Characterization of CD109

by

Joseph Yao Abotsi Prosper

A thesis submitted in conformity with the requirements
Doctor of Philosophy
Graduate Department of Medical Biophysics
University of Toronto.

© Copyright by Joseph Yao Abotsi Prosper 2011
Characterization of CD109

Joseph Yao Abotsi Prosper
Doctor of Philosophy 2011
Department of Medical Biophysics, University of Toronto

ABSTRACT:

CD109 is a 170kD glycosylphosphatidylinositol-anchored protein expressed on subsets of fetal and adult CD34+ haematopoietic stem cells, endothelial cells, activated T cells, and activated platelets. Cloning of the CD109 cDNA by our group identified the molecule as a novel member of the α2M/C3/C4/C5 family of thioester containing proteins. Curiously, CD109 bears features of both the α2M and complement branches of the gene family. Additionally CD109 carries the antigenic determinant of the Gov alloantigen system, which has been implicated in a subset of immune mediated platelet destruction syndromes. In this thesis, the status of CD109 in the evolution and phylogeny of the A2M family has been clarified. First, I elucidated the evolutionary relationships of CD109, and of the other eight human A2M/C3/C4/C5 proteins, using sequence analysis and a detailed comparison of the organization of the corresponding loci. Extension of this analysis to compare CD109 to related sequences extending back to placazoans, defined CD109 as a member of a distinct and archaic branch of the A2M phylogenetic tree. Second, in conjunction with collaborators, the molecular basis of the Gov alloantigen system was identified as an allele specific A2108C; Y703S polymorphism. Utilizing cDNA and genomic sequence we then developed methods to accurately and precisely genotype the Gov system. Finally, the expression kinetics of platelet CD109 was elucidated, in order to obtain basic information regarding its expression and
subcellular localization, and to resolve discrepancies in reported platelet CD109 expression. Quantitative flow cytometry demonstrated that CD109 was expressed on the surface of activated platelets at very low levels in most healthy volunteers. In resting platelets, CD109 was localized to the OCS and intracellular storage granules. CD109 displayed differential agonist induced expression in comparison to GPIIb/IIIa epitope unmasking, and surface expression of CD62P and CD63. CD109 was rapidly expressed on the cell surface in response to low doses of both strong and weak agonists. This early expression is likely the result of CD109’s proximity to the plasma membrane in resting platelets. As such, CD109 is positioned to participate at early stages of primary haemostasis.
ACKNOWLEDGEMENTS:

For those of you who work at, or have visited the Princess Margaret Hospital you may have noticed the inspirational whiteboard on the main floor adjacent to the phlebotomy clinic. Today’s message - “its people not things that are life’s most valued treasures” - is the perfect framework as I reflect on my PhD.

I would like to thank my parents for all the support and sacrifices they have made for me (and my siblings). My siblings, I am fortunate to have such wonderful people I can call family, thank-you for being you.

I would like to thank Dr. Andre Schuh, for taking me on as a student in his lab. Thank-you, for the guidance, support, and sharing of your scientific and clinical knowledge. I would also like to thank-you (on behalf of my family as well) for your compassion during my father’s illness, I am eternally grateful.

This thesis manuscript is the culmination of years of work, with its associated crests and nadirs. There have been a number of people that have been instrumental in helping me along the journey. I would like to take this opportunity to thank a few of them. Wendy, for her technical and intellectual input, and of course our diverse topics of discussion. The evergreen Xiang-Fu, I can only dream of isolating proteins with your yield. Thank-you to all my student colleagues of the Schuh lab, past and present: Denise, Sandra, Abhijit, Martin, Quang, Iris, and Kaustav. Special thanks to Andrea, Boyko, Rachel, Smita, and Joanne.

I have been blessed to have some great people in my life. Your friendship has and will continue to be invaluable to me. Comfort, Ricki, Ivy, Isabel, Kwesi, Bright, Stephen, Imtiaz, Joanne K., Jose, Abhijit, Ryan, Sam, Mandy, Kristina, Joanne S., Neil C., Ockie, Rajal, Teresa, Hyun, Neal, Lan, Sandeep, Rich, Dennis, Fernando, Bart, Jon N., Nosh, Sian, Graeme, Christine, Ron, Monica, and Reza, words cannot express the gratitude I have for you all. My soccer teammates thanks for the good times on and off the pitch. If I have missed anyone I meant no slight.

J.
# TABLE OF CONTENTS:

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>lx</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>x</td>
</tr>
</tbody>
</table>

| Chapter 1: GENERAL INTRODUCTION                  | 1    |
| 1.1 PREFACE                                       | 1    |
| 1.2 CD109                                         | 6    |
| 1.2.1 Historical description of CD109             | 6    |
| 1.2.2 Cloning of CD109                           | 6    |
| 1.2.3 CD109 is a member of the \( \alpha_2 \)macroglobulin/C3, C4, C5 family | 7    |
| 1.2.4 Functional attributes of \( \alpha_2 \)macroglobulin | 7    |
| 1.2.5 Complement C3, C4, and C5                  | 8    |
| 1.2.6 CD109 shares features of both Complement C3 and \( \alpha_2 \)M | 11   |
| 1.3 PLATELET PHYSIOLOGY                          | 12   |
| 1.3.1 Platelet surface receptors                 | 13   |
| 1.3.1.1 GPIb-IX-V                                | 13   |
| 1.3.1.2 GPIib/IIia                               | 15   |
| 1.3.1.3 Collagen receptors                       | 17   |
| 1.3.1.4 Protease activated receptors             | 18   |
| 1.3.1.5 ADP receptor                             | 22   |
| 1.3.1.6 Thromboxane receptor                     | 23   |
| 1.3.1.7 Epinephrine receptor, \( \alpha_{2A} \)-adrenoreceptor | 24   |
| 1.3.1.8 Serotonin receptors                      | 26   |
| 1.3.1.9 Human platelet antigens                  | 27   |
| 1.3.2 Platelet ultrastructure                    | 28   |
| 1.3.2.1 Open canalicular system (OCS)             | 28   |
| 1.3.2.2 Dense tubular system (DTS)                | 30   |
| 1.3.2.3 Dense granules (\( \delta \)-granules)    | 30   |
| 1.3.2.4 Alpha granules (\( \alpha \)-granules)    | 31   |
| 1.3.2.5 Lysosomes                                | 32   |
| 1.3.3 Granule biogenesis                         | 32   |
| 1.3.4 Platelet degranulation                     | 33   |
| 1.3.4.1 Platelet release reaction                | 34   |
| 1.3.4.2 Heterogeneity and differential release of granules | 38   |
| Chapter 2: CD109 GENOMIC STRUCTURE AND PHYLOGENETIC ANALYSIS | 52   |
2.1 INTRODUCTION 52
2.2 OBJECTIVE AND AIMS 56
2.3 EXPERIMENTAL PROCEDURES 57
  2.3.1 Chromosomal localization of CD109 57
  2.3.2 Genomic structure of CD109 58
  2.3.3 Sequence analysis 59
2.4 RESULTS 61
  2.4.1 Chromosomal Mapping of the CD109 locus 61
  2.4.2 Genomic Structure of CD109 61
  2.4.3 The Human A2M Family 62
  2.4.4 Conservation of Genomic Structure in the Human A2M Family 63
  2.4.5 Conservation of Genomic Structure of Mammalian and Avian CD109 64
  2.4.6 Phylogenetic Position of CD109 65
2.5 DISCUSSION 81

Chapter 3: MOLECULAR DETERMINATION OF GOV ALLOANTIGEN DETERMINANTS 91
3.1 INTRODUCTION 91
3.2 OBJECTIVE AND AIMS 94
3.3 EXPERIMENTAL PROCEDURES 95
  3.3.1 Donor panels 95
  3.3.2 Gov phenotyping 95
  3.3.3 cDNA synthesis and analysis of CD109 sequence 95
  3.3.4 DNA sequence analysis of introns flanking Gov^a/b polymorphism-bearing CD109 exon 97
  3.3.5 PCR-SSP analysis of Gova/b alleles 97
  3.3.6 PCR-RFLP analysis of Gov^a/b alleles 97
  3.3.7 Allele-specific real-time PCR analysis 98
3.4 RESULTS 99
3.5 DISCUSSION 106

CHAPTER 4: QUANTIFICATION, SUBCELLULAR LOCALIZATION, AND EXPRESSION OF PLATELET CD109 111
4.1 INTRODUCTION 111
4.2 OBJECTIVES AND AIMS 113
4.3 EXPERIMENTAL PROCEDURES 114
  4.3.1 Materials 114
  4.3.2 Blood collection 114
  4.3.3 Platelet preparation 115
  4.3.4 Flow cytometric determination of epitope binding sites 115
  4.3.5 Flow cytometric analysis of agonist induced responses 116
  4.3.6 Flow cytometric determination of intracellular pools of CD109 116
  4.3.7 Subcellular fractionation of platelet lysates 117
  4.3.8 Immunoprecipitation 118
  4.3.9 Immunofluorescence microscopic analysis of platelets 118
  4.3.10 Immunoelectron microscopic analysis of platelets 119
  4.3.11 Statistical Analysis 120
4.4 RESULTS 121
  4.4.1 Quantification of CD109 levels on PAR1-AP stimulated platelets 121
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4.2 Cell surface CD109 is upregulated on suboptimally stimulated platelets</td>
<td>121</td>
</tr>
<tr>
<td>4.4.3 CD109 activation kinetics</td>
<td>122</td>
</tr>
<tr>
<td>4.4.4 Flow cytometric analysis reveals a subcellular pool of CD109</td>
<td>123</td>
</tr>
<tr>
<td>4.4.5 CD109 codistributes with granular and membrane fractions in sucrose density gradients</td>
<td>124</td>
</tr>
<tr>
<td>4.4.6 Subcellular localization of CD109 by immunofluorescence microscopy</td>
<td>125</td>
</tr>
<tr>
<td>4.4.7 Immunoelectron microscopic detection of platelet CD109</td>
<td>125</td>
</tr>
<tr>
<td>4.5 DISCUSSION</td>
<td>143</td>
</tr>
</tbody>
</table>

CHAPTER 5: SUMMARY AND FUTURE DIRECTIONS  
5.1 Summary  
5.2 Future Studies  
REFERENCES
LIST OF TABLES:

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>pg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Platelet granule contents</td>
<td>50</td>
</tr>
<tr>
<td>1-2</td>
<td>Platelet granule membrane proteins</td>
<td>51</td>
</tr>
<tr>
<td>2-1</td>
<td>Oligonucleotides used for the determination of the intron/exon structure of human CD109</td>
<td>68</td>
</tr>
<tr>
<td>2-2</td>
<td>Intron/exon structure of CD109</td>
<td>69</td>
</tr>
<tr>
<td>2-3</td>
<td>Conservation of intron phase among human A2M family genes</td>
<td>70</td>
</tr>
<tr>
<td>2-4</td>
<td>Conservation of intron phase among mammalian and avian CD109</td>
<td>71</td>
</tr>
<tr>
<td>2-5</td>
<td>Conservation of intron phase among human CD109, <em>C. elegans</em>, and <em>Drosophila TEP</em></td>
<td>72</td>
</tr>
</tbody>
</table>
LIST OF FIGURES:

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 1</td>
<td>CD109 cDNA schematic</td>
<td>39</td>
</tr>
<tr>
<td>1 - 2</td>
<td>CD109 primary structure</td>
<td>40</td>
</tr>
<tr>
<td>1 - 3</td>
<td>$\alpha$2M thioester reaction mechanism</td>
<td>41</td>
</tr>
<tr>
<td>1 - 4</td>
<td>C3 thioester reaction mechanism</td>
<td>42</td>
</tr>
<tr>
<td>1 - 5</td>
<td>Classical and lectin complement pathways</td>
<td>43</td>
</tr>
<tr>
<td>1 - 6</td>
<td>Alternative complement pathway</td>
<td>44</td>
</tr>
<tr>
<td>1 - 7</td>
<td>Overview of haemostasis</td>
<td>45</td>
</tr>
<tr>
<td>1 - 8</td>
<td>Adhesion and activation of platelets</td>
<td>46</td>
</tr>
<tr>
<td>1 - 9</td>
<td>Platelet agonist receptors</td>
<td>47</td>
</tr>
<tr>
<td>1 - 10</td>
<td>Major input signals regulating calcium in platelets</td>
<td>48</td>
</tr>
<tr>
<td>1 - 11</td>
<td>Role of SNAREs in $\alpha$-granule secretion</td>
<td>49</td>
</tr>
<tr>
<td>2 - 1</td>
<td>The genomic organization of CD109</td>
<td>73</td>
</tr>
<tr>
<td>2 - 2</td>
<td>Phylogenetic relationship of human $\alpha$2M family proteins</td>
<td>74</td>
</tr>
<tr>
<td>2 - 3</td>
<td>Conservation of genomic structure among human A2M family members</td>
<td>75</td>
</tr>
<tr>
<td>2 - 4</td>
<td>Conservation of genomic structure among CD109 orthologues</td>
<td>76</td>
</tr>
<tr>
<td>2 - 5</td>
<td>A cross-phyla A2M family phylogenetic tree</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Chapter 2 figure legends</td>
<td>79</td>
</tr>
<tr>
<td>3 - 1a</td>
<td>Gov SNP</td>
<td>101</td>
</tr>
<tr>
<td>3 - 1b</td>
<td>2108 SNP defines the Gov epitope</td>
<td>101</td>
</tr>
<tr>
<td>3 - 1c</td>
<td>CD109 exon containing Gov SNP</td>
<td>102</td>
</tr>
<tr>
<td>3 - 2</td>
<td>Gov genotyping by PCR-SSP</td>
<td>103</td>
</tr>
<tr>
<td>3 - 3</td>
<td>Gov genotyping by PCR-RFLP</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>Chapter 3 figure legends</td>
<td>105</td>
</tr>
<tr>
<td>4 - 1</td>
<td>Flow cytometric analysis of PAR-1-induced platelet surface antigen expression</td>
<td>128</td>
</tr>
<tr>
<td>4 - 2</td>
<td>ADP-stimulated platelet activation marker expression</td>
<td>129</td>
</tr>
<tr>
<td>4 - 3</td>
<td>Expression of activation markers on collagen-stimulated platelets</td>
<td>130</td>
</tr>
<tr>
<td>4 - 4</td>
<td>Activation marker expression on epinephrine-stimulated platelets</td>
<td>131</td>
</tr>
<tr>
<td>4 - 5</td>
<td>Activation-induced marker expression kinetics in ADP-stimulated platelets</td>
<td>132</td>
</tr>
<tr>
<td>4 - 6</td>
<td>Activation-induced marker expression kinetics in PAR-1AP-stimulated platelets</td>
<td>133</td>
</tr>
<tr>
<td>4 - 7</td>
<td>Activation-induced marker expression kinetics in A23187-stimulated platelets</td>
<td>134</td>
</tr>
<tr>
<td>4 - 8</td>
<td>Flow cytometric analysis of intracellular pool of CD109</td>
<td>135</td>
</tr>
<tr>
<td>4 - 9</td>
<td>CD109 codistributes with $\alpha$-granule and dense granule marks on linear sucrose gradient</td>
<td>136</td>
</tr>
<tr>
<td>4 - 10</td>
<td>CD109 codistributes with $\alpha$-granule marks on a discontinuous sucrose density gradient</td>
<td>137</td>
</tr>
<tr>
<td>4 - 11</td>
<td>Confocal microscopy of platelets</td>
<td>138</td>
</tr>
<tr>
<td>4 - 12</td>
<td>Detection of CD109 in resting platelets by immunogold labelling of cryosections.</td>
<td>139</td>
</tr>
<tr>
<td></td>
<td>Chapter 4 figure legends</td>
<td>140</td>
</tr>
</tbody>
</table>
ABBREVIATIONS

aa – amino acid
AA - arachidonic acid
ACD – citric acid, sodium citrate dextrose
A2M – alpha2-macroglobulin
α2MSR – α2-macroglobulin signalling receptor
α-granule – alpha granule
AC – adenylate cyclase
AML – acute myeloid leukemia, acute myelogenous leukemia
APC – antigen presenting cell
AP – activating peptide
ARC - arthrogryposis, renal dysfunction, and cholestasis (ARC) syndrome
ASA - acetylsalicylic acid
bp – base pair
BSA – bovine serum albumin
C3 – complement C3
C4 – complement C4
C5 – complement C5
CT – cytoplasmic tail
Con A - Concanavalin A
COX – cyclo-oxygenase
CWFG – cold water fish gelatin
DTS – dense tubular system
δ-granule – dense granule
ECD – extra-cellular domain
EGF – endothelial growth factor
FCS – fetal calf serum
FGF – fibroblast growth factor
FITC - fluorescein isothiocyanate
GA - glutaraldehyde
GF-Platelets – gel filtered platelets
GP – glycoprotein
GPCR - G-protein coupled receptor
GPI – glycosylphosphatidylinositol
HLA – human leukocyte antigen
5-HT – serotonin
5-HT_{2A} - 5-hydroxtryptamine 2A receptor; serotonin receptor
HHT - 12-l-hydroxy-5,8,10-heptadecatienoic acid
HPA – human platelet antigen
IP_3 - inositol trisphosphate
ITAM - immunoreceptor tyrosine-based activation motif
\( \lambda \)-granule – lysozome
LRP-\( \alpha \)2MR - low density lipoprotein-related protein-\( \alpha \)2M receptor
LRM – lipid rich microdomain
MAC – membrane attack complex
MAIPA – monoclonal antibody immobilization of platelet antigens
MAP - mitogen-activated protein
MDA – malondialdehyde
m\( \phi \) - macrophage
Mk - megakaryocyte
mAb – monoclonal antibody
MPs – microparticles
MVB – multivesicular bodies
NAIT – neonatal alloimmunethrombocytopenia
NO – nitric oxide
NSF – N-ethylmalameide-sensitive fusion protein
nt – nucleotide
OCS – open canalicular system
PAR – protease activated receptor
PAR1-AP – protease activated receptor-1 agonist peptide
PBS – phosphate buffered saline
PCR – polymerase chain reaction
PE – phycoerythrin; also known as R-phycoerythrin (R-PE)
PFA – paraformaldehyde
PGG2 - prostaglandin G2
PGH2 - prostaglandin H2
PHA – phytohemagglutinin
PKC – protein kinase C
Plt – platelets
PLC – phospholipase C
PGE1 – prostaglandin E1
PLA2 – phospholipase A2
PMA - phorbol 12-myristate 13-acetate; also known as TPA
PMP – platelet-derived microparticle
PR – platelet refractoriness
PRP – platelet rich plasma
PS – phosphatidylserine
PSI – plexin/semaphoring/integrin
PTP – post-transfusion purpura
PZP – pregnancy zone protein
RFLP – restriction fragment length polymorphism
RIA – radioimmunoassay
RT – room temperature
SDS – sodium dodecyl sulphate
SEA – staphylococcus enterotoxin A
SERCA - sarcoplasmic endoplasmic reticulum calcium ATPase
SFLLRN – serine, phenylalanine, leucine, leucine, arginine, asparagines; see PAR1-AP
SNARE – soluble NSF Attachment Protein Receptors
SNAP – soluble NSF Attachment Protein
SNP – single nucleotide polymorphism
SSP – single strand polymorphism
TEP – thioester containing protein
TGF – transforming growth factor
TPα – thromboxane-prostanoid receptor
TPA - 12-O-tetradecanoylphorbol 13-acetate; also known as PMA
TxA₂ - thromboxane A₂
TxB₂ – thromboxane B₂
VASP - vasodilator-stimulated phosphorylation
VEGF- vascular endothelial growth factor
vWF – von Willebrand Factor
CHAPTER 1: GENERAL INTRODUCTION

1.1 Preface:

When I began this work, little was known regarding the function of CD109. The molecule had been described as an ~170 kDa GPI-anchored glycoprotein expressed on endothelial cells and in a restricted fashion in the haematopoietic compartment. Our group had identified cDNAs encoding a 1445 aa GPI-anchored peptide recognized by CD109 monoclonal antibodies. Analysis of the translated sequence identified CD109 as a member of the alpha-2-macroglobulin family of thioester containing proteins. Prior to the cloning of CD109, the human A2M family comprised: pregnancy zone protein (PZP), alpha2-macroglobulin (α2M), complement C3 (C3), complement C4 (C4), and complement C5 (C5). In addition, a number of A2M/C3 molecules had been identified in a variety of phylogenetically diverse species. Among these divergent taxa, varying degrees of expansion of the A2M protein family had been observed. Based on phylogenetic analysis utilizing the 5 human A2Ms, an α2M-like molecule was proposed as the ancestral member of this protein family, with the other members arising by repeated gene duplication and subsequent divergence of function (Sottrup-Jensen, Stepanik et al. 1985; Hughes 1994; Dodds and Law 1998). The identification of CD109, and later of three additional A2M members, expanded the array of the A2M/C3 protein family in humans. In view of the central role of thioester-containing proteins in innate and adaptive immunity, clarification of the status of CD109 in the evolution and phylogeny of the A2M family was of considerable interest. In Chapter 2 of this thesis I initially focus on the elucidation of the evolutionary relationships of CD109, the other 5 fully-characterized human A2M proteins α2M, PZP, C3, C4, and C5, and three additional more recently-identified A2M
family members, using sequence analysis and a detailed comparison of the intron/exon structure of the corresponding loci. The analysis is expanded to compare CD109 to related sequences extending back to protostomes, invertebrate deuterostomes, and placazoans, thereby defining CD109 as a member of a distinct and archaic branch of the A2M phylogenetic tree.

Platelet CD109 was known to carry the antigenic determinants of the Gov alloantigen system. Production of Gov alloantibodies was implicated in a subset of alloimmune mediated platelet destruction syndromes (Kelton, Smith et al. 1990; Bordin, Kelton et al. 1997; Berry, Murphy et al. 2000). Additionally, the prevalence of Gov alloantibodies in platelet refractoriness, is equivalent to that of HPA-5 and is exceeded only by that of HPA-1 (Berry, Murphy et al. 2000). In spite of its potential significance in platelet disorders, a precise and accurate method of Gov typing did not exist. Chapter 3 of this thesis describes the studies that identify the single nucleotide polymorphism which defines the Gov phenotype. Identification of this SNP allowed us to develop 3 precise and accurate approaches to genotype the Gov system. Additionally, phylogenetic analysis was carried out to determine the original Gov allele.

Although platelet CD109 was known to carry the Gov system and the antigenic determinant was shown to be defined by a SNP, the relevance of each allele was obscure. Platelets are anucleate cells and as such their proteome is primarily synthesized in their precursor – the megakaryocyte – or in the case of some soluble secretory proteins, is acquired through endocytosis. As such, the quantity of platelet receptors and their downstream
signalling molecules is fixed. Upregulation of cell surface molecules is achieved through the exocytosis of intracellular granules which themselves bear receptors and adhesion molecules on their vesicular membranes. A hierarchy exists with respect to the type of granule released in response to an agonist. With respect to CD109, there existed confusing discrepancies in the literature regarding its expression on platelets. Our position was that CD109 was expressed on activated but not on resting platelets. Other groups insisted however, that CD109 was present in both cases. These differences may conceivably have been related to the platelet isolation method or the panel of monoclonal antibodies used. For example, CD109 may have existed on the cell surface in a conformation which masks mAb epitopes 7D1 and 8A3. To clarify these questions, Chapter 4 describes studies elucidating the quantity, subcellular localization, and regulation of expression of CD109 in platelets. My studies demonstrate the expression of CD109 in α-granules and OCS of resting platelets, and that CD109 is rapidly upregulated upon platelet activation, and resolve the discrepancy in CD109 status of resting platelets among investigators as a direct consequence of its status as an early activation marker. With knowledge of the activation kinetics of CD109, and its subcellular localization in resting platelets, combined with its predicted chemistry, we are better positioned to envisage its physiological role in haemostasis.
1.2 CD109

1.2.1 Historical description of CD109

Improving methods to enrich hematopoietic stem cells from bone marrow is of considerable interest toward treatment of haematological disorders and malignancies, as well as toward potential gene therapeutic modalities. To this end, identification of cell surface markers which are restricted to hematopoietic stem cells is crucial. In an attempt to identify novel haematopoietic stem cell markers, monoclonal antibodies were raised against the CD34\(^+\) AML KG1-derived cell line KG1a (Sutherland, Yeo et al. 1991). A monoclonal antibody designated 8A3, was isolated and shown to react predominately with 170 and 150 kDa proteins, and occasionally with a 120 kD species (Sutherland, Yeo et al. 1991). The smaller species were determined to be derivatives of the 170 kDa species. Later, the 8A3 antigen was demonstrated to be identified by additional mAbs and was designated CDw109, and finally CD109. CD109 is anchored to the plasma membrane via a glycosylphosphatidylinositol (GPI) moiety. Treatment of CD109 bearing cells with phosphatidylinositol-specific phospholipase C (PI-PLC) leads to release of the molecule from the cell surface. The cell surface presentation of CD109 is heterogeneous in some cell types, displaying differential susceptibility to PI-PLC and phosphatidylinositol-specific phospholipase D (PI-PLD) (Smith, Hayward et al. 1995), thereby suggesting variations in anchor structure such as palmitoylation of inositol (Roberts, Kim et al. 1987; Low, Stiernberg et al. 1988; Kinashi, St Pierre et al. 1995). Consistent with its anchorage to the plasma membrane through a GPI-moiety, CD109 resides within lipid rich microdomains. However, given its high solubility relative to other GPI-anchored proteins, CD109 may display a broader cell surface distribution pattern (Solomon, Mallory et al. 1998). CD109 is not directly
phosphorylated after cross-linking and does not associate with Src-like kinases or integrins (Solomon, Mallory et al. 1998).

CD109 was found to be expressed on activated, but not on resting T cells. The vast majority of freshly isolated T cells do not express CD109 (Brashem-Stein, Nugent et al. 1988; Sutherland, Yeo et al. 1991; Haregewoin, Solomon et al. 1994). Following mitogenic stimulation (SEA, Con A, or PHA) of T lymphocytes, CD109 is upregulated significantly by day 3 and is expressed in the majority of T lymphoblasts by day 6, and continues to up to day 29 after initiation (Brashem-Stein, Nugent et al. 1988; Haregewoin, Solomon et al. 1994). Stimulation of T cells through the protein kinase C pathway (via PMA + ionomycin) leads to a rise in expression within 24 hours, which continues to elevate over the following three days. CD109 is expressed on the surface of both activated CD4 and CD8 T cells, and continued expression by day 6 post activation is IL-2 dependent. Although the role of CD109 in T cells is unknown, the observation that mAb LDA1 abrogates T cell helper function, suggests that CD109 plays a role in cell-cell interactions.

In our hands, CD109 is expressed on thrombin-activated, but not on resting platelets. Although the physiological role of CD109 is largely unknown, its presence on the surface of activated platelets suggested that it may play a role in cell-cell or cell-matrix interactions during haemostasis. CD109 does not, however, dimerize with β-integrins (Sutherland, Yeo et al. 1991). The Gov alloantigen system, which has been implicated in a subset of immune mediated platelet destruction syndromes, is carried by CD109 (Kelton, Smith et al. 1990; Smith, Hayward et al. 1995; Bordin, Kelton et al. 1997).
1.2.2 Cloning of CD109

CD109 cDNAs were cloned from KG1a cells in order to facilitate the elucidation of a physiological role of CD109 (Lin, Sutherland et al. 2002). A ~170kD glycoprotein was isolated by 8A3 immunoaffinity purification and digested with endoproteinase Lys-C or Asp-N. After combination of overlapping sequences, 17 independent CD109 derived sequences were obtained and used to query EST databases. The resultant hit — rat EST R4712325 that encoded the CD109 peptide DPKSNLIQQXLSQQ — was used to probe a λ-phage Uni-ZAP HUVEC cDNA library. Subsequently, 8 independent clones comprising 2 overlapping groups were identified, 7 of which were 5’ truncations of the longest clone H6, and 1 clone which contained a longer 3’UTR designated H7. Additionally, a λ-phage Uni-ZAP KG1a library was screened using clone H6 from the previous screen. The resultant 9 independent clones were contiguous with clone H6 but contained an additional ~1.3 kb of additional 5’ sequence, and the assembled contig was designated K1.

Clone K1 comprised a 112bp 5’ UTR, 4335bp ORF, and a 300 bp 3’UTR (Figure 1 - 1). A canonical polyadenylation signal, AATAAA, was found 15 bp upstream of the poly-A tail. The H7 3’UTR was contiguous with K1 but diverged with an additional 1132 bp and contained two poly-A signals found 34 and 19bp upstream of the H7 polyA tail. CD109 antibodies were capable of detecting in vitro or in vivo transcribed/translated K1 cDNA. Additionally, cell surface expression of K1 cDNA-derived CD109 was sensitive to PI-PLC, demonstrating that K1 encoded GPI-anchored CD109 (Lin, Sutherland et al. 2002).
1.2.3 CD109 is a member of the α-2macroglobulin/C3/C4/C5 family

Analysis of the amino acid sequence identified CD109 (Figure 1-2) as a novel member of the alpha2-macroglobulin/C3/C4/C5 family of plasma proteins (Lin, Sutherland et al. 2002). The defining structural feature of this class of proteins is the presence of an intrachain thioester bond between the sulfhydryl group of cysteine and the carbonyl group of glutamine (Figure 1-2) in the sequence –CGEQ– which lies about 2/3 of the way along the pro-molecule (Chu, Howard et al. 1994; Chu and Pizzo 1994). In its native form, this thioester is unreactive except with small nucleophiles such as methylamine, which are able to permeate the core of the protein and directly attack the thioester (Dodds, Ren et al. 1996; Dodds and Law 1998). However, when the protein is proteolytically cleaved, at a site upstream of the thioester, a conformational change ensues leading to the exposure of the thioester to the exterior of the protein (Dodds, Ren et al. 1996; Dodds and Law 1998). The liberated thioester motif is now available for nucleophilic attack, culminating in the covalent crosslinking to target.

1.2.4 Functional attributes of α2-macroglobulin

Alpha2-macroglobulin (A2M) is primarily a protease inhibitor which acts as a scavenger in the plasma. Most prototypical protease inhibitors inactivate proteases by nullifying directly the catalytic domain in an irreversible manner. In contrast to most protease inhibitors which are specific to a particular protease or class of proteases, A2M is capable of inhibiting any protease. The promiscuity of A2M is due to the presence of the bait region, which contains cleavage sites for all 4 mechanistic classes of proteases (Chu and Pizzo 1994). The vast majority of A2Ms are homodimeric or homotetrameric, and protease inhibition is achieved
through a non-covalent mechanism. Essentially, when the A2M is cleaved in the bait region, a conformational change ensues resulting in the entrapment of the protease in a cage composed of the A2M subunits – similar to the Venus fly trap. The entrapped protease is incapable of interacting with most proteins. However, given that the catalytic domain is intact, small proteins that can permeate the cage remain susceptible to degradation. Consequently, A2Ms have evolved a mechanism to limit the degradation of small peptides. The protease induced A2M conformational change leads to the exposure of a binding epitope for the low density lipoprotein-related protein-α2M receptor (LRP-α2MR) and the α2-macroglobulin signalling receptor (α2MSR). LRP-α2MR and α2MSR are displayed on the surface of phagocytic cells which in turn remove the A2M-protease complex from circulation (Misra, Chu et al. 1994; Misra, Chu et al. 1994). This mechanism of A2M-protease complex removal is particularly important for monomeric A2M. Monomeric A2Ms inhibit proteases in a covalent manner. Only amide-containing molecules are sufficiently nucleophilic to directly attack the thioester of A2Ms (Figure 1 – 3). As such, monomeric A2Ms bind to the attacking protease through an amide bond, and the complex is rapidly removed from circulation by receptor mediated endocytosis in phagocytic cells.

1.2.5 Complement C3, C4, and C5

The complement system is a complex cascade of proteins and enzymes that participate in innate and adaptive immunity (Figures 1 – 5 and 1 - 6). Essentially, their primary role is in defence against invading microorganisms. In higher vertebrates the complement cascade is activated primarily via three distinct but overlapping means: [1] the classical pathway, [2] the
alternative pathway, and [3] the lectin pathway. The classical pathway is comprised of a successive cascade of proteolytic complexes (Figure 1-5). At each step, the activated substrate binds to the activating enzyme and changes its specificity, or binds near the activating enzyme and serves as a cofactor for modification of the next substrate. In the classical pathway, C4 is cleaved into C4b and C4a, by C1 (C1 is the classical pathway activator), and C4b subsequently binds to C1 through its thioester. The modification imparted on C1 by C4 allows C1 to activate C2 which together with C4 forms the C3 convertase – C4b2a – which cleaves C3 into C3b and C3a. Complement C3 activation via C3 convertase leads to C3b deposition and covalent attachment to cellular targets, or to the immune complexes on which they are activated. C3 binds to C2a of the C3 convertase to form C4b2a3b – the C5 convertase. The classical pathway culminates in the formation of the membrane attack complex C6-C9 (MAC) which leads to lysis of susceptible cells. The lectin pathway and classical pathway are distinguished from one another by their respective activating molecules: C1 and MBL-MASP complex, respectively. The MBL-MASP complex is formed as a result of the recognition of the large highly repetitive polysaccharides of microbial cell walls. In contrast to the classical and lectin pathways, the alternative pathway proceeds as a result of the natural turnover of C3 rather than by the recognition of a foreign derived activator (Figure 1-6). Recall, that the thioester of C3 is susceptible to direct nucleophilic attack by small nucleophiles, including water. Hydrolyzed C3 (C3(H2O)) has a conformation similar to that of proteolytically activated C3, and C3(H2O), together with factor D activated factor B, forms a soluble C3 convertase. C3(H2O)B converts more C3 into C3b, most of which is rapidly inactivated due to its short half-life. Some C3b deposition may occur, however, if there is a susceptible surface in close proximity. Any C3b
deposited on healthy host tissue is rapidly degraded. Complement C3 is the central component of the complement cascade and is likely phylogenetically ancestral to C4 and C5.

Inflammation and haemostasis are intimately linked. It has been known for a number of years that complement C3 and C5-9 can augment thrombin-induced platelet secretion and aggregation (Polley and Nachman 1978; Polley and Nachman 1983); and C3a can induce platelet activation and aggregation (Polley and Nachman 1983). Developing evidence suggests that platelets are capable of initiating the complement cascade in the absence of traditional inducers of complement. Activated platelets have been shown to bind C1q, C3, C4, and C5-C9 (Del Conde, Cruz et al. 2005; Peerschke, Yin et al. 2006; Gushiken, Han et al. 2009; Hamad, Nilsson et al. 2010). Additionally, complement degradation products such as C3a, C4d, and C5a have also been identified under these conditions (Peerschke and Ghebrehiwet 1994; Del Conde, Cruz et al. 2005; Peerschke, Yin et al. 2006; Gushiken, Han et al. 2009; Hamad, Nilsson et al. 2010). Strikingly, thrombin is capable of acting as a C5 convertase, in the absence of C3 (Huber-Lang, Sarma et al. 2006). As such, thrombin likely contributes to activation of complement during haemostasis. Further, C3 deficient mice have extended bleeding times and a PAR-4 induced-aggregation defect (Gushiken, Han et al. 2009). Taken together, haemostatic and complement components reciprocally activate one another.

Upon activation, platelets release chondroitin sulfate (CS), from their storage granules, which coats the outer leaflet. C1q binds to the platelet membrane through the CS coat (Hamad, Nilsson et al. 2010). As a result, surface bound C1q, allows for the assembly of C1q, C1r, and C1s, and activation of the classical complement pathway. The platelet itself is
protected from complement activity via decay accelerating factor expression and secretion of Factor H (Hamad, Nilsson et al. 2010). Essentially, under healthy physiology, the platelet induced complement cascade occurs in the fluid phase, and is restricted to the vicinity of the adherent activated platelets.

1.2.6 CD109 shares features of both Complement C3 and α2M

Based on overall sequence similarity, CD109 is most closely related to α2M (40 – 45% overall similarity to vertebrate and invertebrate A2Ms) and consistent with this fact a bait region is found ½ way along the promolecule (Lin, Sutherland et al. 2002). However, the reactivity of CD109 is likely similar to that of C3 rather than α2M given that the thioester reactivity determining triplet is VIH rather than LLN. The imidazole of His imparts on the thioester an ability to follow a catalysed transacylation mechanism that culminates in the formation of ester bonds with targets. As a result, CD109 akin to complement C3 and the C4B isotype, is capable of forming covalent bonds with both proteins and carbohydrates (amide and carboxyl containing molecules), whereas the covalent specificity of α2M is restricted to proteins (amide containing molecules).

CD109 was the first monomeric A2M-like protein to be identified in humans. Human α2M is a homotetramer of two disulfide linked dimers of 180kD subunits which together act as a “molecular Venus fly trap” that removes proteases and cytokines from circulation. In multimeric A2Ms, the thioester does not bind directly to targets but instead binds to adjacent subunits, thereby holding the cage tightly in place (Dolmer, Husted et al. 1996). Monomeric A2Ms, however, bind targets directly through their thioester, akin to C3 and C4, and
occasionally independent of the thioester, akin to C5. A2M-target complexes are removed from circulation via receptor-mediated endocytosis. In contrast to α2M, CD109 does not contain epitopes for LRP-α2MR or α2MSR. As such, signalling downstream of CD109 ligation as well as CD109-target complex removal must proceed through a novel mechanism.

1.3 Platelet physiology

Platelets are anucleate cells that survey the vasculature for signs of damage. These small cells are present in blood at 150 – 300 X 10^9 L^-1, with dimensions of 0.5 X 3 μm, and have a lifespan of approximately 7 days (Pujol-Moix, Hernandez et al. 2000). The primary function of platelets is to act as the cellular effector of haemostasis (Figure 1–7). Platelets utilize cell integrins to detect changes in the integrity of vessel walls (Figure 1–8). A number of internal and external controls maintain platelets in a resting state during circulation and prevent intracellular signals from inappropriately activating the integrins. Such controls involve tight regulation of cytosolic Ca^{2+} levels, utilization of intracellular phosphatases that limit signalling through kinase-dependent pathways, and the presence of extracellular ADPases that hydrolyze released ADP (Shattil and Newman 2004).

Haemostasis is a delicate balance between procoagulant and anticoagulant mechanisms. Blood fluidity is maintained by the action of endothelial cell nitric oxide and prostacyclin, which prevent platelet stasis and dilate intact blood vessels. However, when the vascular endothelium is disrupted these mediators are no longer synthesized. This in turn, results in conditions which promote adhesion of platelets to the exposed substratum. At sites of vessel injury, platelets come into contact with components of the extracellular matrix, such as collagen, vWF, fibronectin, and laminin, which are usually excluded from the bloodstream.
of intact blood vessels. Platelets adhere to the site of vessel damage, change shape, and secrete granular contents. The proteins and small molecules thereby secreted by adherent platelets lead to the recruitment and activation of additional platelets which are flowing past the adherent platelets, which then aggregate to the growing platelet plug (Figure 1-8; Table 1-1; Table 1-2). Platelets also provide a surface for the assembly of coagulation factors.

The following section will describe the manner in which the platelet ultrastructure, receptors, and ligands contribute to platelet effector functions.

**1.3.1 Platelet Surface Receptors**

**1.3.1.1 GPIb-IX-V**

GPIb-IX-V, is expressed constitutively on platelet membranes at 25,000 copies per cell (Modderman, Admiraal et al. 1992). As von Willebrand Factor (vWF) is GPIb-IX-V’s main ligand, it is known as the vWF receptor. GPIb-IX-V can also bind, however, to thrombin, high molecular weight kininogen, factor XI, FXII, Mac-1 on neutrophils, and P-selectin (CD62P) on activated platelets and endothelial cells. The binding sites for all named ligands are near the N-terminus of the GPIbα subunit.

vWF is synthesized in megakaryocytes and endothelial cells, and is packaged in α-granules and Weibel-Palade bodies respectively (Ruggeri 1999; Schmugge, Rand et al. 2003). Mature vWF has an apparent molecular weight of 250kDa and circulates in the blood as a dimer, where it behaves as a carrier of Factor VIII. In contrast, multimeric vWF is generally excluded from circulation, and is associated with subendothelial matrix components. The immobilization of multimeric vWF onto the subendothelial matrix, imparts a conformational
change onto vWF that permits its binding to GPIb. At sites of vessel wall damage, immobilized vWF is exposed to circulation and becomes available for interaction with GPIb-IX-V (and GPIIb/IIIa at low shear), leading to platelet adhesion, secretion, activation of GPIIb/IIIa, and aggregation.

The vWF receptor is a complex of GPIbα, disulphide-linked GPIbβ, non-covalently associated GPIX and GPV in a 2:2:2:1 ratio (Ware 1998). A subset of these vWF receptors is associated with lipid rafts (lipid rich microdomain; LRM) in resting platelets, and following platelet activation, additional GPIb-IX-V is recruited to the LRM (Shrimpton, Borthakur et al. 2002). LRM provide a platform for the assembly of receptors and signalling molecules such as Src kinases and Lyn, which have been implicated in vWF induced signalling (Wu, Asazuma et al. 2003).

GPIb-IX-V mediated platelet adhesion is crucial to haemostasis. GPIb-IX-V:vWF binding can occur even at high shear within arteries and arterioles, whereas platelet adhesion mediated through integrin αIIbβ3 (GPIIb/IIIa) occurs exclusively at low shear in larger arteries and veins. The vital role of GPIb-IX-V in haemostasis is highlighted in bleeding diatheses associated with disorders lacking the receptor (Bernard Soulier syndrome) or the ligand (von Willebrand disease). At sites of vascular injury, adherent platelets recruit additional cells which participate in a growing thrombus through fibrinogen-mediated crosslinking of GPIIb/IIIa on adjacent platelet membranes. The multimeric nature of vWF imparts an ability to simultaneously bind more than one platelet. As a result, at conditions of high shear, GPIb-IX-V:vWF interaction rather than GPIIb/IIIa:fibrinogen interaction, is the primary driver of thrombus formation.
Engagement of GPIb-IX-V induces platelet cytoskeletal reorganization (Yuan, Dopheide et al. 1997; Cranmer, Ulsemer et al. 1999; Torti, Bertoni et al. 1999; Yuan, Kulkarni et al. 1999) and rearrangement of associated signalling molecules. Essentially, the cytoskeleton collects and organizes signalling molecules. Conversely, ligand binding to GPIb-IX-V is regulated by cytoskeleton. The cytoplasmic domain of GPIbα is essential for vWF-induced cytoskeletal re-organization (Canobbio, Lova et al. 2002) association with actin and is critical for stable adhesion at high shear (Cranmer, Ulsemer et al. 1999).

1.3.1.2 GPIIb/IIIa

Integrin α_{IIb}β_{3}, also known as GPIIb/IIIa or CD41/CD61, is the most abundant platelet receptor. There are approximately 40 000 – 80 000 α_{IIb}β_{3} molecules and they are densely distributed (<200 Å apart) throughout the resting platelet surface (Wagner, Mascelli et al. 1996). Mature α_{IIb} is composed of 1008 amino acids and forms a heterodimer with the 762 amino acid β_{3}. Integrin α_{IIb}β_{3} serves as the major fibrinogen receptor but can also bind fibronectin, vitronectin, thrombospondin and vWF. The multiplicity of α_{IIb}β_{3} targets is a consequence of an Arg-Gly-Asp (RGD) domain present in each aforementioned ligand (Ruoslahti 1988; Gould, Polokoff et al. 1990; Ojima, Chakravarty et al. 1995). In resting platelets, the affinity of ligands to α_{IIb}β_{3} is quite low, after platelet activation the integrin undergoes a conformational change leading to its activation. This conversion to a high affinity state is mediated by a process known as inside-out signalling. Low affinity α_{IIb}β_{3} is capable of binding immobilized fibrin(ogen) but not soluble fibrin(ogen) whereas high-affinity integrin can bind both.
As with all integrins, the fibrinogen receptor is composed of an α/β heterodimer. Our knowledge of low- and high-affinity α_{IIb}β_3 structure is largely based on the resolution of the crystal structure of integrin α_β_3 and portions of the extracellular domain of α_{IIb}β_3 (Xiao, Takagi et al. 2004). Ligand binding occurs within the extracellular domain of α_{IIb}β_3 (Xiong, Stehle et al. 2001; Xiong, Stehle et al. 2002; Xiao, Takagi et al. 2004). An N-terminal β-propeller domain, the thigh domain, and two calf domains define the extracellular domain (ECD) of α_{IIb} (Xiong, Stehle et al. 2001; Xiong, Stehle et al. 2002). The β_3 ECD consists of an A domain, a PSI (plexin/semaphorin/integrin) domain, a hybrid domain, four EGF domains, and a membrane-proximal βTD domain (Xiong, Stehle et al. 2002). The seven-bladed β-propeller of α_{IIb} interacts with the β_3A domain to form an ellipsoid head. The cytoplasmic domain is a complex of the α_{IIb} (residues 989–1008) and β_3 (residues 716–762) cytoplasmic tails (CTs).

Activation of α_{IIb}β_3 is contingent on the interaction of specific proteins at the cytoplasmic tail, which in turn, transmit signals to the extracellular domain. It is proposed that interaction between the CTs of α_{IIb} and β_3 maintain the low-affinity α_{IIb}β_3 conformation in resting platelets (Vinogradova, Velyvis et al. 2002). Conversely, activation through agonist induction of G-protein coupled receptors (GPCRs), or adhesion induced unclasping of the α_{IIb}-CT and β_3-CT interface, converts α_{IIb}β_3 into its high-affinity state (O’Toole, Katagiri et al. 1994; Hughes, Diaz-Gonzalez et al. 1996; Vinogradova, Velyvis et al. 2002; Podolnikova, O’Toole et al. 2009). The α_{IIb}β_3 cytosolic tail (CT) lacks intrinsic enzymatic activity. As such, initiation and propagation of bidirectional signalling events across α_{IIb}β_3 are dependent on cytoplasmic proteins that bind the α_{IIb}β_3 CT. The vast majority of the >20 identified CT binding proteins have been linked to outside-in signalling (Luo, Carman et al. 2005). Following platelet
activation, the formerly uniform distribution pattern of the cytoplasmic protein, Talin-1, becomes localized to the plasma membrane (Beckerle, Miller et al. 1989). Talin-1 plays a critical role in integrin mediated focal adhesion formation and has been demonstrated to interact with the α\textsubscript{IIb}β\textsubscript{3} CT (Rees, Ades et al. 1990; Pavalko, Otey et al. 1991; Burridge and Chrzanowska-Wodnicka 1996; Knezevic, Leisner et al. 1996; Calderwood, Zent et al. 1999; Calderwood, Yan et al. 2002; Vinogradova, Velyvis et al. 2002; Garcia-Alvarez, Bobkov et al. 2003; Garcia-Alvarez, de Pereda et al. 2003; Tadokoro, Shattil et al. 2003; Ulmer, Calderwood et al. 2003; Wegener, Partridge et al. 2007). It is proposed that Talin-1 binding to α\textsubscript{IIb}-CT and β\textsubscript{3}-CT induces α\textsubscript{IIb}β\textsubscript{3} activation by disrupting the salt bridge between the two tails. Current data suggest that Talin-1 and another cytosolic protein, Kindlin-3, act in concert to activate integrin α\textsubscript{IIb}β\textsubscript{3} and are critical for functional inside-outside signalling (Banno and Ginsberg 2008; Larjava, Plow et al. 2008; Moser, Nieswandt et al. 2008). The current working model of integrin activation is the so-called switch-blade hypothesis in which the extended conformation represents the active state and the bent form is the resting state. Essentially, disruption of α\textsubscript{IIb}-CT and β\textsubscript{3}-CT interaction results in straightening of the extracellular region of the integrin and exposure of the ligand binding site.

1.3.1.3 Collagen Receptors

Platelet-collagen interactions are believed to be most important at medium and high shear forces within arteries and diseased vessels. The main collagen receptors are integrin α\textsubscript{2}β\textsubscript{1} which is involved primarily in adhesion and GPVI which is the collagen signalling receptor. In addition, collagen can indirectly bind to GPIIb/IIIa and GPIb-IX-V through VWF. Integrin α\textsubscript{2}β\textsubscript{1}
mediated platelet adhesion has been demonstrated through collagens type I, II, III, IV, and VI, and except for type IV, these collagens can induce platelet aggregation. Surface expression of $\alpha_2\beta_1$ at 2000 copies per platelet, is considerably lower than that of $\alpha_{IIb}\beta_3$ (Kunicki, Orzechowski et al. 1993). Conversion of either $\alpha_2\beta_1$ or $\alpha_{IIb}\beta_3$ to a high affinity state is required for transition from rolling to stable adhesion. By extension, $\alpha_2\beta_1$ requires GPIIb/IIIa for activation in flow (Van de Walle, Schoolmeester et al. 2007).

GPVI is a member of the immunoglobulin family of receptors and is coupled to the FcR $\gamma$-chain that contains ITAMs 339 a.a. 60-65 kDa. The receptor comprises 2 Ig-2C-like extracellular domains formed by disulfide bonds, a mucin like stalk, a transmembrane domain, and a short 51 amino acid cytoplasmic domain. Positively charged arginine in the transmembrane region is responsible for interaction with FcR $\gamma$-chain (Zheng, Liu et al. 2001; Berlanga, Tulasne et al. 2002). The assembly and initiation of the signalling cascade takes place in LRMIs (Ezumi, Kodama et al. 2002; Locke, Chen et al. 2002; Wonerow, Obergfell et al. 2002). A proline rich cytoplasmic region binds SH3 domain of Src family tyrosine kinases, fyn and lyn (Suzuki-Inoue, Tulasne et al. 2002). In addition, the cytoplasmic domain also contains calmodulin binding sites. Calmodulin is constitutively associated with GPVI and dissociates following platelet activation.

1.3.1.4 Protease Activated Receptors

Haemostasis is the process by which vascular, platelet, and plasma factors arrest bleeding from an injured blood vessel. This physiological process is a delicate balance between procoagulant and anticoagulant mechanisms. By extension, corruption of these
regulatory mechanisms results in haemostatic abnormalities such as bleeding diathesis and thrombosis.

Formation and dissolution of a blood clot are tightly regulated physiological processes. Activation of haemostasis initiates a sequential interaction and proteolytic modification of a retinue of proenzymes, enzymes, and inhibitors of the coagulation and fibrinolysis cascades. Historically, the coagulation cascade has been viewed as consisting of two distinct pathways: [1] the intrinsic pathway and [2] the extrinsic pathway. In reality, these two pathways are closely intertwined in vivo and on their own are simply convenient descriptors of in vitro haemostatic deficiency diagnostic tools. Further, it is self-evident that haemostasis usually follows the extrinsic pathway. Essentially, at sites of vessel wall damage platelets come into contact and are activated by VWF and collagen while the coagulation cascade is initiated by the interaction of plasma factor VII/VIIa (FVII/FVIIa) with extravascular tissue factor (TF). Subsequently, propagation of the thrombus proceeds through the extrinsic pathway recruiting sequential coagulation factors. This amplification and progression is mediated by platelets which provide a thrombogenic surface. The historical intrinsic and extrinsic factors converge to form an activation complex which converts prothrombin to thrombin. Thrombin, in turn, converts fibrinogen to fibrin – the main component of a blood clot/platelet plug – which acts as a platelet cross-linker through integrin $\alpha_{\text{IIb}}\beta_3$. A variety of mechanisms exist to terminate clot formation including the inactivation of thrombin by anti-thrombin III (ATIII).

In addition, to its function as a fibrinogen convertase, thrombin is capable of inducing shape change, degranulation, and aggregation of platelets. Thrombin binds GPIb$\alpha$ and a
unique class of seven-transmembrane domain receptors (G-protein coupled receptors; GPCRs) termed Protease Activated Receptors (PARs). There are four types of protease activated receptors: the prototypical PAR-1, PAR-2, PAR-3, and PAR-4. PAR-2 is primarily a trypsin receptor whereas the other PARs are thrombin receptors. Human platelets display about 1000-2000 PAR-1 per platelet (Brass, Vassallo et al. 1992; Norton, Scarborough et al. 1993). PAR-1 is the predominant thrombin receptor in human platelets whereas PAR-4 is responsive primarily at higher thrombin concentrations (>10 nm whereas PAR1 is activated at picomolar levels). Additionally, GPIbα serves as a cofactor for PAR-1 (and presumably PAR-4) amplifying the reaction 5-7 fold (De Candia, Hall et al. 2001).

Proteolytic cleavage (by thrombin, for example) of the PAR ectodomain, at a single site, generates a new N-terminus which forms a tethered ligand by binding intramolecularly with the receptor, inducing transmembrane signalling (Norton, Scarborough et al. 1993). A subsequent conformational change allows PAR interaction with heterotrimeric G proteins. PAR-1 can couple to G_\text{ai}, G_\text{aq}, and G_\text{a12/13}, while PAR-4 can associate only with the latter two molecules (Coughlin 1999; Faruqi, Weiss et al. 2000; Macfarlane, Seatter et al. 2001). Consistent with heterotrimeric G protein coupled signalling, thrombin induces activation of mitogen-activated protein (MAP) kinases, intracellular Ca^{2+} mobilization, RhoGEF-mediated Rho and Rac signalling, and other cellular effectors (Papkoff, Chen et al. 1994; Wang, Sada et al. 1994; Banno, Nakashima et al. 1995; Hartwig, Bokoch et al. 1995; Saklatvala, Rawlinson et al. 1996; Heijnen, Oorschot et al. 1997). This thrombin-induced signalling cascade promotes diverse cellular responses. In the context of platelet activation, PAR coupling to G_\text{aq} is required for platelet secretion and aggregation, G_\text{a12/13} contributes to shape change, and G_\text{ai} synergizes
signalling by other platelet agonists. Calcium influx is indispensable for platelet degranulation, consistent with liberation of inositol trisphosphate (IP$_3$) and diacylglycerol (DAG) downstream of the activated phospholipase C$_{\beta}$ (PLC$_{\beta}$) isoforms of G$_{aq}$ coupled GPCRs.

While proteolytic activation of PARs is irreversible, there exist, however, several mechanisms which terminate their downstream signals. Phosphorylation and adaptor protein binding play a crucial role in the desensitization of PAR1 and PAR2. G-protein coupled receptor kinases (GRKs), and in some instances, the secondary messengers, protein kinase A and protein kinase C, phosphorylate specific serine and threonine residues within the PAR cytoplasmic tail (PAR-CT). Phosphorylation, at these sites, increases the affinity of the PAR-CT for the adaptor protein $\beta$-arrestin, while concurrently uncoupling the G-protein:PAR interaction, and preventing further G-protein binding (Lohse, Benovic et al. 1990; Krupnick and Benovic 1998). This action essentially prevents further PAR signal transduction. Uncoupling of PARs can also be mediated by GRKs independent of phosphorylation (Ferguson 2007). In platelets, which inherently lack appreciable de novo protein synthesis, PAR responsiveness (i.e. PAR desensitization) is likely a function of the finite amount of secondary messenger generated following PAR activation (Ishii, Hein et al. 1993).

In nucleated cells, trafficking of naïve PARs occurs in both directions, whereas desensitized PARs are internalized in a clathrin-coat-mediated manner, and are directed to lysozomes for degradation. In platelets, however, the vast majority of cleaved PAR-1, although rapidly desensitized, remains on the surface (Molino, Bainton et al. 1997). In contrast, desensitization of PAR-4 is slow and prolonged signal is observed. Nucleated cells,
such as fibroblasts and endothelial cells, possess an intracellular reserve of PARs, that allows for the replenishment of cell surface PAR independent of de novo protein synthesis (Hein, Ishii et al. 1994; Bohm, Khitin et al. 1996), whereas no such pathway exists in platelets.

1.3.1.5 ADP receptors: P2Y₁, P2Y₁₂

ADP is considered a weak agonist, since on its own, ADP induced aggregation, unlike that seen by collagen or thrombin stimulation, is reversible. However, ADP activity is crucial for efficient haemostasis, due to the high amount of ADP present in dense-granules. ADP induced paracrine and autocrine effect on platelets following strong agonist induced degranulation is crucial to haemostasis. Platelets possess two G-protein coupled ADP receptors: P2Y₁ and P2Y₁₂. P2Y₁ is coupled to G₉ and initiates shape change, calcium mobilization and aggregation (Hechler, Eckly et al. 1998; Jin, Daniel et al. 1998; Savi, Beauverger et al. 1998), while P2Y₁₂ is coupled to Gαi and is indispensible for the formation and stabilization of large aggregates. Essentially, P2Y₁₂ is required for the completion of events initiated by P2Y₁. The P2Y₁₂ receptor plays a central role in the amplification of aggregation induced by collagen, thrombin, ADP, serotonin, thromboxane, or epinephrine, regardless of the signalling pathway downstream of the initial agonist. P2Y₁₂ mediated intracellular amplification proceeds through inhibition of cAMP synthesis, vasodilator-stimulated phosphoprotein (VASP) dephosphorylation, phosphoinositide-3-kinase (PI3K) and small GTPase Rap1B activation. There are about 150 P2Y₁ sites per platelet (Baurand, Raboisson et al. 2001). The indispensible nature of P2Y₁₂ is evidenced by the prolonged bleeding times and strong
inhibition of aggregation (at low to medium agonist dose) observed in P2Y_{12} deficient murine or pharmacologically P2Y_{12} antagonized human platelets (Foster, Prosser et al. 2001).

ADP is a weak agonist as a consequence of low expression of its receptors on the cell surface relative to other GPCRs (150/platelet versus 1000-2000/platelet). In fact, overexpression of P2Y_1 leads to degranulation comparable to that observed for thrombin (Hechler, Zhang et al. 2003). P2Y_{12} is directly implicated in the exposure of phosphatidylserine at the activated platelet membrane surface while P2Y_1 is not (Storey, Sanderson et al. 2000; Leon, Freund et al. 2001; Leon, Ravanat et al. 2003).

1.3.1.6 Thromboxane receptor

Agonist-induced activation of platelets leads to thromboxane synthesis via cyclooxygenase dependent metabolism of arachidonic acid. Phospholipase A2 catalyzes arachidonic acid release from membrane phospholipids. Consequently, prostaglandin endoperoxides, prostaglandin G2 (PGG_2) and prostaglandin H2 (PGH_2) are synthesized via COX activity. Thromboxane synthetase converts PGH_2 to thromboxane (TXA_2) (Needleman, Minkes et al. 1976; Needleman, Moncada et al. 1976) 12-l-hydroxy-5,8,10-heptadecatrienoic acid (HHT), and malondialdehyde (MDA) (Haurand and Ullrich 1985; Shen and Tai 1986). Secreted TxA_2 then serves as a platelet agonist, signalling through thromboxane-prostanoid receptor (TP\alpha). TP\alpha is expressed on the cell surface at a density of \sim 1000 – 2000/platelet. TP\alpha is associated with G_q and G_{q12/13} and initiates platelet activation through the PLC\beta1 pathway. In cases in which a GPCR can couple to more than one type of G-protein, the concentration of agonists determines which species will couple. Through G_{q13}, TP\alpha modulates Na^+/H^+
exchange. TxA₂ acts in a localized area since it is extremely unstable and is readily degraded non-enzymatically to its inactive form, TxB₂.

Acetylsalicylic acid (ASA), also known as aspirin, is routinely used by as much as 36% of the adult population in the US for primary and secondary prevention of cardiovascular disease. ASA inhibits platelet activity through irreversible acetylation of Ser⁵²⁹ in COX-1. This in turn, prevents binding of arachidonic acid to the catalytic site, thus blocking Prostaglandin H₂ (PGH₂) synthesis. Consequently, TxA₂ generation and TxA₂-induced platelet aggregation are inhibited for the lifespan of the platelet. ASA treated platelets remain susceptible to shear, thrombin ADP, and epinephrine-induced activation.

1.3.1.7 Epinephrine receptor, α₂A⁻adrenoreceptor

α₂-Adrenoceptors are cell-surface catecholamine receptors widely expressed in the central and peripheral nervous systems. A broad spectrum of physiological functions are attributed to α₂-adrenoceptor ligation, including inhibition of neurotransmitter release, blood pressure regulation, sedation, analgesia, insulin release, lipolysis, renal function, and multiple behavioural and cognitive functions (Ruffolo 1993; Ruffolo, Nichols et al. 1993; Lakhiani, MacMillan et al. 1997; Rohrer, Bernstein et al. 1998; Sallinen, Haapalinna et al. 1998; Altman, Trendelenburg et al. 1999; Beversdorf, Hughes et al. 1999; Hein, Limbird et al. 1999; Hein, Silva et al. 1999; Sallinen, Haapalinna et al. 1999; Willems, Trion et al. 1999). At a cellular level, α₂-adrenoceptor-mediated signalling leads to adenyl cyclase inhibition, activation of K⁺ channels, abrogation of voltage-gated Ca²⁺ channels, phospholipase C activation, intracellular Ca²⁺ release and mitogen activated protein kinase (MAP kinase) activation (Eason and Liggett
The human genome encodes three distinct $\alpha_2$-adrenoceptors, designated: $\alpha_{2A}$, $\alpha_{2B}$ and $\alpha_{2C}$, located on chromosomes 10, 2, and 4, respectively. The predominant isoform in human blood platelets is the $\alpha_{2A}$-adrenoceptor, and its primary agonist is epinephrine.

Epinephrine, also known as adrenalin, 3,4-Dihydroxy-$\alpha$-(methylaminomethyl)benzyl alcohol, 4-[1-Hydroxy-2-(methylamino)ethyl]-1,2-benzenediol, and (±)-3,4-Dihydroxy-$\alpha$-(methylaminomethyl)benzyl alcohol, is a small molecule synthesized by the medulla of the adrenal gland. The molecule was aptly named epinephrine since it was isolated from the adrenal gland which lies above the kidney. Epinephrine secretion is a component of the fight-or-flight response where it acts as a sympathetic catecholamine, accelerating the heartbeat, strengthening cardiac contraction, dilating bronchioles, and mediating numerous additional effects. Historically, epinephrine on its own, is not considered an inducer of platelet aggregation or activation, but instead an enhancer of agonist (thrombin, collagen, ADP, TxA2) induced responses.

$\alpha_2$-adrenoceptors couple to the $G_i$-$G_o$ family of G proteins. Recently, data from $G_z$ protein knockout mice suggest that $\alpha_{2A}$-adrenoceptors prefer $G_{az}$ to other $G_i$ family members (Yang, Wu et al. 2000). Epinephrine activates $G_{az}$, which decreases cAMP levels through inhibition of the adenylate cyclase pathway, and activation of PLC thereby promoting aggregation. Although coupled to $G_{az}$, epinephrine induced platelet effects are independent of the inhibition of cAMP synthesis (Daniel, Dangelmaier et al. 1999; Nakahashi, Kambayashi et al. 2001). It has been shown that adenylyl cyclase inhibition did not restore the response of platelets to epinephrine in $G_{az}$ knockout mice (Yang, Wu et al. 2000).
Epinephrine alone does not cause aggregation. Stimulation of platelets exclusively through $G_{\alpha i}$ or $G_{\alpha12/13}$ does not lead to platelet aggregation. For example, when platelets are stimulated through TP$\alpha$ with low dose TxA$_2$ analogue – conditions in which TP receptors couple exclusively with $G_{\alpha12/13}$ – platelet shape change ensues without elevation calcium levels or aggregation. However, at higher TxA$_2$ analogue doses, TP receptors couple to both $G_{\alpha12/13}$ and $G_{\alpha q}$, leading to platelet shape change, elevation of intracellular calcium and subsequent platelet aggregation. In contrast, when platelets are stimulated simultaneously with low dose TxA$_2$ analogue and epinephrine (or ADP through P2Y$_{12}$ $G_{\alpha i}$) platelets undergo shape change and aggregation. Similar results are observed when platelets are stimulated through PAR1 under $G_{\alpha12/13}$ exclusive conditions, suggesting that $\alpha_2$-adrenergic receptor signals are capable of inducing platelet aggregation when combined with signalling through a $G_{\alpha12/13}$ coupled receptor.

Epinephrine stimulation is synergistic with AA in a PLC/Ca$^{2+}$, MAPK, and cyclooxygenase pathway dependent manner. Additionally the epinephrine/AA synergism is negatively modulated by NO generation in platelets. Subthreshold epinephrine synergises with PLC activity, leading to the formation of IP$_3$ and an associated spike in intracellular calcium levels. Intracellular calcium inturn stimulates PLA$_2$ and COX-1, thereby leading to the synthesis of TxA$_2$.

1.3.1.8 Serotonin receptors

Serotonin (5-hydroxytryptamine, 5-HT; C$_{10}$H$_{12}$N$_{2}$O) is produced by a subset of neurons, and is stored in the pineal gland, digestive tract, the brain, and platelets. Numerous roles have
been attributed to serotonin in the central nervous system. Outside of the CNS, serotonin is instrumental in regulating vascular resistance, blood pressure, haemostasis and platelet function (Kaumann and Levy 2006). Platelets actively take up serotonin from plasma via receptor-mediated endocytosis and package the hormone in dense granules (Ni and Watts 2006). Following platelet activation and degranulation, secreted serotonin promotes platelet aggregation as well as vasoconstriction of blood vessels. This liberated serotonin provides a positive feedback, amplifying platelet responses by stimulating shape change and recruiting additional platelets to the site of vessel damage. The action of serotonin on platelets is transduced through the \( G_q \) coupled 5-hydroxtryptamine 2A (5-HT\(_{2A}\)) receptor.

1.3.1.9 Human platelet antigens

Human platelet antigens (HPA) are biallelic polymorphisms of glycoproteins expressed on the platelet surface. These HPAs are capable of stimulating the production of alloantibodies in recipients of mismatched blood product. The generated alloantibodies, in turn, can lead to immune-mediated destruction of platelets. Alloantibodies directed against HPA-1, -2, -3, -4, -5, and more recently Gov are considered to be the most clinically relevant. CD61 (GPIIIa; integrin \( \beta_3 \)) carries the HPA-1 and HPA-4 epitopes, CD42b\( \alpha \) (GPIb\( \alpha \)) carries HPA-2, CD49b (GPla; integrin \( \alpha_{IIb} \)) carries HPA-5, and CD41 (GPIIb; integrin \( \alpha_{IIa} \)) carries HPA-3. The Gov alloantigen system is a co-dominant biallelic platelet alloantigen carried by CD109 (Kelton, Smith et al. 1990). The prevalence of Gov alloantibodies among neonatal alloimmune thrombocytopenia (NAITP), post-transfusion purpura (PTP) and platelet refractoriness patients is equivalent to that of HPA-5 and exceed only by that of HPA-1, which implied that the Gov
antigen is as immunogenic as HPA-5. In concordance with HPA nomenclature, Gov\textsuperscript{b} has been renamed HPA-15\textsuperscript{a}, since it is the more prevalent allele in most studies (Metcalfe, Watkins et al. 2003).

1.3.2 Platelet Ultrastructure

Although anucleated, platelets are highly metabolic, and contain several mitochondria and membrane demarcated structures. This section will describe some of the ultrastructural components of human platelets.

1.3.2.1 Open canalicular system

The open canalicular system (OCS), also known as the surface connected canalicular system, is thought to arise during the membrane demarcation and compartmentalization of cytoplasm associated with the derivation of platelets from megakaryocytes. The OCS is essentially comprised of invaginations of the plasma membrane which form channels which extend deep into the intracellular environment and are distributed throughout the cytoplasm. These channels serve as bidirectional conduits for the exchange of particles between the platelet and the plasma. Following platelet activation with a strong agonist, the contents of intracellular storage granules are released into the OCS through which they are directed into the plasma.

Ultrastructural studies have demonstrated that the OCS is a network of highly branching and interconnected channels (White and Clawson 1980; Pujol-Moix, Hernandez et al. 2000; van Nispen Tot Pannerden, de Haas et al. 2010). Single channels emanate from
closed apertures at the plasma membrane and almost immediately connect to, and communicate with, other OCS channels within the cytoplasm. Each channel is comprised of dilated regions that alternate with pinched regions, giving the structure an overall sausage-like appearance. The multitude of connections between channels creates an intricate web that can allow for communication between opposing ends of a blood platelet.

The OCS serves as a membrane reserve for the shape change and spreading associated with platelet activation (White and Clawson 1980; Frojmovic and Milton 1983; Morgenstern, Neumann et al. 1987; Escolar, Leistikow et al. 1989). Tether formation – a key step in platelet adhesion under flow conditions - is independent of degranulation and does not require platelet activation, suggesting that additional plasma membrane is derived from the OCS rather than from storage granules (Zucker-Franklin, Benson et al. 1985). Rapid changes in the geometry of the OCS at high flow or following platelet stimulation are required for membrane evagination. It has been proposed that evagination processes such as platelet spreading and tether formation, are driven by the homotypic fusion of OCS membranes (van Nispen Tot Pannerden, de Haas et al. 2010). The OCS, in some instances, interacts with another intracellular membranous tubular structure, the dense tubular system (described in the next section), and it is conceivable that DTS constituents, such as Ca$^{2+}$, may contribute to the induction of OCS membrane evagination (White and Clawson 1980; van Nispen Tot Pannerden, de Haas et al. 2010).
1.3.2.2 Dense tubular system (DTS)

The dense tubular system (DTS) is derived from the rough endoplasmic reticulum of megakaryocytes. Calcium mobilization and thromboxane synthesis are two crucial functions of the DTS. In mature platelets, the DTS is the major site of prostanoid biosynthesis. Cyclooxygenase and thromboxane synthetase are stored within the DTS and these enzymes catabolise arachidonic acid towards thromboxane. The DTS also stores calcium and adenylate cyclase. Following platelet activation, calcium is liberated into the cytoplasm where it activates calcium dependent enzymes such as phospholipase A2, myosin light chain kinase, and calpain. Sarcoplasmic endoplasmic reticulum calcium ATPases (SERCAs) that are tightly coupled with plasma membrane calcium ATPases (PMCAs) located on DTS membranes, regulate calcium storage. The ATPases are in turn regulated by cAMP levels. As such, platelet activation levels are regulated by the balance between calcium and cAMP levels where decreases in cytosolic cAMP favour liberation of calcium. In addition to acting as an enzyme cofactor, calcium is essential for granule movements and secretion.

The OCS and the DTS are closely intertwined membrane systems, particularly at regions that interconnect individual OCS channels. However, although they exist in close proximity at some areas within platelets, the OCS and DTS are not connected (van Nispen Tot Pannerden, de Haas et al. 2010).

1.3.2.3 Dense granules (δ-granules)

Dense granules are the smallest of platelet granules, measuring approximately 150 nm in diameter, and are termed dense because they are both heavy and electron dense. There are
about 3 – 8 δ-granules per platelet. The osmiophilic character of dense granules, together with their inherent density allows for their distinction from other granules by electron microscopy. Dense granule lumens contain primarily non-protein small molecules, such as adenine nucleotides and serotonin. Dense granules are synthesized early during megakaryocyte maturation, where they are filled with adenine nucleotides and serotonin. Serotonin is the major osmiophilic component of δ-granules and exists at a concentration of around 65 mM (1000X greater than the plasma serotonin concentration). Dense granules also contain calcium stores – at 100X the level of that in the platelet cytosol, but this Ca$^{2+}$ is not mobilized following platelet activation. Rather, dense granule calcium serves primarily to stabilize adenine nucleotides and pyrophosphate. Additionally, calcium facilitates serotonin uptake and concentration by dense granules. In addition, δ-granules are rich in ganglioside GM3 and lysolecithin.

1.3.2.4 Alpha granules (α-granules)

Alpha granules are more heterogeneous than are DTS and δ-granules. They are the most abundant of platelet organelles, measure 200 – 400 nm in diameter, and are demarcated by a bilayer. Morphologically, α-granules are spherical to ovoid in structure with an opacity determined by the nature of their constituents. There are two major compartments within α-granules: the dark nucleoid region containing proteoglycans and the electron-luscent gray matrix. The electron-luscent gray matrix can be further subdivided into a: [1] nucleoid region, [2] intermediate zone enriched in plasmatic proteins, and [3] a peripheral zone comprising tubular structures where very large proteins such as vWF, multimerin, and factor C co-localize.
CD62P, GPIIb/IIIa, GPIV (CD36), CD9, and osteonectin are found on the inner membrane of the alpha granule. GLUT-3, PECAM, vitronectin receptor, and GMP33 are also present. Alpha-granules concentrate proteins which play a role in haemostasis, inflammation, and wound healing.

1.3.2.5 Lysosomes

Lysosomes contain a variety of digestive enzymes active under acidic conditions. Their size is intermediate to that of dense and alpha granules, measuring 175 – 225 nm in diameter. Due to their similar size and electron density, lysosomes can only be distinguished from alpha granules by cytochemical staining. Lysosomes contain glycosidases, acid proteases (such as cathepsin D and E, carboxypeptidases, collagenase and elastase), and cationic proteins with bactericidal activity. The lysosomal membrane contains CD63, LAMP-1, and LAMP-2. Following platelet activation CD63 is expressed on the activated platelet surface.

1.3.3 Granule Biogenesis

The mechanisms underlying platelet granule biogenesis are only beginning to be understood. Emerging evidence indicates that α- and δ-granules develop from budding vesicles of the Golgi complex within megakaryocytes. These vesicles subsequently progress into Type I and Type II multivesicular bodies and finally mature α-granules or δ-granules (Heijnen, Debili et al. 1998; Youssefian and Cramer 2000; King and Reed 2002). The biological switch which drives the divergence of α- and δ-granules from a common Type II MVB is unknown. Recent studies using megakaryocytes and platelets from normal and ARC (arthrogryposis, renal dysfunction, and cholestasis) donors defined the novel Sec1/Munc18
protein VPS33B as essential for protein trafficking and α-granule biogenesis (Lo, Li et al. 2005).

In contrast, proteins involved in δ-granule biogenesis include the Rab geranylgeranyl transferase α subunit, VPS33A, Rab27a, β3A AP-3, δ AP-3, and the novel proteins that form Biogenesis of Lysosome-related Organelles Complex (BLOC) (Li, Rusiniak et al. 2004).

The platelet secretome consists of endogenous proteins derived from megakaryocytes and endocytosed plasma proteins and small molecules. The linkage between the endocytic pathway and granule biogenesis is evidenced by the expression of Glut-3 – a constitutive membrane protein of other cells which lacks internalization motifs – on the demarcating membrane of α-granules and the plasma membrane of activated, but not of resting platelets (Pascoe, Inukai et al. 1996; Heijnen, Oorschot et al. 1997).

1.3.4 Platelet degranulation

Stimulation of platelets results in a hierarchal expression of activation induced cell surface markers (Berman, Yeo et al. 1986). Upon treatment of isolated platelets with an agonist, the high affinity-state integrin αIIbβ3 is the first marker to appear followed by α-granule, δ-granule, and lysozome derived proteins (Berman, Yeo et al. 1986). Additionally, only strong agonists are capable of inducing complete degranulation. The agonist-receptor density and their downstream signalling molecules define this dose response relationship. PAR-1, PAR-4, P2Y1, and TPα are all capable of coupling to Gαq which in turn leads to activation of PLCβ (Fig. 1-10). In contrast, GPVI coupling to FcRδ, leads to a signal cascade culminating in the activation of PLCδ2. Receptors for the strong - agonists, thrombin, collagen, and thromboxane - all share the property of inducing DAG and IP3 liberation (via phospholipase C
isoform activity), and the intracellular mobilization of Ca$^{2+}$. The P2Y1 receptor is expressed at 150 sites per platelet (PAR-1 and TPα are 1000 – 2000 sites per platelet) and as result, ADP is incapable of inducing complete degranulation despite being coupled to G$_{i/o}$ (Hechler, Zhang et al. 2003). The differential dose response is also due in part to the fact that for receptors capable of binding more than one type of G-protein - thrombin and thromboxane receptors for example - the concentration of agonists determines which species the GPCR will couple. We can exploit this fact experimentally to determine the subcellular localization of an activation-induced protein by comparing its activation kinetics with that of a well characterized marker.

**1.3.4.1 Platelet release reaction**

Recently, the molecular mechanisms which govern platelet secretion have been elucidated. Activation of platelets, with strong agonists, leads to discernible morphological changes and centralization of intracellular granules. The platelet cytoskeleton provides the contractile force that drives movement and subsequent release of the granules. This critical reorganization of the cytoskeletal is governed by gelsolin and calpain. These cytoskeletal reorganization proteins require micromolar concentrations of Ca$^{2+}$ that are achieved during the late phase of platelet activation. Calpain in particular, plays an important role in platelet granule secretion, aggregation and spreading. Immediately following platelet activation, calpain is translocated from the cytosol to the plasma membrane, where it is activated, and may cleave cytoskeletal proteins. Calpain-mediated events may be attributed to cleavage of various calpain substrates, such as cytoskeletal proteins – cortacin, talin, and actin binding protein, the cytoplasmic tail of integrin β3, and numerous signalling proteins (Fox, Goll et al. 1985; Oda, Druker et al. 1993; Banno, Nakashima et al. 1995; Du, Saido et al. 1995; Cooray,
Yuan et al. 1996; Huang, Tandon et al. 1997; Norris, Atkins et al. 1997; Falet, Pain et al. 1998; Mukhopadhyay, Ramars et al. 2001; Shcherbina, Miki et al. 2001). Disruption of the resting cytoskeletal barrier (along with actin polymerization to some degree) is required for secretion (Flaumenhaft, Dilks et al. 2005). Granular contents are secreted from the platelet, through fusion of the granule membranes with the OCS membrane. In rare cases, granules contents are secreted by direct fusion with the plasma membrane.

Signal transduction downstream of platelet receptor agonists leads to a rise in intracellular calcium levels (Figure 1-10). The platelet release reaction is calcium-dependent. Elevated calcium levels are predominantly the result of calcium influx from the DTS and to a lesser extent from the extracellular milieu. As such, only ligands for $G_{aq}$ or FcR coupled receptors – which lead to phospholipase C mediated elevation of IP3 – are capable of inducing degranulation. Such ligands include the strong agonists, thrombin and collagen, which are capable of inducing complete degranulation. In resting platelets, calcium levels are maintained below 100nM. Following activation, calcium is rapidly mobilized from the DTS, resulting in a profound ultrastructural change in the DTS from an elongated into a round vesicular form, and the centralized movement of granules which come into contact with the open canalicular system (White and Clawson 1980; White and Krumwiede 1987).

The platelet release reaction follows the fundamental secretion events: (1) movements of granules towards the plasma membranes via microtubule constriction and granule bound GTP-binding proteins, (2) docking, and (3) fusion. The widely accepted model posits that vesicles carry specific markers recognized by corresponding receptors on the plasma
membrane. Following the docking step, fusion of membranes requires energy, which is provided and catalyzed by soluble accessory proteins which form the fusion complex.

Granular secretion is tightly regulated, and the underlying molecular mechanisms are beginning to be elucidated. Fusion of granule membranes with membranes of the OCS is driven by soluble NSF attachment protein receptors (SNAREs). SNAREs are membrane associated proteins that are oriented toward the cytoplasm. Generally, SNAREs which are associated with the granule membrane are termed v-SNARES while those associated with the target membrane (OCS or plasma membrane) are termed t-SNARES. The current models of SNARE function state that, vesicle/target membrane fusion (granule membrane-OCS membrane fusion) is governed by the matching of a vesicle SNARE (v-SNARE) with a target membrane SNARE (t-SNARE). Pairing of a v-SNARE with its corresponding t-SNARE, through the ATPase activity of N-ethylmalameide-sensitive fusion protein (NSF), generates the energy that drives fusion (Rothman and Warren 1994; Weber, Zemelman et al. 1998). Human and mouse platelets contain SNARES: VAMP-2, VAMP-3, VAMP-7, and VAMP-8 (Flaumenhaft, Croce et al. 1999; Chen, Bernstein et al. 2000; Chen, Lemons et al. 2000; Polgar, Chung et al. 2002). VAMP-8 is the most abundant v-SNARE in human and mouse platelets, and together with syntaxin-2 associate with actin only following platelet activation (Bernstein and Whiteheart 1999; Polgar, Chung et al. 2002; Schraw, Rutledge et al. 2003; Ren, Barber et al. 2007). Actin polymerization is known to facilitate \( \alpha \)-granule secretion. As such, it is not surprising that VAMP-8 is the dominant effector of degranulation, and by extention, of stable aggregation in platelets. In support of this notion, an SNP in the 3' UTR in VAMP-8 correlates with early onset myocardial infarction (Shiffman, Rowland et al. 2006; Shiffman, O'Meara et al. 2006).
There are two types of t-SNARES: SNAP23/25 (SNAP-23, -25, -29, Sec9p) and syntaxin (syntaxin 1-19). Platelets contain syntaxin-2, -4, -7, and -11 as well as SNAP-23 and -29 (Lemons, Chen et al. 1997; Flaumenhaft, Croce et al. 1999; Reed, Houng et al. 1999; Chen, Bernstein et al. 2000). Syntaxin-4 is involved in alpha granule and lysosomes docking and release. SNAP-23 and syntaxin-2 are each involved in degranulation of all 3 granules.

Fusion of granule membranes with membranes of the OCS may require an initial priming step. Essentially, vesicle priming is a maturation process that enhances SNARE-mediated membrane fusion. In platelets, Munc13-14 appears to serve as a vesicle membrane fusion priming protein primarily for δ-granules, and Munc13-14 deficient mice possess a bleeding diathesis (Ren, Wimmer et al. 2010). α-granule and lysosome release is only partially inhibited in Munc deficient mice, which suggests the existence of another priming mechanism for these granules. Additionally, it is conceivable that a proportion of granule-OCS fusion is governed by mass-action, as is seen in in vitro assays of vacuole fusion, with the cytoskeleton providing the contractile force.

Calpain also plays a role in regulating platelet response to agonist by cleaving SNAP-23. SNAP-23, which is implicated in granule secretion, is cleaved (lost) in SFLLRN induced or Ca^{2+} treated-permeabilized platelets (Rutledge and Whiteheart 2002; Lai and Flaumenhaft 2003). VAMP-3 levels in SFLLRN treated platelets, in the presence or absence of calpain inhibitors, are identical to those of resting platelets. In contrast, SNAP-23 levels decline in SFLLRN treated platelets, and remain at levels identical to resting platelets, when platelets are stimulated in the presence of calpain inhibitors (Rutledge and Whiteheart 2002; Lai and Flaumenhaft 2003).
Calpain-mediated cleavage of SNAREs is slow, 1-5 minutes, and therefore does not contribute to the early events required for secretion (which take seconds). Instead, calpain-mediated cleavage of SNAREs serves as a mechanism to ensure that granule secretion is a single exocytic event (Lai and Flamenhaft 2003).

1.3.4.2 Heterogeneity and differential release of granules

Separate packaging of VEGF (proangiogenic) and endostatin (anti-angiogenic) into distinct α-granules was recently observed in megakaryocytes, pro-platelets, and platelets (Ma, Perini et al. 2005; Italiano, Richardson et al. 2008). Similar distinct distribution patterns have also been described for fibrinogen and VWF as well as thrombospondin-1 and basic FGF (Sehgal and Storrie 2007; Italiano, Richardson et al. 2008). For example, PAR-4 treatment results in loss of endostatin α-granules and retention of numerous VEGF α-granules. Conversely, PAR-1 treatment results in loss of VEGF α-granules and retention of numerous endostatin α-granules. In contrast, stimulation of platelets with ADP leads to increased plasma VEGF and no change in endostatin levels (Bambace, Levis et al. 2010). Additionally, a number of reports have demonstrated that fibrinogen is released more rapidly than is VWF in response to a weak agonist (Koutts, Walsh et al. 1978; Zucker, Broekman et al. 1979; Sehgal and Storrie 2007). Taken together, differential α-granule packaging allows for a mechanism to polarize the platelet secretome’s paratopic effects.
The problem becomes complicated if a conformational change in the native conformation is not exposed to the chemical reactions with all except the smallest nucleophiles (methylthioester reaction mechanism). In the A2H native form, the thioester is borne by the activated thioester directly.
section such that hydroxyl containing carbohydrates or protein ligands can react with the
portion of attack of hydroxyl-bearing molecules including water on the intermediate intermediate
and acyl intermediate. The cysteinyli thiocysteinyli is stimulated by binding of the thioester
of the thioester model is VH. An initial intramolecular transfer of the thioester can
then express the thioester to the aqueous phase. The thioester reaction determining
it C3 thioester reaction mechanism, in the native form, the C3 thioester is buried with

\[
\text{P} \quad \text{Targ} \quad \text{Formation of Intermediate Thioester}
\]

\[
\text{Act} \quad \text{Activated}
\]
Figure 1-5. The classical and lectin pathways of complement. MBL/MAKP replaces C1 in the lectin pathway. In the classical pathway the activator (Act) may be an antibody-antigen complex or other C1 activator, whereas a carbohydrate-coated microorganism serves as the Act in the lectin pathway. Activate enzymes are outlined with solid lines, control peptides (dashed lines), anyphylatoxins (circled).
Figure 1 – 6. The alternative pathway of complement activation. The alternative pathway proceeds as a result of the natural slow turn-over of C3 in the fluid phase. This turnover is initiated by hydrolysis of the thioester and leads to C3 deposition on susceptible surfaces. Any C3 which becomes deposited on self-surfaces (non-activating surfaces, bottom left) is rapidly degraded. In contrast, C3 bound to activating surfaces (bottom right) is protected from inactivation and rapid amplification of the pathway ensues. Active enzymes are outlined with solid lines, control proteins are outlined with dashed lines, and anyphylatoxins are circled.
Figure 1 – 7. Overview of haemostasis.
to adjacent platelets, which in turn aggregate to one another via thrombospondin cross-linking. 45, and release their contents into circulation. Components of the platelet secretome, which converts GpIIb/IIIa to its high-affinity state, additionally interact with intercellular storage granules 25. Interaction directly with collagen. The glycoprotein Iib/IIIa leads to generation of an "open conformational" structure, which interacts with collagen, thereby releasing stored components of the extracellular matrix and vascular wall damage through interaction with components of the extracellular matrix. Adhesion and activation of platelets at denuded endothelium are also crucial in the inflammatory response.
Activated receptor-1 (PAR-1, PAR-4), thrombin (FIIa) produced by the coagulation cascade stimulates platelets in its active conformation. Additionally, thrombin (FIIa) may activate PAR-4, which in turn activates Gq-coupled receptors. ADP synthesized through thromboxane (TXA2) synthesis/release and degranulation secreted ADP and thromboxane (TXA2) induce activation of high sensitivity Gq-coupled receptors (Gq-coupled receptors). Platelet activation is primarily through GPVI or GpIIb-IIIa. Signaling downstream of GPVI-Va. Via Gq-coupled receptors, GPVI-Va induces platelet activation. GPVI-Va activation is also mediated by thrombin. Thrombin activates GplIb-IIIa and ADP, which leads to platelet aggregation.
Figure 1 - 10. Role of SNAREs in α-granule membrane fusion. (A) Platelet α-granule secretion is mediated predominantly by VAMP-8, with VAMP-3 and VAMP-2 serving subordinate functions. Platelet tSNAREs – syntaxin-2 and syntaxin-4 – are involved in α-granule release. Interaction between coiled-coil domains within vSNARE (red) and tSNARE (orange-blue) form a twisted 4-helical bundle. (B)
<table>
<thead>
<tr>
<th></th>
<th>α-granules</th>
<th>Dense Granules</th>
<th>Lysosomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteoglycans:</td>
<td></td>
<td>Nucleotides:</td>
<td>Acid proteases:</td>
</tr>
<tr>
<td>βTG, PF4</td>
<td></td>
<td>ATP, ADP</td>
<td>Cathepsins D, E</td>
</tr>
<tr>
<td>Serglycin, HRGP</td>
<td></td>
<td>GTP, GDP</td>
<td>Carboxypeptidase</td>
</tr>
<tr>
<td>βTG Ag molecules: PBP, CTAP-II, NAP-2</td>
<td></td>
<td></td>
<td>Collagenase</td>
</tr>
<tr>
<td>Adhesive glycoproteins:</td>
<td></td>
<td>Amines:</td>
<td>Arylsulphatase</td>
</tr>
<tr>
<td>Fibronectin, Vitronectin, vWF</td>
<td></td>
<td>Serotonin</td>
<td></td>
</tr>
<tr>
<td>Thrombospondin</td>
<td></td>
<td>Histamine</td>
<td></td>
</tr>
<tr>
<td>Haemostasis factors and cofactors:</td>
<td></td>
<td>Bivalent cations:</td>
<td></td>
</tr>
<tr>
<td>Fibrinogen, FV, FVII, XI, XIII</td>
<td></td>
<td>Calcium</td>
<td></td>
</tr>
<tr>
<td>Kininogens, Protein S, Plasminogen</td>
<td></td>
<td>Magnesium</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>pyrophosphate</td>
<td></td>
</tr>
<tr>
<td>Cellular mitogens:</td>
<td></td>
<td>Glcohydrolases:</td>
<td></td>
</tr>
<tr>
<td>PDGF, TGFβ, ECGF, EGF, VEGF, IGF, IL-β</td>
<td></td>
<td>β-N-acetyl-glucosaminidase</td>
<td></td>
</tr>
<tr>
<td>Protease inhibitors:</td>
<td></td>
<td>β-glucuronidase,</td>
<td></td>
</tr>
<tr>
<td>α2M, α2-antitrypsin, PDCI</td>
<td></td>
<td>β-galactosidase</td>
<td></td>
</tr>
<tr>
<td>α2-antiplasmin, PAI1, TFPI, α1-PI, PIXI</td>
<td></td>
<td>α-D-glucosidase,</td>
<td></td>
</tr>
<tr>
<td>PN-2/APP, C1 inhibitor</td>
<td></td>
<td>β-D-glucosidase,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>β-glycerophosphatase,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>α-L-fucosidase,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>β-D-fucosidase,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>α-L-arabinosidase,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>α-D-mannosidase</td>
<td></td>
</tr>
<tr>
<td>Miscellaneous:</td>
<td></td>
<td>IgG, IgA, IgM</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Albumin, GPIa/Multimerin</td>
<td></td>
</tr>
</tbody>
</table>

Table 1-1 Platelet granule contents
<table>
<thead>
<tr>
<th></th>
<th>α-granules</th>
<th>Dense Granules</th>
<th>Lysosomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GTP binding proteins:</strong></td>
<td>rab4, rab6, rab8 Rap1</td>
<td>Ral, rab27</td>
<td>LAMP1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LAMP2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CD63</td>
</tr>
<tr>
<td><strong>Receptors and antigens:</strong></td>
<td>GPIb</td>
<td>CD63</td>
<td></td>
</tr>
<tr>
<td>P-Selectin (CD62)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPIIbIIIa</td>
<td></td>
<td>LAMP-2</td>
<td></td>
</tr>
<tr>
<td>GPIb-IX</td>
<td></td>
<td>Src</td>
<td></td>
</tr>
<tr>
<td>GPIV (CD36)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p24 (CD9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PECAM (CD31)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLUT-3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitronectin receptor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Osteonectin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMP33</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 – 2 Granule membrane proteins.
CHAPTER 2: CD109 GENOMIC STRUCTURE AND PHYLOGENETIC ANALYSIS

2.1 INTRODUCTION

CD109, a ~170 kDa GPI-anchored cell surface glycoprotein, was originally identified as a marker of activated platelets and T cells, and of early hematopoietic stem and progenitor cells (Suciu-Foca, Reed et al. 1985; Brashem-Stein, Nugent et al. 1988; Sutherland, Yeo et al. 1991; Haregewoin, Solomon et al. 1994; Sutherland 1995). While an early report suggested that T cell CD109 may play a role in T cell helper function (Suciu-Foca, Reed et al. 1985), and platelet CD109 is known to carry the biallelic Gov platelet alloantigen system (Kelton, Smith et al. 1990; Smith, Hayward et al. 1995), the function of CD109 in hematopoietic cells has remained obscure. We recently identified CD109 as a novel member of the α2 macroglobulin (α2M)/C3/C4/C5 superfamily (the A2M family) of thioester-containing protease inhibitor and complement proteins (Lin, Sutherland et al. 2002), and demonstrated that native CD109 contains an intact thioester. Members of this family share a high degree of sequence and structural similarity, and in particular, are characterized by the presence of an intrachain thioester bond formed between a cysteiny1 side chain sulfhydryl and a glutamine side chain carbonyl in the sequence CGEQ (Armstrong, Melchior et al. 1998; Dodds and Law 1998). This bond is typically unreactive in the native state, but following proteolytic cleavage of the protein, the thioester becomes highly reactive towards nucleophiles, such that the proteins become covalently bound to nearby macromolecules via ester or amide bonds (Law, Fearon et al. 1979; Law, Lichtenberg et al. 1979; Law, Lichtenberg et al. 1980; Law, Minich et al. 1981; Law 1983; Law 1983; Dodds, Ren et al. 1996; Law and Dodds 1997; Dodds and Law 1998). In
the case of complement, this leads to the covalent deposition of C3 and C4 on the target cell and on immune aggregates. In the case of the protease inhibitors, covalent binding of the activating protease (or of other adjacent molecules) may similarly occur. Based on sequence similarity (45-50% similarity overall, with particularly high similarity in the region of the thioester and in 11 additional A2M family-specific conserved sequence blocks), size, and overall structural organization, CD109 most closely resembled the protease inhibitors, and appeared to be more distantly related to the complement proteins (Lin, Sutherland et al. 2002). Notably, however, based on the presence of His rather than Asn (characteristic features of the C3-like and α2M-like proteins, respectively) in the terminal position of a conserved hexapeptide lying ~100 amino acids C-terminal to the thioester bond which by protein folding interacts with and modulates the reactivity of the thioester (Ren, Dodds et al. 1995; Nagar, Jones et al. 1998) the biochemical specificity of the CD109 thioester was predicted to resemble that of the complement proteins, preferentially forming amide rather than ester bonds with target substrates. By analogy with the other members of the A2M family, I reasoned therefore, that CD109 may play a conserved role in innate and adaptive immune responses, and that its immune function is likely mediated by thioester-mediated covalent binding to target molecules (Lin, Sutherland et al. 2002). And consistent with such a conserved role, I observed that CD109-like molecules appeared to be evolutionarily conserved as well, with related thioester-containing proteins being found in both vertebrate and invertebrate deuterostomes as well as in arthropod, mollusc, and nematode protostomes.

The A2M superfamily has traditionally comprised two major arms – the α2M protease inhibitors and the complement proteins – that in humans contains α2M and the related
pregnancy zone protein (PZP), and the complement C3, C4, C5 proteins, respectively. Due to its centrality in the immune response, the evolution and phylogeny of the A2M family of thioester proteins has long been of interest. While C3-like proteins have been found in a variety of non-vertebrate deuterostomes (including cephalochordates, urochordates, and echinoderms), jawless fishes, and in jawed vertebrates, C4 and C5 have been found only in jawed vertebrates. In contrast, α2M-like proteins are present not only in all chordates but have also been found in protostomes including various molluscs and arthropods (Quigley and Armstrong 1985; Hergenhahn, Hall et al. 1988; Hall, Soderhall et al. 1989; Enghild, Thogersen et al. 1990; Bender, Fryer et al. 1992; Armstrong, Melchior et al. 1998; Nonaka and Azumi 1999; Nonaka, Azumi et al. 1999). An α2M-like molecule was therefore proposed as the ancestral member of this family, with the other members arising by repeated gene duplication and subsequent divergence of function (Sottrup-Jensen, Stepanik et al. 1985; Hughes 1994; Dodds and Law 1998). The evolutionary divergence of vertebrates from other deuterostomes is believed to have been driven by repeated rounds of whole genome duplication at the boundary between cephalochordates and early vertebrates. Consistent with this, the original α2M/C3 duplication is presumed to have occurred in a common deuterostome ancestor prior to the emergence of vertebrates, with C4 (and possibly C5) arising later in jawed vertebrates (Nonaka 2001).

Several recent observations have added new variables to this evolutionary schema. First, the cloning of CD109, and the identification of additional putative human A2M family genes via the human genome project have increased the complexity of this family. Second, consistent with the rapid expansion of DNA and protein databases, increasing numbers of
phylogenetically diverse A2M proteins bearing similarity to CD109 have been identified in invertebrates. Specifically, CD109 relatives include novel thioester-containing proteins (TEPs) from Dipteran insects (Lagueux, Perrodou et al. 2000; Levashina, Moita et al. 2001), hypothetical TEP-like proteins from the nematode *C. elegans* and an A2M-like protein from the urochordate *C. intestinalis*. Notably, it has recently been proposed that the complement-like TEPs - which can promote bacterial opsonization and phagocytosis *Sua 5.1* cells (a hemocyte-like mosquito cell line) in a thioester-dependent manner - may comprise a phylogenetically distinct, and possibly ancestral, branch of the A2M family (Levashina, Moita et al. 2001). And finally, the recent identification of a C3-like molecule in the coral *Swiftia exserta*, indicates that A2M family is not restricted to protostomes and deuterorostomes, but rather that the origins of the A2M family likely lie prior to the divergence of cnidaria and bilateria - earlier than previously believed.

In view of the central role of thioester-containing proteins in innate and adaptive immunity, clarification of the status of CD109 in the evolution and phylogeny of the A2M family is of considerable interest. In the present study, I initially focussed on the elucidation of the evolutionary relationships of CD109, the other 5 fully-characterized human A2M proteins α2M, PZP, C3, C4, and C5, and three additional more recently-identified A2M family members, using sequence analysis and a detailed comparison of the intron/exon structure of the corresponding loci. I then extended this analysis to compare CD109 to related sequences extending back to prostostomes and invertebrate deuterostomes, thereby defining CD109 as a member of a distinct and archaic branch of A2M phylogenetic tree.
2.2 OBJECTIVE AND AIMS

2.2.1 Rationale:
- If members of a gene family have arisen as a result of gene duplication and subsequent divergence of function, one would expect to see similarities in their gene structure. In addition, the degree of similarity among gene family members is a reflection of their relatedness.

2.2.2 Objectives:
- Develop an evolutionary model of the A2M family

2.2.3 Specific Aims:
- Determine the genomic structure of CD109
- Compare and contrast genomic organization the A2M proteins
- Suggest a A2M family phylogenetic tree

2.2.4 Hypothesis:
Comparison of genomic organization, in conjunction with translated sequence data, will facilitate the generation of a parsimonious A2M phylogenetic tree and delineation of CD109 position amongst the human A2M family.
2.3 EXPERIMENTAL PROCEDURES

2.3.1 Chromosomal Localization of CD109

As early iterations of human genome project data were ambiguous with respect to the location of CD109, I determined its position by FISH and radiation hybrid mapping as well.

i.) FISH Mapping. A ~4.5 kb CD109 cDNA (clone K1) (Lin, Sutherland et al. 2002) was used to screen a P1 derived artificial chromosome (PAC) library in the Canadian Genome Analysis and Technologies (CGAT) Physical Mapping Resource Facility (Hospital for Sick Children, Toronto). The resultant CD109-specific pCYPAC-1 clones - 94J24 and 4L10 - were then used for fluorescence in situ hybridization (FISH) analysis of normal human lymphocyte chromosomes counterstained with propidium iodide and 4',6-diamidino-2-phenylindol-dihydrochloride (DAPI). Following probe biotinylation by nick translation, and cot-1 suppression by preannealing, hybridization was detected with avidin-fluorescein isothiocyanate (FITC). Images of metaphase preparations were visualized by digital imaging microscopy using a thermoelectrically cooled charge coupled camera (Photometrics, Tucson, AZ). Hybridization signals and DAPI banded chromosome images were acquired, and pseudo coloured yellow (FITC) and blue (DAPI) signals were overlaid electronically and merged using Adobe Photoshop 3.0 software. Chromosomal band assignment was determined by measuring the fractional chromosome length and by analysing the banding pattern generated by the DAPI counterstained image.

ii.) Radiation Hybrid Mapping. Two PCR primers (K1UTRs, 5'-GTCACATGTGATTGTATGTTTTCG-3'; K1UTRas, 5'-GGGGAAAATATAGACACACTGC-3') were designed to amplify a 189 bp fragment of the CD109 clone K1 3' UTR. PCR reactions were carried out in 25 μL reaction
volumes with 25 ng of human genomic DNA; 12.5 pmol of each primer; 1.25 units of Taq polymerase; 200 pmol of each dNTP; 1.0 mM MgCl$_2$; 20 mM Tris-HCl, pH 8.4 and 50 mM KCl. A – hot start – was carried out with the addition of Taq polymerase and dNTP after an initial denaturation at 94°C for 5 minutes followed by 30 cycles of 94°C x 30 s, 52°C x 30 s, 72°C x 30 s, followed by a single final extension at 72°C for 15 minutes. All reactions were carried out in a DNA Thermocycler 480 (Perkin Elmer) with an overlay of two drops of mineral oil. At the completion of the PCR run, 5 μL of loading buffer were added to each reaction and a 10 μL aliquot was size-separated electrophoretically on a 2% agarose/TAE gel containing 0.5 μg/ml ethidium bromide, and inspected visually. Negative controls to check for cross contamination were negative, as was the homology control with hamster DNA (A3). Both the GeneBridge 4 RH panel and Stanford G3 RH panel were screened using the K1UTRs/as primer pair. The GeneBridge 4 RH panel controls were HFL (human genomic DNA, positive) and A23 (hamster genomic DNA, negative), while the Stanford G3 RH panel controls were A3 (non-irradiated hamster genomic DNA, negative), and RM (non-irradiated human genomic DNA, positive). Panel results were scored independently by two observers, and were then submitted to the Whitehead Institute/MIT Center for Genomic Research (http://www.genome.wi.mit.edu) and to the Stanford Human Genome Center (http://www-shgc.stanford.edu).

2.3.2 Genomic Structure of CD109

BLAST searches of early release of human genome identified two clones, RP11-553A21 and RP11-251A23, which contained CD109 sequence. These clones, and by extension CD109, were later mapped to position 6q14. Intron/exon junctions were determined by sequencing
EcoRI/HindIII pBluescript subclones of both BACs or from alignment of later human genome releases using the Spidey program (www.ncbi.nih.gov). Intron sizes were originally inferred from Southern blot analysis of the BAC clones (and later confirmed with updated chromosome 6 contig sequences on the NCBI database). Specifically, following test digestions with restriction enzymes, in order to identify rare-cutter RE that cut the CD109 cDNA no more than once, I digested genomic DNAs to completion with KpnI, SacI, and SpeI (alone or in combination), size-separated the DNA fragments electrophoretically on a 0.7% agarose gel, and performed Southern blotting with the exon-specific end-labelled oligonucleotides in Table 2-1.

2.3.3 Sequence analysis

Public databases were searched for sequences similar to CD109 using the position specific iterated BLAST (Altschul, Gish et al. 1990; Altschul, Madden et al. 1997) (PSI-BLAST; www.ncbi.nlm.nih.gov) or Smith-Waterman (Smith and Waterman 1981) (decypher2.stanford.edu) algorithms. Incomplete and redundant sequences were not analyzed further. Multiple sequence alignments (MSA) were performed using Clustal W (Thompson, Higgins et al. 1994) (www.ebi.uk), pairwise alignments used LALIGN (Huang and Miller 1991) (www.ch.embnet.org), and phylogenetic trees were produced using PHYLIP (Felsenstein 1989) or MEGA2.1 (Kumar, Tamura et al. 2001). To confirm selected protein similarities as appropriate, I performed reciprocal best match (RBM) analysis (Zdobnov, von Mering et al. 2002) using BLASTP or Smith-Waterman algorithms to compare specific protein sequences to whole translated genomes.
The intron/exon structures of the non-CD109 human A2M family members were determined using pregnancy zone protein (PZP), alpha-2-macroglobulin (α2M), Ovostatin-like (OVO), VIP/KIAA1283 (VIP), alpha-2-macroglobulin –like (BAC), C3, C4, and C5 sequences to search translated human genomic sequence. In parallel, these sequences were aligned with human CD109 sequence, thereby allowing corresponding exons to be identified. Paralogous exons were identified as those that aligned by MSA and demonstrated >13% sequence identity using the LALIGN algorithm. The phase of introns adjacent to paralogous exons was then determined. In cases in which one exon was paralogous to two exons in another family member, intron position was compared with respect to the organization of CD109. The intron/exon structure and intronic phase of the mammalian (Mus musculus, Canis familiaris, Sus scrofa, Equus caballus, Bos Taurus, and Pan troglodytes) and avian (Gallus gallus, Taeniopygia guttata) CD109, Drosophila melanogaster Mcr and TEP1, 2, 3, 4, Caenorhabditis elegans ZK337.1, and Ciona intestinalis α2M homologue genes were compared to that of human CD109 by a similar approach.
2.4 RESULTS

2.4.1 Chromosomal Mapping of the CD109 locus

Using cDNA clone K1 as a probe, two positive genomic PAC clones - 94J24 and 4L10 - were identified. The chromosomal assignment of each of these clones was then determined by FISH analysis of 20 well-spread metaphases. Both PACs mapped to 6q12-13, with positive hybridization signals being observed in >90% of the cells, and on both homologues in >90% of the positive spreads (not shown). This chromosomal location was then confirmed by radiation hybrid mapping: Using both the Gene Bridge 4 and the Stanford G3 panels and a 3'UTR PCR probe, CD109 was mapped to within 11.09 cR and 6.9 cR of framework markers CHLC.GATA11F10 and SHGC-33186, respectively - a region reportedly corresponding to 6q13. Although the human genome database was initially ambiguous with respect to the location of CD109, its position has now been refined to 6q14.1 (NCBI LocusID: 135228; Genomic contig NT_007158).

2.4.2 Genomic Structure of CD109

As summarized in Figure 2-1, the CD109 locus on chromosome 6q14.1 comprises 33 exons, and spans at least 128kb. Our assignment of intron sizes - determined largely by Southern blot analysis - has been confirmed more recently in the human chromosome 6 sequence assembly NT_007158.10, validating my strategy. All exon/intron junctions conform to the GT-AG rule (Mount 1982) (Table 2-2). Internal exons range in size from 29 bp (exon 3) to 231 bp (exon 4). Intron size is quite variable, ranging in size from 87 bp (intron 5) to 25678 bp (intron 2). Three alternative polyadenylation sites are found in exon 33, resulting in three transcripts containing 473, 1608, or 4851 bp, respectively of exon sequence. All three
transcripts can be detected readily by RT-PCR in a variety of hematopoietic cells (data not shown).

I have assigned each important structural feature of CD109 to a specific exon. Exon 1 encodes the leader peptide, the Gov^a/Gov^b polymorphism is found in exon 19, and exon 33 encodes the GPI anchor addition/cleavage site. Notably, as is the case for all other rodent and human α2M and α2M-like genes analyzed, the bait region of CD109 is encoded by exons 17 and 18 (Matthijs, Devriendt et al. 1992; Overbergh, Hilliker et al. 1994; Overbergh, Lorent et al. 1994). The exons encoding the thioester (exon 23) and the thioester reactivity determining triplet VIH (exon 25) differ numerically from those of other A2M molecules analyzed however (Lundwall, Hellman et al. 1984; Lundwall, Wetsel et al. 1984; Hattori, Hamilton et al. 1989; Carney, Haviland et al. 1991; Haviland, Haviland et al. 1991; Vik, Amiguet et al. 1991; Yu 1991; Matthijs, Devriendt et al. 1992; Overbergh, Hilliker et al. 1994; Overbergh, Lorent et al. 1994), but are similarly separated by one intervening exon.

2.4.3 The Human A2M Family

In addition to the 5 fully characterized non-CD109 human A2M family members α2M (12p13.3-p12.3), Pzp (12p13-p12.2), C3 (19p13.3-p13.2), C4 (6p21.3), and C5 (9q32-q34), Blast searching identified 3 additional human A2M family genes Ovostatin-like (OVO; 12p12.1), VIP/KIAA1283 (VIP; 19p13.12), and α2M-like (BAC; 12p13.31). An additional A2M family gene (LOC 340268; 7p22.1) appeared to be a VIP-related pseudogene. Alignment of the 9 human A2M proteins results in a phylogenetic tree (Figure 2-2) containing three main branches:
CD109 and VIP together form a branch that appears to be distinct and intermediate to that of discrete \( \alpha 2M/PZP/OVO/BAC \) and \( C3, C4, C5 \) branches.

### 2.4.4 Conservation of Genomic Structure in the Human A2M Family

To help clarify the evolutionary relationship of CD109 to other human A2M proteins, I next compared the genomic structure of human \( CD109 \) with that of the other human A2M family genes. This analysis was based on the notion that individual members of gene families that have arisen by gene duplication, as has been proposed for the A2M family (Sottrup-Jensen, Stepanik et al. 1985; Dodds and Law 1998) would be expected to share genomic structural features, reflecting their origin from a common ancestor, with more closely related genes showing greater similarity. Figure 2-3 outlines the exonic organization of the human A2M family genes, and identifies paralogous exons. While total exon numbers vary among family members (\( CD109, 33; \alpha 2M/PZP, 36; OVO, 35; VIP, 42; BAC, 35; C3/C4/C5, 41 \)), marked conservation of genomic organization is nevertheless observed. Notably, and consistent with the three branch human A2M tree shown above (Figure 2-2), the \( \alpha 2M/PZP/OVO/BAC \) and \( C3/C4/C5 \) genes, respectively, define two conserved patterns of genomic organization, with the \( CD109 \) and VIP loci showing features of both (but with both genes diverging from each other and from the conserved \( \alpha 2M \) and \( C3 \) patterns in several cases). I next examined intron phase in the nine human A2M genes, based on the premise that gene family members that have arisen by gene duplication should demonstrate conservation of introns, with the degree of conservation reflecting their relatedness (Table 2-3). As expected, intron phase is conserved between exons that encode conserved structural motifs. For example, complete conservation
of introns is observed among all family members adjacent to the exons encoding the thioester (exon 23 of CD109, exon 24 of α2M/PZP/BAC/C3/C4/C5, exon 25 OVO, exon 26 VIP) and the thioester reactivity determining triplet (exon 25 of CD109, exon 26 of α2M/PZP/BAC/C3/C4/C5, exon 27 of OVO, exon 28 of VIP). Similarly, the bait region-encoding exons (exons 17 and 18 of CD109, PZP, and α2M) are interrupted by a phase 1 intron across the entire gene family. Overall, 25/27 α2M/PZP introns, 27/27 VIP introns, 25/26 BAC introns, 21/24 OVO introns and 25/29 C3/C4/C5 introns are in the same phase as is the paralogous CD109 intron.

2.4.5 Conservation of Genomic Structure of Mammalian and Avian CD109

I next compared the genomic organization of human, chimp, mouse, horse, cow, dog, and chicken CD109 (Figure 2-4). In each sampled taxa, excluding bovine and equine, CD109 is comprised of 33 exons. The bovine CD109 sequence is incomplete in the 5’ end and as such its true size is obscured. Equine CD109 displays examples of intron insertion, the region of the gene which is homologous to exon 1–exon 2 of human, chimp, murine, and chick CD109 is interrupted by 3 introns rather than 1. In accordance, Intron 1 - with respect to human, chimp, murine, and chick - is lost in equine and canine CD109. Introns are largely conserved across the entire subset of analysed taxa with the exception of equine and canine which share a homologous intron at intron 3 and 1, respectively. Notably, intron phase at homologous introns is conserved across the entire set of mammalian and avian genes analysed (Table 2-4).
2.4.6 Phylogenetic Position of CD109

Position specific iterated BLAST (Altschul, Gish et al. 1990; Altschul, Madden et al.
1997) searches of public databases were thereafter used to identify a series of additional
genes encoding proteins with a high degree of sequence similarity to human CD109.
Incomplete and redundant sequences were excluded from further analysis. CD109-related
proteins represented a diverse spectrum of protostome and vertebrate and invertebrate
deuterostome taxa, as well as a single C3-like protein from the cnidarian Swiftia exserta, and a
placazoan. The closest relatives of human CD109 include its primate, canine, equine, swine,
murine, marsupial, and avian orthologs. A phylogenetic tree generated using a subset of these
sequences (Figure 2-5) reveals that the A2M family comprises several distinct branches
corresponding to the $\alpha$2M-related proteins, the complement proteins (with distinct C3, C4,
and C5-specific sub-branches [and additional outlying sub-branches corresponding to C3-like
molecules from cephalochordates, urochordates, and echinodermata, as well as to the
cnidarian, Swiftia exserta]), and a third branch containing CD109. Within this latter branch,
CD109 is grouped most closely with the protease inhibitor HFPI from the hagfish Eptatretus
burgeri (EpBurgPI), $\alpha$2M from the ascidian C. intestinalis (ci$\alpha$2M) and the D. melanogaster
and Anopheles gambiae thioester containing proteins and more distantly with a putative $\alpha$2M-like
protein from the nematode Caenorhabditis elegans (ceZK337.1) and the D. melanogaster TEP
proteins (dmTEP1-4). Additionally, the CD109/TEP branch contains a Trichoplax adhaerens
hypothetical protein. Trichoplax adhaerens is an extant Placazoan, although its evolutionary
position is under some debate. The overall consensus is that Trichoplax is basal to Cnideria
The presence of a placazoan A2M/C3-like protein exclusively in the CD109/TEP branch suggests that CD109 is archaic among human A2M proteins.

As this tree was constructed from the results of PSI-BLAST searches using CD109 as the query, it was possible that it may not reflect true orthologous relationships. I therefore used the reciprocal best match approach to exclude this possibility: Translations of the *D. melanogaster, C. elegans*, and *C. intestinalis* genome databases were searched using human CD109, PZP, α2M, VIP, BAC, OVO, C3, C4, and C5 sequences as queries. Reciprocal searches of the translated human genome using ceZK337.1, ciα2M, dmMcr and the dmTEPs 1-4 as queries were performed in parallel. CD109/ceZK337.1, CD109/ciα2M and CD109/TEP2 combinations yielded the highest BLASTP scores in all cases (not shown). Curiously, despite the closer relatedness of CD109 and the Mcr proteins that is implied by Figure 2-5, CD109/Mcr BLASTP scores were considerably lower than those of CD109/TEP1-4. Moreover, the functional relatedness of the Mcr proteins is questionable as well, as both bear distinct mutations that would be predicted to prevent thioester formation (Isaac and Isenman 1992). I suggest therefore that TEP2 may be the *D. melanogaster* ortholog of CD109.

To clarify these relationships further, I next compared the genomic organization of *hsCD109, ceZK337.1, ciα2M* and *dmTEPs 1-4* (Table 2-5), anticipating that the phylogenetic relatedness of these genes would be reflected by the shared features of genomic organization. The *ceZK337.1, dmMCR* and *dmTEP* genes are considerably smaller than their human or ascidian (ciα2M 32 exons) counterparts (*ceZK337.1* (17 exons); dmMCR (8 exons); *dmTEP1* (9 exons); *dmTEP2* (8 exons), *dmTEP3* (15 exons), and *dmTEP4* (7 exons), consistent with the
lower proportion of non-coding genomic sequence in protostomes compared to chordates. Notably, however, and consistent with the evolutionary relatedness of these genes, varying degrees of conservation of intron position (and of intron phase) were observed, with conservation following the hierarchy \( ci\alpha_2M > ceZK337.1 > dmTEP3 > \)
\( dmMCR/dmTEP1/dmTEP2 > dmTEP4 \).
<table>
<thead>
<tr>
<th>Exon</th>
<th>Oligonucleotide</th>
<th>Clone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>K1-80 5' GTAGCCCAGGCAGACGCC 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>2</td>
<td>K1-193 5' GTTTCTGGTGAACGCCAGCCAG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>3</td>
<td>K1-360 5' GCTCTTTTAAAGACACTTCACTTCCATCA</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>4</td>
<td>411-430 5' GAGATTTATGAGCTACGTGAA 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>5</td>
<td>K1-644 5' CAGTGGTTGTCACAACAAAAGTG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>6</td>
<td>K1-751 5' GACCCAGATTATATCAATCATTTCAGG</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>7</td>
<td>K1-844 5' GCATTTAAATGGTACCATCACGGG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>8</td>
<td>K1-888 5' GGAAGGCCGTAAGAGGACG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>9</td>
<td>K1-1034 5' GGACTTTCTGAATACCTGGATC 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>10</td>
<td>K1-1128 5' GCACATAATGTGTTCTCAAGC 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>11</td>
<td>K1-1277 5' GTATGCATAACAGTGACACAG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>12</td>
<td>K-1485 5' GTCTGTTTAAGTCTCCTAGTAAG</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>13</td>
<td>K-1562 5' GGCTCTGTTTAAGTCTCCTAGTAAG</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>14</td>
<td>K-1622 5' GGACAGTTTGTTGGCTGTAGG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>15</td>
<td>K1-Rev1 5' GGCTGTGTCACAGAGATC 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>16</td>
<td>K1-1982 5' GCGATGCTGATGAATTCTTTTGAGCAGACG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>17</td>
<td>K1-2193 5' GGATTTTGCTAGACACCAACATGGG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>18</td>
<td>CD109GSP3 5' GCCACCCAAGAAGTGATAGA 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>19</td>
<td>CD109GSP1 5'CACCTCTGATAACAGAGTAG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>20</td>
<td>K1-2535 5' GGTCCCGATGAGATTGGGCC 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>21</td>
<td>K1-2724 5' ATAGGCTACAGAGTACCCTG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>22</td>
<td>K1-2858 5' CGGATGCCCTTTATGGCTGG 3'</td>
<td>pCYPAC-109</td>
</tr>
<tr>
<td>23</td>
<td>K1-3014 5' CAGAGGGGAAGATGGCTCTTTCAGG 3'</td>
<td>RP11-553A21</td>
</tr>
<tr>
<td>24</td>
<td>K1-3026 5' GGCTCTTTTCACTGCTATGGG 3'</td>
<td>RP11-553A21</td>
</tr>
<tr>
<td>25</td>
<td>K1-2995 5' CCAAGAGAGATGGCTCTTTCAGG 3'</td>
<td>RP11-553A21</td>
</tr>
<tr>
<td>26</td>
<td>K1-3167R 5' CCGTTGGATTTCTGATGTCC 3'</td>
<td>RP11-553A21</td>
</tr>
<tr>
<td>27</td>
<td>K1-3339 5' GGAGTCTGAATTCAGTAGAGG 3'</td>
<td>RP11-553A21</td>
</tr>
<tr>
<td>28</td>
<td>K1-3570 5' ACAATTTCAGACTTCTGGG 3'</td>
<td>RP11-553A21</td>
</tr>
<tr>
<td>29</td>
<td>K1-3725 5' ATCCAAGTGGCCGTGAGAGCAGG 3'</td>
<td>RP11-553A21</td>
</tr>
<tr>
<td>30</td>
<td>HCD109-8 5' GCTGTGGGTAACAGCAATGGC 3'</td>
<td>RP11-251A23</td>
</tr>
<tr>
<td>31</td>
<td>1-5new 5' CAGCAACATCTAATAACAGACCACCACCAGGG 3'</td>
<td>RP11-251A23</td>
</tr>
<tr>
<td>32</td>
<td>K1-4051 5' GGCTCTTTTCACTGCTATGGG 3'</td>
<td>RP11-251A23</td>
</tr>
<tr>
<td>33</td>
<td>K1-4233 5' CCAAGAGATGGCTCTTTCAGG 3'</td>
<td>RP11-251A23</td>
</tr>
<tr>
<td>34</td>
<td>K1-4378 5' CTCCCATCATCACTTCTCAG 3'</td>
<td>RP11-251A23</td>
</tr>
<tr>
<td>35</td>
<td>7(U2)n 5' ACTCTACTGAGACCCACAGG 3'</td>
<td>RP11-251A23</td>
</tr>
<tr>
<td>36</td>
<td>6(7-1)R 5' GAAGCTGTTGTCACAGAGTAG 3'</td>
<td>RP11-251A23</td>
</tr>
<tr>
<td>37</td>
<td>K1-5337 5' ACTGCAGACCGGACCCTCCTGGC 3'</td>
<td>RP11-251A23</td>
</tr>
</tbody>
</table>

Table 2-1. Oligonucleotides used for the determination of the intron/exon structure of human CD109. Oligonucleotides specific to each of the 33 exons of CD109 and the corresponding genomic PAC and BAC clones are listed.
### Table 2-2. Intron/exon structure of CD109

The sizes of CD109 introns 1-32 and exons 2-33, and the corresponding intronic splice donor and acceptor (lower case) and exonic sequences (upper case) are listed. K1, K1-H7, and K1-3 are three alternative, contiguous exon 33 sequences that have been defined at the mRNA level, with the 3' nucleotides shown corresponding to the 5' ends of the poly(A) tails.

<table>
<thead>
<tr>
<th>Exon</th>
<th>Exon Size (bp)</th>
<th>3' exon donor</th>
<th>Intron Size (bp)</th>
<th>acceptor 5' exon</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1110</td>
<td>TGGCTCCCGG gtaggaacgt</td>
<td>1110 aactctttag GCCTCGGTTT</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>25678</td>
<td>TTTGAAAAAG gtaagatagaa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>32430</td>
<td>TCTTCTCATCA gtaagtagccc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>52664</td>
<td>TCTCATTAAG gtaagttgccaaa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>82857</td>
<td>TCAAGTGAAAG gtaagttcaggt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1110</td>
<td>TCAAGATATGG gtaaggtcctc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2244</td>
<td>TCAGGGCGAA gtaagttgccaa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>64355</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>951</td>
<td>TACTCTTTCAG gtaagtttttc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2244</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>64355</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1110</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1110</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1110</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1110</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1535</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1925</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2317</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>2710</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>3197</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>3678</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>4013</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33a</td>
<td>836</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33b</td>
<td>1608</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33c</td>
<td>4851</td>
<td>TACTCTTTCAG gtaagttttagg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD109 Intron</td>
<td>CD109</td>
<td>VIP</td>
<td>BAC</td>
<td>OVO</td>
</tr>
<tr>
<td>--------------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0(2)</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0(4)</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0(2)</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0(6)</td>
<td>0(5)</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1(8)</td>
<td>1</td>
<td>1(7)</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>2(9)</td>
<td>2</td>
<td>2(8)</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0(10)</td>
<td>0</td>
<td>0(9)</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1(10)</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0(11)</td>
<td>0</td>
<td>0(11)</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0(12)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0(12)</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0(15)</td>
<td>0</td>
<td>0(15)</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0(16)</td>
<td>0</td>
<td>0(16)</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>0(17)</td>
<td>2</td>
<td>0(17)</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>2(19)</td>
<td>2</td>
<td>2(19)</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0(20)</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0(19)</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0(22)</td>
<td>0</td>
<td>0(22)</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
<td>1(25)</td>
<td>1(23)</td>
<td>1(24)</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>1(26)</td>
<td>1(24)</td>
<td>1(25)</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>0(29)</td>
<td>0(27)</td>
<td>0(27)</td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>1(30)</td>
<td>1(28)</td>
<td>1(28)</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>0(32)</td>
<td>0(29)</td>
<td>0(29)</td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>0(33)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>29</td>
<td>0</td>
<td>0(34)</td>
<td>0(30)</td>
<td>0(30)</td>
</tr>
<tr>
<td>30</td>
<td>2</td>
<td>2(35)</td>
<td>2(31)</td>
<td>2(31)</td>
</tr>
<tr>
<td>31</td>
<td>0</td>
<td>0(37)</td>
<td>0(33)</td>
<td>1(33)</td>
</tr>
<tr>
<td>32</td>
<td>1</td>
<td>1(38)</td>
<td>1(34)</td>
<td>2(34)</td>
</tr>
</tbody>
</table>

**Table 2-3. Conservation of intron phase among human A2M family genes.** The phase of introns lying 3’ to paralogous A2M, PZP, VIP, OVO, C3, C4, and C5 exons (as defined by the genomic structure of CD109; see Figure 2) is shown, with the corresponding intron numbers listed in parentheses. The absence of a paralogous intron is denoted by -.
Table 2-4. Conservation of intron phase among mammalian and avian CD109. Paralogous exons were determined by comparing genomic, mRNA, and protein sequences of mammalian and avian CD109 and the phase of adjacent introns was compared. Conserved introns are shown relative to the position of human CD109.
Table 2-5. Conservation of intron phase among human CD109 and C. elegans ZK337.1 and D. melanogaster MCR, TEP1-4 genes, C. intestinalis A2M. Paralogous exons were defined by comparing hsCD109 sequences to those listed in A2M/TEPs from Drosophila Melanogaster, Caenorhabditis elegans, and Ciona intestinalis, and the phase of the adjacent introns was compared. Conserved intron positions are shown, relative to the positions of the 32 CD109 introns, with the phase of the corresponding intron given in parentheses. Corresponding intron positions aligned quite precisely overall, except for four instances (#) in which the alignment was staggered slightly or ambiguous.

C.elegans (ce)ZK337.1 (NM_061213/cosmid) ZK337; Drosophila melanogaster (dm) TEP1 (NM_078854/AE014134); dmTEP2 (AJ973633/AE014134); dmTEP3 (AJ269540/AE014134); dmTEP4 (AJ269541/AE014134); Drosophila melanogaster macroglobulin complement-related (Mcr) protein dmMCR (NM_079949/NT_033779); Ciona ciA2M (NM_001032516/).
Figure 2-1
Figure 2-2
<table>
<thead>
<tr>
<th></th>
<th>Hs</th>
<th>Pt</th>
<th>Mm</th>
<th>Eq</th>
<th>Bt</th>
<th>Cf</th>
<th>Gg</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>12</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>13</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>14</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>15</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>17</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>18</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>19</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>20</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>21</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>22</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>23</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>24</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>25</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>26</td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>26</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>28</td>
<td>26</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>29</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>28</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>30</td>
<td>28</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>31</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>32</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>33</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>32</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td>34</td>
<td>32</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>35</td>
<td>33</td>
<td>33</td>
<td>33</td>
</tr>
</tbody>
</table>
Figure 2-5
Figure Legends:

**Figure 2-1. The genomic organization of CD109.** The intron/exon structure of the CD109 locus on chromosome 6q14.1 is shown, drawn to scale. CD109 encompasses >128kb and comprises 33 exons. The location of the leader peptide (exon 1), the bait region (exons 17/18), the Gov\(^a\)/Gov\(^b\) polymorphism (exon 19), the thioester (exon 23), the reactivity determining triplet (VIH; exon 25), and the GPI cleavage/addition site (exon 33). Filled vertical bars/boxes, coding exon sequence; unfilled boxes, non-coding exon sequence; spaces between adjacent vertical bars/boxes, introns. K1, K1-H7, K1-3; three distinct, contiguous 3’ untranslated regions lying within exon 33, identified at the mRNA level.

**Figure 2-2. Phylogenetic relationship of human A2M family proteins.** The human \(\alpha2\)M, PZP, OVO, BAC, CD109, VIP, C3, C4, and C5 sequences were aligned, and used to generate an unrooted phylogenetic tree. CD109 and VIP form a branch that is intermediate in position to the \(\alpha2\)M/PZP/OVO/BAC and C3/C4/C5 branches. Abbreviations and (protein accession numbers) of sequences used: hs\(\alpha2\)M, *H. sapiens* \(\alpha2\)-macroglobulin (NP_000005); hsPZP, *H. sapiens* pregnancy zone protein (NP_002855); hsOVO, *H. sapiens*, ovostatin-like (XP_090334); hsBAC, *H. sapiens* unnamed protein product (BAC) (assembled from BAC85653, BAC04793, BAC85654 and BAC85494); hsVIP, *H. sapiens* VIP (NP_056507); hsCD109, *H. sapiens* CD109 (NP_598000); hsC5, *H. sapiens* complement component C5 (NP_001726); hsC4, *H. sapiens* complement component C4A/C4B (NP_009224/NP_000583); hsC3, *H. sapiens* complement component C3 (NP_000055). Numbers shown refer to % bootstrap values of 1000 trials.
**Figure 2-3. Conservation of genomic structure among human A2M family members.** The intron/exon structures of human *CD109, VIP, α2M, PZP, BAC, OVO, C3, C4, and C5* are shown schematically (and not to scale), with paralogous exons aligned. Non-aligned exons indicate the absence of paralogous exons in family members. Although the total number of exons varies among family members, marked conservation of genomic organization is evident. The *α2M/PZP/BAC/OVO* and *C3/C4/C5* genes define two conserved patterns of genomic organization, with *CD109* and *VIP* showing features of both. Corresponding mRNA accession numbers/genomic contigs used include *CD109* (AF410459/ NT_007299); *VIP* (NM_015692/ NT_011295); *α2M* (NM_000014/ NT_009714); *PZP* (NM_002864/ NT_009714); *BAC* (assembled from AK123591, AK122624, AK096448, and AK123592/NT_009714); *OVO* (XM_090334/ NT_009714) *C3* (NM_000064/ NT_011255); *C4* (NM_007293 and NM_000592/ NG_000013) and *C5* (NM_001735/ NT_008470).

**Figure 2-4. Conservation of genomic structure among CD109 orthologues.**

The intron/exon structures of human, chimp, mouse, horse, cow, dog, and chicken CD109 are shown schematically (and not to scale), with paralogous exons aligned. Bovine CD109 sequence is incomplete in the 5′ direction and is labelled with respect to human CD109. Corresponding mRNA accession numbers/genomic contigs used include human CD109 (AF410459/NT_007299); chimp (XM_518588/ NW_001236562); mouse (NM_153098/ NW_001581879); horse (XM_001498043/ NC_009153 & NW_001867364); gray short tailed opossum (XM_00137361/ NW_001581879); and chicken (XM_419879/ NW_001471671).
Figure 2-5. A cross-phyla A2M family phylogenetic tree. Human CD109 sequence was used to search public sequence databases for related non-redundant hits were then aligned with Clustal W (Thompson, Higgins et al. 1994) and an unrooted phylogenetic tree was generated using Mega2.1 (Kumar, Tamura et al. 2001). The A2M family comprises several distinct branches corresponding to the $\alpha$2M-related proteins, the vertebrate complement proteins (with distinct C3, C4, and C5-specific sub-branches), a group of outlying non-chordate deuterostome and cnidarian C3-like sequences, and a fourth branch containing CD109, *E. burgeri* HFPI, *C. intestinalis* $\alpha$2M, Anopheles gambiae TEP15, Trichoplax Adhaerens TEP, *C. elegans* ZK337.1, and the *D. melanogaster* TEP proteins. Abbreviations and (protein accession numbers [or *C. intestinalis* genome v1.0 gene model number*]) of sequences used: TgA2M, *Taeniopygia guttata* alpha-2-macroglobulin (XP_002190117); XtMug1, *Xenopus tropicalis* Murinoglobulin-1 protein (AAH91705); EuTTEP, *Euphaedusa tau* thioester containing protein (BAE44110); GgOVO, *Gallus gallus* Ovomacroglobulin (XP_423478); PtA2Mlike, *Pan troglodytes* alpha-2-macroglobulin-like (XP_520828); EqA2M, *Equus caballus* $\alpha$2M (XP_001499173); XlA2M, *Xenopus laevis* $\alpha$2M (NP_001089129); GgA2M, *Gallus gallus* $\alpha$2M (XP_425514); dmTEP4, *Drosophila melanogaster* TEP-4 (CAB87810); BtA2M, *Bos taurus* $\alpha$2M (NP_001103265); RnMUG2. *Rattus norvegicus* murinoglobulin-2(NP_001002826); RnAtI3 *Rattus norvegicus* $\alpha$(1)-inhibitor 3 (NP_001033064); DmTep3, *Drosophila melanogaster* TEP-3 (CAB87809); TgA2M, *Taeniopygia guttata* $\alpha$2M (XP_002190151); PtPZP2, *Pan troglodytes* pregnancy zone protein (XP_520824); MmMUG1, *Mus musculus* murinoglobulin-1 (XP_922625); SsA2M, *Sus scrofa* $\alpha$2M (XP_001925235); RnMUG1, *Rattus norvegicus* murinoglobulin-1 (XP_213006); HsPZP *Homo sapiens* pregnancy zone protein (NP_002855); RnA2M, *Rattus norvegicus* $\alpha$2M (AAA40636); MmA2M, *Mus musculus* $\alpha$2M (NP_783327); MmPZP, *Mus musculus* pregnancy zone protein (NP_031402); GgOVOP *Gallus gallus* ovomacroglobulin precursor (P20740); MjA2M, *Marsupenaeus japonicas* $\alpha$2M (BAC99073); MmCD109, *Mus musculus* CD109 (NP_694738); MdCD109, *Monodelphis domestica* CD109 (XP_001373651); GgCD109, *Gallus gallus* CD109 (XP_419879); TgCD109, *Taeniopygia guttata* CD109 (XP_002190719); EpBurgPI, *Eptatretus burgeri* protease inhibitor (BAD12264); monkeyCD109, *Macaca mulatta* (XP_001112898); CfCD109, *Canis familiaris* CD109
2.5 DISCUSSION

The two major arms of the vertebrate immune system - adaptive and innate immunity - have distinct evolutionary origins. While innate immunity is already present in protostomes, adaptive immunity is found only in jawed vertebrates (Dodds, Ren et al. 1996; Lambris, Reid et al. 1999; Nonaka 2001; Zarkadis, Sarrias et al. 2001). The emergence of this immunological distinctiveness of vertebrates is mirrored by similar changes in the complement system - important in both innate and adaptive immunity: The appearance of jawed vertebrates is paralleled by increases in the number of complement components and activation mechanisms, and by the appearance of a cytolytic (rather than merely opsonic) complement pathway (Dodds, Ren et al. 1996; Lambris, Reid et al. 1999; Nonaka 2001; Zarkadis, Sarrias et al. 2001). For example, the classical pathway of complement activation - intimately linked to adaptive immunity - is restricted to jawed vertebrates, and vertebrate complement can mediate both opsonic and cytolytic effects (Dodds, Ren et al. 1996; Lambris, Reid et al. 1999; Nonaka 2001; Zarkadis, Sarrias et al. 2001). In contrast, components of the alternative and lectin complement activation pathways have been found in a wider range of deuterostomes, including echinoderms, urochordates, and cyclostomes, and thus predate the emergence of adaptive immunity (Dodds, Ren et al. 1996; Lambris, Reid et al. 1999; Nonaka 2001; Zarkadis, Sarrias et al. 2001). Moreover, complement activation in invertebrates is believed to mediate primarily opsonization, with evidence for cytolytic complement in invertebrates limited to date to a single report of a C6-like gene in the cephalchordate amphioxus (Al-Sharif, Sunyer et al. 1998). An understanding of the evolution of the complement system is thus pivotal for understanding the evolution of vertebrate immunity. And in particular, the central position of
C3 in complement activation (the three complement activation pathways converge at C3) and function suggests that an understanding of its evolution is particularly relevant.

As α2M was identified in protostomes including molluscs (Bender, Fryer et al. 1992; Thogersen, Salvesen et al. 1992) and arthropods (Quigley and Armstrong 1985; Hergenhahn, Hall et al. 1988; Iwaki, Kawabata et al. 1996; Kopacek, Weise et al. 2000), while C3 was believed to be present only in deuterostomes (Al-Sharif, Sunyer et al. 1998; Nonaka and Azumi 1999; Nonaka, Azumi et al. 1999; Nonaka 2001; Holland and Lambris 2002), it has been proposed that C3 likely arose from an α2M-like ancestor by gene duplication and subsequent sequence divergence in a common deuterostome ancestor, with C4 and C5 arising later by subsequent gene duplication in early jawed vertebrates (Quigley and Armstrong 1985; Sottrup-Jensen, Stepanik et al. 1985; Hall, Soderhall et al. 1989; Hughes 1994; Bender and Bayne 1996; Al-Sharif, Sunyer et al. 1998; Armstrong, Melchior et al. 1998; Dodds and Law 1998; Nonaka and Azumi 1999; Nonaka, Azumi et al. 1999; Kopacek, Weise et al. 2000; Nonaka 2001; Zarkadis, Sarrias et al. 2001; Marino, Kimura et al. 2002; Suzuki, Satoh et al. 2002). The identification of CD109 and of additional human A2M family members, taken together with the identification of CD109-like proteins in insects, nematodes, urochordates and jawless fish, and additional C3-like proteins in invertebrate deuterostomes and cnidarians, has raised questions about this evolutionary model. I have attempted to clarify the position of CD109 in this schema.

The human A2M family comprises 9 members. By multiple sequence alignment and phylogenetic tree analyses (Figure 2-2) these proteins are shown to fall on three main
branches - an α2M branch that includes α2M, PZP, OVO, and BAC, a complement branch that includes C3, C4, and C5, and an “intermediate” branch that comprises CD109 and VIP. To understand further how this tree may have arisen by gene duplication, and to assess if one family member might be ancestral to the others, I used a number of complementary strategies to clarify the potential relationship of CD109 with other A2M family members. First, I considered the chromosomal locations of various human A2M loci, as close genetic linkage among genes might suggest a close evolutionary relationship based on tandem gene duplication. While α2M, PZP, OVO, and BAC were all located on chromosome 12p13, consistent with the close relatedness illustrated by tree analysis (Figure 2-2), remaining A2M family genes were not linked. Second, I determined the genomic structure of the CD109 locus, and used exon-specific sequence similarity analysis to identify the corresponding paralogous α2M, PZP, OVO, BAC, VIP, C3, C4, and C5 exons. Murine, rat, and human α2M, and the murine and human α2M-like genes MUG 1, 2, 3, and 4, and PZP are known to share a conserved genomic structure (the bait regions are in all cases encoded by exons 17 and 18, for example), consistent with divergence from one gene by gene duplication (Devriendt, Zhang et al. 1989; Hattori, Hamilton et al. 1989; Matthijs, Devriendt et al. 1992; Overbergh, Hilliker et al. 1994; Overbergh, Lorent et al. 1994). The genomic structure of the C3, C4, and C5 genes is also conserved, and in part resembles the organization of the α2M-related genes (Carney, Haviland et al. 1991; Haviland, Haviland et al. 1991; Vik, Amiguet et al. 1991; Yu 1991). I reasoned therefore that comparison of the intron/exon structure and the intron phase of CD109 with that of the A2M and complement genes, might clarify their potential evolutionary relationships. As shown in Figure 2-3 and Table 2-3, these analyses confirmed the overall
relatedness of these genes and demonstrated the existence of highly-conserved characteristic \(\alpha_2M/PZP/OVO/BAC\) and complement \(C_3, C_4, C_5\) patterns of exonic organization and intronic phase conservation, as expected. Notably, the organization of the \(CD109\) and \(VIP\) loci showed features of both the \(\alpha_2M\) and complement patterns, consistent with the intermediate position suggested by tree analysis (Figure 2-2). While a number of evolutionary scenarios could accommodate this pattern, I wondered whether the “intermediate” placement of this branch might reflect an ancestral relationship to the \(\alpha_2M\) and complement groups of proteins, as has been suggested previously for the \(CD109\)-like \(dmTEP1-4\) and \(ceZK337.1\) proteins.

I therefore proceeded to wider phylogenetic analysis. A phylogenetic tree generated using a representative spectrum of bilaterian, cnidarian, and placazooan \(A2M\) family protein sequences (Figure 2-5) revealed that the \(A2M\) family comprises several distinct branches corresponding to the \(\alpha_2M\)-related proteins, the complement proteins (with distinct \(C_3, C_4,\) and \(C_5\)-specific sub-branches and additional outlying sub-branches corresponding to \(C3\)-like molecules from cephalochordates, urochordates, and echinodermata, as well as to the cnidarian, \(Swiftia exserta\)), and a third branch containing \(CD109\). Within this latter branch, \(CD109\) is grouped most closely with the protease inhibitor \(HFPI\) from the hagfish \(E. burgeri\) (EpBurgPI), \(Branchiostoma floridæ\) protein 120995, \(Euphaedusa tau\) TEP, \(Chlamys farreri\) TEP, \(Trichoplax adhaerens\) protein TRIADDRAFT_55768, and more distantly with an \(\alpha_2M\)-like protein from the nematode \(C. elegans\) \(\alpha_2M\) (ceZK337.1), an \(\alpha_2M\) from the ascidian \(C. intestinalis\) (ci\(\alpha_2M\)), \(Ixodes scapularis\) TEP, the \(D. melanogaster\) TEP proteins (dmTEP1-4), \(Nasonia vitripennis\) TEP3, \(Anopheles gambiae\) TEP15, and \(Culex quinquefasciatus\) TEP3. Reciprocal best match analysis (RBM) confirmed that my search strategy had not inadvertently
excluded other potentially related proteins, thereby resulting in spurious measures of relatedness. Indeed, underscoring this relationship, I observed conservation of intron position among CD109, cia2M, dmMCR, ceZK337.1, and dmTEP genes (Table 2-5), suggestive of a common origin. Moreover, consistent with such relatedness, both dmMcr and ceZK337.1 are predicted to be membrane-bound via a GPI anchor, as is mammalian and avian CD109. Notably, in this wider analysis, VIP clusters not with CD109, but together with a VIP-like protein from C. intestinalis (ciVIP) comprises an α2M group sub-branch.

It has previously been suggested that the TEP proteins and the TEP-like ZK337.1 protein from C. elegans may comprise a separate branch of the A2M family (Lagueux, Perrodou et al. 2000; Levashina, Moita et al. 2001). Based on the observation that these proteins demonstrated structural features of both α2M and C3, it was suggested that they might represent prototypes of a putative common α2M/C3 ancestor. The apparent restriction of the TEPs and ZK337.1 proteins to protostomes was consistent with this notion and implied a specific evolutionary hierarchy within the A2M family. In this report, I extend this model by assigning CD109 and related proteins from urochordates, jawless fish, and placazoans to this archaic branch of the A2M family, using both sequence similarity and genomic structural criteria. However, although our A2M tree is unrooted and actual evolutionary relationships are indeterminate, based on the current sequence data set the CD109 branch is archaic, and likely ancestral to the complement and α2M branches. The CD109 branch contains the most phylogenetically diverse set of taxa, and notably, a member of the most basal taxa, Trichoplax adhaerans, lies on the CD109 branch.
While the $\alpha 2\text{M}$ and CD109/TEP branches contain both deuterostome and protostome members, the complement branch (with the exception of the single outlying cnidarian sequence) contains only deuterostome sequences. Moreover, within the complement branch, recognizable C3, C4 and C5 sequences are found only in chordates, while complement-like sequences from cephalochordates, urochordates and echinoderms form distinct branches ("C3-like", Figure 2-5), with the urochordate sequences seemingly the more closely related to those of chordates. And finally, the presence of the most phylogenetically distinct sequence - the cnidarian C3-like protein from the coral *Swiftia exserta* - suggests that the ancestral A2M family protein may have been a C3-like species found in a common ancestor of Cnidaria and Bilateria. In this light, while the actual evolutionary path is not known, a parsimonious explanation of the current sequence data set might be that the A2M family originated prior to the divergence of placazoa and cnidaria, with the later divergence of $\alpha 2\text{M}$-like and CD109-like branches in a common ancestor of protostomes and deuterostomes, and with subsequent ongoing divergence of complement sequences in craniate chordates (possibly coupled with the loss of complement-like sequences in protostomes). The expanded $\alpha 2\text{M}$ and complement families found in some vertebrates (4 $\alpha 2\text{M}$-like genes are found in humans, for example) indicate that ongoing gene duplication and divergence of function has continued during vertebrate evolution.

Can the origin of such an ancestral A2M gene be identified? Searching for additional archaic A2M-like proteins using the Smith-Waterman algorithm, I identified a distantly-related
hypothetical protein (NP_228792.1) from the bacterium *Thermotoga maritima* which shares 21% sequence identity with human CD109 over a 910 amino acid region (CD109 aa127-1027; tm0984 aa207-1016) and which contains putative amino terminal (CDD6592) and carboxy terminal (CDD5952) conserved A2M family domains, as determined by NCBI CDD search (Marchler-Bauer, Panchenko et al. 2002). Consistent with this, reciprocal BLASTP searching using NP_228792.1 as query identified CD109 and several other vertebrate and invertebrate A2M proteins, in addition to other bacterial proteins. However, as NP_228792.1 bears a variant putative thioester motif PYGCVEQ - albeit flanked by appropriately spaced proline and methionine residues known to be required for thioester formation (Isaac and Isenman 1992) - and does not contain an identifiable downstream thioester reactivity determining triplet or a bait region, it is not clear whether it behaves in an A2M-like manner or even if it possesses a functional thioester. Indeed, it is not known whether the bacterial cytosol is even capable of supporting the formation of an A2M-type thioester. Moreover, the other bacterial proteins most closely related to NP_228792.1 contain additional atypical putative thioester sequences, and do not contain flanking methionine residues in the position required for thioester formation (not shown). Thus, while it is possible that the A2M family may have a prokaryotic origin, the relationship of NP_228792.1 to known A2M proteins remains unclear. And in any case, such a putative relationship does not alter the subsequent evolution of A2M proteins within metazoans.

It has previously been suggested that the *D. melanogaster* TEP and *C. elegans* ZK337.1 group of proteins (which is now known to include CD109) might represent prototypes of a putative common α2M/C3 ancestor (Lagueux, Perrodou et al. 2000; Levashina, Moita et al.
The clustering of CD109, together with the only representative TEP from the most primitive organism surveyed – *Trichoplax adhaerans*, supports the notion of a CD109-like molecule as ancestral. However, phylogenetic analysis is dependent on the breadth of available data; as such, absolute identification of the ancestral A2M is dynamic. Nevertheless, in view of the fact that A2M evolution is exemplified by recurrent gene duplication and subsequent sequence divergence, often resulting in the formation of expanded gene families, the striking sequence conservation of the CD109-like proteins across phyla is remarkable, and likely attests to an important, and phylogenetically conserved function. *Anopheles gambiae* TEP 1 has been shown to promote bacterial opsonization and phagocytosis by mosquito Sua 5.1* cells in a thioester-dependent manner (Levashina, Moita et al. 2001). While the role(s) of CD109 in the vertebrate immune response remains unclear, the phylogenetic relationships that I have outlined here suggest that it may play a similar, highly conserved role in vertebrate immunity.
Chapter 3 Acknowledgement:

CHAPTER 3: MOLECULAR DETERMINATION OF GOV ALLOANTIGEN DETERMINANTS

3.1 INTRODUCTION

Platelets are anucleate cells whose primary role is to prevent blood loss at sites of vascular injury. Vascular damage is detected by platelets via the interaction of specialized cell surface glycoproteins with components of the extravascular matrix such as collagen and vWF. For many of these surface proteins, which are crucial to platelet physiology, alternative isoforms exist. As a result, transfusion of mismatched platelets may lead to the production of alloantibodies and immune-mediated destruction of platelets characteristic of post-transfusion purpura (PT) or platelet refractoriness (PR). Similarly, fetal or neonatal alloimmune-thrombocytopenia is caused by the transplacental passage of maternal antibodies directed against paternally-derived glycoproteins expressed on the surface of fetal or neonatal platelets. The majority of alloantibody mediated platelet destruction syndromes are the result of alloantibodies directed against the human platelet antigen (HPA) systems, namely: HPA-1, HPA-2, HPA-3, HPA-4, HPA-5, and Gov (now designated HPA-15) (Kelton, Smith et al. 1990; Bordin, Kelton et al. 1997; Berry, Murphy et al. 2000). Antibodies directed against HPA-1 are responsible for the vast majority of immune mediated platelet destruction syndromes followed by HPA-5, HPA-2, and HPA-3 (Berry, Murphy et al. 2000). The prevalence of Gov alloantibodies in alloantibody mediated platelet destruction syndromes, particularly in PR patients, is exceeded only by that of HPA-1 (Berry, Murphy et al. 2000).

The Gov alloantigen system is a co-dominant biallelic platelet alloantigen carried by CD109 (Kelton, Smith et al. 1990). Gov antigenicity is resistant to neuraminidase and N-
glycanase F/ Endo F treatment (Smith, Hayward et al. 1995). In contrast, antigenicity is lost following SDS denaturation (Smith, Hayward et al. 1995). Taken together, the Gov antigenic determinant was believed to reside within the peptide backbone itself rather than its carbohydrate components. Although the exact location of the Gov epitope was unknown, Kelton’s group described the Gov epitope as a 120kDa calpain digestion product of a 170kDa glycoprotein (which was later identified as CD109), which upon chymotrypsin digestion resides on a 52kDa product (Kelton, Smith et al. 1990; Smith, Hayward et al. 1995).

A variety of robust diagnostic methods existed to immunophenotype HPA-1 to 3 and HPA-5 (Simsek, Faber et al. 1993). However, reliable Gov alloantibody detection methods involved laborious radioimmunoprecipitation assays which are impractical in clinical settings (Kelton, Smith et al. 1990). An alternative method to immunophenotype Gov alloantibodies was the MAIPA assay (Simsek, Faber et al. 1993). Although MAIPA assays were routinely used to immunophenotype integrin and platelet GP, alloantigens Gov phenotyping was problematic. The MAIPA assay is dependent on the utilization of fresh platelets and an abundance of reference antiserum. CD109 expression is quite low (Kelton, Smith et al. 1990; Sutherland, Yeo et al. 1991) and displays marked inter-individual variability. Consequently, the inherent low expression of CD109 may preclude serological detection in some samples (Sutherland 1995; Bordin, Kelton et al. 1997). In addition, reference anti-serum is in limited supply and of variable quality. Further this antiserum frequently contains antibodies directed against ABH, HLA, and occasionally HPA. CD109 also carries ABH antigens on its backbone further compounding serological analysis (Kelton, Smith et al. 1998). As such, genotypic determination of Gov phenotype would alleviate these problems. We reasoned that the Gov
phenotype, akin to other platelet alloantigens, would be governed by a single nucleotide polymorphism. With this in mind, we sequenced cDNAs from donors of known Gov phenotype and identified an SNP at nt 1703 which defines Gov. This information was utilized to develop simple diagnostic assays which can be utilized in research and clinical settings.
3.2 OBJECTIVE AND AIMS

3.2.1 Rationale

- Gov alloantigenic determinants are insensitive to deglycosylation therefore the epitope is carried on the CD109 primary amino acid sequence. As such, we reasoned that as with most human platelet alloantigens, the Gov epitope is defined by single nucleotide polymorphism.

3.2.2 Objectives

- Determine Gov alloantigen determinants and design an accurate and precise molecular diagnostic method.

3.2.3 Specific Aims

- Identify SNPs through sequencing of DNA from serotyped donors
- Develop reliable screening method

3.2.4 Hypothesis

- Gov alloantigenic determinants are defined by a CD109 SNP
3.3 EXPERIMENTAL PROCEDURES

3.3.1 Donor panels

Donor panels consisted of 106 apheresis donors. Panel 1: CD109 cDNAs from 3 Gov\textsuperscript{aa} and 3 Gov\textsuperscript{bb} donors were sequenced to elucidate differences between the two alleles. Panel 2: Blood group O donors whose Gov phenotype was determined 2 or more times by monoclonal antibody-specific immobilization of platelet antigen assay (MAIPA) n=15 (5 Gov\textsuperscript{aa}, 5 Gov\textsuperscript{ab}, 5 Gov\textsuperscript{bb}). Panel 3: 85 donors whose Gov phenotype was determined once by MAIPA.

3.3.2 Gov phenotyping

Gov phenotyping was carried out using a modified MAIPA assay (Kiefel, Jager et al. 1987; Kiefel, Santoso et al. 1992) with well characterized polyclonal antisera and CD109 monoclonal antibody D2 (Berry, Murphy et al. 2000). Transiently transfected Chinese hamster ovarian cells were tested via MAIPA (Berry, Murphy et al. 2000) using the CD109 monoclonal antibody TEA2/16 (BD Pharmingen).

3.3.3 cDNA synthesis and analysis of CD109 sequence

Total RNA was extracted from Panel 1 donor platelets (1 X 10\textsuperscript{10}) using RNA Stat-60 messenger RNA extraction reagent according to the manufacturers description (Tel-Test, Friendswood, Texas, USA). First strand synthesis was carried out utilizing oligo-dT. Platelet cDNA was then used as a PCR amplification template in conjunction with gene specific primers to generate 8 overlapping contigs which spanned the entire CD109 ORF (Table 3-1). PCR amplicons were purified using QIAquick gel extraction kit (Qiagen, Mississauga, ON).
Sequencing reactions were carried out using the original gene specific primer sets and the Thermosequenase Cy5.5 dye-terminator kit (Amersham Biotech), visualized and analysed on the Open Gene Automated sequencing system (Visible Genetics, Toronto, ON). In parallel, PCR products were cloned into the pMAB1 plasmid. Resultant clones were analysed by alkaline lysis/restriction digestion, and sequencing. Taken together, each fragment and the entire ORF was sequenced at least twice in both directions.

Analysis of our K1 CD109 sequence revealed that it was Gov\(^a\). In order to generate a Gov\(^b\) clone, a 572bp region flanking nt 2108 was generated by PCR of cDNA derived from a Gov\(^{bb}\) donor using the following primers: K1-2019 (5’-GTGGACTCTGGGTATTGACAGATGC-3’) and K1-2591R (5’-CCGTTGGATTCTGATGTCC-3’). The amplicons were digested with BstXI and SacI (MBI Fermentas, Burlington, ON) into a 147-bp fragment that was then inserted into SacI (partial)/BstXI-digested pBS/K110, which was then designated pBS/K1b. DNA sequence analysis was used to verify the fidelity of the PCR-derived portion of the resultant Gov\(^b\) cDNA and the presence of a C at position 2108. A 2088-bp BsmBI/BstEII (MBI Fermentas) fragment of pBS/K1b was then inserted into the BsmBI/BstEII-digested expression vector pK1/YFP10, yielding pK1b/YFP. Restriction digestion (BstNI) was used to confirm the presence of the Gov\(^b\) allele in the final construct.

CHO cells were transfected with 10 µg pK1/YFP, pK1b/YFP, or control pIRES-EYFP (Clontech, Palo Alto, CA) plasmid DNA, using Lipofectamine and OPTI-MEM I medium (Life Technologies) as described (Lin, Sutherland et al. 2002) Cells were harvested as described at 40 to 45 hours and analysed by MAIPA.
3.3.4 DNA sequence analysis of introns flanking Gov\(^{a/b}\) polymorphism-bearing CD109 exon

The intron/exon junctions of the exon bearing the putative Gov polymorphism, as well as the nucleotide sequence of the flanking introns, were determined. Specifically, CD109 cDNA-specific oligonucleotides binding in the vicinity of this substitution were used for the direct sequencing of p4L10, a pCYPAC-1-derived PAC genomic clone bearing the human CD109 locus, followed by the alignment of cDNA and corresponding genomic sequences.

3.3.5 PCR-SSP analysis of Gova/b alleles

Gov\(^{a/b}\) allele-specific antisense primers (5'-TTCAAATTCTTGTAATCCTGT-3' [Gov\(^{a}\)]; 5'-TTCAAATTCTTGTAATCCTGG-3' [Gov\(^{b}\)]) were combined with a common sense primer binding in the adjacent intron (5'-ATGACCTATGACCTATTC-3'), yielding a 225-bp product. A 429-bp human growth hormone amplicon was used as a control (sense, 5'-GCCTTCCAACCATTCTGGT-3'; antisense, 5'-TCACGGATTCTGTGTTGTTTTC-3'). Genomic DNA (200 ng) was subject to 35 cycles of PCR (160 mM [NH\(_4\)]\(_2\)SO\(_4\), 15 mM MgCl\(_2\), 670 mM Tris-HCl pH 8.8, 2.5mM dNTPs, and 0.1% wt/vol Tween 20) with 0.35 U Taq DNA polymerase (Sigma, Poole, United Kingdom) using a Hybaid Omnigene thermocycler with heated lid. PCR products were size-separated electrophoretically in a 2.5% agarose/tris borate EDTA (TBE) gel containing 0.2 µg/mL ethidium bromide.

3.3.6 PCR-RFLP analysis of Gov\(^{a/b}\) alleles

Oligonucleotide primers 5'-TTTAGATTATTTTTGGCTT-3' (sense) and 5'-ATGGTTAGTTCCAGGTCA-3' (antisense) were used to generate a genomic DNA derived 448-bp
amplicon containing the Gov polymorphism. Resultant, amplicons were digested with BstNI (New England BioLabs, Pickering, ON) 5U for 90 minutes, 60°C and resolved on a 2.5% agarose/TBE gel.

3.3.7 Allele-specific real-time PCR analysis

Genomic primers and probes were designed for Taqman real-time PCR based genotyping. Reactions were performed with 100 nM each of the oligonucleotide primers 5'-TGTATCAGTTCTTGGTTTTGTGATGTT-3' (sense) and 5'-CCAAGAAGTGATAGAATCAGGTACAGTTAC-3' (antisense) in a total volume of 25 µL containing 1 µL genomic DNA. In addition, each reaction contained 100 nM of the FAM-labeled Gova-specific probe 5'-TATTATCTTGACTTCAGTTACAGGATTTAC CAAGAATTG-3' and 200 nM of the VIC-labeled Govb-specific probe 5'-TATTATCTTGACTTCAGTTCCAGGATTTACCAAGAAT-3'. Reactions were incubated at 50°C for 2 minutes and 95°C for 10 minutes, followed by 40 cycles of 95°C (15 seconds) and 64°C (60 seconds). Allelic discrimination was determined by using a post-PCR plate reader (PE 7700, Perkin Elmer).
3.4 RESULTS

Platelet mRNA from Gov immunophenotyped donors was used to generate CD109 cDNAs. The CD109 ORFs were sequenced in both directions independently at least twice. We reasoned that differences in the nucleotide sequence would define Gov phenotype independent of polymorphisms associated with other platelet GP or ABH antigens. Additionally, the CD109 polymorphism would correspond to the serologically determined Gov phenotype. Essentially, we would expect Gov\textsuperscript{aa} and Gov\textsuperscript{bb} individuals to be homozygous at both the nucleotide and serological level whereas Gov\textsuperscript{ab} individuals would be heterozygous. To this end, we identified a single nucleotide polymorphism at nucleotide 2108, which is either adenine (Gov\textsuperscript{a}) or cytosine (Gov\textsuperscript{b}) (Figure 3-1). This corresponds to a Tyr>Ser substitution at position 703 of the propeptide. In order to confirm that the SNP defines Gov status, putative Gov\textsuperscript{a} and Gov\textsuperscript{b} cDNAs were transiently expressed in CHO cells and assayed by MAIPA. Anti-Gov\textsuperscript{a} serum reacted exclusively with cells expressing Gov\textsuperscript{a} cDNAs and anti-Gov\textsuperscript{b} serum reacted with Gov\textsuperscript{b} expressing CHO cells (Figure 3-1b). As such, the A2108C nt, Y703S amino acid CD109 polymorphism defines the Gov alloantigen system.

Having identified a CD109 SNP associated with Gov phenotype we set our sights on developing a simple diagnostic assay. Although our original strategy to identify the CD109 SNP utilized platelet cDNAs, genomic DNA analysis is amenable to clinical diagnostics given its wider availability and greater stability. To this end the exon bearing nt 2108 and the flanking introns were identified and sequenced (Figure 3-1c). From this, a pair of allele-specific antisense primers were designed and used in conjunction with an intronic common sense
primer to detect the SNP by SSP. As shown in Figure 3-2, the primers amplified an allele specific 225bp product.

The Gov\(^b\) allele contains a BstNI restriction site at the SNP which we utilized to develop an RFLP assay. As shown in Figure 3-3, the primers amplified a 448bp CD109 product from genomic DNA. Digestion with BstNI resulted in an allele specific restriction pattern: Gov\(^{aa}\) 285 and 163; Gov\(^{bb}\) 285, 81, and 82; Gov\(^{ab}\) 285, 163, 81, and 82. Given their similar size the 81 and 82 bp fragments are unresolvable by conventional agarose gel electrophoresis. RFLP and SSP results demonstrated concordance with one another.

Although PCR-SSP and PCR-RFLP are a marked improvement over RIA or MAIPA as an assay to discern Gov status their applicability in a clinical setting is limited. As such, we developed 2108A and 2108C allele specific probes for Taqman based real time PCR. Taqman based genotyping eliminates cumbersome post-PCR steps associated with SSP/RFLP and allows for high throughput screening. We observed complete concordance with SSP and RFLP samples.
Figure 3 – 1a Gov SNP at nucleotide 2108
Figure 3-1c CD109 exon containing the Gov SNP.

101
Figure 3-2 Gov genotyping by PCR-SSP
Figure 3-3 Gov genotyping by PCR-RFLP
Figure 3-1. The Gov alloantigen system is defined by a CD109 SNP. (a) An A2108C; Y703S polymorphism was identified in the ORF of cDNAs derived from donors of known Gov phenotype. (b) Expression of Gov\(^a\)-YFP, Gov\(^b\)-YFP, or control-YFP constructs in CHO cells were analyzed by MAIPA using well characterized anti-Gov serum and control AB serum. (c) The non-coding introns flanking the A2108C were determined through sequencing of genomic DNA.

Figure 3-2. Gov genotyping by PCR-SSP. The Gov SNP (*) lies near the 5' end of the exon (open box). A sense primer was designed utilizing sequence found in the 5' intron. An antisense, allele-specific primer, was designed incorporating the Gov SNP. These two primers generate a 225 bp PCR product.

Figure 3-3. Gov phenotyping by PCR-RFLP. The Gov SNP (*) lies near the 5'end of the exon (open box). The Govb nt-A introduces a new BstNI recognition site (open triangle), in addition to a site common to both alleles (filled triangle). As such, BstNI digestion of the 448 bp amplicon, yields genotype-specific patterns. The position of the PCR primers used is is shown with respect to the locus and restriction map.
3.5 DISCUSSION

Platelet CD109 carries the biallelic codominant Gov alloantigen system. The relevance of the Gov system was demonstrated by the high incidence of anti-Gov alloantibodies in patients presenting with alloantibody-mediated platelet destruction syndromes (Smith, Hayward et al. 1995; Berry, Murphy et al. 2000). A study by our collaborator demonstrated that the incidence of alloimmune syndromes, especially in PR patients, attributed to anti-Gov alloantibodies is equivalent to that of HPA-5 and exceeded only by HPA-1 (Berry, Murphy et al. 2000). As such, inclusion of Gov immunophenotyping into platelet typing panels might be of benefit to subsets of transfusion patients. Traditionally, platelet immunophenotyping has been carried out by serological methods. However, serological approaches are problematic toward elucidation of Gov status for a number of reasons. The most reliable method to determine Gov phenotype was by radioimmunoprecipitation assay which is impractical in clinical diagnostic settings (Kelton, Smith et al. 1990; Smith, Hayward et al. 1995; Bordin, Kelton et al. 1997). As such we sought to develop a simple assay to reliably determine the Gov phenotype of platelets.

Given that the Gov epitope is resistant to enzymatic removal of carbohydrates and sensitive to SDS denaturation (Smith, Hayward et al. 1995) we reasoned that the CD109 peptide backbone itself carries the Gov phenotype determinant. In addition, a number of human platelet antigens were previously defined by single nucleotide polymorphisms (Newman, Derbes et al. 1989; Lyman, Aster et al. 1990; Kuijpers, Faber et al. 1992). Taken
together, we hypothesized that a *CD109* SNP defined the Gov alloantigen system determinants. Based on this notion, we synthesized cDNA from donor platelets of known Gov phenotype and subsequently identified an A>C SNP at nucleotide 2108 that translates to a Tyr>Ser substitution at amino acid 703 in the CD109 propeptide. Expression studies verified that this amino acid substitution does indeed give rise to the Gov alloantigens. PCR-RFLP, PCR-SSP, and Taqman based real time PCR assays distinguished Gov\(^{\alpha}\), Gov\(^{\alpha\beta}\), and Gov\(^{\beta}\).

CD109 expression is quite low in comparison to GPIIb/IIIa, GPIb-IX-V, and \(\alpha_2\beta_1\). In this study, we observed marked interindividual variability in antiserum reactivity. The Gov phenotyping by MAIPA of a few individuals was based on very low absorbance readings. As such, it was unsurprising that we observed a few Gov typing discrepancies. The ability of genotyping to resolve ambiguities is a testament to the sensitivity and accuracy of these approaches.

CD109 is a member of the A2M/C3/C4/C5 family of thioester containing proteins. This family of proteins is evolutionarily conserved across phylogenetically diverse taxa and CD109 is the most archaic among the human paralogs. The CD109 genes of species most closely related to humans are mono-allelic and encode a serine residue at the Gov position. Chimp, monkey, murine, marsupial, canine, equine, and swine CD109 proteins are equivalent to Gov\(^{\beta}\) at the position analogous to residue 708. Our group has searched extensively for a Gov SNP in murine *CD109* to no avail. Sequence analysis of more distal taxa identified additional *CD109*-like genes that are monoallelic Gov\(^{\beta}\). Of this subset, the most intriguing observation was the presence of a serine residue at the Gov position in the placazoan *Trichoplax adhaerens*. 

106
hypothetical protein TRIADDRAFT_55768. Although the evolutionary position of placazoans are subject to debate, overwhelming evidence supports their position as a basal eumetazoan lineage that diverged prior to the emergence of Cnidaria and Bilateria (Srivastava, Begovic et al. 2008). Taken together with the high frequency of $703^{\text{ser}}$ in several species, $\text{Gov}^b$ is likely the original CD109 allele in humans. CD109-like proteins from 2 species, $\textit{Gallus gallus}$ (chicken) and $\textit{Nasonia vitripennis}$ (jewel wasp), however, were paralagous to $\text{Gov}^a$, an observation that implied that the Ser>Tyr substitution has occurred independently in chickens, jewel wasps, and humans. The relevance of this substitution to the functionality of CD109 is as yet unknown.

Based on analogy with characterized A2M/C3/TEPs as well as with its attachment to the plasma membrane through a GPI-moiety, I suggest that CD109 might participate in cell-cell and/or cell-matrix interactions. Alternatively, CD109 may modulate platelet physiology through its interaction with components of the coagulation or inflammatory compartment. Analysis of the Tyr>Ser substitution does not present any obvious alterations to the CD109 peptide. The molecular weight of $\text{Gov}^a$ and $\text{Gov}^b$ peptides are identical, ruling out divergent glycosylation. Additionally, computer modeling suggested that any effect of the substitution on the protein would be minimal. The Tyr>Ser substitution is not predicted to alter CD109 thioester formation or reactivity. The Gov antigenic determinant lies outside of the bait region and C-terminal regions of the propeptide as such is not expected to alter the protease specificity. However, the increased immunogenicity of HPA-15$^b$ ($\text{Gov}^a$) over HPA-15$^a$ ($\text{Gov}^b$) implies that the Tyrosine residue does impart an alteration in secondary structure. Alternatively, the increased immunogenicity of $\text{Gov}^a$ might be influenced by the genetic
background of the responder immune cells (Valentin, Vergracht et al. 1990; Semana, Zazoun et al. 1996; Ahlen, Husebekk et al. 2009). Essentially, the Gov\textsuperscript{a} antigen might be processed and presented by antigen presenting cells more efficiently, particularly by specific HLA-II alleles, than the Gov\textsuperscript{b} antigen. Retrospective studies analysing the MHC class II allele expression of HPA-15 allo-immun thrombocytopenic patients might clarify the existence of such a relationship. Experimental evidence to determine whether HPA-15\textsuperscript{b} specific T cells can be isolated from HPA-15-immun thrombocytopenic patients and whether these clones are responsive to antigen pulsed monocytes/m\(\phi\) or immunocompatible HPA-15\textsuperscript{bb} platelets would further support the significance of the Gov system.

The frequency of HPA systems involved in alloimmune mediated platelet destruction syndromes varies among different populations. For example, in Caucasian (UK & N. America) populations, alloantibodies against HPA-1 are the most common cause of NAIT while HPA-4 is extremely rare (Berry, Murphy et al. 2000; Bhatti, Uddin et al. 2010). However, in Japanese populations, HPA-4 is the predominant alloantigen system involved in NAIT and HPA-1 is extremely rare (Matsuhashi, Tsuno et al. 2009). In 2003, the Gov\textsuperscript{a}/Gov\textsuperscript{b} alloantigen system was renamed HPA-15\textsuperscript{a/b}. Gov\textsuperscript{b} was designated HPA-15\textsuperscript{a} since, at the time, it was the most prevalent allele based on serological analysis of Caucasian populations (Metcalf, Watkins et al. 2003). Since the publication of our work, numerous groups have used our system or derivatives thereof to determine the Gov genotype of various populations across the globe (Cardone, Chiba et al. 2004; Halle, Bach et al. 2004; Higgins, Hughes et al. 2004; Ertel, Al-Tawil et al. 2005; Halle, Bigot et al. 2005; Tomicic, Bingulac-Popovic et al. 2006; Moncharmont, Courvoisier et al. 2007; Matsuhashi, Tsuno et al. 2009; Bhatti, Uddin et al. 2010). In most
instances, the prevalence of each phenotype tended toward a Mendelian distribution. In general, the majority of genotypic studies were in concordance to the original serological Caucasian populations (Brouk, Halle et al. 2010), however, in some ethnic groups, such as a Norwegian study, the Gov\textsuperscript{a} allele is in higher frequency (Randen, Sorensen et al. 2003). The overall immunogenicity of HPA-15 is considered to be low compared to that of other HPAs given the relatively low frequency of HPA-15 alloantibodies in platelet destruction syndromes. However, HPA-15 alloantibodies are more prevalent in PR compared with NAIT, and these are more commonly directed against HPA-15\textsuperscript{b}, consistent not only with the higher prevalence of HPA-15\textsuperscript{a} but also possibly with an inherently higher immunogenicity of HPA-15\textsuperscript{b} over HPA-15\textsuperscript{a}. The higher frequency of HPA-15 alloantibodies among polytransfused patients with PR is also of potential medical importance. The transfusion of Gov matched platelets may be particularly beneficial to patients who receive ongoing platelet transfusions – for example aplastic anaemia, leukaemia, or haematopoietic stem cell transplant patients. CD109 is also expressed on endothelial cells, activated T cells, and subsets of CD34\textsuperscript{+} stem and progenitor cells. Future studies are required to determine the effect, if any, Gov antigen immunization might have on allogeneic tissue transplantation.
CHAPTER 4: QUANTIFICATION, SUBCELLULAR LOCALIZATION, AND EXPRESSION OF PLATELET CD109

4.1 INTRODUCTION

Our group demonstrated CD109 expression on activated but not resting platelets. However, cell surface expression of CD109 on resting platelets has been reported by other groups. This discrepancy may have arisen as a consequence of platelet preparation methods, since platelets readily undergo mechanical activation. Furthermore, platelets are sensitive to pH fluctuation and adenine molecules released from inadvertently lysed red blood cells. Additionally, it is conceivable that the expression discrepancy was a consequence of differences in the epitopes of monoclonal antibodies utilized in each study. In any event, an agonist-induced increase of surface CD109 was observed regardless of CD109 status on untreated platelets.

Platelets are anucleate cells, derived from the fragmentation of megakaryocytes in the bone marrow. As such, the endogenous platelet proteome is predominately synthesized and sorted in megakaryocytes, culminating in cell surface expression of constitutive membrane proteins. In contrast, activation-dependent membrane proteins are expressed on the membranes of intracellular structures – the α-granules, δ-granule, and lysomes – in resting platelets. Following activation of platelets the granules fuse with the open canalicular system and are subsequently directed to the plasma membrane. Alternatively, constitutively expressed membrane proteins - such as the integrin GPIIb/IIIa - may undergo an activation-induced conformational change leading to the appearance of a previously hidden membrane epitope (the unmasking of the PAC-1 epitope of GPIIb/IIIa). A hierarchy of activation-induced antigen expression exists with the following order of appearance: epitope unmasking, α-
granule, δ-granule, and lysosomes. As such, we can determine the subcellular origin of an uncharacterized protein by comparing its activation kinetics and distribution with that of an activation marker of known subcellular origin.

The current study sought to gain basic information regarding the quantity, subcellular localization, and regulation of expression of CD109 in platelets. I reasoned that the discrepancy in CD109 status of resting platelets among investigators was a direct consequence of its status as an early activation marker. As such, I expected to find CD109 associated with α-granules in resting platelets, and perhaps the plasma membrane and OCS.
4.2 OBJECTIVES AND AIMS

4.2.1 Rationale

- Platelet activation-dependent membrane proteins are synthesized in megakaryocytes and sequestered in intracellular structures during megakaryopoiesis/thrombopoiesis. There are 3 types of platelet granules which, in response to agonist, are released in the following hierarchical order: α-granule, δ-granule, and lysosome. We can exploit this fact to infer the subcellular origin of an uncharacterized protein. Additionally, heterogeneity exists with respect to α-granular contents and their release following platelet activation with specific agonists. Taken together, the investigation of agonist-induced CD109 expression and subsequent comparison to that of characterized activation markers is of great interest.

4.2.2 Objectives

- Compare and contrast the expression and subcellular localization of CD109 and platelet activation markers.

4.2.3 Specific Aims

- Quantify platelet CD109 surface expression.
- Determine the activation-induced kinetics of CD109 expression in response to weak and strong agonists.
- Determine the subcellular origin of CD109 in unstimulated platelets.

4.2.4 Hypothesis

- Given that CD109 is expressed on activated but not resting platelets I might find CD109 associated with α-granules in unstimulated platelets. Additionally, I may observe agonist-induced CD109 cell surface appearance with kinetics that resembles that of another α-granule-restricted membrane protein, P-Selectin (CD62P).
4.3 EXPERIMENTAL PROCEDURES

4.3.1 Materials

Antibodies against CD62P (AK-4), CD61 (VI-PL2), CD109 (HU17) were from ebiosciences. PAC-1, CD41 (HIP8), and CD63 (H5C6) were obtained from BD Pharminogen (Mississauga, ON). CD109 moAbs 7D1 and 8A3 were raised against KG1a cells as previously described (Sutherland, Yeo et al. 1991). Anti-biotin IgG was from Cell Signalling Technology Inc (Beverley, MA, USA). P-Selectin, αIIb, LAMP-3, rabbit IgG were from Santa Cruz Biotechnology Inc. Secondary antibodies were from Invitrogen Molecular Probes (Mississauga, ON), BD Pharmingen (Mississauga, ON), Santa Cruz Biotechnology (Santa Cruz, CA, USA) and R & D systems (Minneapolis, MN, USA). Biotin was from Pierce (Rockford, IL, USA). The mounting medium ProLong gold was from Invitrogen Molecular Probes (Mississauga, ON). Prostaglandin E1, apyrase, poly-L-lysine, sepharose 2B, protein G fast flow sepharose, saponin, and A23187 were from Sigma-Aldrich (Oakville, ON). Collagen and bovine serum fraction V were from BioShop (Burlington, ON). Silver enhancement kit and BSAc were from Aurion (Wageningen, Netherlands). Western blots were developed using Amesham ECL Plus reagents (GE Healthcare; Baie d’Urfe, QC).

4.3.2 Blood collection

All subjects gave informed consent to participate in the study which was performed in accordance with local ethics regulations. Phlebotomy techniques were performed under conditions which limit platelet activation including the use of a 21-gauge needle fitted to wide barrel non pyrogenic tubing, tourniquet time < 1 minute, and an initial discard of the first 3 – 5
mL of blood. Blood (7 volumes) was collected into 1 volume of ACD in a sterile plastic centrifuge tube.

4.3.3 Platelet preparation

Whole blood was obtained from healthy donors (53 donors) by venipuncture and collected by free flow into acid citrate dextrose (whole blood to ACD ratio of 1:7). Subjects had not consumed aspirin or any other drug that might affect platelet activity within the past 10 days. Platelet-rich-plasma (PRP) was generated by centrifugation of anticoagulated blood at 110g for 15 minutes slow acceleration and no break. PRP was transferred to a Sepharose 2B column pre-equilibrated with Tyrodes buffer (137 mM NaCl, 2.68 mM KCl, 11.9 mM NaHCO$_3$, 0.42 mM NaH$_2$PO$_4$, 5.5 mM glucose, 1 mM MgCl$_2$, 5mM HEPES, 0.25% BSA) pH 6.5. Platelets were eluted from the column with Tyrodes buffer pH 6.5 or pH 7.3. In most instances, Tyrodes buffer was supplemented with 2 mM CaCl$_2$.

In some cases, platelets were isolated by differential centrifugation (washed platelets). For isolation of resting platelets, blood was collected into ACD containing Prostaglandin E1 (2μM). PRP was diluted 1:1 with PBS-ACD pH 6.5 (2μM PGE1, 7mM theophylline, 100 μM acetylsalicylic acid, apyrase) and held at room temperature for 10 minutes, followed by centrifugation at 900g slow acceleration no brake. The platelet pellet was then carefully washed twice with PBS-ACD.

4.3.4 Flow cytometric determination of epitope binding sites

Gel filtered platelets ($10^6$) were stimulated with thrombin receptor activating peptide (SFLLRN) at 0 – 100 μM as indicated for 15 minutes. Platelets were diluted 10 fold in PBS + 0.1% BSA, and incubated with monoclonal antibodies (5 μg/mL) directed against CD62P (AK-4),
CD63 (H5C6), CD109 (7D1, 8A3), or Isotype controls (IgG1k, IgG2a) for 15 minutes at ambient temperature. Immunolabelled platelets were immediately incubated with GAM-FITC (3.3 \( \mu \text{g/mL} \)) for 10 minutes at room temperature. Calibration beads were labelled in parallel under equivalent conditions. Platelets and calibration beads were diluted with FACS Analysis buffer and analysed immediately on a FACS Calibrator.

4.3.5 Flow cytometric analysis of agonist induced responses

Gel filtered platelets were prepared in Tyrodes buffer pH 6.5 and held at RT for 30 minutes until ready for assay. Platelets were treated with ADP, A23187, Collagen, or PAR1-AP at 37\(^\circ\)C for 10 minutes as indicated, in the presence of saturating concentrations of the appropriate antibodies. The reaction was stopped by the addition of 10 volumes of ice cold 1% paraformaldehyde and fixation was carried out for 20 minutes at RT. Fixed platelets were washed with wash buffer (0.1% BSA in PBS) 3 times and labelled with Alexa 488, FITC, or PE conjugated secondary antibodies for 10 minutes at RT. Platelets were subsequently washed 3 times, resuspended in wash buffer and analysed by flow cytometry.

4.3.6 Flow cytometric determination of intracellular pools of CD109

Resting and agonist treated GF-platelets were fixed for 20 minutes with 2% paraformaldehyde. Platelets were washed 3 times and treated with permeabilization buffer (1% saponin, 0.1% BSA, 0.1% FCS) for 10 minutes. Platelets were washed and incubated with appropriate antibodies diluted in blocking buffer (1% saponin, 1% BSA, 1% FCS) 1 hour at room temperature. Immunolabelled platelets were washed and labelled with Alexa 488, FITC, or PE conjugated goat-anti-mouse IgG diluted in secondary blocking buffer (1% saponin, 1% BSA, 5% FCS) 1 hour at room temperature.
4.3.7 Subcellular fractionation of platelet lysates

Platelets were isolated by differential centrifugation in the presence of inhibitors of platelet activation. Lysis was carried out by sonication, hypotonic lysis, or triton X-100 permeabilization in the presence of protease inhibitors (aprtinin, leupeptin, AEBSF, EDTA, and PMSF). For hypotonic lysis, platelet pellets were resuspended in gradient buffer (0.001M EDTA, 0.01M Tris-HCl pH 7.5, 0.15M NaCl) and passed through a 0-40% glycerol gradient at 1500g for 30 minutes. The resultant platelet pellet was treated with a cold hypotonic buffer (0.25M sucrose, 0.01M Tris-HCl, pH 7.5) to disrupt the plasma membrane while leaving the granules intact (Hermann, Rauch et al. 2001). In other instances, the platelet pellet was resuspended in gradient buffer and treated with 0.1% triton X-100, vortexed for 5 seconds, and immediately centrifuged at 1000g for 10 minutes to remove intact platelets. In all cases, debris and intact platelets were removed by centrifugation of the lysate at 2000 g for 10 minutes at 4°C. The supernatant was then layered on top of a discontinuous sucrose gradient (27% sucrose, d=1.106, 63 500g for 3 hours) or a linear sucrose density gradient (30 – 60% sucrose, 200 000 g for 2hr) and subjected to cold ultracentrifugation as indicated. Fractions were collected from the top of the gradient and dialysed against TBS and held at < -70°C until ready for assay. Aliquots from each sample were treated with SDS-PAGE sample buffer, boiled for 3 minutes, and loaded onto SDS-polyacrylamide gels (6, 8, or 10% w/v) for Western Blot analysis.
4.3.8 Immunoprecipitation

Subcellular fractions were pre-cleared with protein G beads for 1 hour at 4°C. Pre-cleared supernatants were incubated with 7D1, 8A3, HU17, or CD56 antibody for 1 hour on ice. The resultant immune-complex was then precipitated with protein G-sepharose Fast flow (1 hour, rotating, 4°C. The immunoprecipitate was washed with lysis buffer, followed by PBS, and was then treated with SDS-PAGE sample buffer, and loaded onto a SDS-PAGE, followed by Western blot analysis.

4.3.9 Immunofluorescence microscopic analysis of platelets

Gel filtered platelets were fixed in suspension by the addition of 4% paraformaldehyde and washed with TBS followed by PBS. Fixed platelets were allowed to adhere to poly-L-lysine coated slides for 10 minutes. The solid phase was washed with PBS, treated with 2% PFA (5 minutes), and washed 3 times with PBS. Non-specific binding sites were blocked by incubating slides with blocking buffer (PBS, 1% BSA, 1% FCS) for 30 minutes at room temperature. Platelets were subsequently immunolabelled with primary antibodies (α-CD109, 2.5 µg/mL; α-CD63, 10 µg/mL; α-CD61, 10 µg/mL; α-CD56, 2.5 µg/mL; α-IgG1k, α-IgG2a, 2.5 and 10 µg/mL) diluted in blocking buffer for 60 minutes or overnight followed by washing. In most instances, slides were subsequently treated with the appropriate secondary antibody (10 µg/mL) for 30 minutes at RT. Slides were washed three times with wash buffer, fixed with 4% paraformaldehyde (5 minutes, RT), washed, dried, and mounted in ProLong Gold. For labelling of intracellular antigens, platelets were permeabilized prior to immunolabelling with saponin (1% saponin, 0.1% BSA in PBS) for 10 minutes. All blocking, labelling, and washing steps were
carried out in the presence of saponin. Additionally, secondary antibodies were were diluted in buffer containing 1% BSA and 5% FCS.

**4.3.10 Immunoelectron microscopic analysis of platelets**

PRP was mixed 1:1 with PBS-ACD and subjected to centrifugation at 900g with slow acceleration and gradual deceleration at RT. The resultant platelet pellet was fixed with 4% PFA + 0.1% glutaraldehyde for two hours at RT and overnight at 4°C. In some cases, PRP was fixed by the addition of 1 volume 8% PFA + 0.2% glutaraldehyde for 60 minutes. Fixed PRP was washed sequentially with TBS and PBS both supplemented with 0.1% BSA. After the overnight fixation, platelets were washed with 0.15M glycine in PBS and embedded in gelatin (12% w/v) followed by overnight infusion with 2.3M sucrose. Samples were cut into 1mm cubes and frozen in liquid nitrogen onto pins for cryosectioning. Sections were cut with an ultramicrotome and picked up with a sucrose-methyl cellulose solution and placed onto nickel grids. Cryosections were then treated with warm PBS (10 minute; 30 - 40°C) to rehydrate sections and remove gelatin. Samples were subsequently treated with glycine and blocked with 5% cold water fish gelatin (CWFG) for 30 minutes. Cryosections were immediately treated with primary antibody (αCD109, 2.5 μg/mL, αCD63 2.5 μg/mL, αCD56 2.5 μg/mL) 30 minutes (RT) or overnight (4°C). Samples were subsequently washed 5 times with PBS and treated with gold conjugated goat-anti-mouse IgG (< 2nM or 10nM) diluted in 1% CWFG. Immunolabelled grids were washed (5 X 2 minutes, PBS), fixed 2% glutaraldehyde (10 minutes, RT), washed (2 X 2 min, PBS; 5 X 2 min, distilled water), embedded in 2% methylcellulose, and stained with 0.4% uranyl acetate. When ultra small gold (<2 nm) conjugate secondary
antibodies were used, non-specific binding sites were blocked with BSA-c™ (acetylated BSA; AURION Immunogold Reagents, The Netherlands), rather than with CWFG. Additionally, a silver enhancement step was required to visualize ultra small gold particles. Grids were viewed with a FEI Tecnai 20 Transmission Electron Microscope.

4.3.11 Statistical Analysis

All data are expressed as mean ± SEM of n independent experiments. Comparison of the data were performed using Student’s two-tailed T test using GraphPad Prism Instat Version 4.01 for Windows (GraphPad Software, San Francisco, CA, USA). An associated probability value <0.05 was considered to be significant.
4.4 RESULTS

4.4.1 Quantification of CD109 levels on PAR1-AP stimulated platelets

Surface expression of platelet activation markers was determined after stimulation through Proteinase Activated Receptor 1 using a flow cytometric quantification assay. For each activation marker tested, increasing doses of agonist led to a rise in cell surface expression (Figure 4-1). Increased CD109 surface expression was observed at the lowest agonist dose (Figure 4-1c) whereas 30μM of PAR1-AP was required to increase significantly CD62P and CD63 expression (Figure 4-1a, 4-1b). This observation suggested that complete degranulation of platelets is not required for the agonist induced expression of cell surface CD109. Quantitative analysis revealed that in comparison to the cell surface expression of α-granule and δ-granule markers, that of CD109 was quite low. High dose stimulation of platelets resulted in an approximately 60-fold increase over background for CD62P expression, whereas CD109 displayed only about a 5-fold increase.

4.4.2 Cell surface CD109 is upregulated on suboptimally stimulated platelets

Next I investigated the expression of agonist-induced activation epitopes on platelets stimulated with the agonists: collagen, ADP, or epinephrine. In this study, platelets were treated with agonists in the presence of saturating concentrations of antibodies directed against CD62P, CD63, CD109, or fibrinogen. Agonist-induced cell surface expression was subsequently compared to that of platelets stimulated fully with PAR1-AP. When platelets were stimulated with ADP, a weak agonist, resultant CD109 cell surface expression was approximately 75% of that seen in fully activated platelets (Figure 4-2d). In contrast, α-granule
(Figure 4-2b) and δ-granule (Figure 4-2c) marker cell surface expression did not exceed 15% of the PAR1-AP level. Stimulation of platelets with soluble collagen, in contrast, led to activation of GPIIb/IIIa exceeding that observed on PAR1-AP stimulated platelets (Figure 4-3) consistent with the notion of integrin cross-talk (Porter and Hogg 1997; Porter and Hogg 1998; Ortiz-Stern and Rosales 2003). Cross-talk refers to the ability of cell surface molecules to interact with one another and subsequently modulate each other's activity. Essentially, collagen-induced activation of $\alpha_2\beta_1$ can result in the generation of an inside-out signal which promotes activation of $\alpha_{IIIb}\beta_3$. Curiously, half-maximal CD109 expression was achieved at the lowest dose of collagen whereas >10 $\mu$g/mL of collagen was required to obtain a similar increase in CD62P and CD63 expression. Interestingly, epinephrine treatment of platelets resulted in significant rise of cell surface CD109 expression (Figure 4-4). However, epinephrine did not induce an appreciable elevation of granule marker surface expression or of GPIIb/IIIa activation.

4.4.3 CD109 activation kinetics

In order to infer subcellular origin, the activation kinetics of CD109 were determined and compared to those of $\alpha$-granule and δ-granule markers, and markers of PAC-1 epitope unmasking. In this assay, the slope of the curve reflects the strength of the agonist induced response. The hierarchy of ADP induced cell surface expression observed was as follows: PAC-1>CD109>CD62P/CD63 (Figure 4-5). Of note, unlike the expression of the granule markers, CD109 cell surface expression approached the levels displayed on fully-activated platelets. I then stimulated platelets sub-optimally with PAR1-AP, under conditions expected to induce
degranulation of $\alpha$-granules and the secretion of a subset of $\delta$-granules. As expected, CD62P surface expression at 20 minutes (Figure 4-6c) was similar to that observed on 100$\mu$M PAR1-AP-treated platelets at 5 minutes (Figure 2-1a). Consistent with this, CD62P upregulation was faster than that of CD63, while CD109 PAR1-AP induced expression was intermediate to both. GPIIb/IIIa activation was extremely rapid in response to PAR1-AP, with an EC$_{50}$ below 1 minute (Figure 4-6a). Fibrinogen receptor activation, as defined by PAC-1 binding, plateaued at 2.5 minutes, and declined beyond 5 minutes, possibly indicating receptor internalization. Platelets were next stimulated with the calcium ionophore A23187. The A23187 activation kinetics of CD109 cell surface expression followed the CD62P pattern (Figure 4-7). In contrast, only a subset of $\delta$-granules were released during the course of this assay. Taken together, with the ADP and PAR1-AP dose response curves, these data indicate that CD109 likely resides within $\alpha$-granules that are secreted early during platelet activation.

4.4.4 Flow cytometric analysis reveals a subcellular pool of CD109

In order to determine whether there is an intracellular CD109 reserve in platelets, flow cytometric analysis was performed on intact and permeabilized platelets (Figure 4-8). As expected, CD109 expression was absent on resting platelets, and was upregulated following Ca$^{2+}$ ionophore stimulation (Figure 4-8). CD109 mAb binding to unstimulated saponin treated platelets exceeded that of A23187 treated intact platelets, and was equivalent to that of A23187 stimulated, permeabilized platelets. Taken together, therefore, these observations indicate the existence of an intracellular pool of CD109 within platelets.
4.4.5 CD109 codistributes with granular and membrane fractions in sucrose density gradients

Subcellular fractionation of platelet lysates was performed in order to separate platelet organelles from the bulk of plasma membrane and cytosolic proteins, and thereby further elucidate the intra-cellular location of CD109. Three complementary methods were utilized to obtain subcellular fractions which were subsequently analysed by Western blot. In the first method, a sonic lysate was subjected to density centrifugation through a linear sucrose density gradient. CD109 was found predominately in the low-density fractions (0 – 40% sucrose) which were expected to contain cytosolic and membrane proteins (Figure 4-9). A smaller proportion of CD109 was found within the high-density fractions which likely contain α-granules and δ-granules. As expected, the bulk of CD62P and CD63 was found within the high-density fractions (> 45% sucrose) validating my organelle separation technique. In addition to its constitutive cell surface expression, GPIIb/IIIa (αIIb/β3) is also found on α-granule membranes and translocation of this granular reserve is responsible in part for the upregulation of GPIIb/IIIa following platelet activation. Consistent with this notion, integrin αIIb was distributed predominantly in high-density, but also in low-density, fractions.

In order to delineate whether the CD109 found in the low-density fraction of the linear sucrose gradient, was cytosolic or membrane-associated in origin, a hypotonic platelet lysate was subjected to discontinuous sucrose (27% w/v) density gradient centrifugation (Figure 4-10a). In this system, soluble proteins do not enter the gradient, membrane proteins settle at or near the interface, and granules are found in the pellet. CD109 co-distributed with P-selectin and hence likely resides in α-granules. A CD109 signal was also present within the
interface as was P-Selectin, to a lesser extent. The presence of CD109 and CD62P in the membrane fraction may have been the result of activation of a subset of platelets. Alternatively, lysis of a portion of granules may have led to spurious distribution of CD109 and P-Selectin. However, given the presence of protein signal within the discontinuous gradient fraction, the discrepancy was likely the result of lysis of a portion of the granules, rather than of platelet activation. To confirm this, a platelet homogenate was therefore generated by solubilising gently the platelet plasma membrane with detergent. In gently lysed platelet homogenates, CD109 co-distributed predominantly in the granular pellet and was found to a lesser extent in the membrane fraction (Figure 4-10b). In contrast to the hypotonic lysate, CD109 was absent from the discontinuous gradient fraction, confirming that granules were not lysed by gentle detergent solubilization. Taken together, therefore, based on cell fractionation, CD109 resides within intracellular granules and is associated with the membranous open canalicular system (OCS) of resting platelets.

4.4.6 Subcellular localization of CD109 by immunofluorescence microscopy

A more direct approach was utilized to determine the subcellular localization of platelet CD109. I studied the subcellular localization of CD109 by immunofluorescence confocal microscopy. In unstimulated platelets a granular pattern was observed for both CD63 and CD109 (Figure 4-11). Following stimulation with Ca\(^{2+}\) ionophore, CD109 and CD63, were both expressed on the platelet surface (Figure 4-11e).
4.4.7 Immunoelectron microscopic detection of platelet CD109

Although the resolution of confocal microscopy is a marked improvement over conventional immunofluorescence microscopy, the technique remains incapable of resolving ultrastructural details. I therefore used electron microscopy to study the subcellular localization of CD109 in immuno-gold labelled resting platelet cryosections. Platelet cryosections were labelled with mAbs directed against CD109 (8A3, 7D1, W7C5, or HU-17), followed by a gold-conjugated secondary antibody. Frozen sections were used, as this approach is known to preserve the antigenicity of a larger subset of proteins than are protocols involving dehydration and plastic embedding. Platelet ultrastructure was largely preserved as evidenced by the presence of intact membranes, mitochondria, and discernable α-granules. However, in the frozen section approach, the δ-granules lumen contents are lost, resulting in the elimination of their dense character. With this in mind, mAbs against CD63 were used as both a positive control, and as a marker of δ-granules, which as expected under these conditions, appeared as vacuoles (Figure 4 – 12a). In instances in which platelets were labelled with CD109 mAbs, gold particles were observed in clusters adjacent to the inner face of the plasma membrane (Figur 4-12c). Gold particles were also observed in thin channels (Figure 4-12b,d) which in some instances emanated from α-granules toward the plasma membrane (figure 4-12d). We did not observe direct CD109 labelling on α-granule or δ-granule membranes. Given that cryosectioning will sample only a very small portion of an individual granule or platelet (immunogold labeled cryosections are ~50nm thick, whereas
platelets are at least 1μm thick), taken together, with the low levels of expression of CD109 in platelets, it is not at all surprising that immunogold labelling was unsuccessful.
Figure 4-1
Figure 4-2.
Figure 4-3
Figure 4-4
Figure 4-5
Figure 4-6
Figure 4-7
CD109
Figure 4-9
Figure 4-10
Figure 4-12
Figure 4-1. Flow cytometric analysis of PAR-1 activating peptide induced platelet expression of CD62P (a), CD63 (b), and CD109 (c). Gel-filtered platelets were stimulated with (0 – 100 μM) PAR1-AP and labelled sequentially with indicated mAb and FITC-conjugated secondary antibody. Immunolabelled platelets were analysed by flow cytometry and the number of binding sites were determined from a standard curve. Calibration beads were used to generate a standard curve (mean fluorescence intensity vs. antibody binding sites. Histograms represent the number of antibody binding sites of a representative donor. Values are mean ± SEM, n=7 independent experiments, * p<0.05 versus unstimulated controls.

Figure 4-2. Expression of activation expression markers on ADP-stimulated platelets. Gel filtered platelets were stimulated with ADP (0 – 10 μM) in the presence of antibodies directed against fibrinogen (a), CD62P(b), CD63(c), and CD109(d). Platelets were fixed with paraformaldehyde and labelled with secondary antibodies and analysed by flow cytometry. Histograms represent the relative mean fluorescence intensity compared to stimulation with PAR1-AP (100µM) for a representative donor. Values are mean ± SEM, n=7 independent experiments, * p<0.05 versus unstimulated controls.

Figure 4-3. Expression of activation markers on collagen-stimulated platelets. Gel filtered platelets were stimulated with collagen (0 – 20 mg/mL) in the presence of antibodies directed against fibrinogen(a), CD62P(b), CD63(c), and CD109(d). Platelets were fixed with paraformaldehyde and labelled with secondary antibodies and analysed by flow cytometry. Histograms represent the relative mean fluorescence intensity compared to stimulation with PAR1-AP (100µM) for a representative donor. Values are mean ± SEM, n=7 independent experiments, * p<0.05 versus unstimulated controls.

Figure 4-4. Expression of activation markers following epinephrine stimulation of platelets. Gel filtered platelets were stimulated with epinephrine (0 – 10 μM) in the presence of antibodies directed against fibrinogen(a), CD62P(b), CD63(c), and CD109(d). Platelets were fixed with paraformaldehyde and labelled with secondary antibodies and analysed by flow cytometry. Histograms represent the relative mean fluorescence intensity compared to stimulation with PAR1-AP (100µM) for a representative donor. Values are mean ± SEM, n=7 independent experiments, * p<0.05 versus unstimulated controls.

Figure 4-5. Activation induced marker expression kinetics in ADP stimulated platelets. Gel-filtered platelets were stimulated with ADP (25 μM) in the presence of monoclonal antibodies directed against PAC-1(a), CD63(b), CD62P(c), CD109(d). The reaction was stopped with PFA at the indicated time points and platelets were analysed by flow cytometry. Values represent mean ± SEM, n = 7 independent experiments.
Figure 4-6. Activation induced marker expression kinetics in PAR1-AP stimulated platelets. Gel-filtered platelets were stimulated with PAR1-AP (25 µM) in the presence of monoclonal antibodies directed against PAC-1(a), CD63(b), CD62P(c), CD109(d). The reaction was stopped with PFA at the indicated time points and platelets were analysed by flow cytometry. Values represent mean ± SEM, n = 7 independent experiments.

Figure 4-7. Activation induced marker expression kinetics in A23187 stimulated platelets. Gel-filtered platelets were stimulated with A23187 (0.2 µM) in the presence of monoclonal antibodies directed against PAC-1(a), CD63(b), CD62P(c), CD109(d). The reaction was stopped with PFA at the indicated time points and platelets were analysed by flow cytometry. Values represent mean ± SEM, n = 7 independent experiments.

Figure 4-8. Flow cytometric analysis of intracellular pool of CD109. Immunolabelling was performed with unstimulated fixed platelets (red), A23187 stimulated fixed platelets (brown), unstimulated fixed saponin permeabilized platelets (green), and A23187 stimulated fixed saponin permeabilized platelets (blue).

Figure 4-9 CD109 codistributes with α-granule and δ-granule markers on linear sucrose density gradient. Resting platelet sonicates were separated by linear density gradient centrifugation. Fractions were probed with antibodies directed against CD62P, CD63, CD109, and integrin αIIb.

Figure 4-10. CD109 codistributes with α-granule markers on a discontinuous sucrose density gradient. Platelet hypotonic lysates (a) or gently solubilised lysates (b) were separated by discontinuous sucrose gradient (27% w/v sucrose). Five fractions were collected from the top and analyzed by SDS-PAGE and Western Blot.

Figure 4-11 Confocal microscopy of platelets. Activated platelets were immunolabelled for CD109 (green), and CD61 (red), panels A – C. Resting and A23187 platelets were permeabilized with saponin and immunolabelled with CD109 (red). Immunofluorescence detection following saponin treatment demonstrated a granular pattern of CD109 distribution in unstimulated platelets (D). A23187-induced activation prior to saponin mediated permeabilization results in the transition to a cell surface arrangement (E).
Figure 4-12. Detection of CD109 in resting platelets by immunogold labelling of frozen ultrathin sections using monoclonal antibodies. Alpha-granules are visible under this system, δ-granules were detected based on CD63 labelling (A). CD109 was detected in the OCS in clusters in close proximity to the plasma membrane (C); and in thin channels (B, D) which in some cases appeared to bridge α-granules and the plasma membrane (D).
4.5 DISCUSSION

CD109 is a 170 kDa GPI-anchored member of the A2M/complement C3/C4/C5 protein superfamily (Sutherland, Yeo et al. 1991; Lin, Sutherland et al. 2002). CD109 is expressed in fetal and adult CD34+ bone marrow cells subsets, mesenchymal stem cell subsets, activated T lymphoblasts, thrombin-activated platelets, endothelium, epithelium, megakaryoblastic leukemia cells, and human tumour cell lines (Brashem-Stein, Nugent et al. 1988; Haregewoin, Solomon et al. 1994; Murray, Bruno et al. 1999; Giesert, Marxer et al. 2003). Additionally, CD109 is elevated in squamous cell carcinomas of the lung, oesophagus, and uterus (Hashimoto, Ichihara et al. 2004; Zhang, Hashimoto et al. 2005).

At the beginning of my studies, the function of CD109 was largely unknown. What was known, was that CD109 antibodies abrogated responder cell proliferation in a mixed lymphocyte response assay (Suciu-Foca, Reed et al. 1985), and that platelet CD109 carried the antigenic determinant of the Gov alloantigen system in its peptide backbone (Kelton, Smith et al. 1990). This current study sought to elucidate basic information regarding the quantity, expression and subcellular localization of CD109 in platelets. Upregulation of cell surface molecules is achieved through the exocytosis of intracellular granules which themselves bear receptors and adhesion molecules on their vesicular membranes. For example, P-selectin (CD62P) is expressed on the surface of activated but not resting platelets and is found on the vesicular membrane of α-granules in unstimulated platelets (Stenberg, 1985). Although anucleate, platelets are indeed capable of processing pre-mRNA and translating mRNA into protein (Denis, Tolley et al. 2005; Schwertz, Tolley et al. 2006). In this study, I showed that
CD109 is expressed on the surface of activated platelets at very low levels in most healthy volunteers. I demonstrated further that CD109 is localized to the OCS and intracellular storage granules (most likely α-granules) in resting platelets. CD109 is expressed rapidly on the cell surface in response to low doses of both strong and weak agonists, likely reflecting its proximity to the plasma membrane in resting platelets.

As expected, our results confirmed that the expression of CD109 is quite low on PAR1-AP stimulated platelets. Notably, however, PAR1-AP-induced CD109 surface expression differed from both the α-granular marker CD62P and the δ-granule marker CD63. Specifically, the appearance of CD109 on the cell surface at low doses of PAR1-AP suggested either a membrane, OCS, or α-granular origin. In addition, ADP-, collagen-, and epinephrine-induced CD109 surface expression also varied from CD62P, CD63, and GPIIb/IIIa activation. When platelets were stimulated with ADP – a weak agonist – 75% of the total CD109 pool was transferred to the cell surface, whereas CD62P and CD63 expression did not exceed 15%. Interestingly, half maximal CD109 expression was also achieved at the lowest dose of collagen, whereas >10 μg/mL of collagen were required to obtain a similar increase in CD62P and CD63 expression. The divergent CD109 collagen dose-response relationship, compared to those of CD62P, CD63, and PAC-1, was of particular interest as it may indicate specific collagen-receptor colocalization, or cross-talk. Notably, stimulation of platelets with soluble collagen led to activation of GPIIb/IIIa which exceeded that observed in PAR1-AP stimulated platelets, consistent with the notion of integrin cross-talk. The main collagen signalling receptor GPVI, and its associated FcR γ-chain - which provides the cytoplasmic residues to transmit extracellular signals to the cytosol - are localized in lipid rafts on the plasma membrane (Ezumi,
Kodama et al. 2002; Locke, Chen et al. 2002; Wonerow, Obergfell et al. 2002). Lipid rafts provide a platform on which receptors and signalling complexes can co-localize and assemble, thereby enhancing their signalling (Zajchowski and Robbins 2002). For example, in addition to the collagen receptor, a subset of GPIb-IX-V (VWF receptor) is found within lipid rafts in resting platelets, and following platelet activation the majority of GPIb-IX-V is localized in lipid rafts (Shrimpton, Borthakur et al. 2002). The primary physiological roles of GPIb-IX-V and GPVI are to mediate the initial adhesion of platelets to the exposed subendothelium, via endothelial VWF at high shear or via collagen at both low and high shear. In addition to lipid raft-mediated cross-talk between integrins, a number of examples of cross-talk between an integrin and another receptor are also known, and it is possible that lipid raft-restricted CD109 may interact with GPVI or GPIb-IX-V(Blystone, Graham et al. 1994; Pacifici, Roman et al. 1994; Diaz-Gonzalez, Forsyth et al. 1996; Porter and Hogg 1997; Porter and Hogg 1998; Riederer, Ginsberg et al. 2002; Ortiz-Stern and Rosales 2003). Taken together, with its inherent localization in LRM through the GPI-anchor, CD109 is conveniently positioned to participate in and modulate adhesion events. CD109 might modulate adhesion by influencing the protein character of the LRM. Perhaps a CD109 modulatory effect might explain the fact that a fraction of platelets are capable of undergoing stable adhesion through GPVI alone rather than in the presence of additional tethering through $\alpha_{IIb}\beta_3$, $\alpha_2\beta_1$, or GPIb-IX-V (Auger, Kuijpers et al. 2005). Alternatively, proteolytically activated CD109 may form covalent bridges between platelet and the substratum or adjacent platelets. The proteolytic activation of CD109 itself might in turn might lead to the transmission of signals which modulate platelet function.
In order to extract information regarding subcellular origin, I compared the kinetics of agonist-induced CD109 expression to those of α-granule and δ-granule markers, and to an activation-induced membrane protein unmasking marker, in response to weak and strong agonists. Consistent with the previously observed dose-response relationships, CD109 displayed distinct activation-induced kinetics in comparison to the established platelet activation markers, PAC-1, CD62P, and CD63. At the sub-optimal doses of ADP utilized in this study, almost the entire CD109 storage pool was transferred to the cell surface, suggesting sequestration of CD109 in a subset of α-granules that is secreted early. However, the kinetics of PAR1-AP-induced CD109 surface expression, were somewhat intermediate to those of the α- and δ-granule markers. To resolve this ambiguity, I activated platelets with calcium ionophore. Platelet granule secretion is highly dependent on the spike in intracellular calcium levels that follows platelet activation. Treatment of platelets with A23187 ionophore maximally raises intracellular Ca\(^{2+}\) and induces complete degranulation of platelets (Levy-Toledano, Maclouf et al. 1982; Martina, Kilic et al. 1994; Wang, Sada et al. 1994). When platelets were stimulated with A23187, CD109 cell surface display, but not of CD63, closely resembled the A23187-induced expression profile of CD62P, supporting the notion of an α-granule origin for CD109.

A more direct imaging approach was also utilized to determine the subcellular origin of CD109, and to explain the distinct expression pattern observed following stimulation with diverse agonists. Immuno-electron microscopy of resting platelets detected CD109 expression in thin cytoplasmic channels – the OCS - and in regions in close proximity to the plasma membrane. While my flow cytometric, cell fractionation, and confocal microscopy
data suggested that CD109 is also present in intracellular storage granules, I was unable to observe CD109 by EM in the α-granules of platelet cryosections, even when a number of signal amplification techniques were used. The low expression levels of CD109 (~500 – 2000 molecules per platelet), compared to that of CD62P and CD63, likely limits the likelihood of encountering the molecule in a given cryosection, and likely explains this discrepancy. It is also conceivable that granular CD109 exists in a complex/conformation which masks the antibody binding epitopes, whereas CD109 in the OCS is available for binding to our antibodies. The ability to detect granular CD109 by confocal microscopy makes this latter possibility less likely, however.

The perceived primary role of the open canalicular system is to increase the surface area of a platelet following activation and act as a conduit for granule secretion (Klinger 1996; Choi, Karim et al. 2010). Platelets from species lacking OCS display delayed aggregation emphasizing the importance of these surface connected channels to platelet function (Choi, Karim et al. 2010). A number of platelet integrins and receptors are known to be expressed on the OCS as well as within α-granules of resting platelets (Berger, Masse et al. 1996; White, Krumwiede et al. 1999; Nurden, Poujol et al. 2003; Suzuki, Murasaki et al. 2003). The presence of these receptors on the OCS provides a mechanism to rapidly upregulate their cell surface expression at sites of vascular injury. Notably, in resting platelets, the majority of TPα and P2Y₁ are found within the OCS and α-granules, rather than the plasma membrane (Nurden, Poujol et al. 2003). The presence of these receptors on the OCS allows the activated platelet to maintain or elevate receptor density following the increase surface area that is associated with the morphological changes post-platelet activation. Rapid elevation of these receptors is
crucial to promoting stable aggregation given that their ligands – TxA₂ and ADP, respectively - are liberated after the initial adhesion of platelets to the extravascular substratum. By analogy with this notion, expression of CD109 at early stages of platelet activation may similarly ensure that CD109 is expressed on the cell surface prior to secretion of candidate CD109 ligands from α-granules. Additionally, CD109 might play a role in co-ordinating the molecular events on the plasma membrane in a lipid-raft restricted manner. A variety of platelet proteins reside within lipid rich microdomains (Locke, Chen et al. 2002; Wonerow, Obergfell et al. 2002; Zajchowski and Robbins 2002). Some of these proteins undergo increased sequestration in lipid rafts following platelet activation. Prothrombinase assembles on the activated platelet surface and generates thrombin. Thrombin, in turn activates resting platelets, and converts fibrinogen to fibrin. In order, to limit the size and extent of the platelet plug, thrombin activity must be tightly regulated. It is conceivable, therefore, that CD109 is susceptible to cleavage by thrombin (or by some other platelet protease) and may serve as mechanism to limit the activity of thrombin or the platelet response to thrombin. A variety of other platelet derived proteases are also secreted or activated following platelet stimulation, including cathepsins, calpain, matrix metalloproteinases (MMP-1, -2, -3, and -9, and MT1-MMP), disintegrin metalloproteinases (ADAM-10 and ADAM-17), and inhibitors of matrix metalloproteinases (TIMP-1, -2, and -4, and ADAMTS-13). Not surprisingly CD109 is susceptible to shedding by these released proteases (Smith, Hayward et al. 1995). Consistent with this, CD109 expression levels were elevated significantly in purified platelets stimulated with A23187 or PAR1-AP in the presence of protease inhibitors and/or chelating agents (data not shown).
So what is the significance of CD109 with respect to haemostasis? Recall that at sites of vascular injury platelets first come into contact with components of the extravascular milieu. At sites of vessel wall damage, the initial adhesion of platelets occurs through GPIb-IX-V (vWF), α2β1 (collagen), and GPVI (collagen). This interaction in turn transduces signals through GPIb-IX-V that lead to GPIIb/IIIa activation, thereby contributing to stable adhesion. The transition of GPIIb/IIIa to its high affinity state is followed by granule secretion and platelet aggregation. My studies demonstrate that GPIIb/IIIa activation precedes CD109 expression, which in turn precedes complete degranulation. This observation implies that CD109 is an early activation marker. Historically, there has been discrepancy between groups as to whether CD109 is expressed on resting platelets. Given that CD109 expression can be induced by weak agonists such as epinephrine, it is likely that a subset of non-agonist treated platelets express CD109 as a result of mechanical or erythrocyte derived ADP/ATP stimulation. Physiologically, the rapid upregulation of CD109 suggests that it might have an important role in the early phases of primary hemostasis.
CHAPTER 5: SUMMARY AND FUTURE DIRECTIONS

5.1 SUMMARY

At the beginning of my studies, little was known about the function of CD109. The molecule was described as a 170kD GPI-anchored protein (Sutherland, Yeo et al. 1991). Expression of CD109 was demonstrated on endothelial cells, subsets of haematopoietic cells, and subsets of human tumour cell lines. The biallelic platelet alloantigens Gov\(^a\)/Gov\(^b\) (now HPA-15\(^b\)/HPA-15\(^a\)) were known to be carried on the peptide backbone of CD109 (Smith, Hayward et al. 1995). Gov-specific alloantibodies had been implicated in subsets of immune mediated platelet destruction syndromes with a prevalence exceeded only by that of HPA-1 (Berry, Murphy et al. 2000). However, the antigenic determinant and physiological relevance were unknown.

In this study, the evolutionary position of CD109 in the human A2M/C3 family was clarified through comparison of protein sequence and organization of the corresponding loci. This comprehensive phylogenetic analysis defines CD109 as an archaic member of the human A2M family. Through sequencing of well-characterized homozygous group O Gov donors, the Gov phenotype antigenic determinant was defined as a A2108C; Y703S polymorphism. Three genotyping methods – PCR-RFLP, PCR-SSP, and Taqman-based real time PCR were developed to accurately and precisely determine Gov genotype. Finally, the quantitative expression, activation kinetics, and subcellular localization of CD109 in platelets were investigated. CD109 is rapidly displayed following agonist-induced activation of platelets even with weak agonists such as epinephrine. In resting platelets, CD109 resides in \(\alpha\)-granules and within the OCS,
particularly clustering in regions close to the plasma membrane. The agonist-induced dose-response relationship of CD109 differs from that of PAC-1, CD62P, and CD63. This early expression of CD109 is likely reflective of its localization in the OCS and in an early-release subset of α-granules. As an early activation marker, CD109 is conveniently positioned to modulate primary haemostasis.

5.2 FUTURE STUDIES

The described studies have answered a number of questions regarding our understanding of CD109. CD109 is GPI-anchored A2M/C3 protein. This A2M/C3/C4/C5 family is conserved evolutionarily across multiple distant taxa. In humans, the A2M family consists of nine genes, five of which are well characterized in terms of their structure and function. The gene family is predicted to have arisen as the result of gene duplication and subsequent divergence of function. Prior to commencement of my studies, it was postulated that α2-macroglobulin was the most archaic member of the gene family. By this premise, complement C3 would have arisen as a consequence of a gene duplication event of an α2M-like gene and divergence in a common deuterostome ancestor, with complement C4 and C5 arising later due to subsequent gene duplication and divergence of function (Hughes 1994; Dodds and Law 1998). However, the analysis carried out in my studies has modified this evolutionary schema. Here, I posit that CD109 is the most archaic member of the human A2M family. As such, the persistence of the CD109 gene despite the expansion of the gene family, suggests an important physiological role in humans, and by extension, in other mammals. The phylogenetic tree generated in this study places CD109 on a branch with a number of diverse
organisms including: *Drosophila melanogaster, Anopholes gambiae, Ciona intestinalis, C. elegans,* and *Trichoplax adhaerans.* Given its high degree of similarity to these proteins, it is conceivable that the encoded functions might also be conserved. *C. intestinalis* GPI-A2M (CD109) is expressed in sites often associated with host defence such as circulating blood cells, the hepatopancreas and gut. Consistent with this, immune challenge of *C. intestinalis* leads to upregulation of CD109 in larvae and adults (Hammond, Nakao et al. 2005). These observations are reminiscent to the expression of CD109 on activated T cells, and its elevated expression under inflammatory conditions in humans. Analysis of the topology of the Dipteran TEPs which cluster with CD109 on our phylogenetic tree is consistent with species-specific expansion of the TEP families. The *Drosophila* genome contains six TEPs. There are five isoforms of *Drosophila* TEP2 which are derived from alternative splicing within the bait region – allowing for further diversification of function. *Anopholes gambiae* contains fifteen TEPs. Among these AgTEP15 clusters with CD109, contains an intact thioester, is GPI-anchored, and harbours the triplet VPH at the thioester reactivity determination site. AgTEP1, the most well characterized anopheles TEP, can bind and promote the opsonisation of bacteria in a manner reminiscent of complement proteins of higher vertebrates (Lagueux, Perrodou et al. 2000). Given, the simplicity of the Anopheline immune system, and this insect’s short generation time, it would be interesting to use TEP15 as a model for studying CD109 function in immunity.

The Gov alloantigen system, HPA-15, is carried on CD109 and has been implicated in a subset of immune-mediated platelet destruction syndromes. This thesis described the identification of a SNP at nt 2108 which defines each Gov allele, and the subsequent development of genotyping methods. Since the publication of our work, numerous groups
have used our system or derivatives thereof to determine the Gov genotype of various populations across the globe (Cardone, Chiba et al. 2004; Halle, Bach et al. 2004; Higgins, Hughes et al. 2004; Ertel, Al-Tawil et al. 2005; Halle, Bigot et al. 2005; Tomicic, Bingulac-Popovic et al. 2006; Moncharmont, Courvoisier et al. 2007; Matsuhashi, Tsuno et al. 2009; Bhatti, Uddin et al. 2010). In most instances, the prevalence of each phenotype tended toward a Mendelian distribution. In most – but not all – instances, the HPA-15\(^{a}\) allele was predominant.

Currently, alloantibodies directed against HPA-15\(^{b}\) are considered to be more significant to clinical disease (Ertel, Al-Tawil et al. 2005; Moncharmont, Courvoisier et al. 2007). Whether or not this is the result of biological differences or merely a function of the higher prevalence of HPA-15\(^{a}\) is unknown. Isolation and characterization of HPA-15 specific T cells from PR, PTP, or NAIT patients could serve as a means to elucidate a differential HPA-15 response. Briefly, HPA-15 allele specific T cell clones would be expanded \textit{ex vivo} and their response to antigen pulsed monocytes/m\(\phi\) or immune-compatible HPA-15\(^{bb}\) platelets subsequently assessed. Identical experiments would be performed with HPA-1 and/or -5 responder and stimulator cells. Such a study could evaluate the immunogenicity of HPA-15 relative to other HPAs. In addition, these studies could determine whether the perceived stronger immunogenicity of HPA-15\(^{b}\) is a consequence of more efficient processing and presentation of the antigen in APCs.

I have discussed in some detail the Gov alloantigen system and its implications in various platelet disorders. However, CD109 is also expressed on the surface of endothelial
cells, epithelium, and haematopoietic stem cells. As such, HPA-15 mismatches might contribute to poor outcomes in allogeneic transplant settings. This may be particularly relevant to patients who have received blood products prior to transplantation. Under such a scenario, it is highly likely that Gov homozygous patients would have been exposed/immunized to the alternative Gov isoform on multiple occasions. As such, Gov alloantibodies might contribute to tissue rejection, or alternatively support the graft versus leukaemia effect. Retrospective studies analysing Gov status with the overall outcome in terms of engraftment and disease resolution may identify any relationships.

CD109 is expressed on activated but not resting platelets. Upon stimulation of platelets, even at low doses and with weak agonists, CD109 is upregulated rapidly. The swiftness of agonist-induced CD109 expression is likely the result of its position within the OCS and/or within an early-release subset of ρ-granules of resting platelets. Essentially, in most donors, CD109 can be considered an early activation marker. A number of platelet receptors which play a crucial role in the early phases of haemostasis reside in or are enriched in lipid rafts following activation through engagement with components of the extravital milieu. As a GPI-anchored protein, which is readily upregulated, CD109 is conveniently positioned to participate in the platelet-matrix synapse. Given, the complexity of haemostasis, in vivo experiments would provide the best mechanism to test this hypothesis. For example, intravital microscopy of mice which have been genetically rendered either null or augmented at the CD109 locus globally, in platelets, in endothelial cells, or in both endothelial cells and platelets might elucidate CD109 contributions to primary haemostasis. There are two main methods of inducing thrombosis in mouse models: [1] ferric chloride induced and [2] the laser-
induced vessel damage model (Ni, Denis et al. 2000; Ni and Watts 2006). Treatment of animals with ferric chloride results in the disruption of the endothelium and exposure of the subendothelial matrix. As such, this model is dominated by collagen-induced adhesion and activation of platelets. In contrast, laser-induced vessel damage does not alter the integrity of the vascular wall and thrombus formation is dominated by TF mediated events under inflammatory conditions. The response is driven by thrombin and independent of collagen. Any observed variations between mice genetically modified at the CD109 locus and their wild type littersmates under the thrombotic models (as well as incision models of haemostasis) and unstimulated controls, may identify specific physiological contributions.

Alpha-2-macroglobulin-like molecules are involved in protease inhibition as well as in regulation of the bioavailability of cytokines and growth factors and antigen delivery (O'Connor-McCourt and Wakefield 1987; Borth and Luger 1989; Borth 1992; Kurdowska, Alden et al. 2000; Bowers, Horvath et al. 2009). Based on the primary sequence of CD109, it is believed that the CD109 thioester motif is capable of undergoing the catalytic transacylation mechanism culminating in covalent cross linking with both carbohydrate and protein targets (Lin, Sutherland et al. 2002). TGF-β1 has been demonstrated to act as a CD109 ligand in keratinocytes (Finnson, Tam et al. 2006). Additionally CD109 can interact with TβRI thereby negatively modulating TGFβ1 signalling in keratinocytes (Finnson, Tam et al. 2006). In addition to harbouring TGFβ-1, platelets also possess functional TGFβ1 receptor Type I/II as well as the downstream molecules Smad-2 and Smad-3 (Hoying, Yin et al. 1999; Lev, Salim et al. 2007). Ex vivo studies have demonstrated that TGFβ1 augments platelet aggregation in response to a number of agonists including ADP (Soslau, Morgan et al. 1997). Given that activated platelets
express CD109, TGFβ1 (Wakefield, Smith et al. 1988), TβRI/ TβRII (Lev, Salim et al. 2007), and Smad-2 (Lev, Salim et al. 2007), as well as the complete components of intersecting signalling pathways such as Ras/MAPK (Papkoff, Chen et al. 1994; Saklatvala, Rawlinson et al. 1996; Minuz, Gaino et al. 2002), it is feasible that CD109 mediates aggregation and perhaps adhesion (Saklatvala, Rawlinson et al. 1996; Minuz, Gaino et al. 2002; Lev, Salim et al. 2007). Essentially, assuming equivalent CD109 modulation of TGFβ1 signalling in keratinocytes and platelets, one would anticipate that overexpression of CD109 would result in a delayed aggregation response. By the same notion, CD109 deficiency might lead to overactive thrombosis. By extension, elevated levels of TGFβ1 levels might explain the excessive reactivity of platelets in myeloproliferative disorders. Although the levels of TβRI/TβRII have not been quantified, based on mean fluorescence intensities the stoichiometry with CD109 is close to 1:1 on platelets.

There is marked interindividual differences in the level of CD109 expression on platelets (Berry, Murphy et al. 2000). However, the physiological relevance of differential expression is unknown. Does high expression of CD109 lead to enhanced adhesion/aggregation? We could compare and contrast shear- and agonist-induced aggregation of platelets from high and low CD109 expressors. Determination of the molecular mechanism driving overexpression would be interesting from both an academic and clinic perspective. Knowledge of molecular markers will allow for the identification of overexpressors without the need to isolate platelets. Furthermore, given the emerging link between CD109 overexpression and human carcinoma (Hockla, Radisky et al. 2009), elucidation of potential neutralization targets is of great therapeutic interest.
Microparticles (0.1 – 1 µm in diameter) are released from several different cell types in response to stimuli, apoptosis, or resealing following necrosis. Their formation is associated with a shift in membrane polarity, with aminophospholipids transferred from the inner to the outer leaflet. Platelet-derived microparticles (PMP) are mainly thought of as procoagulant due to the presence of phosphatidylserine on their outer leaflet. Essentially platelet MPs promote coagulation by increasing the surface area available for assembly of coagulation factors. In contrast to exosomes, the MP surface proteome is quite heterogeneous (Perez-Pujol, Marker et al. 2007). The heterogeneity and formation of PMPs is governed in part by sequestration of proteins in LRM. CD109 is known to be expressed on platelet derived microparticles (Sutherland, Yeo et al. 1991). Future studies will serve to determine whether CD109 resides on specific subsets of platelet-derived and plasma MPs. Experiments may address whether there is a rise in CD109+ plasma MP, as well as soluble CD109, in cardio-vascular disease and cancer. Additionally analysis of CD109+ megakaryocyte and platelet derived microparticles may elucidate whether CD109 is involved in MP formation (Flaumenhaft, Dilks et al. 2009).

Recent experimental evidence has demonstrated that heterogeneous populations of α-granules exist (Ma, Perini et al. 2005; Italiano, Richardson et al. 2008; Bambace, Levis et al. 2010). It would be interesting to determine whether CD109 is sequestered in specific subsets of platelet or whether it is globally distributed, akin to P-selectin. Platelets selectively take-up, and concentrate serum proteins within their intracellular granules. Following platelet activation, the contents of these intracellular stores are released locally where they can influence inflammation, tissue remodelling, cancer, and angiogensis. Platelets, both in vitro and in vivo, have been demonstrated to segregate pro-angiogenic and anti-angiogenic factors
into distinct α-granule subsets (Italiano, Richardson et al. 2008). vWF and fibrinogen are also segregated, and co-localize with endostatin and predominantly VEGF respectively (Italiano, Richardson et al. 2008). Similarly, thrombospondin-1 and basic fibroblast growth factor (bFGF) are segregated into anti-angiogenic versus pro-angiogenic α-granule subsets. Furthermore, these α-granule subsets are released differentially in a highly regulated fashion. For example, stimulation of platelets through PAR-1 leads to the release of VEGF, a proangiogenic protein, and inhibition of endostatin release, whereas stimulation through PAR-4 exerts the opposite effect. Stimulation of platelets with ADP results in elevated VEGF release, whereas ADP-mediated stimulation (akin to PAR-1-mediated stimulation), does not influence endostatin secretion (Bambace, Levis et al. 2010). Given that ADP-stimulated platelets express CD109, I would expect to find CD109 in pro-angiogenic and not in anti-angiogenic α-granules. Immunolabeling of isolated platelets, followed by confocal microscopy could be utilized to determine whether CD109 colocalizes with VEGF or endostatin, for example. Having determined whether CD109 is sequestered into distinct or global α-granule subsets, we could then move on to investigate the potential contribution of CD109 to angiogenesis and cancer.
REFERENCES


Canobbio, I., P. Lova, et al. (2002). "Proline-rich tyrosine kinase 2 and focal adhesion kinase are involved in different phases of platelet activation by vWF." Thromb Haemost 87(3): 509-17.


