there was still no significant difference. Smoking status and oral hygiene status were not statistically significantly different between the responders and non-responders.

![Figure 3-16: Scatter graph for salivary HNP 1-3 showing lack of statistical difference between (A) pre-treatment and post-treatment (B) responders and non-responders in AgP patients after non-surgical treatment.](image)

3.3.2.3 LL-37

LL-37 levels were not statistically different (t-test, \( p = 0.2 \)) in pre-treatment saliva samples (mean = 21.17ng/ml; SD = 24.16ng/ml; \( n = 10 \)) as compared to post-treatment samples in subjects with AgP (mean = 44.48ng/ml; SD = 57.71ng/ml; \( n = 40 \)) Figure 3-17.
3.3.2.4 HBD-3

The difference in hBD-3 expression levels between pre-treatment versus post-treatment saliva levels were analyzed using t-test with significance set at $p \leq 0.05$. There was no statistically significant difference between the two measurements ($p = 0.78$). The mean salivary hBD-3 levels before and after treatment for the unmatched pairs were 745.4 pg/ml ($n = 9$) and 972.5 pg/ml ($n = 41$) respectively while those for the four-paired samples were 1150 pg/ml versus 2127 pg/ml (Figure 3-18).
Because of sample depletion, not all the saliva samples for both pre-treatment and post-treatment AgP could be measured. In order not to miss any emerging patterns, the data were analyzed in two groups such that all unmatched pre- and post-treatment samples were compared using unpaired t-tests while the fewer paired samples among these were re-analyzed separately using paired t-test. Both analyses yielded a similar pattern of results with a significant reduction of salivary IL-1 alpha post-treatment as compared to pre-treatment (Figure 3-19). For the unmatched pairs mean pre-treatment levels were 2.336 ng/ml (n=12; SD = 1.661 ng/ml) while the post-treatment levels were 1.089 µg/ml (n=46; SD = 0.855 ng/ml) with p value = 0.0006 for the reduction. The matched pairs had a reduction (p value = 0.01) with mean pre-treatment levels of 2.681ng/ml (n=6; SD = 1.552 ng/ml) while the post-treatment levels were 1.285 ng/ml (n=6; SD = 0.871 ng/ml) Figure 3-19.

Figure 3-18: Scatter graph showing the expression levels of salivary hBD-3 in patients with AgP before and after treatment for both unmatched and marched pairs respectively.

3.3.2.5 IL-1 α
Figure 3.19: Scatter plot showing significant reduction of salivary IL-1α levels in post treatment AgP. The first two scatter plots are of unmatched samples while the latter two are matched pairs.
3.4 Discussion

The aim of the current study was to validate the diagnostic utility of the putative diagnostic salivary and GCF biomarkers identified previously using SELDI-TOF-MS to distinguish between patients with gingivitis, chronic periodontitis (CP) and aggressive periodontitis (AgP). Validation was carried out using a different proteomic technique, namely ELISA on the same samples that were analyzed in the initial study for biomarker discovery using SELDI-TOF MS platform as described in Chapter 1.

Saliva and GCF contain molecular products derived not only from the host tissues and serum but also from the oral microbiota including the subgingival microbial plaque (Curtis et al., 1989). It was therefore felt that saliva and GCF could offer a quick non-invasive patient-based as well as site-specific source of an extremely broad range of candidate molecules that could potentially be identified as biomarkers of periodontal disease. These would be valuable both as diagnostic and prognostic markers and could potentially be developed to form part of targeted periodontal therapeutics.

Initial work on saliva and GCF in the diagnostic and prognostic cohorts using SELDI-TOF MS had revealed a total of 41 peaks in saliva and 76 peaks in GCF that were significantly different between gingivitis and periodontitis (Clarke, 2007; Mulli et al., 2008). From previously published data, it was possible to tentatively identify some of the putative biomarkers by their mass/charge ratio (m/z) e.g. α-defensins such as human neutrophil peptide (HNP) – HNP 1 (m/z 3443 Da), HNP 2 (m/z 3372 Da) and HNP3 (m/z 3487 Da), cathelicidin LL-37 (m/z 14KDa) and human β-defensin 3 (hBD-3) with m/z 5060. These markers also had high relative peak intensities and ability to differentiate between different periodontal disease states using SELDI-TOF MS. Due to these unique features, their quantification using ELISA technique was carried out to validate the findings.

The data gathered from ELISA revealed variability of the amounts of measured AMPs in each diagnostic category. This observation may partly be explained by the inherent inter-individual variation in the ability to mount a robust immune defense against the offending microorganisms. A few of the samples had
measured values below detection level of the ELISA kit used and were interpreted as having “zero” amounts of the analytes. The amount of salivary HNP1–3 was two-fold and three-fold higher in patients with AgP and CP respectively than in gingivitis. HNP 1-3 levels were more than 5 fold higher in GCF samples for both CP and AgP compared to gingivitis. The mean levels in GCF were also higher than in saliva Figure 3-7. The presence of these peptides in GCF has been reported before in a study on detection of human defensins secreted by human oral epithelial cells using SELDI-TOF MS (Diamond et al., 2001). They did not however report on the periodontal diagnoses of the subjects from whom they took the GCF samples making it impossible to make any further comparisons. Another study found increased expression of HNP 1-3 in periodontitis than healthy group as confirmed in this work (Dommisch et al., 2009). Our results however are in contrast to two other recent reports. One reported more expression of HNP 1-3 in health than disease using MALDI-TOF (Lundy et al., 2005). The other reported a two-fold and four-fold increase from gingivitis levels when compared to CP and AgP respectively (Puklo et al., 2008). The reason for this discrepancy is not clear. However it can be speculated that genetic variations play a part in the different populations such as genetic polymorphisms or copy number variations (CNVs) of the HNP 1-3 genes that have been associated with differential production of the peptides (Aldred et al., 2005; Ballana et al., 2007; Linzmeier and Ganz, 2005; Nuytten et al., 2009; Redon et al., 2006). In summary, these differences may reflect the heterogeneity of patients in the two sites, the inherent biological complexity and diversity of different sample types, and or differences in sample collection, processing, handling, and analysis techniques used by the two different laboratories.

Interestingly, although mean GCF levels of HNP 1-3 are proportionately higher than in saliva there was no within-subject correlation. Neutrophils permeating into GCF have been postulated to be the main source of HNP 1-3 in saliva (Dale et al., 2001; Dale and Krisanaprapornkit, 2001; Dale and Fredericks, 2005; Pisano et al., 2005; Tao et al., 2005). Our findings however suggest that GCF is not the sole source of HNP 1-3 in saliva. A recent study found that unlike in healthy periodontal tissues, there was no correlation between neutrophil
numbers and HNP 1-3 levels in GCF of subjects with periodontal disease suggesting that there may be other sources of HNP 1-3 especially in inflamed periodontal tissues (Puklo et al., 2008). Such other sources may include lymphocytes (Agerberth et al., 2000) or other yet unknown sources.

Defensins are antimicrobial peptides with a wide range of antimicrobial activity as well as a stimulatory effect on fibroblast proliferation (Abiko and Saitoh, 2007; Abiko et al., 2007; Ganz et al., 1985; Ganz and Lehrer, 1995; Lundy et al., 2008; Miyasaki et al., 1990; Yoshioka et al., 2007). It is interesting to notice that their expression was more pronounced in periodontitis than in gingivitis, which would support our thinking that they are protective. At concentrations higher than 100 µg/ml, HNP-1-3 are able to kill Gram-negative and Gram-positive bacteria, including both intracellular and extracellular organisms, as well as enveloped viruses such as members of the Herpes family (Daher et al., 1986; Ericksen et al., 2005; Schneider et al., 2005). In addition, these peptides have been recently shown to play an anti-inflammatory role probably by inhibiting the secretion of multiple pro-inflammatory cytokines and nitric acid (NO) from macrophages, the main innate immune cell found at sites of chronic inflammation thereby regulating a potentially tissue-destructive inflammatory response (Doss et al., 2009; Miles et al., 2009).

These validation studies corroborate well with previous studies which have demonstrated differences in GCF protein expression between subjects with and without destructive periodontal disease using other methods such as 2-dimensional polyacrylamide gel electrophoresis (2-D PAGE) and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS/PAGE) (Curtis et al., 1990; Mukherjee et al., 1987). Previously 2-D PAGE has been used to show that GCF in a beagle dog model contained proteins with molecular weight 16 kDa or less and that the number of such proteins was greater in samples collected from periodontitis sites rather than the gingivitis sites (Mukherjee et al., 1987). This gives further confirmation that the process of periodontal destruction modifies the GCF composition and therefore GCF could lend itself favourably for the search of diagnostic and probably prognostic markers.
It is well known that neutrophils play a critical role in primary defence against assorted microorganisms in the periodontal tissue (Al-Shibani et al., 2011; Carlsson et al., 2006; Dale et al., 2001; Nussbaum and Shapira, 2011). It is now clear from this work that besides the phagocytic activity of neutrophils, released factors such as these antimicrobial peptides could play a critical role in maintaining the balance between health and disease.

To-date, there has not been any factor in GCF or saliva known that could distinguish between AgP or CP (Loos and Tjoa, 2005). It was exciting to find that salivary LL-37 could differentiate clearly AgP from both CP and gingivitis cases with very high specificity and sensitivity. However, the sample size was small and whether this observation would be replicated in a larger cohort should be looked into further.

Although both forms of destructive periodontal disease had greater mean values of salivary HBD-3 than in gingivitis, the difference was not statistically significant. A larger sample may be able to demonstrate a significant difference but this remains to be tested further. Similar results have been confirmed in a previous study that looked at beta defensin mRNA expression in gingival biopsies (Dommisch et al., 2005). Unlike HNPs that are produced by neutrophils, hBD-3 is produced by epithelial cells locally at the site of assault. The lack of differential expression in periodontitis could be due to background production as a means of priming the mucosa to defend itself whether in disease or health. It might well be that the peak detected by SELDI-TOF MS at m/z 5060 is not hBD-3 but a degradation product of a larger protein or a post-translationally modified peptide. This further highlights the difficulties of interpreting the peak data obtained using SELDI-TOF MS for biomarker discovery as described in chapter 2.

**3.4.1 Conclusion**

This study has validated the preliminary findings that profiling of saliva and GCF by SELDI-TOF-MS can distinguish between gingivitis and periodontitis and between CP and AgP. It is therefore possible to identify specific proteins in saliva and GCF from periodontitis patients that can be used as putative
diagnostic biomarkers. Salivary and GCF analysis may play an important role in the classification of periodontal diseases.

Since periodontal disease is by large a syndrome in which the inflammatory process is out of balance, it would be important to unmask the specific roles of both alpha defensins and LL-37 in this process.

From this preliminary study, it can be noted that SELDI-TOF is capable of detecting discriminating markers between health and periodontitis. However, it is not possible to directly assign specific peptide or protein identities to the peaks based on mass for charge ratio only. As demonstrated with the SELDI-TOF MS peak with m/z 5060 that was thought to be hBD-3, it was not possible to reproduce the profile using ELISA. Also, hBD-3 was not one of the proteins or peptides identified using LC-MS/MS. This raises the need for more robust protein identification work for the putative diagnostic biomarker peaks using standard methods. However, it suffices to note that, even without knowing what the identity of the discriminating peaks are, SELDI-TOF MS can be a useful stand alone tool for generating reliable diagnostic fingerprints for periodontal disease.

Having confirmed that HNP1-3 ELISA assays can discriminate between gingivitis and periodontitis, further validation by testing the assays on a newly collected independent patient population was used to test the robustness of these biomarkers, as described in the next chapter.
Chapter 4 Validation in an independent cohort
4.1 Introduction

As previously discussed, many studies have demonstrated that there are factors at the molecular level in both saliva and GCF that can differentiate between periodontal health and disease. However, most of these biomarkers have commonly been inflammatory markers and as such are not uniquely periodontitis markers (Chapple, 2009). The results presented in the previous chapters are based on demonstrating salivary and GCF markers that can differentiate between two inflammatory conditions i.e. gingivitis and periodontitis. This is useful as any differences in expression levels between the two conditions would not be due to inflammation per se but more due to the uniqueness of the clinical entity. HNP 1-3 has been shown in the previous chapter to fulfill this criterion.

In the previous chapter it was shown that both HNPs and LL-37 have the potential to be clinically useful salivary biomarkers of periodontal disease, based on their ability to distinguish between disease phenotype groups. However for a biomarker to have real utility it needs to be able to identify disease state (diagnosis) in a subject of unknown disease status. Thus to investigate this further, a reference population of samples needs to be tested where subjects were not recruited according to periodontal diagnosis, and therefore may have continuously variable severity of periodontal disease. With a sample cohort such as this, the previously identified cut off point for disease diagnosis can be applied to the new data and tested for its ability to determine periodontal disease status.

In this case, through a collaboration, we were able to obtain saliva samples of over 250 subjects collected to investigate the effects of Type I diabetes mellitus (DM) on periodontal disease and therefore ideal for testing the biomarkers in an independent cohort; in addition it enabled us to investigate the effects of DM on biomarker concentrations.
4.1.1 Aims and Objectives

The aim of this study was to validate two groups of putative biomarkers in an independent cohort:

1. To validate salivary antimicrobial peptides HNP 1-3 and cathelicidin LL-37 as putative diagnostic markers for periodontal disease in an independent cohort selected based on their DM status and not periodontal status.

2. To validate the newly identified salivary protein S100A8 in a cohort selected based on their DM status and not periodontal status.

3. To test the diagnostic utility of combined markers as opposed to single markers.
4.2 Materials & Methods

Saliva samples for the new independent cohort were obtained courtesy of collaboration with Dr. Penny Hodge and colleagues of Glasgow University. The samples had been collected based on the patients’ diabetic and not periodontal status. The protocol for the study was approved by the local research ethics committee, (Glasgow Royal Infirmary Research Ethics Committee, NHS Greater Glasgow and Clyde; Reference number: 05/S0705/70). Three hundred and fifteen patients were recruited from local outpatient diabetes and physiotherapy clinics in Greater Glasgow and Clyde Health Board. Subjects were aged between 20 and 55 (mean = 30.93) years and had been non-smokers for a minimum of 5 years before enrolling in the study.

4.2.1 Exclusion Criteria

The following exclusion criteria were applied: pregnancy; immunosuppression; medication with side effects affecting the oral cavity; patients prescribed antibiotics or anti-inflammatory drugs within the previous 6 weeks and patients with less than 20 teeth and smoking in the last 5 years.

4.2.2 Sample acquisition

Patients with type-1 diabetes had been diagnosed with diabetes for a minimum of 5 years. After giving informed consent, patients attended Glasgow Dental Hospital for sampling and clinical examination. Blood samples were taken from control subjects to confirm that the patients were definitely controls and not in a pre-diabetic state (glycosylated haemoglobin (HbA1c) < 5.9%). HbA1c levels for the patients with diabetes were determined by averaging the HbA1c measurements for each patient over the previous 2 years as recorded in the clinical notes.

Saliva was collected according to a modification of the method described by (Navazesh and Christensen, 1982). The method is described below.

The patient was instructed to refrain from intake of any food or beverage (water excepted) from one hour before the collection period. Chewing gum and intake of coffee were prohibited during this hour. The patient was instructed to minimise all facial movements particularly movements of the mouth. To begin
saliva collection the patient was asked to void the mouth of saliva by swallowing. The patient was then asked to lean slightly forward over the tube and funnel as demonstrated to the patient. The patient was instructed to keep his/her mouth slightly open and to allow saliva to drain into the funnel. The patient was also instructed to keep his/her eyes open. At the end of the 5-minute collection period, the patient was asked to collect any remaining saliva in his/her mouth and expectorate into the funnel and test tube. The saliva volume was measured using micropipettes (Finnpipette, Lab systems). All samples were collected between 08:30 and 11:30 hours to avoid diurnal variation. Samples were stored in minus 80 degrees Celsius freezer until time for analysis.

4.2.3 Total Protein Quantification

The total protein content for the samples was assayed using a commercial kit (Bicinchoninic Acid Protein (BCA) Assay Kit, Sigma-Aldrich, Inc) according to the manufacturers instructions. In brief, samples were diluted by a factor of x10. 25µl of the standards, negative control and samples were applied in duplicates wells of a 96 well plate, then 200µl of the BCA working solution added to the wells followed by a 30 minutes incubation at 37°C. The absorbance of the developed colour was measured using a colorimetric reader (Dynex Revelation® 4.24) at 470nm. The standard curve was plotted using Numbers® spreadsheet software for Apple Macintosh.

4.2.4 ELISAs

ELISA tests for HNP 1-3, LL-37 and Protein S100A8 were run. ELISAs were run using commercial kits for HNP 1-3 and LL-37 (Hycult biotechnology b.v. The Netherlands) while Circulex S100A8 (Human) ELISA Kit (Cyclex Co., Ltd. Terasawaoka, Japan) was used for quantifying protein S100A8 according to the manufacturer’s instructions. In brief, samples and ELISA reagents were thawed at room temperature. Both the test and control samples for saliva were diluted at an optimized dilution factor of x5,000 for HNP 1-3, x5 for LL-37 and x100 for S100A8 (with repeat of the high concentration samples at dilution factor of x250). Colorimetric reading was done using Dynex Revelation® 4.24 at 450nm.
All samples were run in duplicate with an internal control of standards and standard curves plotted using Numbers® spread sheet software (Apple Inc).

4.2.5 Data analysis

4.2.5.1 Case definitions

As outlined earlier, the subjects were recruited based on their DM status. Cases were classified a priori as gingivitis/mild periodontitis (control group), moderate periodontitis or severe periodontitis (disease group) based on number of sites with pre-determined threshold clinical attachment loss (CAL). The criteria for case definitions are summarized in Table 4-1.

<table>
<thead>
<tr>
<th>Diagnostic Group</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gingivitis/Mild disease</td>
<td>Less than 2 sites with CAL ≥ 4mm &amp; BOP</td>
</tr>
<tr>
<td>Moderate disease</td>
<td>2 or more sites CAL ≥ 4mm but &lt; 6mm</td>
</tr>
<tr>
<td>Severe disease</td>
<td>2 or more sites CAL ≥ 6mm</td>
</tr>
</tbody>
</table>

Table 4-1: Table showing the criteria used for case definition (CAL = clinical attachment loss; BOP = bleeding on probing).

4.2.5.2 Data analysis

Statistical analysis was carried out using Prism 5 for Mac OS X (GraphPad Software, Inc). Initially, cut off points of salivary HNP 1-3 and LL-37 concentrations determined from the previous experiments were tested to determine their ability to distinguish between periodontal diagnoses. One-way analysis of variance (ANOVA) with Bonferroni’s post test correction was then used to analyse the difference in expression of the different analytes in the various periodontal conditions. Where only two variables were compared, two-tailed t-test was performed. Statistical significance was set at p ≤ 0.05. Scatter plots were then generated with standard deviation (SD) indicated. ROC constructs were created for those markers showing a significant difference in expression and cut off points were set for optimal sensitivity and specificity. Data were also searched for any correlations between the different analytes and especially with total protein content to rule out the effect of global changes in protein content on the concentrations of the potential biomarkers. Finally, data
for HNP 1-3 and S100A8 were normalized to specific cut-off points and then combined by addition. The new values were then used to generate new ROC curves that determined the specificity and sensitivity of the combined markers.
4.3 Results

4.3.1 Demographic and clinical summary of sample population

After thawing the samples obtained from Glasgow, twenty of the samples tubes did not contain any saliva. 280 samples were thus available for the experiments. Another 11 had no corresponding clinical data or had labeling problems such as double labeling and therefore their true identity could not be determined and were omitted from the analysis. Effectively 269 samples (Table 4-2) were available for the experiments and were analyzed by ELISA for HNP 1-3, LL-37 & Protein S100A8 and total protein content. There was no significant difference between the average ages of the subjects in the different diabetic groups (1 way ANOVA; p = 0.3). CAL and PPD were statistically significantly greater in the poorly controlled diabetics than among the healthy group (1-Way ANOVA with Bonferroni’s post test correction; p = 0.01 and 0.03 respectively). Mean BOP and plaque scores were statistically significantly greater in the poorly controlled diabetic group than among the non-diabetics and well controlled diabetics (1-Way ANOVA with Bonferroni’s post test correction; p < 0.0001 for both BOP and Plaque score). Mean HbA1c levels were statistically significantly different among all groups by 1-Way ANOVA with Bonferroni’s post test correction (p < 0.0001). Table 4-3 shows the summary of the demographic and clinical parameters. Only 29 out of the 269 patients had severe periodontal disease. The majority (178) had gingivitis/mild periodontitis. The proportion of those with moderate to severe periodontal disease was highest in those with poorly controlled diabetes (41%) as compared to 28% in non-diabetics.
Table 4-2: Table showing the distribution of the cohort according to diabetic status and periodontal diagnosis according to the case definition. Numbers in parenthesis represent the proportion in percentage of periodontal diagnosis in each diabetic status followed by proportion of diabetic status in each periodontal severity group.

<table>
<thead>
<tr>
<th>Diabetic Status</th>
<th>Gingivitis/Mild Periodontitis</th>
<th>Moderate Periodontitis</th>
<th>Severe periodontitis</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non Diabetic</td>
<td>70(72.39)</td>
<td>19(20.31)</td>
<td>8(8.28)</td>
<td>97(100.36)</td>
</tr>
<tr>
<td>Well Controlled DM</td>
<td>22(79.12)</td>
<td>3(11.5)</td>
<td>3(11.10)</td>
<td>27(100.10)</td>
</tr>
<tr>
<td>Poorly controlled DM</td>
<td>86(60.48)</td>
<td>40(28.65)</td>
<td>18(13.62)</td>
<td>132(100.54)</td>
</tr>
<tr>
<td>N</td>
<td>178(66.100)</td>
<td>62(23.100)</td>
<td>29(11.100)</td>
<td>269(100.100)</td>
</tr>
</tbody>
</table>

Table 4-3: A table showing summary demographic and basic periodontal clinical parameters according to different diabetic status (M/F = male/female ratio; BOP = Bleeding on probing; PPD = Probing pocket depth; CAL = Clinical attachment loss; HbA1c = glycosylated haemoglobin). P-values (*) are based on one-way ANOVA with Bonferroni’s post test correction.

<table>
<thead>
<tr>
<th>Diabetic Status</th>
<th>Mean age</th>
<th>M/F</th>
<th>BOP mean</th>
<th>Mean PPD</th>
<th>Mean CAL</th>
<th>Mean Plaque score</th>
<th>Mean HbA1c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non diabetics</td>
<td>37.58</td>
<td>0.59</td>
<td>0.40</td>
<td><strong>2.19</strong></td>
<td><strong>1.57</strong></td>
<td><strong>0.63</strong></td>
<td><strong>5.20</strong></td>
</tr>
<tr>
<td>Well Controlled DM</td>
<td>36.85</td>
<td>0.87</td>
<td>0.40</td>
<td><strong>2.21</strong></td>
<td><strong>1.76</strong></td>
<td><strong>0.63</strong></td>
<td><strong>7.08</strong></td>
</tr>
<tr>
<td>Poorly controlled DM</td>
<td>35.68</td>
<td>1.12</td>
<td>0.50</td>
<td><strong>2.33</strong></td>
<td><strong>1.84</strong></td>
<td><strong>0.72</strong></td>
<td><strong>9.05</strong></td>
</tr>
<tr>
<td>Total Average</td>
<td>36.48</td>
<td>0.87</td>
<td>0.46</td>
<td><strong>2.27</strong></td>
<td><strong>1.73</strong></td>
<td><strong>0.68</strong></td>
<td></td>
</tr>
</tbody>
</table>

4.3.2 The application of previously determined HNP 1-3 and LL-37 cut-off points to diagnose periodontal disease among the new cohort.

Firstly, we aimed to apply our predetermined cut-offs for salivary concentrations of HNP 1-3 and LL-37 to evaluate their ability to diagnose periodontal disease in the new independent cohort. These predetermined cut off points including their sensitivity and specificity are indicated in Table 4-4 below. It was apparent that the concentrations in this cohort were completely different from the initial cohort described in Chapter 3 from which we had derived the cut-off points. For instance, the mean salivary HNP 1-3 level in the Gingivitis/mild periodontitis group in this cohort was 1.74 µg/ml which is almost ten times (0.22 µg/ml) that in the initial cohort from which the cut-offs were derived. A similar discrepancy
was noted in the salivary LL-37 concentrations in which the initial cohort mean concentrations were 2.93 ng/ml as compared to 86.3 ng/ml in the gingivitis cases of the Glasgow cohort. This problem persisted even among the diseased cases (Figure 3-5 versus Figure 4-5 and Figure 3-9 versus Figure 4-10). This meant that these cut-off points could not be applied in this new cohort.

<table>
<thead>
<tr>
<th></th>
<th>Cut-off</th>
<th>Specificity (%)</th>
<th>Sensitivity (%)</th>
<th>Area under curve (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNP 1-3</td>
<td>&gt; 0.294 µg/ml</td>
<td>83</td>
<td>64</td>
<td>75</td>
<td>0.007</td>
</tr>
<tr>
<td>LL-37</td>
<td>&gt; 4.39 ng/ml</td>
<td>82</td>
<td>79</td>
<td>81</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Table 4-4: Table showing the cut-off point concentrations of salivary HNP 1-3 that differentiate gingivitis from chronic periodontitis and LL-37 that differentiate gingivitis from periodontitis.

4.3.3 Influence of diabetic status on antimicrobial peptide levels.

One-way ANOVA was carried out on the AMPs levels in different diabetic status to determine any differences regardless of the periodontal status. The data showed no significant difference in salivary HNP 1-3, LL-37, S100A8 and total protein concentrations between the different diabetic statuses (Figure 4-1, Figure 4-2, Figure 4-3 and Figure 4-4).
Figure 4-1: Scatter plot for salivary concentrations of HNP 1-3 in different diabetic status regardless of periodontal diagnosis showing no significant difference by one way ANOVA with Bonferroni’s correction.

Figure 4-2: Scatter plot for salivary concentrations of LL-37 in different diabetic status regardless of periodontal diagnosis showing no significant difference by one way ANOVA with Bonferroni’s correction.
Figure 4-3: Scatter plot for salivary concentrations of S100A8 in different diabetic status regardless of periodontal diagnosis showing no significant difference by one way ANOVA with Bonferroni’s correction.

Figure 4-4: Scatter plot for salivary concentrations of total protein in different diabetic status regardless of periodontal diagnosis showing no significant difference by one way ANOVA with Bonferroni’s correction.
4.3.4 Salivary antimicrobial peptide levels based on periodontal diagnosis

4.3.4.1 HNP 1-3

Considering all the samples regardless of their diabetic status, there was increased concentration of salivary HNP 1-3 with increase of periodontal disease severity with a significant difference between gingivitis/mild periodontitis cases (2.46 µg/ml, SD ± 4.10 µg/ml) versus severe periodontitis (7.12 µg/ml, SD ± 9.73 µg/ml) (ANOVA with Bonferroni’s multiple comparison test; p = 0.0003). Severe periodontitis had higher levels of HNP 1-3 than moderate periodontitis (4.44 µg/ml, SD ± 8.34 µg/ml) though this difference did not reach statistical significance (Figure 4-5).

Figure 4-5: Scatter graph showing differential expression of salivary HNP 1-3 in different periodontal disease severity regardless of the diabetic status. P values are based on 1-way ANOVA with Bonferroni’s post test correction. *** refers to < 0.0001
Figure 4-6: Scatter plots for mean salivary HNP 1-3 concentration in non-diabetics according to periodontal status. Means are significantly different by 1-way ANOVA (p = 0.03) but not with Bonferroni’s post test correction.

Figure 4-7: Scatter plots for mean salivary HNP 1-3 concentration in well controlled diabetics according to periodontal status. P- values are based on 1-way ANOVA with Bonferroni’s post test correction. * refers to p < 0.05
In the non-diabetic group, mean salivary HNP 1-3 concentrations were higher in severe periodontitis (5.74 µg/ml; SD ± 9.04 µg/ml; n = 8) followed by moderate (4.66 µg/ml; SD ± 9.59 µg/ml; n = 19) and least in the gingivitis/mild periodontitis group (1.74 µg/ml; SD ± 2.94 µg/ml; n = 70). However this difference was only statistically significant by 1-way ANOVA without Bonferroni’s correction (p = 0.03). Among the well-controlled diabetics mean salivary HNP 1-3 concentrations was significantly higher by 1-way ANOVA with Bonferroni’s post-test correction (p = 0.03) in moderate periodontitis (12.15 µg/ml; SD ± 13.24 µg/ml; n = 3) than in the gingivitis/mild periodontitis group (3.07 µg/ml; SD ± 4.10 µg/ml; n = 22). In the poorly controlled diabetics, mean salivary HNP 1-3 concentrations were significantly higher (p = 0.005 by 1-way ANOVA with Bonferroni’s correction) in the severe periodontitis group (8.49 µg/ml; SD ± 10.70 µg/ml; n = 18) than in both the gingivitis/mild periodontitis group (2.90 µg/ml; SD ± 4.83 µg/ml, n = 84) and the moderate periodontitis group (3.72 µg/ml; SD ± 7.16 µg/ml; n = 38) Figure 4-8.

ROC curves were constructed for all the cases in which salivary HNP 1-3 concentrations were significantly differentially expressed in different disease stages.
severities. Without considering the diabetic status, HNP 1-3 was able to differentiate gingivitis/mild periodontitis from severe periodontitis with a specificity of 75.6% and sensitivity of 55.2% at a cut-off salivary HNP 1-3 concentration of >2.9 µg/ml (p = 0.02) (Figure 4-9). A similar ROC curve was obtained when data were analyzed in the poorly controlled diabetics. However, among the non-diabetics and well-controlled diabetics, the ROC curves did not produce statistically significant area under the curve.

![ROC Curve](image)

Figure 4-9: A representative ROC construct for salivary HNP 1-3 ability to distinguish between gingivitis/mild periodontitis from severe periodontitis in all cases regardless of their diabetic status.

### 4.3.4.2 LL-37

The overall concentration of salivary LL-37 among all patients did not differ between different diabetic status and periodontal status (Figure 4-10 to Figure 4-13). However, there was a trend towards increased LL-37 with increase of periodontal severity, which did not reach statistical significance. In non-diabetics, mean salivary LL-37 concentration was 86.3 ng/ml, SD ± 125.2 ng/ml in gingivitis/mild periodontitis group, 82.8 ng/ml, SD ± 81.5 ng/ml in moderate group and 110.8 ng/ml ± 136.1 ng/ml in the severe periodontitis group. In the well-controlled diabetic group, mean levels of salivary LL-37 were 55.9 ng/ml, SD ± 70.9 ng/ml in gingivitis/mild periodontitis group, 66.8 ng/ml, SD ± 16.8
ng/ml in moderate group and 78.8 ng/ml ± 23.4 ng/ml in the severe periodontitis group. In the poorly controlled diabetic group, the levels were 77.4 ng/ml, SD ± 85.2 ng/ml in gingivitis/mild periodontitis group, 90.0 ng/ml, SD ± 104.0 ng/ml in moderate group and 121.4ng/ml ± 82.3 ng/ml in the severe periodontitis group. ROC curves did not yield any significant cut-offs.

Figure 4-10: Scatter graph showing no statistically significant differential expression of salivary LL-37 by 1-way ANOVA with Bonferroni’s post-test correction in different periodontal status regardless of the diabetic status.
CHAPTER 4 – VALIDATION IN AN INDEPENDENT COHORT

Figure 4-11: Scatter plot showing no statistically significant differential expression of salivary LL-37 by 1-way ANOVA with Bonferroni’s post-test correction in the non-diabetics.

Figure 4-12: Scatter plot showing no statistically significant differential expression of salivary LL-37 by 1-way ANOVA with Bonferroni’s post-test correction in the well controlled diabetics.
Figure 4-13: Scatter plot showing no statistically significant differential expression of salivary LL-37 by 1-way ANOVA with Bonferroni’s post-test correction in the poorly controlled diabetics.

4.3.4.3 S100A8

Without considering the diabetic status, there was a tendency to increased salivary concentration of S100A8 with increase in periodontal disease severity. The concentrations in the mild group were 124.4 ng/ml; SD ± 126.9 ng/ml; n = 165, in moderate group 154.1 ng/ml; SD ± 146.9 ng/ml; n = 55 and severe group 178.1 ng/ml; SD ± 180.1 ng/ml; n = 25. This difference did not however reach statistical significance (one-way ANOVA p = 0.1) (Figure 4-14). However, when the subjects were classified in their respective diabetic groups, salivary S100A8 was able to differentiate between severe periodontitis (264.3 ng/ml; SD ± 225.6 ng/ml; n = 7) from both gingivitis/mild periodontitis (108.5 ng/ml; SD ± 107.2 ng/ml; n = 66) and the moderate periodontitis group (116.2 ng/ml; SD ± 90.5 ng/ml) in the non-diabetic group (p = 0.005 by one-way ANOVA with Bonferroni’s correction) (Figure 4-15).
Figure 4-14: Scatter graph showing no statistically significant differential expression of salivary S100A8 by 1-way ANOVA with Bonferroni’s post-test correction in different periodontal status regardless of the diabetic status.

Figure 4-15: Scatter plots for mean salivary S100A8 concentration in non-diabetics according to periodontal status. P-values are based on 1-way ANOVA with Bonferroni’s post test correction.
CHAPTER 4 – VALIDATION IN AN INDEPENDENT COHORT

Figure 4-16: Scatter plots for mean salivary S100A8 concentration in well-controlled diabetics according to periodontal status. P-values are based on 1-way ANOVA with Bonferroni’s post test correction showing no statistically significant different.

Figure 4-17: Scatter plots for mean salivary S100A8 concentration in poorly controlled diabetics according to periodontal status. P-values are based on 1-way ANOVA with Bonferroni’s post-test correction showing no statistically significant different.
The ROC curves produced an area under the curve of 0.69 (p = 0.099) for a cut-off salivary S100A8 of >255.4 ng/ml with a sensitivity of 57.14% and a specificity of 81.26% (Figure 4-18).

4.3.4.4 Total protein

Without considering diabetic status, the mean protein content in saliva was significantly increased in moderate periodontitis (1.71 mg/ml; SD ± 1.28 mg/ml; n = 52) compared to gingivitis/mild periodontitis (1.27 mg/ml; SD ± 0.84 mg/ml; n = 165) (one-way ANOVA with Bonferroni’s comparison test, p = 0.01). The amount of salivary protein in the severe periodontitis group (1.64 mg/ml; SD ± 1.21 mg/ml; n = 28) was also greater compared to the gingivitis/mild periodontitis group but this difference did not reach statistical significance. However, when the cases were grouped in their respective diabetic status, the statistical significance of the differential concentration of total salivary proteins seen between the different periodontal statuses was lost (Figure 4-19 to Figure 4-22).
CHAPTER 4 – VALIDATION IN AN INDEPENDENT COHORT

Figure 4-19: Salivary protein content in different periodontal severity showing statistically significant higher salivary concentration of total proteins in the moderate periodontitis group than the mild/healthy group based on one-way ANOVA with Bonferroni’s comparison test, p = 0.01. The cases are not grouped according to diabetic status.

Figure 4-20: Salivary concentrations of total protein in different periodontal disease severities among the non-diabetic group showing no statistical difference.
4.3.5 Correlation between AMPs and total protein concentrations in saliva.

In order to discover whether the expression of AMPs were dependent on the total protein or on each other, and whether this relationship was dependent on periodontal disease severity, correlation coefficient analysis was done for all
analytes. In all cases regardless of diabetic or periodontal status, there was a very weak but statistically significant positive correlation between all the analytes (Table 4-5). Analysis in the different periodontal disease severities was also carried out. In the gingivitis/ mild periodontitis group, there was a very weak, but statistically significant positive correlation between salivary HNP 1-3, LL-37, S100A8 and total protein concentrations. In moderate periodontitis, the weak correlation between the AMPs persisted but not in relation to the total protein concentration. In severe periodontitis, all the correlations disappeared except that between HNP 1-3 and LL-37 (Table 4-6). A representative correlation graph is shown in Figure 4-23 for HNP 1-3 and S100A8 clearly demonstrating the loss of correlation in severe disease.

<table>
<thead>
<tr>
<th></th>
<th>HNP 1-3</th>
<th>LL-37</th>
<th>S100A8</th>
<th>Total Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNP 1-3</td>
<td></td>
<td></td>
<td>p &lt; 0.0001</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>LL-37</td>
<td>R = 0.39</td>
<td></td>
<td>p = 0.001</td>
<td>p = 0.02</td>
</tr>
<tr>
<td>S100A8</td>
<td>R = 0.35</td>
<td>R = 0.21</td>
<td></td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>Total Protein</td>
<td>R = 0.28</td>
<td>R = 0.15</td>
<td>R = 0.36</td>
<td></td>
</tr>
</tbody>
</table>

Table 4-5: Matrix tables for the correlation between salivary HNP 1-3, LL-37, S100A8 and total protein concentration in saliva showing very weak but significant correlation in all pairs of analytes (data are from all cases regardless of their periodontal or diabetic status). (R = correlation coefficient; ‘p’ is the p value).
### (A) Gingivitis/Mild Periodontitis

<table>
<thead>
<tr>
<th></th>
<th>HNP 1-3</th>
<th>LL-37</th>
<th>S100A8</th>
<th>Total Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNP 1-3</td>
<td></td>
<td>p = &lt; 0.004</td>
<td>p = &lt; 0.0001</td>
<td>p = &lt; 0.0001</td>
</tr>
<tr>
<td>LL-37</td>
<td>R = 0.22</td>
<td></td>
<td>p = 0.01</td>
<td>p = 0.05</td>
</tr>
<tr>
<td>S100A8</td>
<td>R = 0.46</td>
<td>R = 0.20</td>
<td></td>
<td>p = &lt; 0.0001</td>
</tr>
<tr>
<td>Total Protein</td>
<td>R = 0.38</td>
<td>R = 0.15</td>
<td>R = 0.56</td>
<td></td>
</tr>
</tbody>
</table>

### (B) Moderate Periodontitis

<table>
<thead>
<tr>
<th></th>
<th>HNP 1-3</th>
<th>LL-37</th>
<th>S100A8</th>
<th>Total Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNP 1-3</td>
<td></td>
<td>p = &lt; 0.0001</td>
<td>p = &lt; 0.0001</td>
<td>p = 0.36</td>
</tr>
<tr>
<td>LL-37</td>
<td>R = 0.63</td>
<td></td>
<td>p = 0.005</td>
<td>p = 0.31</td>
</tr>
<tr>
<td>S100A8</td>
<td>R = 0.58</td>
<td>R = 0.39</td>
<td></td>
<td>p = 0.08</td>
</tr>
<tr>
<td>Total Protein</td>
<td>R = 0.13</td>
<td>R = 0.14</td>
<td>R = 0.25</td>
<td></td>
</tr>
</tbody>
</table>

### (C) Severe Periodontitis

<table>
<thead>
<tr>
<th></th>
<th>HNP 1-3</th>
<th>LL-37</th>
<th>S100A8</th>
<th>Total Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNP 1-3</td>
<td></td>
<td>p = 0.001</td>
<td>p = 0.68</td>
<td>p = 0.15</td>
</tr>
<tr>
<td>LL-37</td>
<td>R = 0.59</td>
<td></td>
<td>p = 0.79</td>
<td>p = 0.38</td>
</tr>
<tr>
<td>S100A8</td>
<td>R = - 0.09</td>
<td>R = 0.06</td>
<td></td>
<td>p = 0.46</td>
</tr>
<tr>
<td>Total Protein</td>
<td>R = 0.28</td>
<td>R = 0.18</td>
<td>R = - 0.16</td>
<td></td>
</tr>
</tbody>
</table>

Table 4-6: Matrix tables for the correlation between salivary HNP 1-3, LL-37, S100A8 and total protein concentration in saliva showing (A) very weak but significant correlation in all pairs of analytes among the gingivitis/mild periodontitis group (B) weak but significant correlation between the AMPs but not with total protein concentration among the moderate periodontitis group and (C) weak correlation between HNP 1-3 and LL-37 and no significant correlation between all the other analytes. (R = correlation coefficient; ‘p’ is the p value.)
Figure 4-23: Correlations coefficient between salivary HNP 1-3 and S100A8 in different periodontal disease severities regardless of the diabetic status. There was weak but statistically positive correlation in the mild and moderate periodontitis groups that was lost in the severe periodontitis group.
4.3.6 Testing the diagnostic utility of combined marker

The purpose of a diagnostic biomarker is to differentiate between disease from health or severity of the disease as well as one disease from another. Its utility is measured on its degree of specificity and sensitivity (discussed in 4.1 above) of detecting the disease. There is always a compromise between these two parameters so that as one value increases, the other decreases. It is up to the user to determine a cut off point with the right trade off between sensitivity and specificity of a test that suits the need for the test. For instance, a test with high sensitivity would be desirable for conditions that lead to high morbidity or death whereas high specificity tests would be useful for conditions whose treatment may cause disability or are very costly ensuring that no unnecessary treatments are provided to the patients.

Each test or biomarker would characteristically have a combination of specificity and sensitivity at each cut-off value of measurement. Recent studies have reported combining the sensitivity and specificity for multiple periodontal biomarkers resulting with a more predictive ROC curve than the use of the individual markers (Gursoy et al., 2010; Gursoy et al., 2011; Kinney et al., 2011; Ramseier et al., 2009). However, the methods used in these papers are complex or not detailed making them difficult to reproduce. In this thesis, a new method is described to combine at least two biomarkers (but the model could also be used for combining more than 2 marker analytes).

The method involves normalizing the values of sample analyte as a proportion (%) of a chosen cut-off value Figure 4-24. These cut-off values can be chosen as the points on which an arbitrarily straight line crosses the x and y-axes in a correlation table for the two markers (Figure 4-25). By normalizing the for each analyte, the line drawn between x and y-axes is described by the formula $x + y = 100$. 
The normalized values for the two markers in each sample are then summed up to make the new combined values that are then handled as a ‘new’ test to plot a ROC construct. This new ROC curve is then used to determine the positive and negative predictive values for the combined markers (Figure 4-26). The results here show an AUC of 65% (p= 0.01), sensitivity of 52% and specificity of 83.5% at a cut off of >188.0%. Results for a different arbitrary cut-off of 500ng/ml for S100A8 and 5µg/ml for HNP 1-3 gave an ROC with AUC of 66% (p=0.006), sensitivity of 50.0% and specificity of 83.1% at a cut off of >104.1% (Figure 4-27 to Figure 4-29).

Another approach to choose the threshold for normalizing the initial marker concentrations were explored by constructing an ROC curve for each marker using the actual measured values and determining the cut-off values at the point of optimal combination of sensitivity and specificity (Figure 4-9 and Figure 4-18). In this case the thresholds were 2.9 µg/ml for HNP 1-3 and 255.4 ng/ml for S100A8. The data were then normalized as before and ROC curve generated using the new values. The ROC curve produced a combined...
sensitivity of 52.0% and specificity of 83.5% with AUC of 65% (p = 0.015) at a combined cut-off point of >97.1% (Figure 4-30 to Figure 4-32). One way ANOVA with Bonferroni’s multiple corrections demonstrated that combining the markers using these thresholds from independent ROC curves was able to differentiate severe periodontitis from both moderate and gingivitis/mild periodontitis (p = 0.0007).

When these data were compared with individual cut-offs for each marker at specificity of around 83%, there was improvement in the sensitivity of the combined test from 25.5% for S100A8 and 48.3% for HNP 1-3 to 52% if the specificity was held constant at around 83%.

Figure 4-25: Representative graph showing an arbitrary choice of x and y-intercepts to be used as the cut-offs for normalizing the two markers (HNP 1-3 on the x-axis and S100A8 on the y-axis). In this case, 12.5µg/ml for HNP 1-3 and 550ng/ml for S100A8 were chosen.
Figure 4-26: ROC curve for the combined normalised values for HNP 1-3 and S100A8 using the cut-offs 12.5µg/ml for HNP 1-3 and 550ng/ml for S100A8.

Figure 4-27: Representative graph showing an arbitrary choice of x and y-intercepts to be used as the cut-offs for normalizing the two markers (HNP 1-3 on the x-axis and S100A8 on the y-axis). In this case, 5µg/ml for HNP 1-3 and 500ng/ml for S100A8 were chosen.
Figure 4-28: Scatter plot of summed value of normalised concentration of HNP 1-3 and S100A8 (normalization cut offs are 5µg/ml and 500ng/ml respectively) showing statistically significant higher mean values in severe periodontitis. P = 0.0003 by 1-way ANOVA with Bonferroni’s correction.

Figure 4-29: ROC curve for the combined normalised values for HNP 1-3 and S100A8 using the cut-offs 5µg/ml for HNP 1-3 and 500ng/ml for S100A8.
Figure 4-30: Representative graph showing choice of x and y-intercepts to be used as the thresholds for normalizing the two markers (HNP 1-3 on the x-axis and S100A8 on the y-axis) based on optimized cut-off points in the individual marker ROC curves (Figure 4-9 and Figure 4-18). In this case, 2.9μg/ml for HNP 1-3 and 255.4ng/ml for S100A8 were chosen.

Figure 4-31: Scatter plot of summed value of normalised concentration of HNP 1-3 and S100A8 (normalization cut offs are 2.9 μg/ml and 255.4ng/ml respectively) showing statistically significant higher mean values in severe periodontitis as compare to the other two. P = 0.0007 by 1-way ANOVA with Bonferroni’s correction.
Figure 4-32: ROC curve for the combined normalised values for HNP 1-3 and S100A8 using the cut-offs 2.9 µg/ml for HNP 1-3 and 255.4 ng/ml for S100A8.
4.4 Discussion

For a biomarker to be established as beneficial, it must be able to retain its utility in the presence of such potential confounders as DM. One of the main objectives in this chapter was to test the validity of previously determined diagnostic cut-off points for HNP 1-3 and LL-37 to correctly diagnose periodontal disease in this new independent cohort. The striking observation was that the levels of previously tested analytes (HNP 1-3 and LL-37) were much more abundant by almost a factor of ten in the saliva samples of the new cohort as compared to the initial samples. This meant that the predetermined diagnostic cut-off points for both HNP 1-3 and LL-37 could not be tested in these new samples. The main reason for this seemed to be in the method of saliva sample collection. The initial samples were collected using Salivettes whereas those for the current cohort were collected by means of spitting. The literature suggests that collection method may affect the retrieval of some specific analytes from saliva (Kozaki et al., 2009; Shirtcliff et al., 2001). At the time of writing this thesis, there was no available data regarding the recovery of these antimicrobial peptides in saliva collected using different methods.

Therefore we had to set new cut-off points for all the analytes. Data were also analyzed to determine whether analytes were able to distinguish between periodontal disease severities as they had previously distinguished between periodontitis from gingivitis in the initial cohort. The data from these experiments displayed variations in the measured amounts of analytes, which may be due to inter-individual characteristics such as genetic determinants of innate immunity. The results showed that HNP 1-3 was still capable of discriminating between periodontitis and gingivitis as well as between different severities of periodontitis regardless of the diabetic status. It also maintained the pattern of increased expression in the more destructive disease as in the initial cohort and other studies (Lundy et al., 2000; 2001; Lundy et al., 2005; Puklo et al., 2008). S100A8 was increased with increase of periodontal disease severity though this did not reach statistical significance except in the non-DM group. Generally, all the biomarkers were not as strongly discriminatory as the original data discussed in Chapter 3. Probably this is because of the
continuously variable disease severities in this cohort and the relatively low overall levels of disease.

The correlation of the biomarkers with each other was very weak and actually disappeared with severity of periodontal disease. For instance, in the gingivitis/mild periodontitis group, there was a very weak, but statistically significant positive correlation between salivary HNP 1-3, LL-37, S100A8 and total protein concentrations. Showing that at baseline, the AMPs may act as a surrogate marker for PMNs, the effector cells. However, this correlation is lost with severity of disease. This may suggest that in periodontal disease these AMPs are expressed independent of each other or total protein as well as possible PMN numbers. It has been recently reported that HNP 1-3 and LL-37 in GCF of periodontitis cases, unlike in healthy cases, was not correlated to the levels of myeloperoxidase, a surrogate marker of PMN numbers (Puklo et al., 2008). The authors of this article hypothesized that there were other sources of these AMPs. Whether there are other sources for these salivary AMPs or they are differentially expressed by the PMNs during inflammation, the fact that they are not correlated especially in severe disease strengthens their potential as periodontal biomarkers and not mere indicators of inflammation. It is possible that multiple pathogenic factors involved in tissue destruction in periodontitis may have an effect on production of these AMPs (Takeuchi et al., 2011). These data suggest that the measurements of these biomarkers are not merely surrogate measurements of neutrophil numbers and raise the possibility that their regulated expression has a causal relationship with disease severity. In summary, it can be concluded that the poor correlation of these factors with each other suggests that they are not purely surrogate markers of neutrophil numbers, reported to be the major source of these peptides. Also the increasing loss of any correlation with disease severity suggests the possibility that these biomarkers may actually indicate relevant pathogenic mechanisms; furthermore it raises the possibility that the expression of these AMPs in periodontal disease is dysregulated in some patients. It is acknowledged that the numbers of diseased cases were low and may have introduced statistical artifact(s). It is known that periodontal disease progresses in bursts of activity accompanied by periods of quiescence (Socransky et al., 1984). The cross-
sectional nature of sampling from disease cases may reflect different levels of
disease activity resulting in differences in AMP levels. At the time of carrying
out this study, it was not possible to determine the activity state of the
periodontal lesion, as there was no established test.

At the time of writing this thesis, there were no published data showing the
correlation between the markers HNP 1-3 and protein S100A8 in subjects
suffering from periodontitis. Moreover, this is the first time to the best of our
knowledge that the utility of the two proteomic markers combined together have
been evaluated to differentiate different periodontitis severities. Unlike in the
recent literature which used regression analysis or cumulative risk scores to
determine the usefulness of the combined markers (Gursoy et al., 2010; Gursoy
et al., 2011; Kinney et al., 2011; Ramseier et al., 2009), a simpler method is
described above that relies on summing the proportions of normalised values of
the markers. Using this method, it was possible to demonstrate some
improvement in the sensitivity of the combined test from 25.5% for S100A8 and
48.3% for HNP 1-3 to 52% if the specificity was held constant at around 83%. It
seems that the gain in sensitivity was more in the marker that had a weaker
sensitivity in the first place. The overall improvement in sensitivity may not be
dramatic probably due to the continuous nature of the data, and the relatively
few severe periodontitis cases in this study.

This study did not reveal any significant difference in the mean concentration of
any of the AMPs nor total protein content in different diabetic status. This
corroborates another study that found no difference in total protein in saliva of
diabetics and non-diabetics (Belazi et al., 1998). However in contrast, some
recent reports have demonstrated up-regulation of HNP 1, 2 and 4, S100A9
(which commonly occurs as a dimer with S100A8) and salivary total protein
associated with diabetes (Cabras et al., 2010; Lopez et al., 2003). However,
these study populations comprised of cases made up of diabetic children who
may have a different mechanism of pathogenesis of diabetes compared to
adults (Kuzuya et al., 2002). The prevalence of periodontitis is also usually low
among children as compared to adults (Califano, 2003; Kinane and Hodge,
2001; Pitts et al., 2011) making it difficult to draw comparisons between these
two studies with the current study. In addition failure to show any difference is
not the same statistically as demonstrating equivalence, although given the relatively high numbers of subjects tested suggests that any undetected differences are minor.

Several systemic conditions are associated with an increased risk for periodontitis. In some cases this appears to correlate with modulated expression of antimicrobial proteins (Cabras et al., 2010; Gorr and Abdolhosseini, 2011; Putsep et al., 2002). The patients in this cohort had been recruited based on their diabetic status and not periodontal status making it ideal to test the biomarkers’ diagnostic utility in the backdrop of a confounder.

Firstly, a case definition was determined based on the severity of periodontal disease as a factor of the number of sites with specific cut-off attachment loss as described earlier. It was noted that this cohort comprised of relatively young subjects (mean age - 36 years) who were non-smokers and even though two thirds of them were known type I diabetics, they did not have a lot of severe periodontal disease. Only 29 out of the 269 patients had severe periodontal disease. The majority (178) had gingivitis/mild periodontitis. The proportion of those with moderate to severe periodontal disease was highest in those with poorly controlled diabetes (41%) as compared to 28% in non-diabetics. This has been described in other studies, which reaffirms the observation that diabetes is a well-known risk factor for periodontal disease (De Silva et al., 2006; Graves et al., 2007; Preshaw et al., 2010).

There are potential linkages between the pathogenesis of DM and periodontal disease (Jones et al., 2007; Kiran et al., 2005; Mealey and Oates, 2006; Sima et al., 2010; Taylor et al., 1996; Taylor, 2001). DM is a metabolic disorder characterized by hyperglycemia as a result of defects in insulin secretion or action (James R, 1998; Kuzuya et al., 2002). It may cause polymorphonuclear leukocyte dysfunction, vascular changes, altered collagen and glycosaminoglycan synthesis, deregulation in cytokine production, and the formation of advanced glycation end products (AGEs), thereby modulating periodontal tissue destruction. AGEs can stimulate monocyte recruitment to the periodontal site, whereas the interaction with the receptors for AGEs (RAGEs) may increase the production of pro-inflammatory cytokines (Cabras et al., 2010;
CHAPTER 4 – VALIDATION IN AN INDEPENDENT COHORT

Costa et al., 2010; Duarte et al., 2007; Sima et al., 2010; Stabholz et al., 2010). One paper has reported increased HNP 1-3 in diabetics as compared to non-diabetics (Cabras et al., 2010).

4.4.3 Conclusion

Diabetes did not seem to affect the expression of the antimicrobial peptides HNP 1-3, LL-37 and S100A8 in the subjects studied. However, HNP 1-3 does show some utility in differentiating mild from both moderate and severe periodontal diseases. It was however not possible to apply previously determined cut-off points to diagnose the new samples. The reason may be differences in collection method wherein the initial samples for validation were collected using Salivettes® while the ones in this cohort were collected through spitting. It is probable that the Salivettes may be the culprits in reducing the amount of measurable analytes in the saliva collected using them. This will be investigated in the next Chapter 5. Combining the markers did add some value over the use of the single markers independently.
Chapter 5 - Translational Studies
5.1 Introduction

Saliva contains not only endogenous proteins and peptides, but also exogenous ones from oral microorganisms. Among these are many enzymes such as proteases, and collagenases (Wong, 2008). These enzymes become a challenge when it comes to the saliva collection and handling methods employed since intact proteins or peptides that are potential disease markers may get degraded during collection, transportation, storage or any other stage (Wong, 2008). In addition, the secretion and composition of saliva is under systemic influences and thus is affected by systemic factors such as circadian rhythms and systemic conditions like Sjögren’s syndrome. Some of the other factors that have been described as influencing saliva composition are temperature, exercise, chewing, drinking and stress (Dawes, 1972; 1975; Dawes and Kubieniec, 2004; Dodds et al., 1991). Most of these factors have been studied based on their effect on specific proteins of interest to the researchers carrying out the studies or on total protein or albumin content as the primary outcome.

There are very few data regarding the effect of saliva collection method, diurnal variations, use of protease inhibitors and storage temperature on the salivary levels of HNP 1-3, LL-37 and S100A8. In the previous chapter, it was described how the salivary concentration of the putative biomarkers HNP 1-3, LL-37 and the total protein content in saliva collected using Salivettes® were far lower than in the samples collected via spitting. It was postulated that the different saliva collection methods in the two studies was the main determinant for this observation. This necessitated further experiments to test the retrievability of the identified putative markers and investigate other factors that affect their salivary concentrations.
5.2 Aims

The aim of this group of experiments was to investigate the factors affecting measured salivary concentration of the putative periodontal markers namely HNP 1-3, LL-37 and S100A8.

5.2.1 Objectives

1. To investigate the effect of saliva collection method on the measured salivary concentration of HNP 1-3, LL-37.

2. To investigate the effect of diurnal variation on the measured salivary concentration of HNP 1-3, LL-37 and S100A8.

3. To investigate the effect of eating on the measured salivary concentration of HNP 1-3, LL-37 and S100A8.

4. To investigate the effect of saliva storage temperature on the measured salivary concentration of HNP 1-3 and LL-37.

5. To determine the usefulness of addition of a protease inhibitor into WS in order to preserve the measured salivary concentration of HNP 1-3 and LL-37.
5.3 Materials and Methods

5.3.1 Ethics approval
The protocol for this part of the study was reviewed and approved by the Outer North East London Research Ethics Committee (ref number 10/H0701/72). All participants signed an informed written consent form.

5.3.2 Patient Cohort and Sample Acquisition

Inclusion criteria
The patients had accepted to give informed consent, were between the ages 18 to 65 years and had either gingivitis, CP or AgP as defined before. For the healthy group, the subjects did not have obvious clinical signs of periodontal inflammation, no bleeding on probing in greater than 10% of the sites, and no greater than one site with > 4mm PPD. All participants had at least 20 natural teeth.

Exclusion criteria
Patients were excluded if they had periodontal treatment in the previous 6 months, were currently taking any antimicrobials or had diabetes mellitus or any other systemic condition known to affect the periodontal tissues.

5.3.3 Sampling and sample collection
Patients were approached in the periodontal consultant clinic at Guy’s Dental Hospital. In all experiments, whole saliva was collected using spitting method as described before. The samples were then centrifuged (3000 RPM for 15 minutes at 4°C) and aliquots of approximately 300µl stored immediately at minus 80°C. In 27 of the patients, an extra saliva sample was collected at the same sitting using Salivettes® (Sarstedt Ltd) in order to compare recovery of AMPs using the two methods. There was no wash out period between saliva collection. Several experiments were planned and the unique details of sample handling are described in each experiment below:
5.3.3.1 Effect of collection method on salivary AMPs recovery

Paired WS samples were collected from 27 chronic periodontitis patients by either spitting for 2 minutes first followed by use of Salivettes® for 2 minutes or vice versa. The sequence of saliva collection method in every individual was randomized following a random list (see Table 8-7 in appendix) generated by using Microsoft Excel® 2007. This was to control for the saliva stimulating effect of introducing the Salivette®. These data were also analyzed to see if the sequence changed the AMPs recovered in the samples collected using either of the methods. Saliva was collected between 09:30 and 11:00 hrs and between 13:30 and 15:00 hrs.

5.3.3.2 Effect of diurnal variation and eating on salivary AMPs concentration

Unstimulated whole saliva was collected among non-smokers who were periodontally and systemically healthy using the spitting method for 5 minutes. The subjects did not have obvious clinical signs of periodontal inflammation, no bleeding on probing in greater than 10 % of the sites, and no more than one pocket greater than 4mm. Samples were collected daily on five consecutive days (Monday to Friday) at 10:00 hours. On the first day, additional samples were collected at 12:00, 14:00 and 16:00 hours. Also on the first day, all subjects had a cheese sandwich with plain bottled water between 12:25 and 13:30 hours. Subjects were not allowed to eat or drink anything for one hour prior to sample collection at all other times.

5.3.3.3 Salivary AMP stability under different storage temperature and proteinase inhibitor.

WS samples were collected using the spitting method for 5 minutes in the consultant led periodontology clinic at Guy’s Dental Hospital, London. The samples were then centrifuged (3000 RPM, 15 minutes and 4°C) and the supernatant divided in to aliquots of approximately 300µl. One aliquot per patient was immediately frozen at minus 80°C. All the other aliquots were left at room temperature (RT) and one aliquot per original sample frozen at minus 80°C at the time points 6, 24, 48 and 72 hours post-collection.

At time zero a proteinase inhibitor (PI), Calbiochem® PI cocktail set 1 (a mixture of 5 protease inhibitors with broad specificity for the inhibition of various
proteases and esterases) was added to a further aliquot per patient. 1X cocktail contains 500µM AEBSF, 500µM EDTA, 1µM E-64, 1µM Leupeptin and 1µg/ml Aprotinin. The PI cocktail was reconstituted with 1 ml deionized water (DI) to yield a 100X stock solution. 3µl of the stock PI was added to 297µl of saliva supernatant and vortexed briefly to make a 1:100 sample (‘saliva + PI’). This sample was then left at room temperature (RT) for at least 72 hours with the other RT aliquots when it was also finally frozen with the other RT samples until time for analysis.

5.3.4 Total Protein Quantification
The total protein content for the samples was assayed using a commercial kit (Bicinchoninic Acid Protein (BCA) Assay Kit, Sigma-Aldrich, Inc) according to the manufacturers instructions. In brief, samples were diluted by a factor of x10. 25µl of the standards, negative control and samples were applied in duplicates on a 96 well plate, then 200µl of the BCA working solution added to the wells followed by a 30 minutes incubation at 37 degrees Celsius. The absorbance of the developed colour was measured using a colorimetric reader (Dynex Revelation® 4.24) at 470nm. The standard curve was plotted using Numbers® spread sheet software for Apple Macintosh.

5.3.5 ELISAs
ELISA tests for HNP 1-3 and LL-37 were run on all samples. In addition, S100A8 was also measured in the diurnal variation, eating and storage effect experiments. ELISAs were run using commercial kits for HNP 1-3 and LL-37 (Hycult biotechnology b.v. The Netherlands) while Circulex S100A8 (Human) ELISA Kit (Cyclex Co., Ltd. Terasawaoka, Japan) was used for quantifying protein S100A8 according to manufacturer’s instruction. Colorimetric reading was done using Dynex Revelation® 4.24 at 450nm. All samples were run in duplicates with an internal control of standards and standard curves plotted using Numbers® spread sheet software for Apple Macintosh.

5.3.6 Statistical analysis
Statistical analysis was carried out using Prism 5 for Mac OS X (GraphPad Software, Inc). One-way analysis of variance (ANOVA) with Bonferroni’s post
test correction was used to analyse the difference in concentrations of the different analytes in the various collection or storage conditions. Where only two variables were compared, two tailed t-test was performed with significance set at $p = 0.05$. Two-way analysis of variance (ANOVA) with Bonferroni’s post test correction was used to determine the contribution of the various parameters including collection method, duration of sample storage at room temperature and inter-individual variation on the total variance observed in the data. Linear regression analysis was carried out to determine the correlation between relevant data points. Data for the diurnal variation study were normalized to the individual mean values to determine the degree of change in concentrations at different time points assuming that the concentrations oscillated about the individual’s true mean value. Data for stability studies were normalized as a proportion to concentration at time zero.
5.4 Results

5.4.1 Effect of collection method on the recovery of AMPs in whole saliva.

HNP 1-3 and LL-37 ELISA were performed for all the pairs of samples in the 27 subjects as previously described. Mean subject age was 38 Years (22 – 63 years). The M/F ratio was 0.86. All subjects had moderate to severe periodontal disease with ≥ 2 sites with CAL > 5 mm.

5.4.1.1 HNP 1-3

Mean salivary concentration of HNP 1-3 was significantly higher (20.8 fold) in the samples obtained using the spitting method than in the Salivette® method (9.4µg/ml versus 0.5µg/ml respectively, t-test p = 0.0001) Figure 5-1 and Figure 5-2. However, the concentrations were weakly but significantly positively correlated in the two collection methods by linear regression analysis, R = 0.65, r² = 0.42, p = 0.0002 (Figure 5-3). Saliva collection method significantly accounted for 26.6% of the total variance observed based on two-way ANOVA (F = 19.92, DF = 1, n = 27, p = 0.0001). Inter-individual variation accounted for 38.7% of the total variance observed though this did not reach statistical significance, (F = 1.11, DF = 26, n = 27, p = 0.39).

5.4.1.2 LL-37

Mean salivary concentration of LL-37 was significantly higher (22.9 fold) in the samples obtained using the spitting method than in the Salivette® method (126.3ng/ml versus 5.1ng/ml respectively, t-test p = 0.0002) Figure 5-4 and Figure 5-5. However, the concentrations were weakly but significantly positively correlated in the two collection methods by linear regression analysis, R = 0.63, r² = 0.40, p = 0.0005 (Figure 5-6). Saliva collection method significantly accounted for 25.2% of the total variance using two-way ANOVA (F = 18.67, DF = 1, n = 27, p = 0.0002). Inter-individual variation accounted for 39.8% of the total variance observed though this did not reach statistical significance, (F = 1.14, DF = 26, n = 27, p = 0.39).
Figure 5-1: A graph of salivary HNP 1-3 concentration showing a 20.8 fold higher mean concentration in saliva collected by spitting as compared to Salivette®. T-test p = 0.0001.

Figure 5-2: A plot showing the relationship between salivary HNP 1-3 concentration in saliva collected by spitting as compared to Salivette®. The concentrations were higher in the spitting sample than in the Salivette sample for each paired sample.
Figure 5-3: Regression analysis between salivary concentration of HNP 1-3 collected using Salivette® versus spitting method showing a moderate but significant positive correlation between the two. \( r^2 = 0.42 \) (\( p = 0.0002\))

Figure 5-4: A graph of salivary LL-37 concentration showing a 22.9 fold higher mean concentration in saliva collected by spitting as compared to Salivette®. T-test \( p = 0.0002\).
Figure 5-5: A plot showing the relationship between salivary LL-37 concentration in saliva collected by spitting as compared to Salivette®. The concentrations were higher in the spitting sample than in the salivette sample for each paired sample.

Figure 5-6: Regression analysis between salivary concentration of LL-37 collected using salivette® versus spitting method showing a moderate but significant positive correlation between the two. $r^2 = 0.4$ ($p = 0.0005$)
To determine whether the sequence of sample collection (i.e. spitting first followed by Salivette® and vice versa) had an effect, the data were expressed as a ratio (Salivette®/spitting) and the ratios compared using unpaired t-test. There was no significant difference in the HNP 1-3 and LL-37 recovery regardless of the order of collection method used (Figure 5-7).

5.4.2 The effect of sample storage temperature and proteinase inhibitor on salivary AMPS.

The effect of storage temperature was tested in samples from 25 patients with moderate to severe periodontal disease as described before. After processing, each sample had been divided into several aliquots of 300µl that were frozen at minus 80°C after being left at RT for 0, 6, 24, 48 and 72 hours. Of the original 25 samples, 15 had one additional aliquot preserved with a proteinase inhibitor (PI) cocktail Calbiochem® PI cocktail set 1 (AEBSF, EDTA, E-64, Leupeptin and Aprotinin) before being frozen after 72 hours at RT. One sample did not have aliquots for 24, 48 and 72 hours due to sample depletion and so was excluded from the analysis. Effectively, there were 24 samples with all time points of which 14 had PI added.

5.4.2.1 The effect of sample storage temperature on salivary AMPS.

The mean salivary HNP 1-3 remained relatively stable although there was a tendency to decrease over time. This observation however was not statistically significant, 1-way ANOVA with Bonferroni’s post test comparisons, \( p \) value = 0.16 (Figure 5-8). Two-way ANOVA did not reveal any significant effect of duration of storage at room temperature and inter-individual variation to the total variation observed (Table 5-1). When the data were normalized as a proportion to the baseline ‘time zero’, there was no significant change in the mean levels except between sample ‘72Hrs+PI’ which had almost 50% reduction in HNP 1-3 levels as compared to all the other time points, one-way ANOVA, \( p < 0.0001 \) (Figure 5-9).
Figure 5-7: Scatter graphs for (A) HNP 1-3 and (B) LL-37 showing Salivette®/spitting ratios with no significant difference regardless of the order of sample collection method. T-test p value = 0.45 and 0.72 for HNP 1-3 and LL-37 respectively.
Figure 5-8: Scatter graph of salivary HNP 1-3 concentrations for 24 samples left at room temperature for 0, 6, 24, 48 and 72 hours. Notice one outlier at time zero which corresponds to sample ID number 2118 that had mucin residuals that probably affected HNP 1-3 measurement. There was no statistical difference in the concentrations at different time points (1-way ANOVA with Bonferroni’s post test comparisons, p value = 0.16)

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>% of total variation</th>
<th>F</th>
<th>DF</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage time at room temperature</td>
<td>5.00</td>
<td>1.69</td>
<td>4</td>
<td>0.1583</td>
</tr>
<tr>
<td>Inter-Individual variation</td>
<td>27.05</td>
<td>1.59</td>
<td>23</td>
<td>0.0625</td>
</tr>
</tbody>
</table>

Table 5-1: Two-way ANOVA with Bonferroni’s multiple comparison summary table for salivary HNP 1-3 showing no significant contribution by RT storage time and inter-individual variation to the total variation observed.
Mean salivary LL-37 generally decreased over time. This reduction was statistically significant from baseline and 6 hours compared to both 48 and 72 hours, 1-way ANOVA with Bonferroni’s post test comparisons, p value = 0.0001 (Figure 5-10). Two-way ANOVA to test for the effect of duration of storage at room temperature and inter-individual variation revealed statistically significant effect for both factors but more for the latter (Table 5-2). Normalized data as a proportion to the baseline ‘time zero’ showed a significant reduction in LL-37 at 72 hours with or without PI as compared to both ‘time zero’ and 6 hours, one-way ANOVA, \( p = 0.001 \) (Figure 5-11). Samples containing PI had almost half the amount of HNP 1-3 and LL-37 compared to the baseline and 6Hrs (Figure 5-9 and Figure 5-11).

Figure 5-9: Bar graph showing proportion change of salivary HNP 1-3 from baseline levels (Time zero). Error bars are SD. There is significant reduction in HNP 1-3 at ‘72Hrs+PI’ compared to all other time points (\( p \) value < 0.0001, one-way ANOVA with Bonferroni’s post test correction).
Figure 5.10: Scatter graph of salivary LL-37 concentrations for 24 samples left at room temperature for 0, 6, 24, 48 and 72 hours. There was a significant reduction in the concentration between time zero and both 48 and 72 hours, and also between 6 hours and both 48 and 72 hours, 1-way ANOVA with Bonferroni’s post test comparison, p value = 0.0001.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>% of total variation</th>
<th>F</th>
<th>DF</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage time at room temperature</td>
<td>3.08</td>
<td>6.46</td>
<td>4</td>
<td>0.0001</td>
</tr>
<tr>
<td>Inter-Individual variation</td>
<td>85.96</td>
<td>31.38</td>
<td>23</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

Table 5.2: Two-way ANOVA with Bonferroni’s multiple comparison summary table for salivary LL-37 showing significant effect of both RT storage time and inter-individual variation to the total variation observed.
5.4.3 The effect of diurnal variations on salivary AMP concentrations.

Seven systemically and periodontally healthy subjects were recruited for this study. Only one subject had BoP on one site and another with a PPD of 4 mm with no CAL loss. One other subject had recession of 2 mm on 4 sites caused by aggressive oral hygiene practices. Whole saliva samples were collected from the subjects for five consecutive days (i.e. Monday to Friday) and also at 4 times with 2-hour intervals between collections during the first day. Four subjects however did not give samples on the final day. Therefore only 3 samples were available for day 5 (Friday) due to lateness/absence or other factors. Day 5 samples were therefore excluded from the analysis to avoid introducing bias. The data for each analyte for the first 4 days were analyzed by linear regression to detect any correlations between the different days and within the different times in the first day. 1-way ANOVA was performed to establish any differences and 2-way ANOVA to determine the amount of effect time and day of the week as well as inter-individual variation had on the observed total variance. The mean levels with SD are presented in Table 5-7 and Table 5-8.
5.4.3.1 Total protein

There was no statistically significant difference in salivary total protein content at different time points even when the data were normalized as proportions to the intra-subject’s mean levels (1-way ANOVA with Bonferroni’s post test comparisons, p value = 0.144) Figure 5-12 and Figure 5-13. Two-way ANOVA analysis with Bonferroni’s multiple comparisons did not reveal any significant effect of both time or day of sample collection as well as inter-individual variation to the total variation observed in the total salivary protein concentration (Table 5-3).

![Figure 5-12: Scatter plot of salivary total protein concentration normalised to the mean value for the total samples collected. No significant difference in different time points (ANOVA with Bonferroni’s post test comparison, p value = 0.144)](image)

<table>
<thead>
<tr>
<th></th>
<th>Effect accounting for total variation observed</th>
<th>F</th>
<th>DF</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time and day of sample collection</td>
<td>12.71 %</td>
<td>1.17</td>
<td>6</td>
<td>0.35</td>
</tr>
<tr>
<td>Inter-individual variation</td>
<td>21.95%</td>
<td>2.02</td>
<td>6</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 5-3: Two-way ANOVA summary table for total salivary protein showing lack of significant effect of time and day of sample collection as well as inter-individual variation.
There was variation observed in the measured salivary HNP 1-3 concentration at different time points but this did not reach statistical significance Figure 5-15 and Table 5-7. Normalized data revealed intra-individual variation at different time points with a tendency to peak at 12:00 and 16:00 hours but again it was not statistically significant (Figure 5-14). Two-way ANOVA analysis with Bonferroni’s multiple comparisons revealed a significant moderate contribution of inter-individual variation to the total variance observed. Both time and day of sample collection did not contribute significantly to this observed variation (Table 5-4).
5.4.3.3 LL-37

There was variation observed in the absolute salivary LL-37 concentration at different time points but this did not reach statistical significance (Figure 5-17 and Table 5-7). Normalized data also revealed intra-individual variation at different time points that peaked at 12:00 and 16:00 hours but this was not statistically significant (Figure 5-16). Two-way ANOVA analysis with Bonferroni’s multiple comparisons revealed a moderate but statistically significant contribution of inter-individual variation to the total variance observed (2-way ANOVA, p value = <0.0001). Both time and day of sample collection did not contribute significantly to this observed variation (Table 5-5).
Figure 5-15: Bar graph showing variations in mean salivary HNP 1-3 concentrations at different time points. However, variation did not reach statistical significance using 1-way ANOVA with Bonferroni’s post test comparison.

Figure 5-16: Scatter plot of salivary LL-37 concentration normalised to the intra-subject mean value. No significant difference in different time points is noted (ANOVA with Bonferroni’s post test comparison, p value = 0.07).
<table>
<thead>
<tr>
<th>Effect accounting for total variation observed</th>
<th>F</th>
<th>DF</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time and day of sample collection</td>
<td>9.82 %</td>
<td>1.62</td>
<td>6</td>
</tr>
<tr>
<td>Inter-individual variation</td>
<td>53.74%</td>
<td>8.85</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 5-5: Two-way ANOVA (Bonferroni’s multiple comparisons) summary table for salivary LL-37 showing a moderate but significant contribution of inter-individual variation to the total variation observed. There was lack of significant effect of time and day of sample collection.

Figure 5-17: Bar graph showing variations in actual mean salivary LL-37 concentrations at different time points. However, variation did not reach statistical significance by 1-way ANOVA with Bonferroni’s post test comparison.

5.4.3.4 S100A8

There was variation observed in the absolute salivary S100A8 concentration at different time points with a statistically significant peak at 16:00 hours, 1-way ANOVA, p value = 0.001 (Figure 5-19). Normalized data confirmed this peak
was significantly greater than all the other time points within the day or between different days, one-way ANOVA, p value = 0.0001 (Figure 5-18). Both time and day of sample collection significantly accounted for 20.45% effect of the total variance observed (2-way ANOVA, p value = <0.001, F = 4.79, DF = 6) while inter-individual variation accounted for a moderately large and extremely statistically significant effect to the total variance observed (2-way ANOVA, effect = 53.95%, p value = <0.0001, F = 12.64, DF = 6) Table 5-6.

Figure 5-18: Scatter plot of salivary S100A8 concentration normalised to the intra-subject mean value. Notice significant peak at 16:00 hours compared to all other time points (ANOVA with Bonferroni’s post test comparison, p value = 0.0001)

<table>
<thead>
<tr>
<th>Effect accounting for total variation observed</th>
<th>F</th>
<th>DF</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time and day of sample collection</td>
<td>20.45 %</td>
<td>4.79</td>
<td>6</td>
</tr>
<tr>
<td>Inter-individual variation</td>
<td>53.95%</td>
<td>12.64</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 5-6: Two-way ANOVA with Bonferroni’s multiple comparison summary table for salivary S100A8 showing a moderate but significant contribution of inter-individual variation to the total variation observed as well as a significant effect of time and day of sample collection.
**Figure 5-19:** Bar graph showing variations in mean salivary S100A8 concentrations at different time points. Note the statistically significant peak at 16:00 hours compared to 10:00, 14:00 hours and days 2 and 3 (One-way ANOVA with Bonferroni’s post test comparison, p = 0.001).

### 5.4.4 The effect of eating on salivary AMP concentrations.

The above subjects were given a cheese sandwich and plain bottled water for their lunch between 12:30 and 13:00 hours. Total protein, HNP 1-3, LL-37 and S100A8 did not significantly vary before and after eating a cheese sandwich in the subjects studied (one-way ANOVA, p value > 0.05). However eating seemed to contribute slightly but significantly to the variation of salivary S100A8 concentrations observed before and after eating with an effect factor of 13.02% of the total variance (2-way ANOVA, F = 6.34, DF = 1, p value = 0.045). Interpersonal variation also contributed a 74.65% effect of the total variance for S100A8 observed between before and after eating (2-way ANOVA, F = 6.06, DF = 6, p value = 0.023).
### Table 5.7: Summary table showing the mean levels (SD) of different salivary analytes at different time points. There was no statistical difference between the different time points in total protein, HNP 1-3 and LL-37 concentrations. S100A8 was significantly higher at 16:00 hrs (day 1) than 10:00 hrs (days 1,2 and 4) and 14:00 hrs (day 1), ANOVA with Bonferroni’s post test comparison, p value = 0.001.

<table>
<thead>
<tr>
<th></th>
<th>10:00hrs (Day 1)</th>
<th>12:00hrs (Day 1)</th>
<th>14:00hrs (Day 1)</th>
<th>16:00hrs (Day 1)</th>
<th>10:00hrs (Day 2)</th>
<th>10:00hrs (Day 3)</th>
<th>10:00hrs (Day 4)</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Protein (mg/ml)</strong></td>
<td>1.05 (0.21)</td>
<td>1.18 (0.20)</td>
<td>1.06 (0.15)</td>
<td>1.37 (0.55)</td>
<td>1.23 (0.35)</td>
<td>1.37 (0.43)</td>
<td>1.27 (0.36)</td>
<td>1.22 (0.32)</td>
</tr>
<tr>
<td><strong>HNP 1-3 (µg/ml)</strong></td>
<td>0.94 (0.68)</td>
<td>1.51 (1.22)</td>
<td>0.68 (0.52)</td>
<td>1.27 (0.94)</td>
<td>1.13 (0.95)</td>
<td>1.40 (2.06)</td>
<td>0.89 (1.36)</td>
<td>1.12 (1.10)</td>
</tr>
<tr>
<td><strong>LL-37 (ng/ml)</strong></td>
<td>17.75 (10.01)</td>
<td>28.25 (20.14)</td>
<td>15.48 (9.94)</td>
<td>31.56 (23.12)</td>
<td>23.78 (19.61)</td>
<td>23.47 (22.07)</td>
<td>17.45 (17.14)</td>
<td>22.53 (17.43)</td>
</tr>
<tr>
<td><strong>S100A8 (ng/ml)</strong></td>
<td>0.52 (0.49)</td>
<td>0.76 (0.63)</td>
<td>0.38 (0.42)</td>
<td>1.10 (0.46)</td>
<td>0.45 (0.39)</td>
<td>0.60 (0.58)</td>
<td>0.45 (0.45)</td>
<td>0.61 (0.49)</td>
</tr>
</tbody>
</table>

*
**
<table>
<thead>
<tr>
<th></th>
<th>10:00 Hrs (Day 1)</th>
<th>12:00 Hrs (Day 1)</th>
<th>14:00 Hrs (Day 1)</th>
<th>16:00 Hrs (Day 1)</th>
<th>10:00 Hrs (Day 2)</th>
<th>10:00 Hrs (Day 3)</th>
<th>10:00 Hrs (Day 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total protein</td>
<td>89.64 (10.25)</td>
<td>94.21 (7.66)</td>
<td>90.16 (18.51)</td>
<td>112.4 (38.72)</td>
<td>92.03 (13.27)</td>
<td>121.1 (25.60)</td>
<td>100.5 (28.66)</td>
</tr>
<tr>
<td>HNP 1-3</td>
<td>82.58 (48.79)</td>
<td>135.9 (59.49)</td>
<td>87.21 (57.87)</td>
<td>133.2 (56.29)</td>
<td>94.78 (47.01)</td>
<td>100.7 (72.49)</td>
<td>65.62 (44.39)</td>
</tr>
<tr>
<td>LL-37</td>
<td>84.22 (15.95)</td>
<td>128.6 (59.68)</td>
<td>80.97 (34.30)</td>
<td>144.2 (41.26)</td>
<td>92.61 (64.11)</td>
<td>98.42 (48.47)</td>
<td>71.01 (26.21)</td>
</tr>
<tr>
<td>S100A8</td>
<td>81.39 (34.84)</td>
<td>114.5 (44.09)</td>
<td>55.27 (29.12)</td>
<td>232.4 (124.8)</td>
<td>78.08 (45.00)</td>
<td>76.36 (45.11)</td>
<td>62.02 (35.67)</td>
</tr>
</tbody>
</table>

Table 5-8: Summary table showing mean (SD) of the different analytes normalised as the percentage proportion of the individual mean over the 7 time points.
5.5 Discussion

Despite its success story, saliva still poses some major challenges in the search for biomarkers (Esser et al., 2008; Helmerhorst et al., 2010; Schipper et al., 2007c; Thomadaki et al., 2011). There are major issues that must be addressed before a salivary-based biomarker can reach clinical usefulness. These include controlling for the many variables that can affect the measurement of the potential biomarkers such as possible diurnal variations, change in composition affected by exercise, temperature, age, medication, stress, salivary gland diseases/conditions and salivary flow rate. (Dawes, 1972; 1993; Elishoov et al., 2008; Garde et al., 2009; Kariyawasam and Dawes, 2005); (Morris et al., 2002; Ng et al., 2003; Schipper et al., 2007b). Another factor that must be addressed is what type of saliva sample is to be collected e.g. glandular or whole saliva and what method of collection will be employed such as unstimulated versus stimulated and so on (Dabbs, 1991). Protein breakdown caused by microorganisms and proteases especially in WS samples poses a serious problem for quantitative measurements of proteins (Esser et al., 2008).

Every salivary constituent is unique in the way it responds to the factors mentioned above (Thomadaki et al., 2011). This demands that standard operating procedures must be established for the acquisition, processing and storage of saliva samples as well as the biomarker measurement for each biomarker. Some factors may be more critical for specific biomarkers than others. Failure to establish these standard operating procedures makes it difficult to use salivary biomarkers (Crouch, 2005; Henson and Wong, 2010).

In Chapter 4, it was reported that the average amount of salivary HNP 1-3 and LL-37 were extremely high in the Glasgow cohort as compared to the initial samples collected at Royal London Hospital making it impossible to compare results from the two cohorts. One of the obvious differences between the two samples besides being from different populations was the collection method. The Glasgow samples had been collected using the spitting method while those from the Royal London Hospital had been collected using Salivettes®. It was postulated that this could be the main cause for the discrepancy in the salivary
concentrations of the AMPs. An experiment was set to measure the impact on AMP recovery from whole saliva using the two methods in a new cohort.

The results confirmed that despite a slight positive correlation between the 2 collection methods, there was greater than 20-fold increase in detected levels of the AMPs using the spitting method in these subjects compared to the Salivettes®. Some samples collected using Salivettes® did not have any detectable levels of AMPs whereas the paired spitting sample had detectable measurable levels. This implies that the use of Salivettes® alone in the search for biomarkers may introduce errors of measurement by suppressing low concentration analytes.

The mechanism by which the proteins are retained by the Salivettes® is not clear. In fact some studies on testosterone have demonstrated the opposite with the testosterone levels increasing in samples collected by the use of Salivettes® compared to spitting. One study demonstrated that Salivettes® without saliva had detectable levels of testosterone (Dabbs, 1991). Salivettes® are made out of cotton. Some literature report that cotton based collection methods do adsorb the proteins in the cotton material and therefore caution should be used before using this method for saliva collection (Shirtcliff et al., 2001; Wong, 2008). Considering that AMPs are cationic (positively charged) molecules, they could easily be attracted by the negatively charged cotton fibers in the Salivette® through electrostatic means. The poor correlation between the two collection methods suggests that the Salivettes® are not reliable for collecting saliva samples for the analysis of these AMPs.

This study demonstrated that HNP 1-3 and LL-37 are highly stable in saliva at room temperature for up to 48 hours. This stability may partly be attributed to the fact that the samples had been centrifuged to remove the cellular component, which includes microorganisms known to produce proteolytic enzymes. However, at 72 hours it seems there was some loss of reactivity of these AMPs. This may be the result of renewed exponential growth of microorganisms in the saliva over time producing proteolytic byproducts. Paradoxically, the addition of the proteinase inhibitor seemed to reduce the levels of detectable AMPs. It may be that the PI masked specific epitopes in
the AMPs blocking them from being detected by the ELISA. The possibility that addition of PI into the samples diluted the AMPs would not have played any significant role as only 3µl were added into 297µl of sample. The PI was added only on one sample that was kept for up to 72 hours. It was assumed that PI would stabilize the samples further (Liu et al., 2010) and so there was no need to have samples with PI tested at time points before 72 hours. It is not possible with this data to detect at what time point the observed PI effect started between ‘time zero’ to 72 hours. Recent studies based on SELDI-TOF MS demonstrated the emergence of new peptides or increased peaks of existing proteins/peptides 3 to 4 hours after storage in room temperature. These peaks were thought to be degradation products of larger proteins (Esser et al., 2008; Schipper et al., 2007a). These studies also showed the use of proteinase inhibitors to be partly beneficial in stabilizing specific proteins and peptides meaning that not all proteins and peptides are subject to endogenous proteinases (Esser et al., 2008; Schipper et al., 2007a). In fact, Esser and group showed that a product with the MW of 3370 Da was stable in room temperature and suggested that this product can be used as a marker for sample integrity (Esser et al., 2008). Interestingly, this product’s MW corresponds to that of HNP-2 using the same platform they used i.e. SELDI-TOF MS. HNP-2 is thought to be a by-product of either HNP-1 or 3 by the loss of one amino acid since no mRNA has been identified for it. Whether or not HNP1 or 3 degrades into HNP-2, this would not affect the measurement of HNP 1-3 using the current ELISA, as the total detectable HNP1-3 epitopes would theoretically remain constant.

These results raise the possibility that saliva could be collected away from the clinical or laboratory setting and posted by mail for analysis of these AMPs 24 to 48 hours later without losing their utility. The slight intra-individual variation observed during the day would need to be confirmed in more studies comparing the same time points in multiple days. It would be recommended though to take samples at the same time to improve reproducibility.

The data also showed that there was a wide range of inter-individual variation as observed for other salivary analytes (Harmon et al., 2008; Larsen et al., 1999; Quintana et al., 2009). This casts doubt on the usefulness of pooling
samples from different subjects before analysis in search of salivary biomarkers employed in many studies (Hardt et al., 2005; Hu et al., 2008; Ryu et al., 2006; Streckfus et al., 2006) since the results may not reflect the true analyte levels in the individual. There was also a slight but non-significant degree of variation within the day for total protein, HNP 1-3 and LL-37. S100A8 showed a significantly distinct peak at 16:00 hours. Whereas this suggests that sample collection time for testing HNP 1-3 and LL-37 are not critical, it would be advisable to collect samples for testing S100A8 at 16:00 hours. This further lends to the idea that these AMPs are independently regulated despite being products of neutrophils. However these experiments on variation with time should be regarded as only preliminary. Firstly there were relatively few samples included and thus might not give sufficient power to measure true variations. Secondly, failure to show a difference is not the same statistically as showing equivalence, which requires many more samples. Thirdly as the highest levels of all AMPs were at 1600h it would have been useful to include a 1600h time-point in day 2 and day 3 too.

Although these data did not demonstrate any effect of eating on the salivary concentration of AMPs, the results are not conclusive due to lack of controls. An appropriate control would include collecting the samples in alternate days at the same time point (14:00hrs) in the same subjects with and without eating before sample collection.

5.5.1 Conclusion

Saliva collection method is critical in determining the integrity of the sample for AMPs analysis. Spitting method is recommended as opposed to use of Salivettes®. Since HNP 1-3 and LL-37 seemed stable in RT up to 48 hours, it may be possible to collect WS samples in field settings and post them via mail to the clinicians or laboratory for AMPs assay. On a public-health point of view this may be useful to screen and even monitor diseases of public health importance at the community level. However, it would be important to set the collection time carefully so that the measured levels of AMPs do not vary due to diurnal variations. The lack of correlation between the salivary concentrations of different AMPs with each other suggests that each of the AMPs is under a different regulatory pathway. It follows then that their expression in saliva may
not be a surrogate marker of inflammation per se making them potentially useful as biomarkers for periodontal disease so long as precautions are made to maintain sample integrity.
Chapter 6 - Genomic Studies on DEFA1A3
6.1 Introduction

Previous chapters have described the finding that anti-microbial peptides, most particularly HNP 1–3 and LL-37, are potentially useful biomarkers of periodontitis, because of their relatively high specificity for the destructive forms of periodontal disease. The major source of HNPs and LL-37 both in GCF and in saliva are neutrophils, which raises the possibility that measuring AMPs is largely a surrogate measure of neutrophil numbers. However in Chapter 4 it was found that there is a rather weak correlation between HNP and LL-37 levels, which would suggest that their salivary concentrations are not simply a reflection of neutrophil numbers present. In addition, in Chapter 3 it was observed that LL-37 levels (unlike HNP levels) were elevated specifically in AgP, again emphasizing the weak association between these analytes. These observations not only suggest the utility of these analytes as biomarkers, but raise the possibility that AMPs may have important mechanistic roles in determining disease susceptibility.

The pathophysiology of periodontal disease is multifaceted with both host and environmental factors coming into play to produce the array of pathognomonic features observed in these conditions. The innate immune system is one of the main inherent factors that determine the balance between health and disease. For instance, it has been shown that people with Morbus Kostmann syndrome suffered from severe periodontitis. This is a genetic disorder, characterized by a recessive mutation, severe congenital neutropenia and severe periodontitis (Dale et al., 2000; Putsep et al., 2002). Even with treatment using recombinant granulocyte colony-stimulating factor (G-CSF) to restore neutrophil counts, the patients still suffer from severe periodontitis. However, those treated using bone marrow transplant (BMT) have normal periodontal health. The only difference found between the G-CSF and the neutrophils in the BMT patients is the very low levels (less than 30% of the normal) of LL-37 in G-CSF (Aprikyan and Dale, 2001; Dale et al., 2000; Putsep et al., 2002).

Genetic variations such as copy number variations (CNVs) play a role in the expression of protein and peptides in different populations. Such variations involving immune genes could contribute to the differences in innate immunity
between individuals as well as influence predisposition and susceptibility to diseases. CNVs is not a frequent feature of the innate immune genome as they appear to be restricted to a small subset of this genome (Linzmeier and Ganz, 2006). Some of the well characterised ones include genetic CNVs of the DEFA1A3 gene, that encode HNP 1-3 peptides (Aldred et al., 2005; Hollox et al., 2003) and have been reported to be associated with differential production of the specific peptides. For instance the amount of CNVs for DEFA1A3 is proportional to the levels of HNP 1-3 (Aldred et al., 2005; Ballana et al., 2007; Linzmeier and Ganz, 2005; Nuytten et al., 2009; Redon et al., 2006).

In a case–control study in a Chinese population of more than 200 patients, a high copy number for DEFA1A3 gene was shown to be associated with more than two-fold increase of risk of severe sepsis (Chen et al., 2010) as well as Danish Crohn’s Disease (Jespersgaard et al., 2011). However in a different study, DEFA1A3 copy number variation could not be linked to the carrier rate for Staphylococcus aureus (S.aureus) (van Belkum et al., 2007).

The DEFA1A3 gene is located in chromosome 8 at loci 8p23.1 close to other defensin genes, Figure 6-1 (Linzmeier et al., 1993; Linzmeier et al., 1999; Sparkes et al., 1989). The DEFA1 gene encodes for the peptide HNP 1 while DEFA3 gene encodes for HNP 3 and these two differ only in one nucleotide, C3400A which is also responsible for the N-terminal amino acid difference between the two peptides (Ballana et al., 2007). The gene for HNP 2 has never been described and it is thought that HNP 2 is a proteolytic product of one or both of HNP 1 or 3 by losing the N-terminal amino acid that differentiates the two (Ganz, 2003).

![Figure 6-1: Schemata of Chromosome 8 showing the arrangement of the DEFA1A3 and its neighbouring defensin genes. Adapted from (Nuytten et al., 2009)](image-url)
In Chapter 3 and Chapter 4, it was noted that there was inter-individual variation in the expression of salivary HNP 1-3 as well as an increased concentration of the peptides in subjects suffering from periodontal disease. One of the explanations for this observation may have been due to CNVs relating to DEFA1A3. This chapter sets to investigate the role of DEFA1A3 CNVs on periodontal disease.
6.2 Aims

The aim of the experiments in this chapter is to investigate the hypothesis that CNV for DEFA1A3 is associated with periodontal disease.

6.2.1 Objectives

1. To determine the profile of DEFA1A3 CNVs among periodontitis and control subjects

2. To investigate any relationship between DEFA1A3 CNVs and periodontal status.
6.3 Methodology

6.3.1 Sample description

Banked blood samples for 263 subjects who had been recruited from Guys Hospital, London were used to extract the genomic DNA. Ethics was granted by the NHS and Guy’s research ethics committee REF 01/04/14. Subjects had signed an informed written consent to take part. The mean age was 40 years (Median 39 years, range 8 to 76 years) with 176 Caucasians and the rest non-Caucasians. 132 were non-smokers, 56 previous smokers and 75 current smokers. Average number of teeth per subject was 27 (median = 28, range 11 to 32). Only 10 subjects had < 20 teeth and all of them were above 34 years of age except one who was 20 years with 15 teeth.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>16</td>
</tr>
<tr>
<td>Gingivitis</td>
<td>49</td>
</tr>
<tr>
<td>CP</td>
<td>161</td>
</tr>
<tr>
<td>AgP</td>
<td>37</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>263</strong></td>
</tr>
</tbody>
</table>

Table 6-1: Table showing the frequency of the different periodontal status.

6.3.2 Genomic DNA extraction and quality control

Genomic DNA was extracted from the blood using a commercial kit, DNeasy Blood & Tissue Kit (QIAGEN Gmbh) according to the manufacturer’s instructions and flow through as shown in Figure 6-2. 200µl of DNA was eluted and its quality and concentration determined using Qubit® fluorometer (QIAGEN Gmbh) according to the manufacturer’s instruction.

Samples were then stored at minus 20 degrees Celsius until time for analysis by Real Time PCR (RT-PCR).
6.3.3 qPCR

Custom made primers and probes were designed by blasting the sequence for DEFA1A3 gene using Pubmed Blast tool, Figure 6-3 and Table 6-4. In this TaqMan copy number quantitation reaction, purified genomic DNA was combined with the TaqMan Copy Number Assay (contains two primers and a FAMTM dye- labeled MGB probe to detect the genomic DNA target sequence) the TaqMan Copy Number Reference Assay (contains two primers and a VIC® dye-labeled TAMRATM probe to detect the genomic DNA reference sequence) and the Genotyping Master Mix (contains AmpliTaq Gold® DNA Polymerase, UP (Ultra Pure) and dNTPs) required for the PCR reactions in the proportions shown in Table 6-2. The Reference assay detects the Ribonuclease P RNA component H1 (H1RNA) gene (RPPH1) on chromosome 14, cytoband 14q11.2.

263 samples of gDNA were used in this assay. All reactions per sample were carried out in a single well making a total PCR reaction mix volume of 20µl. Each 20µl assay contained 20 ng of gDNA so that the final concentration was 1ng/µl of gDNA in each well.
Figure 6-2: Flow through schemata for DNA extraction using DNeasy extraction kit (Qiagen Group)

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Vol (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaqMan Genotyping Master Mix</td>
<td>10</td>
</tr>
<tr>
<td>TaqMan CNV Assay x20 stock (Assay ID-DEFA1_CCWR1LZ, Lot No. P101113-000G01)</td>
<td>1</td>
</tr>
<tr>
<td>TaqMan Reference Assay x20 stock</td>
<td>1</td>
</tr>
<tr>
<td>Nuclease Free Water</td>
<td>4</td>
</tr>
<tr>
<td>gDNA Sample (5 ng/μl)</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total volume</strong></td>
<td><strong>20</strong></td>
</tr>
</tbody>
</table>

Table 6-2: This table shows the composition of the RT-PCR reaction mix (volume in microliters).

Individual samples were run in triplicates including the ‘No Template Controls’ (NTC). NTC samples did not contain gDNA and were used as internal controls for background fluorescence. Thermal-cycling conditions (ABI 7900HT Fast
System) were set as shown in Table 6-3. Absolute quantification standard real-time data were captured using the SDS 2.3 software. Manual Ct threshold was set at 0.2 and the autobaseline set on. The data were then exported to CopyCaller™ Software for analysis to determine the copy numbers for DEFA1A3.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hold</td>
<td>95°C</td>
<td>10 min</td>
</tr>
<tr>
<td>Cycle (40 Cycles)</td>
<td>95°C</td>
<td>15 sec</td>
</tr>
<tr>
<td></td>
<td>60°C</td>
<td>60 sec</td>
</tr>
</tbody>
</table>

Table 6-3: RT-PCR run parameters for DEFA1A3 CNV assay using ABI 7900HT Fast system on Absolute quantification standard setting.
<table>
<thead>
<tr>
<th>Forward Primer Sequence</th>
<th>Reverse Primer Sequence</th>
<th>Reporter 1 Sequence</th>
<th>Reporter 1 Quencher</th>
<th>Reporter 1 Dye</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGAGTAACATAGCAGGGTCTTTT</td>
<td>CTGTTCCTCCCAGTTTCATGCTGA</td>
<td>CCAACCTAGCACAGACATTAA</td>
<td>NFQ</td>
<td>FAM</td>
</tr>
</tbody>
</table>

Table 6-4: Table showing Custom TaqMan® Copy Number Assays details used to perform the RT-PCR experiment. The primers are shown in 5’ to 3’ direction except for the reverse primer which is in 3’ to 5’. The assay ID is DEFA1_CCWR1LZ from Applied Biosystems.
**Figure 6-3:** NCBI Primer blast for DEFA1 PCR template input >ref|NT_023736.17|:c6827602 (6825171-6827602) Homo sapiens chromosome 8 genomic contig, GRCh37 reference primary assembly (Range 6827602-6825171) Green = Forward primer, Pink = Probe and Blue = Reverse primer all in 5' to 3'.
6.4 Results

Two hundred and sixty three samples were analysed by RT-PCR using a custom made CNV Assay (Applied Biosystems). All samples achieved excellent amplification curves as can be seen in a representative plot (Figure 6-7). As expected the NTC samples did not amplify. Analysis using the CopyCaller® software generated the CNV scores for all the samples.

The general pattern of distribution of the copy numbers (CN) for this cohort was a bell-shaped histogram (Figure 6-4) with the most frequent CN being 6 (31.9%) followed by 5 (20.5%) and 7 (14.1%) respectively. The least CN was 3 while the maximum was 13. Since the average DEFA1A3 CNV copy number in the United Kingdom and probably in the rest of Europe’s general population is six (Aldred et al., 2005; Nuytten et al., 2009), a threshold was set at ≤ 6 or > 6 CN in order to calculate the proportion of CN within each periodontal diagnostic status above or below this threshold. The results show that for all disease cases, the proportion for CN ≤ 6 was greater than for CN > 6. The pattern was the opposite in the healthy cases, (Figure 6-5 and Figure 6-6) although this difference did not reach statistical significance using Chi square distribution test (p = 0.20, df = 3). A post-hoc exploratory analysis using Fischer’s exact test was done comparing our data with a UK (Aldred et al., 2005) and a Belgium (Nuytten et al., 2009) reference population DEFA1A3 data (n = 111 and 334 control subjects in the 2 populations respectively). These two cohorts were of unknown periodontal status. Each of the periodontal disease categories in our study was statistically significantly different from that of the independent UK reference population (p = 0.03, 0.004 and 0.03 for CP, AgP and G respectively) but not against the Belgian population. Interestingly, the healthy group was not statistically different from the control population (p value = 1.0).
<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>CN ≤ 6</th>
<th>CN &gt; 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gingivitis</td>
<td>32</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>97</td>
<td></td>
<td>64</td>
</tr>
<tr>
<td>AgP</td>
<td>27</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Healthy</td>
<td>7</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>163</strong></td>
<td></td>
<td><strong>100</strong></td>
</tr>
<tr>
<td><em>(UK Population)</em></td>
<td></td>
<td><em>(51)</em></td>
<td><em>(60)</em></td>
</tr>
<tr>
<td><em>(Belgium Population)</em></td>
<td></td>
<td><em>(105)</em></td>
<td><em>(95)</em></td>
</tr>
</tbody>
</table>

Table 6-5: Total counts of subjects with DEFA1A3 copy numbers ≤ 6 or > 6 in different periodontal diagnostic categories with no significant difference between the groups by Chi square distribution (p = 0.20, df = 3). The UK and Belgian populations’ DEFA1A3 CNV data are shown for comparison, adapted from (Aldred et al., 2005; Nuytten et al., 2009) respectively.

Figure 6-4: Frequency histogram showing the frequency (%) of the various copies of the DEFA1A3 CNV. The most frequent CN was 6 (31.9%).
Figure 6-5: Bar graph showing the proportion of cases with copy numbers ≤6 versus > 6 within each periodontal diagnosis.

Figure 6-6: Bar graph of the absolute numbers of cases with copy numbers ≤6 versus > 6 within each periodontal diagnosis showing statistically significant difference between the UK reference population and the disease cases but not the periodontally healthy group.
Figure 6-7: A representative print-out for the RT-PCR amplification plot showing excellent amplification plots for samples and no amplification for NTC samples.
6.5 Discussion

The data show that the cohort studied had DEFA1A3 CNV pattern that was similar to the general UK and Belgium populations with the most frequent CN being 6 (Aldred et al., 2005; Nuytten et al., 2009). When our data were analyzed against the UK reference population of unknown periodontal status and dichotomized on the basis of DEFA1A3 copy numbers ≤ 6 or > 6, they revealed that the disease cases had proportionately more DEFA1A3 gene ≤ 6 copies than those > 6 copies and this difference was statistically significant. The healthy group had the opposite pattern with proportionately more DEFA1A3 CNV > 6 copies compared to those ≤ 6 copies, which is similar to the pattern in the reference population consisting of controls (Aldred et al., 2005).

These are very interesting findings as they suggest that in the UK population, possessing ≤ 6 copy numbers of DEFA1A3 may confer risk for periodontal disease. On the other hand the finding that in the healthy group the proportion of CN > 6 were significantly more than those ≤ 6, a pattern that is similar to that in the UK general population, may be suggestive that possessing > 6 CN may confer protective characteristics against periodontal disease. However, the data should be treated as preliminary as the numbers involved are not adequate to make generalizations for genetic studies. A larger sample size in each diagnostic category may provide better representation and comparison between the different periodontal statuses. Further, this relationship could not be demonstrated when the data were analysed against the Belgian population. It is hard to decipher why the discrepancy between the two populations was observed but the fact that the numbers in the Belgian population where higher than the UK reference population may mean that any chance findings of differences might have been eliminated.

Another striking observation was that gingivitis seemed to cluster with the disease cases and not with the healthy. Typically gingivitis is thought of as a possible control population for genetic studies. Our current data seems to suggest either that the results are too small to interpret or that risk of gingivitis itself is also partly genetically determined.
It has previously been shown that the DEFA1A3 copy numbers are roughly proportional to the amount of HNP 1-3 in neutrophils (Linzmeier and Ganz, 2005). Neutrophils are the first line of defence against periodontal pathogens in the periodontal lesion (Schroeder et al., 1975). Their increased numbers partly explain the increased amount of HNP 1-3 in both GCF and saliva. However the correlation between the neutrophils and HNP 1-3 levels in GCF is poor and cannot account for the total increase in HNP 1-3 (Puklo et al., 2008). Unfortunately it was not possible to compare the DEFA1A3 CNV data with the salivary levels of the peptides HNP 1-3, as saliva samples had not been collected when the blood samples were acquired. Notwithstanding, we have already established in previous chapters that salivary HNP 1-3 are generally elevated in periodontal disease as compared to healthy cases.

This is somewhat paradoxical in that it would be expected that the CNV copies would be proportional to the salivary HNP 1-3 levels in which case, there would have been proportionately more > 6 copies than ≤ 6 copies in the disease cases as compared to the healthy cases. The reason for this is not clear although some reports suggest that DEFA1A3 copy numbers do not necessarily correlate to the amount of peptide produced (Aldred et al., 2005). Others have also demonstrated that the amount of copy number can have an opposite effect to the protein expression (Guryev et al., 2008) including the observation that genes in CNV regions show reduced expression. It has also been observed that CNV may exert control over other genes in their vicinity that may include their own copies (Henrichsen et al., 2009; Stranger et al., 2007) therefore dampening or amplifying their expression.

A study aimed to determine a possible correlation between caries prevalence in 149 middle school children and salivary concentrations of AMPs found that high HNP1-3 levels were protective against caries (Tao et al., 2005). There was no correlation between salivary HNP 1-3 levels and salivary levels of the caries-causing Streptococcus mutans (S. mutans), which suggests that the HNP 1-3 is capable of suppressing caries development, by some other means. Based on the finding of that study, our findings seem to support the thinking that relatively high baseline levels of HNP 1-3 as may be expected in subjects with > 6
DEFA1A3 gene copies may be protective against initiation of periodontal disease.

On the other hand, it could be that the inflammatory process in periodontal disease may lead to neutrophil dysfunction in patients with fewer copies of DEFA1A3 that may dysregulate the release of HNP 1–3. Differences in innate immune responses, including antimicrobial peptides, play a role in determining the microbial balance of oral tissues and the subsequent outcome of bacterial colonization. Consequently, population studies have identified population subgroups with different levels of susceptibility to periodontal disease (Gorr, 2009). One can imagine a scenario in which the baseline levels of salivary HNP 1-3 or at least in the gingival crevice in subjects at risk (with ≤ 6 DEFA1A3 copies) lead to the initiation and establishment of periodontal disease by allowing selective microbial colonization. The inflammatory process may then set off a cascade of feedback signals leading to the overexpression of HNP 1-3 that occurs in periodontal disease as demonstrated in previous chapters (Vankeerberghen et al., 2005). This argument is graphically represented in the model below (Figure 6-8).
6.5.1 Conclusion

In conclusion, globally the data suggest that DEFA1A3 CN distribution in this cohort mirrors that in the general population in two European countries with the commonest copy number being 6. Although we acknowledge that the sample size especially for the healthy group was relatively small, the study suggests that ≤ 6 DEFA1A3 gene copy numbers in a UK population may confer risk for periodontal disease as demonstrated by the highly significant difference from the reference control population for the CNV. It would be interesting to find out in similar but larger studies in each periodontal diagnostic category what the true relationships between DEFA1A3 copy numbers, salivary HNP 1-3 levels and the corresponding periodontal status is. It would be important in a longitudinal study to relate baseline HNP 1-3 levels with DEFA1A3 gene CN in subject without periodontal disease, and follow up the subjects for a long period in order to determine their role in periodontal disease development.
Chapter 7 - Final Discussion, Future Studies and Conclusions
7.1 Final Discussion

The principle findings from these studies as described in previous chapters suggest the diagnostic utility of antimicrobial peptides particularly HNP 1-3 and LL-37 and to some extent, S100A8. Firstly we identified these biomarkers by LC-MS/MS from a list of putative differential peaks screened using SELDI-TOF MS, although not all of them could be identified. Then we were able to test at length those we were able to identify using ELISA and obtained very good sensitivity and specificity in both GCF and saliva. For instance, using HNP 1-3 we were able to differentiate between gingivitis and periodontitis while LL-37 was able to differentiate between AgP and both CP and gingivitis.

We then wanted to validate the findings in an independent sample cohort, which had been selected, not based on its periodontal status but upon a potential confounder of periodontal disease, namely diabetic status (Kinane, 1999). Again the basic findings appeared to support our initial findings with both HNP 1-3 and LL-37 being able to differentiate between gingivitis and periodontitis. It is acknowledged though that the initial cut-off points of the biomarkers could not be validated in this new cohort owing to the different sample collection methods. It was interesting to find that diabetic status did not affect the diagnostic ability of these novel biomarkers to differentiate between periodontal statuses. Generally our results compare relatively well with other biomarker studies such as MMP-8 (Gursoy et al., 2010; Gursoy et al., 2011; Hernandez et al., 2010; Mantyla et al., 2003; Mantyla et al., 2006; Sorsa et al., 2010; Sorsa et al., 2011; Yakob et al., 2011).

In additional translational studies, we were able to demonstrate that these salivary markers were relatively stable for up to 48 hours at room temperature, especially the HNP 1-3 and LL-37. This is important especially if the biomarkers are to be used in epidemiologic studies in field conditions, or at home for monitoring periodontal disease. The samples would have to be stable from the time of collection to the time they are assayed in the analyzing laboratory.

One of the unique things about this work is the biomarker discovery process approach that we employed to define new biomarkers. This has not been
thoroughly explored previously for periodontal disease biomarker discovery. We have shown the potential of using different platforms including SELDI-TOF MS to screen for peaks corresponding to specific biomarkers that can differentiate between disease statuses. We have further demonstrated that it is possible to identify what some of these flagged markers would be using other Proteomic techniques such as LC-MS/MS. Those markers have then been validated in the same samples using other techniques such as ELISA confirming their ability to differentiate between different periodontal statuses as described in the MS studies.

The correlation between samples collected using different methods revealed the discrepancies in the measurable concentrations in the same subjects. This meant it was impossible to validate the cut off points set in initial validation studies. This raises the important question of standardizing sample collection in future studies.

It is notable that all the biomarkers we tested are regarded as antimicrobial peptides, and furthermore it is expected that the main source of these AMPs in saliva are likely to be from neutrophils, particularly those present in the gingival crevicular fluid. This raised the possibility that the concentrations of the AMPs were merely surrogate markers of neutrophil numbers. However it was found that the correlation between the AMPs was very weak in individual subjects. These findings suggest that salivary AMP concentrations are not merely surrogate markers for PMN numbers but that their production may be regulated by a number of complex regulatory mechanisms, as has been previously observed in another study (Puklo et al., 2008). Some of the regulatory mechanisms appear to be at the genome level, especially related to the copy numbers of DEFA1A3 gene for the HNP 1-3 peptide. Some studies have previously suggested that the level of HNP 1–3 peptide may be proportionate to the copy number of the DEFA1A3 gene present (Aldred et al., 2005). Although our study found ≤ 6 DEFA1A3 copy numbers related with periodontal disease in a UK population, in which there tends to be higher levels of salivary HNP 1-3, it is possible that the baseline levels of the peptide are usually lower than in these susceptible subjects before the disease starts (Tao et al., 2005). High levels of the gene seemed to be associated with periodontal health in this population. It
was however not possible to directly demonstrate the relationship between DEFA1A3 and salivary HNP 1-3 in this study. Also, it was not possible to demonstrate the same effect when the data were compared with a Belgian population.

We recognize that in the current studies, there were some limitations. For instance it still remains challenging to identify all the markers flagged using the SELDI-TOF MS. However, with advancement of technology such as the advent of the Lucid Proteomics SystemTM (Bio-Rad and Bruker Daltonic) this seems feasible in the not far future as techniques to couple SELDI-TOF MS with other quadruple mass spectrometry continue to emerge. We also acknowledge that the sample size was small in the “before and after treatment AgP” study and that for measuring LL-37 in gingivitis, CP and AgP. Further, for genomic studies, more samples would be necessary. We also observed that relatively, there was not much severe disease in the independent validation cohort selected based on their diabetic status. Notwithstanding these challenges, the results were reasonably reproducible in different experiments and sample cohort settings. This does however mean that the results need to be treated as preliminary.

7.1.1 Future studies

Firstly, we would recognize that not all the flagged potential markers during the biomarker screening using SELDI-TOF MS were identified. It would be useful to further identify these using other fractionation techniques such as molecular weight cut-off columns and iTRAQ tags. The newly identified markers would then be validated as we have done in the studies above. The data should be analysed to determine the usefulness of combining multiple markers in an effort to increase the overall diagnostic utility of the test.

In the current studies, we have not been able to cover all aspects of the relationship between salivary and GCF AMP concentration with periodontal status. For instance, the cut-off points for salivary AMP concentration that differentiate between periodontal disease and health or gingivitis are not validated yet. Secondly, does the salivary AMP concentration correlate with previously published putative diagnostic markers such as matrix
metalloproteinase 8 (MMP-8) (Costa et al., 2010; Furuholm et al., 2006; Gursoy et al., 2010; Leppilahti et al., 2011; Sorsa et al., 2011) and does combining the AMP and MMP-8 improve the diagnostic utility of the test? Thirdly, it remains to be known whether baseline levels of AMPs in saliva can predict the outcome of periodontal therapy. Lastly, it is important to establish the relationship between DEFA1A3 gene and salivary HNP 1-3 concentration and any possible link with susceptibility to periodontal disease.

To address the first question of establishing the diagnostic cut-off points, a cross-sectional study in a new cohort comprising of periodontally healthy and untreated gingivitis, CP and AgP subjects would be set up to measure the salivary AMP concentrations. Cut-off points would be determined using ROC constructs from these data. The new cut-off points would then be tested for their ability to accurately diagnose gingivitis, CP and AgP from healthy subjects in a new independent cohort. The data would be analyzed to determine whether salivary LL-37 levels would retain their ability to differentiate AgP from CP and gingivitis as observed in this current report.

To compare the levels of AMPs and those of other promising putative salivary biomarkers that have been shown to have high predictive value for periodontitis such as MMP-8, the latter would be measured in the saliva samples collected in the cohorts described above. The data would further be explored to determine whether combining salivary MMP-8 and HNP 1-3 tests would improve their overall diagnostic value than using a single biomarker test, as the pilot data obtained here and elsewhere (Gursoy et al., 2010; Gursoy et al., 2011; Kinney et al., 2011; Ramseier et al., 2009) suggests might be the case.

To address the third question, a prospective longitudinal study would be carried out. A suitable study population would be a cohort of periodontal patients with CP and AgP. These would have baseline pre-treatment AMP concentration in saliva measured followed by other measurements a day after the last treatment, and weekly thereafter for 6 consecutive weeks. Further measurements would be carried out at time 3 months, 6 months and 12 months post-treatment to determine the stability of any observed changes in salivary AMP concentration after periodontal therapy. In addition, a longitudinal follow up of patients not
accessing specialist periodontal care would be carried out to look for association of salivary AMP concentration with progression of disease. This group would also act as an untreated control for the treatment group.

Lastly, larger population studies would be useful both to determine if DEFA1A3 CNVs are associated with risk of periodontal disease, and secondly to investigate the relationship between CNVs and HNP 1-3 concentration in saliva.

### 7.1.2 Conclusions

This thesis has made marked progress in identifying novel salivary biomarkers, which appear to have considerable potential for diagnostic purposes in periodontal disease. Specifically, salivary antimicrobial peptides HNP 1-3 and LL-37 have been shown to consistently differentiate between periodontitis and gingivitis. Preliminary data have also demonstrated the potential for salivary LL-37 to differentiate CP from AgP, a feat that has not been demonstrated hitherto before by any other biomarker.

We have also clearly demonstrated that the putative salivary diagnostic markers HNP 1-3 and LL-37 are stable for up to 48 hours at room temperature. The implications are that they can possibly be employed in field studies for periodontal disease studies. Samples could be mailed through regular postal services from the field back to the analysing lab. This also opens the potential to package a salivary test for these biomarkers for use as at-home diagnostic kits.

In addition, it seems possible that these potential biomarkers may be useful in informing the treatment planning process. For instance, those patients with high levels of these biomarkers may need a more aggressive approach of management. In addition, it may be possible to use them for monitoring the treatment progress.

Finally, we have demonstrated that these factors that predominantly constitute part of the innate system are closely related to the presence and severity of periodontal disease. This potentially gives new insights into the pathogenic mechanism of periodontitis with the innate system playing a major role through the expression of HNP 1-3 and LL-37.
The future looks bright for the use of salivary HNP 1-3 and LL-37 as molecular descriptors in the classification of periodontal diseases. A molecular description of a diagnosis would potentially be more definitive and quantifiable than the current use of indices and clinical measurements that rely on the clinician’s judgement and experience.

Finally, this study has raised important questions that need appropriate future studies as highlighted in the previous section. In particular, further work is needed to identify other potential salivary protein biomarkers and to further validate salivary AMP cut-off points. Longitudinal intervention studies are also needed to determine the prognostic value of the use of salivary AMP testing in the treatment of patients with destructive periodontal disease and the role of DEFA1A3 copy numbers in the pathogenesis of periodontal disease.
Chapter 8 – Appendices
8.1 Chapter 2: Protein Identification

8.1.1 Biomarker discovery using SELDI-TOF MS

Preliminary studies in both GCF and saliva from subjects with gingivitis, CP and AgP using SELDI-TOF generated a list of potential biomarkers based on their differential expression in each of the periodontal status.

<table>
<thead>
<tr>
<th>Biomarker (Da)</th>
<th>P-Value</th>
<th>ROC Integral</th>
<th>Intensity Av</th>
<th>Putative Marker</th>
</tr>
</thead>
<tbody>
<tr>
<td>3307</td>
<td>0.00</td>
<td>0.76</td>
<td>42.8</td>
<td></td>
</tr>
<tr>
<td>3328</td>
<td>0.00</td>
<td>0.79</td>
<td>12.6</td>
<td></td>
</tr>
<tr>
<td>3347</td>
<td>0.00</td>
<td>0.79</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>3372</td>
<td>0.00</td>
<td>0.78</td>
<td>55.0</td>
<td>HNP-2</td>
</tr>
<tr>
<td>3378</td>
<td>0.00</td>
<td>0.78</td>
<td>50.4</td>
<td></td>
</tr>
<tr>
<td>3399</td>
<td>0.00</td>
<td>0.78</td>
<td>16.1</td>
<td></td>
</tr>
<tr>
<td>3423</td>
<td>0.01</td>
<td>0.72</td>
<td>38.9</td>
<td></td>
</tr>
<tr>
<td>3443</td>
<td>0.01</td>
<td>0.73</td>
<td>59.7</td>
<td>HNP-1</td>
</tr>
<tr>
<td>3457</td>
<td>0.04</td>
<td>0.69</td>
<td>24.4</td>
<td></td>
</tr>
<tr>
<td>3463</td>
<td>0.03</td>
<td>0.67</td>
<td>36.7</td>
<td></td>
</tr>
<tr>
<td>3466</td>
<td>0.00</td>
<td>0.74</td>
<td>12.4</td>
<td></td>
</tr>
<tr>
<td>3483</td>
<td>0.00</td>
<td>0.79</td>
<td>11.2</td>
<td></td>
</tr>
<tr>
<td>3487</td>
<td>0.04</td>
<td>0.69</td>
<td>36.7</td>
<td>HNP-3</td>
</tr>
<tr>
<td>3531</td>
<td>0.03</td>
<td>0.69</td>
<td>41.2</td>
<td></td>
</tr>
<tr>
<td>3536</td>
<td>0.00</td>
<td>0.78</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>3552</td>
<td>0.00</td>
<td>0.76</td>
<td>14.9</td>
<td></td>
</tr>
<tr>
<td>3574</td>
<td>0.00</td>
<td>0.78</td>
<td>9.8</td>
<td></td>
</tr>
<tr>
<td>3575</td>
<td>0.02</td>
<td>0.69</td>
<td>31.9</td>
<td></td>
</tr>
<tr>
<td>3597</td>
<td>0.00</td>
<td>0.83</td>
<td>11.4</td>
<td></td>
</tr>
<tr>
<td>3606</td>
<td>0.00</td>
<td>0.79</td>
<td>9.3</td>
<td></td>
</tr>
<tr>
<td>3640</td>
<td>0.00</td>
<td>0.79</td>
<td>5.2</td>
<td></td>
</tr>
<tr>
<td>3646</td>
<td>0.00</td>
<td>0.78</td>
<td>18.0</td>
<td></td>
</tr>
<tr>
<td>3684</td>
<td>0.01</td>
<td>0.72</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>3688</td>
<td>0.01</td>
<td>0.74</td>
<td>9.6</td>
<td></td>
</tr>
<tr>
<td>3756</td>
<td>0.03</td>
<td>0.67</td>
<td>7.4</td>
<td></td>
</tr>
<tr>
<td>4492</td>
<td></td>
<td></td>
<td>10.4</td>
<td>Cathelicidin (LL-37)</td>
</tr>
<tr>
<td>6424</td>
<td>0.00</td>
<td>0.78</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>7369</td>
<td>0.02</td>
<td>0.71</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>9097</td>
<td>0.00</td>
<td>0.78</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>10116*</td>
<td>0.00</td>
<td>0.78</td>
<td>1.3</td>
<td>hCAP18/LL-37</td>
</tr>
<tr>
<td>10163*</td>
<td>0.00</td>
<td>0.85</td>
<td>3.6</td>
<td>hCAP18/LL-37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>10206</td>
<td>0.00</td>
<td>0.76</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>10209</td>
<td>0.00</td>
<td>0.79</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>10237</td>
<td>0.02</td>
<td>0.69</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>10336</td>
<td>0.01</td>
<td>0.76</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>10421</td>
<td>0.00</td>
<td>0.79</td>
<td>7.2</td>
<td></td>
</tr>
<tr>
<td>10470</td>
<td>0.03</td>
<td>0.69</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>10576</td>
<td>0.00</td>
<td>0.83</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>10606</td>
<td>0.00</td>
<td>0.81</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>10638</td>
<td>0.01</td>
<td>0.73</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>10859</td>
<td>0.01</td>
<td>0.72</td>
<td>24.4</td>
<td></td>
</tr>
<tr>
<td><strong>11060</strong></td>
<td>0.03</td>
<td>0.69</td>
<td>6.7</td>
<td>hCAP18/LL-37</td>
</tr>
<tr>
<td>12689</td>
<td>0.01</td>
<td>0.72</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>12890</td>
<td>0.03</td>
<td>0.69</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>13264</td>
<td>0.00</td>
<td>0.76</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>13295</td>
<td>0.04</td>
<td>0.68</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>13322</td>
<td>0.01</td>
<td>0.71</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td><strong>14694</strong></td>
<td>0.00</td>
<td>0.79</td>
<td>2.1</td>
<td>hCAP18/LL-37</td>
</tr>
<tr>
<td>14750</td>
<td>0.00</td>
<td>0.77</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>14939</td>
<td>0.00</td>
<td>0.83</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>15076</td>
<td>0.00</td>
<td>0.78</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>15124</td>
<td>0.00</td>
<td>0.88</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>21514</td>
<td>0.00</td>
<td>0.81</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>21708</td>
<td>0.03</td>
<td>0.68</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>23972</td>
<td>0.01</td>
<td>0.74</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>26153</td>
<td>0.00</td>
<td>0.78</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>26246</td>
<td>0.00</td>
<td>0.83</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>38353</td>
<td>0.02</td>
<td>0.69</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>2641</td>
<td>0.00</td>
<td>0.16</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>2796</td>
<td>0.01</td>
<td>0.29</td>
<td>5.8</td>
<td></td>
</tr>
<tr>
<td>2867</td>
<td>0.00</td>
<td>0.19</td>
<td>12.4</td>
<td></td>
</tr>
<tr>
<td>4109</td>
<td>0.00</td>
<td>0.24</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td><strong>4132</strong></td>
<td>0.00</td>
<td>0.14</td>
<td>8.0</td>
<td>HBD-2</td>
</tr>
<tr>
<td>4221</td>
<td>0.04</td>
<td>0.35</td>
<td>17.1</td>
<td></td>
</tr>
<tr>
<td>4989</td>
<td>0.00</td>
<td>0.14</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td><strong>5068</strong></td>
<td>0.00</td>
<td>0.22</td>
<td>4.0</td>
<td>HBD-3</td>
</tr>
<tr>
<td>5414</td>
<td>0.02</td>
<td>0.31</td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>5443</td>
<td>0.03</td>
<td>0.30</td>
<td>9.6</td>
<td></td>
</tr>
<tr>
<td>5466</td>
<td>0.00</td>
<td>0.26</td>
<td>6.6</td>
<td></td>
</tr>
</tbody>
</table>
Table 8-1: Partial list of putative GCF & Salivary markers for periodontal disease as flagged by SELDI-TOF MS. The markers with tentatively known identity are in bold font. Different suspected fragments of partially processed hCAP38/LL-37 are shown with an asterisk (*).

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6369</td>
<td>0.03</td>
<td>0.30</td>
<td>6.0</td>
</tr>
<tr>
<td>6766</td>
<td>0.00</td>
<td>0.21</td>
<td>4.4</td>
</tr>
<tr>
<td>6794</td>
<td>0.00</td>
<td>0.24</td>
<td>3.5</td>
</tr>
<tr>
<td>13536</td>
<td>0.00</td>
<td>0.25</td>
<td>3.6</td>
</tr>
<tr>
<td>13774</td>
<td>0.02</td>
<td>0.30</td>
<td>2.4</td>
</tr>
<tr>
<td>33545</td>
<td>0.00</td>
<td>0.21</td>
<td>0.1</td>
</tr>
</tbody>
</table>

8.1.2 Prefractionation spin column protocols

8.1.2.1 IMAC 30 Spin Column

IMAC support is supplied in a slurry/suspension in 1 M sodium chloride, 20% ethanol (v/v).

Before immobilizing a metal ion, the support is washed extensively with deionized (DI) water to eliminate ethanol and sodium chloride preservatives in the storage solution.

Recommended sample binding time is 20–40 minutes

Charging Solutions

Copper (Cu2+) — 100 mM copper sulfate

Binding and Equilibration Buffers

Binding buffer (BB) — 0.1M NaH2PO4, 0.5 M NaCl, adjusted to pH 8.0 with NaOH

Elution Buffers

Elution was achieved by increasing concentrations of imidazole.

Elution buffer (EB) — 10mM, 40mM, 70mM and 100mM imidazole

Tap the ProteinChip IMAC spin column lightly to settle the support to the bottom (near the tapered end) of the column.

Remove top and bottom caps on the spin column.
Set the spin column upright in a 1.5 ml microcentrifuge tube.

Centrifuge the spin column at approximately 80 x g (1,000 rpm) for 30 seconds to remove the buffer.

Add 500 µl of DI water to the column. Centrifuge the column to remove water. Repeat two more times. Replace bottom cap on the column.

Add 200 µl of 100 mM copper sulfate to the column, then replace the top cap and incubate with mixing for 20 minutes at room temperature (RT).

Remove the top and bottom caps; set the spin column upright in a new 1.5 ml microcentrifuge tube. Centrifuge the column for 30 seconds to remove the charging solution. Repeat steps 6 and 7 once.

Add 200 µl of DI water to the spin column. Centrifuge the column to remove the water. Repeat two more times. Replace the bottom cap.

Add 200 µl of BB to the column; replace top cap. Mix and centrifuge the column to remove the buffer. Replace the bottom cap. Repeat two more times.

Re-suspend the support with 90 µl of BB.

**Sample Preparation**

For GCF, Dilute 100–150 µl of sample 1:3 in BB1. For saliva, no need to dilute.

Add sample to the spin column, then replace the top cap. Mix for 20 to 40 minutes at 4°C.

Remove top and bottom caps; set column upright in a new 1.5 ml microcentrifuge tube. Centrifuge column for 30 seconds and collect the flow through (unbound fraction). Replace the bottom cap.

Add 200 µl of BB to wash unbound fraction and add to first flow through.

Add 100µl of 10mM EB to the spin column, then replace the top cap and incubate with mixing for 5 minutes at RT. Remove the top and bottom caps; set the column upright in a new 1.5 ml microcentrifuge tube. Centrifuge for 30 seconds and collect the fraction. Replace the bottom cap.
Repeat eluting step and pool the fractions. Label the tube with the buffer used i.e. 10mM EB

Repeat eluting step with an increasing molarity (40mM, 70mM and 100mM) of the EB and pool the duplicate fractions. Label the tube with the buffer used.

The fractions are ready to be profiled on ProteinChip® arrays.

**8.1.2.2 CM10 spin column**

Recommended sample binding time is 20–40 minutes

Each column containing CM ceramic HyperD F sorbent has a minimal protein binding capacity of 5.4 mg.

**Column Equilibration**

ProteinChip CM spin columns contain a sorbent that requires buffer equilibration prior to use.

This equilibration buffer should be the same as the intended sample binding buffer i.e. CM10 Binding buffer (ProteinChip CM low-stringency buffer (catalog #K20-00003) 0.1 M sodium acetate, pH 4.0 antimicrobial preservatives, 200 ml). The salt concentration in this buffer should be <100 mM and the pH should be approximately 4.5 to ensure that most proteins will bind to the column. Tap the spin column lightly to settle the sorbent to the bottom (near the tapered end) of the column. Remove the caps on the top and bottom of the spin column. Set the column upright in a 1.5 ml microcentrifuge tube. Centrifuge the spin column at ~80 x g (1,000 rpm) for 30 seconds to remove the buffer. Replace the bottom cap of the spin column.

Add 200µl buffer, pH 4.5 to the column, and then replace the top cap on the column. Vortex the column to mix the sorbent with the buffer. Remove the top and bottom caps of the spin column; set the column upright in a new 1.5 ml microcentrifuge tube. Centrifuge the spin column for 30 seconds to remove the buffer.

Repeat the washing steps for a total of three buffer washes. Replace the bottom cap on the spin column. The sample can now be added to the spin
Acetic acid was used to adjust the pH of the buffer to pH 4.5, 6.0, 7.0, 8.0 and 9.0.

**8.1.3 Confirmation of peaks using SELDI-TOF MS**

**8.1.3.1 ProteinChip array treatment and buffers**

1. Add 5 µl per spot of 0.1 M copper sulfate and incubate for 10 minutes with shaking (on MicroMix shaker setting 20/4).
2. Remove the metal solution and replace with 5 µl of DI water and incubate for 1 minute.
3. Remove the water and replace it with 0.1 M sodium acetate buffer, pH 4 (neutralization buffer) and incubate for 5 minutes.
4. Remove the neutralization buffer and replace with 5 µl of DI water and incubate for 1 minute.
5. Remove the water and replace with 5 µl of binding buffer on each spot for 5 minutes. Repeat once.
6. Remove the binding buffer and replace with 5 µl of sample. Do not allow the spot to air-dry during sample application.
7. Incubate in a humid chamber for 30 minutes with shaking (on MicroMix shaker setting 20/4).
8. Wash each spot with 5 µl of binding buffer with shaking, and remove buffer. Repeat two more times.
9. Wash each spot with 5 µl of DI water. Repeat once.
10. Air-dry the array for 15–20 minutes.
11. Apply 1 µl of ProteinChip SPA EAM in solution to each spot. Air-dry for 5 minutes and apply another 1 µl of EAM in solution. Allow to air-dry.
12. Analyse the array using the ProteinChip SELDI system.

---

**Table 8-2: Protocol for GCF Profiling Using a Copper-Enriched Array On-Spot.**
1. Pre-wet the spots with 5 µl of binding buffer for 5 minutes. Repeat once.
2. Remove the pre-wetting solution and replace with 5 µl of sample. Do not allow the spot to air-dry during sample application.
3. Incubate in a humid chamber for 30 minutes with shaking (on Micromix shaker setting 20/4).
4. Wash each spot with 5 µl of binding buffer with shaking and remove buffer. Repeat two more times.
5. Wash each spot with 5 µl of DI water. Repeat once.
6. Air-dry the array for 15–20 minutes.
7. Apply 1 µl of EAM in solution to each spot. Air-dry for 5 minutes and apply another 1 µl of EAM in solution. Allow to air-dry.
8. Analyze the array using the ProteinChip SELDI system.

Table 8-3: Protocol for GCF Profiling On-Spot with CM10.

8.1.4 Stacking Gel Preparation for Membrane Proteomics
Using Invitrogen X Cell ‘Sure-lock’ electrophoresis system, 1mm plastic cassettes and 1mm 10well combs. (Refer to the Invitrogen X Cell ‘Sure-lock’ instruction manual). TEMED from Sigma-Aldrich; Acrylamide, APS and gel running buffers from National Diagnostics.

<table>
<thead>
<tr>
<th>Constituents</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>0.38ml</td>
</tr>
<tr>
<td>Resolving gel buffer</td>
<td>1.625ml</td>
</tr>
<tr>
<td>Acrylamide</td>
<td>4.125ml</td>
</tr>
<tr>
<td>APS (10% w/v)</td>
<td>62.5ul</td>
</tr>
<tr>
<td>TEMED</td>
<td>6.25ul</td>
</tr>
</tbody>
</table>

Table 8-4: Table showing the composition of 20% resolving gel used in 1D-PAGE before LC-S/MS.

Layer a small amount of 50% isopropanol over the top of the gel so it levels out while setting and leave to set for approximately 30 minutes. Mark with a permanent pen the level of the resolving gel on the cassette. Wash out the isopropanol and rinse several times with water. Pour stacking gel to give 1-2 cm stacking gel height. Leave to set for 30-40 minutes. Remove the comb and
the white tap from the bottom of the gel and securely fix into the gel tank. Fill the tank with Tris-glycine buffer (~800 mL) and check gel running.

<table>
<thead>
<tr>
<th>Constituents</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>1.83ml</td>
</tr>
<tr>
<td>Stacking gel buffer</td>
<td>0.75ml</td>
</tr>
<tr>
<td>Acrylamide</td>
<td>0.39ml</td>
</tr>
<tr>
<td>APS</td>
<td>45ul</td>
</tr>
<tr>
<td>TEMED</td>
<td>2.5ul</td>
</tr>
</tbody>
</table>

Table 8-5: Table showing the composition of 4% stacking gel using 10 well, 1mm comb used before LC-S/MS.

**Sample preparation:**
Boil the sample in Laemmli buffer for 10 minutes before loading on the gel. A maximum of 40µl sample can be loaded per well. However it is advisable to spread samples out across all wells and keep the volume per well to a minimum.

Run the gel at 150V for 20-30 minutes or until the dye front just passes the interface between the stacking and resolving gel. Carefully remove the gel from the cassette and place in a sterile plastic tray.

**Colloidal Coomassie Staining.**
Fix proteins for 30 minutes in a solution of 7% acetic acid, 40% methanol. It is important to keep the acetic acid concentration low so as not to over fix the proteins in the gel. For 20mls of distaining solution, make 8ml Methanol, 1.4ml acetic acid and 10.6ml de-ionised water (di-water). Dilute the colloidal concentrate with di-water; 1 part colloidal concentrate to 4 parts di-water. Immediately before staining combine 4 parts of the above solution with 1 part methanol. Mix well by vortexing. Place the gel in the stain for 30minutes or until bands are clear. (This is sample dependent but you may need a reduced staining time compared to normal Coomassie protocols because stacking gel bands can stain intensely).
De-stain with 25% methanol, 7% acetic acid for 10 minutes. For 20mls of distaining solution, make 5mls methanol, 1.4ml acetic acid and 13.6ml di-water. Remove and continue de-staining with 25% methanol, 2% acetic acid for a further 30 minutes. For 20mls of distaining solution, make 5mls methanol, 0.4ml acetic acid and 14.6ml di-water. Store gel in de-ionised water at 4°C.

**Figure 8.1:** Apparent molecular weights of SeeBlue Plus2 Pre-stained standard on a NuPAGE Novex 4-12% Bis-Tris Gel w/MES
8.1.5 Protein identification using LC-MS/MS

GCF and saliva samples were analyzed using 1D-PAGE coupled to LC-MS/MS to identify the putative biomarkers.

<table>
<thead>
<tr>
<th>Protein ID</th>
<th>Accession Number</th>
<th>Molecular Weight</th>
<th>Taxonomy</th>
<th>2_3480 (F020168)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1AT_HUMAN Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3</td>
<td>P01009</td>
<td>47 kDa</td>
<td>unknown</td>
<td>75%</td>
</tr>
<tr>
<td>ACBP_HUMAN Acyl-CoA-binding protein OS=Homo sapiens GN=DBI PE=1 SV=2</td>
<td>P07108</td>
<td>10 kDa</td>
<td>unknown</td>
<td>65%</td>
</tr>
<tr>
<td>ACTB_HUMAN Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1</td>
<td>P60709 (+1)</td>
<td>42 kDa</td>
<td>unknown</td>
<td>100%</td>
</tr>
<tr>
<td>ALBU_HUMAN Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2</td>
<td>P02768</td>
<td>69 kDa</td>
<td>unknown</td>
<td>100%</td>
</tr>
<tr>
<td>Alloalbumin Venezia</td>
<td>gi</td>
<td>178345 (+5)</td>
<td>69 kDa</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>Alpha-1-antichymotrypsin</td>
<td>AACT_HUMAN</td>
<td>48 kDa</td>
<td>Homo sapiens</td>
<td>24%</td>
</tr>
<tr>
<td>Alpha-1-antitrypsin</td>
<td>A1AT_HUMAN</td>
<td>47 kDa</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>Alpha-1-antitrypsin precursor (Alpha-1 protease inhibitor) (Alpha-1-antiproteinase) (AAT), gi</td>
<td>68742</td>
<td>pir</td>
<td></td>
<td>ITBA alpha-1-antitrypsin precursor - baboon (fragment), gi</td>
</tr>
<tr>
<td>Alpha-amylase 1</td>
<td>AMY1_HUMAN</td>
<td>58 kDa</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>AMY1_HUMAN Alpha-amylase 1 OS=Homo sapiens GN=AMY1A PE=1 SV=2</td>
<td>P04745</td>
<td>58 kDa</td>
<td>unknown</td>
<td>100%</td>
</tr>
<tr>
<td>Amylase, pancreatic, alpha-2B precursor</td>
<td>gi</td>
<td>10280622 (+12)</td>
<td>58 kDa</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>B2MG_HUMAN Beta-2-microglobulin OS=Homo sapiens GN=B2M PE=1 SV=1</td>
<td>P61769</td>
<td>14 kDa</td>
<td>unknown</td>
<td>65%</td>
</tr>
<tr>
<td>C21orf18 isoform B protein</td>
<td>gi</td>
<td>17046385 (+2)</td>
<td>14 kDa</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession</td>
<td>Mass (kDa)</td>
<td>Species</td>
<td>Identity (%)</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------</td>
<td>-----------</td>
<td>------------</td>
<td>---------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Calcypophosphine 2</td>
<td>gi</td>
<td>14211947</td>
<td>Homo sapiens</td>
<td>51%</td>
</tr>
<tr>
<td>Casein alphaS1</td>
<td>gi</td>
<td>225632 (+2)</td>
<td>Bos taurus</td>
<td>99%</td>
</tr>
<tr>
<td>Chain , Modified Beta Trypsin (Monoisopropylphosphoryl Inhibited) (E.C.3.4.21.4) (Neutron Data)</td>
<td>gi</td>
<td>230196</td>
<td>Bos taurus</td>
<td>100%</td>
</tr>
<tr>
<td>Chain A, Bovine Beta-Trypsin Bound To Meta-Amidino Schiff Base Magnesium(II) Chelate, gi</td>
<td>13096613</td>
<td>pdb</td>
<td>1G3C</td>
<td>A Chain A, Bovine Beta-Trypsin Bound To Para-Amidino Schiff Base Iron(III) Chelate, gi</td>
</tr>
<tr>
<td>Chain A, Crystal Structure Of A Serpin:protease Complex</td>
<td>gi</td>
<td>11514321 (+16)</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>Chain A, Trypsin (E.C.3.4.21.4) Complexed With The Inhibitor Diisopropyl-Fluorophosphofluoridate (Dfp), gi</td>
<td>1064991</td>
<td>pdb</td>
<td>1TRN</td>
<td>B Chain B, Trypsin (E.C.3.4.21.4) Complexed With The Inhibitor Diisopropyl-Fluorophosphofluoridate (Dfp)</td>
</tr>
<tr>
<td>CRNN_HUMAN Cornulin OS=Homo sapiens GN=CRNN PE=1 SV=1</td>
<td>Q9UBG3</td>
<td>54 kDa</td>
<td>unknown</td>
<td>65%</td>
</tr>
<tr>
<td>cystatin SA-III=potential precursor of acquired enamel pellicle</td>
<td>gi</td>
<td>235948 (+2)</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>Cystatin-B</td>
<td>CYTB_HUMAN</td>
<td>11 kDa</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>Cystatin-D</td>
<td>CYTD_HUMAN</td>
<td>16 kDa</td>
<td>unknown</td>
<td>32%</td>
</tr>
<tr>
<td>Cystatin-SN</td>
<td>CYTN_HUMAN</td>
<td>16 kDa</td>
<td>unknown</td>
<td>100%</td>
</tr>
<tr>
<td>Cysteine-rich secretory protein 3</td>
<td>CRIS3_HUMAN</td>
<td>28 kDa</td>
<td>unknown</td>
<td>50%</td>
</tr>
<tr>
<td>CYTA_HUMAN Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1</td>
<td>P01040</td>
<td>11 kDa</td>
<td>unknown</td>
<td>100%</td>
</tr>
<tr>
<td>CYTB_HUMAN Cystatin-B OS=Homo sapiens GN=CSTB PE=1 SV=2</td>
<td>P04080</td>
<td>11 kDa</td>
<td>unknown</td>
<td>100%</td>
</tr>
<tr>
<td>CYTC_HUMAN Cystatin-C OS=Homo sapiens GN=CST3 PE=1 SV=1</td>
<td>P01034</td>
<td>16 kDa</td>
<td>unknown</td>
<td>65%</td>
</tr>
<tr>
<td>CYTD_HUMAN Cystatin-D OS=Homo sapiens GN=CST5 PE=1 SV=1</td>
<td>P28325</td>
<td>16 kDa</td>
<td>unknown</td>
<td>87%</td>
</tr>
<tr>
<td>CYTN_HUMAN Cystatin-SN OS=Homo sapiens GN=CST1 PE=1 SV=2</td>
<td>P01037</td>
<td>16 kDa</td>
<td>unknown</td>
<td>100%</td>
</tr>
<tr>
<td>CYTS_HUMAN Cystatin-S OS=Homo sapiens GN=CST4 PE=1 SV=3</td>
<td>P01036</td>
<td>16 kDa</td>
<td>unknown</td>
<td>100%</td>
</tr>
<tr>
<td>CYTT_HUMAN Cystatin-SA OS=Homo sapiens GN=CST2 PE=1 SV=1</td>
<td>P09228</td>
<td>16 kDa</td>
<td>unknown</td>
<td>100%</td>
</tr>
<tr>
<td>DEF1_HUMAN Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1</td>
<td>P59665 (+1)</td>
<td>10 kDa</td>
<td>unknown</td>
<td>100%</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Description</td>
<td>OS</td>
<td>GN</td>
<td>PE</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>-------</td>
<td>--------</td>
<td>----</td>
</tr>
<tr>
<td>DMBT1_HUMAN</td>
<td>Deleted in malignant brain tumors 1 protein</td>
<td>Homo sapiens</td>
<td>DMBT1</td>
<td>1</td>
</tr>
<tr>
<td>ENO1_HUMAN</td>
<td>Alpha-enolase</td>
<td>Homo sapiens</td>
<td>ENO1</td>
<td>1</td>
</tr>
<tr>
<td>FABP5_HUMAN</td>
<td>Fatty acid-binding protein, epidermal</td>
<td>Homo sapiens</td>
<td>FABP5</td>
<td>1</td>
</tr>
<tr>
<td>FIBA_HUMAN</td>
<td>Fibrinogen alpha chain</td>
<td>Homo sapiens</td>
<td>FGA</td>
<td>1</td>
</tr>
<tr>
<td>G3P_HUMAN</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
<td>Homo sapiens</td>
<td>GAPDH</td>
<td>1</td>
</tr>
<tr>
<td>G6P1_HUMAN</td>
<td>Glucose-6-phosphate isomerase</td>
<td>Homo sapiens</td>
<td>GPI</td>
<td>1</td>
</tr>
<tr>
<td>GELS_HUMAN</td>
<td>Gelsolin</td>
<td>Homo sapiens</td>
<td>GSN</td>
<td>1</td>
</tr>
<tr>
<td>GSTP1_HUMAN</td>
<td>Glutathione S-transferase P</td>
<td>Homo sapiens</td>
<td>GSTP1</td>
<td>1</td>
</tr>
<tr>
<td>H2A1_HUMAN</td>
<td>Histone H2A type 1</td>
<td>Homo sapiens</td>
<td>HIST1H2AG</td>
<td>1</td>
</tr>
<tr>
<td>H2B1K_HUMAN</td>
<td>Histone H2B type 1-K</td>
<td>Homo sapiens</td>
<td>HIST1H2BK</td>
<td>1</td>
</tr>
<tr>
<td>H31_HUMAN</td>
<td>Histone H3.1</td>
<td>Homo sapiens</td>
<td>HIST1H3A</td>
<td>1</td>
</tr>
<tr>
<td>H4_HUMAN</td>
<td>Histone H4</td>
<td>Homo sapiens</td>
<td>HIST1H4A</td>
<td>1</td>
</tr>
<tr>
<td>HBA1_HUMAN</td>
<td>Hemoglobin subunit alpha</td>
<td>Homo sapiens</td>
<td>HBA1</td>
<td>1</td>
</tr>
<tr>
<td>HBB_HUMAN</td>
<td>Hemoglobin subunit beta</td>
<td>Homo sapiens</td>
<td>HBB</td>
<td>1</td>
</tr>
<tr>
<td>HPT_HUMAN</td>
<td>Haptoglobin</td>
<td>Homo sapiens</td>
<td>HP</td>
<td>1</td>
</tr>
<tr>
<td>HSP71_HUMAN</td>
<td>Heat shock 70 kDa protein 1A/1B</td>
<td>Homo sapiens</td>
<td>HSPA1A</td>
<td>1</td>
</tr>
<tr>
<td>Ig alpha-1 chain C region</td>
<td>Homo sapiens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry</td>
<td>Description</td>
<td>Accession</td>
<td>Mass (kDa)</td>
<td>Species</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>-----------</td>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>Ig gamma-2 chain C region</td>
<td>IGHG2_HUMAN</td>
<td>36</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>Ig gamma-4 chain C region</td>
<td>IGHG4_HUMAN</td>
<td>36</td>
<td>Homo sapiens</td>
<td>98%</td>
</tr>
<tr>
<td>Ig kappa chain C region</td>
<td>IGKC_HUMAN</td>
<td>12</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>Ig lambda chain C regions</td>
<td>LAC_HUMAN</td>
<td>11</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>IGHA1_HUMAN Ig alpha-1 chain C region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OS=Homo sapiens GN=IGHA1 PE=1 SV=2</td>
<td>P01876</td>
<td>38</td>
<td>unknown</td>
</tr>
<tr>
<td>IGHG1_HUMAN Ig gamma-1 chain C region</td>
<td>OS=Homo sapiens GN=IGHG1 PE=1 SV=2</td>
<td>P01857</td>
<td>36</td>
<td>unknown</td>
</tr>
<tr>
<td>IGKC_HUMAN Ig kappa chain C region</td>
<td>OS=Homo sapiens GN=IGKC PE=1 SV=1</td>
<td>P01857</td>
<td>12</td>
<td>unknown</td>
</tr>
<tr>
<td>IL1RA_HUMAN Interleukin-1 receptor antagonist protein</td>
<td>OS=Homo sapiens GN=IL1RN PE=1 SV=1</td>
<td>P18510</td>
<td>20</td>
<td>unknown</td>
</tr>
<tr>
<td>ILEU_HUMAN Leukocyte elastase inhibitor</td>
<td>OS=Homo sapiens GN=SERPINB1 PE=1 SV=1</td>
<td>P30740</td>
<td>43</td>
<td>unknown</td>
</tr>
<tr>
<td>Immunoglobulin heavy chain constant region</td>
<td>gi</td>
<td>10799664 (+6)</td>
<td>36</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>Keratin 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keratin, type I cytoskeletal 10</td>
<td>K1C10_HUMAN</td>
<td>60</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>Keratin, type I cytoskeletal 9 (Cytokeratin-9) (CK-9) (Keratin-9) (K9), gi</td>
<td>435476</td>
<td>emb</td>
<td>CAA82315.1</td>
<td>cytokeratin 9</td>
</tr>
<tr>
<td>Keratin, type II cytoskeletal 1</td>
<td>K2C1_HUMAN</td>
<td>66</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>Keratin, type II cytoskeletal 2 epidermal</td>
<td>K22E_HUMAN</td>
<td>66</td>
<td>Homo sapiens</td>
<td>51%</td>
</tr>
<tr>
<td>KIAA1607 protein</td>
<td>gi</td>
<td>10047289 (+1)</td>
<td>143</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>KV302_HUMAN Ig kappa chain V-III region SIE</td>
<td>OS=Homo sapiens PE=1 SV=1</td>
<td>P01620 (+3)</td>
<td>12</td>
<td>unknown</td>
</tr>
<tr>
<td>LAC_HUMAN Ig lambda chain C regions</td>
<td>OS=Homo sapiens GN=IGLC1 PE=1 SV=1</td>
<td>P01842</td>
<td>11</td>
<td>unknown</td>
</tr>
<tr>
<td>LCN1_HUMAN Lipocalin-1</td>
<td>OS=Homo sapiens GN=LCN1 PE=1 SV=1</td>
<td>P31025</td>
<td>19</td>
<td>unknown</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession</td>
<td>MW (kDa)</td>
<td>Species</td>
<td>Identity (%)</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-----------</td>
<td>----------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>LDHA_HUMAN L-lactate dehydrogenase A chain</td>
<td>P00338</td>
<td>37</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>Leukocyte elastase inhibitor</td>
<td>ILEU_HUMAN</td>
<td>43</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>LRIQ1_HUMAN Leucine-rich repeat and IQ domain-containing protein 1</td>
<td>Q96JN4</td>
<td>155</td>
<td>Homo sapiens</td>
<td>19%</td>
</tr>
<tr>
<td>LYSC_HUMAN Lysozyme C</td>
<td>P61626</td>
<td>17</td>
<td>Homo sapiens</td>
<td>65%</td>
</tr>
<tr>
<td>MUC5B_HUMAN Mucin-5B</td>
<td>Q9HC84</td>
<td>590</td>
<td>Homo sapiens</td>
<td>65%</td>
</tr>
<tr>
<td>mutant keratin 9</td>
<td>gi</td>
<td>1890020 (+2)</td>
<td>26</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>NC2A_HUMAN Dr1-associated corepressor</td>
<td>Q14919</td>
<td>22</td>
<td>Homo sapiens</td>
<td>31%</td>
</tr>
<tr>
<td>Neutrophil gelatinase-associated lipocalin</td>
<td>P80188</td>
<td>23</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>neutrophil granule peptide HP1</td>
<td>gi</td>
<td>228797 (+4)</td>
<td>03</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>NGAL_HUMAN Neutrophil gelatinase-associated lipocalin OS=Homo sapiens</td>
<td>gi</td>
<td>1890020 (+2)</td>
<td>26</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>PERM_HUMAN Myeloperoxidase</td>
<td>P05164</td>
<td>84</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>PIGR_HUMAN Polymeric immunoglobulin receptor</td>
<td>P01833</td>
<td>83</td>
<td>Homo sapiens</td>
<td>65%</td>
</tr>
<tr>
<td>GN=PIGR PE=1 SV=4</td>
<td>P12273</td>
<td>17</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>PLSL_HUMAN Plastin-2</td>
<td>P13796</td>
<td>70</td>
<td>Homo sapiens</td>
<td>65%</td>
</tr>
<tr>
<td>Polymeric immunoglobulin receptor</td>
<td>PIGR_HUMAN</td>
<td>83</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>polyubiquitin</td>
<td>gi</td>
<td>627129 (+41)</td>
<td>68</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>PPIA_HUMAN Peptidyl-prolyl cis-trans isomerase A</td>
<td>P62937</td>
<td>18</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
<tr>
<td>PRD15_HUMAN PR domain zinc finger protein 15</td>
<td>P57071</td>
<td>169</td>
<td>Homo sapiens</td>
<td>24%</td>
</tr>
<tr>
<td>PREDICTED: keratin 6A</td>
<td>gi</td>
<td>114644568 (+6)</td>
<td>60</td>
<td>Pan troglodytes</td>
</tr>
<tr>
<td>PROF1_HUMAN Profilin-1</td>
<td>P07737</td>
<td>15</td>
<td>Homo sapiens</td>
<td>100%</td>
</tr>
</tbody>
</table>
CHAPTER 8 – APPENDICES

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Accession</th>
<th>Molecular Mass</th>
<th>Molar Mass</th>
<th>Mass Percentage</th>
<th>Molar Mass Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>28 kDa</td>
<td>65%</td>
<td></td>
</tr>
<tr>
<td>P24156</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>81 kDa</td>
<td>25%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>83 kDa</td>
<td>87%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>11 kDa</td>
<td>70%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>11 kDa</td>
<td>70%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>305 kDa</td>
<td>87%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>10 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>13 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>11 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>43 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>10 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>27 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>45 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>12 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>21 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>246 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>16 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>77 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>97 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>12 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>68 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>77 kDa</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>P24158</td>
<td>gi</td>
<td>21614544 (+2)</td>
<td>97 kDa</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

Legend:

- **PRTN3_HUMAN**: Myeloblastin
- **PUS7L_HUMAN**: Pseudouridylate synthase 7 homolog
- **QSOX1_HUMAN**: Sulfhydryl oxidase 1
- **RETN_HUMAN**: Resistin
- **RYR2_HUMAN**: Ryanodine receptor 2
- **S100 calcium-binding protein A8**: gi|21614544 (+2)
- **S100P_HUMAN**: Protein S100-P
- **S10A7_HUMAN**: Protein S100-A7
- **S10A8_HUMAN**: Protein S100-A8
- **S10A9_HUMAN**: Protein S100-A9
- **SERPINB3**: Serpin B3
- **SERPINB3**: Serpin B3
- **SH3BGRL3**: SH3 domain-binding glutamic acid-rich-like protein 3
- **SPINK1**: Serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 1
- **SPLUNC2**: Short palate, lung and nasal epithelium carcinoma-associated protein 2
- **TFF1**: TFF1
- **THIO_HUMAN**: Thorodoxin
- **TKT_HUMAN**: Transketolase
- **TRY1_HUMAN**: Trypsin-1
<table>
<thead>
<tr>
<th>Accession</th>
<th>Name</th>
<th>OS</th>
<th>GN</th>
<th>PE</th>
<th>SV</th>
<th>MW (kDa)</th>
<th>Status</th>
<th>Matched %</th>
</tr>
</thead>
<tbody>
<tr>
<td>P62988</td>
<td>Ubio HUMAN Ubiquitin</td>
<td>Homo sapiens</td>
<td>RPS27A</td>
<td>1</td>
<td></td>
<td>248</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>gi</td>
<td>28317</td>
<td>unnamed protein product</td>
<td>Homo sapiens</td>
<td>unknown</td>
<td></td>
<td></td>
<td>09</td>
<td>100%</td>
</tr>
<tr>
<td>gi</td>
<td>96900</td>
<td>ZG16B_HUMAN Zymogen granule protein 16 homolog B</td>
<td>Homo sapiens</td>
<td>ZG16B</td>
<td>1</td>
<td></td>
<td>34</td>
<td>100%</td>
</tr>
<tr>
<td>gi</td>
<td>96904</td>
<td>Zinc-alpha-2-glycoprotein</td>
<td>Homo sapiens</td>
<td>ZA2G</td>
<td>3</td>
<td></td>
<td>34</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 8.6: Complete list of all the proteins and peptides identified using LC-MS/MS in both GCF and saliva.

CHAPTER 8 – APPENDICES
8.2 Chapter 3: Preliminary biomarker validation

8.2.1 ELISA protocols

8.2.1.1 HNP 1-3
All the reagents were brought to room temperature (20 - 25°C) before use. 100µl of standard and samples were applied into appropriate wells in duplicate in a 96 well plate pre-coated with antibody. The plate was incubated for 1 hour at room temperature before the solutions were carefully emptied by inverting plate and shaking contents out then dried by tapping on a thick layer of tissues. Approximately 300 µl of wash buffer was added to each well and kept for 20 seconds before emptying as above. The washing process was repeated for 3 times. Then 100 µl of reconstituted biotinylated tracer antibody was added to each well using the same pipetting order as applied in previous steps and again incubated for 1 hour at room temperature. Washing was done as described above before 100 µl of diluted streptavidin-peroxidase was added to each well and incubated for 1 hour at room temperature. Washing as above was repeated and 100 µl of TMB substrate added to each well and incubated for 20–30 minutes at room temperature in darkness. The reaction was stopped by adding 100 µl of stop solution (oxalic acid) and the plates read immediately at 450 nm using a plate reader (Optima BMG LABTECH. Germany).

8.2.1.2 HBD3
All the reagents were brought to room temperature (20 - 25°C) before use. All the steps were performed at room temperature and all samples and standards done in duplicate in an antibody pre-coated 96 well plate. 200-300µl of working wash solution was added to each well and let to stand for about 5 minutes before sample addition. The solutions were carefully emptied by inverting plate and shaking contents out then dried by tapping on a thick layer of tissues. Approximately 300 µl of wash buffer was added to each well and kept for 20 seconds before emptying as above. The washing process was repeated for 4 times. 100µl of standards and samples were applied to wells, incubated for 60 minutes, and washing done as described previously. 100µl of working detection antibody was then added to each well and incubated for 1 hour then washing process repeated as above. 100µl of working streptavidin-
HRP conjugate was added to each well and incubated for 30 minutes then washed 5 times. 100 µl TMB substrate was then added to each well and incubated for 15 minutes in the dark before adding the 100 µl of stop solution. Optical absorbance was immediately read at 450nm and the read out used to plot a standard curve to generate an equation to calculate the actual concentration of the samples.

8.2.1.3 IL-1α

ELISA for IL-1alpha was done using the DuoSet ELISA Development kit (RandD Systems. Cat No. DY200). In brief, Capture Antibody was diluted to the working concentration in PBS without carrier protein. Ninety-six well microplates were immediately coated with 100 µl per well of the diluted capture antibody. The plates were incubated overnight at room temperature then washed 3 times with wash buffer (0.05% Tween 20-PBS) in an awtawasher and blocked using 300 µl reagent diluent (1% BSA in PBS, pH 7.2 - 7.4, 0.2 mm filtered) at room temperature for a minimum of 1 hour. Samples were diluted using reagent diluent (RD) by a factor of x10 and those found to have readings above the high standard were repeated with a dilution factor of x20. 100 µl of sample and standards in RD were applied per well and incubated for 2 hours at room temperature. Wells were aspirated and washed 3 times as described above. 100 µl of the working dilution of detection antibody (biotinylated goat anti-human IL-1a) was added to each well and incubated for 2 hours at room temperature and washing repeated as before. 100 µl of the working dilution of Streptavidin-HRP was added to each well then incubated for 20 minutes at room temperature in darkness followed by washing as before. 100 µl of substrate solution (1:1 mixture of H2O2 and Tetramethylbenzidine) was added to each well and incubated for 20 minutes at room temperature in darkness. The reaction was stopped by adding 50 µl of stop solution (2N H2SO4) to each well and immediately the optical density determined using a microplate reader set to 450nm (Optima BMG LABTECH. Germany). The optical density readings were then converted into actual IL-1α concentrations using a 2nd order polynomial equation. Two tailed student’s T-test for p value and receiver operating characteristic curves for sensitivity and specificity were generated.
using GraphPad Prism version 5.00b for Macintosh (GraphPad Software, San Diego California USA) Results were expressed as the mean and SEM.

8.2.1.4 LL-37
All the reagents were brought to room temperature (20 - 25°C) before use. All the steps were performed at room temperature and all samples and standards done in duplicate in an antibody pre-coated 96 well plate according to manufacturer’s instructions. Optical density of the plates was read immediately at 450 nm using a plate reader (Optima BMG LABTECH. Germany) and the values were used to derive the equation used to calculate the final concentrations of the analyte in the samples.

8.3 Chapter 5: Translational Studies

<table>
<thead>
<tr>
<th>Subject ID</th>
<th>Randomization number</th>
<th>Subject ID</th>
<th>Randomization number</th>
</tr>
</thead>
<tbody>
<tr>
<td>M2001</td>
<td>1</td>
<td>M2015</td>
<td>1</td>
</tr>
<tr>
<td>M2002</td>
<td>1</td>
<td>M2016</td>
<td>1</td>
</tr>
<tr>
<td>M2003</td>
<td>2</td>
<td>M2017</td>
<td>2</td>
</tr>
<tr>
<td>M2004</td>
<td>2</td>
<td>M2018</td>
<td>1</td>
</tr>
<tr>
<td>M2005</td>
<td>2</td>
<td>M2019</td>
<td>1</td>
</tr>
<tr>
<td>M2006</td>
<td>1</td>
<td>M2020</td>
<td>1</td>
</tr>
<tr>
<td>M2007</td>
<td>1</td>
<td>M2021</td>
<td>1</td>
</tr>
<tr>
<td>M2008</td>
<td>1</td>
<td>M2022</td>
<td>2</td>
</tr>
<tr>
<td>M2009</td>
<td>2</td>
<td>M2023</td>
<td>2</td>
</tr>
<tr>
<td>M2010</td>
<td>1</td>
<td>M2024</td>
<td>1</td>
</tr>
<tr>
<td>M2011</td>
<td>1</td>
<td>M2025</td>
<td>1</td>
</tr>
<tr>
<td>M2012</td>
<td>1</td>
<td>M2026</td>
<td>1</td>
</tr>
<tr>
<td>M2013</td>
<td>2</td>
<td>M2027</td>
<td>1</td>
</tr>
<tr>
<td>M2014</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**KEY**
- Spitting 1
- Salivette® 2

Table 8-7: Randomization list for collecting saliva samples using either salivette or by spitting first.
Chapter 9 - References


Curtis MA, Sterne JAC, Price SJ, Griffiths GS, Coulthurst SK, Wilton JMA *et al.* (1990). The protein composition of gingival crevicular fluid sampled from male adolescents with no
DESTRUCTIVE PERIODONTITIS: BASELINE DATA OF A LONGITUDINAL STUDY.


DAHER KA, SELSTED ME, LEHRER RI (1986). DIRECT INACTIVATION OF VIRUSES BY HUMAN GRANULOCYTE DEFENSINS.


Guerini V (1909). A history of dentistry from the most ancient times until the end of the eighteenth century Philadelphia: Lea & Febiger.


REFERENCES


