HAL
open science

Genesis of immune diversity and selection of catalytic antibodies : a new investigation

Melody Shahsavarian

To cite this version:

Melody Shahsavarian. Genesis of immune diversity and selection of catalytic antibodies : a new investigation. Biotechnology. Université de Technologie de Compiègne, 2015. English. NNT : 2015COMP2215 . tel-01258767

HAL Id: tel-01258767

https://theses.hal.science/tel-01258767v1

Submitted on 22 Jan 2016

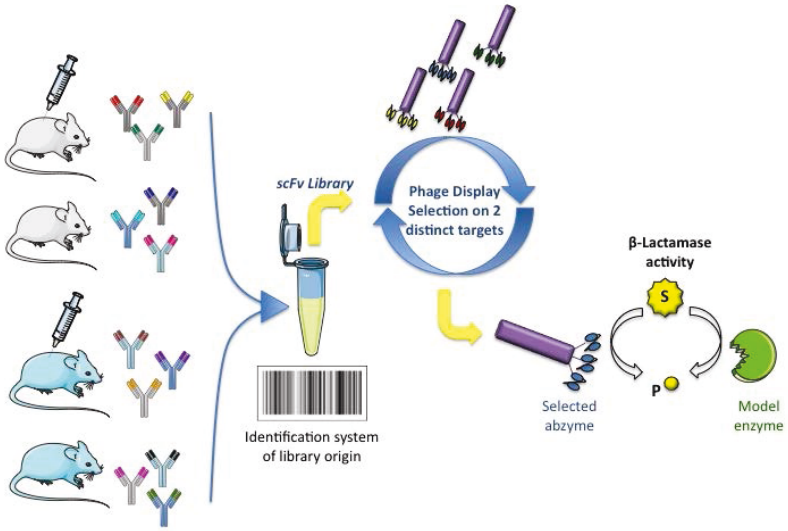
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Par **Melody SHAHSAVARIAN**

Genesis of immune diversity and selection of catalytic antibodies : a new investigation

Thèse présentée pour l'obtention du grade de Docteur de l'UTC



Soutenue le 16 octobre 2015
Spécialité : Biotechnologie

D2215

UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE

THÈSE

Présentée pour l'obtention du grade de

Docteur de l'Université de Technologie de Compiègne

Spécialité : Biotechnologie

Unité de Recherche : FRE 3580 GEC

Par

Melody SHAHSAVARIAN

Genèse de la diversité immune et la sélection des anticorps catalytiques : une nouvelle investigation

Soutenue le 16 octobre 2015

Devant le jury composé de :

- M. GABIBOV Alexander	Professeur, Russian Academy of Sciences, Moscou
- M. MATAGNE André	Professeur, Center for Protein Engineering, Liège
- M. DUMAS Bruno	Senior Scientist, SANOFI, Vitry sur Seine
- M. FRIBOULET Alain	Directeur de recherche, CNRS/FRE 3580, UTC
- Mme AVALLE Bérangère	Professeur des universités, CNRS/FRE 3580, UTC
- Mme PADIOLLEAU-LEFEVRE Séverine	Maître de conférences, CNRS/FRE 3580, UTC

Abstract

Catalytic antibodies (or abzymes) have been the focus of numerous studies for some decades and have been produced with the ability to catalyze a wide range of reactions. They have also been discovered naturally in normal physiological and pathological conditions, notably on the background of autoimmune disease. Some have beneficial effects and others are detrimental to individuals' health. Hence, the origin of abzymes and their role in the immune response are ambiguous and must be enhanced.

We have developed a novel strategy for the study of abzymes based on the phage display technology. We have constructed 4 libraries representing 4 murine immune repertoires with different genetic backgrounds and immunological states: *healthy and naive*, *healthy and immunized*, *autoimmune and naive*, and *autoimmune and immunized*. The strategies for the amplification and cloning of the immunoglobulin (Ig) variable regions have been designed to optimize the size and diversity of the libraries. We have been able to pool the four libraries to create a large repertoire of size 2.7×10^9. After sequence analysis, we have found a number of statistically significant differences between the libraries. We have then used two strategically chosen targets to select for antibodies endowed with β lactamase activity: a cyclic peptide and a penam sulfone, both inhibitors of the enzyme. We have selected for a total of 5 Igs endowed with β lactamase activity. The selected abzymes have different amino acid sequences. 3D modeling has provided insights on potential active sites demonstrating the ability of different structures to maintain the β lactamase activity and confirming the flexibility of the active site.

Keywords: Autoimmunity, Abzymes, β-lactamase, Catalytic antibodies, Immune repertoires, Molecular diversity, Phage display, single chain Fragment variable.

Résumé

Les anticorps catalytiques (ou abzyme) ont fait l'objet de nombreuses recherches et ont été produits pour réaliser de nombreuses réactions. Ces protéines ont été ensuite découvertes dans le sérum d'individus sains ou atteints de pathologies, dont les pathologies autoimmunes. Les études suggèrent que ces abzymes peuvent avoir des effets bénéfiques ou délétères sur la santé des individus. L'origine des anticorps catalytiques et leur rôle restent ambigus et doivent être approfondis. Nous avons développé une nouvelle stratégie visant à étudier les abzymes, basée sur la technologie du phage display.

Nous avons construit 4 banques de fragments d'anticorps, chacune présentant un répertoire immun différent (fond génétique et état d'immunitaire): *saine et naïve, saine et immunisée, autoimmune et naïve*, et *autoimmune et immunisée*. Les stratégies d'amplification et de clonage des régions variables des immunoglobulines ont été conçues afin d'optimiser la taille et la diversité des banques. Nous avons rassemblé les 4 banques en une banque unique élargie contenant 2.7×10^9 séquences. L'analyse des séquences a mis en évidence des différences dans les profils d'expression des sous-groupes de gènes selon la banque. Nous avons ensuite procédé à la sélection d'abzymes à activité β-lactamase en utilisant deux cibles : un peptide cyclique, et un dérivé de sulfone pénam, inhibiteurs de l'enzyme. Nous avons sélectionné 5 abzymes. Chacun de ces immunoglobulines ont des séquences protéiques propres, incluant un potentiel site actif. Ces résultats montrent que différents motifs peuvent assurer la fonction catalytique de la β-lactamase, confirmant la flexibilité moléculaire de cette enzyme.

Mots clés: Abzymes, Anticorps catalytiques, Autoimmunité, β-lactamase, Diversité moleculaire, Repertoire immuns, Phage display, single chain Fragment variable.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my PhD supervisor, my mentor, my role model, and my friend, Dr. Séverine Padiolleau-Lefèvre, without whom I would not have been able to accomplish this work and would not have gotten to where I am today. I will eternally be grateful for your endless support, in professional as well as personal matters, your scientific direction, your cheerful sense of humor, and your attitude toward life in general. I would also like to thank my co-supervisor, Pr. Bérangère Avalle, who was also an essence not only for the completion of this work, but also for my survival during these three years. These two lovely women are probably the most kind, humane, warm, supportive, and dynamic duos that I will meet in my entire life. Words cannot describe to which extent I appreciate you two.

I would like to thank Dr. Alain Friboulet for his extensive scientific input into this work and throughout the three years of my PhD. I would also like to thank Dr. Didier Boquet for his scientific support throughout this work and specially for sharing his knowledge on the IMGT database and the design of immunoglobulin-specific primers. Without his help this work would have never gone on the right track. I would like to thank Dr. Véronique Giudicelli for her help, time and time again, on the correct use of the IMGT tools. I want to thank Pr. Bernard Offmann for his valuable contribution to the structural modeling part of this work. I also thank my jury members, Dr. Bruno Dumas, Pr. Alexandre Gabibov, and Pr. André Matagne for their time and input in reviewing this work. I would also like to thank the late Dr. Daniel Thomas for his warm and inviting spirit and his heart-felt enthusiasm for science.

I would like to thank my team members Dr. Claire Loussouarn, soon to be Dr. Hassan Isber, and Dr. Raouia Ben Naya for their friendship and the memorable moments we shared together during these years.

I would like to thank my family, mom, dad, and my little sister for their unconditional love and support no matter what situation life would put in my way. My life would not be the same without these three figures by my side.

Last but not least, I would like to thank the love of my life and my soon to be husband, Dr. Paolo Bonomi, who has been by my side for the majority of this journey, supporting me just as much during the difficult times as the happy times. I love you, I thank you for being who you are and am incredibly excited for the future that stands in front of us.

Table of Contents

GENERAL INTRODUCTION **7**

I. MOLECULAR DIVERSITY AND THE IMMUNE REPERTOIRE **9**

I.1. The Immune Repertoire, a natural source of diversity **9**
I.1.A. Innate Immunity 9
I.1.B. Adaptive Immunity 11
I.1.C. Immunoglobulins 16
I.1.D. Defects of the immune system 27

I.2. Display Technologies, a synthetic source of diversity **29**
I.2.A. Phage Display 29

I.3. Immune Repertoire Studies **33**

I.4. Generation of diversity: Construction of antibody libraries representing four immune repertoires **34**
I.4.A. Choice of immune repertoires 35
I.4.B. Choice of the antigen 36
I.4.C. Library construction 40

I.5. Analysis of immune repertoires **54**

II. CATALYTIC ANTIBODIES WITH β-LACTAMASE ACTIVITY **62**

II.1. Catalytic Antibodies **62**
II.1.A. Strategies for the production of catalytic antibodies 63
II.1.B. Naturally occurring catalytic antibodies 71
II.1.C. Genetic study manifesting certain characteristics of abzymes 76
II.1.D. Applications of catalytic antibodies 80

II.2. β-lactamase: a model catalytic activity **83**
II.2.A. The enzyme 83
II.2.B. The catalytic antibody 9G4H9 89
II.2.C. Pep90: a cyclic peptide inhibitor of 9G4H9 94

II.3. Selection of target antibodies by phage display **96**
II.3.A. Selection procedure 96
II.3.B. Applications 97

II.4. Selection of catalytic antibodies with β-lactamase activity **98**
II.4.A. Selection against Pep90 99
II.4.B. Selection against penam sulfone 104

II.5. Study of catalytic activity **111**
II.5.A. Colorimetric assay 111
II.5.B. Fluorometric assay 113

II.6. Structural modeling **118**

III. DISCUSSION AND PERSPECTIVES **124**
III.1. Library Construction **124**
III.2. Characterization of the library **127**
III.3. Selection of catalytic antibodies **128**
III.4. Structural analysis **131**

ADDENDUM **133**

REFERENCES **134**

ANNEXES ..... **i**

LIST OF FIGURES

Figure 1. *T cells derived from naive CD4+ T lymphocytes.* 12
Figure 2. *Activation of a CD8+ T lymphocyte by a CD4+ helper T cell.* 13
Figure 3. *Activation of a B lymphocyte by a helper T lymphocyte.* 14
Figure 4. *The different pathways of immune response resulting from T_H1 or T_H2 lymphocytes.* 15
Figure 5. *Structure of IgG.* 18
Figure 6. *Antibody variable regions.* 19
Figure 7. *Antibody fragments with different formats. Orange and green segments represent light chains, wheras blue segments represent heavy chains.* 20
Figure 8. *Germ-line organization of human immunoglobulin (Ig) loci.* 21
Figure 9. *Immunoglobulin VDJ gene rearrangement producing complete heavy and light chains.* 22
Figure 10. *The mechanistic pathway of the addition of P- and N- nucleotides showing the different enzymes involved.* 23
Figure 11. *Mechanisms involved in class-switch recombination.* 24
Figure 12. *Mechanisms involved in somatic hypermutation.* 26
Figure 13. *Structure of filamentous phage.* 30
Figure 14. *The possible phenotypes of phagemid particles.* 31
Figure 15. *Four mouse models representing four immune repertoires of different genetic backgrounds and immunological states.* 35
Figure 16. *Penicillin sulfone derivative connected to a biotin by a spacer containing a disulfide bond.* 36
Figure 17. *Strategy for the synthesis of the penam sulfone derivative.* 37
Figure 18. *Inhibition of the enzyme β-lactamase by the penam sulfone derivative.* 38
Figure 19. *ELISA test confirming the induction of an immune response of both strains against KLH.* 39
Figure 20. *Construction of scFv library by phage display.* 40
Figure 21. *Two-step PCR strategy for the amplification of V_L and V_H.* 41
Figure 22. *Testing of RNA integrity by Experion™ RNA StdSens Analysis Kit (Biorad).* 43
Figure 23. *Analysis of the amplified variable regions V_L and V_H on 1.3% agarose gel after PCR2.* 44
Figure 24. *SOE-PCR allowing the assembly of V_L and V_H into scFv.* 45
Figure 25. *Assemby of V_L and V_H by SOE-PCR.* 45
Figure 26. *Rolling Circle Amplification of scFv fragments.* 47
Figure 27. *Sequence of the restriction site of the enzyme SfiI.* 47
Figure 28. *RCA results after SfiI digestion.* 48
Figure 29. *Schematic representation of phagemid pAK100.* 48
Figure 30. *Phagemid vectors for library construction.* 50
Figure 31. *Experimental restriction digestion profiles of phagemid vectors and library clones.* 51
Figure 32. *Western Blot analysis of the library.* 53
Figure 33. *PCR on colony on 24 randomly selected individual clones for each of the 4 libraries.* 54
Figure 34. *Frequencies of the light chain κ subgroups reported in IMGT.* 55
Figure 35. *Frequencies of the heavy chain γ subgroups reported in IMGT.* 56
Figure 36. *Frequencies of the light chain λ subgroups reported in IMGT in July, 2015.* 57
Figure 37. *Analysis of the amplified V_L λ of immunized Balb/C mice on 1.3% agarose gel after PCR2.* 58
Figure 38. *Analysis of PCR on colony results on 1.3% agarose gel of 10 pGEMT clones containing λ sequences using the PCR2 primer mix.* 58
Figure 39. *Immunoglobulin gene subgroup distribution of Kappa light chain.* 59
Figure 40. *Immunoglobulin gene subgroup distribution of Gamma heavy chain.* 60
Figure 41. *Diagram representing energy states along a reaction pathway with or without the presence of a catalyst.* 64
Figure 42. *Hapten design using the structure of penam sulfone (a suicide inhibitor of the enzyme β-lactamase).* 66
Figure 43. *Comparison of active site cavities of natural and de novo created biocatalysts.* 67
Figure 44. *Schematic of the anti-idiotypic network.* 69
Figure 45. *Dual role of catalytic antibodies in physiopathology.* 74
Figure 46. *The frequency of IGKV gene subgroups referenced in IMGT.* 78

Figure 47. *Comparison between the expression frequencies of IGKV rare gene subgroups between binding and catalytic antibodies.* 79
Figure 48. *Structure of some β-lactam antibiotics.* 83
Figure 49. *The 3D structure of the class A β-lactamase from E coli TEM-1 (pdb 1BTL).* 84
Figure 50. *Two-dimensional schematic representation of active-site residues, the conserved hydrolytic water and the penicillin substrate.* 85
Figure 51. *General catalytic pathway of serine β-lactamases.* 86
Figure 52. *The two debated pathways for the acylation step of the β-lactamase reaction.* 87
Figure 53. *Stereoview of the active site of TEM-1 β-lactamase.* 88
Figure 54. *Amino acid sequence of antibody 9G4H9 variable regions.* 90
Figure 55. *Structural comparison of the active sites of κ-light chain of 9G4H9, β-lactamase, and DD peptidase by molecular modeling.* 91
Figure 56. *3D modeling of 9G4H9.* 92
Figure 57. *Chemical structures of β-lactam substrates.* 93
Figure 58. *Structural modelization of Pep90.* 95
Figure 59. *Selection cycle of phage display library against a specific target.* 96
Figure 60. *Scheme of the support structure for the selection procedure.* 99
Figure 61. *Calibration curve of Pep90 for the determination of the extinction coefficient performed in PBS (pH 7.4).* 99
Figure 62. *Enrichment of phage specific to Pep90.* 101
Figure 63. *Phage-ELISA assay of the Pep90 selected phage.* 102
Figure 64. *The 22 positive clones selected against Pep90.* 102
Figure 65. *The alignment of Pep90 selected clones by the Multalin software.* 103
Figure 66. *Restriction digestion profiles of the Pep90 selected clones.* 104
Figure 67. *Calibration curve of penam sulfone for the determination of the extinction coefficient in PBS (pH 7.4).* 105
Figure 68. *Enrichment of phage specific to penam sulfone.* 106
Figure 69. *Phage-ELISA assay of the penam sulfone selected phage.* 107
Figure 70. *The 17 positive clones selected against penam sulfone.* 107
Figure 71. *The alignment of penam sulfone selected clones by the Multalin software.* 108
Figure 72. *Restriction digestion profiles of the penam sulfone selected clones.* 109
Figure 73. *Phage-ELISA assay on Pep90 selected clones using penam sulfone as target.* 110
Figure 74. *Phage-ELISA assay on penam sulfone selected clones using Pep90 as target.* 110
Figure 75. *The structure of nitrocefin and its product after hydrolysis of the β-lactam ring.* 112
Figure 76. *Colorimetric assay for the determination of β-lactamase activity of the five selected clones using nitrocefin as substrate.* 113
Figure 77. *Hydrolysis of Fluorocillin™.* 113
Figure 78. *Fluorocillin™ hydrolysis by the different controls over 200 minutes.* 114
Figure 79. *Fluorocillin hydrolysis by the five selected clones and the control scFv 9G4H9 over 200 minutes.* 116
Figure 80. *Alignment of the five selected clones and 9G4H9.* 117
Figure 81. *3D model of antibody P90C1 with superposition of β-lactamase active residues.* 119
Figure 82. *3D model of antibody P90C2 with superposition of β-lactamase active residues.* 120
Figure 83. *3D model of antibody P90C3 with superposition of β-lactamase active residues.* 120
Figure 84. *3D model of antibody PSC1 with superposition of β-lactamase active residues.* 121
Figure 85. *Superposition of antibodies PSC1 and 9G4H9.* 122
Figure 86. *3D model of antibody PSC2 with superposition of β-lactamase active residues.* 122
Figure 87. *Pep90 and related protein catalysts.* 131

LIST OF TABLES

Table 1. *Myeloid cells in innate and adaptive immunity.* _____ *10*
Table 2. *Immonoglubulin isotypes with their different characteristics and functions.* _____ *17*
Table 3. *Some common autoimmune diseases* _____ *28*
Table 4. *Primers for RT-PCR.* _____ *41*
Table 5. *Size of each of the 4 libraries.* _____ *52*
Table 6. *Catalytic antibodies reported in different pathological diseases.* _____ *71*
Table 7. *Therapeutic applications of catalytic antibodies.* _____ *81*
Table 8. *Kinetic parameters of different catalyst with β-lactamase activity.* _____ *94*
Table 9. *Saturation experiment of the magnetic beads with Pep90.* _____ *100*
Table 10. *Pep90-selected clones with their library origin and gene subgroup identification.* _____ *103*
Table 11. *Saturation experiment of the magnetic beads with penam sulfone.* _____ *105*
Table 12. *Penam sulfone-selected clones with their library origin and gene subgroup identifiation.* _____ *108*
Table 13. *Characteristics of the five selected antibodies and 9G4H9.* _____ *117*

LIST OF ABBREVIATIONS

Ab: Antibody
AβP: Amyloid-β-Protein
ABTS: 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid
AD: Alzheimer's Disease
ADAPT: Antibody Directed Abzyme Prodrug Therapy
ADCC: Antibody Dependent Cell Cytotoxicity
ADCP: Antibody Dependent Cell Phagocytosis
ADEPT: Antibody Directed Enzyme Prodrug Therapy
AID: Activation Induced cytidine Deaminase
APC: Antigen Presenting Cell
APE: Apurinic/Apyrimidinic Endonuclease
BCR: B Cell Receptor
BSA: Bovine Serum Albumin
BZ: Benzylpenicillin
CD: Cluster of Differentiation
CDC: Complement Dependent Cytotoxicity
cDNA: complementary Deoxyribo Nucleic Acid
CDR: Complementarity Determining Region
CTX: Cefotaxime
D: Diversity
DMSO: DiMethyl Sulfoxide
DNA: Deoxyribo Nucleic Acid
ds: double stranded
DTT: DiThioThreitol
EAE: Experimental Allergic Encephalomyelitis
EBV: Epstein Bar Virus
ELISA: Enzyme Linked Immuno Sorbent Assay
ESBL: Extended Spectrum Beta Lactamase
Fab: Fragment antibody
Fc: Fragment crystallizable
Ff: Filamentous phage
FR: Framework
gp: glycoprotein
GSP: Gene Specific Primer
HIV: Human Immunodeficiency Virus
HRP: Horseradish peroxide
HTS: Hight Throughput Sequencing
IFN: Interferon
Ig: Immunoglobulin
IL: Interleukin
IMGT: IMmunoGeneTics
J: Joining
KLH: Keyhole Limpet Hemocyanin
MBP: Myelin Binding Protein
MHC: Major Histocompatibility Complex
MS: Multiple Sclerosis
MSH: MutS Protein Homolog
NGS: New Generation Sequencing
NK: Natural Killer
OD: Optical Density
PAR: Protein Activated Receptor
PBP: Penicillin Binding Protein
PBS: Phosphate Buffered Saline
PCR: Polymerase Chain Reaction
PDB: Protein Database
PRR: Pathogen Recognition Receptor
PpL: Peptostreptococcus protein L
PS: Packaging Signal
QM/MM: Quantum Mechanical/Molecular Mechanical
RAG: Recombination Activating Gene
RCA: Rolling Circle Amplification
RF: Replicative Form
RNA: Ribonuclease Amino Acid
RSS: Recombination Signal Sequence
RT-PCR: Reverse Transcription Polymerase Chain Reaction
sc: single chain
scFv: single chain Fragment variable
SLE: Systemic Lupus Erythematosis
SOE-PCR: Single Overlap Extension Polymerase Chain Reaction
ss: single stranded
TdT: Terminal deoxynucleotidyl Transferase
Tg: Thyroglobulin
TI: Thymus Independent
TSA: Transition State Analog
UNG: Uracil-DNA Glycosylase
V: Variable
VH: Variable Heavy
VIP: Vasoactive Intestinal peptide
VL: Variable Light
WT: Wild Type

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Catalytic antibodies (or abzymes) have, in recent decades, been the subject of study of many scientists in search of new biocatalysts for a variety of applications. The first two catalytic antibodies were reported in 1986 (Tramontano *et al.*, 1986; Pollack *et al.*, 1986), and since, many more have been produced in the laboratory. These antibodies have proven to be able to catalyze a wide range of reactions, some of which do not exist in nature (Blackburn and Garçon, 2008). In parallel to the advances in the chemical production of abzymes, these proteins have also been discovered naturally both in normal physiology and pathology, notably on the background of autoimmune disease (Wootla *et al.*, 2011). Studies have shown that some of these abzymes have beneficial effects for the patients' health whereas others are detrimental. Therefore, despite the advances in the field, the origin and nature of catalytic antibodies as well as the role they play in the immune response are yet to be understood.

In recent years, developments in protein engineering and molecular biology techniques have provided researchers with the possibility to produce, *in vitro*, molecular diversities close to natural immune repertoires. The phage display technology, for example, has led to the construction of antibody libraries mimicking natural immune repertoires and the selection of antibodies against a large range of molecular entities (Bradbury and Marks, 2004).

In the present work, we describe a novel strategy for the study of catalytic antibodies, in an attempt to reveal some answers to the many questions still remaining on the characteristics of these fascinating proteins. In this strategy, we want to utilize the powerful technique of phage display to construct antibody libraries originating from four murine models with different genetic backgrounds and different immunological states: *(i)* healthy and naive[1], *(ii)* healthy and immunized, *(iii)* autoimmune-prone and naive, and *(iv)* autoimmune-prone and immunized. As such, we attempt to increase the molecular diversity represented by our repertoire of immunoglobulin variable regions, leading to an improved probability of selecting for antibodies with any desired characteristics.

We then want to use this large and diverse immune repertoire in order to select for antibodies endowed with a model catalytic activity, namely β-lactamase. We, therefore, want to use two strategically chosen inhibitors of the enzyme β-lactamase in order to select for the desired antibodies. As a first point of analysis, we want to explore whether the genetic background or the immunization state of the murine immune repertoire influences the tendency of the expression of catalytic antibodies. Results of such analysis will potentially reveal some information on the previously hypothesized link between an autoimmune background and an increased expression level of abzymes (Wootla *et al.*, 2011b). As a second point of analysis, we want to study the genetic characteristics of the potentially selected catalytic antibodies, such as immunoglobulin gene subgroup usage, germline percentage of identity, etc. This analysis will potentially illustrate whether there

[1] In the context of the current study, what we consider naive is the non-immunized IgG repertoire.

are fundamental differences between these characteristics and those held by binding antibodies.

In addition to the genetic analysis of catalytic antibodies, our disposition of four antibody libraries based on different murine models will allow us to perform a statistical study on the potential underlying genetic differences between these B cell repertoires. Furthermore, the abzymes endowed with β-lactamase activity selected from the repertoires, will potentially provide us with a number of different structural models capable of performing such activity. We, therefore, want to perform comparative analyses of the selected abzymes' catalytic sites with that of the natural enzyme β-lactamase, in order to provide more information on the previously discussed structural plasticity of this catalyst (De Wals *et al.*, 2009).

Together, we hope that our results will help us better understand the origin of catalytic antibodies and their potential link to autoimmunity. Additionally, we pursue to discover whether an autoimmune background or an immunization event lead to fundamental changes in the B cell repertoire diversity of an organism. Finally, we hope to expand our knowledge on the potential structural possibilities capable of holding a β-lactamase catalytic function.

I. Molecular Diversity and the Immune Repertoire

I. MOLECULAR DIVERSITY AND THE IMMUNE REPERTOIRE

This chapter focuses on the theme of molecular diversity, a recurring concept in biology and chemistry, and one whose importance cannot be denied for forming theories about the universe as well as for developing applications for the benefit of mankind. A remarkable example of molecular diversity created by nature is the immune repertoire, an almost unlimited reservoir of molecules protecting us from infectious agents, without which no superior living organism would survive. Of course, as for any other example of nature, the immune repertoire is an inspiration for scientists to attempt to recreate a similar source of diversity in the laboratory. The first part of this chapter is dedicated to this outstanding creation of nature. The second part presents a description of current technologies attempting to approximate this vast diversity and finally our manipulation of such technologies for the ultimate objective of producing and studying catalytic antibodies.

I.1. The Immune Repertoire, a natural source of diversity

The immune system is one of evolution's most fascinating and remarkable examples of the means of survival in vertebrates. Through the immense molecular diversity of the immune repertoire, this system is capable of recognizing and defending the organism against any number of foreign invaders. The complex machinery of the immune system is able to distinguish between self and non-self entities. This distinction accounts not only for disease-causing pathogens including viruses, bacteria, fungi, parasites, and non-conventional transmissible agents such as prions (Dormont, 1998), but also for a wide range of chemical structures such as metals or any chemical drugs. Upon trespassing of a non-self molecule, not only does the immune system launch a series of extremely efficient and highly specific events in order to identify and eliminate this intruder, but it is also capable of keeping an immunological memory of the trespasser's identity and the occurrences leading to its destruction. Hence, should the trespasser reappear, the immune system would be able to re-launch its response with a higher speed and intensity.

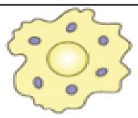
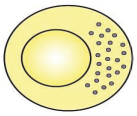
The immune system achieves its defensive objective through four elaborate tasks: *(i)* immunological recognition, *(ii)* immune effector functions, *(iii)* immune regulation, and *(iv)* immunological memory. These tasks are performed by the interplay and regulation between two distinct but closely related arms of the immune system, namely the innate and adaptive immunity (Dempsey *et al.*, 2003). The mechanisms involved in these two components of the immune response are well summarized in publications of Kindt *et al.* and Janeway *et al.* (Janeway *et al.*, 2001; Kindt *et al.*, 2008; Murphy, 2012).

I.1.A. Innate Immunity

The innate immune response occurs rapidly but non-specifically (within a few hours) upon exposure to a foreign molecule. This first line of defense involves processes such as

phagocytosis, inflammation and activation of the complement system for the elimination of a microbe. It engages a group of plasma proteins, known as pathogen recognition receptors (PRRs), which induce various inflammatory cytokines, chemokines, and type I interferons and cover the infectious microbe by a cascade of proteolytic reactions (Kumar *et al.*, 2011). Most of the leukocytes responsible for the innate immune response, with one exception, derive from the myeloid lineage, which in turn derives from pluripotent hematopoietic stem cells originating in the bone marrow. These cells include macrophages, granulocytes including neutrophils, eosinophils and basophils, mast cells, and dendritic cells (Table 1). The dentritic cells and the macrophages provide a link between the innate and adaptive immune response by acting as antigen presenting cells (APCs) to T lymphocytes (Lipscomb and Masten, 2002) (described in I.1.B.). Natural killer cells (NKs) are the one exception of the leukocytes of the innate immune response, which derive from the lymphoid lineage (Vivier *et al.*, 2008).

Table 1. Myeloid cells in innate and adaptive immunity.

Cell type	Characteristics	Location	Image
Mast cell	Dilates blood vessels and induces inflammation through release of histamines and heparin. Recruits macrophages and neutrophils. Involved in wound healing and defense against pathogens but can also be responsible for allergic reactions.	Connective tissues, mucous membranes	
Macrophage	Phagocytic cell that consumes foreign pathogens and cancer cells. Stimulates response of other immune cells.	Migrates from blood vessels into tissues.	
Natural killer cell	Kills tumor cells and virus-infected cells.	Circulates in blood and migrates into tissues.	
Dendritic cell	Presents antigens on its surface, thereby triggering adaptive immunity.	Present in epithelial tissue, including skin, lung and tissues of the digestive tract. Migrates to lymph nodes upon activation.	
Monocyte	Differentiates into macrophages and dendritic cells in response to inflammation.	Stored in spleen, moves through blood vessels to infected tissues.	
Neutrophil	First responders at the site of infection or trauma, this abundant phagocytic cell represents 50-60 percent of all leukocytes. Releases toxins that kill or inhibit bacteria and fungi and recruits other immune cells to the site of infection.	Migrates from blood vessels into tissues.	
Basophil	Responsible for defense against parasites. Releases histamines that cause inflammation and may be responsible for allergic reactions.	Circulates in blood and migrates to tissues.	
Eosinophil	Releases toxins that kill bacteria and parasites but also causes tissue damage.	Circulates in blood and migrates to tissues.	

Source: inspired by Murphy, 2012

I.1.B. Adaptive Immunity

When an antigen escapes the innate immune response, the adaptive immune response reacts with a longer delay (upto a few days), but with more efficiency and higher specificity. Antigen specific lymphocytes are the key components of the adaptive immune response. They recognize a quasi infinite number of antigenic species with a remarkably high specificity. This capability is due to the large diversity of their highly variable antigen receptors. These lymphocytes, including B and T cells, derive from the lymphoid lineage of the pluripotent hematopoietic stem cells originating in the bone marrow (Murphy, 2012).

B lymphocytes also mature in the bone marrow, whereas T lymphocytes migrate to the thymus for their maturation. After maturation, naive lymphocytes circulate in the blood stream until they are met with an antigen or an APC, particularly macrophages and dendritic cells, at the peripheral lymphoid tissues. At this point, the lymphocytes begin to proliferate and differentiate into clones of identical progeny, a process known as clonal expansion (Von Boehmer, 1994).

There are a number of different pathways, which the adaptive immune system can take to develop its response. These include activation of the cells of the innate immune system and B and T lymphocytes of the adaptive immune system, as well as the regulation of the immune response. The pathway to be taken depends on many different factors, notably the nature of the pathogen, and is orchestrated by the CD4+ T lymphocytes. The resulting effector functions depend on the differentiation of different progeny of CD4+ T lymphocytes, namely T_H1, T_H2, T_H17, follicular helper T cells T_{FH}, and regulatory T cells T_{reg} (Swain *et al.*, 2012) (Figure 1).

These effector T lymphocytes are activated by APCs expressing Major Histocompatibility Complex class II (MHC II) receptors and promote different types of responses targeted to different sources of infection. They function by releasing different signaling molecules known as cytokines (Luckheeram *et al.*, 2012). T_H1 and T_H2 cells are the most documented of the effector T lymphocytes and have been thought to be responsible for the two major tasks of the adaptive immune response: activation of the cytotoxic CD8+ T lymphocytes (T_H1) and the B lymphocytes (T_H2).

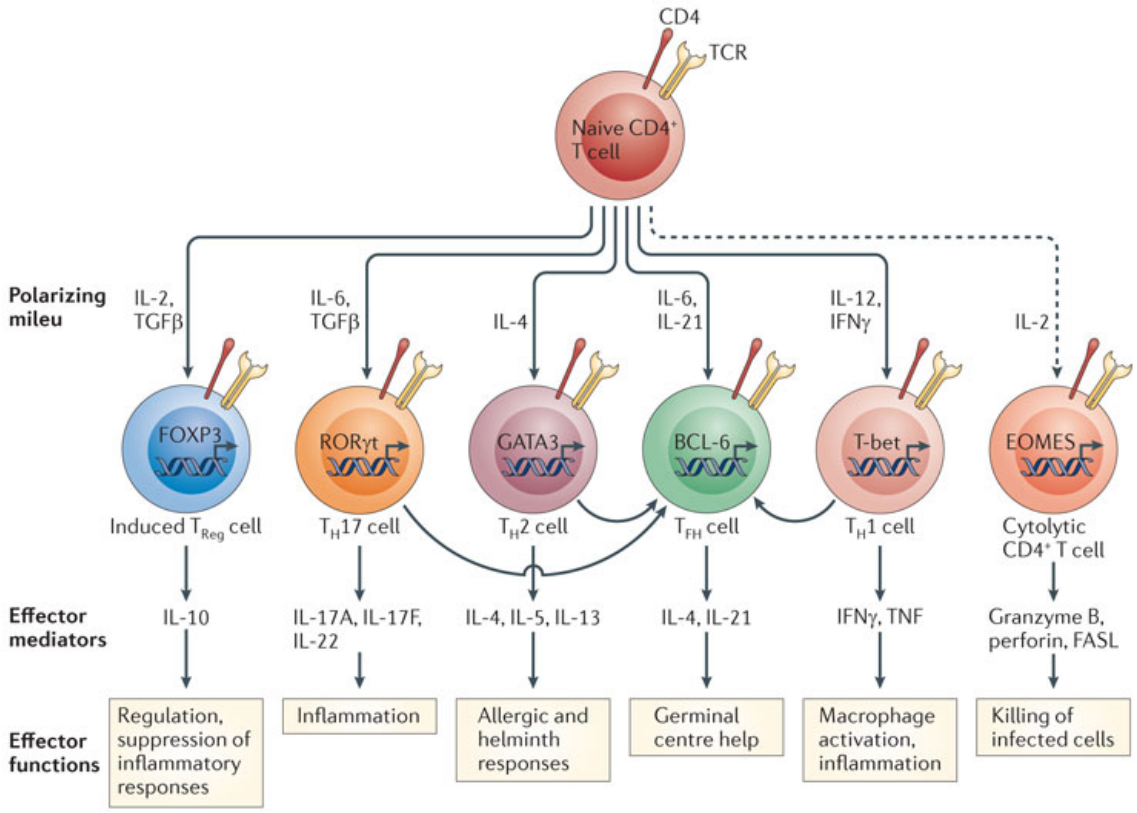


Figure 1. T cells derived from naive CD4+ T lymphocytes.

The different cytokines driving the differentiation of different types of CD4+ T cells and the cytokines produced by the mature cells leading to their different effector functions are shown. Cytolytic CD4+ T cell has been recently proposed to derive from an additional T cell differentiation pathway. Source: Swain *et al.*, 2012

The secretion of interleukin 12 (IL12) and interferon γ (IFNγ) orient the immune response toward the T_H1 pathway, where two signals are required: *(i)* the presentation of viral proteins to the CD8+ T cells by Major Histocompatibility Complex class I (MHCI) molecules expressed on the surface of infected cells, and *(ii)* the secretion of a number of cytokines, notably, IL2 and IFNγ, by activated T_H1 cells. This activation is due to the previous presentation of the antigenic determinants of the pathogen to T_H1 cells via MHCII molecules expressed on the surface of APCs. Both signals lead to the activation and proliferation of the cytotoxic CD8+ T lymphocytes (Figure 2).

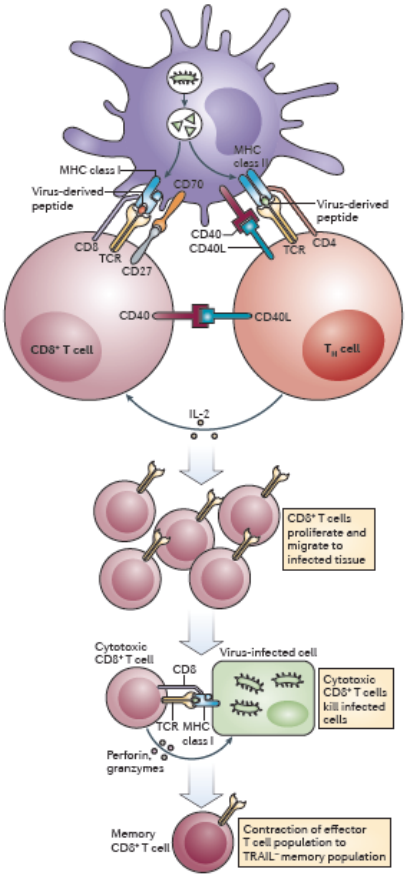


Figure 2. Activation of a CD8+ T lymphocyte by a CD4+ helper T cell.

In the particular case of this figure, the infected cell is presented by an APC. Source: Swain *et al.*, 2012

It is worthy of mention, however, that more recent studies have shown that in the presence of a strong antigenic signal, the CD8+ T lymphocytes are capable of activation directly by MHCII antigen presentation of infected cells, independent of CD4+ helper T cells (Bevan, 2004). The help of CD4+ T lymphocytes is, however, necessary for the differentiation of the CD8+ cells into memory cells. In both cases, the activation of CD8+ T lymphocytes results in the local secretion of effector molecules such as granzymes and perforins, which ultimately lead to the lysis of the infected cell. This activation can also lead to the secretion of cytokines such as TNFα or β, as well as to caspase-dependent apoptosis due to the presence of Fas receptor on the surface of the CD8+ T lymphocyte (Suzuki and Fink, 2000).

In addition to their role in the activation of CD8+ T lymphocytes, T_H1 cells are involved in activating macrophages to destroy bacteria present in membrane-enclosed vesicles and are associated with organ-specific autoimmunity.

The secretion of IL4 and IL10 by the APC, on the other hand, orients the immune response toward the T_H2 pathway, in which also two signals are required. Historically, it has been

thought that the action of this effector T lymphocyte is the main pathway for the activation of B cells. The first signal, required for this activation, is the recognition of an antigen by a B cell receptor (BCR), followed by its internalization, digestion, and presentation on the surface of the B lymphocyte by MHCII molecules. The second signal is produced by the interaction of T_H2 cells with the antigenic determinants presented on the B cell surface by MHCII, following the previous internalization and digestion. The interaction of the CD40 receptor present on the B cell surface and its corresponding ligand is also involved in the activation of the B lymphocyte (Wykes, 2003). These signals lead to the secretion of specific cytokines, notably IL4, 5, 6, and 10, ultimately leading to the differentiation of B lymphocytes to immunoglobulin-secreting plasmocytes (Figure 3).

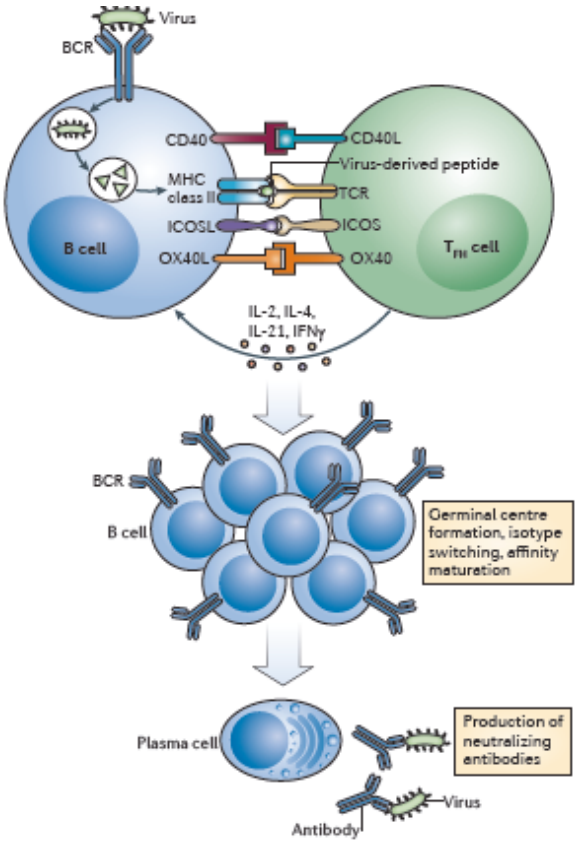


Figure 3. Activation of a B lymphocyte by a helper T lymphocyte.

The helper T cell is presented by a T_{FH} cell rather than a T_H2 cell (described below). Source: Swain *et al.*, 2012

More recently, however, it has been demonstrated that the adoptive transfer of either T_H1 or T_H2 cells provide efficient help for the generation of neutralizing IgG responses (Maloy *et al.*, 2000). The functions of these cells are interconnected and exert a level of control on one another (Figure 4).

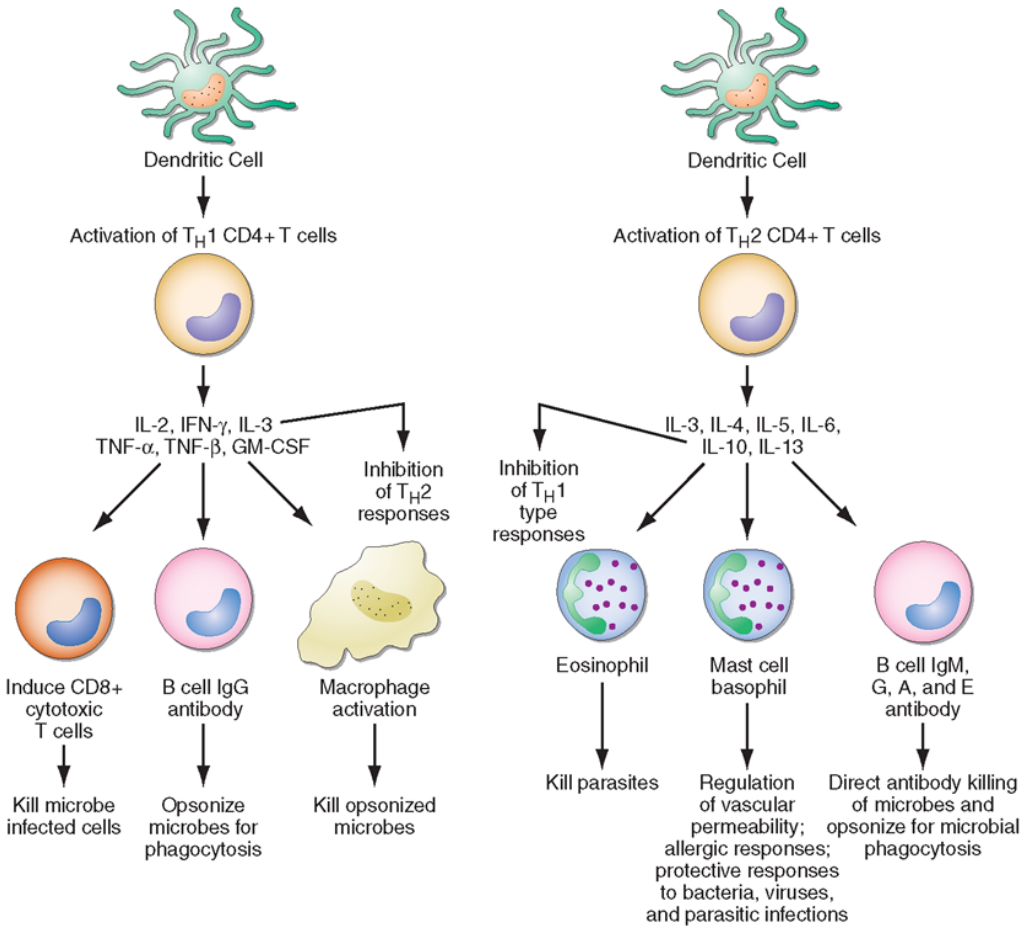


Figure 4. The different pathways of immune response resulting from T_H1 or T_H2 lymphocytes.

Source : what-when-how.com/rheumatology/introduction-to-the-immune-system-the-immune-system-in-health-and-disease-rheumatology-part-3

The role of B cell activation is now given mainly to a different subset of effector T lymphocytes, namely the T_{FH} cells (Hatzi *et al.*, 2015). These lymphocytes enter B cell follicles and provide help to B cells, resulting in germinal center formation and plasmocyte differentiation (Figure 3). In turn, T_H2 cells are now thought to be involved in immune responses to extracellular parasitic infections in mucosal surfaces and to play a major role in the induction and persistence of asthma as well as other allergic diseases.

It is also worthy of note, however, that some microbial constituents such as bacterial polysaccharides are able to induce an antibody response without the interaction T helper cells (Murphy, 2012). This phenomenon is known as the Thymus Independent (TI) mechanism.

In addition to the previously described effector T cells, T_H17 cells are involved in recruiting neutrophils in a response to extracellular bacteria and fungi and are also involved in the

generation of autoimmune diseases (Wilke *et al.*, 2011). T_{reg} cells are involved in the regulation of inflammation and the suppression of the T cell response and play a major role in the maintenance of immunological tolerance to self (Swain *et al.*, 2012).

A final subset of CD4+ effector T cells has recently been discovered, capable of directly performing cytolytic activity mediated by perforin and Fas proteins (Brown *et al.*, 2006). These cells are termed cytolytic CD4+ T cells (Figure 1).

As mentioned above, one of the principle roles of CD4+ T lymphocytes is the activation of B lymphocytes. The main function of these B lymphocytes, in turn, is the production of highly specialized proteins, namely immunoglobulins, responsible for recognizing antigenic molecules with exceptionally high specificity. These proteins, also known as antibodies, are described below.

I.1.C. Immunoglobulins

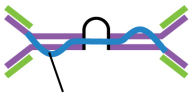
Immunoglobulins are glycoproteins present in both membrane-bound (BCR) and secreted (antibody) forms. They are the end products of B lymphocyte differentiation. BCRs are responsible for antigen recognition whereas antibodies have a number of different effector functions:

- Neutralization
- Opsonization
- Activation of the complement system and Complement Dependent Cytotoxicity (CDC)
- Induction of Antibody Dependent Cell Cytotoxicity (ADCC) by NKs
- Induction of Antibody Dependent Cell Phagocytosis (ADCP)

There are five classes of immunoglobulins, known as isotypes, namely IgA, IgD, IgE, IgG, and IgM. They each display different structures and functions (Table 2), with IgG being the most abundant. Each immunoglobulin is composed of a heavy chain and a light chain. Each chain is divided into two parts: a variable region, including the complementarity determining regions (CDRs) responsible for antigen recognition, interspersed by Framework regions (FR), and a constant region responsible for the effector functions. Immunoglobulin isotypes are determined by segments of the constant region of the heavy chains designated as α, δ, ε, γ, or μ, respectively. Similarly, there are two types of light chains: κ or λ. In humans the proportion of these two types of light chains κ:λ is 2:1, whereas in mice this proportion is of 20:1 or more (possibly 53:1 according to Dildrop *et al.*, 1987; Zocher *et al.*, 1995; and Almagro *et al.*, 1998). Through a number of different mechanisms, the B lymphocytes are able to express the remarkably vast diversity of variable regions needed to recognize the entirety of the immunogens individuals are faced with throughout their life time. Some of these mechanisms, including gene rearrangement and nucleotide addition and subtraction, are responsible for the initial diversity of the B

cells developing in the bone marrow. The following is a description of the mechanisms leading to diversification of the B lymphocytes prior to any contact with an antigen.

Table 2. Immonoglubulin isotypes with their different characteristics and functions.

The Five Immunoglobulin (Ig) Classes					
	IgM pentamer	IgG monomer	Secretory IgA dimer	IgE monomer	IgD monomer
					
Heavy chains	μ	γ	α	ε	δ
Number of antigen binding sites	10	2	4	2	2
Molecular weight (Daltons)	900,000	150,000	385,000	200,000	180,000
Percentage of total antibody in serum	6%	80%	13%	0.002%	1%
Crosses placenta	no	yes	no	no	no
Fixes complement	yes	yes	no	no	no
Fc binds to		phagocytes		mast cells and basophils	
Function	Main antibody of primary responses, best at fixing complement; the monomer form of IgM serves as the B cell receptor	Main blood antibody of secondary responses, neutralizes toxins, opsonization	Secreted into mucus, tears, saliva, colostrum	Antibody of allergy and antiparasitic activity	B cell receptor

Source: http://www.microbiologyinfo.com/antibody-structure-classes-and-functions/

As mentioned previously, out of the five immunoglobulin isotypes, IgG is by far the most abundant, representing 75-80% of the immunoglobulins present in the serum. The general structural features of all immunoglobulin isotypes are similar. The following section, therefore, focuses on a more detailed description of only the IgG isotype, which from here on out will also be referred to as an antibody.

I.1.C.a. Structure of IgG

Immunoglobulin G (IgG) has several subclasses (IgG1, IgG2, IgG3, and IgG4 in humans), with different levels of expression (Murphy, 2012). This molecule is a Y-shaped heterodimer protein of size 150 kDa, consisting of two light L (25 kDa) and two heavy H

(50 kDa) polypeptide chains (Figure 5). Each chain consists of a series of 110 amino acid long sequences, compactly folded, known as antibody domains. The heavy chain of IgG is made up of four, whereas the light chain has two of these domains. The two heavy chains are linked to each other by disulfide bonds, and each heavy chain is linked to a light chain by a disulfide bond. As mentioned before, there are two types of light chains, κ and λ, which can be found on any of the five isotypes. The isotype and effector functions of the antibody are defined by the structure of its heavy chain.

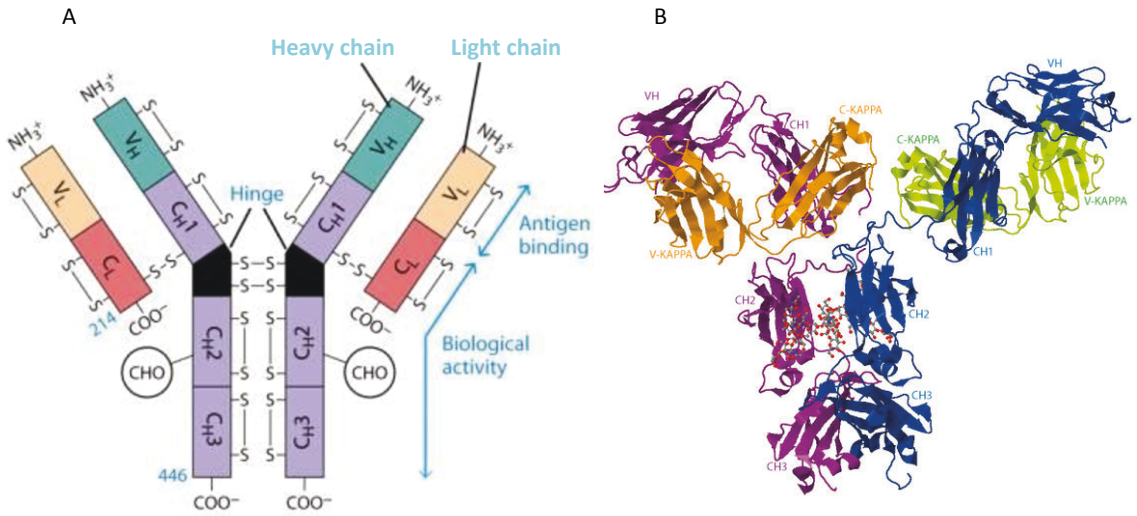


Figure 5. Structure of IgG.

A) 2D and B) 3D structures. Each IgG is a heterodimer protein consisting of 2 light chains and 2 heavy chains. Each chain is divided in turn into constant and variable regions. The variable regions are responsible for antigen recognition and contain the 3 Complementarity Determining Regions (CDRs). The hinge region in A is shown in black. CHO: N-linked carbohydrate.

Source: A) http://img.docstoccdn.com/thumb/orig/129870650.png. B) IMGT www.imgt.org

The amino-terminal domain of the light chain and the heavy chain are known as the variable domains (V domains V_L and V_H, respectively) and together make up the V region of the antibody responsible for the antigen recognition function. The remaining domains are known as the constant domains (C domains) and together make up the C region of the antibody. The multiple heavy chain C domains, C_H1, C_H2, and C_H3, are numbered from the amino-terminal to the carboxy-terminal of the antibody. There are sites on the antibodies that are glycosylated, with the majority being N-glycosylation but some with O-glycosylation. The sites and types of glycosylation depend on the immunoglobulin isotype. On IgG antibodies, for example, there are two sites of N-glycosylation, one on either C_H2 domain (Arnold *et al*., 2007).

The variability of the antibody V region is not evenly distributed, but is concentrated in three distinct segments designated as the hyper-variable regions (CDRs). There are three CDRs on each of the heavy and light chains, namely CDR1, CDR2, and CDR3, making a total

number of six hyper-variable loops. These six hyper-variable loops are brought together giving rise to the combinatorial diversity of antibodies, which is responsible for the ability to bind a wide range of antigens. CDR1 and CDR2 of both heavy and light chains are encoded by the V gene segments. CDR3 of the light chain is encoded by the joint segments between V and J gene segments, whereas the CDR3 of the heavy chain is encoded by the D gene segment (Figure 6A). The regions between the hyper-variable loops, which comprise the rest of the V domain, are known as framework regions and contain less variability. There are four of these segments for either of the light and heavy variable regions, namely FR1, FR2, FR3, and FR4 (Wu and Kabat, 1970) (Figure 6B).

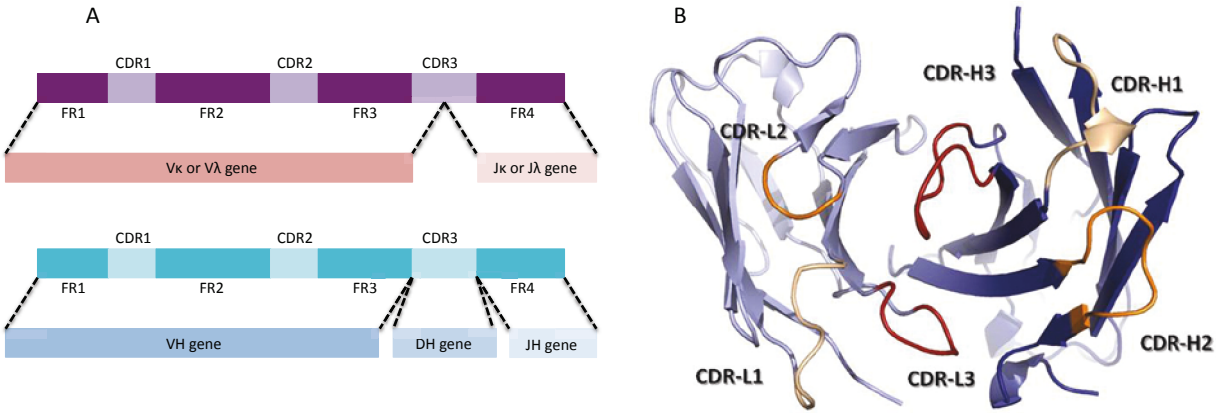


Figure 6. Antibody variable regions.

A) genetic association of the different segments, and B) The structure of CDRs indicated in loops. The regions between the CDRs are the framework regions. Source: Finlay and Almagro, 2012

The protein domains of antibodies associate to form three larger globular domains joined together by a stretch of flexible polypeptide chain known as the hinge region (Figure 5A). The different domains can be separated by enzymatic proteolysis or antibody engineering to produce different antibody fragments (Figure 7). The isolated form of the two antigen-binding domains of the molecule, for example, is known as a Fab fragment for *Fragment antibody binding*. The isolated constant region of the antibody, on the other hand, is known as the Fc fragment for *Fragment crystallizable*, and is responsible for the effector function of the antibody. The dimer form of a Fab fragment, in which the two antigen binding arms of the antibody remain linked, can also be isolated by enzymatic digestion. This is known as the $F(ab')_2$ fragment. Such a fragment can also be engineered as a bispecific antibody (Moran, 2011). Several other antibody fragments can be engineered. These include the single chain Fragment variable or scFv, which consists of a single chain with the V domain of the heavy chain linked to the V domain of the light chain with a short flexible linker. The scFv fragment is approximately 28 kDa and is widely used due to its small size and ease of manipulation (Monnier *et al.*, 2013).

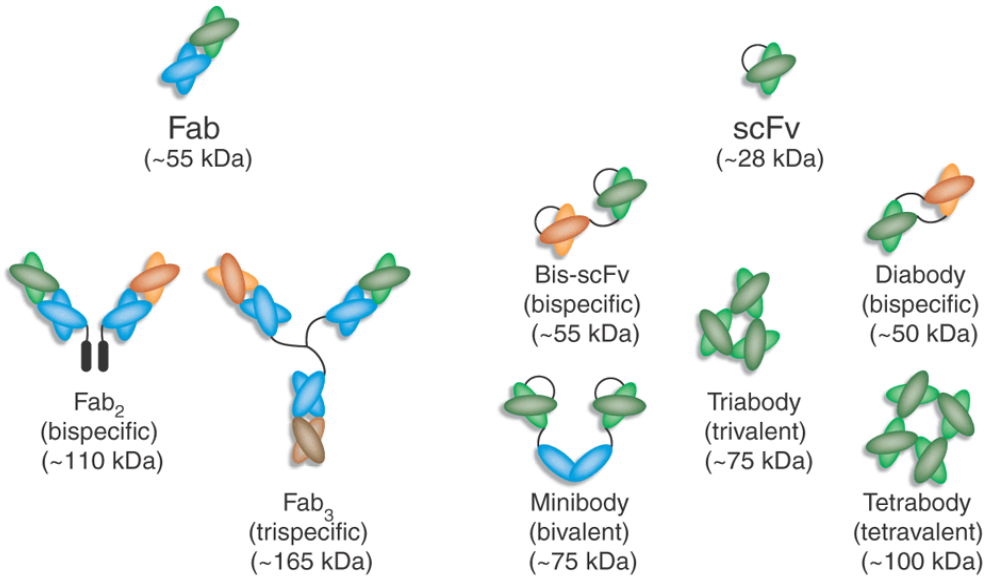


Figure 7. Antibody fragments with different formats. Orange and green segments represent light chains, wheras blue segments represent heavy chains.

Source: Moran, 2011

I.1.C.b. Gene rearrangement

Gene rearrangement, or somatic recombination, is the main event responsible for the diversity of mature B lymphocytes prior to their encounter with an antigen (Tonegawa 1983; Berek and Milstein, 1988; Neuberger and Milstein, 1995; Murphy, 2012). This event is based on the organization of gene segments encoding the variable κ and λ, and constant C regions of immunoglobulins (Kindt *et al.*, 2008). The variable region of the light chains κ and λ are encoded by two separate DNA segments, namely the variable (V) gene and the joining (J) gene. The variable region of the heavy chain is encoded by three DNA segments: V and J genes similar to the light chain, and an additional diversity (D) gene. These gene segments exist in multiple copies and are organized into three genetic loci present on different chromosomes (Figure 8). The κ light chain locus in humans is located on chromosome 2 and is composed of a cluster of approximately 34 to 37 functional V_κ gene segments followed by a cluster of 5 J_κ gene segments. The λ light chain locus in humans is located on chromosome 22 and is composed of a cluster of 29 to 35 functional V_λ gene segments, followed by a cluster of 4-5 J_λ gene segments. Finally, the heavy chain locus is located on chromosome 14, and resembles that of the κ locus, with 38 to 46 V_H, 23 D_H, and 6 J_H functional gene segments. The number of functional gene segments in mice is different from that of humans. In these animals, on the light κ chain, there are approximately 85 V_κ, 4 J_κ gene segments, whereas on the λ light chain, there are only 3 V_λ and 3 J_λ gene segments. For the heavy chain of these animals, there are around 134 V_H, 13 D_H, and 4 J_H gene segments. The number of gene segments in humans and in mice could vary slightly due to the polymorphism in different individuals.

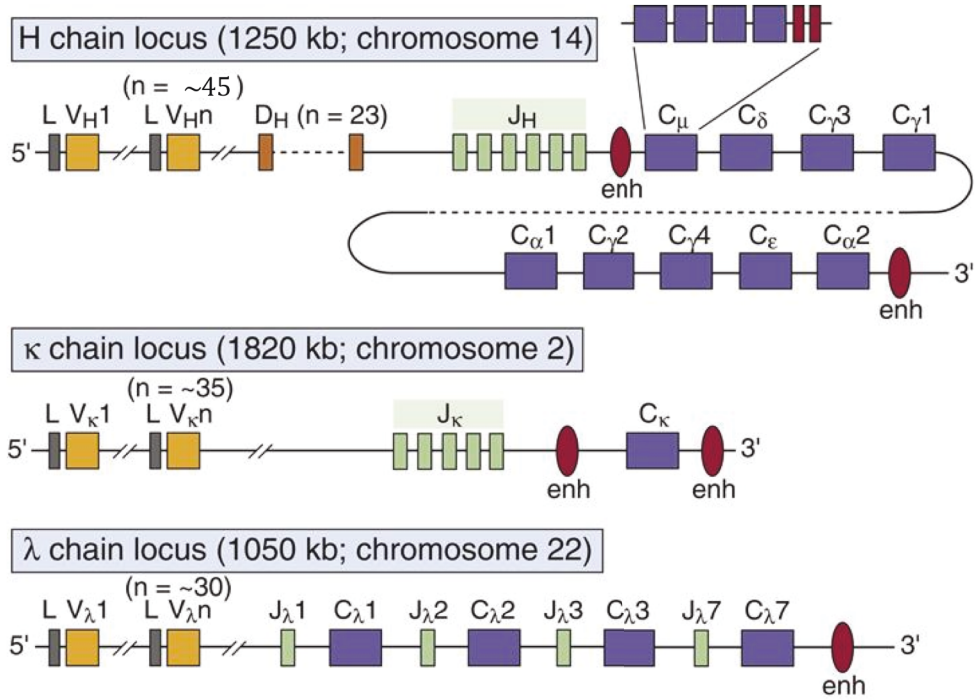


Figure 8. Germ-line organization of human immunoglobulin (Ig) loci.

The human heavy chain, κ-light chain, and λ-light chain loci are shown. Only functional genes are shown; pseudogenes have been omitted for simplicity. Exons and introns are not drawn to scale. Each C_H gene is shown as a single box but is composed of several exons, as illustrated for C_μ. Gene segments are indicated as follows: *L*, leader (often called signal sequence); *V*, variable; *D*, diversity; *J*, joining; *C*, constant; *enh*, enhancer. There exist multiple copies of V, D, and J gene segments for V_H and multiple copies of V and J for V_κ and V_λ. In contrast to V_H and V_κ, the copies of J gene segments of V_λ are separately located preceding each C gene, instead of being all clustered together in a single locus. Source: Abbas *et al.*, 2010

In premature B lymphocytes, the majority of the V region genes are far away from the C region genes. In mature B lymphocytes, however, gene rearrangement is responsible for the random selection of one segment per each gene, and bringing them in close proximity in order to produce a complete immunoglobulin chain (Figure 9). Gene rearrangement is achieved by specialized enzymes RAG1 and RAG2, named after their encoding recombination activating genes *RAG1* and *RAG2* (Recombination Activating Gene), and terminal deoxynucleotidyl transferase (TdT) (Melek *et al.*, 1998). These enzymes recognize conserved non-coding DNA sequences known as recombination signal sequences (RSS) present adjacent to the points at which recombination takes place. Gene rearrangement is the first source of the great diversity of immunoglobulins leading to 185 possible rearrangements for V_κ, 140 rearrangements for V_λ and 6300 rearrangements for V_H. In addition, the random association of different light chains with different heavy chains brings this diversity to approximately 2×10^6 different antibody molecules, known as the combinatorial diversity.

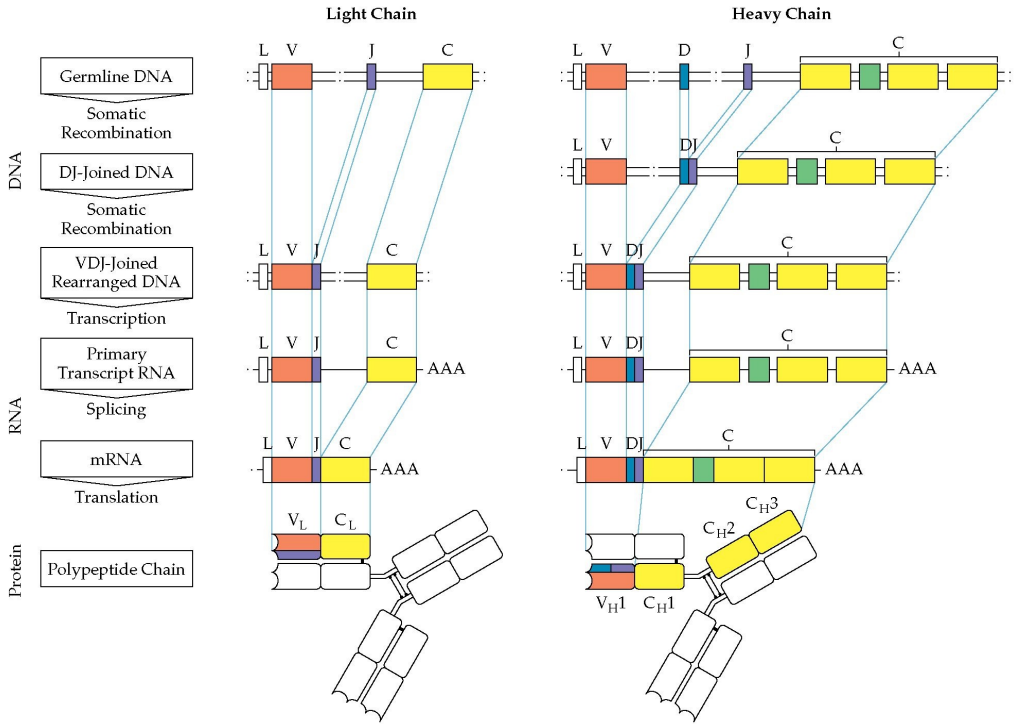


Figure 9. Immunoglobulin VDJ gene rearrangement producing complete heavy and light chains.

Source: Murphy, 2012

I.1.C.c. Nucleotide addition and subtraction

The CDR3 region of immunoglobulins has a significantly higher diversity introduced by the addition and subtraction of nucleotides at the junctions of different gene segments; this is known as junctional diversity (Murphy, 2012). The addition of nucleotides is performed by two processes: P-nucleotide addition or N-nucleotide addition (Sanz, 1991). P-nucleotides are so called due to the formation of palindromic sequences added to the ends of gene segments. This occurs when DNA cleavage leads to the formation of single stranded tails from a number of nucleotides on the coding sequence and the complementary nucleotides on the other strand. DNA repair enzymes then fill in complementary nucleotides on the tails, leaving short palindromic sequences. The RAG enzyme is continuously expressed in various stages of B cell maturation and P-nucleotide addition occurs in both light and heavy chain gene rearrangements. At later stages of B cell maturation, however, for the heavy chain and a small portion of the human light chain rearrangements, there is a second mechanism of N-nucleotide addition, so called due to their non-template-encoding procedure (Li *et al.*, 1993). N-nucleotides are added by the TdT enzyme to the single stranded ends of the coding DNA produced by the hairpin cleavage of the RAG1/2 complex. After the addition of N-nucleotides on the two strands of DNA, repair enzymes synthesize the complementary strands and ligate the new DNA to the palindromic region (Figure 10).

Nucleotides can also be randomly deleted at junctions of gene segments, leading to further diversity of the CDR3 regions. The addition of this junctional diversity increases the diversity of the heavy chain to 1.6×10^3 and of the light chain to 3×10^2, bringing the overall diversity of immunoglobulins to 10^{11}.

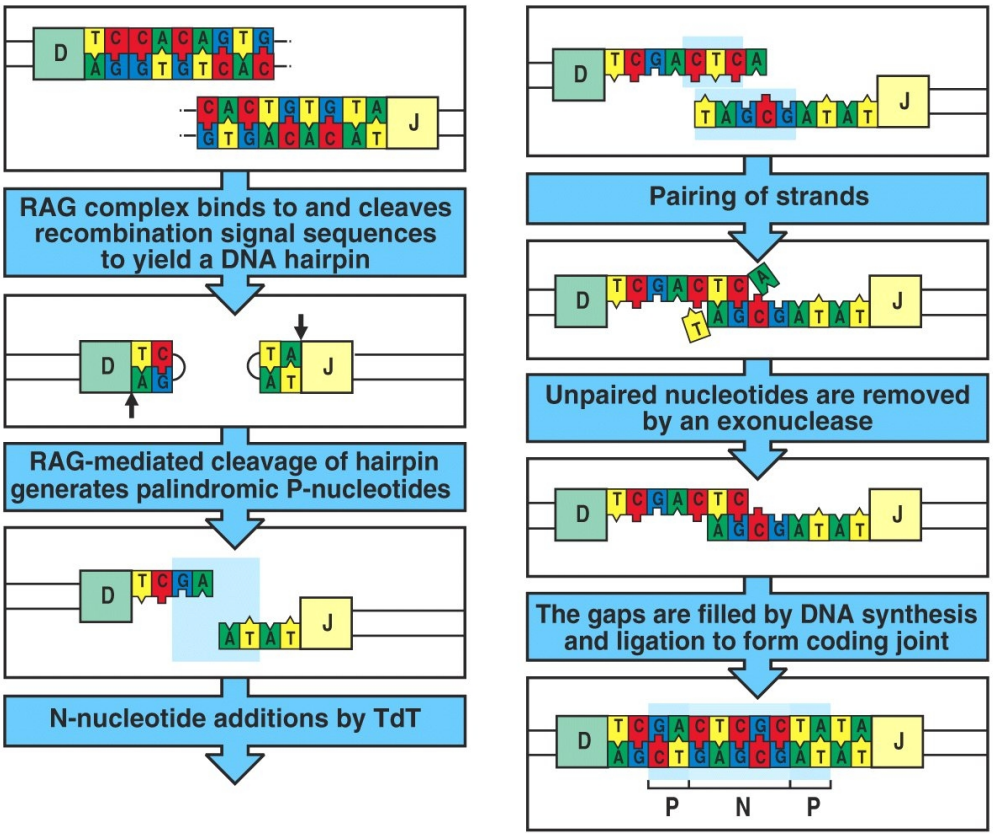


Figure 10. The mechanistic pathway of the addition of P- and N- nucleotides showing the different enzymes involved.

Source: Janeway *et al.*, 2001

Further diversification of the B lymphocytes, however, occurs only upon activation of the B cells by an antigen, which has been recognized by a BCR (Schroeder and Cavacini, 2010). The secondary mechanisms responsible for this further diversification lead to enhanced recognition and effector capacities of the produced immunoglobulins. They include class switching and somatic hypermutation as described below.

I.1.C.d. Isotype switching

IgM is the first immunoglobulin to be produced during B cell development, existing before maturation of the cell (Murphy, 2012). Once the B lymphocyte has completed maturation,

IgD is coexpressed with IgM on the surface of the cell. Upon activation of the B lymphocyte by an antigen, however, the cell goes through a process of class switching, leading to the expression of IgG, IgA, and IgE immunoglobulins. The enzyme responsible for the mechanisms involved in class switch recombination is AID, an activation-induced cytidine deaminase (Muramatsu *et al.*, 1999). This enzyme acts on cytidine residues present in sequences of repetitive DNA known as switch regions (S), which exist upstream from all of the constant region gene segments. Working in concert with the enzyme UNG and an additional apurinic/apyrimidinic endonuclease 1 (APE1) enzyme, it forms single strand nicks on the constant region gene segments. Multiple single strand nicks on both strands eventually lead to a double strand break in the DNA. Cellular mechanisms for repairing double strand breaks consequently come into play and lead to the nonhomologous recombination between switch regions leading to the deletion of IGHC genes located between the VDJ genes and the final IGHC gene to be used, and ultimately to class switching (Honjo *et al.*, 2002) (Figure 11).

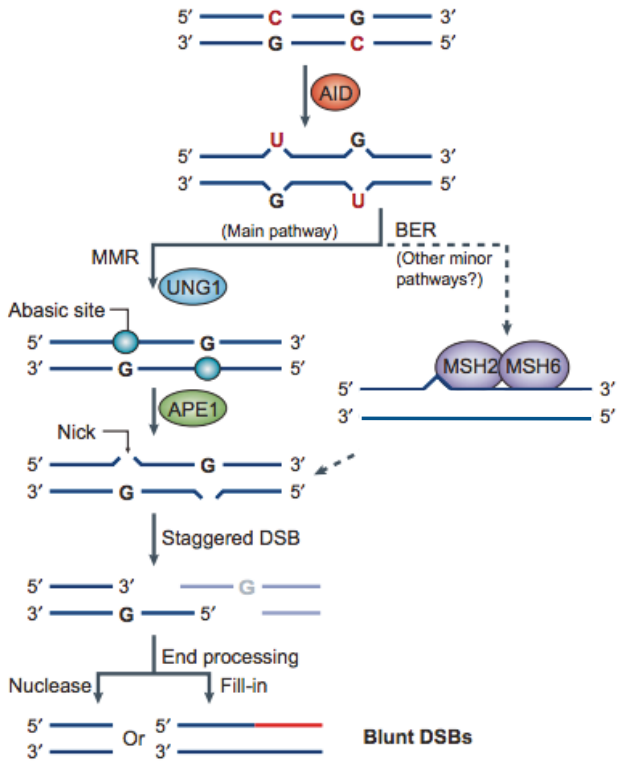


Figure 11. Mechanisms involved in class-switch recombination.

Activation-induced cytidine deaminase (AID) deaminates cytidine residues in DNA, converting them to uridine residues. The U:G mismatch can then be processed by either uracil DNA glycosylase (UNG), a component of the base excision repair pathway, or the mismatch-repair machinery (MSH2, MSH6)- resulting in gaps or nicks in DNA. During class-switch recombination the nicks induced by the BER pathway are thought to be generated by the following process: UNG removes the AID-introduced deoxyuridine in S-region DNA, creating an abasic site that is processed by the apurine/apyrimidine endonuclease 1 (APE1), which created the nick. Processing of the staggered ends by exonucleases or by error-prone end-filling reactions can lead to blunt or double-stranded breaks that can be ligated to similarly created breaks on downstream S-region DNA to complete class-switch recombination. Source: inspired by Nussenzwieg and Alt, 2004

The class switching mechanism is the only diversifying mechanism in maturing B cells described so far, which acts on the constant region of immunoglobulin heavy chains, as opposed to their antigen binding sites. Interestingly, a recent study has shown that the selection of B cells also acts upon another non-antigenic binding region of the heavy chain, namely the heavy chain variable "elbow" located between the variable and constant domains (Kaplinsky *et al.*, 2014). This study has demonstrated that B cells whose VH genes encode a more flexible elbow are more likely to go through maturation.

A second mechanism occurring in maturing B lymphocytes following activation by antigen presentation is somatic hypermutation. This mechanism occurs on the level of immunoglobulin variable regions and is described below.

I.1.C.e. Somatic hypermutation

Somatic hypermutation operates only on the actively transcribed rearranged variable regions of activated B lymphocytes, occurring mainly in germinal centers (Weill and Reynaud, 1996). It is responsible for performing point mutations into the V regions, which lead to changes in the affinity of the BCR for its antigen. These point mutations occur on segments known as "hot spots", located about 150 bp downstream from the V region promoter and extended over 2000 base pairs. The frequency of these point mutations is up to 100 000 times higher than the spontaneous mutations occurring on other genes (Pavri and Nussenzweig, 2011). Since in the germinal centers, the B lymphocytes are in competition with one another to produce clones that bind the antigen with the highest affinity, favorable mutations lead to the preferential selection of the B cells expressing them. This preferential selection leads to the phenomenon known as affinity maturation.

The introduction of point mutations is also performed by the enzyme AID (Muramatsu *et al.*, 1999), preferentially targeting hot spots located in the CDR regions. This enzyme binds only to single stranded DNA, therefore acts only on genes that are being transcribed. Its mechanism occurs in two phases: the first phase, engaging a base-excision repair pathway, is more predominant and targets C:G bases, whereas the second phase, engaging a mismatch repair pathway, targets A:T pairs. In the first phase, the AID enzyme transforms a cytidine into a uridine. The resulting U:G mismatch will be recognized by the enzyme uracil-DNA glycosylase (UNG), which will remove the uracil base from the uridine to create an abasic site in the DNA (Di Noia and Neuberger, 2007). Consequently, in the next round of replication, a random nucleotide will be inserted by DNA polymerase opposite the abasic site leading to a mutation. In the second phase, the presence of uridine is detected by mismatch repair proteins MSH2 and MSH6. The MSH2/6 proteins result in the complete removal of the uridine along with several adjacent nucleotides on the damaged DNA strand. A DNA polymerase then fills in the complementary strand with a highly error-prone synthesis, which introduces point mutations specially at nearby A:T base pairs (Roa *et al.*, 2010) (Figure 12).

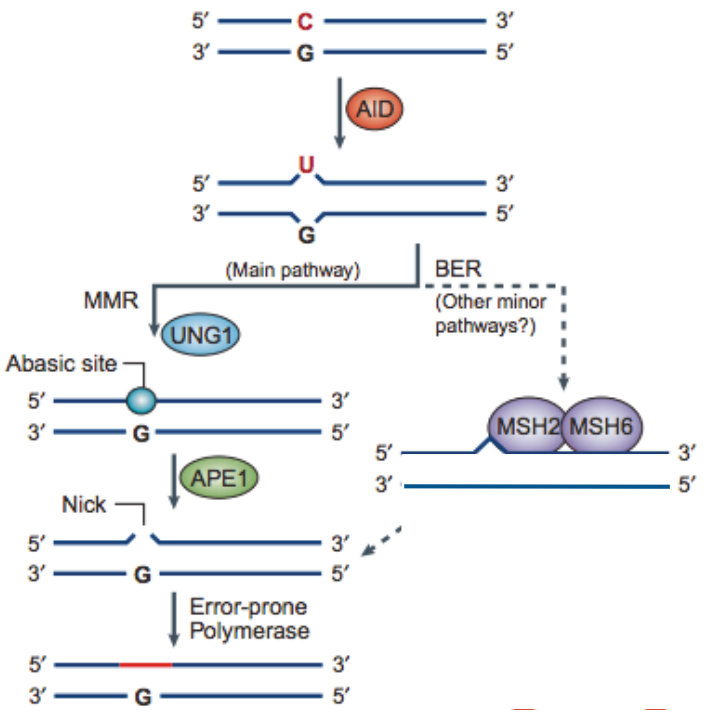


Figure 12. Mechanisms involved in somatic hypermutation.

During somatic hypermutation the U:G mismatch can simply be replicated to produce a C to T mutation. In a first phase, processing the nick ny UNG and the mismatch repair machinery can produce an abasic site that will produce a C to A or C to T change; in a second phase, a gap can be filled in by error-prone polymerase to produce mutations in nucleotides other than the targeted C. Source: inspired by Nussenzwieg and Alt, 2004.

The combination of somatic hypermutation and isotype switching mechanisms in humans increase the diversity of immunoglobulins up to 10^{14}-10^{15}.

As described in this section, the clonal selection of the B lymphocytes is much dependent on the events of somatic hypermutation. This selection pressure drives the B cell repertoire diversity toward the maturation of clones, which exhibit a high affinity to a particular pathogen. In the case of a vaccination event, following the evolution of the B cell diversity provides a valuable perspective on the mechanisms of clonal selection. The following section describes two such studies revealing information on the tendencies of somatic selection.

I.1.C.f. Immunization and clonal expansion

As previously discussed, the ability of the adaptive immune system to react to foreign pathogens is based on the diversity of its immune repertoire, which in the case of B lymphocytes is generated in two phases. The first phase of the generation of diversity involves the recombination of germline gene segments to create BCRs, whereas the second

phase involves somatic hypermutation mechanisms during an immune response. In an event of an immunization, or vaccination, the diversification processes leading to immune repertoire diversity are coupled with somatic selection pressures based on the binding affinity of the BCRs (Casola *et al.*, 2004). Nevertheless, studies have shown that the initial germline diversity has in fact a direct influence on the final results of B cell diversification and selection.

A study performed by Schwartz and Hershberg has compared the germline diversity of 12 individuals to their immune repertoire diversity following vaccination with the influenza virus (Schwartz and Hershberg, 2013). The authors have demonstrated a strong correlation between the diversity of the germline repertoire at a particular position and the number of B cell clones that have an amino acid change at that same position. The diversity of amino acids used in the mutated position, however, has logically been observed to be higher than the germline. They conclude that germline diversity is indeed a good indicator of the likelihood of survival following mutation. They assess, nevertheless, that this diversity cannot account for the specific amino acid whose usage accounted for the survival of a specific clone.

In a different study performed by Wang *et al.*, the authors have explored the influence of genetic background on the B cell repertoire and clonal expansions in response to attenuated varicella-zoster vaccination in identical twins (Wang *et al.*, 2015). The results indicated that identical twins show increased correlation in antibody gene segment usage, junctional features, and mutation rates in their antibody pools, however they demonstrate little similarity in clonal responses to an acute stimulus. Once again, the authors conclude that a shared germline genome sequence is correlated with overall convergence of antibody repertoires, even though the particular antibody response to a given vaccination is less predictable.

Therefore, we now know that the germline variable gene diversity could provide us with some predictive information on the final result on clonal expansion and an immunization-induced immune response. Indeed, the size and diversity of the immune repertoire clearly impact the clonal selection process, but the precise determination of the specific clone to be selected is not predictable. Nevertheless, for conducting immune repertoire studies, a representative window of the immune diversity is required in order to be able to mimic such a repertoire and to better understand a given immune response.

I.1.D. Defects of the immune system

The immune system is designed to protect the organism against foreign infectious agents, however, sometimes this intricate system leads to a detrimental response for the host organism. The instances in which the immune response is directed to a non-infectious agent include (i) allergies and hypersensitivity, when the antigen is an innoculous foreign substance, (ii) autoimmune diseases, when the response is to a self-antigen, and (iii) graft rejections, when the antigen is borne by a transplanted foreign cell.

Autoimmune diseases are numerous (Table 3) and affect approximately 5% of the population of the Western countries (Davidson and Diamond, 2001). They occur as a result of the failure of the immune system to successfully perform the necessary checkpoints in order to ensure self-tolerance. In a normal state, the immune system recognizes a number of surrogate markers in order to distinguish between self and non-self molecules. This recognition involves the mechanisms of central tolerance, leading to the elimination of clones with a high affinity toward self-antigens, as well as peripheral tolerance, making certain clones non-responsive or inhibiting their response via T_{reg} cells (Hogquist *et al.*, 2005; Kindt *et al.*, 2008; Rudensky, 2011). In an autoimmune state, however, as a result of failure in self-tolerance, some self-reactive lymphocytes escape clonal deletion, leave the thymus, and are activated to cause autoimmune diseases. Furthermore, most circulating lymphocytes have a low affinity to self-antigens but do not react to them. Certain situations such as an infection leading to a higher presence of inflammatory cytokines (Silverstein and Rose, 2000), however, can disrupt this homeostasis and activate these lymphocytes. Such cells then go through maturation and develop a higher affinity to a particular self-antigen, hence leading to an aberrant response. A more recent theory explaining the occurrence of autoimmunity suggests the implication of the idiotypic network (described in detail in section II.1.A.b.i) and a malfunction of the homeostasis in the concentrations of idiotypic antibodies *versus* anti-idiotypic antibodies (Sherer and Shoenfeld, 2000). It is now understood that there are both genetic and environmental factors leading to the pathological development of autoimmunity, but the understanding of the mechanisms involved remains one of the major challenges in the field of immunology today.

Table 3. Some common autoimmune diseases

Disease	Disease mechanism	Consequence	Prevalence
Psoriasis	Autoreactive T cells against skin-associated antigens	Inflammation of skin with formation of scaly patches or plaques	1 in 50
Rheumatoid arthiriris	Autoreactive T cells against antigens of joint synovium	Joint inflammation and destruction causing arthritis	1 in 100
Grave's disease	Autoantibodies against the thoroid-stimulating-hormone receptor	Hyperthyroidism: overproduction of thyroid hormones	1 in 100
Hashimoto's thyroiditis	Autoantibodies and autoreactive T cells against thyroid antigens	Destruction of thyroid tissue leading to hypothyroidism: underproduction of thyroid hormones	1 in 200
Systemic lupus erythrmatosus	Autoantibodies and autoreactive T cells against DNA, chromatin proteins, and ubiquitous ribonucleoprotein antigens	Glomerulonephritis, vasculitis, rash	1 in 200
Sjogren's syndrome	Autoantibodies and autoreactive T cells against ribonucleaoprotein antigens	Lymphocyte infiltration of exocrine glands, leading to dry eyes and/or dry mouth; other organs may be ivolved, leading to systemic disease	1 in 300
Crohn's disease	Autoreactive T cells against intestinal flora antigens	Intestinal inflammation and scarring	1 in 500
Multiple sclerosis	Autoreactive T cells against brain antigens	Formation of sclerotic plaques in brain with destruction of myelin sheaths surrounding nerve cell axons, leading to muscle weakess, ataxia, and other symptoms	1 in 700
Type 1 diabetes (insluin-dependent diabetes mellitus, IDDM)	Autoreactive T cells against pancreatic islet cell antigens	Destruction of pancreatic islet β cells leading to nonproduction of insulin	1 in 800

To conclude, the different mechanisms described in this section all play a role in the ability of the immune system to produce the vast diversity of immunoglobulin variable regions necessary for an individual's survival. This fascinating power of the immune system to recognize a quasi-limitless number of foreign antigens is a remarkable example of molecular diversity existing in nature. Inspired by nature, researchers have also tried to produce such levels of molecular diversity in the laboratory. The following section describes the technologies developed to date for this purpose.

I.2. Display Technologies, a synthetic source of diversity

Over the past few decades, a number of *in vitro* technologies have been developed for the discovery of new molecules. These technologies, together known as display technologies, are based on the ability to create a largely diverse pool, or library, of molecular structures, such as ligands, nucleic acids, peptides, proteins, antibodies, etc. and to employ a number of different criteria in order to isolate the desired candidate. Some display technologies are able to produce molecular diversities of up to or even superior to the immune repertoire diversity (Bradbury *et al.*, 2011). In general, all display methods use the principle of coupling the phenotype and genotype of the displayed molecules.

A number of display techniques have been developed to date. These include cell-based systems such as phage display and cell surface display (Georgiou *et al.*, 1997), including yeast display (Boder and Wittrup, 1997; Shusta *et al.*, 1999), or cell free systems such as ribosome display (He and Taussig, 2002) and mRNA display (Roberts and Szostak, 1997). Of these, the phage display technology is the most established and has been applied to numerous technological problems including the discovery of novel drugs and target antibodies. The following section is a more detailed description of this technology.

I.2.A. Phage Display

The phage display technology involves the expression of peptides, proteins, or antibodies on the surface of filamentous phage through fusion to endogenous phage coat proteins. The great advantage of this technology is the packaging of the genetic sequence of the molecule of interest within the phage, expressing the molecule of interest on the phage surface, and linking phenotype to genotype selection. Phage display was originally conceived by George Smith as a means to identify the sequence of the protein target binding to an antibody (Smith, 1985). Interestingly, this application has remained one of the most common uses of this technology, 30 years after its development.

I.2.A.a. Filamentous phage biology

Bacteriophages are viruses that infect a variety of Gram-negative bacteria using their pili as receptors. The Ff (Filamentous phage) particles (strains M13, f1, and fd), which infect *E coli* through their F pili, are the most extensively studied phage (Barbas *et al.*, 2001). They consist of a single-stranded (ss) DNA enclosed in a cylindrical capsid containing a number of different proteins. The genome of the phage is about 6400 nucleotides long and consists of 11 genes. The length of the cylinder consists of approximately 2700 copies of the 50 amino acid major coat protein (g8p or pVIII). At one end of the particle, there are about five copies of each of the 33-residue gene VII protein (g7p or pVII) and the 32-residue gene IX protein (g9p or pIX). The other end contains approximately five copies of each of the 406-residue gene III protein (g3p or pIII) and 112-residue gene VI protein (g6p or pVI). The DNA is oriented within the virion so that a 78-nucleotide hairpin region known as the packaging signal (PS) is always located at the end of the particle containing the pVII and pIX proteins (Figure 13a).

The Ff phage do not kill their host during productive infection, but only use their cellular components to replicate their DNA. Infection is initiated by the attachment of the phage g3p to the f pilus of a male *E coli*. Only the circular phage ssDNA enters the bacterium where it is converted by the host DNA replication machinery into the double-stranded (ds) plasmid like replicative form (RF). The RF undergoes rolling circle amplification and the result of this replicative process is a newly synthesized viral ss DNA in a complex with many copies of phage-encoded proteins. During the assembly process, the viral DNA is extruded through the membrane-associated assembly site, the phage DNA binding proteins are removed and the capsid proteins are package around the DNA. The phage coat proteins pIII and pVIII are generally used as sites of fusion with the gene sequence of interest. In this way, the recombinant molecules are expressed by the phage and displayed on its surface fused to the coat proteins (Figure 13b).

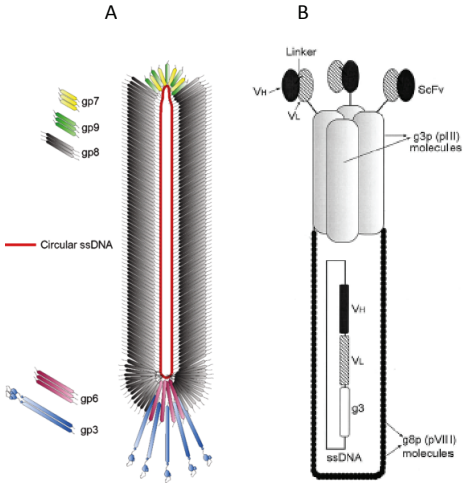


Figure 13. Structure of filamentous phage.

A) The bacteriophage particle showing the different coat proteins and the single stranded genomic DNA B) Simplified schematic of a filamentous phage displaying a scFv molecule. Source: A) Ploss and Kuhn, 2010 B) Azzazy and Highsmith, 2002

I.2.A.b. Phage display vectors

There are two types of vectors used for phage display: phage and phagemid. In phage vectors, the gene encoding the recombinant protein is included in the phage genome. Therefore, the infected bacteria produce phage particles that all display the protein of interest and each clonal population of phage are phenotypically and genotypically homogeneous. Phagemid vectors contain only the origins of replication for the phage and its host, the gene of the coat protein (i.e. pIII) fused to the recombinant gene of interest, and an antibiotic resistance gene. These vectors, therefore, require the assistance of helper phage (i.e. M13KO7 or VCSM13) to supply all of the structural proteins required for generating a complete virion with ssDNA, including pIII (Bazan *et al.*, 2012). Helper phages are normal Ff phages, which contain an additional origin of replication, an antibody resistance gene, and a severely defective packaging signal. When a bacterium is infected with both phagemid and helper phage, the phagemid DNA is packaged in preference. In this system, therefore, both wild type pIII and recombinant gene-fused pIII proteins will be present on the phage surface. In general phagemid particles are able to display from 0 to 4 or 5 copies of the recombinant molecule (Chasteen *et al.*, 2006) (Figure 14). Theoretically, the defective packaging signal should significantly reduce the number of helper phage particles, but a percentage still remains. Hence, the use of different antibiotic resistance genes by the helper phage and phagemid genomes allows the selection of phagemid particles displaying the recombinant molecule.

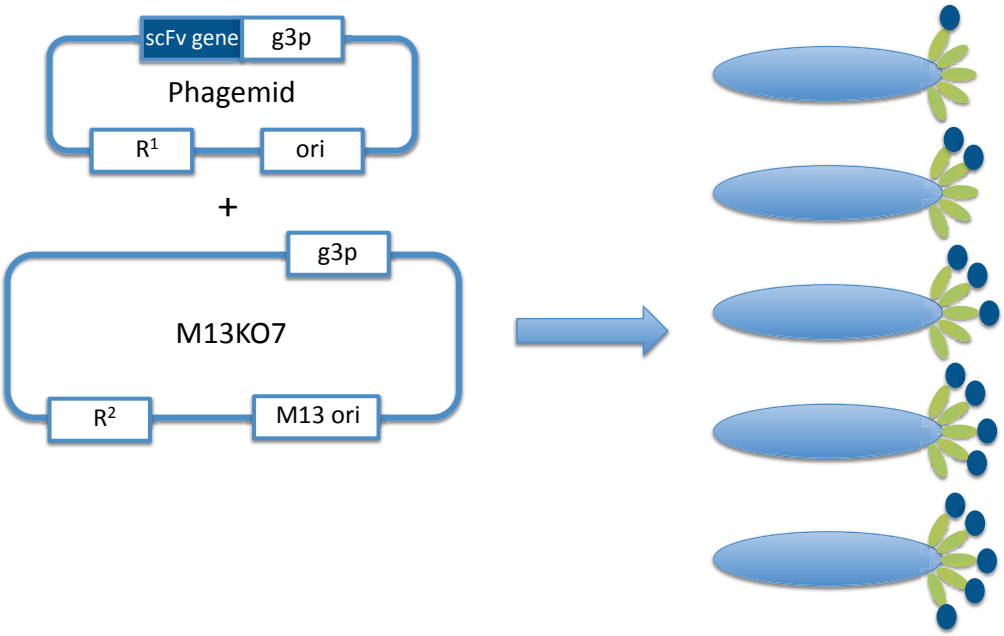


Figure 14. The possible phenotypes of phagemid particles.

Due to the heterogeneity of the phage including both the phagemid and helper phage genomes, the resulting phage particle can display from 1-5 scFv fragments. g3p represents the gene for protein 3, one of the surface proteins of the phage. R represents the resistance cassette in both plasmids.

There are a number of advantages and disadvantages associated with the two systems. Phagemid vectors facilitate the DNA manipulation and cloning steps, leading to the production of much larger libraries as compared with phage vectors. Insertion of an amber stop codon in combination with the high copy number[2] of the phagemid vector has also made this system better suitable for the soluble expression of the recombinant protein (Hoogenboom *et al.*, 1991). This is in contrast with phage vectors, which have low copy numbers and a weak g3p promoter. Another advantage of the phagemid is its genetic stability and relative resistance to deletions of exogenous genetic material. Phage vectors, on the other hand, tend to delete foreign DNA due to the selective growth advantage of a smaller phage genome. The major advantage of phage vectors is their operational facility. This is because they do not require helper phage and therefore eliminate all the technical procedures involved with helper phage infection. This makes them more suitable for high-throughput selection procedures (Hallborn and Carlsson, 2002). Secondly, using pIII for fusion and not taking into account eventual proteolysis, phage particles are able to display five copies of the recombinant molecule. This is contrary to phagemid vectors, whose heterogeneous genetic nature does not guarantee the expression of the five copies. As a result, in a selection process, phage vectors lead to the recovery of a greater number and a more diverse pool of binders, whereas phagemid vectors lead to the selection of fewer unique binders. As a counterbalance, however, the resulting binders from phage selection have lower affinity, whereas phagemid selection leads to the isolation of higher affinity binders (Huie *et al.*, 2001).

I.2.A.c. Display of antibody libraries

The use of phage display technology for the display of a vast antibody repertoire for the robust screening or selection of a variety of target antigens and rapid recovery of specific antibody candidates, first developed by Greg Winter and Richard Lerner (McCafferty *et al.*, 1990; Barbas *et al.*, 1991), is arguably one of the most successful uses of this technique today. Antibodies with affinities comparable to those obtained using traditional hybridoma technology can be selected from phage display libraries, and their affinity can be further increased by using the selected antibodies for subsequent library construction and selection. In this way, phage display has produced antibodies with optimized affinities (Bradbury and Marks, 2004).

There are two types of natural phage display libraries: naive and immune. The naive libraries are derived from natural unimmunized rearranged V genes containing mostly IgM isotypes (Marks *et al.*, 1991; Sblattero and Bradbury, 2000), synthetic V genes (Griffiths *et al.*, 1994; Nissim *et al.*, 1994) or shuffled V genes (Soderlind *et al.*, 2000). Immunized libraries are created from immunized humans or animals, contain mostly IgG isotypes (Barbas *et al.*, 1991; Orum *et al.*, 1993), have a strong bias toward antibodies of a certain specificity, and have much higher affinities for the antigen used. In addition to the natural

[2] A high copy number plasmid has >200 copies per cell *versus* a low copy number plasmid having 1-20 copies per cell (Friehs, 2004)

libraries, there now exist synthetic or semi-synthetic phage display libraries in which the diversity of CDR regions is increased synthetically by site-directed mutagenesis (Barbas *et al.*, 1993; Weber *et al.*, 2014).

The success of any phage display experiment to yield high quality antibody candidates depends on the size of the library and the diversity of the presented antibodies. The size of antibody phage display libraries generally ranges from 10^6 to 10^{10} (Christ *et al.*, 2006). Antibody libraries are constructed by the specific amplification of immunoglobulin V_L and V_H genes from B lymphocytes via polymerase chain reaction (PCR). The design of primer sets targeting the most possible immunoglobulin sequences is therefore highly influential to the library diversity. The amplified V_L and V_H are subsequently assembled together, cloned into an appropriate vector and transformed into bacteria for final phage production. The cloning and transformations steps are also of crucial importance for the maintaining of library diversity (Christ *et al*, 2006).

In general, the assembled antibody fragment can be in either of two forms: Fab or scFv, and the choice of which form to use is dependent on practical considerations. The general advantage of using Fabs is the possibility of their conversion into full length IgGs for further applications. However, Fabs are more difficult to assemble, more likely to be degraded, have lower yields as soluble fragments, and lower DNA stability due to their larger size (Barbas *et al.*, 2001; Rothlisberger *et al.*, 2005). Considering these parameters, scFvs are better candidates for library construction. There is however a main disadvantage of using scFvs, which is their tendency to (i) dimerize and (ii) aggregate into inclusion bodies once expressed in soluble form. The dimerization occurs when the V_H of one scFv interacts with the V_L of another. This effect can be reduced by increasing the length of the linker to more than 20 amino acids, however, it is difficult to eliminate competely and the purification of monomeric scFv is usually required (Bradbury and Marks, 2004). The formation of inclusion bodies during the soluble expression of scFv, on the other hand, is sequence-dependent and can be circumvented with appropriate expression systems (Ben Naya *et al.*, 2013).

The ability of the phage display technology to produce antibody libraries approximating the natural size and diversity of immune repertoires has lead researchers to take advantage of the availability of such libraries to perform immune repertoire studies. These studies usually involve the sequencing of large samples of antibodies via high throughout sequencing. The following section describes a number of such studies.

I.3. Immune Repertoire Studies

The total number of antibody specificities of an individual is known as the immunoglobulin or immune repertoire. As it was discussed in section I.1.C.b, this diversity is produced by gene rearrangement in the immunoglobulin variable regions and *via* a series of events

occurring after the recognition of an antigen by an immunoglobulin. There is increasing evidence, however, that the immune repertoire diversity of each individual is dependent on many different factors influencing the individual's immune response. These factors range from age and sex (Wang *et al.*, 2014; Giefing-Kroll *et al.*, 2015), susceptibility to autoimmune diseases (Ferreira *et al.*, 2009), immunogenic events such as infection or vaccination (Parameswaran *et al.*, 2013; Jackson *et al.*, 2014), and cancer (Zuckerman *et al.*, 2010). Sequencing of the immune repertoire under different immunological states provides a deeper understanding of the expressed and constantly evolving nature of the antibody repertoire and enables the characterization of this complex system. New generation sequencing (NGS) technologies for performing high-throughput sequencing (HTS) have contributed a great deal to immune repertoire studies (Fischer, 2011; Jackson *et al.*, 2013).

To give a number of examples, the group of Gregg Silverman has studied the influence of the B cell superantigen protein L (PpL) from *Finegoldia magna* on the host immune response and has shown by NGS that this V_L targeted microbial superantigen induces a global shift in the B cell repertoire (Gronwall *et al.*, 2012). A different study has used NGS to show that there are a number of differences in the V_H gene repertoire of patients with chronic chagas' heart disease as compared with the repertoire of healthy individuals (Grippo *et al.*, 2009). As a final example, the group of Dimiter Dimitrov has used this method to demonstrate a number of differences in the antibody repertoire of human cord blood cells in neonates as compared with the repertoires of adult subjects (Prabakaran *et al.*, 2012).

Not only has NGS expanded our knowledge about the nature of different immune repertoires, but it has also opened new doors to numerous other applications. It has led to the discovery of new alleles of the V region of immunoglobulins (Boyd *et al.*, 2010; Gadala-Maria *et al.*, 2015). It has enabled the bypassing of screening procedures after antibody selection by the sequencing of the pool of selected antibodies and the visualization of the highly polarized clones (Reddy *et al.*, 2010), or the bypassing of *in vitro* selection altogether by performing an *in silico* selection by molecular docking methods (Ravn *et al.*, 2010). NGS has also led to the better characterization of the content and quality of phage display libraries by a more thorough validation of the library diversity (Glanville *et al.*, 2009). It is important to note, however, that NGS platforms are prone to a higher error rate compared to conventional Sanger sequencing methods, which needs to be taken into consideration during the analysis of the data (Kircher and Kelso, 2010). Nevertheless, NGS has become an extremely powerful tool in repertoire analyses and will continue to revolutionize antibody research and discovery.

I.4. Generation of diversity: Construction of antibody libraries representing four immune repertoires

In the present work, we have utilized the phage display technology in building four antibody libraries from murine models of different genetic background and immune status,

hence creating an increased molecular diversity encompassing all four immune repertoires. Similarly to the studies mentioned in the last section, we have performed an immune repertoire analysis, exploring existing genetic differences between the four repertoires.

I.4.A. Choice of immune repertoires

In order to select antibodies endowed with catalytic activity from a vast diversity of immunoglobulin genes and to study the potential influence of the nature of the immune repertoire on the expression of abzymes, we have constructed four antibody libraries with different genetic backgrounds and immunological states. The libraries are of murine origin and consist of i) healthy and naive, ii) autoimmune-prone and naive, iii) healthy and immunized, and iv) autoimmune-prone and immunized mice (Figure 15). In order to be consistent across the libraries, all four are focused only on IgG isotypes, and therefore, what we consider naive is a non-immunized repertoire. The healthy repertoires are represented by the murine strain BALB/C, whereas the autoimmune-prone repertoires are represented by the strain SJL/J, an experimental allergic encephalomyelitis (EAE) model (Brown and McFarlin, 1981). The SJL/J strain is different from the Balb/C due to a mutation in the dysferlin protein inducing an active myopathy in these mice. Furthermore, it has been shown that this strain shows an elevated T_{H1} response following an infection due to an overproduction of IL-12. Also, the SJL/J strain is subject to autoimmune disease due to some still unknown anomalies of suppressor genes participating in immune tolerance (Amagai and Cinader, 1981; Alleva *et al.*, 2001). This strain is used as a murine model for the study of multiple sclerosis, an autoimmune disease (Martin and McFarland, 1995). Finally, the SJL/J strain has a high tendency to develop B cell lymphomas and was used some time ago for the study of Hodgkins lymphomas (Kumar, 1983). By including this model, we aim to examine whether an autoimmune background, in line with previous observations, does indeed increase the possibility of the formation of catalytic antibodies (Tawfik *et al.*, 1995; Gabibov *et al.*, 2002). Our objective is to construct a library with the highest possible size and diversity. Considering the hypothesis that the Balb/C and SJL/J strains have different immune repertoires independently of autoimmunity, including the SJL/J repertoire in the library allows the collection of a different set of genes, some of which might not exist in the Balb/C strain. Furthermore, the inclusion of the immunized repertoires, produced by a strategically chosen target antigen, also satisfies this goal. Thus, we are able to improve our chances of obtaining the desired catalytic antibodies or any other antibodies of interest.

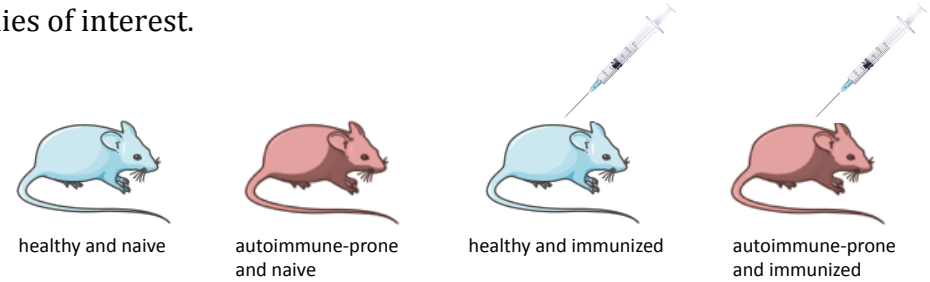


Figure 15. Four mouse models representing four immune repertoires of different genetic backgrounds and immunological states.

I.4.B. Choice of the antigen

We have performed a reactive immunization procedure for the production of the immunized libraries (Wirsching *et al.*, 1995). This method promotes the clonal expansion of antibodies, which have been selected via a chemical reactivity toward the antigen, and therefore, are capable of performing a chemical modification of the antigenic compound (Tanaka and Barbas, 2002) (described in more detail in chapter 2). In line with previous work done in our laboratory and the availability of a well-characterized suicide inhibitor of the enzyme β-lactamase, we have focused on this activity, and have therefore chosen this inhibitor, namely a penam sulfone derivative (Figure 16), as the antigenic entity used for immunization (Vanwetswinkel *et al.*, 1995). Furthermore, considering that the β-lactamase activity involves the hydrolysis of an amide bond (described in more detail in section II.2.A.), we approach the possibility to also generate antibodies with peptidase or protease activity. It is important to note that this method has been previously used by other groups for the selection of abzymes (Tanaka *et al.*, 1999) and enzymes (Avalle *et al.*, 1997) endowed with β-lactamase activity.

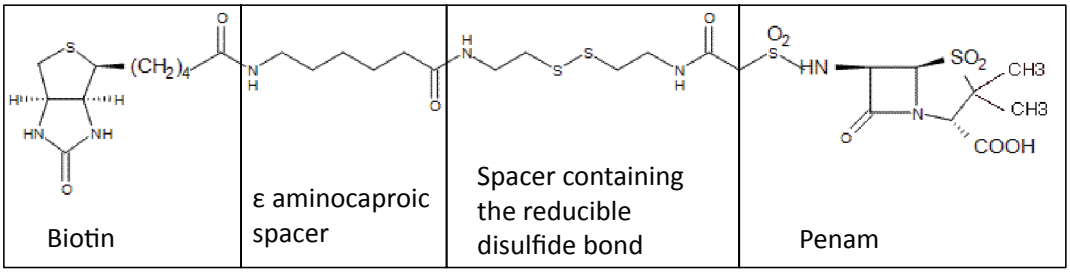


Figure 16. Penicillin sulfone derivative connected to a biotin by a spacer containing a disulfide bond.

I.4.B.a. Synthesis

The synthesis of this molecule was outsourced to the company Alpha Chimica (Chatenay Malabry, France). The strategy based on the publication of Vanwetswinkel and collaborators (1995) is to synthesize a biotinylated spacer arm **A** and to conjugate it to a β-lactam derivative **B** previously synthesized. The product **C** is then used in a reaction to produce the final penam derivative **D.** This strategy is very complex and involves over 10 steps starting from the commercially available precursors and has a reported yield of less than 8% (Figure 17).

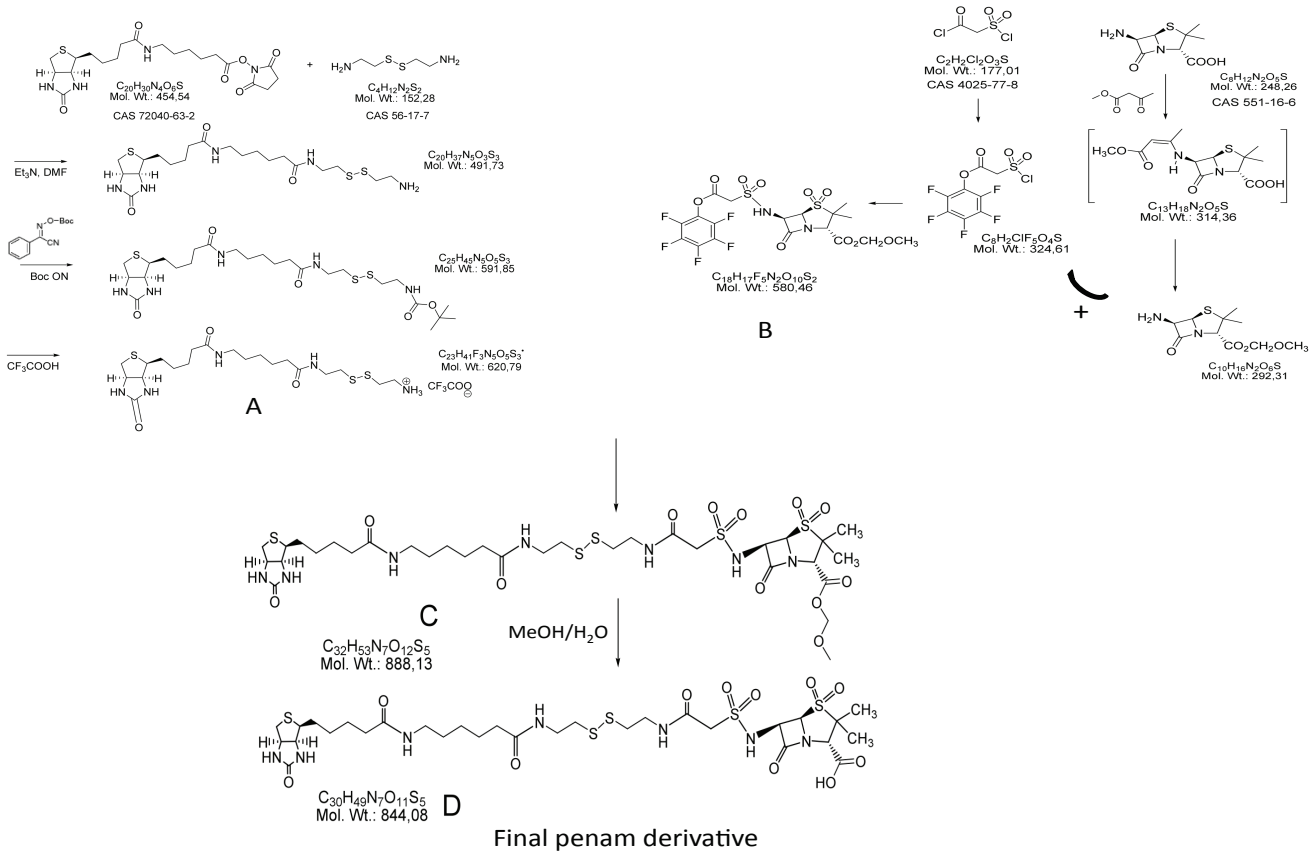


Figure 17. Strategy for the synthesis of the penam sulfone derivative.

The synthesis of the spacer **A** is performed by 4 steps starting from an N-hydroxysuccinimide ester derivative of biotinyl-N-ε-aminocarpoic acid, a compound provided by Sigma Aldrich (CAS: 72040-63-2). The yield of this procedure is reported to be 20-30% with the most difficult part being the purification of the products by chromatography on silica columns. The synthesis of the penam derivative **B** is also performed by 4 steps. This process is particularly difficult because the pathway intermediates are extremely unstable and easily degraded during the purification of the products. During the production steps, it is absolutely necessary for the reactions to be performed in an anhydrous environment and an inert atmosphere. Due to the sensitivity of the reactions, it was difficult to reproduce the same yield of product after each reaction. The average yield, however, was estimated to be less than 10%.

I.4.B.b. Inhibition test

An inhibition test by the synthesized penam sulfone derivative was performed on the enzyme β-lactamase I of *Bacillus cereus* 569/H/9. Penam sulfone at 1.67 mM was incubated with the enzyme at 110.8 nM during 30min at 4°C, making the ratio [Inhibitor]/[Enzyme] ≈

15 000 (Vanwetswinkel *et al*, 2000). In the final reaction, with the presence of ampicillin, the enzyme is at 3.18 nM and the inhibitor at 48 µM. The kinetics of the enzymatic reaction were followed with and without the presence of inhibitor. In the absence of inhibitor, the kinetic parameters are V_{max} = (1.61±0.19)×10^{-4} M.min^{-1} and K_M = (4.40± 0.99)×10^{-4} M, which gives a k_{cat} value of 5.09×10^4 min^{-1} coherent with a previous publication (Padiolleau-Lefèvre *et al.*, 2006). With the presence of the inhibitor, the kinetic parameters are V_{max} = (1.66± 0.16)×10^{-4} M.min^{-1} and K_M = (6.61± 0.97)×10^{-4} M. Therefore, in the presence of inhibitor, we observe an increase in the K_M value (even though small) while the V_{max} value remains unchanged, confirming that the penam sulfone molecule indeed displays an inhibitory capability resembling a competitive inhibition mechanism. We have calculated the K_i to be approximately 9.56×10^{-5} M. This value is quite low, the reason for which might be the fact that we have overestimated the concentration of inhibitor used, due to the lack of its efficient solubilization in DMSO. The inhibition is more noticeable with a Lineweaver Burk linearization, characteristic of a competitive inhibition (Figure 18). We did not perform further kinetic tests due to the high cost of the penam sulfone derivative.

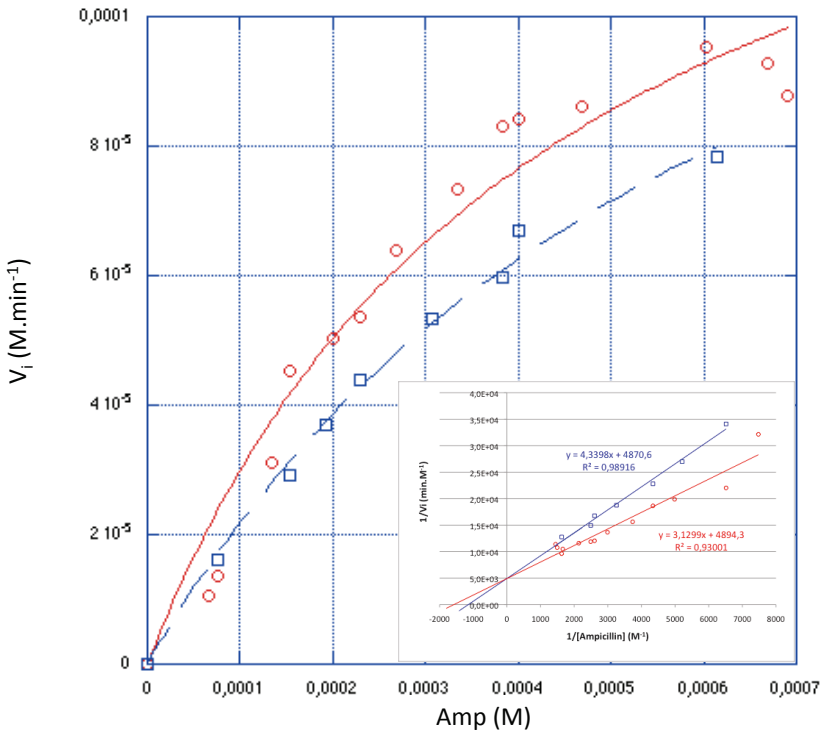


Figure 18. Inhibition of the enzyme β-lactamase by the penam sulfone derivative.

The kinetics of the enzyme with (indicated in blue) and without (indicated in red) the inhibitor are shown both as following a Michaelis-Menten behavior and as a Lineweaver Burk linearization. The concentration of enzyme without inhibitor is 3.18 nM. For the inhibition test, the enzyme is at 3.18 nM and the inhibitor at 48 µM, making the ratio inhibitor/enzyme=15000.

I.4.B.c. Immunization

The immunization of mice was outsourced to the company In-Cell-Art (Nantes, France). The immunization procedure was a short immunization, inspired by the work of Dan Tawfik (Tawfik *et al.*, 1995). Tawfik demonstrated that a long immunization process induces the expression of clones with high affinities toward the antigenic target, but reduces the expression of catalytic clones. This is in line with the hypothesis that the mechanisms of somatic hypermutation distant the final immunoglobulin genes from their germline sequence for high affinity antibodies, whereas catalytic antibodies have higher homologies to their germline, with lower mutations leading to lower structural constraints and therefore higher catalytic potential (described more in detail in section II.1.B.c.). Thus, only one injection of the immunogen was performed and spleens were collected after 21 days. The immunogen was the penam sulfone derivative conjugated to a Keyhole Limpet Hemocyanin (KLH) carrier, by use of the biotin/avidin interaction. In order to ensure the quality of the immune response, an enzyme linked immunosorbent assay (ELISA) was performed against the KLH carrier molecule using different dilutions of sera (Figure 19). These results confirmed the development of a similar immune response against KLH for both strains BALB/C and SJL/J. The results, however, do not confirm that an immune response was also developed toward the penam sulfone derivative. In order to confirm this, we need to perform a similar ELISA test against this molecule. However, due to the absence of sufficient quantities of sera (provided by In-Cell-Art), we were not able to perform this analysis.

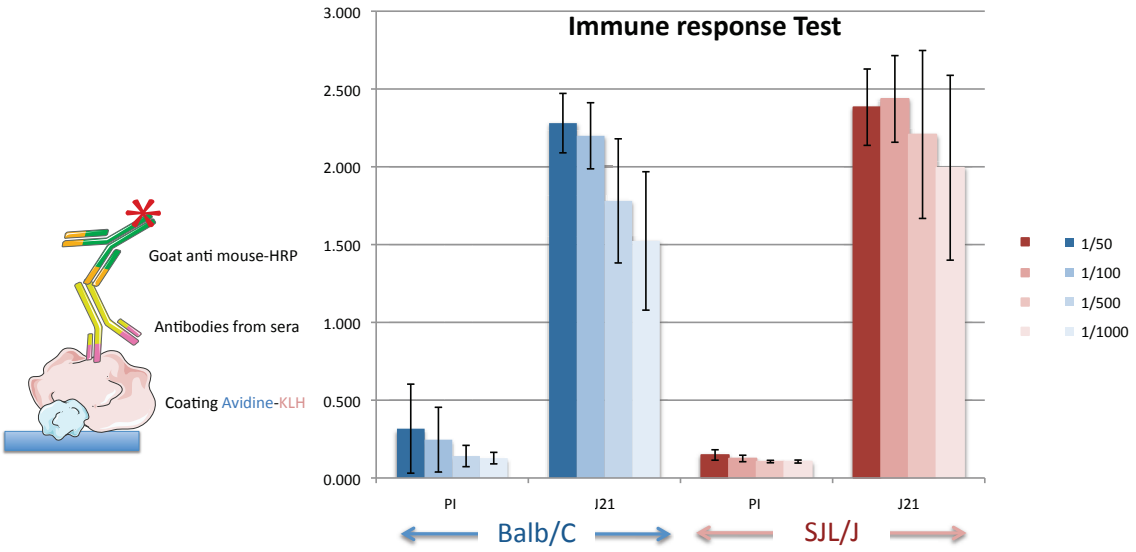


Figure 19. ELISA test confirming the induction of an immune response of both strains against KLH.

The Balb/C strain is shown in blue and the SJL/J strain in red. The different shades of each color show the different dilutions of sera. The graph clearly illustrates an increase in the response of both strains on day 21 as compared with the pre-immunization (PI) response. There is no significant difference between the immune response of the two strains.

I.4.C. Library construction

The four immune repertoires are constructed as single chain fragment variable (scFv) libraries by using the phage display technology. The global strategy is illustrated in figure 20. Briefly, total RNA is isolated from the mouse spleens. A reverse transcription is performed by using isotype-specific primers to produce only immunoglobulin G cDNA. The immunoglobulin variable regions V_L and V_H are amplified separately by using specific mixes of primers and are joined together by a linker sequence to form scFv fragments. These fragments are then cloned into phagemid vectors, transformed into *E coli*, and finally expressed on the surface of phage.

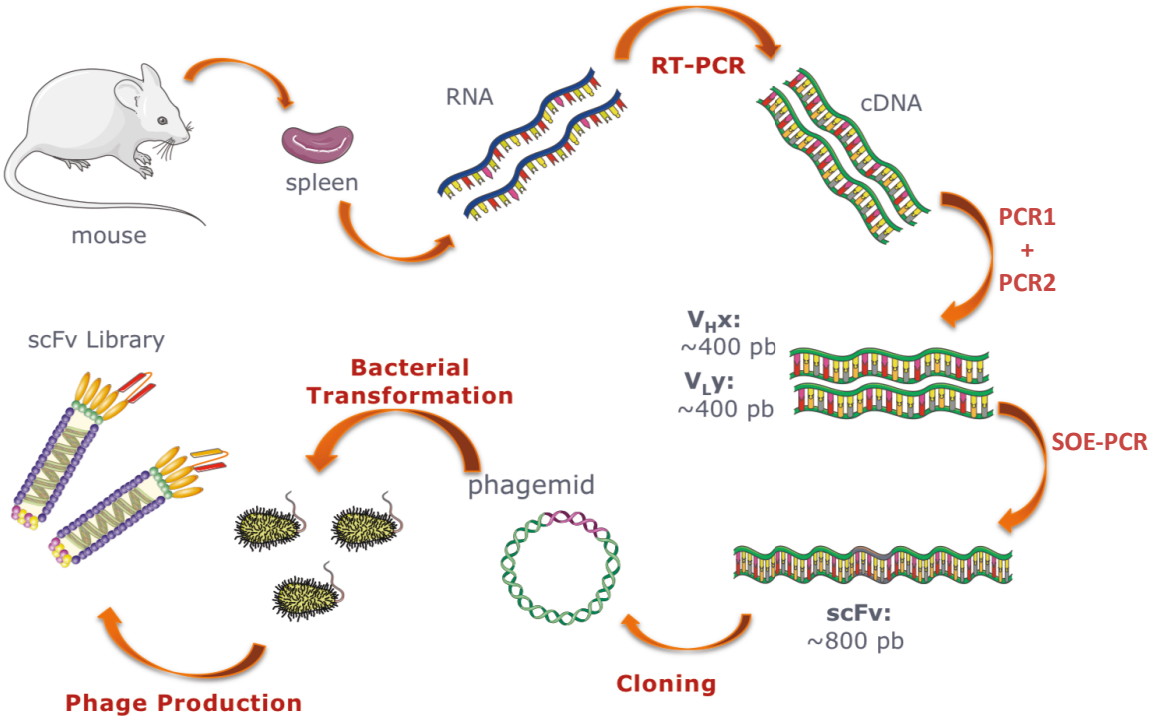


Figure 20. Construction of scFv library by phage display.

I.4.C.a. Design of primers

In an attempt to increase the yield and specificity of the amplification of immunoglobulin variable regions for library construction, we have decided to reduce our pool of DNA sequences starting from the reverse transcription reaction. We have used isotype-specific primers to target only rearranged IgG RNA. We have designed primers specific to the constant region CH3 of the heavy chain gamma 1, 2a, 2b, and 3 for the amplification of V_H, and primers specific to the constant regions of the light chains κ and λ for V_L. These primers are listed in table 4 and allow for the production of heteroduplexes following the RT-PCR reaction.

Table 4. Primers for RT-PCR.

Primer	Sequence
IgK	GGG TGA AGT TGA TGT CTT GTG AGT GGC
IgL	GCA GGA GAC AGA CTC TTC TCC AC
IgG1	GTG GGA GAG GCT CTT CTC AGT ATG GTG G
IgG2a	ACC CGG AGT CCG GGA GAA GCT CTT AGT CGT GTG G
IgG2b	ACC CGG AGA CCG GGA GAT GGT C
IgG3	GTT ATG GAG AGC CTC ATG CAC CAC GG

The primers target constant regions of immunoglobulin light κ and λ, and immunoglobulin heavy γ sequences, used in RT-PCR and PCR1.

In line with the previously mentioned objective, we have applied a two-step PCR strategy (Figure 21). The approach of targeting L-PART regions in order to keep the integrity of the 5' ends of immunoglobulin sequences has previously been reported in literature (Chardès *et al.*, 1999; Schroeder and Cavacini, 2010). Inspired by this strategy, we have designed a novel set of oligonucleotide primers targeting the murine leader sequences L-PART1 to be used as the 5' outer primer set for the first amplification step (PCR1).

Our design strategy is similar to that applied by the group of Didier Boquet (Essono *et al.*, 2003) and takes into account a newly updated and thorough analysis of the murine immunoglobulin variable gene sequences listed in the IMGT/LIGM-DB database, consulted in June, 2013 (Giudicelli *et al.*, 2006). The system of annotation in the IMGT® database permits the selection of data according to different criteria such as species (but not the strains), isotype, and functionality, set by the user. It also allows access to either germline or cDNA (sequences obtained from cloning of recombinant antibodies) immunoglobulin sequences.

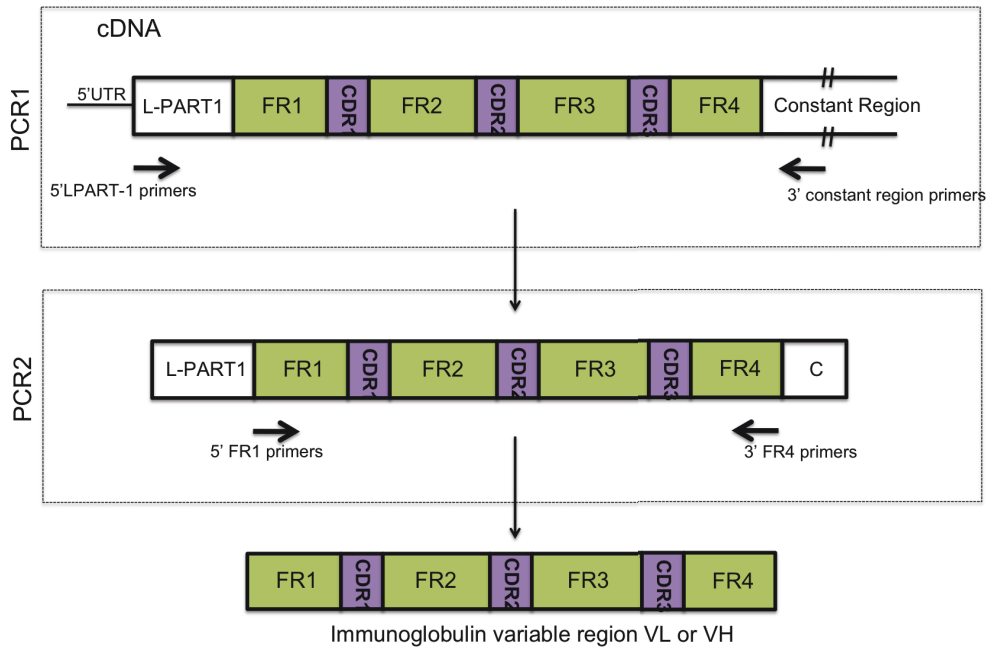


Figure 21. Two-step PCR strategy for the amplification of V_L and V_H.

We have chosen to include all referenced germline sequences in our alignment in order to avoid the exclusion of any potential sequences that have not yet been cloned. We have selected, however, only functional or productive sequences (open reading frame without stop codons and with normal splicing). Each of the immunoglobulin variable gene subgroups, IGKV (19 distinct subgroups), IGLV (2 distinct subgroups), and IGHV (15 distinct subgroups), is retrieved independently. This is performed by selecting the species of *Mus musculus* in the database, selecting the option of germline sequences, then selecting either Ig-Heavy or Ig-Light and doing a primary search. Then, we focalize the search output by selecting the "more selection" option and specifying a subgroup at each time.

Finally, we select for the sub-sequence of L-PART1 and confine our search output to this region. The sequences outputted by the database for each subgroup are aligned and subdivided into smaller clusters based on their percentage of identity (>68%). A number of degenerate oligonucleotide primers are designed for each of these clusters based on the following criteria: *i)* the length of the sequences is between 15-20 base pairs, *ii)* more than 50% of each sequence consists of GC base pairs, and *iii)* the level of degeneracy for each primer is less than 32%. In this manner, the ensemble of the designed oligonucleotide primers covers the entirety of the immunoglobulin sequences present in the IMGT-LIGM-DB database.

For the amplification of IGHV genes, we have designed 19 primers with a total degeneracy factor of 89. For IGKV we have designed 30 primers with a total degeneracy factor of 240. Finally, for IGLV we have used 2 primers with a degeneracy factor of 3. The designed L-PART1 primers are included in the Annexes (Table A.1.). The 3' outer primers for PCR1 are the same primers used for the reverse transcription of cDNA targeting the constant regions of immunoglobulin genes (Table 4).

The inner primers for the nested amplification step (PCR2) target sequences in FR1 at the 5' extremity, and sequences in FR4 at the 3' extremity. The 5' FR1 primers for the amplification of the heavy chain and the light chain κ and λ are inspired by a previous report (Essono *et al.*, 2003). These primers have been nevertheless modified in order to reflect an updated analysis of the sequences available in the IMGT® database and to include a segment of the linker sequence used to assemble the scFv fragments and the SfiI restriction site needed for the cloning process (Table A.2. in the Annexes). The 3' FR4 primers for the amplification of heavy chain and light chains κ and λ are inspired from the oligonucleotides reported by Plueckthun's group (Krebber *et al.*, 1997). These primers have also been redesigned and previously published by us (Shahsavarian *et al.*, 2014) (Table A.3. in the Annexes). Therefore, in the nested PCR reaction, specific to FR1, we have designed a total number of 35 IGKV primers with a total degeneracy factor of 460, a single IGLV with a degeneracy factor of 1, and 15 IGHV primers with a degeneracy factor of 56. The primers specific to FR4 have unique sequences and consist of 5 primers for the light chains κ and λ combined, and 4 primers for the heavy chain.

I.4.C.b. Amplification of immunoglobulin variable genes and assembly of scFv

The first step of building an scFv library is to obtain the coding sequences of antibody variable regions of the immune repertoire. Total RNA was isolated from the spleen tissue of each of the four mouse models and was observed to have good integrity and expected 28S:18S ratio of 1.8 (Figure 22).

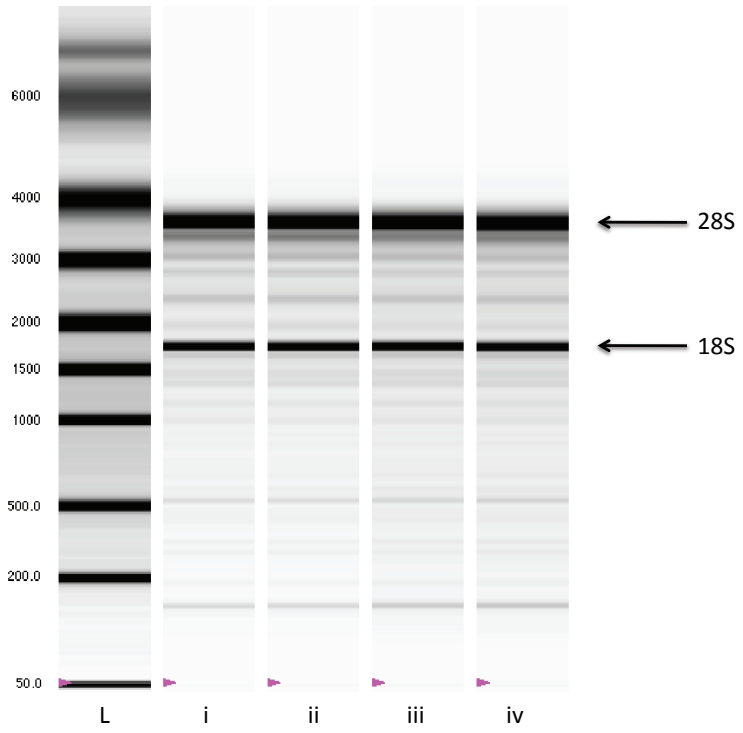


Figure 22. Testing of RNA integrity by Experion RNA StdSens Analysis Kit (Biorad).

Migration profiles of isolated total RNA from the 4 mouse models i) naive Balb/C, ii) naive SJL/J, iii) immunized Balb/C, and iv) immunized SJL/J are shown in comparison with standard RNA ladder (L). Distinct peaks of 18S rRNA and 28S rRNA are observed in the electrophoregram for both aliquots indicating good RNA integrity.

The IgG RNA is reverse transcribed into cDNA by reverse transcription polymerase chain reaction (RT-PCR), using isotype-specific primers, which can hybridize with the constant regions of immunoglobulins (I.4.C.a.). The variable regions are subsequently easily targeted and amplified, by using a mix of specific primers, in this restricted pool of cDNA containing only immunogobulin sequences. This amplification step is performed in two steps via a nested-PCR strategy. After the first reaction (PCR1), targeting L-PART1 at the 5' end and constant regions at the 3' end, the amplificates are not visible by DNA gel electrophoresis

analysis due to their low concentration. Following a second nested amplification (PCR2), however, the variable region amplificates are clearly observed indicating bands at the expected size of ~400 bp for V_L and V_H (Figure 23). We have used a model antibody 9G4H9, an IgG2b, which has previously been produced in our laboratory and readily available, as a positive control. It would be interesting to use antibodies of various isotypes as control, however, we have contented to the used of 9G4H9 due to practical reasons.

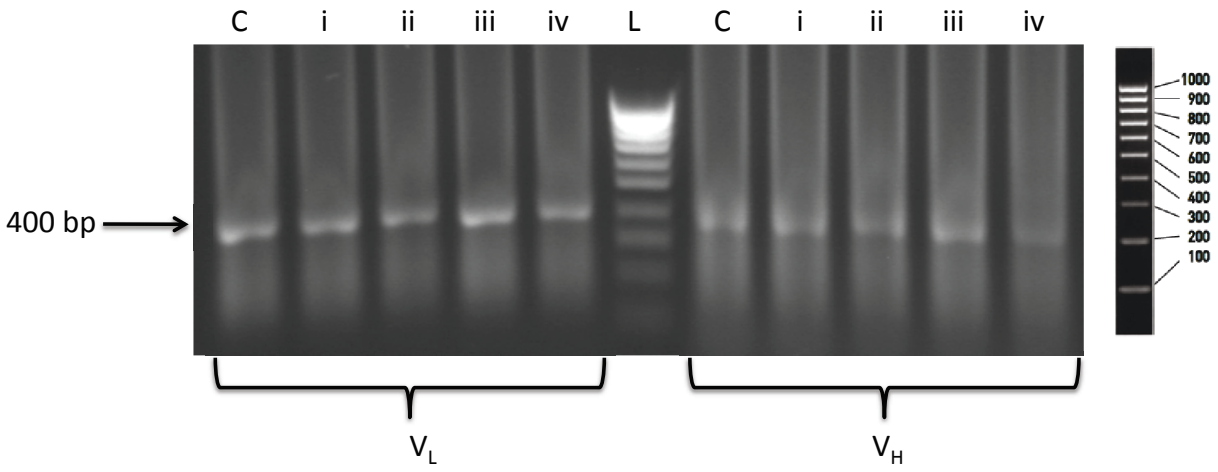


Figure 23. Analysis of the amplified variable regions V_L and V_H on 1.3% agarose gel after PCR2.

Distinct bands at ~400bp were noticed for both V_L and V_H in comparison with marker L (SmartLadder® SF, Eurogentec). The different lanes represent C: 9G4H9 cloned into pAk100 as control, i: naive Balb/C, ii: naive SJL/J, iii: immunized Balb/C, and iv: immunized SJL/J variable regions.

The random combination of variable region encoding sequences (V_L and V_H) was performed by splicing by overlap extension (SOE) PCR. The V_L3' and V_H5' primers contain sequences encoding an scFv linker required to assemble the variable regions together by complementary base pairing. The hybridization of V_L and V_H is therefore achieved by the overlapping of these sequences. The most commonly used linker for constructing an scFv is the linker type $(G_4S)_n$ (Huston *et al.*, 1993; Krebber *et al.*, 1997). We have chosen to use the linker $(G_4S)_4$ for the assembly of the two variable region sequences. This linker sequence is long enough to avoid the dimerization of scFv sequences. SOE-PCR is performed in two steps (Figure 24).

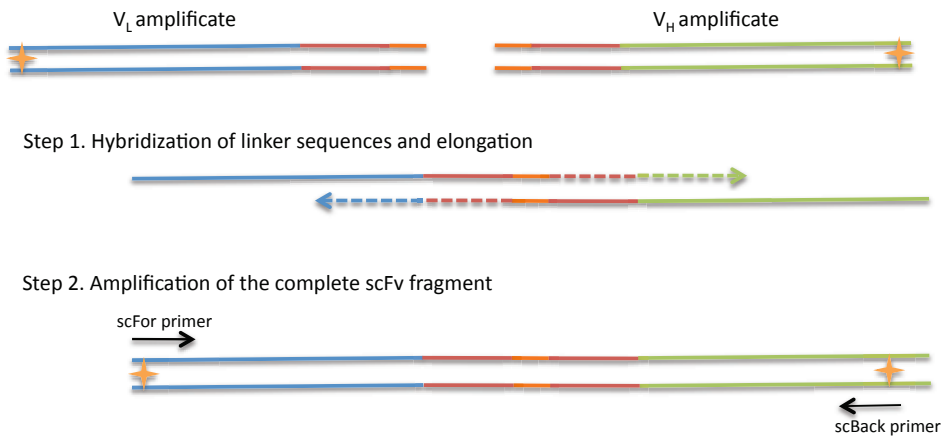


Figure 24. SOE-PCR allowing the assembly of V_L and V_H into scFv.

Blue and green segments represent V_L and V_H amplificates, respectively. The red/orange segments represent the linker sequence. The first step involves the denaturation and hybridization of the overlapping segments of the linker sequence indicated in orange, followed by the elongation of the single strands by Taq polymerase. The second step involves the amplification of scFv fragments by using sc primers. The stars indicate the SfiI restriction sites, which were inserted by previously used primers in PCR2.

The first step is to associate the V_L and V_H via the overlapping linker sequences at the 5' end of one DNA strand and the 3' end of the other. This step results in the formation of a double-stranded region at the center but with a single strand protruding towards each end. These single-stranded sequences are then filled in by the action of PWO polymerase enzyme by synthesizing the complementary strand using the 3' free OH. In the second step, a conventional PCR is performed using single chain (sc) primers, namely, scFor and scBack. These primers hybridize at the SfiI site inserted by the primers V_L5' and V_H3' and are phosphorylated in order to enable the circularization of scFv (required for the next step, described below). The SOE-PCR amplificates of ~800 bp representing the V_L-linker-V_H association were observed on agarose gel (Figure 25). The SOE-PCR results were purified and concentrated for further use.

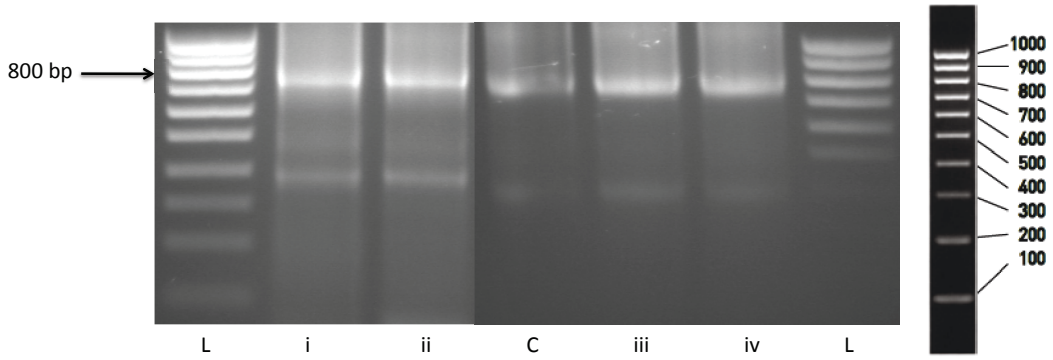


Figure 25. Assemby of V_L and V_H by SOE-PCR.

scFv fragments are observed at ~800 bp on 1.3% agarose gel. The different lanes represent C: 9G4H9 as control, i: naive Balb/C, ii: naive SJL/J, iii: immunized Balb/C, and iv: immunized SJL/J, and L: SmartLadder® SF (Eurogentec®) as reference.

In an attempt to increase the size of our libraries, to be able to have a better approximation of the mouse antibody repertoire, we have focused on improving the efficiency of the restriction digestion of the scFv fragments by using the technique of rolling circle amplification (RCA). Indeed, the digestion step of scFv fragments is a limiting step in library construction. This is due to the inefficiency of the digestion reaction due to the restriction sites being at the extremities of the scFv fragments. This means that the restriction enzymes do not have sufficient sequence length in order to bind to their sites. Using RCA to amplify scFv sequences in tandem repeats, places the restriction sites in the middle of the long sequence and therefore increases the efficiency of the digestion reaction. RCA is an isothermal DNA amplification procedure mediated by some DNA polymerases, which can generate a long single-stranded (ss) linear DNA molecule from a short circular ssDNA template by using a single DNA primer (Fire and Xu, 1995). This type of replication involves the generation of a site-specific nick by the plasmid encoded initiator (Rep) proteins at the plasmid leading strand origin, followed by covalent extension of the 3' OH end by DNA polymerase (Khan, 1997). Since its discovery in the 1990's, RCA has been used as a method for the amplification of small circular DNA templates (<200 nucleotides in size) with wide-ranging applications in genomics (Zhang *et al.*, 1998), proteomics (Cheng *et al.*, 2010), diagnostics (Ding *et al.*, 2012), biosensing (Banér *et al.*, 1998), drug discovery and nanotechnology (Zhao *et al.*, 2008). In multiply-primed RCA, the polymerization process is primed by exonuclease-resistant random hexamers that bind to multiple locations on the circular template DNA, generating multiple replication forks (Dean *et al.*, 2001). The method involves the phi29 DNA polymerase from bacteriophage, a stable, high-fidelity enzyme with high processivity and proofreading activity. By using multiply-primed RCA, circular DNA templates can be exponentially amplified up to 10^7-fold (Nelson *et al.*, 2002).

A circular template is required for the RCA reaction (Blanco *et al.*, 1989). To make a linear scFv into circular form, the scFv fragments are ligated onto themselves. The ligase enzyme must bind to the phosphodiester bond between the 5' and 3' sequences of the two double strands. Hence, it is necessary to use Sc primers (scFor and scBack) phosphorylated at their 5' end during SOE-PCR. Circularized scFv sequences act as templates for amplification by the phi29 polymerase enzyme. This polymerase produces a single stranded linear amplificate containing a number of repeats of the same sequences. Use of random hexameric primers allows the initiation of the amplification at several loci in the circular template sequence (Figure 26).

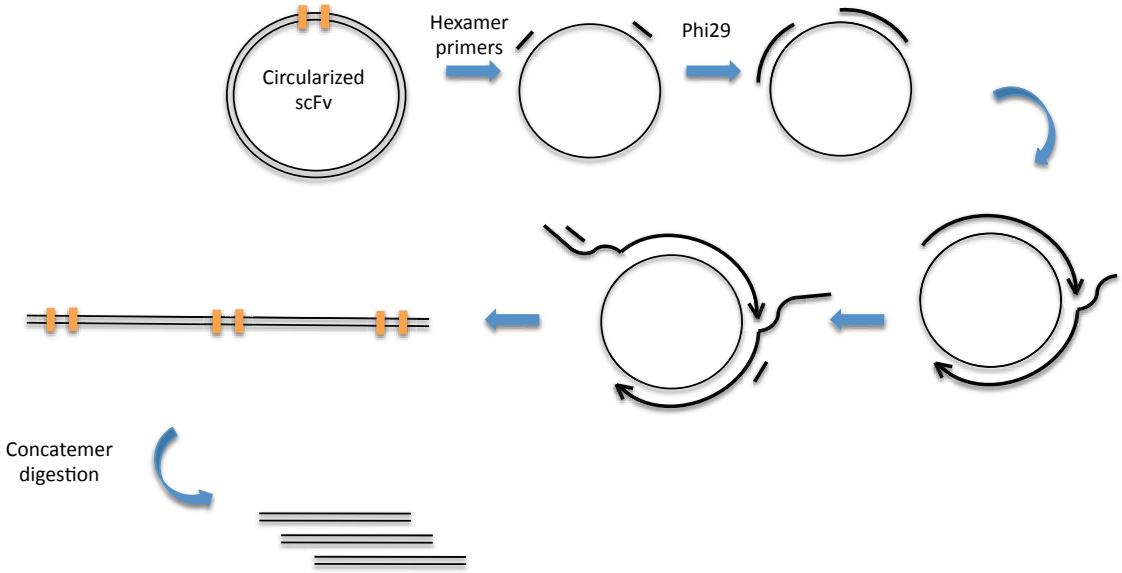


Figure 26. Rolling Circle Amplification of scFv fragments.

After denaturation of ds circularized scFv, random hexamer primers are annealed and phi29 polymerase amplifies the circular fragments into long linear concatemers. These extended molecules are then digested by the SfiI restriction enzyme into single scFv fragments. SfiI sites are represented by thick orange lines.

SfiI restriction enzyme is conventionally used in directional cloning of scFv sequences. SfiI restriction site is discontinuously organized in 2 series of 4 nucleotides separated by 5 random palindrome nucleotides (Figure 27). This feature allows avoiding the potential digestion of certain immunoglobulin sequences that could result in a loss of diversity. This enzyme has a unique characteristic where it attaches to two restriction sites to perform simultaneous cleavage (Wentzell *et al.*, 1995), therefore offering the advantage of orienting the cloning of recombinant genes in the vector despite the use of a unique enzyme. Moreover, the sequence of the SfiI restriction site is rarely epressed on genes encoding for immunoglobulins, making the SfiI site a well-suited enzyme for antibody cloning (Krebber *et al.*, 1997).

Figure 27. Sequence of the restriction site of the enzyme SfiI.

The arrows indicate where the enzyme cuts the DNA sequence.

The scFv fragments of each of the four libraries are finally recovered by efficient restriction digestion with SfiI enzyme. With this strategy, a higher percentage of the inserts should suitably be digested. Thus, the following ligation step into the phagemid is also expected to be more efficient. Indeed, the incorporation of RCA in the cloning process increases the size of the library by a factor of 10^3 (Shahsavarian *et al.*, 2014). Electrophoretic profiles of the

scFv fragments after restriction digestion by SfiI enzyme after RCA are as shown in figure 28. It is clearly observed that a crisp band at ~800 bp is the product of SfiI digestion on the scFv fragments. 9G4H9 cloned in pAK100 is used as a control. Digestion of this plasmid by SfiI leads to the release of 9G4H9 scFv (800 bp) and the linearized vector (4 kb).

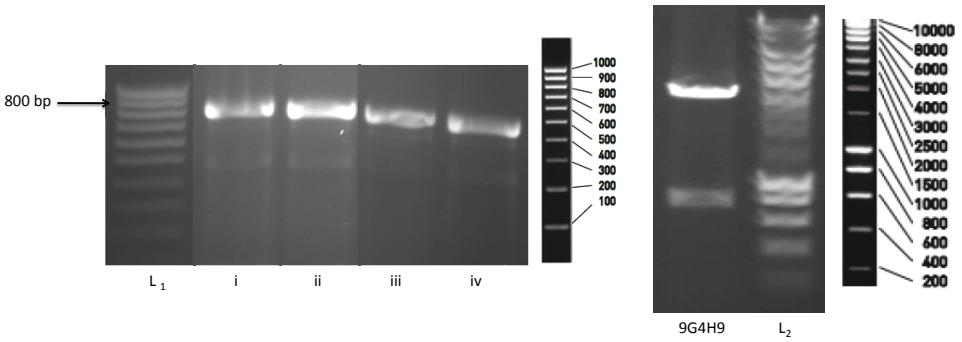


Figure 28. RCA results after SfiI digestion.

Restriction enzyme digestion of RCA products yielded bands at ~800bp. The different lanes represent i: naive Balb/C, ii: naive SJL/J, iii: immunized Balb/C, and iv: immunized SJL/J. For the 9G4H9 control, an additional band at ~4kb is observed due to linear pAK100 released by SfiI digestion. L_1: SmartLadder® SF (Eurogentec®) and $L_{2:}$ SmartLadder® 10kb (Eurogentec®) are used as reference.

I.4.C.c. Construction of uniquely "restriction bar-coded" host vectors and cloning

We have used a phagemid vector system for the construction of the libraries. The phagemids are based on the vector pAK100, which has been kindly given to us by Professor Andreas Plueckthun (Krebber, *et al.*, 1997). The components of this phagemid include the gene III leader peptide used to direct the expression of scFv at the surface of phage, and the chloramphenicol resistant gene used for the selection and maintenance of the phagemid (Figure 29).

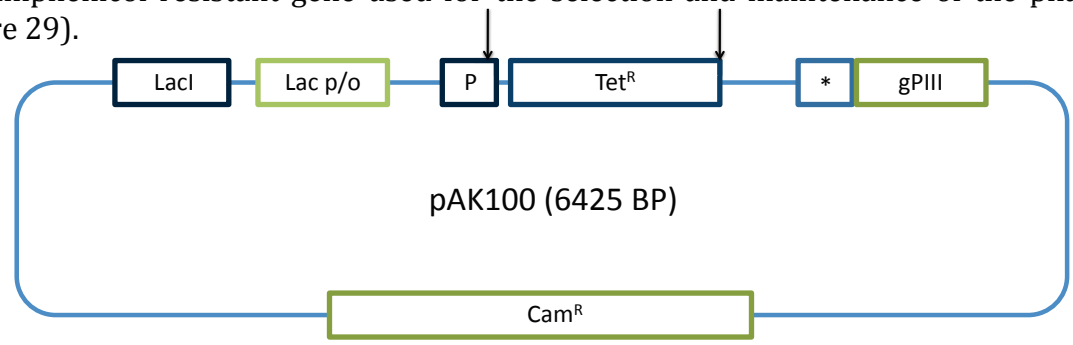


Figure 29. Schematic representation of phagemid pAK100.

LacI: inhibitor of the *lac* operon; Lacp/o: promoter and operator of the *lac* operon; P: pelB signal sequence for the periplasmic expression of proteins; Tet^R: resistance cassette for tetracycline; *: amber codon (UAG) in order to provide a switch between the expression of protein fused to pIII and the expression of the soluble form; gPIII: protein pIII gene; Cam^R: resistance cassette for chloramphenicol. The arrows indicate the position of SfiI sites. Source: inspired by Krebber *et al.*, 1997.

We have pooled the four libraries to produce a single large immune repertoire and simplify further selection steps. In order to be able to identify the library origin of each potentially selected clone in this pool, we have modified the vector pAK100 in order to have four slightly different but distinguishable phagemids to use for cloning the antibody genes. We have constructed 3 modified vectors (pAK101, pAK102, and pAK103) based on the wild type host vector pAK100. For each modified vector a single restriction site is eliminated via nucleotide deletion. The restriction site KpnI is eliminated for the phagemid pAK101, EcoRI for pAK102, and BstEII for pAK103. The naive Balb/C library is cloned into the host phagemid pAK101, the immunized Balb/C library into pAK102, the naive SJL/J library into pAK100, and finally the immunized SJL/J library into pAK103. With this methodology, the size of the modified phagemids remains intact, which prevents any undesired influence on the transformation step during library construction or on phage amplification. This strategy will simplify further selection procedures or any other manipulations of the antibody library, by allowing a unique selection on a single pooled library. This is achievable due to the possibility of distinguishing between the four vectors by a number of simple restriction digestion steps. Moreover, for a first screening, no sequencing is required. This is contrary to the insertion of nucleotide sequences in the vectors, commonly known as "bar-codes".

In this way, the eliminated restriction sites act as unique "restriction bar-codes", which can lead to the identification of the original phagemid and therefore the identification of the library to which a selected antibody belongs. The theoretical profiles of the restriction digestion of each phagemid with and without an antibody insert are shown in figure 30. The experimental profiles are shown with isolated clones from the pool of the four libraries, demonstrating an identifiable profile and revealing library origin (Figure 31). Three restriction digestion reactions are performed for each sample, using the enzyme HindIII in addition to one of the three enzymes corresponding to the eliminated restriction sites. In this way, the profile with a single DNA band reveals which of the restriction sites have been eliminated and therefore to which library the clone belongs.

For example, for a clone originating from the naive Balb/C library, for which the phagemid pAK101 with the eliminated KpnI restriction site has been used, the restriction digestion profile with the enzymes KpnI and HindIII will demonstrate a single band corresponding to the linearized phagemid. The restriction digestion of clones selected from the 3 other libraries using the same restriction enzymes will lead to 2 separate cuts in the phagemid and therefore to the observation of 2 distinct bands. If there are any ambiguities in the obtained restriction profiles, due to the presence of the restriction sites KpnI, EcoRI, or BstEII in the scFv sequence itself, sequencing of the scFv will reveal the presence of these restriction sites and lead to the clarification of the digestion profiles and therefore the identification of the library of origin.

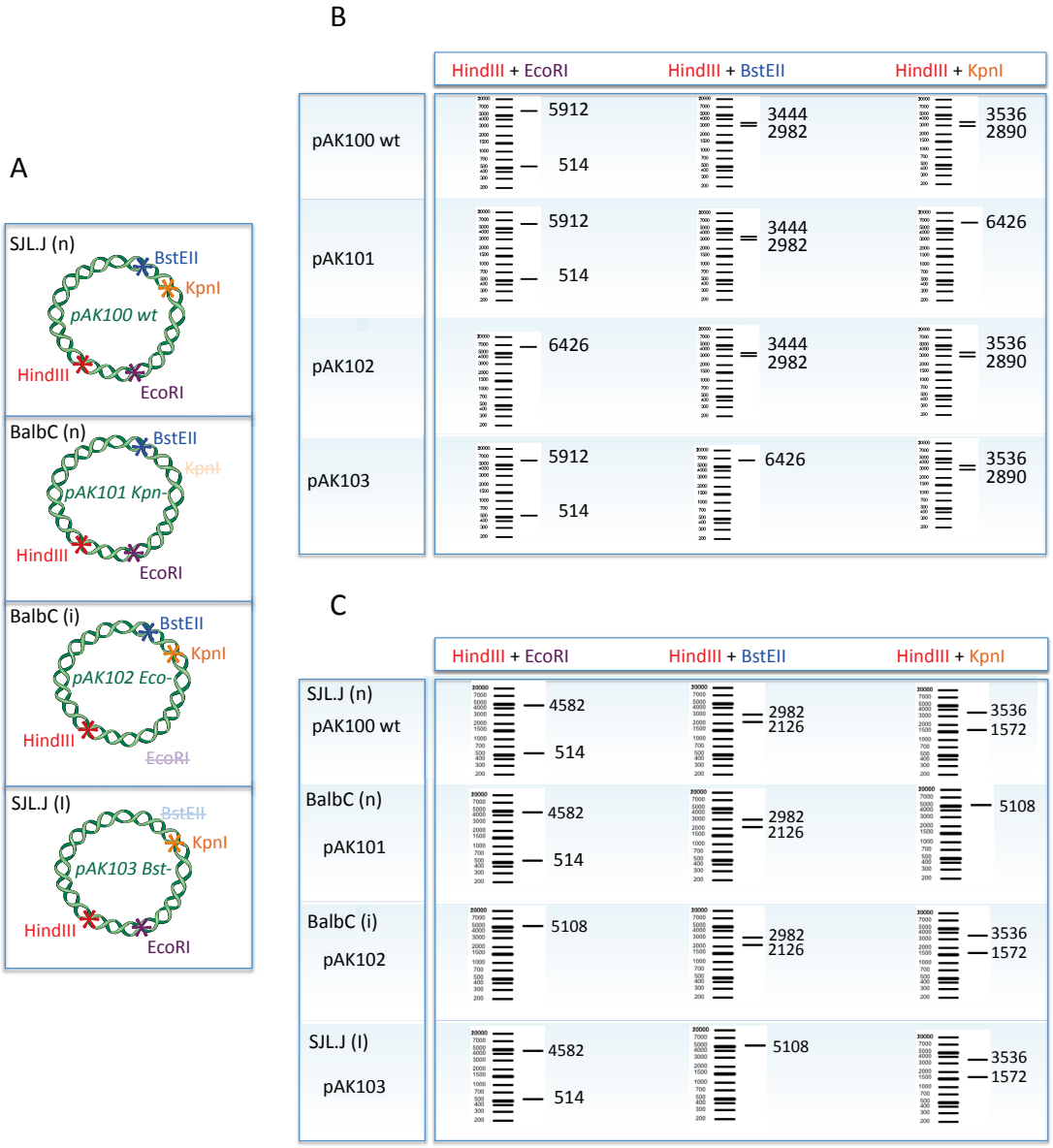


Figure 30. Phagemid vectors for library construction.

The four modified phagemid vectors A and their theoretical restriction digestion profiles without any insert B and with an scFv insert C. The eliminated restriction site for each of the vectors are the following: pAK101: KpnI, pAK102: EcoRI, and pAK103: BstEII. The phagemid pAK100 has all the restriction sites conserved.

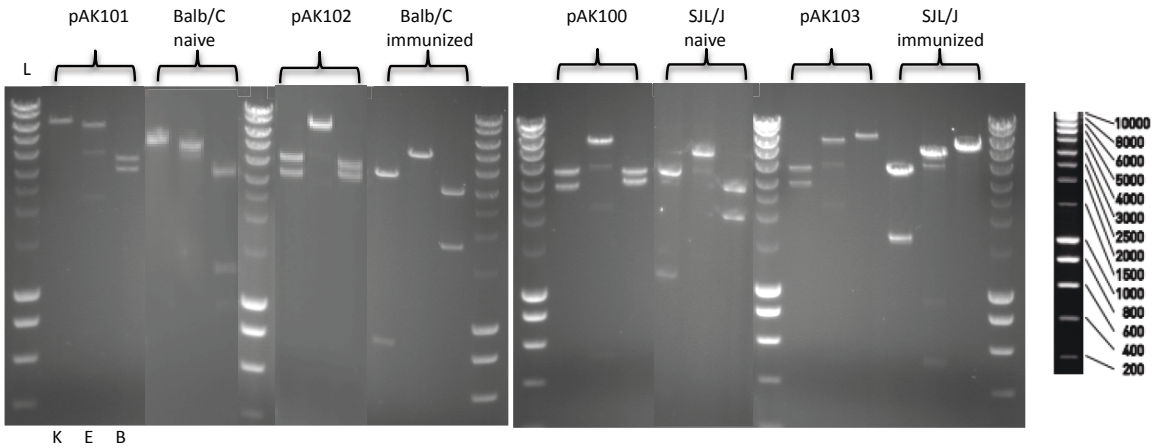


Figure 31. Experimental restriction digestion profiles of phagemid vectors and library clones.

Profiles include digestion of the four phagemid vectors without any insert and clones isolated from the pool of the four libraries demonstrating an identifiable profile of library origin. L: SmartLadder® 10kb (Eurogentec®). For each of the samples, the first lane corresponds to the digestion by enzymes HindIII and KpnI, the second lane corresponds to the digestion by enzymes HindIII and EcoRI, and the third lane corresponds to the digestion by enzymes HindIII and BstEII.

For the cloning procedure, the vectors need to be digested by the SfiI enzyme in order to have complementary ends to the previously SfiI-digested scFv fragments. Thus, the insert will replace the teracyclin resistance marker (Figure 29). Before digesting the phagemids with SfiI, however, we first treat the vectors to a process of multi-digestion by the restriction enzymes SacII, MfeI, NspI, and EcoNI. All of these restriction sites are located in the resistance cassette for tetracyclin. This process eliminates the tetracyclin resistance marker located between the two SfiI restriction sites, which will ensure that the vector will not re-ligate back onto itself. The Tet^R resistance cassette is approximately 2 kbp, hence the difference of size observed between the wild type and recombinant phagemid vectors (without the 2kbp from the resistance marker fragment but containing the 800 bp from the scFv fragment) is about 1200 bp. The vectors and scFv inserts are then ligated, transformed successfully into freshly prepared electrocompetent *E. coli* bacteria, and stored.

I.4.C.d. Library size, diversity and functionality

The size of each of the 4 libraries is shown in table 4. The combined pool of the 4 libraries leads to the creation of a large repertoire of size 2.68×10^9.

Table 5. Size of each of the 4 libraries.

Library	Size
Balb/C (naive)	4.8×10^8
Balb/C (immunized)	1×10^9
SJL/J (naive)	4×10^8
SJL/J (immunized)	8×10^8
Pooled library	2.7×10^9

The diversity of antibody libraries, according to the literature, is conventionally confirmed by obtaining about 20 unique sequences of individual clones (Kay *et al.*, 1996). Further investigation of the quality of the libraries can be performed by sequencing up to several hundred clones (Ravn *et al.*, 2010). We have analyzed a population of 400 sequences from the pooled library (a window of 100 sequences per library) *via* conventional Sanger sequencing. The determination of the percentage of identity of all sequences *via* the ClustalW2 tool (Larkin *et al.*, 2007) has demonstrated that for the immunized Balb/C library, two V_L sequences as well as two of the V_H sequences have been observed twice. Similarly, one V_H sequence of the naive SJL/J library and one V_H sequence of the immunized SJL/J library have been observed twice (data not shown). Looking at the pairing of V_L and V_H, however, all of the scFv sequences for the four libraries are unique. These results have therefore confirmed the diversity of the libraries. Considering each immunoglobulin gene segment (IGKV, IGKJ, IGHV, IGHJ, and IGHD) separately, we have obtained and analyzed a total number of 1934 gene sequences. Within this population, we have found 10 pseudogenes (0.52%), which is in accord with findings of other teams (Glanville *et al.*, 2009). The presence of pseudogenes in our library could be explained by the fact that even though we have designed our primers based on "functional" sequences reported in IMGT, some primers could have a high enough homology to pseudogene sequences to be able to hybridize them in the PCR conditions used.

We have confirmed the functionality of the pooled library by a Western Blot analysis. We look at the association of the M13 phage coat protein pIII to the scFv antibody by comparing the size of the wild-type phage (pAK100) to the library phage displaying the scFvs (Figure 32). The results show a difference of 30kDa between the two samples corresponding to the correct fusion of the scFv fragments to the phage coat protein pIII.

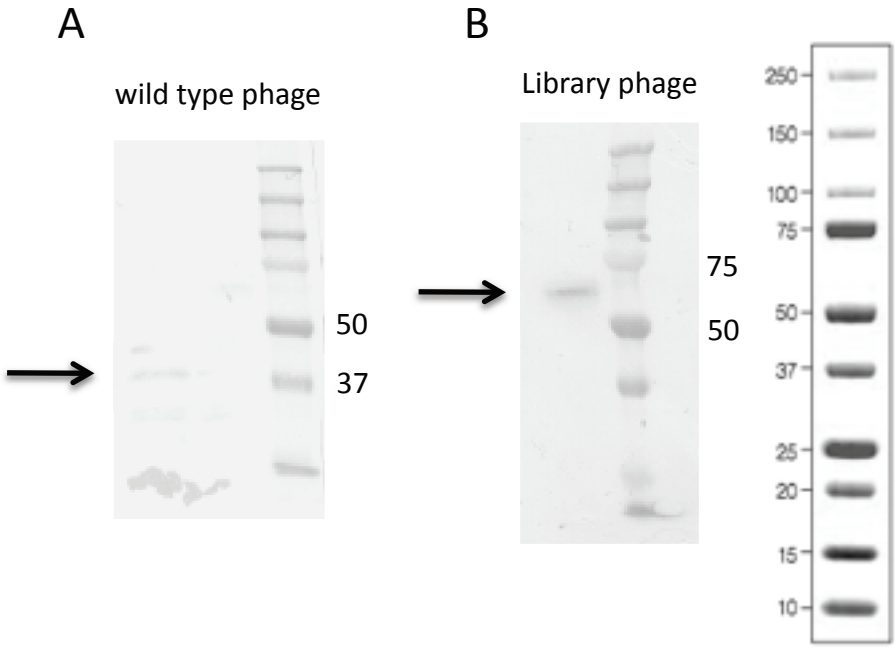


Figure 32. Western Blot analysis of the library.

Revelation was performed using an anti-mouse antibody conjugated to HRP as secondary antibodies. A) The wild type phage pAK100 (5×10^{11} particles) with no insert. B) Phage from a pooled sample of the 4 libraries (5×10^{11} particles). Arrows indicate the specific revelation of pIII protein alone or fused to scFv fragments. Individual molecular weight of pIII and scFv are expected at 70 and 30 kDa, respectively. The ladder is the Precision Plus Protein Standard (BioRad).

We have further analyzed the percentage of successful cloning by performing PCR on colony on 24 randomly selected individual clones from each of the four libraries (Figure 33). The results confirm that all of the screened clones have an inserted fragment, but that the sequences are of different size (400-1000 base pairs), corresponding to single variable chain or complete scFv fragments. The cloning of only one of the two variable chains has already been observed in literature (Taguchi *et al.*, 2008c). We consider that both complete and incomplete scFv sequences contribute to the diversity of our libraries. We have shown that 100% of the screened samples have been successfully cloned, where the percentage of clones bearing a complete scFv fragment (>800 bp) is 96% for the naive libraries Balb/C and SJL/J, and 88% for the immunized libraries Balb/C and SJL/J.

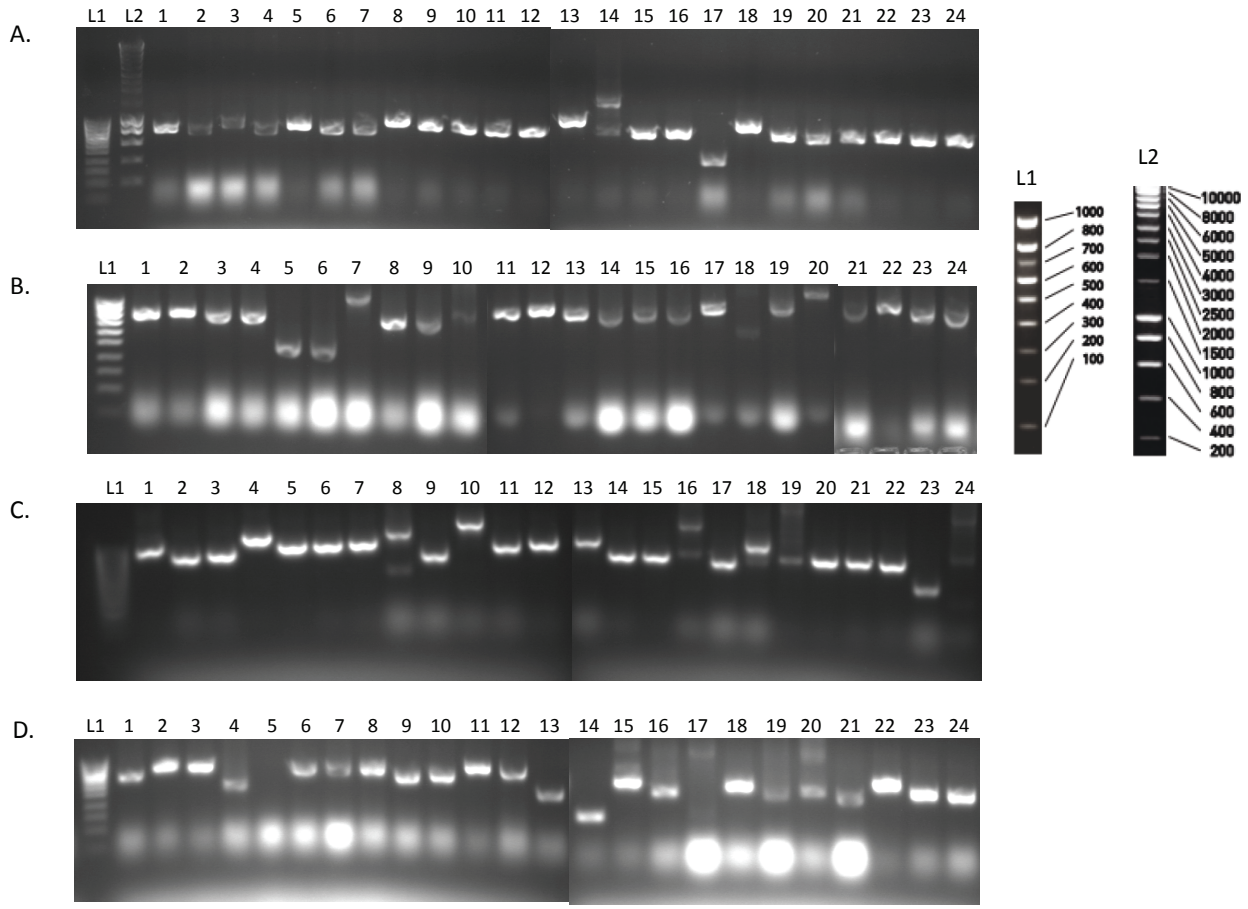


Figure 33. PCR on colony on 24 randomly selected individual clones for each of the 4 libraries.

A) Balb/C naive library. B) Balb/C immunized library. C) SJL/J naive library. D) SJL/J immunized library. L1: SmartLadder® SF (Eurogentec). L2: SmartLadder® 10kb (Eurogentec).

Once the functionality of the library has been confirmed, we go on to analyze its diversity. In parallel to the sequence analysis of individual clones in order to validate the diversity, we have performed a study of the potential genetic differences existing between the four immune repertoires.

I.5. Analysis of immune repertoires

A previous analysis performed by our group has focused on the immunoglobulin κ gene subgroups and their reported frequencies in the IMGT® database (IMGT®, the international ImMunoGeneTics information system® http://www.imgt.org), using the GeneFrequency tool (Le Minoux *et al.*, 2012; Lefranc *et al.*, 2009). This tool is only available for the light chain κ. We have extended this analysis to include all reported gene subgroups of light and heavy V, D, and J regions, this time using the GeneDB tool (Giudicelli *et al.*,

2005). We have, therefore, calculated the percentage of the occurrence of each of the gene subgroups IGKV, IGKJ, IGHV, IGHD, and IGHJ reported in the IMGT® database. These results demonstrate that for the light chain κ, the subgroups IGKV7, IGKV11, IGKV13, IGKV15, IGKV17, IGKV19, and IGKJ3 are reported with a frequency of less than 1%. Similarly, for the heavy chain γ, the subgroups IGHV4, IGHV8, IGHV9, IGHV12, IGHV13, IGHV15, IGHV16, and IGHD5 are reported with a frequency of less than 1% (Figures 34 and 35). It is important to note that this analysis of the sequences reported in the IMGT database was performed in April of 2014, at which time the 400 sequences from our library were submitted for the analysis described below. However, a more recent look at the sequences available in IMGT, in July of 2015, has demonstrated slightly different results. Firstly, the total number of sequences available for each of the gene segments has changed (for example for IGKV, there are 1587 sequences in 2015, whereas in 2014 this number was 1571). Moreover, there are a number of gene subgroups, i.e. IGKV15, IGKV16, IGKJ3, and IGHD5, which are no longer presented in IMGT. These sequences have possibly been placed under a different gene subgroup or they have potentially been completely removed from the database. The modified data is presented in the Annexes, Section V.

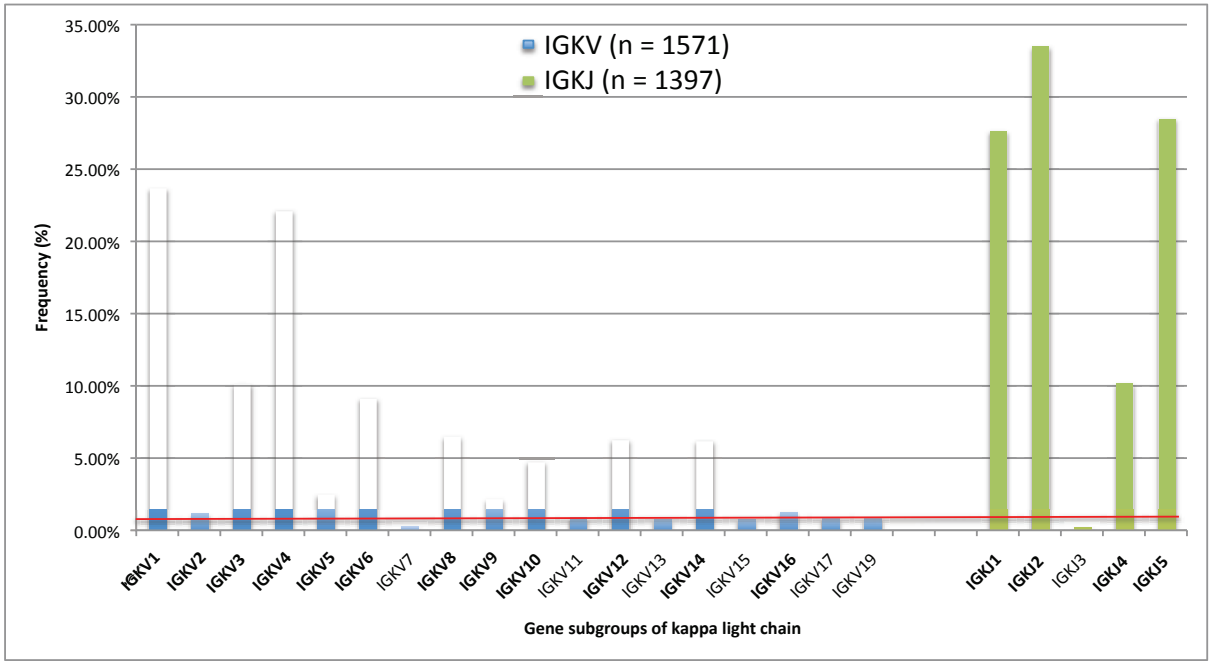


Figure 34. Frequencies of the light chain κ subgroups reported in IMGT.

The red line represents a frequency of 1%. The gene subgroups with a frequency of more than 1% are in bold.

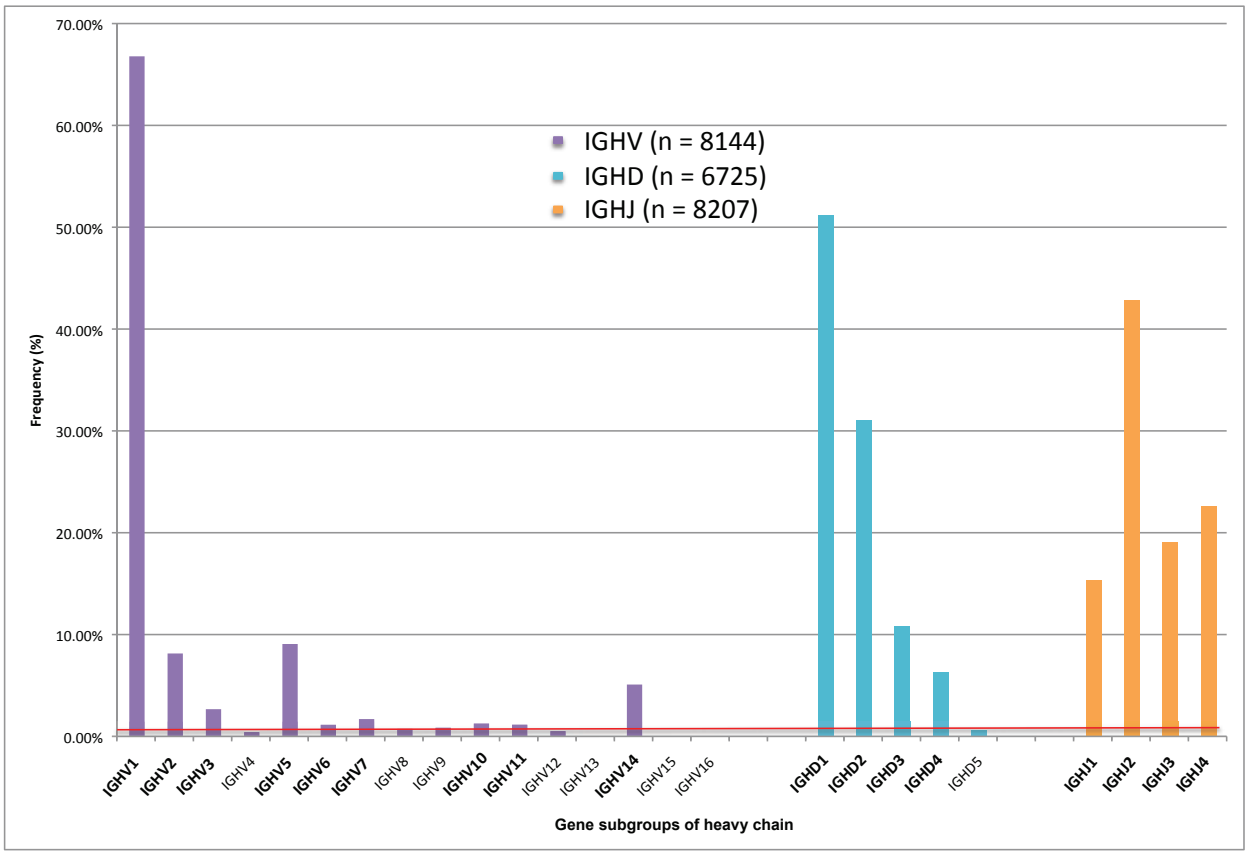


Figure 35. Frequencies of the heavy chain γ subgroups reported in IMGT.

The red line represents a frequency of 1%. The gene subgroups with a frequency of more than 1% are in bold.

These discrepancies illustrate the dynamic nature of IMGT, which is regularly updated as a function of submissions of new sequences to the database. As evidenced by this analysis, there are also sequences that are removed or displaced from certain subgroups (i.e. sequences belonging to IGHD are fewer in number considering the data from 2015 as compared to 2014.

In order to investigate the extent of the representation of different immunoglobulin genes present in the four libraries, we have performed a similar analysis, using the IMGT®/V-QUEST tool (Giudicelli *et al.*, 2011) for the determination of gene subgroups of the 400 sequenced scFv fragments. It is important to note, once again, that this analysis was done in 2014, and therefore the comparisons with IMGT takes into account the subgroup representation frequencies present in this database at that time. First, we note the presence of different gene subgroups, which demonstrates the high level of diversity of the libraries. The represented genes include 11 out of the 19 IGKV subgroups, 4 out of the 5 IGKJ subgroups, 7 out of the 15 IGHV subgroups, and finally all of the IGHD and IGHJ subgroups. Taken together, 32 out of the 49 gene subgroups referenced in IMGT® are represented in the 400 sequences analyzed (65.3%). Second, if we take into account the overall low representation of the gene subgroups with frequencies of less than 1% (15 out

of 49), and the fact that we are focusing on a window of 100 sequences per library, we can expect the representation of only 34 gene subgroups. This means that 91.1% of the gene subgroups expected to be present in our small sample size are represented.

Despite the high level of diversity of the four libraries, we have not observed any λ gene subgroups in the pool of 400 analyzed sequences. This could either be due to the failure of our designed primer set to hybridize and amplify immunoglobulin lambda genes, or it could mean that the actual ratio of κ:λ immunoglobulin light chains in mice are in fact lower than 95:5, in line with the observation of several other groups (Dildrop *et al.*, 1987; Zocher *et al.*, 1995; and Almagro *et al.*, 1998). In order to examine this, we have performed a similar analysis on the reported lambda gene subgroups available in the IMGT GeneDB database (Figure 36).

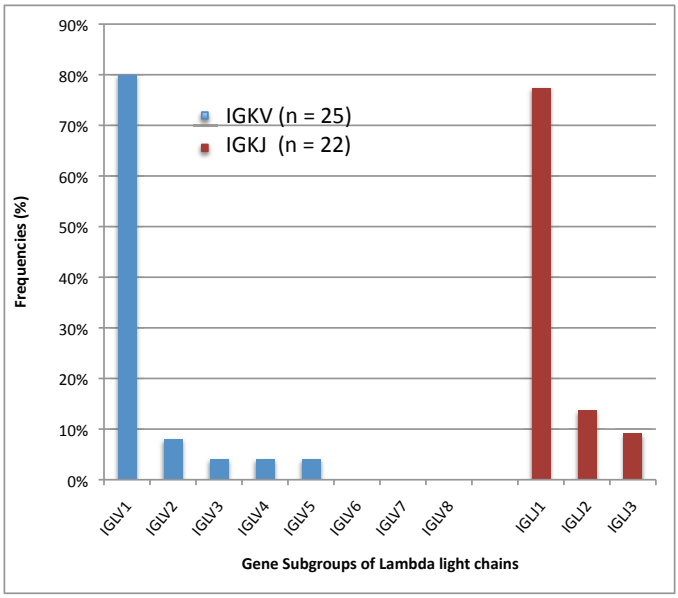


Figure 36. Frequencies of the light chain λ subgroups reported in IMGT in July, 2015.

The total number of sequences reported for IGLV is 25, and for IGKJ is 22.

The results indicate that indeed the ratio of IGKV:IGLV is 98.45:1.55 and the ratio of IGKJ:IGLJ is 98.39:1.61. Considering these ratios, we should still observe at least one lambda gene per library. We have, therefore, performed an experiment in order to ensure the capability of our primer set to amplify murine lambda genes. We have first used only the lambda primers present in the mix, namely LPL1, LPL2, and IgL for PCR1, and OVL and VL3'C for PCR2, and performed a PCR amplification similar to the strategy described in section I.4.C.b. on a previously prepared cDNA sample of immunized Balb/C mice. Results show that the lambda primers isolated from the mix are indeed capable of amplifying immunoglobulin lambda genes (Figure 37).

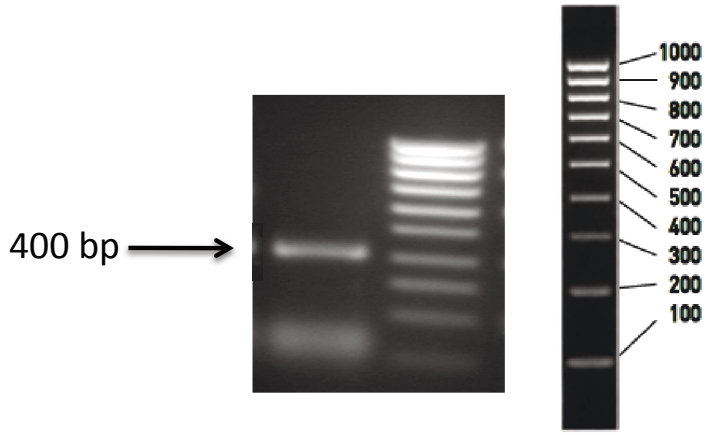


Figure 37. Analysis of the amplified V_L λ of immunized Balb/C mice on 1.3% agarose gel after PCR2.

A distinct band at ~400bp was noticed in comparison with the marker (SmartLadder® SF, Eurogentec).

We have sequenced seven isolated clones and determined that all are of the subgroups IGLV1 and IGLJ1. This is not surprising, considering the high representatively (~80%) of these subgroups and our small sample size. In order to confirm the capability of the totality of our primer mix to target and amplify lambda sequences, we have cloned 10 isolated sequences into pGEMT vectors and have performed a PCR on colony using the primer mix of PCR2 (i.e. including all the primers previously used for the amplification of κ and λ genes) (Figure 38). The results demonstrate successful amplification for 8 out of the 10 clones, illustrating the ability of the PCR2 primer mix to target and amplify immunoglobulin lambda sequences. We have not tested the complete PCR1 primer mix, however, given the similarity of design conditions and the fact that the lambda primers isolated from the mix are able to successfully target lambda sequences, we can be fairly certain of the absence of any biases in these primers.

Figure 38. Analysis of PCR on colony results on 1.3% agarose gel of 10 pGEMT clones containing λ sequences using the PCR2 primer mix.

A distinct band at ~600 bp was observed in comparison with the marker (SmartLadder® SF, Eurogentec).

We have compared the level of representation of different immunoglobulin gene subgroups between the 4 libraries in order to investigate whether the genetic background and/or immunological status of mice influence the immunoglobulin gene expression pattern. The expression profiles for V_L and V_H for each of the 4 libraries are shown in figures 39 and 40, respectively. A statistical test developed in collaboration with Pr. Nikolaos Limnios has allowed us to validate the significance of our observations (detailed in the Materials and Methods section of the Annexes).

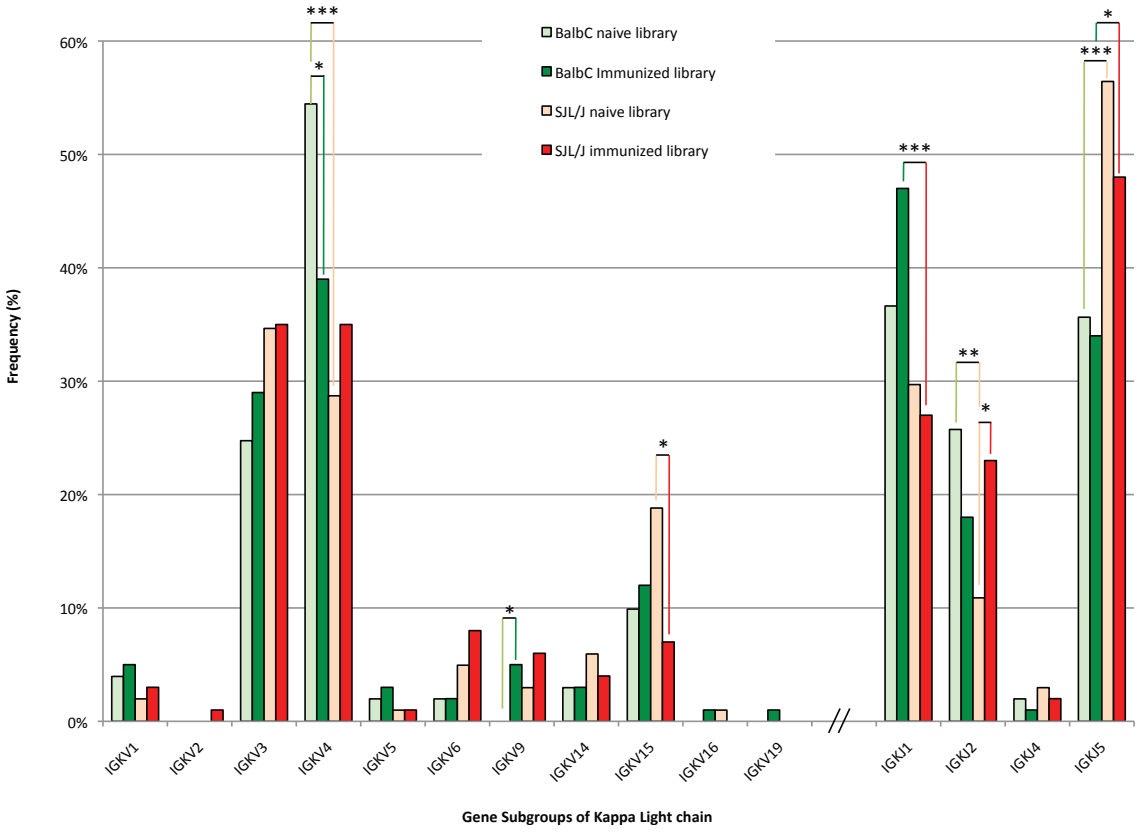


Figure 39. Immunoglobulin gene subgroup distribution of Kappa light chain.

Only subgroups represented in the libraries are shown. Statistical differences with $p<0.05$ are represented by *, $p<0.01$ by **, and $p<0.005$ by ***.

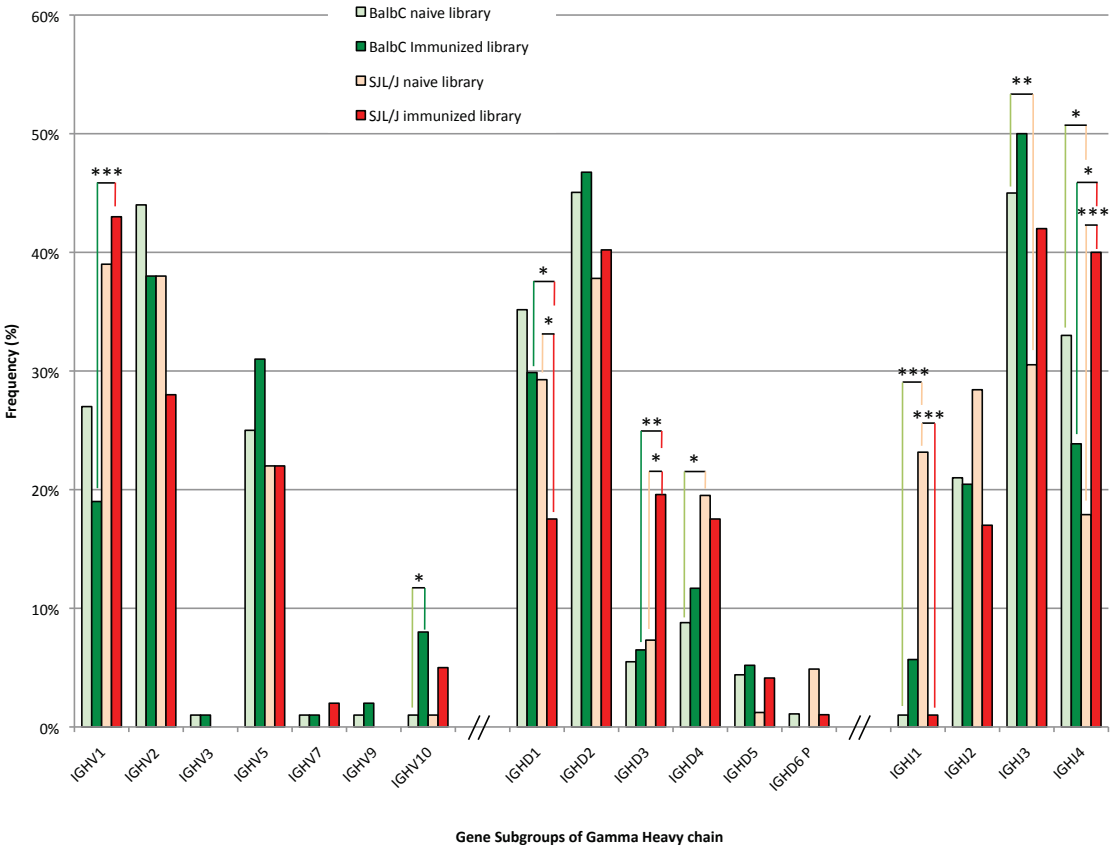


Figure 40. Immunoglobulin gene subgroup distribution of Gamma heavy chain.

Only subgroups represented in the libraries are shown. Statistical differences with $p<0.05$ are represented by *, $p<0.01$ by **, and $p<0.005$ by ***.

Our results show that there are significant ($p>0.05$-0.005) differences in the expression frequencies of a number of gene subgroups. Under the effect of immunization, certain gene subgroups are significantly over-expressed, while others are significantly under-expressed. We have observed that immunization leads to an over-expression of IGKV9, IGKJ2, IGHV10, IGHD3, and IGHJ4, and an under-expression of IGKV4, IGKV15, IGHD1, and IGHJ1. We have also observed that there are significant differences in the immunoglobulin gene expression profiles between the two strains, which have different immunological backgrounds but the same status of immunization. The subgroups IGKV4, IGKJ2, IGHJ3, and IGHJ4 have a significantly higher percentage of representation in naive Balb/C mice compared to naive SJL/J mice. The subgroups IGKJ1 and IGHD1 are over-represented in immunized Balb/C mice compared to immunized SJL/J mice. The subgroups IGKJ5, IGHD4, and IGHJ1 are under-represented in naive Balb/C mice compared to naive SJL/J mice. And finally, the subgroups IGKJ5, IGHV1, IGHD3, and IGHJ4 have a lower percentage of representation in immunized Balb/C mice compared to immunized SJL/J mice. Our results also indicate that despite their lower diversity, the observed differences are more frequent in J (light chain κ)

and D/J (heavy chain γ) segments compared to V sequences. It is important to note, however, that these differences are observed on a very small sampling size compared to the *in vivo* immune repertoire, and that we do not imply that these profiles and related differences are a general rule for the models under study. The observed differences, however, lead to an increase in the overall diversity of the pooled library, which provides us with improved chances of selecting antibodies with the desired characteristics.

Another observation worthy of mention, resulting from this analysis is the representation of the subgroup IGHD6 in the pool of 400 sequences, notably in the strain SJL/J. According to IMGT, this subgroup belongs only to pseudogenes, and in the GeneDB database, there are no sequences belonging to this subgroup. Whether this discrepancy results from a bias in our primer mix urging it to target pseudogenes or whether there are functional genes that belong to this subgroup remains to be discovered.

With the objective of enhancing our fundamental understanding of the nature of catalytic antibodies, and having to our disposal a largely diverse antibody library containing numerous variable region possibilities, we want to continue with the selection of potential catalytic antibody candidates and their characterization. As mentioned previously in chapter one, with our group's expertise and the availability of an enzyme suicide inhibitor, using β-lactamase as a model activity provides us with several advantages. Hence, we want to exploit the pool of the four scFv libraries, taking advantage of its enlarged diversity, to search for and isolate such antibody candidates. Before describing our strategy and results for achieving this objective, the following chapter gives a general overview on catalytic antibodies and the β-lactamase enzyme.

II. Catalytic Antibodies with β-lactamase Activity

II. CATALYTIC ANTIBODIES WITH β-LACTAMASE ACTIVITY

The first section in this chapter focuses on a description of catalytic antibodies, also known as abzymes, beginning with the hypothesis of their existence to their experimental production in the laboratory and the discovery of their natural existence in both physiology and disease. The section provides a framework on what has been observed about the nature and characteristics of catalytic antibodies, knowing that no concrete conclusions about general rules on the subject have been drawn to date.

The second section focuses on one of the activities displayed by a number of previously reported catalytic antibodies, namely the β-lactamase activity. This section provides a description of the natural enzyme and its role in the development of bacterial antibiotic resistance, a major concern in the medical field today. Then it goes over the catalytic mechanism performed by the enzyme and the residues involved in the catalytic pocket. Finally, the section describes a catalytic antibody endowed with β-lactamase activity, which has been produced as a result of a pioneer work by our group in the 1990s.

The final section in this chapter illustrates the results of the present work in the selection and characterization of catalytic antibodies endowed with this model activity.

II.1. Catalytic Antibodies

In 1948, Linus Pauling introduced the postulate that enzymes are capable of catalyzing a chemical reaction by stabilizing the transition state of the conversion of substrates into products (Pauling, 1948). He also proposed the idea that with the immense diversity of the immunoglobulin repertoire, there is the possibility that some antibody variable regions mimic the active sites of enzymes. Hence, if these antibodies are able to stabilize the transition state of the reaction, then they can also act as catalysts. Following the postulates of Pauling, William Jencks proposed in 1969 that antibodies generated against a transition state analog (TSA) of an enzymatic reaction, a stable compound which mimics the structure of the transition state, could potentially be endowed with the same enzymatic activity (Jencks, 1969). It wasn't until 1986, after the development of the hybridoma technology (Kohler and Milstein, 1975), that this hypothesis became a reality when two teams, independently of each other, simultaneously reported the first catalytic antibodies. They were the groups of Richard Lerner and Peter Shultz, who successfully used TSAs as haptens to produce and characterize monoclonal antibodies capable of hydrolyzing aryl esters and carbonates, respectively (Tramontano *et al.*, 1986; Pollack *et al.*, 1986). These discoveries began an exciting new field of research in biocatalysis focused on the production of catalytic antibodies, or abzymes, capable of catalyzing numerous enzymatic reactions.

II.1.A. Strategies for the production of catalytic antibodies

Since 1986, many more antibodies have been generated being able to catalyze a large spectrum of enzymatic activities, some of which do not exist in nature (Xu *et al.*, 2004; Rao and Wootla, 2007). The methodologies for acquiring catalytic antibodies also have since been evolved. The different strategies have incorporated the fields of chemistry, biochemistry, immunology, and molecular biology, in order to produce the best candidates of abzymes. These strategies can be categorized into two different groups: *(i)* chemical strategies, and *(ii)* biological strategies.

II.1.A.a. Chemical strategies

The different strategies under this category involve the rational design and chemical synthesis of haptens, in order to induce the production, through immunization, of antibodies capable of reactively interacting with the hapten and facilitating its transition through a chemical reaction. In order to provide sufficient immunogenecity, this hapten is then conjugated to a carrier protein, such as bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH), and used for the immunization of animals. After the immunization, the identification of catalytic clones is performed by screening methods in multiple ways: *(i)* by performing inhibition tests to show that the hapten inhibits catalysis in a competitive fashion, and *(ii)* by demonstrating that it binds the antibody with a higher affinity than the corresponding ground state substrate, and finally *(iii)* by measuring the kinetic parameters. The following are brief descriptions of the current existing chemical strategies for the production of abzymes.

II.1.A.a.i. Transition State Analogs (TSA)

This technique has historically been the most widely used strategy in the elicitation of catalytic antibodies. It involves the design and preparation of the antigenic species as the transition state analog of the targeted reaction. The transition state is defined as a short-lived theoretical species, which exists only at the energetic peak in a reaction pathway. As it is practically impossible to isolate such a molecule, a stable analog of the transition state, mimicking this theoretical structure and geometry as close as possible, is often used. Just as enzymes catalyze a reaction by lowering the energy barrier of the transition state (Figure 41), an antibody raised against a TSA should bind this transition more tightly than the substrate or products and thus should also accelerate the formation of products.

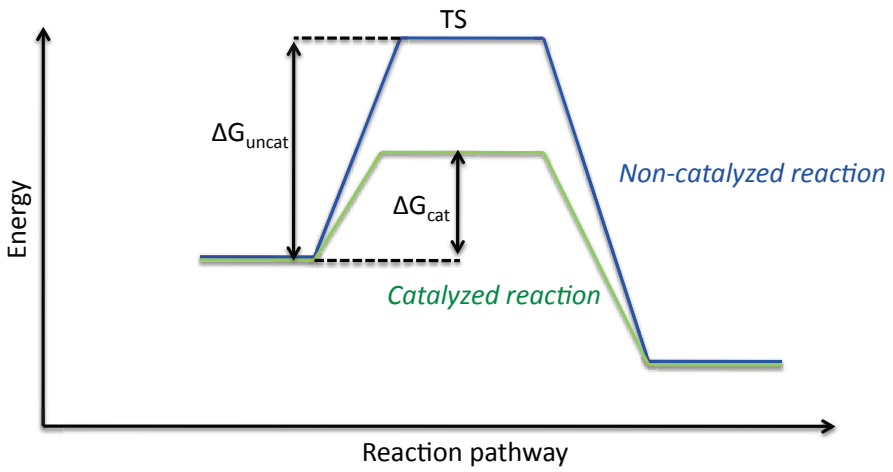


Figure 41. Diagram representing energy states along a reaction pathway with or without the presence of a catalyst.

TS: transition state.

The majority of the antibodies generated with this strategy to date catalyze hydrolytic reactions of acyl species. Antibodies with amidase and esterase activities have previously been reviewed (Tanaka, 2002). The amidase activity is particularly difficult to target by the TSA approach, due to limitations of an amine as a leaving group. Examples of such abzymes are antibody 43C9 and antibody 312D6 (Gao *et al.*, 1994; Aggarwal *et al.*, 2003). A second interesting class of antibodies produced by TSA immunization is the antibodies generated by the group of Peter Shultz, which hydrolyze unactivated alkyl esters in a stereospecific manner (Pollack *et al.*, 1989). Catalytic antibodies have also been elicited by this method to catalyze Diels-Alder reactions, which is of particular interest to chemists because it proceeds via an entropically disfavored transition state and its existence in nature has not been confirmed. A few examples of these abzymes are antibody 1E9 (Hilvert *et al.*, 1989), antibody 39-A11 (Braisted *et al.*, 1990), and antibody 13G5 (Yli-Kauhaluoma, *et al.*, 1995).

There are many other reactions performed by antibodies produced by the TSA approach, however, as discussed later in section II.1.A.b.iii, the catalytic parameters of such abzymes remain much lower than their natural enzyme counterparts. The design of better TSAs for the elicitation of antibodies with higher catalytic turnovers suitable for various applications remains one of the major challenges of the catalytic antibody field today. There are a number of modifications of this approach developed to date, which incorporate different strategies. These are described below.

II.1.A.a.i.i. Bait and switch

In this strategy, a hapten is designed carrying a point charge, in order to recruit antibodies with a complementary charged amino acid in their active site, leading to catalysis (Janda *et al.*, 1990). The point charge is placed in close proximity or in direct substitution of a chemical functional group expected to transform the substrate. The complementary

charged amino acid residues then are able to catalyze the reaction as general acid/base or nucleophilic catalysts. The first antibody generated by this strategy was antibody 43D4-3D12 performing β-elimination (Shokat *et al.*, 1989). Among other antibodies elicited with this strategy are antibody MATT.F-1 able to hydrolyze DNA or RNA phosphodiester bonds and antibody HA8-25A10 capable of catalyzing a cationic cyclization reaction (Wentworth *et al.*, 1998; Hasserodt *et al.*, 2000).

II.1.A.a.i.ii. Entropic trapping

Control of translational and rotational entropy of a reaction's transition state is a key factor of catalysis. This control provides the means for a catalyst to position the substrate in proximity with the catalytic residues and configure it in such a way to favor the interactions leading to the transformation of the substrate to products (Petsko and Ringe, 2004). In this way, antibodies should be able to catalyze reactions with unfavorable activation entropies, by acting as "entropy traps", providing the necessary favorable configuration. Examples of antibodies generated by this method are antibodies that catalyze the Claisen rearrangement, involved in the conversion of chorismate to prephenate in the bacterial synthesis of aromatic amino acids (Hilvert *et al.*, 1988), and antibodies catalyzing a Diels-Alder reaction (Romesberg *et al.*, 1998).

II.1.A.a.i.iii. Reactive immunization

Most of the hapten design strategies described above make use of chemically inert molecules. Reactive immunization, however, as evident by the term, uses reactive immunogens to induce the formation of catalytic antibodies. Richard Lerner and Kim Janda pioneered this method by using a highly reactive organophosphorus diester hapten, which underwent a chemical reaction in the antibody combining site during immunization (Wirsching *et al.*, 1995). This led to the elicitation of antibodies with the target acyl transferase activity reaching k_{cat} of 31 min^{-1}. Later, this strategy was also used to generate antibodies with aldolase activity (Wagner *et al.*, 1995; Barbas *et al.*, 1997; Zhong *et al.*, 1999).

Mechanism based inhibitors have proven to be promising candidates in the design of reactive immunogens, due to their ability to covalently react with the active site in target proteins and inhibit their activity (Silverman, 1988). Since these inhibitors form covalent adducts with the enzymes that process them along a defined reaction pathway, they should allow for the direct selection of catalytic antibodies that utilize particular features of a designed mechanism. In this way, the irreversible covalent binding aspect allows for kinetic selection, favoring antibody variable sites that accelerate bond formation by the stabilization of the transition state. The innate reactivity of such molecules can, therefore, be used to elicit highly evolved biocatalysts that conserve an essential residue for covalent catalysis (Soumillion and Fastrez, 2001)

Richard Lerner and Carlos Barbas combined this knowledge with the technique of reactive immunization to target the hydrolysis of β-lactam rings (Tanaka *et al.*, 1999). They used the structure of penam sulfones, previously shown to be potent mechanism-based inhibitors of the enzyme β-lactamase (Fisher *et al.*, 1981; Soumillion *et al.*, 1994) by forming an acyl-enzyme intermediate, for the hapten design (Figure 42). In this project, a phage display scFv library of immunized mice with the hapten-KLH immunogen led to the generation of two antibodies, FT6 and FT12, which hydrolyzed the substrate ampicillin with k_{cat}/k_{uncat} values of 5200 and 320, respectively. Furthermore, using this technique, Sudhir Paul's group has obtained abzymes against the gp120 coat protein of HIV-1 virus (Paul *et al.*, 2003) and against the amyloid-β peptide (Taguchi *et al.*, 2008b).

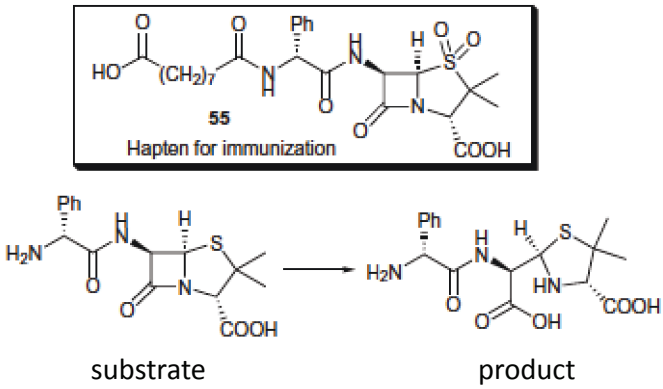


Figure 42. Hapten design using the structure of penam sulfone (a suicide inhibitor of the enzyme β-lactamase).

This hapten is used in reactive immunization to induce the production of catalytic antibodies with β-lactamase activity. Source: Tanaka *et al.*, 1999

The group of Gabibov has also used this strategy and a semisynthetic phage display library to isolate catalytic antibodies, notably antibody A.17, against covalently reactive phosphonate esters (Reshetnyak *et al.*, 2007; Smirnov *et al.*, 2011). The authors have used the term reactibody to denote this protein template selected for chemical reactivity rather than for ground state binding. They have suggested that the combined strategy of reactive immunization and kinetic selection allows for the eliciation of antibodies with the flexibility, plasticity, and dynamic structure suitable for catalytic activity. They have demonstrated, for example, that the active site cavity of the reactibody A.17 resembles more that of the cholinesterase enzymes, rather than other catalytic or non-catalytic antibodies (Figure 43).

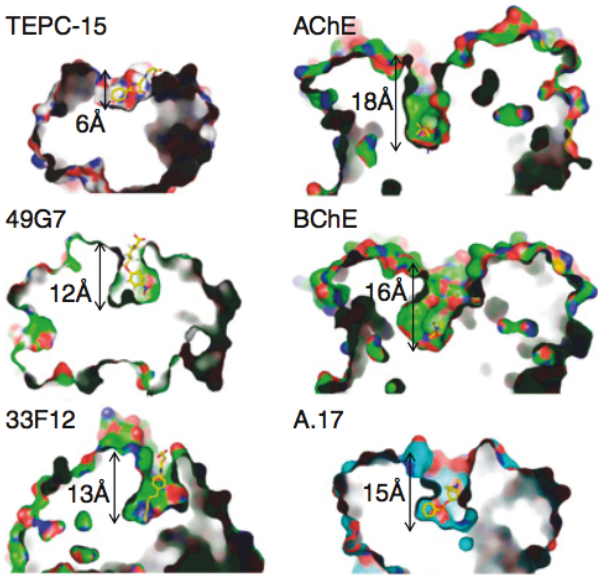


Figure 43. Comparison of active site cavities of natural and *de novo* created biocatalysts.

Reactibody A.17 has a deep substrate binding cavity, which is different from esterolytic antibodies 49G7 and TEPC15, and aldolase antibody 33F12. Its depth resembles more the cavities of choline esterase enzymes AChE and BChE. Source: Smirnov *et al.*, 2011.

The reactive immunization technique has proven most successful among the chemical methods for the elicitation of catalytic antibodies so far, producing abzymes approaching the catalytic efficiencies of natural enzymes (Tanaka and Barbas, 2002).

Aside from the TSA technique and its optimizations, two other approaches have also been employed for the chemical elicitation of catalytic antibodies. These are described below.

II.1.A.a.ii. Cofactor approaches

Cofactors are nonpeptidyl catalytic groups, such as ions, hemes, thiamine, flavins and pyridoxal, which aid enzyme catalysis (Blackburn and Garçon, 2000). Combining catalytic antibodies with cofactors has led to the improvement of the abzyme's catalytic activity in terms of substrate selectivity, turnover, and efficiency. This technique has also led to modifications of the abzyme active site in a way to completely alter its function. Using metal-coordinated natural enzymes as models, researches have put much effort into the development of catalytic antibodies that recruit metals at their active site. As an example, Kim Janda's group has used the antibody aldolase 38C2 (Barbas *et al.*, 1997), the only existing commercialized catalytic antibody, and bis-imidazolyl ligand coordinated copper complexes as cofactors to derive the semisynthetic metalloantibody 38C2-58-$CuCl_2$, which catalyzes the hydrolysis of picolic acid ester (Nicholas *et al.*, 2002). Additionally, the incorporation of palladium(II) with antibody aldolases 38C2 and 33F2 has shown to

accelerate their reaction rates while keeping the same catalytic mechanism (Finn *et al.*, 1998). Aside from metallic cofactors, external nucleophilic cofactors have also been employed to improve abzyme catalytic activity (Ersoy *et al.*, 1998).

II.1.A.a.iii. Inclusion of catalytic groups

Another method that has been used to produce catalytic antibodies is the direct inclusion of catalytic residues in the antibody binding site by site-directed mutagenesis. In a pioneer report, the group of Schultz introduced a significant esterase activity into an antibody combining site by a single amino acid change (Baldwin and Schultz, 1989). Later, Fletcher has used this approach to introduce ribonuclease activity in a scFv specific for RNA (Fletcher *et al.*, 1998). As a final example, Liu and colleagues have introduced protease activity in a scFv specific for a bacterial protein HPr (Liu *et al.*, 1998). Nevertheless, there are very few catalytic antibodies produced by this strategy reported in the literature. This is potentially due to the difficulty of creating a catalytic function by point mutations, independently of a structure-driven selection pressure.

Parallel to the efforts performed to chemically construct the best transition state analogs, and possibly as an attempt to overcome the difficulties in the hapten design process, strategies were developed to biologically generate high efficiency catalytic antibodies. These are described below.

II.*1.A.b. Biological strategies*

The biological methods use a biological antigen for the immunization of animals and the elicitation of catalytic antibodies. They are divided into two categories: *(i)* the manipulation of the idiotypic network and *(ii)* the utilization of protein inhibitors. The following are brief descriptions of these approaches.

II.1.A.b.i. Idiotypic network

The idiotypic network theory was first proposed by Niels Jerne in 1974 (Jerne, 1974a; Jerne, 1974b). This theory provides a distinction between idiotypic antibodies and anti-idiotypic antibodies. Idiotypic antibodies recognize an antigen *via* the structure of their idiotype (the ensemble of the antigen binding sites) or idiotopes present in the antibody variable region (Oudin and Michel, 1963; Kunkel *et al.*, 1963). Anti-idiotypic antibodies, on the other hand, recognize and bind the idiotypes of idiotypic antibodies. It has been suggested that this idiotypic network of antibodies provides a regulatory mechanism of idiotypic suppression in the adaptive immune system by either neutralizing secreted

idiotypic antibodies or providing a control on their secretion by B lymphocytes (Hampe, 2012).

The exploitation of the idiotypic network to elicit catalytic antibodies, known as the internal image theory, was performed for the first time by our group in 1993 (Izadyar *et al.*, 1993) (Figure 44).

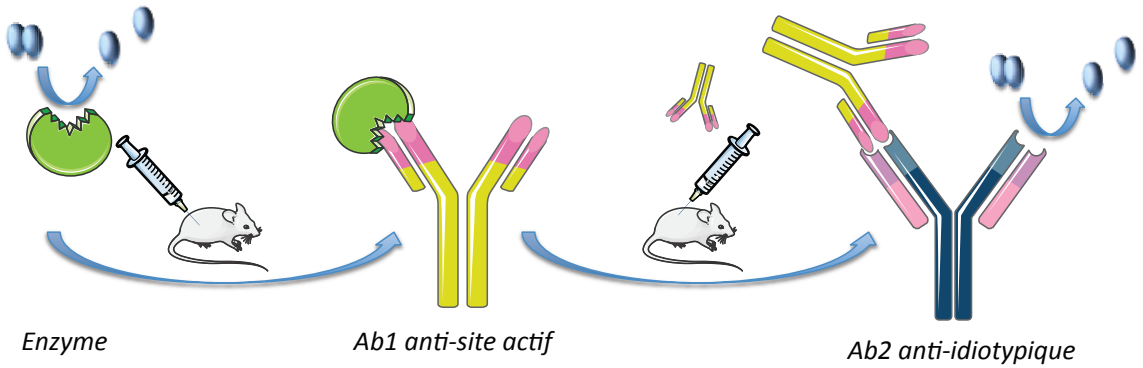


Figure 44. Schematic of the anti-idiotypic network.

In a first immunization event, a mouse is immunized with a model enzyme and an antibody specific to the active site of the enzyme leading to its inhibition is selected (Ab1). In a second immunization event, a second mouse is immunized with Ab1 and antibodies specific to the variable regions of Ab1 (anti-idiotype) are selected, which show catalytic activity similar to the original enzyme.

In this work, the antibody AE2 (Ab1), an anti-acetylcholinesterase idiotypic antibody able to inhibit the enzyme and commercially available as a hybridoma, was used in an immunization to isolate an anti-idiotypic antibody 9A8 (Ab2). This Ab2 was able to mimick the active site of the original enzyme, therefore, producing an internal image of the enzyme and displaying cholinesterase activity. Since this pioneering study, different enzymes have been used to elicit antibodies with different activities *via* the idiotypic network. The enzymes β-lactamase and subtilisin Carlsberg were used by our group to produce the antibody 9G4H9, displaying one of the highest amidase activities reported by catalytic antibodies (k_{cat} = 2.1 min^{-1}, K_m = 4300 μM) (Avalle *et al.*, 1998), and the antibody 6B8E12 (Pillet *et al.*, 2002; Ponomarenko *et al.*, 2007), respectively. The idiotypic network strategy for the isolation of the antibody 9G4H9 is described in more detail later in this chapter (Section II.2.B). Other examples of this technique include the production of an antibody with carboxypeptidase activity (Hu *et al.*, 1998) and an antibody with alliinase activity (Li *et al.*, 2012). It is of interest to note that antibodies generated by this approach have higher catalytic efficiencies than abzymes produced by others (Tellier, 2002).

II.1.A.b.ii. Protein inhibitors

An alternative strategy to the idiotypic network methodology described above is the utilization of protein inhibitors as immunogens to elicit antibodies endowed with catalytic

activity of the corresponding enzyme. This method is also an implication of the internal image theory, where the protein inhibiting an enzyme in a competitive manner holds similar structural information as the enzyme active site. Therefore, antibodies directed against this protein should have the internal image of the original enzyme active site and hence be potentially endowed with the same catalytic activity. An example of antibodies developed with this technique is the utilization of Tendamistat, an inhibitor of the enzyme pancreatic α-amylase present in porc (Goncalves *et al.*, 2002). The polyclonal antibodies obtained by this work demonstrated a significant catalytic activity.

II.1.A.b.iii. Conclusions on deliberately produced abzymes

As mentioned previously, despite the initial enthusiasm raised toward the biocatalysis capabilities of catalytic antibodies, they have proven to be less efficient than their enzymatic counter-parts, with enzymes having catalytic efficiencies (k_{cat}/K_M) in the range of 10^6-10^8 $M^{-1}.sec^{-1}$ *versus* abzymes' catalytic efficiencies of the order 10^2-10^4 $M^{-1}.sec^{-1}$ (Hilvert, 2000; Wentworth and Janda, 2001). For the chemical production of abzymes, this is partially due to the fact that the design of stable but reactive analogs is problematic. Additionally, the TSA mimicking moieties of haptens are positioned near the entrance of the antibody-combining site, whereas in natural enzymes, the catalytic machinery is buried in pockets of the enzyme active site, isolating chemical transformations from the solvent (Petsko and Ringe, 2004; Herschlag and Natarajan, 2013).

The proximity effect, proposed by Pauling, responsible for bringing the substrate in close proximity with the catalytic groups of the enzyme in order to ultimately lower the entropy of the reaction transition state is a characteristic developed by enzymes over a million years of evolution. Abzymes, on the other hand, are induced to develop a catalytic activity in the laboratory over a period of weeks. In contrast to enzymes, antibodies have been evolved through a process of affinity maturation in order to develop high affinities for their corresponding antigens. These affinities can reach upto sub-picomolar dissociation constants, much higher than the affinity of enzymes to their substrates (Rathanaswami *et al.*, 2005; Rispens *et al.*, 2013).

Furthermore, the mechanism involved in antigen internalization for the activation of B lymphocytes and leading to the production of antibodies can take upto 30 min to 1 h (Jang *et al.*, 2010). This means that for a catalytic B cell to be activated, the antigen retention time has to be inferior to the time required for catalytic activation, corresponding to an antibody with a low catalytic turnover (Padiolleau-Lefèvre *et al.*, 2014). All of these characteristics of the nature of catalytic antibodies are potential reasons for the discrepancies between the catalytic capabilites of enzymes and abzymes.

Shortly after the first successes of the experimental production of catalytic antibodies in the 1980's, reports of the observation of naturally occurring abzymes in human sera produced a second wave of enthusiasm for researchers in the field. These observations led

the study of catalytic antibodies toward another exciting direction: the analysis of the implications of these antibodies in physiopathology.

II.1.B. Naturally occurring catalytic antibodies

In 1989 Sudhir Paul and colleagues became the first group to report the presence of naturally occurring catalytic antibodies in human sera (Paul *et al.*, 1989). These antibodies were capable of hydrolyzing the vasoactive intestinal peptide (VIP) in the sera of asthmatic patients. Since then, the discovery of many other naturally occurring catalytic antibodies, both in physiological and pathological conditions, having both positive and negative effects on patients' health (Wootla *et al.*, 2011b), has stimulated the interest of researchers in the field, as well as raised many questions about the role of these proteins in the immune system. The following are examples of naturally occurring catalytic antibodies reported to date in physiopathology.

II.1.B.a. Abzymes in pathology

Since Paul's discovery, numerous studies have demonstrated the prevalence of catalytic antibodies in different pathological conditions. Interestingly, many of them exist on the background of an autoimmune disease (Table 6).

Table 6. Catalytic antibodies reported in different pathological diseases.

Pathology		Targeted Antigen	References	Suspected Effect
Inflammatory disorders	Asthma	Vasoactive intestinal peptide (VIP)	Paul *et al.*, 1989	-
	Sepsis	FVIII, FIX	Lacroix-Desmazes *et al.*, 2005	+
Autoimmune disorders	Hashimoto's thyroiditis	Thyroglobulin	Li *et al.*, 1995	+/-
	Systemic lupus erythematosus	DNA, RNA	Shuster *et al.*, 1992; Nevinsky and Buneva, 2002	-
	Scleroderma	DNA, RNA	Vlassov *et al.*, 1998	unknown
	Rheumatoid arthritis	DNA/RNA	Baranovskii *et al.*, 1998	unknown
	Multiple sclerosis	DNA, RNA, MBP	Baranovskii *et al.*, 1998; Ponomarenko *et al.*. 2002; Gabibov *et al.*, 2011	-
	Acquired hemophilia	FVIII, FIX	Wootla *et al.*, 2008a	+/-
	Alzheimer's disease	Aβ-peptide	Paul *et al.*, 2005; Taguchi *et al.*, 2008a,b	+
Metabolic disorders	Diabetes	Grp94	Pagetta *et al.*, 2007	-
Infectious disorders	Immune thrombocytopenia in HIV-1	Platelet GPIIIa	Nardi and Karpatkin, 2000	+
Neoplastic disorders	Multiple myeloma	Prothrombin	Mastuura *et al.*, 1994; Paul *et al.*, 1995	-
	Waldenstrom's Macroglobulinemia	Aβ-peptide	Planque *et al.*, 2004; Taguchi *et al.*, 2008a	-
Alloimmune disorders	Hemophilia A	FVIII	Lacroix-Desmazes *et al.*, 1999	-
Other	Renal rejection in transplantation	FVIII, FIX	Wootla *et al.*, 2008b	+

II.1.B.a.i. Abzymes with nuclease activity

A large number of the reported naturally occurring catalytic antibodies are endowed with DNA- and RNA- hydrolyzing activity. Alexander Gabibov's group has reported the presence of antibodies with DNase activity in the sera of patients with systemic lupus eythematosus (SLE) and demonstrated a correlation between the activity of these antibodies and the stage of development of the disease (Shuster *et al.*, 1992). A different group has also reported an elevated RNase activity by polyclonal antibodies isolated from sera of patients with SLE as well as hepatitis B (Vlassov *et al.*, 1998). Polyclonal antibodies with both DNase and RNase activity have also been discovered in the blood and cerebrospinal fluid of patients with multiple sclerosis (MS) (Baranovskii *et al.*, 1998). In these examples, the catalytic antibodies were suspected to have deleterious effects on the patients' health. On the contrary, as an example of catalytic antibodies with a beneficial influence for patients' health, Gabibov's group has made the observation that anti-DNA catalytic antibodies from sera of SLE and chronic lymphocytic leukemia patients have anti-tumor properties (Kozyr *et al.*, 2000). The authors examined the cytotoxicity of DNA-specific autoantibodies from sera of patients and healthy donors on different tumor cell lines. Their results demonstrated a correlation between a caspase-dependent cytotoxicity and the DNA-hydrolyzing properties of the autoantibodies. All of these examples taken together illustrate that despite the relevance of catalytic antibodies enable to hydrolyze nucleic acids in numerous pathologies, their exact role, whether beneficial or detrimental, is not yet understood.

II.1.B.a.ii. Abzymes with protease activity

A second category of catalytic antibodies, reported naturally in subjects, are abzymes with protease activity. Proteolytic antibodies, for example, have been proposed to play a role in certain hormonal dysfunctions. Pagetta and colleagues have argued that the presence of anti-idiotypic complexes of anti-Grp94 idiotypic antibodies, with serine-like amydolytic activity, and anti-idiotypic antibodies is the reason for the absence of proteolytic activity in plasma of patients with type 1 diabetes (Pagetta *et al.*, 2007). Furthermore, Li and colleagues have shown that anti-thyroglobulin (Tg) antibodies in patients with Hashimoto's thyroiditis are responsible for the cleavage of Tg, the precursor of thyroid hormones and a proven autoantigen in autoimmune thyroiditis (Li *et al.*, 1995). The role of these antibodies in Hashimoto's thyroiditis, however, remains unclear. The elimination of Tg by anti-Tg catalytic antibodies might lead to the reduction of autoimmune response. On the other hand, these antibodies have been shown to access the thyroid and might potentially lead to the depletion of thyroid hormones, making them detrimental to the patient's health.

A number of other experiments have demonstrated, however, that these autoantibodies are present only in SLE and autoimmune thyroiditis, and not in healthy individuals (Paul *et al.*, 1997). The authors have suggested that the development of the autoimmune state is responsible for the transition from polyreactive proteolytic activity, present in all subjects,

to autoantigen-directed activity in patients, potentially due to the chemical reactivity of the antigen promoting the expression of catalytic activity by autoantibodies. The observations around these anti-Tg catalytic antibodies, therefore, provide another example of the potential duality of the role of abzymes in pathology. Secondly, this example illustrates a recurring observation in the study of catalytic antibodies, which is a potential link between their expression and the existence of an autoimmune state.

Abzymes have also been demonstrated to play a role in a number of neurodegenerative diseases. Rangan and colleagues have reported the hydrolysis of amyloid β peptide (AβP), involved in the development of Alzheimer's disease (AD), by cross-reactive antibodies to neuropeptide vasoactive intestinal peptide (Rangan *et al.*, 2003). AβP-hydrolyzing IgM antibodies have also been observed to substantially degrade AβP and block its aggregation and toxicity (Paul *et al.*, 2005; Taguchi *et al.*, 2008a). Catalytic antibodies hydrolyzing the myelin basic protein (MBP) involved in multiple sclerosis (MS) have been reported first in experimental allergic encephalomyelitis (EAE) mouse models (Ponomarenko *et al.*, 2002) and later isolated from patients (Ponomarenko *et al.*, 2006a). This work has given a diagnostic value to the presence of serum autoantibodies able to recognize and degrade MBP peptides, acting as novel biomarkers for MS (Belogurov *et al.*, 2008). The heightened incidence of the expression of catalytic antibodies in the ensemble of the above mentioned autoimmune neurodegenerative diseases further confirms the hypothesis of an existing link between abzymes and autoimmunity.

This link is also evidenced by the presence of catalytic antibodies involved in the coagulation cascade. There have been reports of autoantibodies hydrolyzing prothrombin in SLE (Bertolaccini *et al.*, 1998), thrombosis (Palosuo *et al.*, 1997), and multiple myeloma (Thiagarajan *et al.*, 2000). The latter study has suggested that these antibodies might induce a pro-coagulant state by cleaving prothrombin and generating thrombin-like activities, but the precise connection of these antibodies and hypercoagulability remains unclear. The anti-prothrombin autoantibodies are, thus, involved in the development of disease and provide examples of abzymes with a detrimental role to the patient's health. Autoantibodies hydrolyzing factor VIII (FVIII) and factor IX (FIX) have shown to be involved in acquired hemophilia A (Wootla *et al.*, 2008a; Wootla *et al.*, 2011a). Anti-FIX antibodies have been suggested to have IgG-mediated FIX activation, representing an anti-hemorrhagic mechanism and providing a beneficial role to the patient's health. The presence of such antibodies has also been associated with a reduced risk of chronic allograft nephropathy in renal trasnplants (Wootla *et al.*, 2008b), providing another example of abzymes with positive influences to patients' health.

II.1.B.a.iii. Regardless of activity, abzymes display a complex dual behavior...

To conclude, the numerous examples of catalytic antibodies reported in different pathological states, demonstrate two critical observations. First is the dual role exhibited by these proteins, where the abzymes play a beneficial function in some instances, whereas

they have a detrimental effect in others. In a report published by the group of Alexander Gabibov, this dual role has cleverly been compared to the fictional characters of Dr. Jeckyl and Mr. Hyde, hinting to the complexity of abzyme function (Belogurov *et al.*, 2009) (Figure 45).

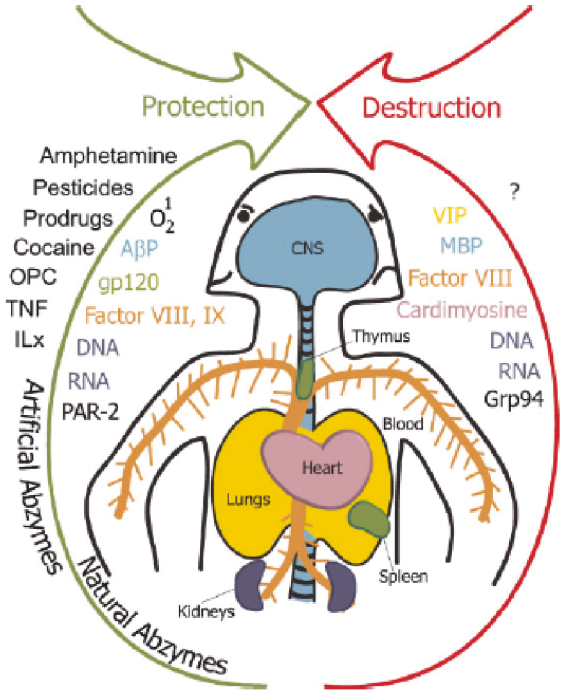


Figure 45. Dual role of catalytic antibodies in physiopathology.

Source: Belogurov *et al.*, 2009

A second observation, which cannot be overlooked, is the extensive number of reported catalytic antibodies on the background of autoimmune disease, suggesting a potential link between the expression of abzymes and autoimmunity. This is an important question facing the researchers in the field today. The hypothesis of the existence of this potential link has been further strengthened by studies showing that immunization of autoimmune-sensitive mice (SJL/L, an EAE model) leads to a significantly larger rate of expression of catalytic antibodies as compared with healthy mice (BALB/C) (Tawfik *et al.*, 1995) or a higher proteolytic activity of polyclonal antibodies (Gabibov *et al.*, 2002).

Almost a decade after the discovery of catalytic antibodies in disease, reports of their presence in normal physiological states added even further to the ambiguities around the question of the role of these antibodies in the immune system. The following are examples of the instances of abzymes in normal physiology.

II.1.B.b. Abzymes in normal physiology

Sudhir Paul and colleagues reported the first instance of the presence of catalytic antibodies in healthy individuals in 1995 (Kalaga *et al.*, 1995). The authors observed IgG antibodies with peptidase activity in the serum of healthy subjects, which were absent in patients with rheumatoid arthritis. This discovery demonstrated that the immune system is capable of spontaneously generating catalytic antibodies in the absence of immunization by an exogenous molecule.

Since this first observation, naturally occurring catalytic antibodies have been reported in other normal physiological conditions, contributing to the uncertainty of the role of abzymes in physiopathology. The same group in 2004, for example, reported the presence of catalytic IgM antibodies in healthy individuals capable of cleaving the gp120 protein of the HIV, potentially providing a protection against this virus (Paul *et al.*, 2004). Moreover, IgG and IgA antibodies with protein kinase and DNase activity have been reported in the milk of healthy lactating mothers (Kit *et al.*, 1996; Kanyshkova *et al.*, 1997). A more recent study linking the presence of proteolytic IgA antibodies to the activation of Proteinase Receptor-Activated 2 (PAR-2) receptors in intestinal epithelial cells responsible for the expression of an antimicrobial peptide, suggested that the catalytic antibodies present in milk might play a beneficial role in innate immunity (Barrera *et al.*, 2009).

Wentworth and colleagues have proposed that all antibodies, regardless of their antigen specificity, have bactericidal activity (Wentworth *et al.*, 2002). This is due to the ability of antibodies to catalyze the generation of hydrogen peroxide (H_2O_2) from water and single molecular oxygen, the latter provided by activated neutrophils that bind immunoglobulins. This observation ties into a previous hypothesis raised by our group that naturally occurring catalytic antibodies play a direct role in the elimination of the biochemical wastes released by the metabolism of the organism, providing an intrinsic protective property under physiological conditions (Friboulet *et al.*, 1999).

II.1.B.c. Conclusions on naturally occuring abzymes

The numerous discoveries of catalytic antibodies soon after their production as biochemical tools, in both physiological and pathological conditions, leading to both beneficial and detrimental outcomes for the organism, has made it difficult for researchers to draw any concrete conclusions about the place of these peculiar proteins in physiopathology. What is certain, however, is the complexity of the role of abzymes in the immune system. It is now known that autoreactivity exists naturally in the innate as well as the adaptive immune response (Bendelac *et al*, 2001). In healthy physiological conditions, it seems that this autoreactivity remains in a state of homeostasis. This control is exerted by negative selection of the B lymphocytes expressing autoreactive antibodies, by mechanisms of the regulatory T lymphocytes, or by an anti-idiotypic network. Under pathological conditions, however, this homeostasis is disrupted, autoreactive B lymphocytes are activated, and clonal expansion leads to a higher expression of particular

autoreactive antibodies. This hypothesis is validated by the recognition by certain catalytic antibodies of the autoantigens directly responsible for the development of disease, such as the cases of hemophilia A (Wootla *et al.*, 2008a) and MS (Ponomarenko *et al.*, 2006a). There are other contradicting instances, however, when catalytic antibodies are targeted to antigens not at all involved with the disease (i.e. DNA-hydrolyzing antibodies in MS (Baranovskii *et al.*, 1998) or FVIII-hydrolyzing antibodies in sepsis (Wootla *et al.*, 2006)). In these cases, the existence of the catalytic antibody is most likely not due to clonal expansion of B lymphocytes directly resulted by the disease-causing antigen. This idea is further confirmed by studies showing the prevalence of catalytic IgM antibodies in certain conditions (Planque *et al.*, 2004). These studies have suggested that somatic mutations disfavor catalytic activity and that the structure of IgM prior to undergoing class comutation is favorable for catalysis (Sapparapu *et al.*, 2012). These hypotheses are, once again, contradictory with the fact that most of the naturally occurring catalytic antibodies in different pathologies are IgG. The observation of IgGs in the majority of the reported cases, however, could potentially be simply due to experimental biases or to the choice of authors to focus only on this isotype.

The different examples of the instances of naturally occurring catalytic antibodies leading to different, and sometimes contradictory, hypotheses on the nature and role of these immunoglobulins, is evidence of the complexity of the immune system and the convoluted function of abzymes in this system. We are a long way from understanding the place of these special proteins in physiopathology and many more studies need to be performed in order to bring us closer to this objective. The following is an example of a theoretical study performed by our group in order to reveal more information on the nature and characteristics of catalytic antibodies.

II.1.C. Genetic study manifesting certain characteristics of abzymes

Focusing on the study of the immunoglobulin genes encoding for catalytic antibodies can potentially reveal useful information on the nature of these proteins as well as provide hints as to how the immune system leads to their generation. This is why many researchers have focused on such analyses. Numerous studies have focused on sequence alignment of abzyme-encoding genes and their germline counterparts, the level of amino acid modification following B cell maturation, and both of their effects on the abzyme's catalytic activity. A study performed by Sudhir Paul's group generating a protease catalytic antibody against VIP, has identified the mature antibody's mutated residues, and has performed point mutations to convert the antibody to its germline-coded form (Gololobov *et al.*, 1999). The authors observed that both germline and mature forms of the antibody demonstrate similar catalytic activity, and hence proposed that the catalytic activity of certain antibodies is an innate function as opposed to somatic processes. A similar study by Peter Shultz's group reported that converting the antibody AZ-28 to its germline version improved the abzyme's catalytic efficiency by 35-fold (Ulrich *et al.*, 1997). The hypotheses arising from the latter study are in line with the results acquired by Tawfik and collaborators, demonstrating that a long immunization protocol (>3 months),

corresponding to a higher level of affinity maturation, as opposed to a short immunization protocol (21 days), leads to a decrease in the expression of catalytic antibodies from 75% to 6% (Tawfik *et al.*, 1995). A number of other studies done by Peter Shultz and Donald Hilvert, however, have observed a contradictory phenomenon. These studies have suggested that the acquisition of amino acid modifications following somatic mutations leads to an improved catalytic activity by the antibody (Romesberg *et al.*, 1998; Xu *et al.*, 1999). Therefore, catalytic antibodies have once again proven to have a complex nature, even at the level of genetic sequence, resulting in a great difficulty of deciphering their systematic characterization.

The idea that catalytic antibodies with higher catalytic activities are encoded by the germline or have a higher germline homology is in line with numerous instances of the reported polyreactivity of abzymes (Kalaga *et al.*, 1995; Tawfik *et al.*, 1995; Planque *et al.*, 2004; Zein *et al.*, 2010b). Indeed, a number of studies have suggested that polyreactive antibodies generally possess a low specific affinity for a variety of antigens and their sequences are close to their germline (Casali and Notkins, 1989; Zhou *et al.*, 2007). This is a characteristic quite different from that of high-affinity mono-reactive binding antibodies, which acquire their affinities by undergoing processes of somatic mutation and affinity maturation.

In order to provide a better look on the genetic characteristics of catalytic antibodies, a recent study done by our group has focused on an analysis of 40 abzymes reported in the literature with a variety of activities (i.e. amidase, esterase, DNase, decarboxylase, transaminase, and Diels-Alder activity) and different methods of generation (Le Minoux *et al.*, 2012). Performing the study on a combined pool of these antibodies helps to avoid any bias brought by the experimental methods leading to the production or the particular activity of abzymes. A first point of analysis focused on the level of identity of the catalytic antibodies with their corresponding germline genes. Using the V-QUEST tool available in the IMGT database (Brochet *et al.*, 2008), it was observed that the percentage of identity of the immunoglobulin variable regions of these antibodies was 97.5±1.7% for V_L and 97.1±2.5% for V_H. These values demonstrate that the studied catalytic antibodies have a high level of sequence similarity in their nucleotide sequence as compared to their germline.

As a second point of analysis, the mutation rate was explored at the amino acid level. Excluding silent mutations, it was observed that catalytic antibodies on average contained 6.4 mutated residues on V_L and 11.1 on V_H, which leads to a total of 17.5 mutated residues on the V_L/V_H pair. A similar study performed by Poulsen *et al.* on a total of 48 binding (i.e. non-catalytic) antibodies reported that an average of 21 mutations is produced during affinity maturation on the V_L/V_H pair (Poulsen *et al.*, 2007), which is 16.7% higher than the value observed for catalytic antibodies.

As a third point of analysis, the study focused on a qualitative measure of the physicochemical modifications resulting from the mutated amino acids. In order to perform a valid analysis for this point, the window of study was restricted to catalytic antibodies with the same activity and a common and well-identified catalytic motif. This study was,

therefore, focused on 23 catalytic antibodies endowed with amidase activity (i.e. protease, peptidase, β-lactamase). This type of activity is largely described in literature and it was the largest group of referenced catalytic antibodies with a similar activity, making the group a good candidate for study. It was observed that 64.5% of the mutations in V_L and 65.1% of the mutations in V_H induced important physicochemical modifications. In this analysis, the level of physicochemical modifications of amino acids are measured as a function of side chain properties such as hydropathy, volume, chemistry, charge, polarity, and being a hydrogen donor or acceptor. Such modifications are ranked by the DomainGapAlign tool of IMGT as similar, dissimilar, or very dissimiliar (Ehrenmann *et al.*, 2010; Ehrenmann and Lefranc, 2011). It was thus considered that a "very dissimilar" qualification according to IMGT critera induced important physicochemical modifications. Furthermore, the study focused on 15 of the amidase antibodies for which the active residues were identified and published. It was observed that 50% of the antibodies contained a mutated residue in their active site and that 71% of these mutations induced important modifications of the physicochemical properties of the amino acids involved. These results demonstrated that catalytic antibodies contain a low level of mutation as compared to their germline sequence, but these mutations provide the antibody with significantly modified physicochemical properties potentially influencing the acquisition of the catalytic activity.

In an attempt to further explore potential characteristics shared by catalytic antibodies, the study performed a parallel analysis on the frequency of immunoglobulin gene subgroups expressing abzymes. Using the GeneFrequency tool of the IMGT database (Lefranc *et al.*, 2009), the authors have defined a number of the murine κ gene subgroups as rare, based on their low representation of less than 5% (Figure 46). In this way, they have demonstrated that 16.1% of the V_κ sequences available in IMGT are expressed by rare gene subgroups.

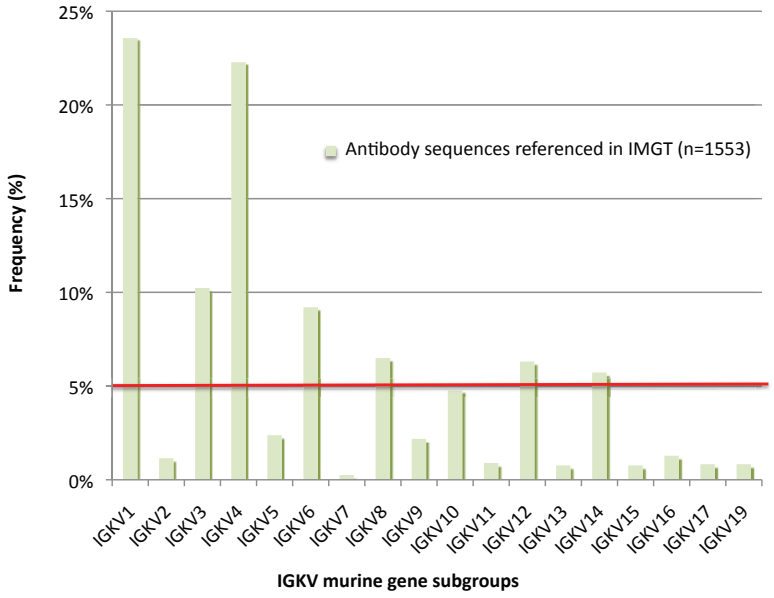


Figure 46. The frequency of IGKV gene subgroups referenced in IMGT.

Subgroups below the red line are considered as rare subgroups. Date of consultation of database: 2011.

The authors have then proceeded with a similar analysis on the 33 murine catalytic V_κ sequences assembled from the literature and observed that 33.3% are encoded by rare subgroups, which is significantly higher than the binding antibodies referenced in IMGT ($p<0.01$) (Figure 47).

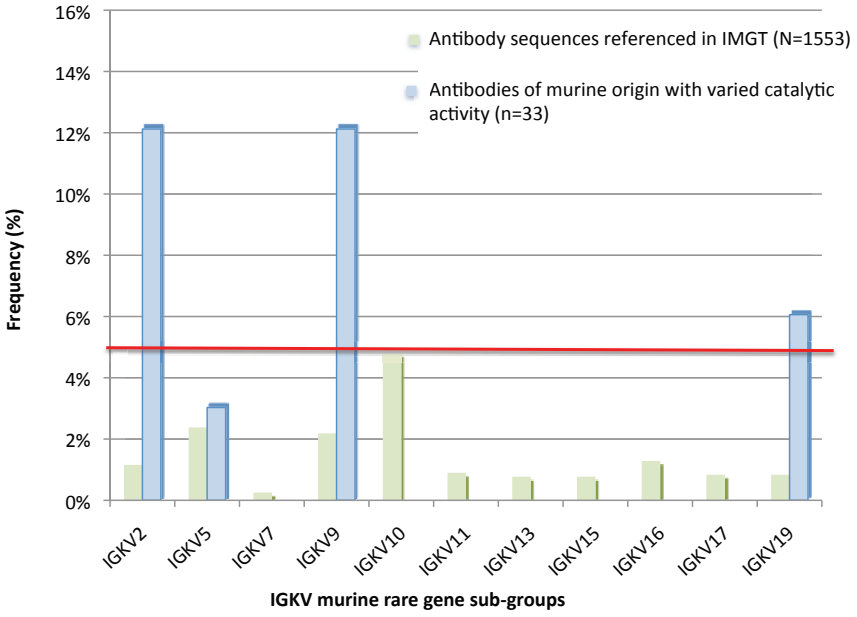


Figure 47. Comparison between the expression frequencies of IGKV rare gene subgroups between binding and catalytic antibodies.

Date of consultation of database: 2011.

Interestingly, by expanding the analysis in order to include also human V_κ genes, the same tendency was observed. The results demonstrated that 6.9% of the 3664 sequences available in IMGT *versus* 28.9% of the 38 sequences from literature ($p<0.001$) are encoded by rare gene subgroups.

The authors have suggested that the reason for this discrepancy is explained through an evolutionary standpoint: the primary role of binding antibodies is to acquire a high affinity for their corresponding antigen through a process of somatic mutations leading to affinity maturation. Therefore, it follows that the gene subgroups that are more susceptible to mutational mechanisms are more frequently expressed by binding antibodies. And similarly, it also follows that abzymes that favor a lower mutation rate and whose main function is catalysis are more frequently expressed by rare gene subgroups. These results also raise the question of the possibility of a "catalytic maturation" process of abzymes *versus* the affinity maturation of binding antibodies.

The question of whether the catalytic activity of abzymes is an innate characteristic of the immune system or produced by antigen activation and B cell maturation is a recurring one still facing researchers in the field today. The observations provided by the study of Le Minoux and colleagues (2012) seem to be more in favor of the former hypothesis. The

proposed idea that a high level of mutations occurring in the variable region of an antibody reduces the degree of conformational freedom of its CDR regions (Venkateshrao *et al.*, 2004) provides support for these observations. This rigidity on the recognition site of the antibody enforces its strong interaction with the antigen. In the case of catalysts, however, a high rigidity of the active site is unfavorable for the release of products and the catalytic turnover (Petsko and Ringe, 2004; Hammes-Schiffer and Benkovic, 2006), which is also probable for the CDR regions of catalytic antibodies. Therefore, the gene subgroups less susceptible to mutation, and hence leading to a lower level of affinity maturation, are more favorable for the catalytic function of antibodies. It is important to note that the mentioned hypotheses are based on a small sampling of well-characterized catalytic antibodies. Further studies need to be conducted in order to provide a more concrete understanding of the genetic origins and characteristics of catalytic antibodies, and to identify the specific events in the immune response that favor their generation.

II.1.D. Applications of catalytic antibodies

Despite their lower turnover, there are certain characteristics of abzymes as compared with enzymes, which have made them interesting candidates for a number of clinical applications:

- Their higher half-life corresponding to longer serum circulation periods
- Their use in lower doses as compared to other drugs
- Their lower immunogenecity and peripheral toxicity
- Their relevance in autoimmune disease
- Their multivalency, allowing to diversify their functionalities (i.e. to couple a binding and a catalytic activity)

Table 7 summarizes the various therapeutic applications provided by abzymes to date.

One particular application of interest to note is the antibody directed enzyme prodrug therapy (ADEPT). This is a two-component strategy combining an antibody-enzyme conjugate and an anticancer prodrug. The antibody binds to a tumor-associated antigen and brings the enzyme in proximity of the tumor surface. A non-toxic prodrug of an anticancer agent is administered, which the enzyme hydrolyzes and converts into its active form at the local environment of the tumor. The immunogenecity of the enzyme, however, is a limitation of this technique. Replacing the antibody-enzyme conjugate by a humanized catalytic antibody, antibody directed abzyme prodrug therapy (ADAPT), is therefore an interesting and viable solution. A bispecific abzyme combining an anti-tumor function with a catalytic one is an example of this concept (Wentworth *et al.*, 1996). The abzyme 38C2 displaying aldolase activity, which activated the prodrug forms of doxorubicin and camptothecin is another example (Shamis *et al.*, 2004; Sinha *et al.*, 2004). It is important to note, however, that this antibody reaches its target either by a local injection or by conjugation with a targeting compound. Three humanized variants of 38C2 have entered clinical trials (Goswami *et al.*, 2009; Rader, 2014).

Table 7. Therapeutic applications of catalytic antibodies.

Application	Antibodies	Activity	References
Abzymes against drug addiction	15A10	Cocaine esterase	Briscoe *et al.*, 2001;
	3F5, 3H9		McKenzie et al., 2007
	TD1-10E8, TD1-36H10	Oxidative degradation of nicotine	Dickerson *et al.*, 2004
Abzymes against cancer			
Prodrug activation	38C2	Aldolase	Abraham *et al.*, 2007
	84G3, 85H6, 90G8	Aliinase	Goswami *et al.*, 2009
	VHHC10		Li *et al.*, 2012
Gene silencing	3D8-VL	mRNA of HER2 hydrolysis	Lee *et al.*, 2010
Abzymes against chemical warfare agents	A17	Degradation of organophosphate compounds	Smirnov *et al.*, 2011
Abzymes against infectious agents			
HIV	YZ18, YZ20, YZ24	Cleavage of gp120	Paul *et al.*, 2003; Planque *et al.*, 2008
Rabies virus	Clones 1 and 18 (Light chains)	Peptidase	Hifumi *et al.*, 2013
Helicobacter pylori	UA15-L	Protease	Hifumi et al., 2008
Abzymes against inflammatory diseases			
Alzheimer disease	c23.5, Polyclonal autoAb	Proteolytic cleavage of Aβ-peptide aggregates	Rangan *et al.*, 2003; Taguchi *et al.*, 2008 a,b; Paul *et al.*, 2010
Autoimmune inflammatory disorders	ETNF-6-H	Hydrolysis of TNF-α	Hifumi *et al.*, 2010
Abzymes for the protection of oxydative damage	Se-scFv-B3; scFv 2D8	Glutathion peroxidase activity	Xu *et al.*, 2010
Abzymes against coagulation factor	Polyclonal antibodies	Hydrolysis of FIX	Wootla *et al.*, 2011a

Another valuable application of catalytic antibodies lies in their potential as catalytic vaccines. An example of this is the presence of catalytic antibodies specific to the glycoprotein gp120 of the HIV-1 virus, which have been reported in the sera of healthy individuals (Paul *et al.*, 2004; Hifumi *et al.*, 2002; Gorny *et al.*, 2005; Choudhry *et al.*, 2007; Gach *et al.*, 2008), SLE patients (Karle *et al.*, 2004), and induced by animal immunization (Ponomarenko *et al.*, 2006b). The group of Alexander Gabibov has vaccinated SJL/J mice in an active immunization procedure by using DNA combined with gp120 fragments incorporated into liposomes, and have demonstrated the induction of catalytic antibodies with virus neutralizing capability, which can further be humanized and used for passive vaccination (Durova *et al.*, 2009). This has shown that the immune repertoire of antigen-recipient animals is a great source for the development of such epitope-specific antibodies, which can be used for further passive vaccination.

The frequent appearance of abzymes in numerous autoimmune diseases also provides these proteins with potential therapeutic and diagnostic value (Ponomarenko *et al.*, 2002). The isolation of efficient catalytic domains directed against pathogenetically and clinically relevant autoimmune epitopes offers great therapeutic potential, whereas detection of abzymes showing catalytic activity against organ or tissue specific antigenic targets can lead to future diagnostic measures.

Abzymes have also been used in the fight against cocaine addiction. Anti-cocaine abzymes have been reported, which are able to bind cocaine and hydrolyze it to its non-psychoactive produces (Deng *et al.*, 2002; McKenzie *et al.*, 2007; Gorelick, 2012). Therefore, such catalytic antibodies have great potential for the development of vaccines (Padiolleau-Lefèvre *et al.*, 2014).

More marginally, catalytic antibodies have been used in a number of industrial applications. They have, for example, been used in the pharmaceutical industry for the large-scale preparation of enantiomerically pure intermediates and products (Schmid *et al.*, 2001). One of the interesting industrial applications of enzymes is their immobilization on solid surfaces in order to retain their continuous process. A main problem with this technique, however, is the attachment of enzymes without causing a loss of activity. Therefore, the structure of antibodies, notably the presence of their constant region which does not attribute to the catalytic activity, and their ability to maintain activity and specificity in organic solvents (Rao and Wootla, 2007) make them interesting candidates for such applications.

The focus of the current study is, however, a more fundamental question: the elucidation of the nature of catalytic antibodies, determination of potential characteristics unique to these proteins, and the influence of the immune repertoire properties on their expression. Historically, our group has pioneered the use of the idiotypic network in the production of catalytic antibodies. This technology has been used in our laboratory to produce antibodies with esterase, peptidase, and β-lactamase activities. Taking into consideration the previous experience of our group with the β-lactamase activity and the vast amount of information available on this enzyme, we consider it a suitable model for the study of catalytic antibodies. Additionally, the fact that β-lactamases are unique to bacterial species and do not exist in the eukaryotic world makes their activity a desirable model for the study of catalytic antibodies in animals. For example, in studies using the hybridoma technology or the idiotypic network, we are certain that potential contamination with endogenous enzymatic activity can be avoided. Furthermore, the nature of this activity is of interest in this field, due to the general low activation energy of the amidase activity, necessary for the hydrolysis of the amide bond of the β-lactam ring. This is why during the past two decades, our group has focused on this model for the elicitation and characterization of catalytic antibodies and we have chosen to focus on this model catalytic activity for the realization of our study.

II.2. β-lactamase: a model catalytic activity

II.2.A. The enzyme

β-lactamases are bacterial enzymes discovered for the first time in 1940 by Edward Abraham and Ernst Chain (Abraham and Chain, 1940). They provide bacteria with resistance to β-lactam antibiotics by hydrolyzing their conserved β-lactam moiety and hence destroying their antibiotic activity. Several β-lactam antibiotic structures have been developed as therapeutic measures against bacterial infection (Figure 48). These antibiotics form stable acyl-enzymes with penicillin-binding proteins (PBPs) in the membrane, inhibiting peptidoglycan biosynthesis and leading to bacterial cell death (Vollmer *et al.*, 2008). A major challenge of antibiotic design is that despite the molecular modifications of penams and cephems to yield newer drugs with better affinity for PBPs and poorer substrates for β-lactamases, new bacteria continue to emerge as a result of natural selection, resulting in the development of more effective β-lactamases adapted to the hydrolysis of newly commercialize β-lactam antibiotics (Kluge and Petter, 2010).

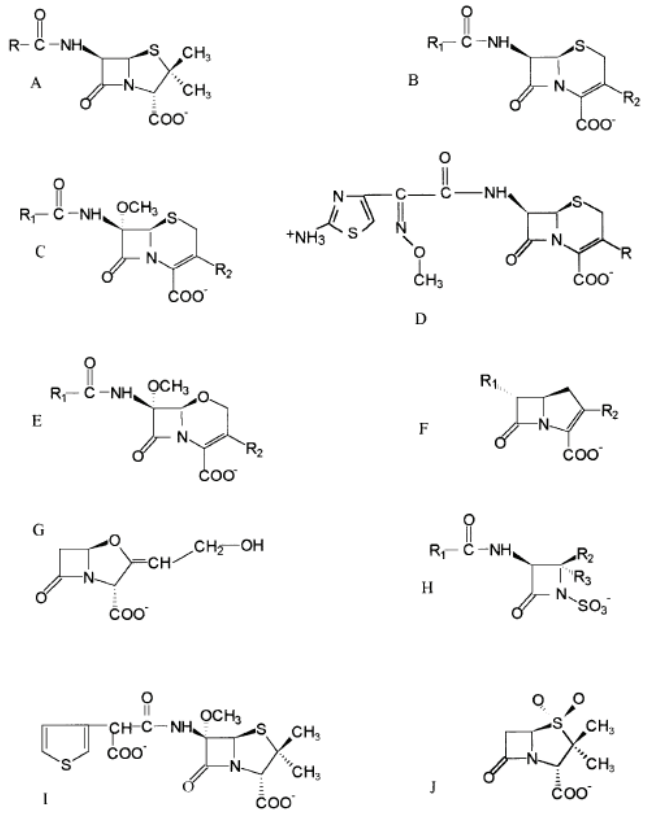


Figure 48. Structure of some β-lactam antibiotics.

(A) Penams (e.g. benzylpenicillin, ampicillin, ampxycillin), (B) Cephems (cephalosporins), (C) Cephamycins (e.g. cefoxitin), (D) Cefotaxime (oximino cephalosporin), (E) Oxacephamycins (e.g. moxalactam), (F) carbapenems (e.g. imipenem), (G) calvulanate (oxapenam), (H) Monobactams (e.g. aztreonam), (I) Temocillin (6-α-methoxy penam) and (J) sulbactam (penam sulfone). Source: Matagne *et al.*, 1998.

β-lactamases are grouped into four familial classes based on their molecular structure: A, B, C, and D (Bush and Jacoby, 2010). Classes A, C, and D enzymes share a common serine-based mechanism, whereas the activity of class B enzymes, also known as metallo-β-lactamases, relies on a metal cofactor. Of these, class A β-lactamases exhibit a highly diverse substrate profile and are most commonly encountered in clinical isolates (Tomanicek *et al.*, 2010). The following section details the structure and mechanisms of class A β-lactamases.

Class A β-lactamases are composed of two distinct domains: an α domain and an α/β domain, with the active site situated in a groove between the two domains. The active site pocket, in turn, consists of four motifs (Figure 49). The first motif, Ser^{70}-X-X-Lys (X representing any amino acid), includes the active serine, which is located on the first helix of the α domain (α-2). Due to the helical structure of this motif, the Ser^{70} and Lys^{73} residues are located inside the active cavity, whereas the two variable residues in between point away from the active site. The second motif, Ser^{130}-Asp-Asn, is located on a loop in the α domain. The side chains of the Ser and Asn make up one side of the active site, whereas the Asp residue points toward the core of the protein. The third domain, Lys^{234}-Thr/Ser-Gly, is located on the inner most strand of the β sheet and forms the opposite side of the active-site cavity. The Gly residue is strictly conserved in class A enzymes, since any side chain at this position would protrude in the active site and sterically hinder the interaction of substrates with the active serine (Galleni *et al.*, 1995). The final motif, known as the Ω-loop (residues 164 to 179), is a flexible structure, which controls access to the active site pocket. Interestingly, PBPs share a level of structural and functional similarity with β-lactamases, which suggests that they are derived from the same ancestor. The difference between these two enzymes is that PBPs do not contain the Ω-loop motif and their deacylation step is much slower than that of β-lactamases. These differences provide evidence for the theory that PBPs have evolved into class A β-lactamases, with the most significant step of the evolutionary pathway being the formation of the Ω-loop with the essential Glu^{166}, thus ensuring an efficient deacylation (Lobkovsky *et al.*, 1994). Consequently, β-lactam antibiotics are substrates of β-lactamases but can be viewed as mechanism-based (or suicide) inhibitors of the PBPs (Vanwetswinkel *et al.*, 2000).

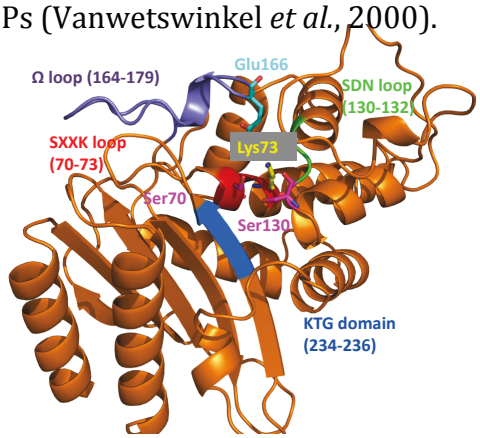


Figure 49. The 3D structure of the class A β-lactamase from *E coli* TEM-1 (pdb 1BTL).

The Ser-X-X-Lys motif is indicated in red, the Ser-Asp-Asn motif indicated in green, the Lys-Ser/Thr-Gly motif indicated in blue, and finally the Ω-loop indicated in purple. The two serines are shown in magenta. The figure was generated using PyMOL.

As illustrated in figure 50, several residues in the active site pocket and the surrounding environment play important roles in the efficiency and specificity of substrate binding.

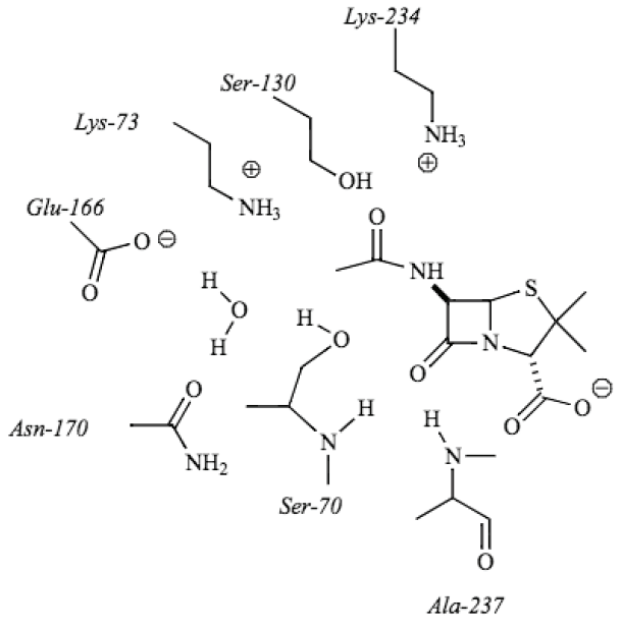


Figure 50. Two-dimensional schematic representation of active-site residues, the conserved hydrolytic water and the penicillin substrate.

Source: Fisher and Mobashery, 2009.

Similar to other serine hydrolases, the oxyanion hole formed by the tetrahedral intermediate needs to be stabilized for the reaction to occur. In class A β-lactamases, it is the residues Ser^{70} and Ala^{237}, which form an oxyanion hole responsible for the polarization of the β-lactam carbonyl group in order to stabilize the transition state (Drawz and Bonomo, 2010). The Lys^{234} and Lys^{73} residues stabilize Ser^{130} through hydrogen-bonding interactions. Ser^{130} in turn interacts with the carboxyl group to ensure substrate binding by the stabilization of the carboxylic acid (Matagne and Frère, 1995). Residue Asn^{132} interacts *via* hydrogen bonding with the amide group of the substrate in order to correctly position the substrate in the active-site pocket (Lamotte-Brasseur *et al.*, 1991; Galleni *et al.*, 1995). There are numerous other reports underlining the importance of several other residues in substrate binding. The Asp^{131} residue has been shown to be responsible for the stabilization of the active-site structure *via* the transfer of a proton (Jacob *et al*, 1990). The hydroxyl group of the residue at position 235, either a Ser or a Thr, was shown to play a role in substrate specificity (Dubus *et al.*, 1994). Finally, there are reports demonstrating the role of the Arg^{244} residue as an electrostatic catalyst in the active-site cavity (Maveyraud *et al.*, 1998). More recently, the group of Hoshino has demonstrated that the residues Ser^{130}, Asn^{132}, Ser^{235}, Gly^{237}, and Arg^{244} are all involved in cooperatively restricting the mobility of the substrate by making salt bridges with the carboxyl and amide groups of substrates (Hata *et al.*, 2006).

The general mechanism of class A β-lactamase enzymes involves a typical serine hydrolase activity with the following steps: (*i)* activation of a serine by a base to form the corresponding alcoholate, *(ii)* nucleophilic attack by the alcoholate on the electrophilic carbon of the amide bond (the carbonyl) and formation of a tetrahedral intermediate (general base catalysis), *(iii)* collapse of the tetrahedral intermediate by the leaving of the amino group via N-protonation and formation of the acylated serine (general acid catalysis), and finally *(iv)* nucleophilic attack of a water molecule on the carbonyl of the serine ester and release of the serine (Fisher and Mobashery, 2009) (Figure 51).

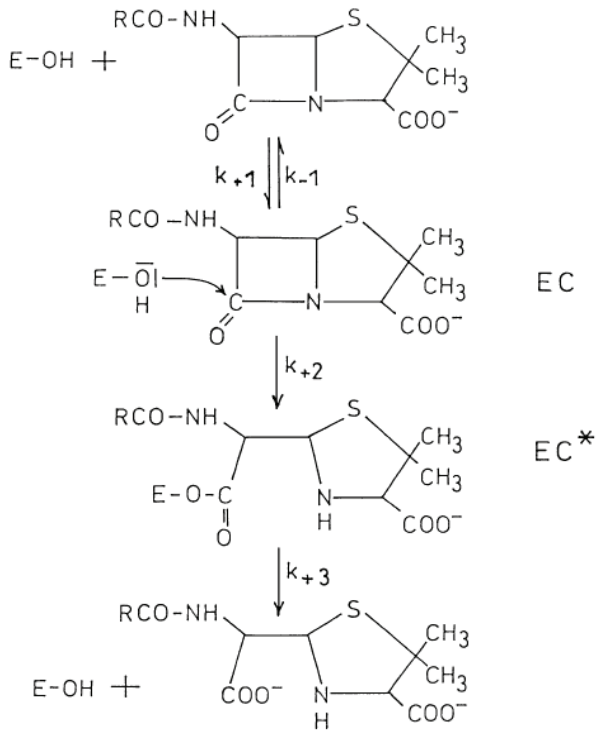


Figure 51. General catalytic pathway of serine β-lactamases.

EC represents the non-covalent acyl enzyme, EC* represents the covalent acyl-enzyme. The catalytic constant of the acylation reaction is represented by k_{+2} and that of the deacylation reaction is represented by k_{+3}. Source: Matagne *et al.*, 1998.

In the active site of these enzymes, the nucleophilic attack is performed by Ser^{70} on the carbonyl carbon of the β-lactam ring. This serine is able to act as a nucleophile due to activation by a general base, resulting in an acyl-enzyme intermediate (Matagne *et al.*, 1998). There are controversies, however, as to which of the highly conserved residues, Glu^{166} or Lys^{73}, is the catalytic base in the acylation mechanism of the enzyme (Figure 52).

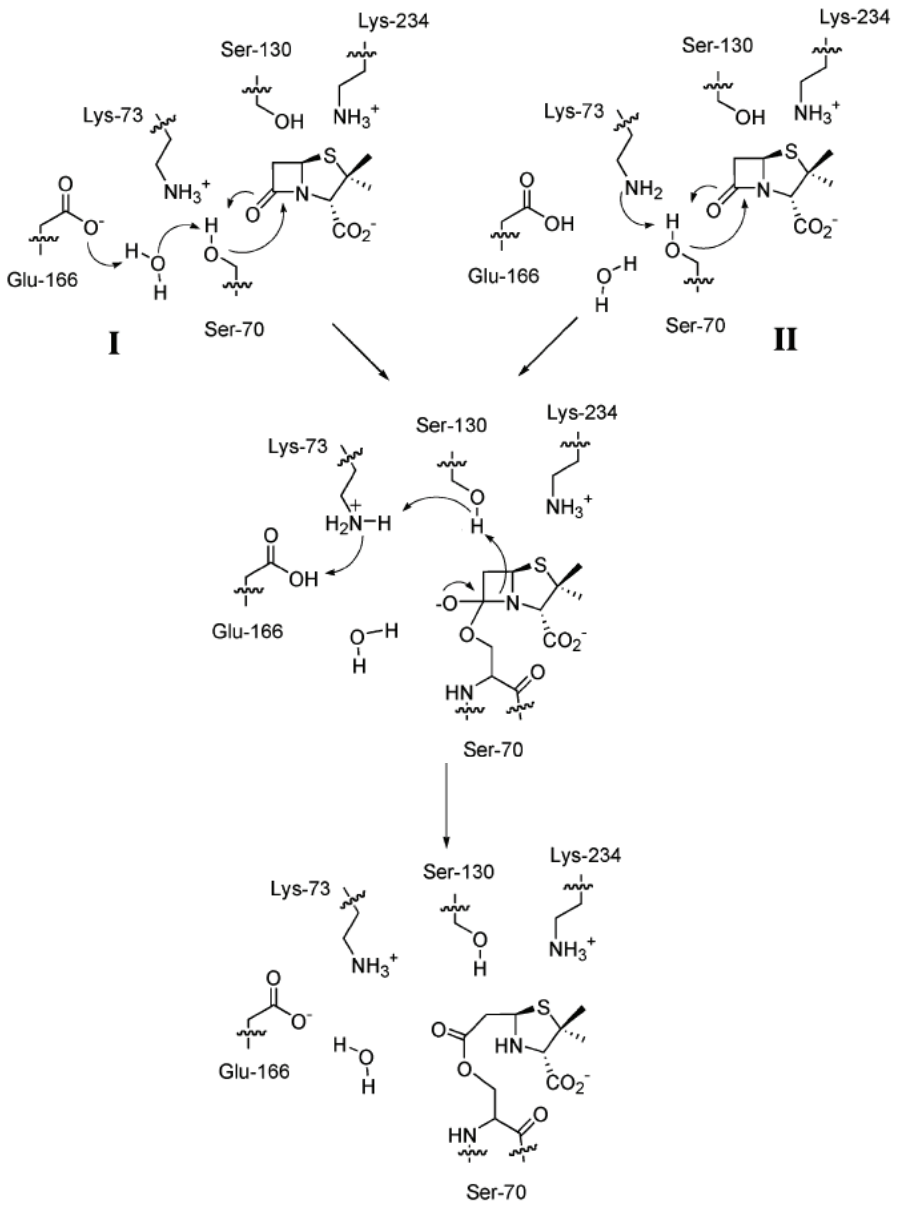


Figure 52. The two debated pathways for the acylation step of the β-lactamase reaction.

I represents the pathway with Glu[166] as the general base and II demonstrates the pathway with Lys[73] as the general base. Source: Meroueh *et al.*, 2005.

Studies considering pH effects on enzyme activity and the observation of proton positions in the active-site region from ultra high resolution X-ray structures, including the work of J.M. Frère and colleagues, propose a symmetric mechanism in which Glu[166] acts as the general base in both acylation and deacylation reactions (Matagne and Frère, 1995; Minasov *et al.*, 2002; Tomanicek *et al.*, 2010). Other studies based on mutagenesis and structural evidence, including the work of N. Strynadka, suggest an asymmetric mechanism in which the NH_2 extremity of the side chain of Lys[73] plays the role of the general base in the acylation step (Strynadka *et al.*, 1992; Golemi-Kotra *et al.*, 2004). The final deacylation step for the release of the free Ser residue has historically been less controversial. In this

step, the abstracted proton is back-donated to the nitrogen of the substrate β-lactam ring through a proton transfer between Glu[166], Ser[130], and finally the substrate *via* a water molecule (Matagne *et al.*, 1998).

More recently, however, reports have demonstrated evidence of the involvement of both residues Glu[166] and Lys[73] in both acylation and deacylation mechanisms. The work of Fink and colleagues, for example, has suggested that both of these residues are important for maintaining the electrostatic environment necessary for efficient catalytic activity and the mechanism promoted by the enzyme (the use of Glu *versus* Lys as the catalytic base) is dependent on the nature of the substrate (Lietz *et al.*, 2000). Mobashery and his group have also reported a duality of the acylation mechanism involving both residues, *via* quantum mechanical/molecular mechanical (QM/MM) calculations (Meroueh *et al.*, 2005). This suggested mechanism involves the transfer of proton from Lys[73] to Glu[166], through a catalytic water molecule and Ser[70], making both residues deprotonated. Tetrahedral formation then follows with both residues acting as a general base in competition for the promotion of the nucleophilic attack of Ser[70] (Figure 51). Hoshino and colleagues have suggested that Lys[73] enhances the base activity of Glu[166] by reducing the ionic interaction between the two residues *via* a concerted double proton transfer from Lys[73] to Ser[130] and from Ser[130] to the substrate carboxylate (Hata, *et al.*, 2006). Furthermore, their study proposes Lys[73] as the residue responsible for the deacylation mechanism, where this residue donates a proton to the Ser[70] leaving group *via* a proton back transfer from Ser[130] to Lys[73]. What is clear and irrefutable is that all active-site residues discussed above interact with each other and water molecules to form a dense network of hydrogen bonds, crucial for the catalytic reaction (Figure 53).

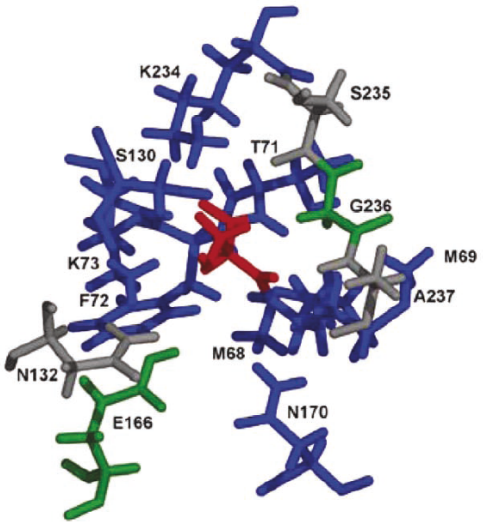


Figure 53. Stereoview of the active site of TEM-1 β-lactamase.

The Ser[70] residue is shown in red. Source: Savard and Gagné, 2006.

As a result of the wider clinical exposition spectrum of third generation antibiotics, such as oxyimino-cephalosporins, class A extended-spectrum β-lactamases (ESBLs) have come to existence (Tomanicek *et al.*, 2010). ESBLs are sub-categorized into two subgroups, with the first group including the more studied penicillinases TEM-1 (Drawz and Bonomo, 2010). The residues responsible for ESBL activity are generally near the substrate-binding site but do not directly contribute to catalysis. These include residues at positions 104, 238, and 240. Point mutation studies on these residues have demonstrated an increase in ESBL activity by providing modifications such as an expansion of the active site cavity in order to accommodate larger substrates or an improvement in the formation of electrostatic bonds with these substrates (Sowek *et al.*, 1991; Petit *et al*, 1995; Cantu and Palzkill, 1998). Interestingly, the acquisition of resistance toward the extended spectrum antibiotics provided by these mutations has led to a concomitant decrease of resistance toward the classical substrates and first generation antibiotics. For example, point mutation studies on a number of different residues of the TEM-1 β-lactamase lead to an increase of up to 1000-fold of the hydrolysis of Cefotaxime (CTX), an extended spectrum cephalosporin, and a concomitant decrease of up to 100-fold of the hydrolysis of benzylpenicillin (BZ) (De Wals *et al.*, 2009). These studies have demonstrated a certain level of structural plasticity of the active site of β-lactamases, providing a strong tolerance to different mutations and allowing for substrate spectrum diversification. More recently, a mutation on the Glu^{166} residue (E166Y) was shown to retain activity by using the replaced Tyr residue as the general base (Stojanoski *et al.* 2015). These results provided further evidence for the robustness of the β-lactamase fold and the plasticity of the active site pocket.

In addition to active site plasticity, the conformational dynamics of the ESBL active site has also been studied. In a study performed by Nicolas Doucet and Joelle Pelletier, a molecular dynamics simulated annealing experiment demonstrated that the Tyr^{105} residue, in the free TEM-1 enzyme, holds two different conformations simultaneously in solution, each binding preferentially a different substrate (Doucet and Pelletier, 2007). As this residue plays an important role in substrate specificity (Doucet *et al.*, 2004), these results provide additional explanation as to how conformational dynamics of the ESBL enzymes provides them with binding specificity for such a wide spectrum of substrates. Thus, various sequential or structural motifs held by these enzymes seem to be able to support the catalytic function.

As mentioned previously, the β-lactamase activity has been used by our group as a model activity for the study of catalytic antibodies. Using the β-lactamase enzyme and taking advantage of the idiotypic network, an antibody was produced endowed with this model activity. The following is a description of the generation and characterization of this antibody.

II.2.B. The catalytic antibody 9G4H9

As one of the pioneering studies in the field, β-lactamase from the bacteria *Bacillus cereus* and the anti-idiotypic network strategy were used in the 1990s by our group to generate a

catalytic antibody with the model activity (Avalle *et al.*, 1998). This was performed by the immunization of BALB/C mice with the enzyme, preparing monoclonal hybridoma cell lines, and screening for antibodies raised against the active site of the enzyme. The Ab1 antibody 7AF9 was isolated by this method, proving its specificity to the active site by significantly reducing the catalytic efficiency of the enzyme in a number of kinetic and inhibition tests. In the second immunization step, following the anti-idiotypic network, Biozzi mice were immunized by the Ab1 7AF9 and Ab2 hybridoma clones were tested for their affinity to 7AF9 $F(ab')_2$ as well as their potential catalytic activity. The Ab2 9G4H9 was able to hydrolyze the β-lactam rings of both a penicillin substrate (ampicillin) and a cephalosporin substrate (PADAC), with catalytic constants of 2.1 min^{-1} and 2.3×10^{-3} min^{-1}, respectively (Padiolleau-Lefèvre *et al.*, 2003).

It was later described that 9G4H9 functions both as a β-lactamase by breaking the cyclic amide bond of β-lactams (Avalle *et al.*, 2000), and a PBP since it is rapidly inactivated by some β-lactam molecules (Lefèvre *et al.*, 2001). This behavior of 9G4H9 pointing to its weak deacylation capability is potentially the reason for the discrepancies between the catalytic parameters (using benzylpenicillin as substrate) of this antibody (k_{cat} = 2.1 min^{-1}, K_M = 4300 μM) and the β-lactamase enzyme (k_{cat} = 7.3×10^4 min^{-1}, K_M = 656 μM) (Padiolleau-Lefèvre *et al.*, 2003). The variable regions V_L and V_H of 9G4H9 were sequenced (Figure 54), and interestingly did not display any significant homology to the primary structure of the model enzyme (Débat *et al.*, 2001). The formation of a stable complex between 9G4H9 and benzylpenicillin demonstrated by Western blotting revealed the involvement of the light chain of the antibody (Lefèvre *et al.*, 2001). However, the heavy chain influences the maintenance of this activity, possibly by affecting the overall structure of the light chain, since the light chain alone (without the presence of the heavy chain) did not exhibit any activity (Hélène Débat, personal communication).

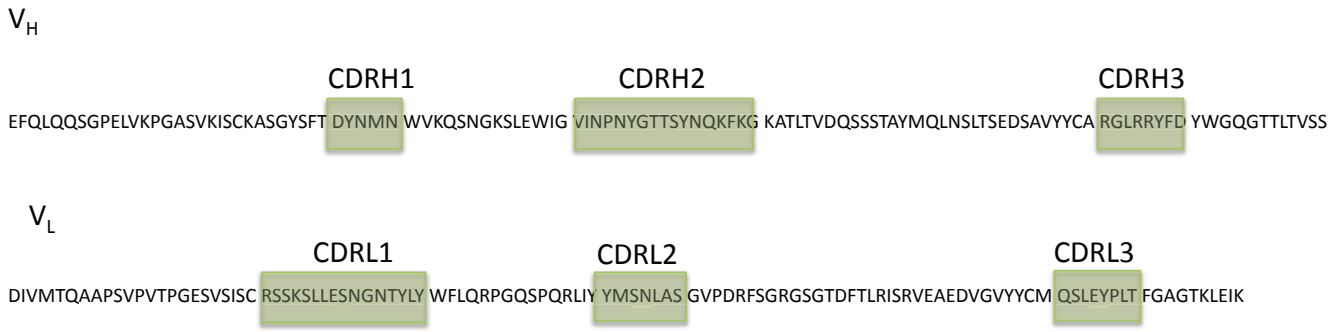


Figure 54. Amino acid sequence of antibody 9G4H9 variable regions.

(EMBL reference: aj277812 VL; aj277813 VH).

Three-dimensional structural modeling of this light chain showed a certain level of structural homology with both β-lactamase and PBP (Padiolleau-Lefèvre *et al.*, 2003). The residues Ser^{26}, Ser^{28}, and Lys^{27} in the light chain CDR1 and Glu^{98} and Gln^{95} in CDR3 were identified as potentially belonging to the active site (Figure 55). A number of observations were drawn from the structural positioning of these residues. First, it was observed that contrary to the native enzyme, there was no residue available to induce the deprotonation of Lys^{27} leading to a decrease in nucleophilicity of Ser^{26} or Ser^{28}, which is a potential explanation for the lower catalytic efficiency of the abzyme. Second, the positioning of the two serine residues and their average distance to the deacylating group (Glu^{98} and Gln^{95}) potentially explains the suicide substrate behavior of 9G4H9. The binding of the substrate on Ser^{26} might prevent the attack by Ser^{28} and therefore rapidly lead to a decrease of hydrolysis.

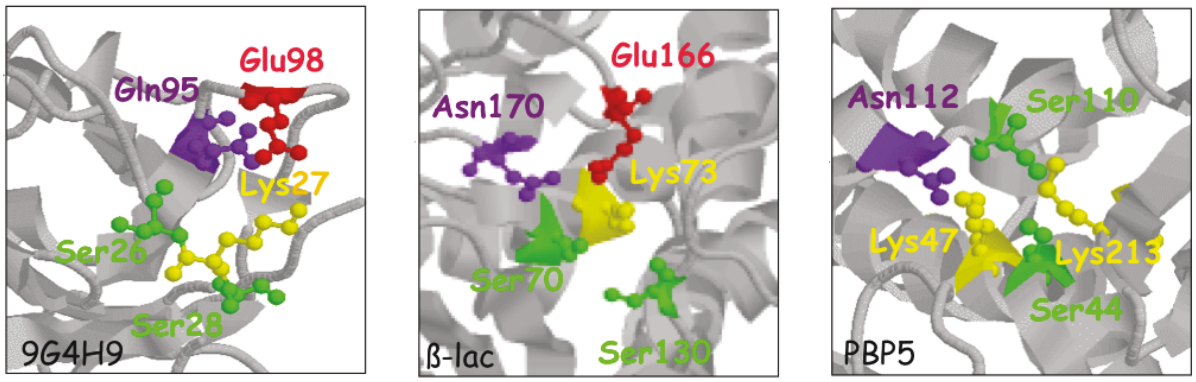


Figure 55. Structural comparison of the active sites of κ-light chain of 9G4H9, β-lactamase, and DD peptidase by molecular modeling.

Source: Padiolleau-Lefèvre *et al.*, 2003

Later in 2006, the single chain variable fragment (scFv) format of 9G4H9 was cloned in a V_L-linker-V_H orientation using a linker described by Whitlow *et al.* (Padiolleau-Lefèvre *et al.*, 2006). The efficient periplasmic expression and correct folding of the scFv by using the chaperone FkpA demonstrated similar catalytic efficiency as compared to the whole antibody format of 9G4H9. Furthermore, structural modeling of the scFv and its superposition on β-lactamase active site seemed to confirm the involvement of the residues previously described. The superposition of Lys^{27} and Glu^{166} from β-lactamase with Lys^{27} and Glu^{98} of 9G4H9 light chain, as well as Ser^{70} and Ser^{130} of β-lactamase with Ser^{26} and Ser^{28} of 9G4H9 light chain revealed a similar positioning of these residues. Finally, the observation of the proximity of Arg^{24} to the putative active site suggested a possible role of this residue in the catalytic mechanism of 9G4H9 (Figure 56).

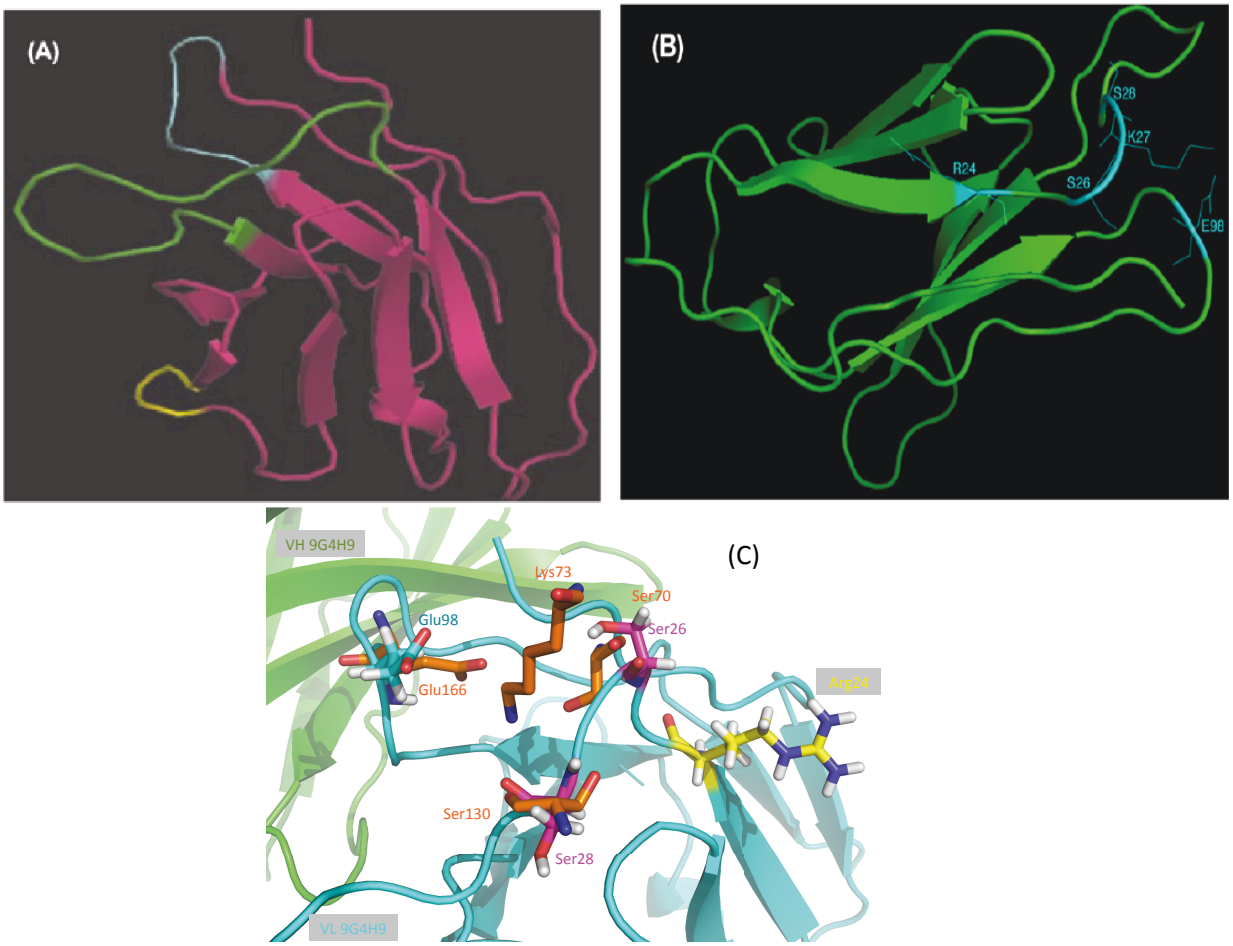


Figure 56. 3D modeling of 9G4H9.

A) The 3D model of 9G4H9 light chain showing CDR1 in green, CDR2 in yellow and CDR3 in cyan. B) scFv9G4H9 light chain backbone with the catalytic residues. Source: Phichith *et al.*, 2009. C) Superposition of scFv9G4H9 and the enzyme β-lactamase active site (in orange). The numbering attributed to these models is sequential (different from that of IMGT).

Moreover, the flexibility of the Ω-loop protects the β-lactamase active site, which is buried in to a cavity. This is comparable to the light chain CDR3 structural properties of the scFv 9G4H9. However, the potential catalytic residues of the scFv are exposed to the solvent, which might be another reason for its lower catalytic efficiency. Lastly, a point mutation study was performed on the scFv 9G4H9 in order to experimentally confirm the involvement of the suspected residues in the catalytic mechanism. The global structure of the mutated scFv was analyzed by circular dichroism showing no modification in the tridimensional configuration of each mutant. The authors thus confirmed the implication of Ser^{26}, Ser^{28}, and Glu^{98}, but they have also demonstrated that Arg^{24}, instead of Lys^{27}, is implicated in the catalytic activity. Indeed, it was demonstrated that alternatively replacing each of these residues with Ala, with the exception of Lys^{27}, led to a loss of activity, confirming the participation of Ser^{26}, Ser^{28}, Arg^{24}, and Glu^{98} in the catalytic process as either nucleophile or general acid/base catalyst (Phichith *et al.*, 2009).

The mutation results revealed, however, that contrary to the potential role of Lys[73] in the acylation mechanism of β-lactamase, the Lys[27] residue of 9G4H9 couldn't partake this role. It was suggested, therefore, that due the positioning of Arg[24] within hydrogen bonding distance of the serines and its crucial implication in catalysis revealed by site-directed mutagenesis, the role of the general base could potentially be attributed to this residue. The role of Arginine acting as a general base has been previously reported in several enzymes (Guillén Schlippe and Hedstrom, 2005) and an amidase catalytic antibody (Benedetti *et al.*, 2004). The microenvironment of this residue in the active site of 9G4H9 (i.e. its contact with a hydrophilic and charged substrate and the solvent) could potentially alter the pKa of this residue and lead to its suggested function.

Finally, the remote position of Glu[98] from the putative nucleophilic serine (Ser[26] or Ser[28]) is potentially the reason for the limiting deacylation rate of 9G4H9 and its inefficiency in hydrolyzing the acyl-enzyme complex. The scFv 9G4H9 light chain and its superposition with the β-lactamase active site are shown in figure 55D. The kinetic parameters were further measured for β-lactamase and different formats of 9G4H9, using ampicillin (Padiolleau-Lefèvre *et al.*, 2006) and fluorocillin (Ben Naya *et al.*, 2013), a fluorogenic β-lactam substrate with higher detection sensitivity than ampicillin (described in more detail below). The structures of these substrates along with PADAC are shown in figure 57. A comparison between the catalytic efficiencies of β-lactamase and different formats of 9G4H9 is resumed in table 8.

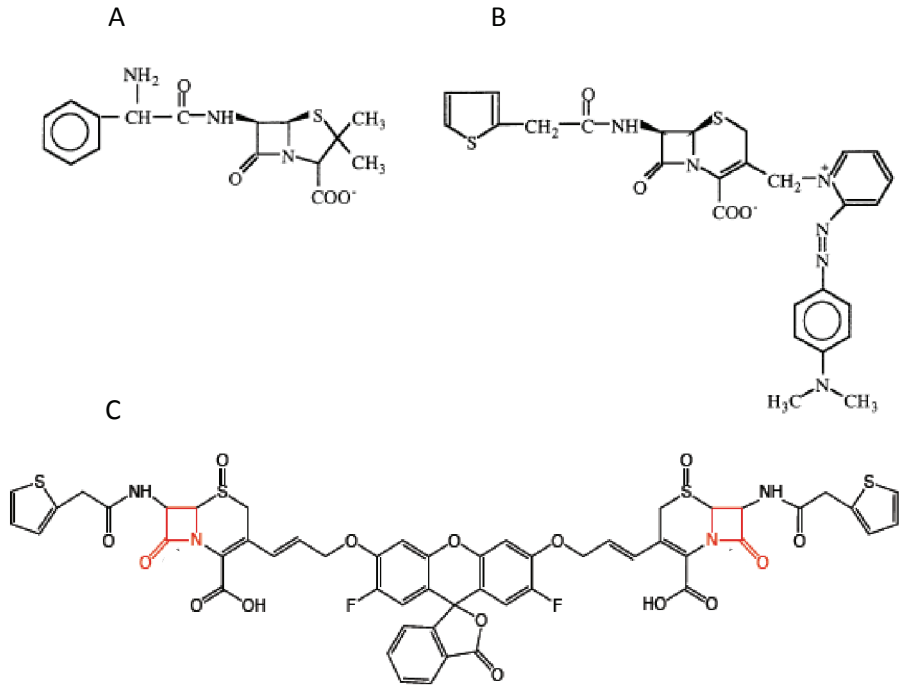


Figure 57. Chemical structures of β-lactam substrates.

A) Ampicillin, B) PADAC, and C) Fluorocillin.

Table 8. Kinetic parameters of different catalyst with β-lactamase activity.

Catalyst	Substrate	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($M^{-1}.min^{-1}$)	k_{cat}/k_{uncat}
β lactamase	Ampicillin	665	7.3×10^4	1.1×10^8	1.3×10^9
IgG-9G4H9		4300	2.3±0.2	4.9×10^2	3.8×10^4
scFv-9G4H9		2300±600	3.3±0.2	1.4×10^3	5.9×10^4
β lactamase	Fluorocillin	-	-	$(2.017\pm0.007)\times10^5$	-
scFv-9G4H9		-	-	$(1.01\pm0.01)\times10^3$	-

Soluble scFv 9G4H9 has been produced by periplasmic expression.
Sources: Padiolleau-Lefèvre *et al.*, 2006; Phichith *et al.*, 2009; Ben Naya *et al.*, 2013

To further investigate the catalytic mechanism of 9G4H9, the phage display technology was used in our laboratory in order to select for a peptide inhibitor of this antibody. This process and the characteristics of this peptide are described below.

II.2.C. Pep90: a cyclic peptide inhibitor of 9G4H9

In an attempt to elucidate further structure-function relationships associated with β-lactamase activity and to investigate mechanisms governing the transfer of catalytic information throughout the idiotypic network, our group has selected a peptide inhibitor specific for the activity of 9G4H9 (Yribarren *et al.*, 2003). This peptide named Pep90, was isolated from a random library of 4.5×10^9 cyclic heptapeptides displayed on the surface of bacteriophage by a two step selection procedure. It binds the antibody 9G4H9 with K_D=74±16 nM and inhibits its catalytic activity with IC_{50}=90 μM. The sequence of this inhibitory peptide Tyr-His-Phe-Leu-Trp-Gly-Pro boarded by two cysteines is mainly composed of central hydrophobic residues, similar to most of the other selected peptide candidates, and contains a proline in the C-terminal. This commonality has suggested the requirement of a basic structural feature able to bind the active site of 9G4H9.

Pep90 also inhibits the scFv format of 9G4H9 with IC_{50}=482 μM. Interestingly, the point mutation study, mentioned previously, revealed that the replacement of Lys^{27}, which did not affect the catalytic capability of scFv 9G4H9, does not alter the inhibitory capability of Pep90. This confirmed that this residue is not implicated in the β-lactamase activity of scFv 9G4H9 (Phichith *et al.*, 2009). Furthermore, site-directed mutagenesis on the peptide Pep90 revealed that the Tryptophan residue has a critical role in the interaction between the peptide and scFv 9G4H9, and the Proline residue ensures the necessary rigidity of the peptide. A different study on the inhibitory abilities of Pep90 demonstrated that this peptide not only inhibits 9G4H9, but also binds to and inhibits TEM-1 β-lactamases and several PBPs from various bacterial species with IC_{50} values ranging from 6 to 100 μM (Phichith *et al.*, 2010). The particular ability of Pep90 to inhibit both β-lactamases and PBPs

could potentially open new avenues for this peptide as a potential candidate for the development of antibiotics. A possible structure of this peptide and a modelization of its interaction with scFv 9G4H9 and TEM-1 β-lactamase are shown in figure 58.

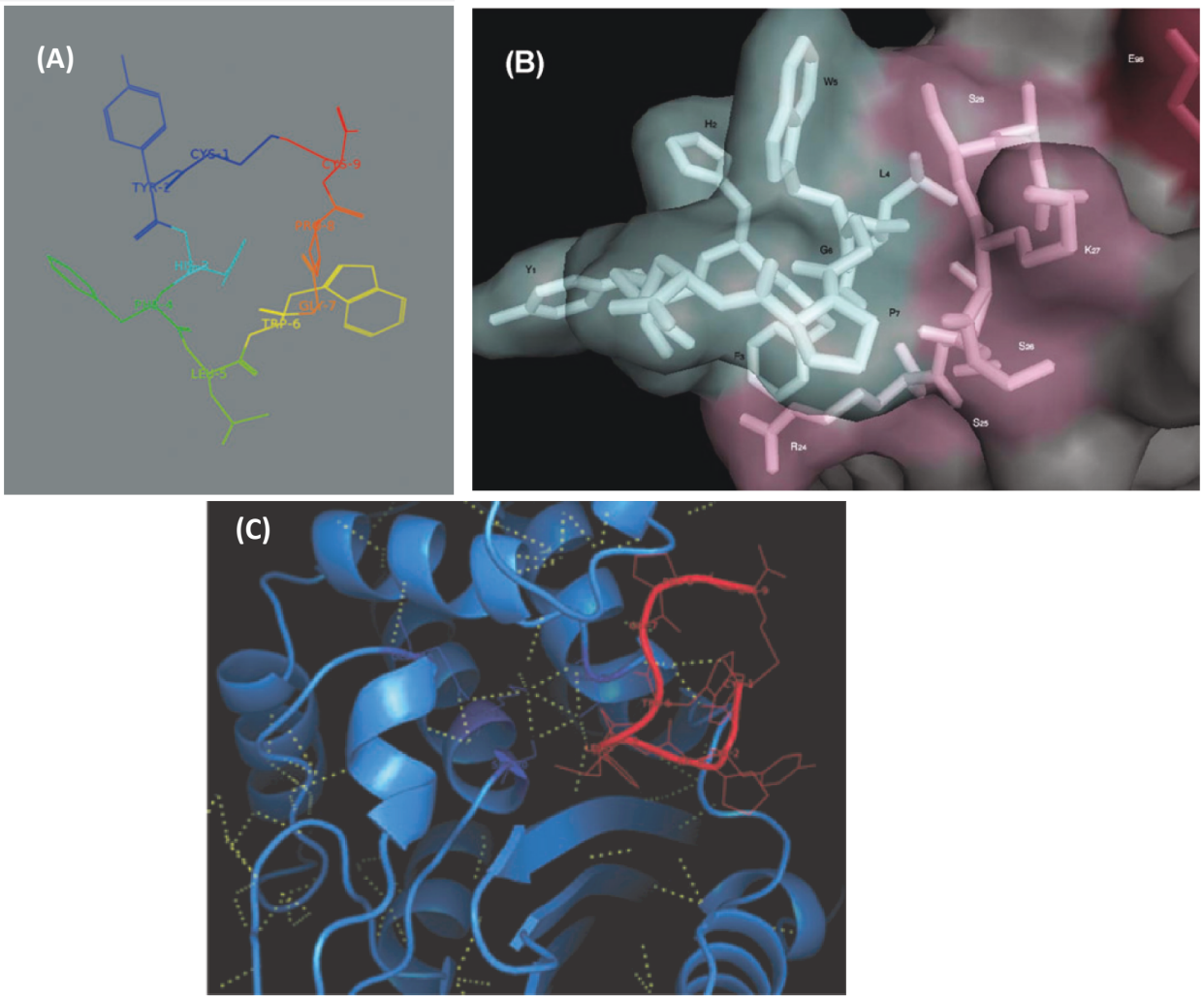


Figure 58. Structural modelization of Pep90.

A) 3D structure of Pep90 B) 3D interaction model of Pep90 (blue) docked within scFv9G4H9 active site ($R^{24}SSKS^{28}$ in pink and E^{98} in red. C) 3D interaction model of Pep90 (red) in the active site of β-lactamase (blue). Source: Phichith *et al.*, 2009; Phichith *et al.*, 2010.

As mentioned previously in the manuscript, one of the biological techniques for the production of catalytic antibodies is the utilization of protein inhibitors of enzymes in order to recruit antibodies that are internal images of these enzymes. Following this methodology, the peptide inhibitor Pep90 could potentially be an appropriate candidate for the isolation of antibodies displaying a functional mimicry of the enzyme β-lactamase. This is why, in the current work, we have taken advantage of the large diversity of our constructed library and have used Pep90 as one of the targets to select for antibodies from

this library, which are endowed with β-lactamase activity. Our methodology for selection is the phage display technique, which is introduced herein.

II.3. Selection of target antibodies by phage display

Phage display is the most widely used selection technique for the isolation of antibodies against any desirable target. The biology of phage and the procedures involved in this technique have been previously described (I.2.A.). The phage display technology takes advantage of the direct linkage of the genotype of the phage to its phenotype, allowing for facile and rapid identification of the gene encoding for the protein of interest displayed on the surface of phage. The selection procedure consists of several different steps, which are detailed below.

II.3.A. Selection procedure

The *in vitro* selection process of antibodies from phage display libraries based on target affinity is termed "panning" (Kretzschmar and von Rüden, 2002). It consists of several cycles of binding, washing, and elution, leading to the enrichment of target-specific antibodies (Figure 59) in the library.

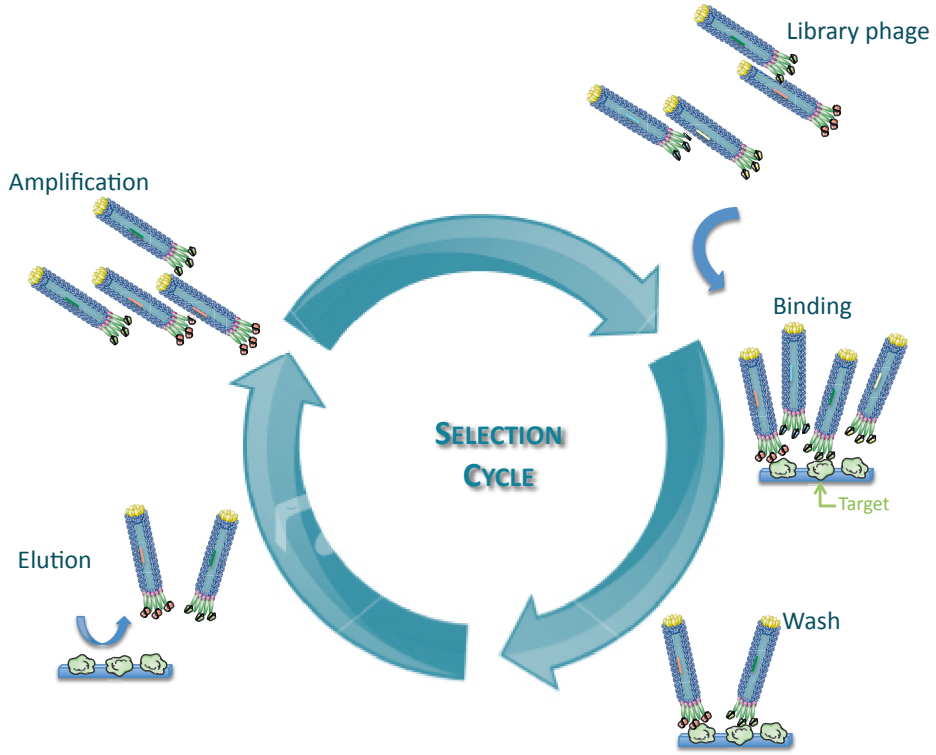


Figure 59. Selection cycle of phage display library against a specific target.

The cycle includes the following steps: binding, wash, elution, and amplification.

Common support systems used for selection are nitrocellulose, magnetic beads, agarose columns, monolithic columns, polystyrene tubes and 96-well microtiter plates (Kontermann and Dubel, 2010). The panning includes the following steps: the target is coated or captured on the support surface and is then incubated with phages from the library. This incubation will allow the binding of phage-displayed entities to the target. Subsequently, the unbound phages are eliminated by a number of washing steps. Finally, an elution step is performed either by enzymatic digestion, reduction, competition with soluble target, or pH shift. The eluted phage particles are then reamplified and the cycle is repeated.

II.3.B. Applications

The phage display techology offers a rapid and efficient method to obtain large amounts of proteins or peptides for a variety of purposes. It allows the investigation of protein-protein interactions, receptor binding sites, and antigenic epitopes as routine procedures. Additionally, when antibody libraries are used, this technique provides the advantage of eliminating the necessity of immunization, allowing for the production of antibodies against any non-immunogenic as well as toxic targets. The production of such antibodies is not possible by *in vivo* antibody production techniques. It is therefore not surprising that phage display has become a valuable tool in numerous biomedical applications.

Among the growing applications of phage display in the medical field is the production of antibodies in hematological therapeutics (Marks *et al.*, 1993). These antibodies against red blood cell antigens are used for hemagglutination assays and are needed in large amounts. Therefore, phage display has been the method of choice to produce large amounts of these antibodies in a short time. Moreover, phage display has been used to produce antibodies against the white blood cells such as B and T cells such as (Maeda *et al.* 2009), or dendritic cells (Fitting *et al.*, 2011), and against paraproteins (O'Nuallain *et al.*, 2007).

The ability of phage display to represent human immune libraries has led to the application of this technology in the study of autoimmune disease. To date, many autoimmune diseases have been investigated by phage display for the design of potential diagnostic and therapeutic agents. In the case of rheumatoid arthritis, for example, this technology has enabled the development of a new drug based on anti-TNFα antibodies (Rau, 2002). Another example is the anti-CD52 phage display produced antibody for the treatment of multiple sclerosis (Klotz *et al.*, 2012). Antibodies for autoimmune disorders generated by phage display are increasingly introduced to the market.

Another major advantage of phage display is the possibility for the directed evolution of antibodies for increased stability and higher affinity. This technology has also been made important contributions to vaccine development, molecular imaging and tumor development (Bazan *et al.*, 2012). Furthermore, phage display has been used in combination with NGS to mimic and follow *in vitro* the effect of clonal selection occurring *in*

vivo (Ravn *et al.*, 2010). Due to the high flexibility of this technique, it is likely that phage display will remain useful to a broad spectrum of therapeutic and diagnostic applications.

Finally, phage display has also been used for the study of catalytic antibodies. The ease of the manipulation of this technique has lead to the facile screening methods of antibody libraries for the selection of candidates with catalytic activity. This has been done for example by the immunization of animals with a reactive TSA and selecting for the corresponding activity (Tanaka *et al.*, 1999) or by the selection from a library derived from the peripheral blood lymphocytes of patients with a particular pathology (Gabibov *et al.*, 2011).

In the present study, we have taken advantage of the phage display technique and have used two different strategically chosen targets in order to isolate, from our constructed library, antibodies with potential β-lactamase activity. This strategy and the acquired results are described below.

II.4. Selection of catalytic antibodies with β-lactamase activity

In order to study the expression and perform a characterization of catalytic antibodies, we focused on a β-lactamase model activity. We have therefore chosen two strategic targets in order to select for antibodies with β-lactamase activity: *i)* the cyclic peptide Pep90, and *ii)* the penam sulfone structure used as the antigen for immunization. As previously described (sesction II.2) both of these molecules are inhibitors of the enzyme β-lactamase.

- Pep90 has been selected against the idiotope of 9G4H9, a catalytic antibody with β-lactamase activity, and therefore, is potentially the structural mirror image of the active site of the enzyme. As a result, antibodies selected against this peptide can potentially mimic the active site of the enzyme and be endowed with β-lactamase activity.

- Penam sulfone, on the other hand, has been used for the reactive immunization of mice in order to deliberately induce the production of catalytic antibodies with β-lactamase activity. It is a suicide inhibitor of β-lactamase, and therefore, antibodies with affinity or reactivity toward this molecule can also potentially mimic the enzyme active site and be endowed with the same activity.

Both targets are biotinylated and the support structure of the selection method consists of magnetic beads conjugated with streptavidin molecules. The procedure is as follows: first, the target is immobilized on the beads via the biotin/streptavidin interaction, with a sufficient concentration in order to ensure the saturation of the beads (Figure 60). The saturation of the beads by the target is necessary in order to reduce the exposure of the undesirable sites of the support structure to the library phage. Second, the pooled library is

put in contact with the target and several rounds of panning are performed in order to enrich the pool of high affinity binders. Finally, a sufficient number of the selected phage is screened in order to isolate positive clones, exhibiting a high affinity toward the target.

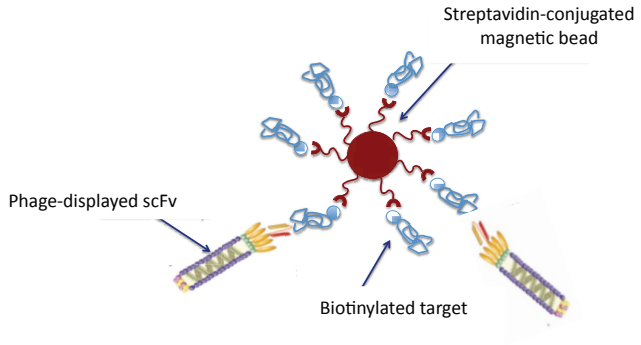


Figure 60. Scheme of the support structure for the selection procedure.

II.4.A. Selection against Pep90

The cyclic peptide Pep90 used for selection has been synthesized by the company Eurogentec (Liège, Belgium). It has a molecular weight of 2061.5 g.mol^{-1} and the following N-terminal to C-terminal amino acid sequence (underlined) : NH_2 - GAC (S-) <u>YHFLWGPC</u> (S-) *GAAGAEK* (Biotin) - $CONH_2$, where the sites of the disulfide bond leading to the circularization are indicated by (S-), and the spacer providing a level of flexibility between the cyclic peptide and the biotin is indicated in italic . We have performed an optical density spectral scan for Pep90 and have determined maximal absorption to be at wavelength of 210 nm. We have also produced a calibration curve with different dilutions of the peptide in water and determined its extinction coefficient to be ε = 221 425 L.mol^{-1}.cm^{-1} (Figure 61).

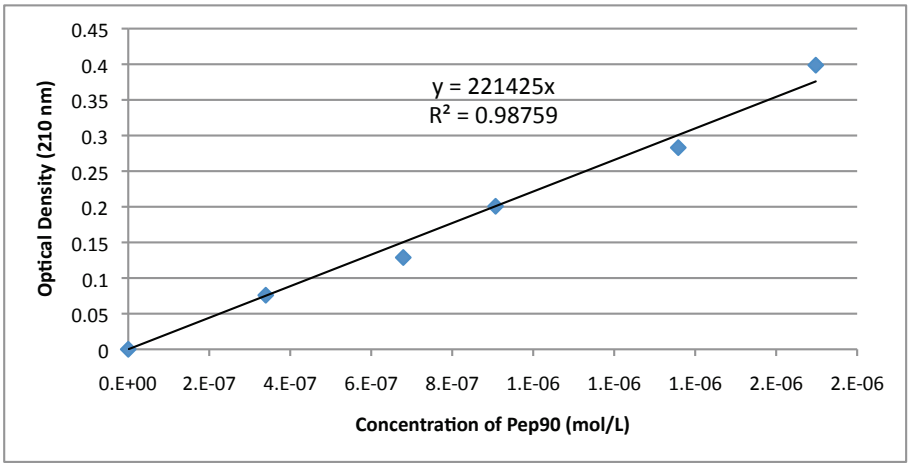


Figure 61. Calibration curve of Pep90 for the determination of the extinction coefficient performed in PBS (pH 7.4).

As shown by the linearization formula ε = 221425 L.mol^{-1}.cm^{-1}. The R^2 value shows a good estimation for the linearization curve.

In order to determine the quantity of the peptide necessary for saturation of the beads, we have placed different dilutions of Pep90 in contact with 500 µg of the magnetic beads, measured the optical density of the unbound fraction, and hence deduced the maximal quantity of the immobilized peptide on the beads. The results are not coherent and we do not observe the expected behavior of saturation of the beads with increased concentration of the peptide (Table 9).

Table 9. Saturation experiment of the magnetic beads with Pep90.

Sample	Quantity (µg)	Concentration (nmol/mL)	OD (original)	OD (eluted)	ΔOD	Immobilized peptide (pmol)
1	4	1.36	0.35	0.38	-	-
2	3	0.92	0.31	0.25	0.06	436
3	2	0.68	0.36	0.22	0.15	964
4	1	0.33	0.14	0.07	0.07	449

Different dilutions from a stock solution of 2.8 mg/mL are put in contact with 200 µg magnetic beads for 2 hours at room temperature. From the obtained results, the capacity of the beads is determined to be 2000 pmol of peptide per 1 mg of beads.

This could be due to a number of factors such as the insensitivity of the spectrophotometer, or the fact that we might be detecting other species in the eluted fractions such as portion of the beads broken off due to extensive vortexing. Nevertheless, we consider the capacity of the beads at 2000 pmol of peptide per 1 mg of beads, the observed maximal quantity of immobilization. This is probably an over-estimation, considering the manufacturer's claim that the capacity of the beads is 400 pmol of peptide per 1 mg beads, but in this way we ensure the saturation of the beads and avoid any undesirable interaction of the library phage with species other than the target.

We have performed seven rounds of panning of the pooled library against Pep90. For each round, phage are incubated with 200 µg of immobilized beads in order for target specific binding to occur, the non-bound phage are eliminated by a number of wash steps, the bound phage are eluted by an acidic solution (pH = 2.2) and finally amplified to be used for the following selection round. The percentages of eluted phage after each selection round, indicating an overall enrichment of binder phages are shown in figure 62. These percentages are calculated by the ratio of eluted phage over the original quantity of phage put in contact with the beads. The selection parameters are kept constant for rounds 1 to 4, where for each round 5×10^{12} phage particles are placed in contact with the beads. For rounds 5 to 7, we have increased the selection pressure by decreasing the number of phage in each round to 5×10^{11} particles. Finally, a negative selection is performed by placing 5×10^{11} phage particles after round 7 in contact with the magnetic beads without the immobilization of the target. This step is to ensure that the selected phages are indeed specific to the Pep90 target not to other parts of the support structure such as the beads or the streptavidin molecule. The results showed that 86% of the selected phage did not bind

to the empty beads, indicating that the selection was successfully performed specific to the Pep90 target.

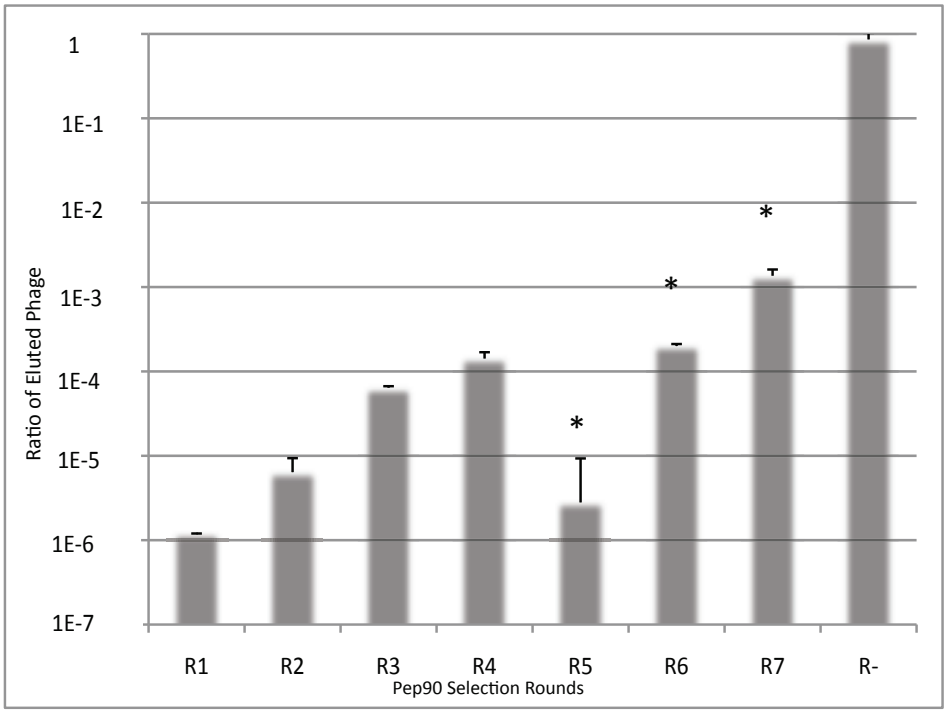


Figure 62. Enrichment of phage specific to Pep90.

Results are calculated as the percentage of eluted phage i.e. the ratio titer of infectious phages incubated with the beads: titer of infectious phages eluded, after each round of selection. The asterisks illustrate the selection rounds in which the selection pressure was increased by lowering the quantity of phage put in contact with the beads to 5×10^{11}, from 5×10^{12}. R- represents the final negative selection round in which also 5×10^{11} phage particles were added to the beads.

In order to confirm the selection of high-affinity binders and to visualize the increase in affinity after each round of selection, we have performed a Phage-ELISA assay. The procedure consists of a sandwich ELISA, where anti-biotin antibodies are first immobilized on the surface of the 96-well plate, the biotinylated target previously incubated with 5×10^{12} phage particles of each selection round are added to each well in duplicates. Finally, anti-pVIII antibodies conjugated with horseradish-peroxide (HRP) are added followed by the addition of ABTS substrate and revelation at 405 nm. The schematic of the sandwich Phage-ELISA and the results are shown in figure 63. The results show an overall increase in the affinity of the selected phage toward Pep90. After round 5, however, the signal decreases, which could potentially mean that five rounds of panning are sufficient for the isolation of high affinity binders.

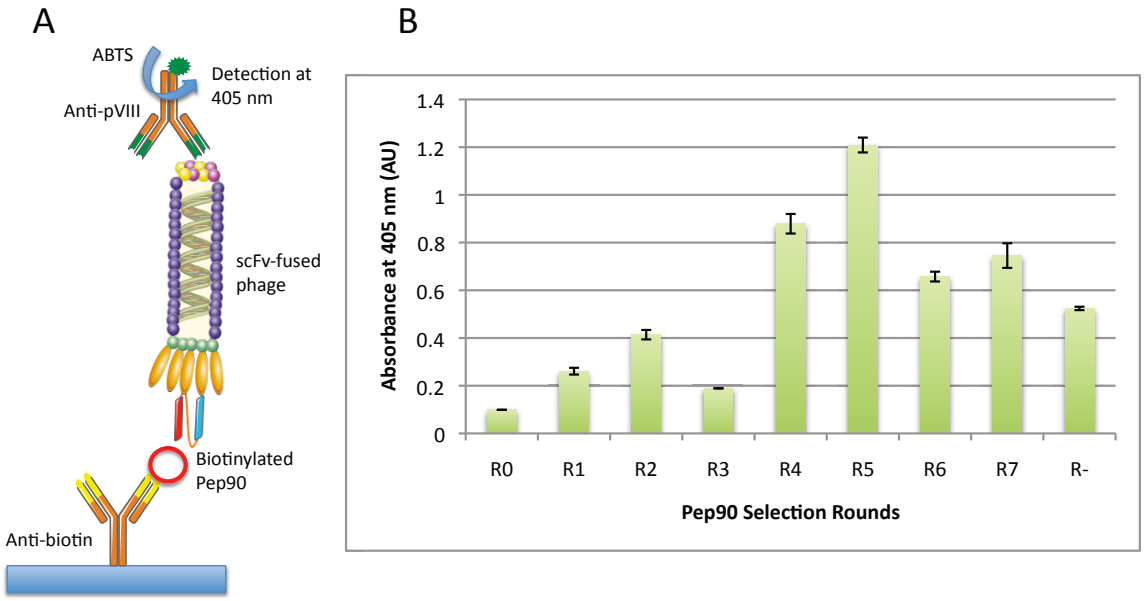


Figure 63. Phage-ELISA assay of the Pep90 selected phage.

A) Schematic of the sandwich ELISA: anti-biotin antibodies are immobilized on plate, biotinylated target and phage are added, and anti-pVIII antibodies-HRP are added to the plate. The signal is revealed with the addition of ABTS B) Results of the seven selection rounds indicating an overall increase in the affinity of the selected phage toward Pep90.

We have screened 10 isolated clones per round, starting from round 4 to the negative selection round, making the total number of screened clones 50. Screening by a Phage-ELISA assay has revealed that 22 out of the 50 clones have affinity (with the ELISA signal >10x greater than the negative control) toward the target Pep90 (Figure 64).

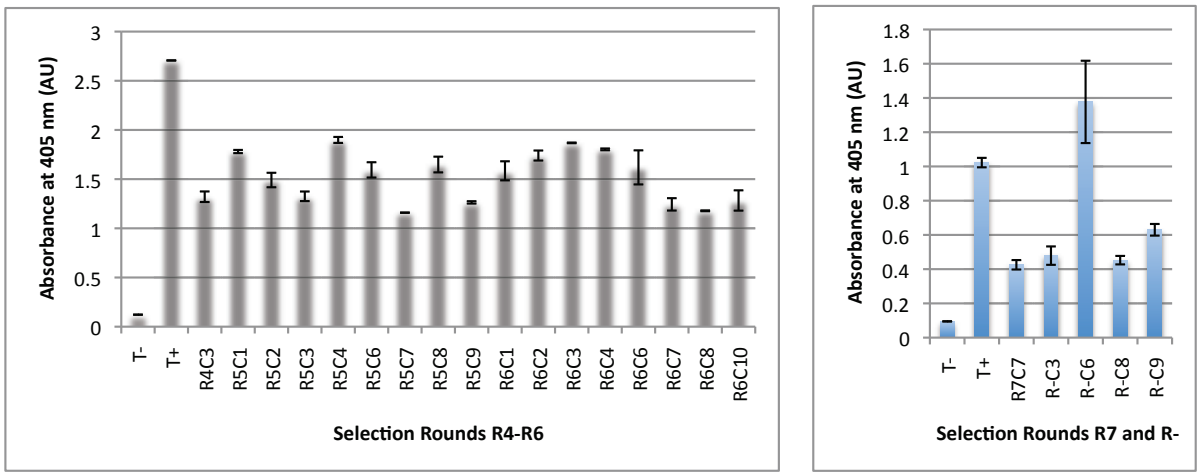


Figure 64. The 22 positive clones selected against Pep90.

The results are shown in two different Phage-ELISA assays. Each sample is performed in duplicates with the error bars indicating the standard deviation. Samples are indicated by RxCy, where x is the selection round and y is the clone number. R- indicates the final negative selection round. For each test, T- indicates the negative control, where library phage prior to selection have been used, and T+ indicates the positive control, where the scFv 9G4H9 displayed on the surface of phage has been used.

The 22 clones were sequenced and it was revealed that 20 out of the 22 clones are identical. This clone has been named P90C1. Clones R-C6 and R5C7 are unique clones with different sequences and have been named P90C2 and P90C3, respectively. It is important to note that P90C2 showed the highest signal in the Phage-ELISA assay. The sequences of these clones are shown in figure 65.

```
Light chain
        1         11        21        31        41        51        61        71        81        91        101       111
P90C1   EIVLTQSPAI MSASPGEKVT MTCSASSSIS ....Y..MHW YQQKPGFSPK LLIYRTSNLA SGVPARFSGS GSGTSYSLTI GTMEAEDVAT YYCQQGSSIP LTFGAGTKLE LKR
P90C2   QIVLSQSPAL MAASPGEKVT ITCSVSSSIS ....SSNLHW YQQKPGSSPK LWIYSTSNLA SGVPVRFSGS GSGTSYSLTI SSMEAEDAAT YYCQQWSSYP LTFGAGTKLE LKR
P90C3   DIVMTQSPAI MSASPGEKVT MTCSASSSVS ....Y..MHW YQQKSGTSPK RWIYDTSKLA SGVPARFSGS GSGTSYSLTI SSMEAEDVAT YYCQQGSSIP LTFGAGTKLE LKR

Heavy chain
        1         11        21        31        41        51        61        71        81        91        101       111
P90C1   EVKLVESGGG LVKPGGSLKL SCAASGFTF. SDYYMYWVRQ TPEKRLEWVA TISDGGSYTY YPDSVKGRFT ISRDNAKNNL YLQMSSLKSE DTAMYYCVRE GRRREGYAMD YWGQGTSLTVS
P90C2   EVMLVESGGD LVKPGGSLKL SCAASGFTF. SSYGMSWVRQ TPDKRLEWVA TISSGGSYTY YPDSVKGRFT ISRDNAKNTL YLQMSSLKSE DSAVYYCTRP TKRR.GFA.. YWGQGTLVTVS
P90C3   EVQLQESGPD LVKPSQSLSL TCTVTGYSIT SGYSWHWIRQ FPGNKLEWMG YIHYSGSTNY NP.SLKSRIS ITRDTSKNQF FLQLNSVTTE DTATYYCASL ARG....AMD YWGQGTSVTVS
```

Figure 65. The alignment of Pep90 selected clones by the Multalin software.

The CDRs are indicated in yellow. The sequential numbering scheme attributed to this alignment is different from that used by IMGT.

The CDR3 sequence of P90C1 and P90C3 are identical, however there is no significant sequence similarity for any other segment between the three antibody sequences. We have submitted these sequences to the IMGT V-QUEST tool in order to determine to which immunoglobulin gene subgroups they belong. As for the light chain of these antibodies, all three belong to the IGKV4 and IGKJ5 light chain κ subgroups. As for the heavy chain γ V gene subgroups, clones P90C1 and P90C2 belong to the IGHV5 subgroup whereas P90C3 belongs to IGHV3 subgroup. Furthermore, for the heavy chain γ J gene subgroups, clones P90C1 and P90C3 belong to the IGHJ4 subgroup, whereas P90C2 belongs to IGHJ3. Finally, all three clones belong to the IGHD2 subgroup. There are no significant differences observed in the percentage of germline identity between the three clones (Table 10).

Table 10. Pep90-selected clones with their library origin and gene subgroup identification.

Clone	Library origin	IGKV	% identity	IGKJ	% identity	IGHV	% identity	IGHJ	% identity	IGHD
P90C1	Balb/C(i)	IGKV4	95.29	IGKJ5	100	IGHV5	99.31	IGHJ4	92.00	IGHD2
P90C2	Balb/C(i)	IGKV4	97.16	IGKJ5	100	IGHV5	97.92	IGHJ3	91.11	IGHD2
P90C3	SJL/J(n)	IGKV4	96.01	IGKJ5	100	IGHV3	99.65	IGHJ4	84.00	IGHD2

(i) represents an immunized library, (n) represents a naive library.

In order to identify the library of origin of each of the three selected clones, we have performed the restriction digestion reactions previously described (Section I.4.C.c) (Figure 66). The results show that the clones P90C1 and P90C2 originate from the immunized

Balb/C library, evident from the single DNA band in lane E, whereas clone P90C3 originates from the naive SJL/J library.

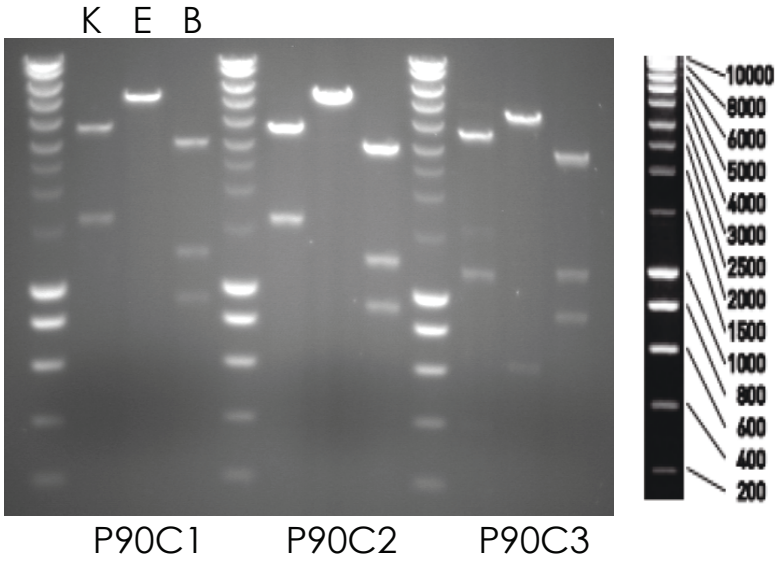


Figure 66. Restriction digestion profiles of the Pep90 selected clones.

For each clone, the first lane K corresponds to the digestion with enzymes HindIII and KpnI, the second lane E corresponds to the digestion with enzymes HindIII and EcoRI, and the third lane B corresponds to the digestion with enzymes HindIII and BstEII. The marker is the SmartLadder® 10kb (Eurogentec).

The observation of three DNA bands in lane B is due to the fact that the antibody sequences of all three clones contain a BstEII restriction site (confirmed by sequencing). The library origins of the clones, their gene subgroup identification and percentage of germline identity are summarized in table 10.

II.4.B. Selection against penam sulfone

A similar selection strategy has been followed using the synthesized penam sulfone molecule as the target. This molecule has a molecular weight of 843.2 $g.mol^{-1}$and is soluble in 100% DMSO. We have performed an optical density spectral scan of this molecule and have determined maximal absorption to occur at wavelength 208 nm. We have produced a calibration curve with different dilutions of the molecule in phosphate buffer saline (PBS) and determined the extinction coefficient to be $\varepsilon=3\times10^{6}$ $L.mol^{-1}.cm^{-1}$ (Figure 67).

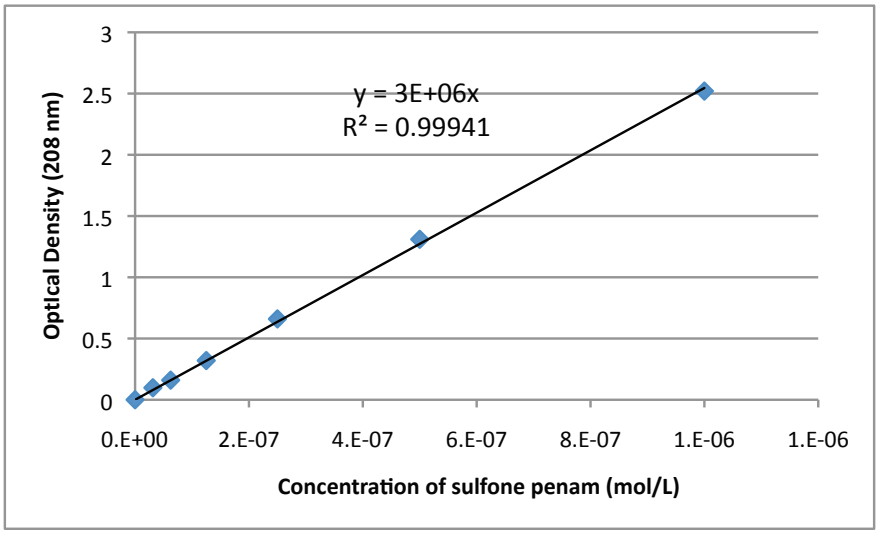


Figure 67. Calibration curve of penam sulfone for the determination of the extinction coefficient in PBS (pH 7.4).

The stock solution of penam sulfone is at 20mM in 100% DMSO, however, the dilutions of factors 1000 upto 80000 are made in PBS. As shown by the linearization formula $\varepsilon = 3\times10^{6}$ $L.mol^{-1}.cm^{-1}$. The R^2 value shows a good estimation for the linearization curve.

We have also performed a saturation experiment, similar to that done for Pep90, in order to determine the optimal quantity of the target to be immobilized on the magnetic beads. The results are shown in table 11.

Table 11. Saturation experiment of the magnetic beads with penam sulfone.

Sample	Quantity (pmol)	OD (eluted)	Immobilized (pmol)
1	8000	3.3	7560
2	800	1.1	653
3	400	0.8	293
4	200	0.55	127
5	100	0.38	49

Different dilutions from a stock solution of 16.86 mg/mL are put in contact with 200 µg magnetic beads. The value corresponding to the first dilution seems to be too high to draw a valid conclusion. The capacity of the beads is determined to be 3500 pmol of the molecule per 1 mg of beads.

Once again we do not observe a typical behavior of saturation of the beads, however we consider the capacity of the beads to be approximately 3500 pmol of the penam sulfone molecule per 1 mg of beads, which was the maximal observed quantity of immobilization.

In this second selection procedure, the suicide inhibitor nature of the penam sulfone potentially promotes the formation of covalent bonds between antibody candidates and the target. Therefore, we have used a modified elution buffer as compared with the previous selection procedure, in order to incorporate 50 mM DTT and elute any potentially covalently bonded antibody by the reduction of the disulfide bond present in the structure of the penam sulfone derivative (Figure 16 in section I.4.B.). However, this modified procedure did not lead to any results. We were not able to observe any enrichment of phage after a three rounds of selection. Furthermore, the eluted phage after each round were hardly amplifiable. We have therefore reverted to the same elution procedure as that performed for the selection against Pep90, which has been validated and proven functional.

In this procedure, we have performed six rounds of panning of the pooled library against the penam sulfone target. The strategy is similar to that used for the Pep90 selection. The percentages of eluted phages after each selection round, indicating an overall enrichment of binder phages are shown in figure 68. Similar to the Pep90 selection, the parameters are kept constant for rounds 1 to 4, where for each round 5×10^{12} phage particles are placed in contact with the beads. For rounds 5 and 6, we have increased the selection pressure by decreasing the number of phage in each round to 5×10^{11} particles.

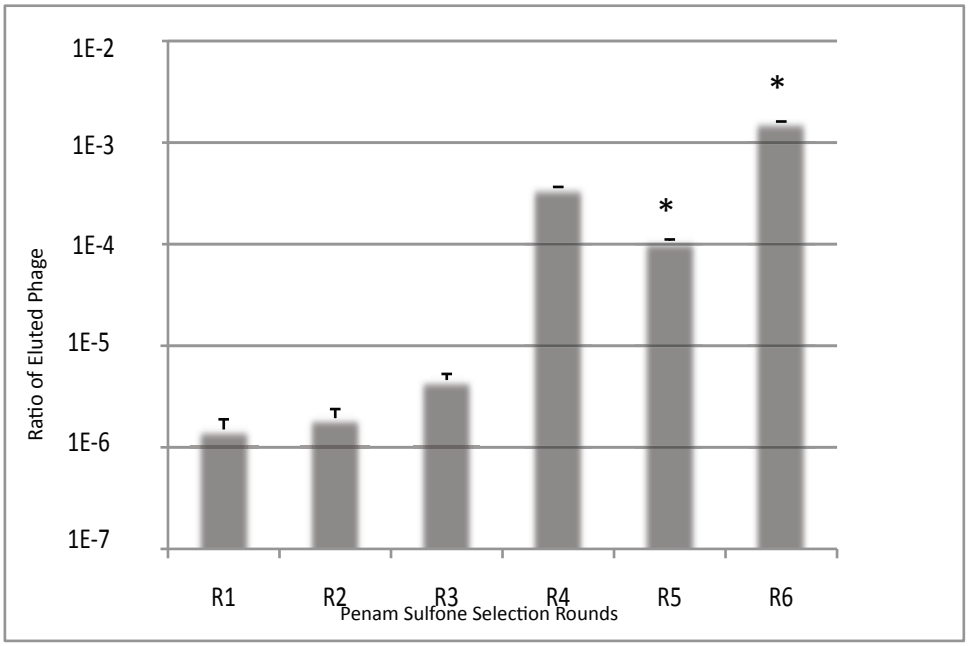


Figure 68. Enrichment of phage specific to penam sulfone.

Results are calculated as the percentage of eluted phage i.e. the ratio titer of infectious phages incubated with the beads: titer of infectious phages eluted, after each round of selection. The asterisks illustrate the selection rounds in which the selection pressure was increased by lowering the quantity of phage put in contact with the beads to 5×10^{11}, from 5×10^{12}.

We have performed a Phage-ELISA assay to confirm the affinity of the selected phage after each panning round toward the penam sulfone target. The strategy of this test is similar to

that used for Pep90. The results show an overall increase in the affinity of the selected phages toward penam sulfone (Figure 69).

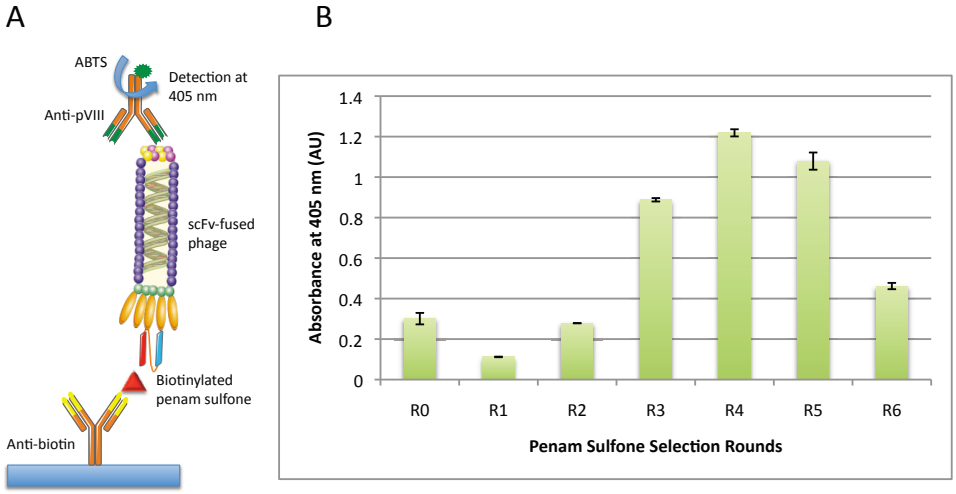


Figure 69. Phage-ELISA assay of the penam sulfone selected phage.

A) Schematic of the sandwich Phage-ELISA B) Results of the six selection rounds indicating an overall increase in the affinity of the selected phage toward penam sulfone.

Once again, after round 5, we observe a decrease in the signal. This could be due to the lower quality of the preparation of phage after round 6, but it, again, could potentially mean that five rounds of panning are sufficient for the isolation of high affinity binders.

We have screened a total number of 30 isolated clones, 10 per round starting from round 4. Screening by Phage-ELISA has revealed that 17 out of the 30 clones recognize the target penam sulfone (Figure 70).

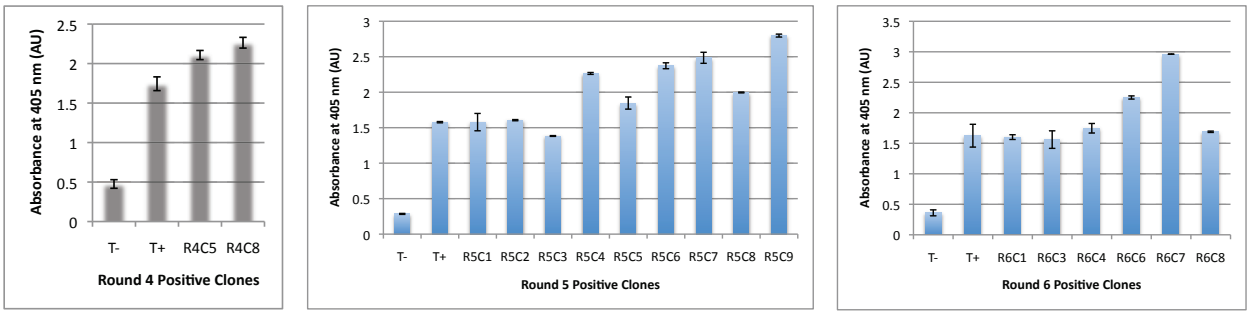


Figure 70. The 17 positive clones selected against penam sulfone.

The results are shown in three different Phage-ELISA assays. Each sample is performed in duplicates with the error bars indicating the standard deviation. Samples are indicated by RxCy, where x is the selection round and y is the clone number. For each test, T- indicates the negative control, where library phage prior to selection have been used, and T+ indicates the positive control, where the scFv 9G4H9 displayed on the surface of phage has been used.

In the absence of having an antibody specific to the penam sulfone molecule, we have used 9G4H9 as the positive control and have observed that this antibody does indeed recognize the target molecule. The 17 clones have been sequenced and it was discovered that 16 were identical. This clone has been named PSC1, whereas the other unique sequence arising from round 6 (R6C1) has been named PSC2. The differences of signal in Phage-ELISA for the samples that are in fact the same clone (PSC1) could be due to differences in the quality of phage preparation. The sequences of these clones are shown in figure 71.

```
Light chain
      1          11         21         31         41         51         61         71         81         91         101        111
PSC1  ETTVTQSPAS LAVSLGQRAT ISCRASKSVS TSR.YSYMHW YQQKPGQPPR LLIKYASNLE AGVPTRFSGS GSRTDFTLNI HPVEEDDAAT YYCQQSREYP WTFGGGTKLE IKR
PSC2  NIVLTQSPAS LAVSLGQRAT ISCRASKSVS TSS.YSYMHW YQQKPGQPPK LLIYHASNLE TGVPARFSGS GSRTDFTLTI DPVEEDDVAI YYCLQSRKIP WTFGGGTKLE IKR

Heavy chain
      1          11         21         31         41         51         61         71         81         91         101        111
PSC1  QVQLQQPGAE LVKPGASVKM SCKASGYTF. TSYWITWVKQ RPGQGLEWIG EIDPSDSYTN YNQKFKGKAT LTVDKSSSTA YMQLSSLTSE DSAVYYCARR DGKK.GAWFA YWGQGTLLTVS
PSC2  QVQLQQSGAE LVKPGASVKL SCKASGYTF. TSYWMHWVKQ RPGRGLEWIG RIDPNSGGTK YNEKFKSKAT LTVDKPSSTA YMQLSSLTSE DSAVYYCARS GYSN.P..FD YWGQGTTVTVS
```

Figure 71. The alignment of penam sulfone selected clones by the Multalin software.

The CDRs are indicated in yellow. The sequential numbering scheme attributed to this alignment is different from that used by IMGT.

The CDR1 of the heavy chain of PSC1 and PSC2 are identical, however the rest of the two sequences do not have any significant similarity. We have performed an immunoglobulin gene subgroup analysis for these two clones, similarly to the Pep90 selected antibodies. For the light chain κ, both clones belong to subgroups IGKV3 and IGKJ1. For the heavy chain γ, they belong to subgroups IGHV1 and IGHD2. However, the heavy chain J segment of PSC1 belongs to IGHJ2, whereas that of PSC2 belongs to IGHJ3. There are again no significant differences observed in the percentage of germline identity between the three clones (Table 12).

Table 12. Penam sulfone-selected clones with their library origin and gene subgroup identifiation.

Clone	Library origin	IGKV	% identity	IGKJ	% identity	IGHV	% identity	IGHJ	% identity	IGHD
PSC1	SJL/J(n)	IGKV3	92.78	IGKJ1	97.37	IGHV1	99.65	IGHJ2	95.45	IGHD2
PSC2	SJL/J(n)	IGKV3	88.66	IGKJ1	97.37	IGHV1	98.61	IGHJ3	100	IGHD2

We have performed restriction digestion reactions in order to identify the library of origin of the two selected clones (Figure 72).

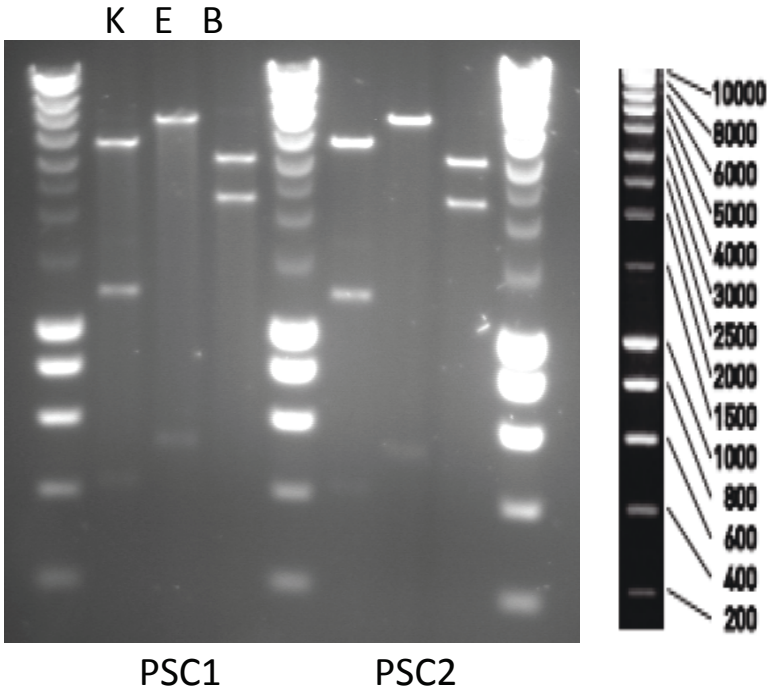


Figure 72. Restriction digestion profiles of the penam sulfone selected clones.

For each clone, the first lane K corresponds to the digestion with enzymes HindIII and KpnI, the second lane E corresponds to the digestion with enzymes HindIII and EcoRI, and the third lane B corresponds to the digestion with enzymes HindIII and BstEII. The marker is the SmartLadder® 10kb (Eurogentec).

The results show that both clones originate from the naive SJL/J library. The observation of the three DNA bands in lane K is due to the presence of a KpnI restriction site in the sequences of these two clones, confirmed by sequencing. The immunoglobulin gene subgroup identification and percentage of germline identity of the selected clones, as well as their library origin are shown in table 12.

We have also performed a Phage-ELISA assay, using penam sulfone as target, for the clones selected against the Pep90 target. The purpose of this test is to examine whether any of the Pep90 selected clones also recognize the penam sulfone molecule. This is of course conceivable since both of these targets are mechanistically inhibitors of the enzyme β-lactamase. The results are shown in figure 73. Interestingly, we have observed that only the clone P90C2, similarly to the model antibody 9G4H9 (positive control), recognizes the penam sulfone molecule.

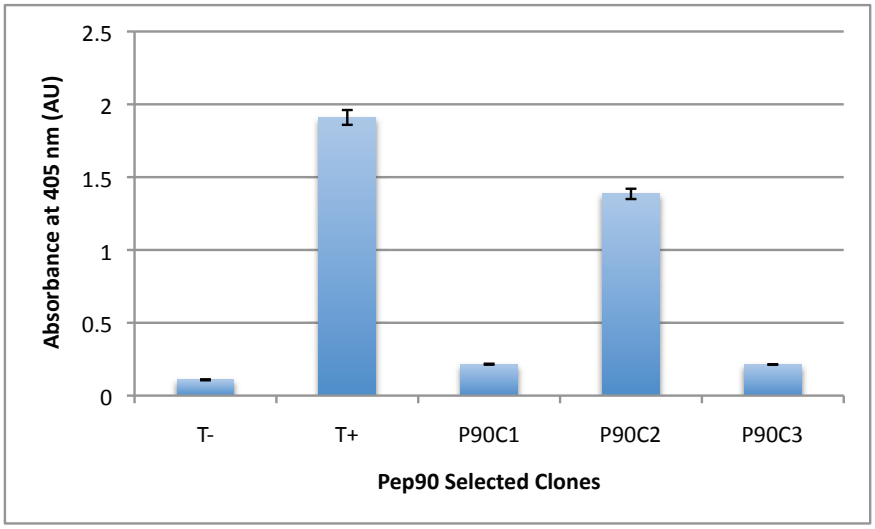


Figure 73. Phage-ELISA assay on Pep90 selected clones using penam sulfone as target.

The capture system is similar to the previous Phage-ELISA experiments. T- indicates the negative control, where library phage prior to selection have been used, and T+ indicates the positive control, where the scFv 9G4H9 displayed on the surface of phage has been used.

Similarly, we want to examine whether the clones selected against penam sulfone also recognize the peptide Pep90. In order to achieve this, we have performed a final Phage-ELISA assay on PSC1 and PSC2, this time using Pep90 as target and including 9G4H9 and P90C2 as controls. The results indicate that the penam sulfone selected clones do indeed recognize Pep90, with a signal comparable to P90C2 (Figure 74).

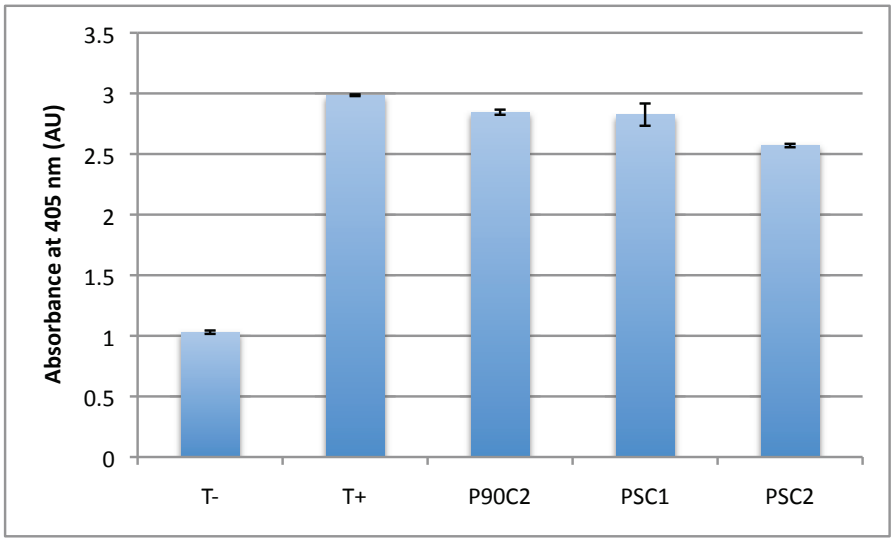


Figure 74. Phage-ELISA assay on penam sulfone selected clones using Pep90 as target.

The capture system is similar to the previous Phage-ELISA experiments. T- indicates the negtive control, where library phage prior to selection have been used. 9G4H9 (T+) and P90C2 are used as positive controls.

Following the selection of five scFv against the two strategically chosen targets, Pep90 and sulfone penam, we want to examine whether any of them are catalytically active.

II.5. Study of catalytic activity

We have chosen two assays in order to test for the presence of β-lactamase activity: *i)* a colorimetric assay using nitrocefin as the substrate, and *ii)* a fluorometric assay using fluorocillin as the substrate bearing the β-lactam ring. Both of these tests are qualitative and are used only for the detection of the presence of β-lactamase or β-lactamase-like activity. We are not able to measure the kinetic parameters of the reactions for the species under study. This is partly due to the fact that, at this step, our subjects of study are scFvs displayed on the surface of phage. As it has been described in the introduction of this manuscript, these phages are able to express from 1 to 5 copies of the scFv on their surface, without considering eventual proteolytic events, which can also occur during the different steps of phage preparation. Therefore, we do not have the means of determining structurally intact phage or the number of scFv species in each sample. Hence, we are unable to determine the number of catalytic sites for each sample, information necessary for the calculation of catalytic constants.

II.5.A. Colorimetric assay

Nitrocefin is the chromogenic cephalosporin developed by Glaxo Research Limited (code 87/312). This compound exhibits a distinctive color change from yellow (max at pH 7.0 = 390nm) to red (max at pH 7.0 = 486 nm) as the amide bond in the β-lactam ring is hydrolyzed by a β-lactamase enzyme (Figure 75). Nitrocefin is sensitive to hydrolysis by all known β-lactamases produced by both Gram-positive and Gram-negative bacteria (O'Callaghan *et al.*, 1972).

We have used nitrocefin-embedded solid disks, provided by the company Oxoid (Dardilly, France), to qualitatively test the five selected clones (P90C1, P90C2, P90C3, PSC1, nad PSC2) for β-lactamase activity. We have used the penicillinase enzyme from the bacteria *Bacillus cereus* (Sigma PO389) at 45nM and the model scFv 9G4H9 as positive controls. It is important to note that the enzyme is in free form, whereas 9G4H9, similar to the five selected scFv under study, is conjugated on the surface of phage. As a negative control, the wild-type phage pAK100 without a conjugate scFv is used. The concentration of phage deposited on the disks is at 5×10^{13} phage particles/mL (determined by OD measurement at 265 nm) and the volume of samples deposited on each disk is 40 µL.

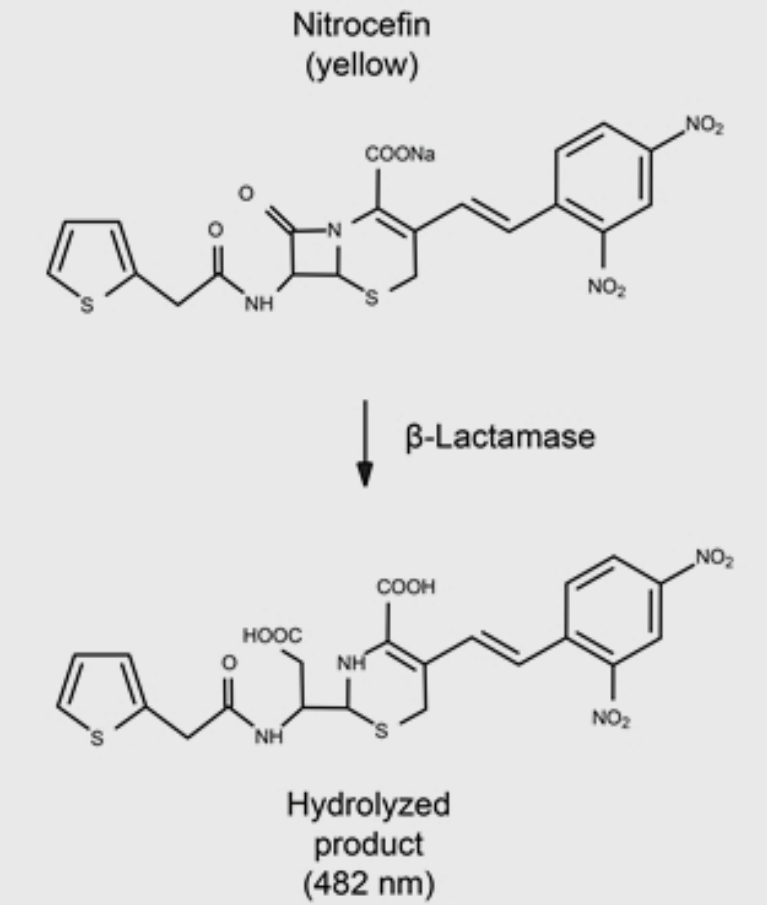


Figure 75. The structure of nitrocefin and its product after hydrolysis of the β-lactam ring.

The hydrolyzed product displays a red color, which provides for a facile identification of the presence of β-lactamase activity.

The disks after 1 hour of the deposition of samples are shown in figure 76. The results show that the disk containing penicillinase enzyme, as expected, has turned into a dark red color. The disk containing the model antibody 9G4H9 has slightly changed color and displays a light pink color, whereas the disk containing the wt phage has remained yellow. As for the selected antibodies, we observe that the clone P90C2 is apparently endowed with the highest catalytic activity, clones PSC1 and PSC2 display a similar phenotype to 9G4H9, whereas clones P90C1 and P90C3 have not contributed to any significant color change.

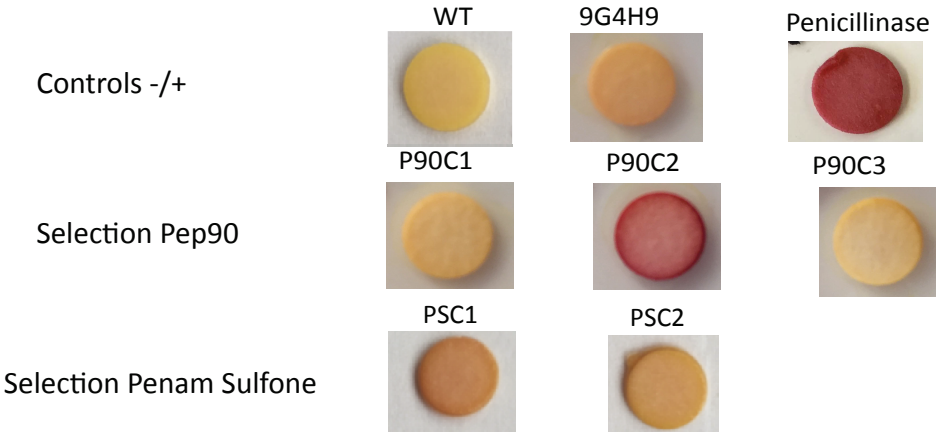


Figure 76. Colorimetric assay for the determination of β-lactamase activity of the five selected clones using nitrocefin as substrate.

The absence of activity is indicated by a yellow color. The presence of activity induces a change of disk color to slightly pink up to red color.

We have therefore established that three out of the five selected clones are able to significantly hydrolyze nitrocefin as a substrate.

II.5.B. Fluorometric assay

Fluorocillin is a new fluorogenic substrate developed in 2009 containing two β-lactam rings (Patent N°: US 2009/0047692 A1, 02/19/2009). The molecule is a difluorofluorescein cephalosporin composed of a central fluoroscein derivative with two fluor atoms in positions 2 and 7, on which two cephalosporin molecules are grafted. The hydrolysis reaction of Fluorocillin and the corresponding products are shown in figure 77.

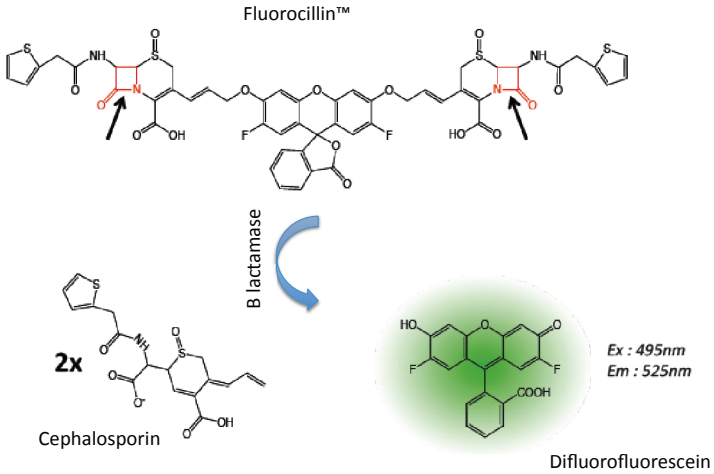


Figure 77. Hydrolysis of Fluorocillin .

The β-lactam rings are indicated by black arrows. The fluorophore, difluorofluorescein, is liberated after the reaction occurs and fluoresces at 525 nm by an excitation at 495 nm.

We have used Fluorocillin to test the activity of the five selected antibodies. We have used similar experimental conditions as those reported in a previous publication by our group (Ben Naya *et al*, 2013). Reactions are performed in 500 µL PBS-1% BSA buffer with the addition of Fluorocillin at 5 µM. The solutions are excited at 495 nm, emissions are detected at 525 nm, and the kinetics is followed for 200 minutes. Similarly to the tests performed on nitrocefin, the penicillinase enzyme is used as a positive control with a concentration of 45 nM. The wild type phage is once again used as a negative control, as well as the substrate in the absence of any catalyst to examine spontaneous hydrolysis. A final negative control reaction has been included with the wild type phage in the absence of substrate in order to confirm the absence of fluorescence caused by the phage components. The volume of phage used for each reaction is 445 µL, where final concentrations are approximately at 5×10^{13} phage particles/mL. The results for the control reactions are shown in figure 78. The penicillinase enzyme displays a Michaelis-Menten behavior with an initial slope of 19.7 min^{-1}, whereas the negative controls, WT phage as well as the spontaneous hydrolysis of Fluorocillin , do not. The final negative control confirms that the phage particles do not fluoresce in the absence of substrate.

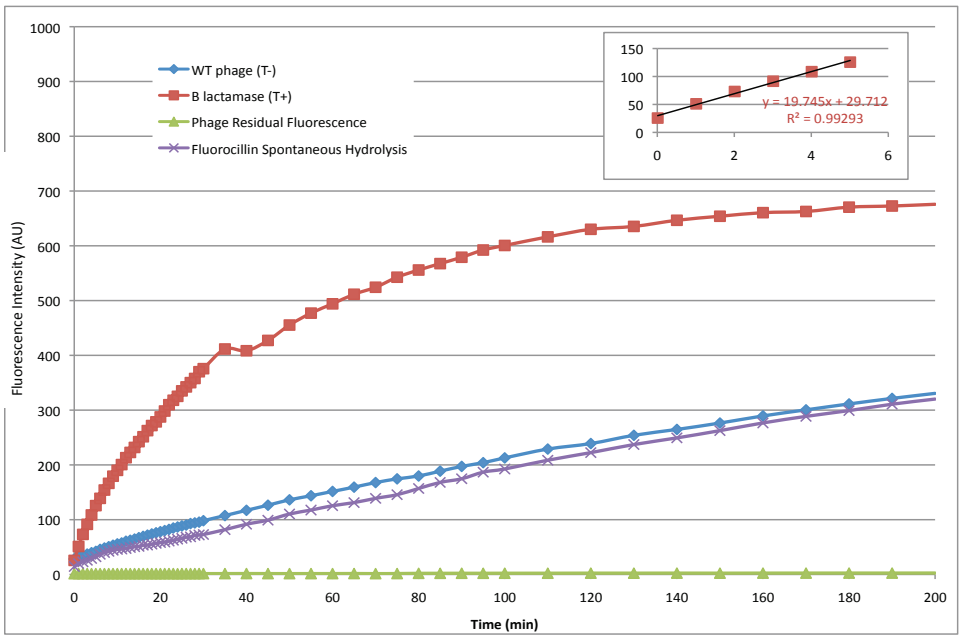


Figure 78. Fluorocillin hydrolysis by the different controls over 200 minutes.

Penicillinase enzyme (T+) and the three negative controls, the wild-type phage (T-), the residual fluorescence of wild-type phage without the inclusion of Fluorocillin, and finally the spontaneous hydrolysis of Fluorocillin in the absence of any catalyst. The inset on the top right of the figure demonstrates the initial slope of the penicillinase positive control (45 nM) to be 19.7 min^{-1}.

We have tested the catalytic antibody 9G4H9 in a similar reaction as a second control. This antibody displays a weak but non-negligible signal as compared with the WT negative control with an initial slope of 3.3 min^{-1}. Under soluble form, this antibody has previously demonstrated to be able to hydrolyze Fluorocillin (Ben Naya *et al.*, 2013). The results for 9G4H9 along with the five selected antibodies are demonstrated in figure 79, where each

scFv tested is compared with the negative control. We observe that clone P90C2 (Figure 79.C) is capable of hydrolyzing Fluorocillin with the highest initial slope (29.4 min^{-1}) at the concentration used for all of the phage reactions (5×10^{13} phage particles/mL). The signal for P90C2 at this concentration is in fact higher than the detection limit of the spectrometer with the set parameters, as we see a saturation of the signal after 50 minutes (curve P90C2 in figure 79.C). In order to confirm that the behavior of this antibody fragment follows a Michaelis-Menten model, we have repeated the test with a lower concentration of phage (3×10^{13} phage particles/mL), which has indeed confirmed this point (curve P90C2* in figure 79.C). This observation is consistent with the results of the colorimetric assay, where P90C2 demonstrated a much stronger activity than the others.

Antibodies P90C1 and P90C3 display weak activities comparable to 9G4H9 with initial slope values of 6.0 and 7.3 min^{-1}, respectively (Figures 79.B and D). The penam sulfone selected antibodies PSC1 and PSC2 demonstrate higher activities with initial slope values of 24.9 and 22.8 min^{-1}, respectively (Figures 79.E and F). This assay has illustrated that all five selected antibodies demonstrate some activity using Fluorocillin as substrate. These results are in general coherent with the results of the colorimetric assay, with the exception that clones P90C1 and P90C3 show positive results, when they did not do so on the nitrocefin substrate. It is important to note that the concentrations measured by optical density are not indicative of the precise concentration of intact phage in each solution, therefore, there might be sight differences between the quantity of intact scFv-displaying phage in each of the samples tested. The difference between initial slope values of 9G4H9 and both P90C1 and P90C3 is not significant, especially due to this lack of precision. However, the fact that all of the three catalytsts are able to hydrolyze the Fluorocillin substrate but only 9G4H9 hydrolyzes the nitrocefin substrate, suggests that the specificities of P90C1 and P90C3 are potentially different from that of 9G4H9. The selected antibodies need to be expressed in soluble form in order to confirm the preliminary results shown in the two assays described here, and in order to be able to quantify the precise number of catalytic site to quantitatively calculate their kinetic parameters.

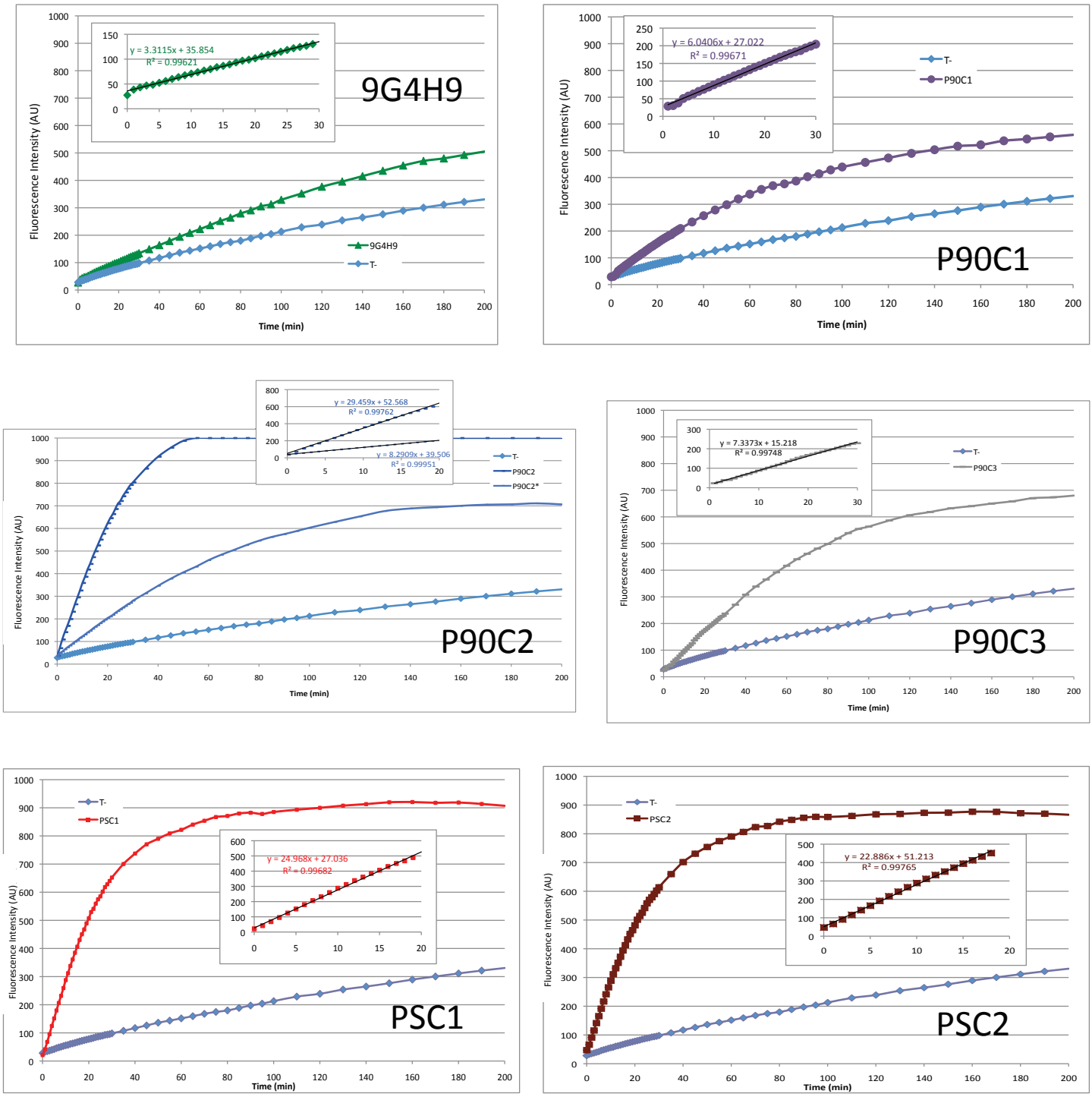


Figure 79. Fluorocillin hydrolysis by the five selected clones and the control scFv 9G4H9 over 200 minutes.

The antibody fragments are displayed on the surface of phage whose concentration in each reaction is 5×10^{13} phage particles per mL, except for the assay marked with a * where the concentration is 3×10^{13} phage particles per mL. The initial slope values for each scFv-phage are displayed by a linear trendline on the first 5-30 minutes of each plot. A) 9G4H9, B) P90C1, C) P90C2, D) P90C3, E) PSC1, F) PSC2.

Table 13 summarized the characteristics of the selected antibodies as compared with the model scFv 9G4H9. Figure 80 demonstrates a sequence alignment of the five selected antibodies along with our model scFv 9G4H9. One of the selected catalytic antibody fragments, namely PSC1, displays the same residues in its amino acid sequence as those determined to be involved in the active site of 9G4H9 (demonstrated in the figure by red letters in bold). Interestingly, the other antibody fragments do not share this similarity, at least at the amino acid sequence level. We have further examined the selected antibodies at a three-dimensional structural level via 3D modeling, in order to determine whether there is any structural similarity between the catalytic antibodies, the model 9G4H9, as well as the parent enzyme β-lactamase.

Table 13. Characteristics of the five selected antibodies and 9G4H9.

Clone	Library origin	IGKV	% identity	IGKJ	% identity	IGHV	% identity	IGHJ	% identity	IGHD	Affinity to Pep90	Affinity to PS	Catalytic (Nitrocefin)	Catalytic (Fluorocillin)
P90C1	Balb/C(i)	IGKV4	95.29	IGKJ5	100	IGHV5	99.31	IGHJ4	92.00	IGHD2	+	-	+	+
P90C2	Balb/C(i)	IGKV4	97.16	IGKJ5	100	IGHV5	97.92	IGHJ3	91.11	IGHD2	+	+	+++	+++
P90C3	SJL/J(n)	IGKV4	96.01	IGKJ5	100	IGHV3	99.65	IGHJ4	84.00	IGHD2	+	-	+	+
PSC1	SJL/J(n)	IGKV3	92.78	IGKJ1	97.37	IGHV1	99.65	IGHJ2	95.45	IGHD2	+	+	+	++
PSC2	SJL/J(n)	IGKV3	88.66	IGKJ1	97.37	IGHV1	98.61	IGHJ3	100	IGHD2	+	+	+	++
9G4H9	Biozzi	IGKV2	96.60	IGKJ5	97.37	IGHV1	99.65	IGHJ2	95.45	IGHD1	+	+	+	+

The immunoglobulin gene subgroups, the percentage of germline identity, affinity to the target molecules Pep90 and penam sulfone, and catalytic capability on the substrates nitrocefin and fluorocillin are indicated.

```
        Light chain
        1          11         21         31         41         51         61         71         81         91         101        111
9G4H9   DIVMTQAAPS VPVTPGESVS ISCRSSKSLL HSNGNTYLYW FLQRPGQSPQ RLIYYMSNLA SGVPDRFSGR GSGTDFTLRI SRVEAEDVGV YYCMQSLEYP LTFGAGTKLE LKR
PSC1    ETTVTQSPAS LAVSLGQRAT ISCRASKSVS TSR.YSYMHW YQQKPGQPPR LLIKYASNLE AGVPTRFSGS GSRTDFTLNI HPVEEDDAAT YYCQQSREYP WTFGGGTKLE IKR
PSC2    NIVLTQSPAS LAVSLGQRAT ISCRASKSVS TSS.YSYMHW YQQKPGQPPK LLIYHASNLE TGVPARFSGS GSRTDFTLTI DPVEEDDVAI YYCLQSRKIP WTFGGGTKLE IKR
P90C1   EIVLTQSPAI MSASPGEKVT MTCSASSSIS ....Y..MHW YQQKPGFSPK LLIYRTSNLA SGVPARFSGS GSGTSYSLTI GTMEAEDVAT YYCQQGSSIP LTFGAGTKLE LKR
P90C2   QIVLSQSPAL MAASPGEKVT ITCSVSSSIS ....SSNLHW YQQKPGSSPK LWIYSTSNLA SGVPVRFSGS GSGTSYSLTI SSMEAEDAAT YYCQQWSSYP LTFGAGTKLE LKR
P90C3   DIVMTQSPAI MSASPGEKVT MTCSASSSVS ....Y..MHW YQQKSGTSPK RWIYDTSKLA SGVPARFSGS GSGTSYSLTI SSMEAEDVAT YYCQQGSSIP LTFGAGTKLE LKR

        Heavy chain
        1          11         21         31         41         51         61         71         81         91         101        111
9G4H9   EVQLQQSGPE LVKPGASVKI SCKASGYSF. TDYNMNWVKQ SNGKSLEWIG VINPNYGTTS YNQKFKGKAT LTVDQSSSTA YMQLNSLTSE DSAVYYCARG LRRY....FD YWGQGTTLTVS
PSC1    QVQLQQPGAE LVKPGASVKM SCKASGYTF. TSYWITWVKQ RPGQGLEWIG EIDPSDSYTN YNQKFKGKAT LTVDKSSSTA YMQLSSLTSE DSAVYYCARR DGKK.GAWFA YWGQGTLLTVS
PSC2    QVQLQQSGAE LVKPGASVKL SCKASGYTF. TSYWMHWVKQ RPGRGLEWIG RIDPNSGGTK YNEKFKSKAT LTVDKPSSTA YMQLSSLTSE DSAVYYCARS GYSN.P..FD YWGQGTTVTVS
P90C1   EVKLVESGGG LVKPGGSLKL SCAASGFTF. SDYYMYWVRQ TPEKRLEWVA TISDGGSYTY YPDSVKGRFT ISRDNAKNNL YLQMSSLKSE DTAMYYCVRE GRRREGYAMD YWGQGTSLTVS
P90C2   EVMLVESGGD LVKPGGSLKL SCAASGFTF. SSYGMSWVRQ TPDKRLEWVA TISSGGSYTY YPDSVKGRFT ISRDNAKNTL YLQMSSLKSE DSAVYYCTRP TKRR.GFA.. YWGQGTLVTVS
P90C3   EVQLQESGPD LVKPSQSLSL TCTVTGYSIT SGYSWHWIRQ FPGNKLEWMG YIHYSGSTNY NP.SLKSRIS ITRDTSKNQF FLQLNSVTTE DTATYYCASL ARG....AMD YWGQGTSVTVS
```

Figure 80. Alignment of the five selected clones and 9G4H9.

The CDR regions are highlighted in yellow. The letters in bold red represent the residues, which have been proposed by the 3D modeling to be implicated in the abzyme active site (next section). The letters in regular font red represent multiple possibilities of an active site residue suggested by the 3D model and mutagenesis. The dots represent the gaps inserted in the alignment due to the Multalin software. The sequential numbering scheme attributed to this alignment is different from that used by IMGT.

II.6. Structural modeling

We have used the *RosettaAntibody* modeling software (Sircar *et al.*, 2009) with the collaboration of Pr. Bernard Offmann (Université de Nantes, CNRS UMR 6286), in order to develop three-dimensional models of the selected antibody fragments. This program is part of the Rosetta softwares and is suitable for the analysis of numerous fundamental questions related to different subjects including structural biology, conception of new enzymatic proteins, and the prediction for the structures of large non-coding RNA molecules. *RosettaAntibody* is able to produce 3D models for the variable regions of antibodies from the amino acid sequences of the light and heavy chains.

Following the construction of the 3D models for the five selected scFv, we have attempted to find structural motifs present in the antibodies that are similar to the active site of the enzyme β-lactamase from TEM-1 (PDB: 1BTL; Jelsch *et al.*, 1993). As described previously in the manuscript, the active site of this enzyme contains the residues Ser^{70}, Lys^{73}, Ser^{130}, and Glu^{166}, all of which are involved in the catalytic mechanism. Similarly, in our 3D models we search for structural motifs that resemble the structural positioning of these active residues or have an adequate superposition of the model enzyme active site. For each model, we have fixed the two serine residues in similar position to the β-lactamase active site, but for the remaining two residues, we have considered the availability of amino acids with similar positioning as well as physicochemical properties as the lysine and the glutamate. For example, in place of the lysine residue, we have also considered potential arginine residues, whose role has been confirmed in the model antibody 9G4H9 as well as in previous work on a different catalytic antibody (Benedetti *et al.*, 2004). Additionally, in place of a glutamate residue, we have also considered potential tyrosine residues, whose involvement in a β-lactamase variant has also been previously demonstrated (Stojanoski *et al.* 2015).

The structure of the antibody fragment P90C1 demonstrates two potential serine residues on the light chain at positions 65 and 69 (Figure 81). It is important to note that the numbering of residues in the 3D models are given according to the IMGT criteria. The residues identified in P90C1 hold similar positions as the serine residues present in the β-lactamase active site, however their distance from each other is further than that of the enzyme. A tyrosine residue, also on the light chain, at position 55 is accessible to the serines and could potentially be involved in the catalytic process. The only available residues in this proximity, which could potentially play the role of Lys^{73} in the enzyme active site, are four arginine residues located on a loop in the heavy chain at positions 106, 109, 110, and 111.

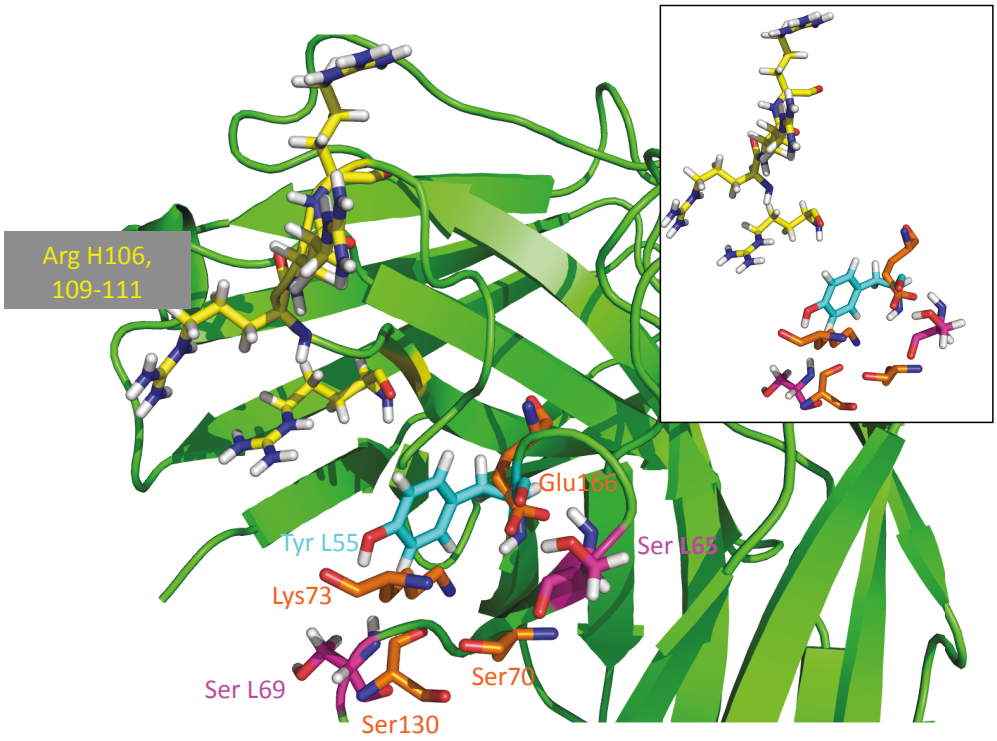


Figure 81. 3D model of antibody P90C1 with superposition of β-lactamase active residues.

The active residues of β-lactamase are in orange. The residues Ser^{65}, Ser^{69}, and Tyr^{55} are located on the light chain of the scFv. There are four possible arginine residues completing the active site pocket, which are all located on the heavy chain. The box on the top right simplifies the visualization of the structure by only depicting the superposition of catalytic residues.

All of these residues, however, are much further away from the serines than the active site pocket of the enzyme. The light chains of scFv P90C2 and P90C3 also contain the residues Ser^{65}, Ser^{69}, and Tyr^{55}, similar to scFv P90C1, potentially contributing to their activity (Figures 82 and 83). In the structure of antibody fragment P90C2, however, the residue Arg^{113} located on the CDR3 loop of the heavy chain is much closer to the serines in the three dimensional structure than the three arginines of P90C1. For scFv P90C3, Arg^{109} is the only positively charged residue in the vicinity being able to potentially participate in the reaction. This residue is nevertheless much further away from the serines as compared to the structure of P90C2. This might be a reason why the antibody P90C2 is much more active than the other two antibodies, namely P90C1 and P90C3.

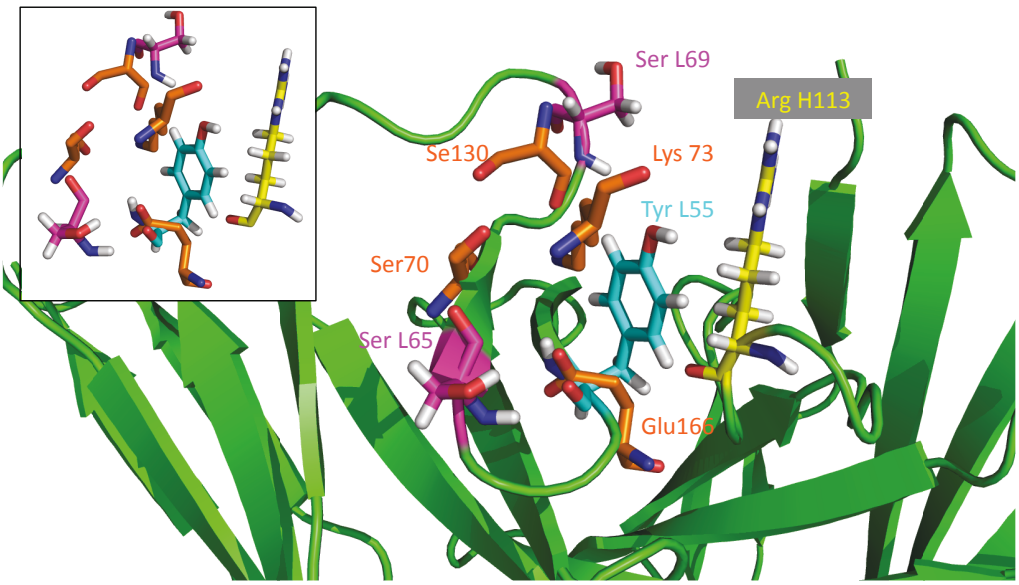


Figure 82. 3D model of antibody P90C2 with superposition of β-lactamase active residues.

The active site residues of β-lactamase are indicated in orange. The residues Ser^{65}, Ser^{69}, and Tyr^{55} are located on the light chain of the scFv. The Arg^{113} residue completing the quartet is located on the heavy chain. The box on the top left simplifies the visualization of the structure by only depicting the superposition of catalytic residues.

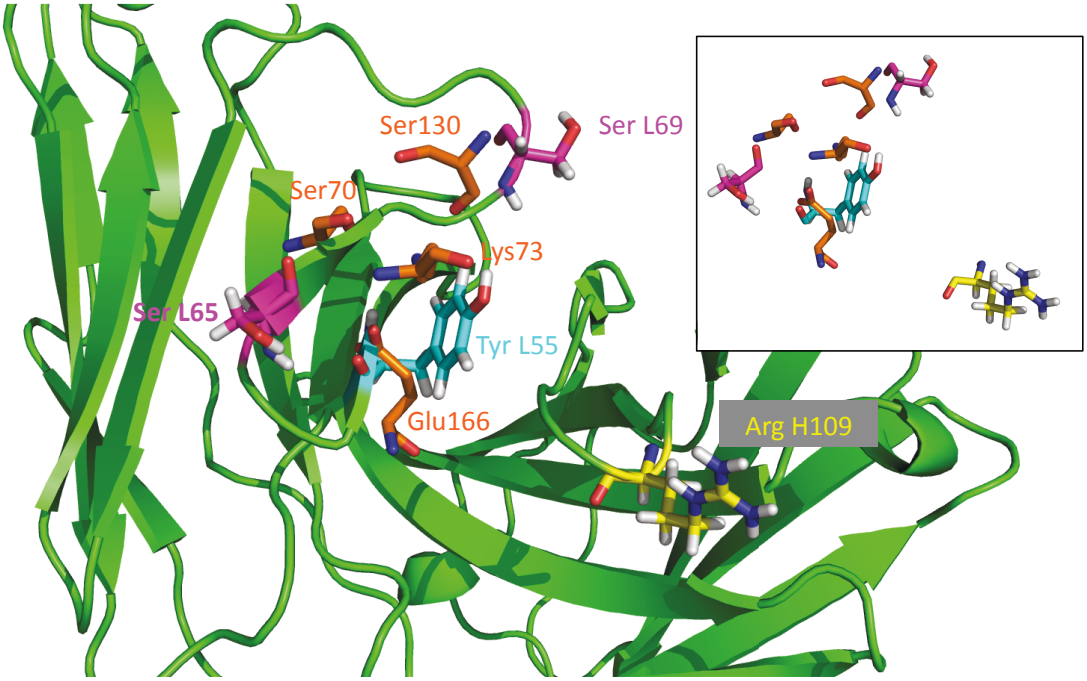


Figure 83. 3D model of antibody P90C3 with superposition of β-lactamase active residues.

The active site residues of β-lactamase are indicated in orange. The residues Ser^{65}, Ser^{69}, and Tyr^{55} are located on the light chain of the scFv. The Arg^{109} residue completing the quartet is located on the heavy chain. The box on the top right simplifies the visualization of the structure by only depicting the superposition of catalytic residues.

The structure of the antibody PSC1 is very similar to the model antibody 9G4H9 (Figure 84).

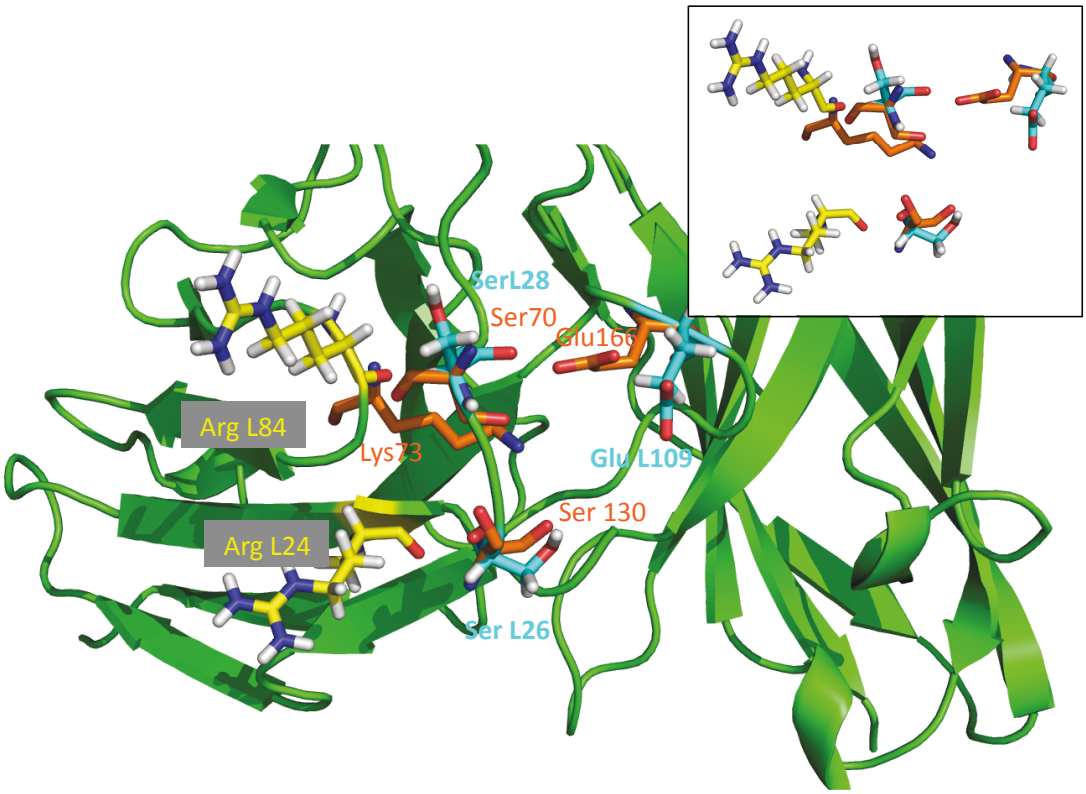


Figure 84. 3D model of antibody PSC1 with superposition of β-lactamase active residues.

The active site residues of β-lactamase are indicated in orange. All of the assigned potential active site residues are located on the light chain. These include residues Ser^{26}, Ser^{28}, and Glu^{109}. There are two arginine residues that could potentially be involved in the reaction: Arg^{24} and Arg^{84}. The box on the top right simplifies the visualization of the structure by only depicting the superposition of catalytic residues.

The active site residues Ser^{26}, Ser^{28}, and Glu^{98} of 9G4H9 are also present on the light chain of antibody PSC1. However, the numbering of these residues on the latter antibody fragment, namely PSC1, have been attributed according to the IMGT criteria, whereas that of 9G4H9 have been determined sequentially, similar to the referenced publications (Phichith *et al.*, 2009). The residue Arg^{24}, which has been confirmed by point mutation studies to be involved in the catalytic activity of 9G4H9 is also present in PSC1. There is, however, an additional possibility of the residue Arg^{84} taking this role in PSC1, which does not exist in 9G4H9. Arg^{84} seems to have a more favorable structural orientation toward Ser^{28} than the orientation of Arg^{24}. This might be the reason for the higher activity of PSC1 as compared to 9G4H9. A superposition of the two antibody fragments PSC1 and 9G4H9 is shown in figure 85.

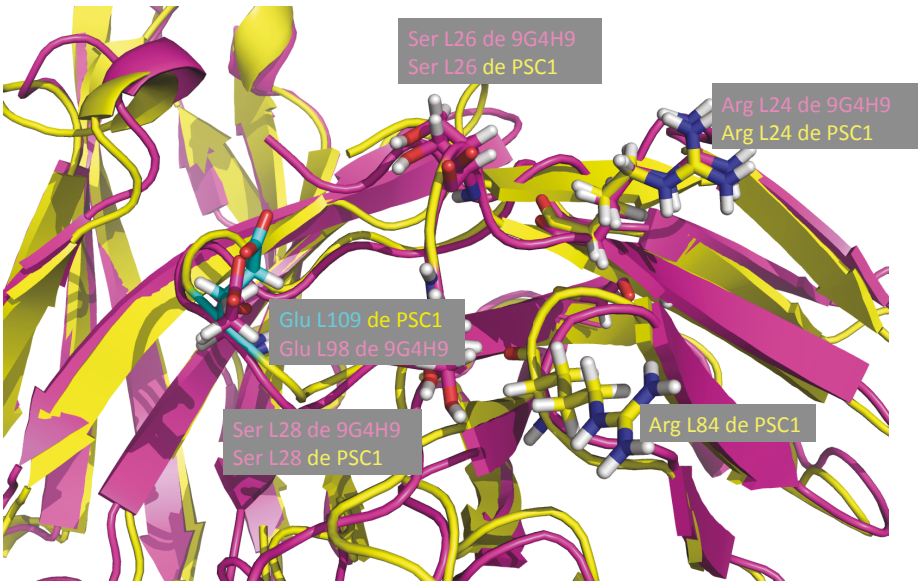


Figure 85. Superposition of antibodies PSC1 and 9G4H9.

PSC1 is indicated in yellow and 9G4H9 in pink.

Antibody fragment PSC2 is the only antibody whose best structural superposition to the β-lactamase active site is entirely located on the heavy chain (Figure 86). The proposed active pocket of this antibody contains an aspartate residue instead of the Glu^{166} residue of the enzyme.

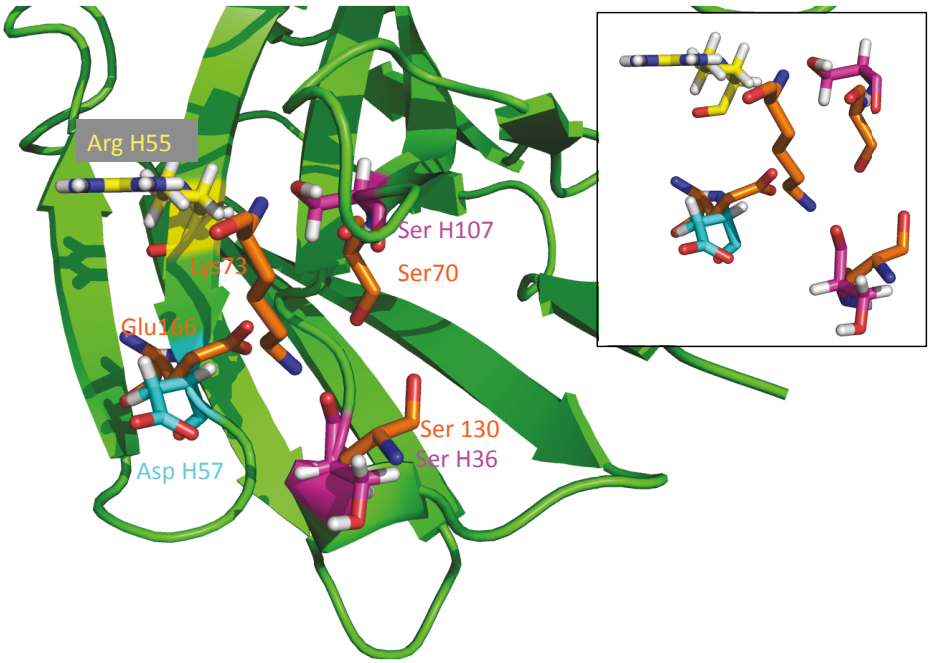


Figure 86. 3D model of antibody PSC2 with superposition of β-lactamase active residues.

The active site residues of β-lactamase are indicated in orange. All of the assigned potential active site residues are located on the heavy chain. These include residues Ser^{36}, Ser^{107}, and Asp^{57}, and Arg^{55}. The box on the top right simplifies the visualization of the structure by only depicting the superposition of catalytic residues.

We have therefore found possible structural motifs for three out of the five selected antibodies, containing residues that are potentially responsible for the demonstrated catalytic activity of these abzymes. The absence of such motifs for P90C1 and P90C2, resembling the β-lactamase enzyme active site pocket is potentially the reason for the lower detected activity of these antibodies. The seemingly out of reach arginine residues situated further away from the serines that what is expected for a β-lactamase active site pocket potentially offers an explanation. It is also likely that these catalysts have a completely different catalytic structure, which does not resemble the active site pocket of the enzyme β-lactamase and has not been detected in our structural study. To provide a clearer answer to all these speculations, and to confirm the presently provided preliminary models, it is necessary to acquire a crystal structure of these antibodies in the future.

DISCUSSION AND PERSPECTIVES

III. DISCUSSION AND PERSPECTIVES

In the past few decades, catalytic antibodies have become the focus of interest in studying a new aspect of the immune response: the tendency to induce the maturation of B lymphocytes capable of catalyzing numerous enzymatic reactions. This is a different function, attributed to mature B cells, than their previously understood role of antigen binding with high affinity and high specificity for the recruitment of effector cells. The involvement of these peculiar antibodies in numerous diseases highlighted the importance of their implication in physiopathology. However, little is yet known about the origin and nature of catalytic antibodies and their exact role in the immune system. In the present work, we have constructed four phage display libraries representing immune repertoires with different genetic backgrounds and immunological states. Due to a novel "restriction bar-coding" strategy of the host vectors, which allows us to determine the library origin of any selected antibody, we have been able to pool these 4 repertoires together in order to create a final library of size 2.7×10^9. We have analyzed the immunoglobulin gene expression patterns of each library, which has revealed a number of statistical differences between and confirmed the large diversity across the four immune repertoires. These libraries are, therefore, an excellent reservoir of diversity and can be used for the selection of antibodies against any targets of potential interest. In an attempt to study the nature of catalytic antibodies and the influence of the nature of immune repertoires on their expression, we have subsequently taken advantage of this elevated diversity in order to select for and characterize catalytic antibodies with a model β-lactamase activity.

III.1. Library Construction

Phage display technology has provided scientists with a powerful tool allowing the *in vitro* reproduction of immune repertoires. With the use of carefully designed primer sequences and a well adapted cloning strategy for the amplification of immunoglobulin variable regions, this technology has the potential to create antibody libraries close in size and diversity to their *in vivo* natural repertoires. This has enabled the isolation of antibodies targeted toward almost any conceivable antigen. Furthermore, the capability of this technology has prompted researchers to construct models of natural immune repertoires in order to study their characteristics such as immunoglobulin expression patterns and changes in B cell repertoires influenced by immunization or disease.

We have utilized the phage display technology to construct four scFv antibody libraries of murine origin, representing different genetic backgrounds and immunological states. We have decidedly chosen to focus on IgG repertoires, in order to limit our study to the same isotype originating from the four populations. This means that we have excluded the IgM immunoglobulins present in natural naive immune repertoires (Sidhu, 2005) so that we can perform a valid analysis on comparable populations. What we consider naive in our represented populations, therefore, are non-immunized IgG repertoires.

For the amplification of immunoglobulin variable genes, we have used a novel strategy slightly different from the commonly used RACE-PCR (Frohman, 1990; Ozawa *et al.*, 2008). In the RACE-PCR strategy, mRNA are reverse-transcribed by targeting the 3' poly-A tails of these sequences *via* oligo-dT primers. Following the synthesis of this first strand cDNA, a known adaptor sequence is ligated to the 5' ends of the immunoglobulin cDNA in order for the following PCR amplification to target this adaptor sequence at the 5' and the poly-A tail at the 3' end. In this manner, RACE-PCR is thought to lead to the amplification of variable genes in an unbiased fashion and to avoid the introduction of exogenous sequences due to the mismatch hybridization of primers. We, however, originally utilized this technique but were not able to produce libraries of sufficient size and diversity. This led us to design a new, more efficient, strategy of library construction in order to produce larger libraries with higher diversity.

In our novel strategy, we have first performed an RT-PCR reaction using isotype specific primers, hence reducing the pool of cDNA, in order to focus only on the amplification of IgG genes, and therefore increasing the specificity of the following amplification step. We have then performed a nested PCR strategy, where we target, in a first reaction, L-PART1 and constant regions of V_L and V_H, and in a second nested reaction, the FR1 and FR4 sequences of the previously amplified immunoglobulin genes. The degenerated oligonucleotide primer mixes have been newly designed based on an updated and thorough analysis of the IMGT database in order to include all of the murine immunoglobulin gene sequences reported up to date. In this way, we have been able to target the maximum possible number of immunoglobulin gene sequences present in the murine repertoires. This novel technique has three advantages: *(i)* by using isotype-specific primers, we are able to reduce the pool of cDNA in one single reaction, increasing the specificity of the overall procedure, *(ii)* by targeting the L-PART1 regions in the first PCR reaction, we maintain the integrity of the 5' regions of immunoglobulin sequences (FR1), leading to a more accurate representation of the naive repertoires, and *(iii)* by redesigning a more complete mix of degenerated primer sequences used for the amplification steps, we ensure the amplification of a higher number of immunoglobulin sequences available in the naive repertoire, increasing the diversity of the libraries.

The analysis of the total number of gene subgroups observed in the 4 libraries reported in this study shows a large representation of all the subgroups available in IMGT®. Indeed, we have shown that 91.1% of the gene subgroups expected to be present in our sample are represented, indicating the high quality of the newly designed primers. Previously, we have constructed a naive SJL/J library by use of a different set of primers, in which we observed an overexpression of the gene subgroups IGKV2 and IGKV6 compared to the data available in IMGT® (Shahsavarian *et al.*, 2014). We concluded at the time that this discrepancy was either due to a bias in the design of primers used or to the fact that the IMGT® sequences referenced in GeneFrequency are based mostly on monoclonal antibodies induced by immunization and might not be an appropriate reference. Comparing these results with the data obtained for the naive SJL/J library reported here (constructed by the same technique but a different and newly designed set of primers), we observe differences in the two gene expression profiles. This indicates that regardless of the suitability of the IMGT® GeneFrequency tool as a reference for comparing expression profiles, the design strategy of

primers has a crucial influence on the final representation of the immunoglobulin repertoire. Therefore, the broader the range of immunoglobulin sequences covered by the primer mix, the more accurate the representation of the produced libraries of the *in vivo* repertoire. It is important to note, however, the absence of λ genes in the pool of 400 sequences analyzed. If we consider the murine immunoglobulin pool to be composed of 95% κ chains and only 5% λ chains (Murphy, 2012) and considering the 400 immunoglobulin light chains we have analyzed, we should in fact expect the presence of λ sequences. We have, therefore, speculated that there might be a bias in our amplification procedure originating from our primer design. In order to investigate this, we have used the designed λ primers separately in an amplification procedure and have indeed confirmed that they are capable of amplifying λ genes. We have then cloned a number of the λ genes into pGEMT vectors and have also confirmed the capability of our total PCR2 mix of primers in amplifying the λ genes, using the same conditions applied for the construction of libraries. These results suggest that the true ratio of κ to λ genes in our murine immune repertoire might be higher than 95:5, as suggested previously by other groups (Dildrop *et al.*, 1987; Zocher *et al.*, 1995; Almagro *et al.*, 1998).

It is also important to mention that it seems that a different IgG subclass, namely IgG2c, is present in the SJL/J strain replacing the IgG2a subclass present in Balb/C mice. This is confirmed by the information available in the IMGT database, where the isotype IgG2c is observed in the immunoglobulin sequences of SJL/J mice, whereas the isotype IgG2a is absent in this strain. This additional isotype has not been included in our primer design, which suggests the absence of this particular isotype in our immunoglobulin pool. The reason for this discrepancy is the absence of reported IgG2c sequences in the IMGT LIGM/DB database, which we have used for the design of primers. Sequences belonging to this isotype, however, are present in the IMGT GENE/DB database. This suggests to us that for future primer design strategies, the Gene/DB database is a more appropriate and complete source of sequences. Due to this realization and the dynamic nature of the sequences available in the IMGT database, we are aware that for the future construction of libraries, a new set of amplification primers needs to be designed taking into account updated immunoglobulin sequence information and based on the database Gene/DB.

Another optimization of our library construction strategy is the incorporation of RCA in the cloning procedure in order to improve the diversity of the libraries. The construction of antibody libraries by phage display technology involves a number of different molecular and cell biology techniques, each having an influence on the size and diversity of the libraries and requiring optimization. Inappropriate digestion of restriction sites on the extremities of scFv sequences, due to insufficient length of the flanking base pairs on which the restriction enzymes are to bind, can lead to a decrease in the number of successful ligations. Indeed, this can lead to a reduced number of transformed bacteria, affecting the final size of the library. It is important to note that RCA has been previously used by the group of Greg Winter for the improvement of library diversity (Christ *et al.*, 2006). The authors have used RCA following the ligation of scFv fragments into the corresponding vector, in order to amplify the ligation reactions and improve their chances of transformation into bacteria. The use of RCA at this late stage, however, does not exclude the loss of diversity introduced in the previous restriction digestion reaction. We have

therefore incorporated the RCA technique immediately following the assembly of the scFv fragments. In this way, we ensure the efficient digestion of all scFv fragments, leading to an increase in library size (from 10^5 to 10^{8-9}), percentage of successful cloning, and an overall improved cloning procedure (Shahsavarian *et al.*, 2014).

III.2. Characterization of the library

We have performed a sequence analysis on a total number of 400 immunoglobulin sequences (100 sequences per library), in order to confirm the diversity of our libraries and to compare the immunoglobulin gene expression profiles between the 4 immune repertoires. We have achieved this by performing conventional Sanger sequencing in order to avoid the potential introduction of mutations in the sequencing results, which is practically inevitable by the currently existing high throughput sequencing (HTS) techniques. Additionally, our level of analysis requires the sequencing of the entire scFv segment (~800bp) in order to study the pairing of V_L and V_H, which HTS would not allow us to do. HTS allows the sequencing of V_L and V_H chains independently (~300 to 400 bps) preventing a direct analysis of a combination of the two (Kircher and Kelso, 2010; Ravn *et al.*, 2010). This preliminary sequencing analysis focused on a relatively small number of clones is indispensable for the validation of the amplification spectrum endowed by the primers. For further and more in length analyses of the immune repertoires, NGS needs to be performed.

For the analysis of the gene expression profiles of the immune repertoires, we have focused on a comparison between the 4 libraries rather than using the IMGT® database (IMGT®, the international ImMunoGeneTics information system® http://www.imgt.org) (Lefranc *et al.*, 2009; Giudicelli *et al.*, 2011) as a reference. The immunoglobulin sequences available in IMGT® are compiled from different murine models, regardless of their genetic background, and are mostly issued from immunized mice since the annotations are deduced from monoclonal antibodies. In addition, the primers used for sequencing and the methods used for cloning are different for each reported antibody. These differences are sources of bias when we want to study the representation of different gene subgroups. Performing a comparison between the currently reported immune repertoires issued from Balb/C or SJL/J mice, however, is not subject to these biases since the amplification primers and the cloning method used for all 4 are identical. In our analysis, therefore, we report a relevant study of the differences in immunoglobulin gene expression between two strains of mice with different genetic backgrounds and distinct immunological states. It is well understood that an immunization event induces a series of somatic mutations in the immunoglobulin sequences of the same gene subgroup. Our results show that immunization may also lead to modifications of the distribution of expressed gene subgroups. This latter consequence of immunization is yet to be explored and is an interesting subject for further studies. Our data equally show that the genetic background of mice also influence their B cell repertoires. Regardless of the state of immunization, we observe a difference in the distribution profiles of immunoglobulin genes, depending on the murine strain. Finally, we have shown that the observed significant differences are more frequent in J and D/J rather

than V segments. Knowing that these segments are located in the CDR3 sequences of immunoglobulins, this observation is coherent with the understanding that the CDR3 hypervariable region plays a crucial role in immunoglobulin diversity. In order to expand this study to a larger population of immunoglobulin sequences per library and to validate the current results on a larger scale, we need to perform a high throughput sequencing (HTS) study using the New Generation Sequencing (NGS) technology. We have already designed a number of primers needed for this analysis and the experimental steps are underway. We have also developed a collaboration with Professor Offmann (Université de Nantes, CNRS UMR 6286) in order to treat the sequencing results by automated bioinformatic tools.

It is important to mention that our comparative analysis of the four repertoires lacks the inclusion of within-group comparisons. Thus, despite the availability of the statistical test validating our results, we do not affirm that the observed gene subgroup distributions are rigorously representative of a specific strain nor a given immune status. In order to limit the variability existing in different individuals of the same population (i.e. different mice of the same strain with the same immunological background), we need to construct several libraries representing the same repertoire, originating from mice with the same genetic background and immunological state, and to perform an ANOVA statistical test including a within group study. This is, however, a difficult task due to the fact that building phage display libraries is a complex and time-consuming process. A second option would have been to pool a number of spleen samples from mice of the same population and build each of the libraries from these pools. In this way, the question of within group variability would have been resolved. However, this methodology also brings a certain practical difficulty due to the extreme fragility and instability of working with RNA. Due to the fact that we do not have access to animal research facilities on site, and that we needed to acquire the animal samples from external sources, we did not want to risk the extensive loss of RNA, and hence, focused the library construction on spleen samples from a single mouse.

III.3. Selection of catalytic antibodies

We have used two strategic targets, both inhibitors of the enzyme β-lactamase, for the purpose of selecting antibodies with a similar activity. The original goal of these selection procedures was to determine whether the different genetic backgrounds and immunological states of the four repertoires have an influence on the tendency to express catalytic antibodies. The theoretical objective was to explore whether there are particular characteristics of an immune repertoire (autoimmunity, immunization, etc.), which provide a favorable environment for the expression of catalytic antibodies. The experimental results, however, lead to the selection of five antibody sequences. This population number does not allow for a valid and statistically significant analysis of the capability of the four immune repertoires in the expression of catalytic antibodies. Despite the success of selecting five catalytic scFv, no conclusions related to this particular objective, therefore, could be drawn. The number of selected antibodies is due to an inherent bias existing in the

phage display selection procedure, which favors the selection of only a small number of fit candidates out of the many, possibly thousands, of possibilities.

In theory, such a methodology should have led to the selection of all of the antibody candidates, in an exhaustive manner, with different nucleotide sequences but the same structure or amino acid sequences. This was, however, not the case. This bias is exemplified also by a selection performed on a peptide library, which is a much simpler model of this phage display selection procedure. In a previous study done by our group, a selection procedure performed against 9G4H9 on a peptide phage display library also produced only a small number of 9G4H9-specific peptide candidates (Yribarren *et al.*, 2003). This observation is therefore not linked to this particular study but reflects a general assessment, based on the experience of our group on this technology. A possible solution to this existing bias might be the reduction of the number of selection cycles. By performing fewer selection rounds and starting the screening process earlier on in the section procedure, we might be able to reduce this bias and discover a larger number of potential candidates.

The relative small number of selected antibodies is also limiting for the potential elucidation of any particular characteristics inherent to catalytic antibodies. We cannot observe any statistically significant differences in the percentage of germline identity or the expression of rare gene subgroups between the selected catalytic antibodies. None of the selected scFv is expressed by rare gene subgroups. This is potentially due to our small sampling size. It is worth of mention, however, that thepercentage of germline identity of all selected clones are relatively high. The average germline identity of IGKV genes is 93.98±3.38%, whereas that of IGHV genes is 99.03±0.75%. Similarly, the average germline identity of IGKJ genes is 99.34±1.32%, whereas that of IGHJ is 93.91±3.85%.

The analysis of sequence homologies between the antibody fragments selected against the two targets reveals some interesting information. The average sequence homology between the clones selected against Pep90, namely P90C1, P90C2, and P90C3, determined by using the ClustalW identity matrix tool, is 82.52±7.07. The same homology between the clones selected against the penam sulfone target, namely PSC1 and PSC2, is 81.20±6.81. Therefore, no significant difference is observed between these two values. Looking at the average sequence homology between the Pep90-selected clones *versus* the penam sulfone-selected clones, however, we obtain a value of 66.60±1.45, significantly lower than either of the former values. This observation could mean that the target against which the selection has been performed potentially does introduce a certain level of sequence similarity. It is also important to note the strong homology of the light chain CDR3 for P90C1 and P90C3, even though these antibody fragments originate from two different libraries, namely the immunized Balb/C and the naive SJL/J libraries, respectively.

Following the shortcoming of our sample size of catalytic antibodies resulted from the phage display *in vitro* selection procedure, we want to develop an *in silico* strategy to achieve our original objective of studying catalytic antibodies. In this strategy, we want to perform 3D modeling, also by collaboration with Pr. Offmann, of the 400 scFv sequences presently in our disposition. We then want to perform docking experiments on the 3D

models using the same targets Pep90 and penam sulfone. Antibody fragments displaying structures with a sufficient fit to the targets used in this procedure are potentially endowed with the same β-lactamase activity. We hope to acquire a larger pool of antibodies from this *in silico* selection in order to be able to perform relevant statistical analyses on the genetic sequences of these antibodies. It is important to note, however, that such results are only based on structural models and the true catalytic nature of the *in silico* selected antibodies needs to be verified experimentally. If this experiment is successful, we can ofcourse expand the sample size of our study to the larger pool of sequences acquired from the NGS analysis.

One of the two targets of selection we have used is the penam sulfone derivative, a suicide inhibitor of the enzyme β-lactamase. In order to isolate antibody candidates that have potentially formed covalent interactions with this target, we need to eliminate such interactions in the elution step. We, therefore, need to ideally use a reducing agent in this step, which will reduce the disulfide bond present in the spacer arm of the penam sulfone derivative structure and liberate the covalently linked antibodies. We have attempted to achieve this by using an elution buffer with 50 mM DTT. However, after three rounds of selection, we did not observe an enrichment of phages and more over we observed some difficulty in the amplification of phage after elution by this buffer. We then proceeded with following the same protocol of selection used for the first procedure against Pep90, with an elution step using a pH drop. With this strategy we succeeded in the selection of two antibodies with the desired catalytic activity. Due to the shortage of time, we did not go further to try different elution buffers capable of reducing covalent interactions. However, in order to select for antibodies with potentially stronger catalytic capabilities, such experiments need to be conducted to find a strategy for isolating covalently bound antibodies.

We have tested the selected clones for β-lactamase activity by using two assays: a colorimetric assay and a fluorometric one. Three out of the five antibodies, namely P90C2, PSC1, and PSC2 have demonstrated activity in the colorimetric assay, on the substrate nitrocefin. All five antibody fragments have demonstrated activity in the fluorometric assay, on the substrate fluorocillin. Due to the nature of the targets used and the β-lactamase activity, acting with a hydrolysis of an amide bond, the selected antibodies might also demonstrate peptidase or protease activity. These possibilities need to be further examined experimentally.

An interesting point about these results is that the clones selected against the peptide Pep90, namely P90C1, P90C2, and P90C3, elicit a β-lactamase activity. This gives Pep90 a great deal of importance, indicating the ability of this small peptide to transfer important functional information between 9G4H9 and the newly selected antibodies (Figure 87). Considering the way 9G4H9 has been produced and the relation between the enzyme β-lactamase, 9G4H9, Pep90, and P90C2, this observation has important implications for the idiotypic network and the transfer of structural memory and catalytic function through this cycle of molecules. This type of cross-reactivity caused by a peptide, also exemplified by the cross-reactivity of antibodies with MBP and EBV illustrated by the group of Alexander Gabibov (Gabibov *et al.*, 2011), has implications for the importance of the role of such small

structures as peptides in different pathologies and their ability to induce the expression of catalytic antibodies. This transfer of functional information could for example explain why in certain disease states, such as MS, a high concentration of a self molecule elicits the maturation of a catalytic antibody.

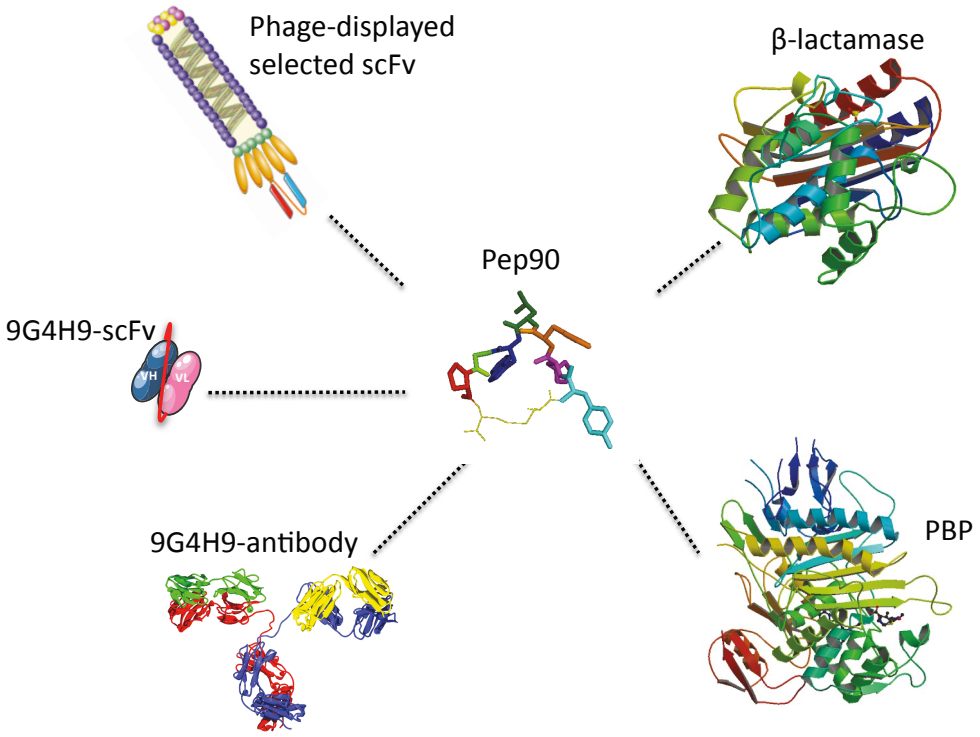


Figure 87. Pep90 and related protein catalysts.

β-lactamase pdb: 1BTL; PBP (Penicillin-Binding Protein) pdb: 1QMF.

It is also of interest to note the capability of a naive immune repertoire in inducing the production of catalytic antibodies, which is the case for the naive SJL/J repertoire leading to the isolation of catalytic clones PSC1 and PSC2. This leads to the hypothesis that naive repertoires in their natural form hold sufficient diversity to produce antibodies with any particular catalytic activity. This confirms the idea that autoreactivity exists naturally in the immune response and is in line with the other studies, which have shown that even the germline genes of immunoglobulins are endowed with the ability to produce catalytic antibodies (Uda and Hifumi, 2004; Le Minoux *et al.*, 2012).

III.4. Structural analysis

Several structural analyses of catalytic antibodies have shown that they do not possess any similarity to their enzyme counterpart at the level of amino acid sequence, but resemble

these enzymes by their three dimensional structure (Kolesnikov *et al.*, 2000; Ponomarenko *et al.*, 2007; Phichith *et al.*, 2009). Therefore, comparing the selected antibodies to the enzyme β-lactamase at a structural level could potentially elucidate some structure-function relationships.

We have performed 3D modeling of the selected scFv fragments with our collaborator Pr. Offmann. We have superposed these structures onto the active site pocket of the enzyme β-lactamase from TEM-1 (PDB: BTL1). Interestingly, we have found structural motifs similar to the active site of this enzyme for antibodies P90C2, PSC1, and PSC2, the only three antibodies demonstrating activity on nitrocefin and demonstrating the three highest activities on fluorocillin. We have not found any such structure for the antibodies P90C1 and P90C3. This might potentially be the reason for their lower catalytic activity. The reason could also be that there are other structural motifs able to catalyze this reaction, which are different from that of the enzyme β-lactamase. We admit, however, that these results are only based on models and there might be other potential structures responsible for the observed activities. The validity of these results needs to be confirmed by further crystallography studies.

In conclusion, we have constructed an antibody fragment library of size 2.7×10^9, which is a valuable reservoir of great molecular diversity and can be applied to the selection of any potential antibody candidates of interest. We have used this great diversity to select for five scFv endowed with β-lactamase activity, each having unique conformations and active-site structures. These results have further demonstrated the plasticity of the structures capable of catalyzing the β-lactam ring hydrolysis. The analysis and further characterization of these structures will provide valuable information on the different mechanisms involved in antibiotic resistance performed by different ESBL enzymes and is of potential value for the further design of novel antibiotic candidates.

ADDENDUM

Addendum

As we have seen in Section II.4, the selected scFv sequences have a relatively high conservation of sequence identity as compared with their germline. We have further analyzed these sequences, by submitting them to the IMGT/DomainGapAlign tool (Ehrenmann *et al.*, 2010; Ehrenmann and LeFranc, 2011), in order to examine the number of mutated amino acids and the extent of the modification of physicochemical properties introduced by those amino acid changes. This is similar to the analysis previously performed by our group on a larger population of catalytic antibodies reported in the literature (Le Minoux *et al.*, 2012). Table 14 summarizes these results.

Table 14. Number of somatic mutations of selected catalytic scFv.

	Light chain		Heavy chain		
scFv	Number of AA changes	% dis&vdis AA changes	Number of AA changes	% dis&vdis AA changes	VL/VH pairing AA Changes
P90C1	6	50%	2	50%	8
P90C2	7	43%	4	75%	11
P90C3	3	33%	2	50%	5
PSC1	13	69%	3	33%	16
PSC2	13	61%	1	100%	14

The number of mutated amino acids at the level of V_L/V_H pairing is relatively low, which is coherent with the results previously reported by our group (Le Minoux *et al.*, 2012). The percentage of important physicochemical modifications, however, in contrast to the previously obtained results on catalytic antibodies, does not seem to follow a particular trend. The percentage of dissimilar and very dissimilar amino acid changes ranges from 33-100%. Once again, the current population size is not large enough in order for us to be able to draw any concrete conclusions on the general nature of catalytic antibodies.

REFERENCES

REFERENCES

Abbas, A.K. Lichtman, A.H. Pillai, S. (2010). *Cellular and Molecular Immunology*, 6th edit. Saunders, Philadelphia.

Abraham, E.P. Chain, E. (1940). An enzyme from bacteria able to destroy penicillin. *Nature.* **46**, 837.

Abraham, S. Guo, F. Li, L.S. Rader, C. Liu, C. Barbas, C.F. III. Lerner, R.A. Sinha, S.C. (2007). Synthesis of the next-generation therapeutic antibodies that combine cell targeting and antibody-catalyzed prodrug activation. *Proc Natl Acad Sci U S A.* **104**, 5584-9.

Aggarwal, R. Benedetti, F. Berti, F. Buchini, S. Colombatti, A. Dinon, F. Galasso, V. Norbedo, S. (2003) A catalytic antibody programmed for tortional activation of amide bond hydrolysis. *Chemistry.* **9**, 3132-42.

Alleva, D. G., Johnson, E. B., Wilson, J., Beller, D. I., Conlon, P. J. (2001). SJL and NOD macrophages are uniquely characterized by genetically programmed, elevated expression of the IL- 12 (p40) gene, suggesting a conserved pathway for the induction of organ-specific autoimmunity. *Journal of Leukocyte Biology*, **69**, 440–8.

Almagro, J.C. Hernandez, I. Ramirez, M.C. Vargas-Madrazo, E. (1998). Structural differences between the repertoires of mouse and human germline genes and their evolutionary implications. *Immunogenetics*. **47**, 355-63.

Amagai, T., Cinader, B. (1981). Resistance of MRL / Mp-lpr / lpr mice to tolerance induction. *European Journal of Immunology*, **11**, 923–6.

Arnold, J.N. Wormald, M.R. Sim, R.B. Rudd, P.M. Dwek, R.A. (2007). The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu Rev Immunol*. **25**, 21-50.

Avalle, B. Débat, H. Friboulet, A. Thomas, D. (2000). Catalytic mechanism of an abzyme displaying a beta-lactamase-like activity. *Appl Biochem Biotechnol.* **83**, 163-71.

Avalle, B. Thomas, D. Friboulet, A. (1998). Functional mimicry: elicitation of a monoclonal anti-idiotypic antibody hydrolysing β-lactams. *FASEB J*. **12**, 1055-60.

Avalle, B. Vanwetswinkel, S. Fastrez, J. (1997). *In vitro* selection for catalytic turnover from a library β-lactamase mutants and penicillin-binding proteins. *Bioorganic and Medicinal Chemistry.* **7**, 479-84.

Azzazy, H.M. Highsmith, W.E. Jr. (2002). Phage display technology: clinical applications and recent innovations. *Clinc Biochem*. **35**, 425-45.

Baldwin, E. Schultz, P.G. (1989). Generation of a catalytic antibody by site-directed mutagenesis. *Science.* **245**, 1104-7.

Banér, J. Nilsson, M. Mendel-Hartvig, M. Landegren, U. (1998). Signal amplification of padlock probes by rolling circle replication. *Nucleic Acids Res.* **26**, 5073-8.

Baranovskii, A.G. Kanyshkova, T.G. Mogelnitskii, A.S. Naumov, V.A. Buneva, V.N. Gusev, E.I. Boiko, A.N. Zargarova, T.A. Favorova, O.O. Nevinsky, G.A. (1998). Polyclonal antibodies from blood and cerebrospinal flui of patients with multiple sclerosis effectively hydrolyze DNA and RNA. *Biochemistry (Mosc).* **63**, 1239-48.

Barbas, C.F. III. Bain, D. Hoekstra, D.M. Lerner, R.A. (1993). Semisynthetic combinatorial antibody libraries: A chemical solution to the diversity problem. *Proc Natl Acad Sci U S A.* **89**, 4457-61.

Barbas, C.F. III. Burton, D.R. Scott, J.K. Silverman, G.J. (2001). *Phage Display A Laboratory Manual.* Cold Spring Harbor, New York.

Barbas, C.F. III. Heine, A. Zhong, G. Hoffman, T. Gramatikova, S. Bjornestedt, R. List, B. Anderson, J. Stura, E.A. Wilson, I.A. Lerner, R.A. (1997). Immune versus natural selection: antibody aldolases with enzymatic rates but broader scope. *Science.* **278**, 2085-92.

Barbas, C.F. III. Kang, A.S. Lerner, R.A. Benkovic, S.J. (1991). Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc Natl Acad Sci U S A.* **88**, 7978-82.

Barrera, G.J. Portillo, R. Mijares, A. Rocafull, M.A. del Castillo, J.R. Thomas, L.E. (2009). Immunoglobulin A with protease activity secreted in human milk activates PAR-2 receptors of intestinal epithelial cells HT-29, and promotes beta-defensin-2 expression. *Immunol Lett.* **123**, 52-9.

Bazan, J. Calkosinski, I. Gamian, A. (2012). Phage display-a powerful technique for immunotherapy: 1. Introduction and potential therapeutic applications. *Hum Vaccin Immunother*. **8**, 1827-28.

Belogurov, A.A. Jr. Kozyr, A. Ponomarenko, N. Gabibov, A.G. (2009). Catalytic antibodies: balancing between Dr. Jeckyl and Mr. Hyde. *BioEssays.* **31**, 1161-71.

Belogurov, A.A. Jr. Kurkova, I.N. Friboulet, A. Thomas, D. Misikov, V.K. Zakharova, M.Y. Suchkov, S.V. Kotov, S.V. Alehin, A.I. Avalle, B. Souslova, E.A. Morse, H.C. III. Gabibov, A.G. Ponomarenko, N.A. (2008). Recognition and degredation of myelin basic protein peptides by serum autoantibodies: novel biomarker for multiple sclerosis. *J Immunol.* **180**, 1258-67.

Ben Naya, R. Matti, K. Guellier, A. Matagne, A. Boquet, D. Thomas, D. Friboulet, A. Avalle, B. Padiolleau-Lefèvre, S. (2013). Efficient refolding of a recombinant abzyme: Structural and catalytic characterizations. *Appl Microbiol Biotechnol.* **97**, 7721-31.

Bendelac, A. Bonneville, M. Kearney, J.F. (2001). Autoreactivity by design: innate B and T lymphocytes. *Nat Rev Immunol.* **1**, 177-86.

Benedetti, F. Berti, F. Brady, K. Colombatii, A. Pauletto, A. Pucillo, C. Thomas, N.R. (2004). An unprecedented catalytic motif revealed in the model structure of amide hydrolyzing antibody 312d6. *Chembiochem.* **5**, 129-31.

Berek, C. and Milstein, C. (1988). The dynamic nature of the antibody repertoire. *Immunol Rev*. **105**, 5-26.

Bertolaccini, M.L. Atsumi, T. Khamashta, M.A. Amengual, O. Hughes, G.R. (1998). Autoantibodies to human prothrombin and clinical manifestation sin 207 patients with systemic lupus erythematosus. *J Rheumatol.* **25**, 1104-8.

Bevan, M.J. (2004). Helping the CD8+ T-cell response. *Nature Rev Immunol*. **4**, 595-602.

Blackburn, G.M. Garçon, A. (2000). Catalytic antibodies. In *Biotechnology: Biotransformations.* Volume 8b. 2nd Edit. John Wiley & Sons Ltd. Chichester, UK.

Blanco, L. Bernad, A. Lázaro, J.M. Martin, G. Garmendia, C. Salas, M. (1989), Highly efficient DNA sybthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. *J Biol Chem.* **264**, 8935-40.

Boder, E.T. Wittrup, K.D. (1997) Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol*. **15**, 553-7.

Boyd, S.D. Gaeta, B.A. Jackson, K.J. Fire, A.Z. Marshall, E.L. Merker, J.D. Maniar, J.M. Zhang, L.N. Sahaf, B. Jones, C.D. Simen. B.B. Hanczaruk, B. Nguyen, K.D. Nadeau, K.C. Egholm, M. Miklos, D.B. Zehnder, J.L. Collins, A.M. (2010). Individual variation in the germline Ig gene repertoire inferred from variable region gene rearrangements. *J Immunol*. **184**, 6986-92.

Bradbury, A.R. Marks, J.D. (2004). Antibodies from phage antibody libraries. *J Immunol Methods.* **290**, 29-49.

Bradbury, A.R.M. Sidhu, S. Dubel, S. McCafferty, J. (2011). Beyond natural antibodies: the power of *in vitro* display technologies. *Nat Biotechnol*. **29**, 245-54.

Braisted, A.C. Schultz, P.G. (1990). An antibody-catalyzed biomolecular Diels-Alder reaction. *J Am Chem Soc.* **112**, 7430-1.

Briscoe, R.J. Jeanville, P.M. Cabrera, C. Baird, T.J. Woods, J.H. Landry, D.W. (2001). A catalytic antibody against cocaine attenuates cocaine's cardiovascular effects in mice: a dose and time course analysis. *Int Immunopharmacol.* **1**, 1189-98.

Brochet, X. LeFranc, M.P. Giudicelli, V. (2008) IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res.* **36**, W503-8.

Brown A.M., McFarlin D.E. (1981). Relapsing experimental allergic ecephalomyelitis in the SJL/J mouse. *Laboratory Investigation*, **45**, 278-84.

Brown, D.M. Dilzer, A.M. Meents, D.L. Swain, S.L. (2006). CD4 T cell mediated protection from lethal influenza: perforin and antibody-mediated mechanisms give a one-two punch. *J Immunol.* **177**, 2888-98.

Bush, K. Jacoby, G.A. (2010). Updated functional classification of beta-lactamases. *Antimicrob Agents Chemother.* **54**, 969-76.

Cantu, C. III. Palzkill, T. (1998). The role of residue 238 of TEM-1 beta-lactamase in the hydrolysis of extended-spectrum antibiotics. *J Biol Chem.* **273**, 26603-9.

Casali, P. Notkins, A.K. (1989). CD5+ B lymphocytes, polyreactive antibodies and the human B-cell repertoire. *Immunol Today.* **10**, 364-8.

Casola, S. Otipoby, K.L. Alimzhanov, M. Humme, S. Uyttersprot, N. Kutok, J.L. Carroll, M.C. Rajewsky, K. (2004). B cell receptor signal strength determines B cell fate. *Nat Immunol*. **5**, 317-27.

Chardès, T. Villard, S. Ferrières, G. Piechaczyk, M. Cerutti, M. Devauchelle, G. Pau, B. (1999). Efficient amplification and direct sequencing of mouse variable regions from any immunoglobulin gene family. *FEBS Lett*. **452**, 386-94.

Chasteen, L. Auriss, J. Pavlik, P. Bradbury, A.R. (2006). Eliminating helper phage from phage display. *Nucleic Acids Res.* **34**, e145.

Cheng, W. Yan, F. Ding, L. Ju, H. Yin, Y. (2010). Cascade signal amplification strategy for subattomolar protein detection by rolling circle amplification and quantum dots tagginf. *Anal Chem.* **82**, 3337-42.

Choudhry, V. Zhang, M.Y. Sidorov, I.A. Lois, J.M. Harris, I. Dimitrov, A.S. Bouma, P. Cham, F. Choudhary, A. Rybak, S.M. Fouts, T. Montefiori, D.C. Broder, C.C. Quinnan, G.V. Jr. Dimitrov, D.S. (2007). Cross-reactive HIV-1 neutralizing monoclonal antibodeis selected by screening of an immune human phage library against an envelope glycoprotein (gp140) isolated from a patient (R2) with broadly HIV-1 neutralizing antibodies. *Virology.* **363**, 79-90.

Christ, D. Famm, K. Winter, G. (2006). Tapping diversity lost in transformations-*in vitro* amplification of ligation reactions. *Nucleic Acids Res.* **34**, e108.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research*. ***16***, 10881–90.

Davidson, A. Diamond, B. (2001). Autoimmune diseases. *N Engl J Med*. **345**, 340-50.

De Wals, P.Y. Doucet, N. Pelletier, J.N. (2009). High tolerance to simultaneous active-site mutations in TEM-1 beta-lactamase: Distinct mutational paths provide more generalized beta-lactam recognition. *Protein Sci.* **18**, 147-60.

Dean, F.B. Nelson, J.R. Giesler, T.L. Lasken, R.S. (2001). Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* **11**, 1095-9.

Débat, H. Avalle, B. Chose, O. Sarde, C.O. Friboulet, A. Thomas, D. (2001). Overpassing an aberrant V_{κ} gene to sequence an anti-idiotypic abzyme with β-lactamase-like activity that could have linkage with autoimmune diseases. *FASEB J.* **15**, 815-22.

Dempsey, P.W. Vaidya, S.A. Cheng, G. (2003). The art of war: Innate and adaptive immune responses. *Cell Mol Life Sci*. **60**, 2604-21.

Deng, S.X. de Prada, P. Landry, D.W. (2002). Anticocaine catalytic antibodies. *J Immuno Methods.* **269**, 299-310.

Dickerson, T.J. Yamamoto, N. Janda, K.M. (2004). Antibody-catalyzed oxidative degradation of nicotine using riboflavin. *Bioorg Med Chem.* **12**, 4981-7.

Dildrop, R. Gause, A. Muller, W. Rajewsky, K. (1987). A new V gene expresses in lambda-2 light chains of the mouse. *Eur J Immunol*. **5**, 731-4.

Di Noia, J.M. Neuberger, M.S. (2007). Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem*. **76**, 1-22.

Ding, C. Liu, H. Wang, N. Wang, Z. (2012). Cascade signal amplification strategy for the detection of cancer cells by rolling circle amplification and nanoparticles tagging. *Chem Communications.* **48**, 5019-21.

Dormont, D. (1998). Biology of non-conventional transmissible agents or prions. *Rev Neurol*. **154**, 142-51.

Doucet, N. De Wals, P.Y. Pelletier, J.N. (2004). Site-saturation mutagenesis of Tyr-105 reveals its importance in substrate stabilization and discrimination in TEM-1 beta-lactamase. *J Biol Chem.* **279**, 46295-303.

Doucet, N. Pelletier, J.N. (2007). Simulated annealing exploration of an active-site tyrosine in TEM-1 beta-lactamase suggests the existence of alternate conformations. *Proteins.* **69**, 340-8.

Drawz, S.M. Bonomo, R.A. (2010). Three Decades of β-lactamase inhibitors. *Clin Microbiol Rev.* **23**, 160-201.

Dubus, A. Normark, S. Kania, M. Page, M.G. (1994). The role of tyrosine 150 in catalysis of beta-lactam hydrolysis by AmpC beta-lactamase from Escherichia coli investigated by site-directed mutagenesis. *Biochemistry.* **33**, 8577-86.

Durova, O.M. Vorobiev, I.I. Smirnov, I.V. Reshetnyak, A.V. Telegin, G.B. Shamborant, O.G. Orlova, N.A. Genkin, D.D. Bacon, A. Ponomarenko, N.A. Friboulet, A. Gabibov, A.G. (2009). Strategies for induction of catalytic antibodies toward HIV-1 glycoprotein gp120 in autoimmune prone mice. *Mol Immunol.* **47**, 87-95.

Ehrenmann, F. Kaas, Q. Lefranc, M.P. (2010). IMGT/3Dstructure-DB and IMGT/DomainGapAlign: a database and a tool for immunoglobulins or antibodies, T cell receptors, MHC, IgSF and MhcSF. *Nucleic Acids Res.* **38**, D301-7.

Ehrenmann, F. Lefranc, M.P. (2011). IMGT/DomainGapAlign: IMGT standardized analysis of amino acid sequences of cariable, constant, and groove domains (IG, TR, MH, IgSF, MhSF). *Cold Spring Harb Protoc.* **2011**, 737-49.

Ersoy, O. Fleck, R. Sinskey, A. Masamune, S. (1998). Antibody catalyzed cleavage of an amide bond using an external nucleophilic cofactor. *J Am Chem Soc.* **120**, 817-8.

Essono, S., Frobert, Y., Grassi, J., Creminon, C., & Boquet, D. (2003). A general method allowing the design of oligonucleotide primers to amplify the variable regions from immunoglobulin cDNA. *J Immunol Methods,* **279**, 251–66.

Ferreira, C. Singh, Y. Furmanski, A.L. Wong, F.S. Garden, O.A. Dyson, J. (2009). Non-obese diabetic mice select a low-diversity repertoire of natural regulatory T cells. *Proc Natl Acad Sci U S A.* **106**, 8320-5.

Finlay, W.J. Almagro, C. (2012). Natural and man-made V-gene repertoires for antibody discovery. *Front Immunol*. **3**, 342.

Finn, M.G. Lerner, R.A. Barbas, C.F. (1998). Cofactor induced refinement of catalytic antibody activity: A metal-specific allosteric effect. *J Am Chem Soc.* **120**, 2963-4.

Fire, A. Xu, S.Q. (1995). Rolling replication of short DNA circles. *Proc Natl Acad Sci U S A.* **92**, 4641-5.

Fischer, N. (2011). Sequencing antibody repertoires. *mAbs.* **3**, 17-20.

Fisher, J. Charnas, R.L. Bradley, S.M. Knowles, J.R. (1981). Inactivation of the RTEM beta-lactamase from Escherichia coli. Interaction of penam sulfones with enzyme. *Biochemistry.* **20**, 2726-31.

Fisher, J.F. Mobashery, S. (2009). Three decades of the class A beta-lactamase acyl-enzyme. *Curr Protein Pept Sci.* **10**, 401-7.

Fitting, J. Killian, D. Junghanss, C. Willenbrock, S. Murua, E.H. Lange, S. Nolte, I. Barth, S. Tur, M.K. (2011). Generation of recombinant antibody fragments that target canine dendritic cells by phage display technology. *Vet Comp Oncol.* **9**, 183-95.

Fletcher, M.C. Kuderova, A. Cygler, M. Lee J.S. (1998). Creation of a ribonuclease abzyme through site-directed mutagenesis. *Nat Biotechnol.* **16**, 1065-7.

Friboulet, A. Avalle, B. Débat, H. Thomas, D. (1999). A possible role of catalytic antibodies in metabolism. *Immunol Today.* **20**, 474-5.

Friehs, K. (2004). Plasmid copy number and plasmid stability. *Adv Biochem Eng Biotechnol*. **86**, 47-82.

Frohman, M.A. (1990). *PCR protocols: A guide to methods and applications.* Academic Press. San Diego.

Gabibov, A.G. Belogurov, A.A. Jr. Lomakin, Y.A. Zakharova, M.Y. Avakyan, M.E. Dubrovskaya, V.V. Smirnov, I.V. Ivanov, A.S. Molnar, A.A. Gurtsevitch, V.E. Diduk, S.V. Smirnova, K.V. Avalle, B. Sharanova, S.N. Tramontano, A. Friboulet, A. Boyko, A.N. Ponomarenko, N.A. Tikunova, N.V. (2011). Combinatorial antibody library from multiple sclerosis patients reveals antibodies that cross-react with myelin basic protein and EBV antigen. *FASEB J.* **25**, 4211-21.

Gabibov, A.G. Friboulet, A. Thomas, D. Demin, A.V. Ponomarenko, N.A. Vorobiev, I.I. Pillet, D. Paon, M. Alexandrova, E.S. Telegin, G.B. Reshetnyak, A.V. Grigorieva, O.V. Gnuchev, N.V. Malishkin, K.A. Genkin, D.D. (2002). Antibody proteases: induction of catalytic response. *Biochemistry.* **67**, 1168-79.

Gach, J.S. Quendler, H. Ferko, B. Katinger, H. Kunert, R. (2008). Expression, purification, and in vivo administration of a promising anti-idiotypic HIV-1 vaccine. *Mol Biotechnol.* **39**, 119-25.

Gadala-Maria, D. Yaari, G. Uduman, M. Kleinstein, S.H. (2015). Automated analysis of high-throughput B cell sequencing data reveals a high frequency of novel immunoglobulin V gene segment alleles. *Proc Natl Acad Sci U S A.* **112**, E862-70.

Galleni, M. Lamotte-Brasseur, J. Raquet, X. Dubus, A. Monnaie, D. Knox, J.R. Frère, J.M. (1995). The enigmatic catalytic mechanism of active-site serine beta-lactamases. *Biochem Pharmacol.* **49**, 1171-8.

Gao, Q.S. Sun, M. Tyutyulkova, S. Webster, D. Rees, A. Tramontano, A. Massey, R.J. Paul, S. (1994). Molecular cloning of a proteolytic antibody light chain. *J Biol Chem.* **269**, 32389-93.

Georgiou, G. Stathopoulos, C. Daugherty, PSP. Nayak, A.R. Iverson, B.L. Curtiss, R. 3rd. (1997). Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat Biotechnol*. **15**, 29-34.

Giefing-Kroll, C. Berger, P. Lepperdinger, G. Grubeck-Loebenstein, B. (2015). How sex and age affect immune responses, susceptibility to infections, and response to vaccination. *Aging Cell.* **14**, 309-21.

Giudicelli, V. Brochet, X. Lefranc, M.P. (2011). IMGT/V-QUEST: IMGT standardized analysis of the immunoglobulin (IG) and T cell receptor (TR) nucleotide sequences. *Cold Spring Harbor protocols*. **6**, 695–715.

Giudicelli, V Chaume, D. Lefranc, M.P. (2005). IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res.* **33**, D256-61.

Giudicelli, V. Duroux, P. Ginestoux, C. Folch, G. Jabado-Michaloud, J. Chaume, D. Lefranc, M.P. (2006). IMGT/LIGM-DB, the IMGT comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. *Nucleic Acids Res.* **34**, D781-4.

Glanville, J. Zhai, W. Berka, J. Telman, D. Huerta, G. Mehta, G.R. Ni, I. Mei, L. Sundar, P.D. Day, G.M.R. Cox, D. Rajpal, A. Pons, J. (2009). Precise determination of the diversity of a combinatorial library gives insight into the human immunoglobulin repertoire. *Proc Natl Acad Sci U S A.* **106**, 20216-21.

Gololobov, G. Sun, M. Sudhir, P. (1999). Innate antibody catalysis. *Mol Immunol.* **36**, 1215-22.

Golemi-Kotra, D. Meroueh, S.O. Kim, C. Vakulenko, S.B. Bulychev, A. Stemmler, A.J. Stemmler, T.L. Mobashery, S. (2004). The importance of a critical protonation state and the fate of the catalytic steps in class A beta-lactamases and penicillin-binding proteins. *J Biol Chem.* **279**, 34665-73.

Goncalves, O. Dintinger, T. Blanchard, D. Tellier, C. (2002). Functional mimicry between anti-tendamistat antibodies and α-amylase. *J Immunol Methods.* **269**,29-37.

Gorelick, D.A. (2012). Pharmacokinetic strategies for treatment of drug overdose and addiction. *Future Med Chem.* **4**, 227-43.

Gorny, M.K. Stamatatos, L. Volsky, B. Revesz, K. Williams, C. Wang, X.H. Cohen, S. Staudinger, R. Zolla-Pazner, S. (2005). Identification of a new quaternary neutralizing epitope on human immunodeficiency virus type 1 virus particles. *J Virol.* **79**, 5232-7.

Goswami, R.K. Huang, Z.Z. Forsyth, J.S. Felding-Habermann, B. Sinha, S.C. (2009). Multiple catalytic alsolase antibodies suitable for chemical programming. *Bioorg Med Chem Lett.* **19**, 3821-4.

Griffiths, A.D. Williams, S.C. Hartley, O. Tomlinson, I.M. Waterhouse, P. Crosby, W.L. Kontermann, R.E. Jones, P.T. Low, N.M. Allison, T.J. (1994). Isolation of high affinity human antibodies directly from large synthetic repertoires. *EMBO J.* **13**, 3245-60.

Grippo, V. Mahler, E. Elias, F.E. Cauerhff, A. Gomez, K.A., Tentori, M.C. Ruiz, A. Vigliano, C.A. Laguens, R.P. Berek, C. Levin, M.J. (2009). The heavy chain variable segment gene repertoire in chronic Chagas' heart disease. *J Immunol*. **183**, 8015-25.

Gronwall, C. Kosakovsky Pond S.L. Young, J.A. Silverman, G.J. (2012). *In vivo* VL-targeted microbial superantigen induced global shifts in the B cell repertoire. *J Immunol.* **189**, 850-9.

Guillén Schlippe, Y.V. Hedstrom, L. (2005). Guanidine derivatives rescue the Arg418Ala mutation of Tritrichomonas foetus IMP dehydrogenase. *Biochemistry.* **44**, 16695-700.

Hallborn, J. Carlsson, R. (2002). Automated screening procedure for high-throughput generation of antibody fragments. *Biotechniques*. Suppl: 30-7.

Hammes-Schiffer, S. Benkovic, S.J. (2006). Relating protein motion to catalysis. *Annul Rev Biochem.* **75**, 519-41.

Hampe, C.S. (2012). Protective role of anti-idiotypic antibodies in autoimmunity - Lessons for type 1 diabetes. *Autoimmunity.* **45**, 320-31.

Hasserodt, J. Janda, K.D. Lerner, R.A. (2000). A class of 4-aza-lithocholic acid-derived haptens for the generation of catalytic antibodies with steroid synthase capabilities. *Bioorg Med Chem.* **8**, 995-1003.

Hata, M. Fujii, Y. Tanaka, Y. Ishikawa, H. Ishii, M. Neya, S. Tsuda, M. Hoshino, T. (2006). Substrate deacylation mechanisms of serine-β-lactamases. *Biol Pharm Bull.* **29**, 2151-9.

Hatzi, K. Nance, J.P. Kroenke, M.A. Bothwell, M. Haddad, E.K. Melnick, A. Crotty, S. (2015). *J Exp Med*. **212**, 539-53.

He, M. Taussig, M.J. (2002). Ribosome display: cell-free protein display technology. *Brief Funct Genomic Proteomic*. **1**, 204-12.

Herschlag, D. Natarajan, A. (2013). Fundamental challenges in mechanistic enzymology: progress toward understanding the rate enhancements of enzymes. *Biochemistry.* **52**, 250-67.

Hifumi, E. Fujimoto, N. Arakawa, M. Saito, E. Matsumoto, S. Kobayashi, N. Uda, T. (2013). Biochemical features of a catalytic antibody light chain, 22F6, prepared from human lymphocytes. *J Biol Chem.* **288**, 19558-68.

Hifumi, E. Higashi, K. Uda, T. (2010). Catalytic digestion of human tumor necrosis factor-α by antibody heavy chain. *FEBS J.* **277**, 3823-32.

Hifumi, E. Mitsuda, Y. Ohara, K. Uda, T. (2002). Targeted destruction of the HIV-1 coat protein gp41 by a catalytic antibody light chain. *J Immunol Methods.* **269**, 283-98.

Hifumi, E. Morihara, F. Hatiuchi, K. Okuda, T. Nishizono, A. Uda, T. (2008). Catalytic features and eradication ability of antibody ligh chain UA15-L against *H. pylori. J Biol Chem.* **283**, 899-907.

Hilvert, D. (2000). Critical analysis of antibody analysis. *Annu Rev Biochem*. **69**, 751-93.

Hilvert, D. Carpenter, S.H. Nared, K.D. Auditor, M.T. (1988). Catalysis of concerted reactions by antibodies: the Claisen rearrangement. *Proc Natl Acad Sci.* **85**, 4953-5.

Hilvert, D. Hill, K.W. Nared, K.D. Auditor, M.T. (1989). Antibody catalysis of the Diels-Alder reaction. *J Am Chem Soc.* **111**, 9261-2

Hogquist, K.A. Baldwin, T.A. Jameson, S.C. (2005). Central tolerance: learning self-control in the thymus. *Nat Rev Immunol*. **5**, 772-82.

Honjo, T. Kinoshita, K. Muramatsu, M. (2002). Molecular mechaism of class switch recombination: linkage with somatic hypermutation. *Annu Rev Immunol*. **20**, 165-96.

Hoogenboom, H.R. Griggiths, A.D. Johnson, K.S. Chiswell, D.J. Hudson, P. Winter, G. (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Res*. **19**, 4133-7.

Hu, R. Xie, G.Y. Zhang, X. Guo, Z.Q. Jin, S. (1998). Production and characterization of monoclonal anti-idiotypic antibody exhibiting a catalytic activity similar to carboxypeptidase. *J Biotechnol*. **61**, 109-15.

Huie, M.A. Cheung, M.C. Muench, M.O. Becerrill, B. Kan, K.W. Marks, J.D. (2001). Antibodies to human fetal erythroid cells from a nonimmune phage antibody library. *Proc Natl Acad Sci U S A.* **98**, 2682-7.

Huston, J.S. McCartney, J. Tai, M.S. Mottola-Harshorn, C. Jin, D. Warren, F. Keck, P. Oppermann, H. (1993). Medical applications of single-chain antibodies. *Int Rev Immunol*. **10**, 195-217.

Izadyar, L. Friboulet, A. Remy, M.H. Roseto, A. Thomas, D. (1993). Monoclonal anti-idiotypic antibodies as functional internal images of enzyme active sites: production of a catalytic antibody with a cholinesterase activity. *Proc Natl Acad Sci U S A.* **90**, 8876-80.

Jackson, K.J. Kidd, M.J. Wang, Y. Collins, A.M. (2013). The shape of the lymphocyte receptor repertoire: lessons from the B cell receptor. *Front Immunol*. **4**, 263.

Jackson, K.J. Liu, Y. Roskin, K.M. Glanville, J. Hoh, R.A. Seo, K. Marshall, E.L. Gurley, T.C. Moody, M.A. Haynes, B.F. Walter, E.B. Liao, H.X. Albrecht, R.A. Garcia-Sastre, A. Chaparro-Riggers, J. Raipal,A. Pons, J. Simens, B.B. Hanczaruk, B. Dekker, C.L. Laserson, J. Koller, D. Fire, A.Z. Boyd, S.D. (2014). Human responses to influenza vaccination show seroconversion signatues and convergent antibody rearrangements. *Cell Host Microbe.* **16**, 105-14.

Jacob, F. Joris, B. Lepage, S. Dusart, J. Frère, J.M. (1990). Role of the conserved amino acids of the 'SDN' loop (Ser130, Asp131, Asn132) in a class A beta-lactamase studied by site-directed mutagenesis. *Biochem J.* **271**, 399-406.

Janda, K.D. Weinhouse, M.I. Schloeder, D.M. Lerner, R.A. Benkovic, S.J. (1990). Bait and switch strategy for obtaining catalytic antibodies with acyl-transfer capabilities. *J Am Chem Soc.* **112**, 1274-5.

Janeway, C.A. Jr. Travers, P. Walport, M. Shlomchik, M.J. (2001). *Immunobiology*, 5th edit. Garland Science, New York.

Jang, C. Machtaler, S. Matsuuchi, L. (2010). The role of Ig-α/β in B cell antigen receptor internalization. *Immunol Lett*. **134**, 75-82.

Jelsch, C. Mourey, L. Masson, J.M. Samama, J.P. (1993). Crystal structure of Escherichia coli TEM1 beta-lactamase at 1.8 A resolution. *Proteins.* **16**, 364-83.

Jencks, W.P. (1969). *Catalysis in chemistry and enzymology*. McGraw Hill Book Co., Inc., New York.

Jerne, N.K. (1974a). Clonal selection in a lymphocyte network. *Soc Gen Physiol Ser.* **29**, 39-48.

Jerne, N.K. (1974b). Towards a network theory of the immune system. *Annu Immunol*. **125C**, 373-89.

Kalaga, R. Li, L. O'Dell, J.R. Paul, S. (1995). Unexpected presence of polyreactive catalytic antibodies in IgG from unimmunized donors and decreased levels in rheumatoid arthritis. *J Immunol.* **155**, 2695-702.

Kanyshkova, T.G. Semenov, D.V. Khlimankov, D.Y. Buneva, V.N. Nevinsky, G.A. (1997). DNA-hydrolyzing activity of the light chain of IgG antibodies from milk of healthy human mothers. *FEBS Lett.* **416**, 23-6.

Kaplinsky, J. Li, A. Sun, A. Coffre, M. Koralov, S.B. Arnaout, R. (2014). Antibody repertoire deep sequencing reveals antigen-independent selection in maturing B cells. *Proc Natl Acad Sci U S A*. **111**, E2622-9.

Karle, S. Planque, S. Nishiyama, Y. Taguchi, H. Zhou, Y.X. Salas, M. Lake, D. Thiagarajan, P. Arnett, F. Handon, C.V. Paul, S. (2004). Cross-clade HIV-1 neutralization by an antibody fragment from a lupus phage display library. *AIDS.* **18**, 329-31.

Kay, B.K. Winter, J. McCafferty, J. (1996). *Phage display of peptides and proteins.* 44th edit. United Kingdom Press. San Diego.

Khan, S.A. (1997). Rolling -circle replication of bacterial plasmids. *Microbiol Mol Biol Rev*. **61**, 442-55.

Kindt, T.J. Goldsby, R.A. Osborne, B.A. (2008). *Immunology*, 6th edit. W.H. Freeman and Company, New York and Basingstoke.

Kircher, M. Kelso, J. (2010). High-throughput DNA sequencing--concepts and limitations. *Bioessays.* **32**, 524-36.

Kit, Y.Y. Semenov, D.V. Nevinsky, G.A. (1996). Phosphorylation of different human milk proteins by human catalytic secretory immunoglobulin A. *Biochem Mol Biol Int.* **39**, 521-7.

Klotz, L. Meuth, S.G. Wiendl, H. (2012). Immune mechanisms of new therapeutic strategies in multiple sclerosis-A focus on alemtuzumab. *Clin Immunol.* **142**, 25-30.

Kluge, A.F. Petter, R.C. (2010). Acylating drugs: redesigning natural covalent inhibitors. *Curr Opin Chem Biol.* **14**, 421-7.

Kohler, G. Milstein, C. (1975). Continuous culture of fused cells secreting antibody or predefined specificity. *Nature.* **256**, 495-7.

Kolesnikov, A.V. Kozyr, A.V. Alexandrova, E.S. Koralewski, F. Demin, A.V. Titov, M.I. Avalle, B. Tramontano, A. Paul, S. Thomas, D. Gabibov, A.G. Friboulet, A. (2000). Enzyme mimicry by the antiidiotypic antibody approach. *Proc Natl Acad Sci U S A.* **97**, 13526-31.

Kontermann, R. Dubel, S. (2010). *Antibody Engineering.* 2nd edit. Springer. Heidelberg, Dordrecht, London, New York.

Kozyr, A.V. Kolesnikov, A.V. Zelenova, N.A. Sashchenko, L.P. Mikhalap, S.V. Bulina, M.E. Ignatova, A.N. Favorov, P.V. Gabibov, A.G. (2000). Autoantibodies to nuclear antigens: correlation between cytotoxicity and DNA-hydrolyzing activity. *Appl Biochem Biotechnol.* **83**, 255-69.

Krebber, A. Plueckthun, A. Bornhauser, S. Burmester, J. Honegger, A. Willuda, J. Bosshar, H. R. (1997). Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J Immunoll Methods.* **201**, 35–55.

Kretzschmar, T. von Rudem T. (2002). Antibody discovery: phage display. *Curr Opin Biotechnol.* **13**, 598-602.

Kumar, H. Kawai, T. Akira, A. (2011). Pathogen recognition by the innate immune system. *Int Rev Immunol*. **30**, 16-34.

Kumar, R.K. (1983). Hodgkin's disease SJL/J murine lymphomas. *Am. J. Pathol.* **110**, 393-6.

Kunkel, H.G. Mannik, M. Williams, R.C. (1963). Individual antigenic specificities of isolated antibodies. *Science.* **140**, 1218-9.

Lacroix-Desmazes, S. Bayry, J. Kaveri, S.V. Hayon-Sonsino, S. Thorenoor, N. Charpentier, J. Luyt, C.E. Mira, J.P. Nagaraja, V. Kazatchkine, M.D. Dhainaut, J.F. Mallet Vincent, O. (2005). High levels of catalytic antibodies correlate with favourable outcome in sepsis. *Proc Natl Acad Sci U S A.* **102**, 4109-13.

Lacroix-Desmazes, S. Moreau, A. Sooryanarayana, Bonnemain, C. Stieltjes, N. Pashov, A. Sultan, Y. Hoebeke, J. Kazatchkine, M.D. Kaveri, S.V. (1999). Catalytic activity of antibodies against factor VIII in patients with hemophilia A. *Nat Med.* **5**, 1044-7.

Lamotte-Brasseur, J. Dive, G. Dideberg, O. Charlier, P. Frère, J.M. Ghuysen, J.M. (1991). Mechanism of acyl transfer by the class A serine beta-lactamase of Streptomyces albus G. *Biochem J.* **279**, 213-21.

Larkin, M.A. Blackshields, G. Brown, N.P. Chenna, R. Mcgettigan, P.A. Mcwilliam, H. Valentin, F. Wallace, I.M. Wilm, A. Lopez, R. Thompson, J.D. Gibson, T.J. Higgins, D.G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947-8.

Laserson, J. Koller, D. Davis, M.M. Fire, A.Z. Boyd, S.D. (2014). Human responses to influenza vaccination show seroconversion signatures and convergent antibody rearrangements. *Cell Host Microbe.* **16**, 105-14.

Lee, W.R. Jang, J.Y. Kim, J.S. Kwon, M.H. Kim, Y.S. (2010). Gene silencing by cell-penetrating, sequence-selective and nucleic-acid hydrolyzing antibodies. *Nucleic Acids Res.* **38**, 1596-609.

Lefèvre, S. Débat, H. Thomas, D. Friboulet, A. Avalle, B. (2001). A suicide-substrate mechanism for hydrolysis of β-lactams by an anti-idiotypic catalytic antibody. *FEBS Lett.* **489**, 25-8.

Lefranc, M.P. Giudicelli, V. Ginestoux, C. Jabado-Michaloud, J. Folch, G. Bellahcene, F.Wu, Y. Gemrot, E. Brochet, X. Lane, J. Regnier, L. Ehrenmann, F. Lefranc, G. Duroux, P. (2009) IMGT®, the international ImMunoGeneTics information system®. *Nucleic Acids Res.* **37**, D1006-D1012.

Le Minoux, D. Mahendra, A. Kaveri, S. Limnios, N. Friboulet, A. Avalle, B. Boquet, D. Lacroix-Desmazes, S. (2012). A novel molecular analysis of genes encoding catalytic antibodies. *Mol Immunol*, **50**, 160–8.

Li, J.W. Xia, L. Su, Y. Liu, H. Xia, X. Lu, Q. Yand, C. Reheman, K. (2012). Molecular imprint of enzyme active site by camel nanobodies: rapid and efficient approach to produce abzyme with aliinase activity. *J Biol Chem.* **287**, 13713-21.

Li, L. Paul, S. Tyutyulkova, S. Kazatchkine, MLDL Kaveri, S. (1995) Catalytic activity of anti-thyroglobulin antibodies. *J Immunol.* **154**, 3328-32.

Li, Y.S. Hayakawa, K. Hardy, R.R. (1993). The regulated expression of B lineage associated genes during B cell differentiation in bone marrow and fetal liver. *J Exp Med.* **178**, 951-60.

Lietz, E.J. Truher, H. Kahn, D. Hokenson, M.J. Fink, A.L. (2000). Lysine-73 is involved in the acylation and deacylation of beta-lactamase. *Biochemistry.* **39**, 4971-81.

Lipscomb, M.F. Masten, B.J. (2002). Dendritic cells: immune regulators in health and disease. *Physiol Rev.* **82**, 97-130.

Liu, E. Prasad, L. Delbaere, L.T. Waygood, E.B. Lee, J.S. (1998). Conversion of an antibody into an enzyme which cleaves the protein HPr. *Mol Immunol.* **35**, 1069-77.

Lobkovsky, E. Billings, E.M. Moews, P.C. Rahil, J. Pratt, R.F. Knox, J.R. (1994). Crystallographic structure of a phosphonate derivative of the Enterobacter cloacae P99 cephalosporinase mechanistic interpretation of a beta-lactamase transition-state analog. *Biochemistry.* **33**, 6762-72.

Luckheeram, R.V. Zhou, R. Verma, A.D. Xia, B. (2012). CD4+ T cells: Differentiation and functions. *Clin Dev Immunol*. **2012**, 925135.

Maeda, M. Ito, Y. Hatanaka, T. Hashiguchi, S. Torikai, M. Nakashima, T. Sugimura, K. (2009). Regulation of T cell response by blocking the ICOS signal with the B7RP-1-specific small antibody fragment isolated from human antibody phage library. *mAbs.* **1**, 453-61.

Maloy, K.J. Burkhart, C. Junt, T.M. Odermatt, B. Oxenius, A. Piali, L. Zinkernagel, R.M. Hengartner, H. (2000). CD4(+) T cell subsets during virus infection. Protective capacity depends on effector cytokine secretion and on migratory capability. *J Exp Med*. **191**, 2159-70.

Marks, J.D. Hoogenboom, H.R. Bonnert, T.P. McCafferty, J. Griffiths, A.D. Winter, G. (1991). By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J Mol Biol.* **222**, 581-97.

Marks, J.D. Ouwehand, W.H. Bye, J.M. Finnern, R. Gorick, B.D. Voak, D. Thorpe, S.J. Hughes-Jones, N.C. Winter, G. (1993). Human antibody fragments specific for human blood group antigens from a phage display library. *Biotechnology (NY).* **11**, 1145-9.

Martin R. McFarland H.F. (1995). Immunological aspects of experimental allergic encephalomyelitis and multiple sclerosis. *Crit Rev Clin Lab Sci.* **32**, 121-82.

Matagne, A. Frère, J.M. (1995). Contribution of mutant analysis to the understanding of enzyme catalysis: the case of class A beta-lactamases. *Biochim Biophys Acta.* **1256**, 109-27.

Matagne, A. Lamotte-Brasseur, J. Frère, J.M. (1998). Catalytic properties of class A beta-lactamases: efficiency and diversity. *Biochem J.* **330**, 581-98.

Matsuura, K. Yamamoto, K. Sinohara, H. (1994). Amidase activity of human Bence Jones proteins. *Biochem Biophys Res Commun.* **204**, 57-62.

Maveyraud, L. Pratt, R.F. Samama, J.P. (1998). Crystal structure of an acylation transition-state analog of the TEM-1 beta-lactamase. Mechanistic implications for class A beta-lactamases. *Biochemistry.* **37**, 2622-8.

McCafferty, J. Griggiths, A.D. Winter, G. Chiswell, D.J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*. **348**, 552-4.

McKenzie, K.M. Mee, J.M. Rogers, C.J. Hixon, M.S. Kaufmann, G.F. Janda, K.D. (2007). Identification and characterization of single chain anti-cocaine catalytic antibodies. *J Mol Biol.* **365**, 722-31.

Melek, M. Gellert, M. Van Gent, D.C. (1998). Rejoining of DNA by the RAG1 and RAG2 proteins. *Science.* **280**, 301-3.

Meroueh, S.O. Fisher, J.F. Schlegel, H.B. Mobashery, S. (2005). Ab initio QM/MM study of class A beta-lactamase acylation: dual participation of Glu166 and Lys73 in a concerted base promotion of Ser70. *J Am Chem Soc.* **127**, 15397-407.

Minasov, G. Wang, X. Shoichet, B.K. (2002). An ultrahigh resolution structure of TEM-1 beta-lactamase suggets a role for Glu166 as the general base in acylation. *J Am Chem Soc.* **124**, 5333-40.

Monnier, P.P. Vigouroux, R.J. Tassew, N.G. (2013) *In vivo* applications of single chain Fv (variable domain) (scFv) fragments. *Antibodies*. **2**, 193-208.

Moran, N. (2011). Boehringer splashes out on bispecific antibody platforms. *Nature Biotech*. **29**, 5-6.

Muramatsu, M. Sankaranand, V.S. Anant, S. Sugai, M. Kinoshita, K. Davidson, N.O. Honjo, T. (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center b cells. *J Biol Chem*. **274**, 18470-6.

Murphy, K.M. (2012). *Janeway's Immunobiology*, 8th edit. Garland Science, New York.

Nardi, M. Karpatkin, S. (2000). Antiidiotypic antibody against platelet anti-GPIIIa contributes to the regulation of thrombocytopenia in HIV-1-ITP patients. *J Exp Med.* **191**, 2093-100.

Nelson, J.R. Cai, Y.C. Giesler, T.L. Farchaus, J.W. Sundaram, S.T. Ortiz-Rivera, M. Hosta, L.P. (2002). TempliPhi, phi29 DNA polymerase based rolling circle amplification of templates for DNA sequencing. *BioTechniques.* Suppl, 44-7.

Neuberger, M.S. Milstein, C. (1995). Somatic hypermutation. *Curr Opin Immunol*. **7**, 248-54.

Nevinsky, G.A. Buneva, V.N. (2002). Human catalytic RNA- and DNA-hydrolyzing antibodies. *J Immunol Methods.* **269**, 235-49.

Nicholas, K.M. Wentworth, P. Jr. Harwig, C.W. Wentworth, A.D. Shafton, A. Janda, K.D. (2002). A cofactor approach to copper-dependent catalytic antibodies. *Proc Natl Acad Sci U S A.* **99**, 2648-53.

Nissim, A. Hoogenboom, H.R. Tomlinson, I.M. Flynn, G. Midgley, C. Lane, D. Winter, G. (1994). Antibody fragments from a 'single pot' phage display library as immunochemical reagents. *EMBO J.* **13**, 692-8.

Nussenzweig, M.C. Alt, F.W. (2004). Antibody diversity: one enzyme to rule them all. *Nat Med*. **10**, 1304-5.

O'Callaghan, C.H. Morris, A. Kirby, S.M. Shingler, A.H. (1972). Novel method for detection of beta-lactamases by using a chromogenic cephalosporin substrate. *Antimicrob Agents Chemother.* **1**, 283-8.

O'Nuallain, B. Allen, A. Ataman, D. Weiss, D.T. Solomon, A. Wall, J.S. (2007). Phage display and peptide mapping of an immunoglobulin light chain fibril-related conformational epitope. *Biochemistry.* **46**, 13049-58.

Orum, H. Andersen, P.S. Oster, A. Johansen, L.K. Riise, E. Bjornvad, M. Svendsen, I. Engberg, J. (1993). Efficient method for constructing comprehensive murine Fab antibody libraries displayed on phage. *Nucleic Acids Res.* **21**, 4491-8.

Oudin, J. Michel, M. (1963). Une nouvelle forme d'isiotypie des globulines γ du sérum de lapin, apparement liée à la fonction et à la spécificité des anticorps. *C R Acad Sci Paris.* **257**, 805-8.

Ozawa, T. Tajuru, K. Kishi, H. Muraguchi, A. (2008). Comprehensive analysis of the functional TCR repertoire at the single-cell level. *Biochem Biophys Res Commun.* **367**, 820-5.

Padiolleau-Lefèvre, S. Ben Naya, R. Shahsavarian M.A. Friboulet, A. Avalle, B. (2014). Catalytic antibodies and their applications in biotechnology: state of the art. *Biotechnol Lett.* **36**, 1369-79.

Padiolleau-Lefèvre, S. Débat, H. Phichith, D. Thomas, D. Friboulet, A. Avalle, B. (2006). Expression of a functional scFv fragment of an anti-idiotypic antibody with a beta-lactam hydrolytic activity. *Immunol Lett.* **103**, 39-44.

Padiolleau-Lefèvre, S. Débat, H. Thomas, D. Friboulet, A. Avalle, B. (2003). *In vivo* evolution of a β-lactamase-like activity throughout the idiotypic pathway. *Biocatal Biotransfor.* **21**, 79-85.

Pagetta, A. Tramentozzi, E. Corbetti, L. Frasson, M. Brunati, A.M. Finotti, P. (2007). Characterization of immune complexes of idiotypic catalytic and anti-idiotypic inhibitory antibodies in plasma of type 1 diabetic subjects. *Mol Immunol.* **44**, 2870-83.

Palosuo, T. Virtamo, J. Haukka, J. Taylor, P.R. Aho, K. Puurunen, M. Vaarala, O. (1997). High antibody levels to prothrombin imply a risk of deep venous thrombosis and pulmonary embolism in middle-aged men--a nested case-control study. *Thromb Haemost.* **78**, 1178-82.

Parabakaran, P. Chen, W. Singarayan, M.G. Stewart, C.C. Streaker, E. Feng, Y. Dimitrov, D.S. (2012). Expressed antibody repertoires in human cord blood cells: 454 sequencing and IMGT/HighV-QUEST analysis of germline gene usage, junctional diversity and somatic mutations. *Immunogenetics.* **64**, 337-50.

Parameswaran, P. Liu, Y. Roskin, K.M. Jackson, K.L. Dixit, V.P. Lee, J.Y. Artiles, K. Zompi, S. Vargas, M.J. Simen, B.B. Hanczaruk, B. McGowan, M.R. Tariq, M.A. Pourmand, N. Koller, D. Balmaseda, A. Boyd, S.D. Harris, E. Fire, A.Z. (2013). Convergent antibody signatures in human dengue. *Cell Host Microbe.* **13**, 691-700.

Paul, S. Karle, S. Planque, S. Taguchi, H. Salas, M. Nishiyama, Y. Handy, B. Hunter, R. Edmundson, A. Hanson, C. (2004). Naturally occurring proteolytic antibodies. *J Biol Chem.* **279**, 39611-9.

Paul, S. Li, L. Kalaga, R. O'Dell, J. Dannenbring, R.E. Jr. Swindells, S. Hinrichs, S. Caturegli, P. Rose, N.R. (1997). Characterization of thyroglobulin-directed and polyreactive catalytic antibodies in autoimmune disease. *J Immunol.* **159**, 1530-6.

Paul, S. Li, L. Kalaga, R. Wilkins-Stevens, P. Stevens, F.J. Solomon, A. (1995). Natural catalytic antibodies: Peptide-hydrolyzing activities of Bence Jones Proteins and V_L fragments. *J Biol Chem.* **270**, 15257-61.

Paul, S. Nishiyama, Y. Planque, S. Karle, S. Taguchi, H. Hanson, C. Weksler, M.E. (2005). Antibodies as defensive enzymes. *Springer Semin Immun.* **26**, 485-503.

Paul, S. Planque, S. Nishiyama, Y. (2010). Immunological origin and functional properties of catalytic autoantibodies to amyloid beta peptide. *J Clin Immunol.* Suppl 1: S43-9.

Paul, S. Planque, S. Zhou, X.Y. Taguchi, H. Bathia, G. Karle, S. Hanson, C. Nishiyama, Y. (2003). Specific HIV gp120-cleaving antibodies induved by covalently reactive analog of gp120. *J Biol Chem.* **278**, 20429-35.

Paul, S. Volle, D.J. Beach, C.M. Johnson, D.R. Powell, M.J. Massey, R.J. (1989). Catalytic hydrolysis of vasoactive intestinal peptide by human autoantibody. *Science.* **244**, 1158-62.

Pauling, L. (1948). Chemical achievement and hope for the future. *Am Sci.* **36**, 51-8.

Pavri, R. Nussenzweig, M.C. (2011). AID targeting in antibody diversity. *Adv Immunol*. **110**, 1-26.

Petit, A. Maveyraud, L. Lenfant, F. Samama, J.P. Labia, R. Masson, J.M. (1995). Multiple substitutions at position 104 of beta-lactamase TEM-1: assessing the role of this residue in substrate specificity. *Biochem J.* **305**, 33-40.

Petsko, G.A. Ringe, D. (2004). *Protein structure and function.* Chapter 2. New Science Press. USA and Canada.

Phichith, D. Bun, S. Padiolleau-Lefèvre, S. Banh, S. Thomas, D. Friboulet, A. Avalle, B. (2009). Mutational and inhibitory analysis of a catalytic antibody. Implication for drug discovery. *Mol Immunol.* **47**, 348-56.

Phichith, D. Bun, S. Padiolleau-Lefèvre, S. Guellier, A. Banj, S. Galleni, M. Frère, J.M. Thomas, D. Friboulet, A. Avalle, B. (2010). Novel peptide inhibiting both TEM-1 beta-lactamase and penicillin-binding proteins. *FEBS J.* **277**, 4965-72.

Pillet, D. Paon, M. Vorobiev, I.I. Gabibov, A.G. Thomas, D. Friboulet, A. (2002). Idiotypic network mimicry and antibody catalysis: lessons for the elicitation of efficient anti-idiotypic protease antibodies. *J Immunol Methods.* **269**, 5-12.

Planque, S. Bangale, Y. Sont, X.T. Karle, S. Taguchi, H. Poindexter, B. Bick, R. Edmundson, A. Nishiyama, Y. Paul, S. (2004). Ontogeny of proteolytic immunity: IgM serine proteases. *J Biol Chem.* **279**, 14024-32.

Planque, S. Nishiyama, Y. Taguchi, H. Salas, M. Hanson, C. Paul, S. (2008). Catalytic antibodies to HIV: physiologica role and potential clinical utility. *Autoimmun Rev.* **7**, 473-9.

Ploss, M. Kuhn, A. (2010). Kinetics of filamentous phage assembly. *Phys Biol*. **7**, 045002.

Pollack, S.J. Jacobs, J.W. Shultz, P.G. (1986). Selective chemical catalysis by an antibody. *Science.* **234**. 1570-3.

Pollack, S.J. Hsiun, P. Schultz, P.G. (1989). Stereospecific hydrolysis of alkyl esters by antibodies. *J Am Chem Soc.* **111**, 5961-2.

Ponomarenko, N.A. Durova, O.M. Vorobiev, I.I. Aleksandrova, E.S. Telegin, G.B. Chamborant, O.G. Sidorik, L.L. Suchkov, S.V. Alekberova, Z.S. Gnuchevf, N.V. Gabibov, A.G. (2002). Catalytic antibodies in clinical and experimental pathology: human and mouse models. *J Immunol Methods.* **269**, 197-211.

Ponomarenko, N.A. Durova, O.M. Vorobiev, I.I. Belugorov, A.A. Kurkova, I.N. Petrenko, A.G. Telegin, G.B. Suchkov, S.V. Kiselev, S.L. Lagarkova, M.A. Govorun, V.M. Serebryakova, M.V. Avalle, B. Tornatore, P. Karavanov, A. Morse, H.C. III. Thomas, D. Friboulet, A. Gabibov, A.G. (2006a). Autoantibodies to myelin basic protein catalyze site-specific degradation of their antigen. *Proc Natl Acad Sci U S A.* **103**, 281-6.

Ponomarenko, N.A. Pillet, D. Paon, M. Vorobiev, I.I. Smirnov, I.V. Adenier, H. Avalle, B. Kolesnikov, A.V. Kozyr, A. Thomas, D. Gabibov, A.G. Friboulet, A. (2007). Anti-idiotypic antibody mimics proteolytic function of parent antigen. *Biochemistry.* **46**, 14598-609.

Ponomarenko, N.A. Vorobiev, I.I. Aleksandrova, E.S. Reshetnyak, A.V. Telegin, G.B. Khaidudov, S.V. Avalle, B. Karavanov, A. Morse, H.C. III. Thomas, D. Friboulet, A. Gabibov, A.G. (2006b). Induction of a protein-targeted catalytic response in autoimmune prone mice: antibody-mediated cleavage of HIV-1 glycoprotein gp120. *Biochemistry.* **45**, 324-30.

Poulsen, T.R. Meijer, P.J. Jensen, A. Nielsen, L.S. Andersen, P.S. (2007). Kinetic, affinity, and diversity limits of human polyclonal antibody responses against tetanus toxoid. *J Immunol.* **179**, 3841-50.

Rader, C. (2014). Chemically programmed antibodies. *Trends Biotechnol.* **32**, 186-97.

Rangan, S.K. Liu, R. Brune, D. Planque, S. Paul, S. Sierks, M.R. (2003). Degradation of beta-amyloid by proteolytic antibody light chains. *Biochemistry.* **42**, 14328-34.

Rao, D. Wootla, B. (2007). Catalytic antibodies: Concept and promise. *Resonance (Springerlink).* **12**, 6-21.

Rathanaswami, P. Roalsad, S. Roskos, L. Su, Q.J. Lackie, S. Babcook, J. (2005). Demonstration of an in vivo generated sub-picomolar affinity fully human monoclonal antibody to interleukin-8. *Biochem Biophys Res Commun.* **334**, 1004-13.

Rau, R. (2002). Adalimumab (a fully human anti-tumour necrosis factor alpha monoclonal antibody) in the treatment of active rheumatoid arthritis: the initial results of five trials. *Ann Rheum Dis.* **61**, 70-3.

Ravn, U. Gueneau, F. Baerlocher, L. Osteras, M. Desmurs, M. Malinge, P. Magistrelli, G. Farinelli, L. Kosco-Vilbois, M.H. Fischer, N. (2010). By-passing *in vitro* screening--next generation sequencing technologies applied to antibody display and *in silico* candidate selection. *Nucelic Acids Res.* **38**, e193.

Reddy, S.T. Ge, X. Miklos, A.E. Hughes, R.A. Kang, S.H. Hoi, K.H. Chrysostomou, C., Hunicke-Smith, S.P. Iverson, B.L. Ticker, P.W. (2010). Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. *Nat Biotechnol.* **28**, 965-9.

Reshetnyak, A.V. Armentano, M.F. Morse, H.C. Friboulet, A. Makker, S.P. Tramontano, A. Knorre, V.D. Gabibov, A.G. Ponomarenko, N.A. (2007). Mechanism-dependent selection of immunoglobulin gene library for obtaining covalent biocatalysts. *Dokl Biochem Biophys*. **415**, 179-82.

Rispens, T. Ooijevaar-de Heer, P. Derksen, N.I. Wolbink, G. Van Schouwenburg, P.A. Kruithof, D. Aalberse, R.C. (2013). Nanomolar to sub-picomolar affinity measurements of antibody-antigen interactions and protein multimerizations: fluorescence-assisted high-performance liquid chromatography. *Anal Biochem.* **437**, 118-22.

Roa, S. Li, Z. Peled, J.U. Zhao, C. Edelmann, W. Scharff, M.D. (2010). MSH2/MSH6 complex promotes error-free repair of AID-induced dU:G mispairs as well as error-prone hypermutation of A:T sites. *PLoS One*. **5**, e11182.

Roberts, R.W. Szostak, J.W. (1997). RNA-peptide fusions for the *in vitro* selection of peptides and proteins. *Proc Natl Acad Sci U S A*. **94**, 12297-302.

Romesberg, F.E. Spiller, B. Schultz, P.G. Stevens, R.C. (1998). Immunological origins of binding and catalysis in a Diels-Alderase antibody. *Science.* **279**, 1929-33.

Rothlisberger, D. Honegger, A. Plueckthun, A. (2005). Domain interactions in the Fab fragment: A comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. *J Mol Biol.* **347**, 773-89.

Rudensky, A.Y. (2011). Regulatory T cells and Foxp3. *Immunol Rev*. **241**, 260-8.

Sanz, I. (1991). Multiple mechanisms participate in the generation of diversity of human H chain CDR3 regions. *J Immunol.* **147**, 1720-9.

Sapparapu, G. Plamque, S. Mitsuda, Y. McLean, G. Nishiyama, Y. Paul, S. (2012). Constant domain-regulated antibody catalysis. *J Biol Chem.* **287**, 36096-104.

Savard, P.Y. Gagné, S.M. (2006). Backbone dynamics of TEM-1 determined by NMR: evidence for a highly ordered protein. *Biochemistry.* **45**, 11414-24.

Sblattero, D. Bradbury, A. (2000) Exploiting recombination in single bacteria to make large phage antibody libraries. *Nat Biotechnol*. **18**, 75-80.

Schmid, A. Dordick, J.S. Hauer, B Kiener, A. Wubbolts, M. Witholt, B. (2001) Industrial biocatalysis today and tomorrow. *Nature.* **409**, 258-68.

Schroeder, H.W. Jr. Cavacini, L. (2010). Structure and function of immunoglobulins. *J Allergy Clin Immunol*. **125**, S41-52.

Schwartz, G.W. Hershberg, U. (2013). Germline amino acid diversity in B cell receptors is a good predictor of somatic selection pressures. *Front Immunol*. **4**, 357.

Shahsavarian, M. A. Le Minoux, D. Matti, K. M. Kaveri, S. Lacroix-Desmazes, S. Boquet, D. Friboulet, A. Avalle B. Padiolleau-Lefèvre, S. (2014). Exploitation of rolling circle amplification for the construction of large phage-display antibody libraries. *J Immunol Methods*, **407**, 26–34.

Shamis, M. Lode, H.N. Shabat, D. (2004). Bioactivation of self-immolative dendritic prodrugs by catalytic antibody 38C2. *J Am Chem Soc.* **126**, 1726-31.

Sherer, Y. Shoenfeld, Y. (2000). The idiotypic netwoek in antinuclear-antibody-associated diseases. *Int Arch Allergy Immunol*. **123**, 10-5.

Shokat, K.M. Leumann, C.J. Sugasawara, R. Schultz, P.G. (1989). A new strategy for the generation of catalytic antibodies. *Nature*. **338**, 269-71.

Shusta, E.V. Kieke, M.C. Parke, E. Kranz, D.M. Wittrup, K.D. (1999). Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency. *J Mol Biol*. **292**, 949-56.

Shuster, A.M. Gololobov, G.V. Kvashuk, O.A. Bogomolova, A.E. Smirnov, I.V. Gabibov, A.G. (1992) DNA hydrolyzing autoantibodies. *Science.* **256**, 665-7.

Sidhu, S.S. (2005). *Phage display in biotechnology and drug discovery.* CRC Press, New York.

Silverman, R.B. (1988). *Mechanism-based enzyme inactivation: Chemistry and enzymology.* CRC Press: Boca Raton, FL.

Silverstein, A.M. Rose, N.R. (2000). There is only one immne system! The view from immunopathology. *Semin Immunol*. **12**, 173-8.

Sinha, S.C. Li, L.S. Miller, G.P. Dutta, S. Rader, C. Lerner, R.A. (2004). Prodrugs of dynemicin analogs for selective chemotherapy mediated by an aldolase catalytic Ab. *Proc Natl Acad Sci U S A.* **101**, 3095-9.

Sircar, A. Kim, E.T. Gray, J.J. (2009). RosettaAntibody: antibody variable region homology modeling server. *Nucleic Acids Res.* **37**, W474-9.

Smirnov, I. Carletti, E. Kurkova, I. Nachon, F. Nicolet, Y. Mitkevich, V.A. Débat, H. Avalle, B. Belogurov, A.A. Jr. Kuznetsov, N. Reshetnyak, A. Masson, P. Tonevitsky, A.G. Ponomarenko, N. Makarov, A.A. Friboulet, A. Tramontano, A. Gabibov, A.G. (2011). Reactibodies generated by kinetic selection couple chemical reactivity with favorable protein dynamics. *Proc Natl Acad Sci.* **108**, 15954-9.

Smith, G.P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*. **228**, 1315-7.

Soderlind, E. Strandberg, L. Jirholt, P. Kobayashi, N. Alexeiva, V. Aberg, A.M. Nilsson, A. Jansson, B. Ohlin, M. Wingren, C. Danielsson, L. Carlsson, R. Borrebaeck, C.A. (2000). Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nat Biotechnol*. **18**, 852-6.

Soumillion, P. Fastrez, J. (2001). Novel concepts for selection of catalytic activity. *Curr Opin Biotechnol.* **12**, 387-94.

Soumillion, P Jespers, L. Bouchet, M. Marchand-Brynaert, J. Winter G. Fastrez, J. (1994). Selection of beta-lactamase on filamentous bacteriophage by catalytic activity. *J Mol Biol.* **237**, 415-22.

Sowek, J.A. Singer, S.B. Ohringer, S. Maley, M.F. Dougherty, T.J. Gougoutas, J.Z. Bush, K. (1991). Substitution of lysine at position 104 or 240 of TEM-1pTZ18R beta-lactamase enhances the effect of serine-164 substitution on hydrolysis or affinity for cephalosporins and the monobactam aztreonam. *Biochemistry.* **30**, 3179-88.

Stojanoski, V. Chow, D.C. Hu, L. Sankaran, B. Gilbert, H.F. Prasad, B.V. Palzkill, T. (2015). A triple mutant in the Ω-loop of TEM-1 β-lactamase changes the substrate profile via a large conformational change and an altered general base for catalysis. *J Biol Chem.* **290**, 10382-94.

Strynadka, N.C. Adachi, H. Jensen, S.E. Johns, K. Sielecki, A. Betzel, C. Sutoh, K. James, M.N. (1992). Molecular structure of the acyl-enzyme intermediate in beta-lactam hydrolysis at 1.7 A resolution. *Nature.* **359**, 700-5.

Suzuki, I. Fink, P.J. (2000). The dual functions of fas ligand in the regulation of peripheral CD8+ and CD4+ T cells. *Proc Natl Acad Sci U S A.* **97**, 1707-12.

Swain, S.L. McKinstry, K.K. Strutt, T.M. (2012). *Nature Rev Immunol*. **12**, 136-48.

Taguchi, H. Planque, S. Gopal, S. Boivin, S. Hara, M. Nishiyama, Y. Paul, S. (2008c). Exceptional Amyloid β peptide hydrolyzing activity of nonphysiological immunoglobulin variable domain scaffolds. *J Biol Chem.* **283**, 36724-33.

Taguchi, H. Planque, S. Nishiyama, Y. Symersky, J. Boivin, S. Szabo, P. Friedland, R.P. Ramsland, P.A. Edmundson, A.B. Weksler, M.E. Sudhir, P. (2008a). Autoantibody-catalyzed hydrolysis of amyloid β peptide. *J Biol Chem.* **283**, 4714-22.

Taguchi, H. Planque S. Nishiyama, Y. Szabo, P. Weksler, M.E. Friedland, R.P. Sudhir, P. (2008b). Catalytic antibodies to amyloid β peptide in defense against Alzheimer disease. *Autoimmun Rev.* **7**, 391-7.

Tanaka, F. (2002). Catalytic antibodies as designer proteases and esterases. *Chem Rev*. **102**, 4885-906.

Tanaka, F. Almer, H. Lerner, R.A. Barbas, C.F. III. (1999). Catalytic single-chain antibodies possessing β-lactamase activity selected from a phage displayed combinatorial library using a mechanism-based inhibitor. *Tetrahedron Lett.* **40**, 8063-6.

Tanaka, F. Barbas, C.F. III. (2002). Reactive immunization: a unique approach to catalytic antibodies. *J Immunol Methods.* **269**, 67-79.

Tawfik, D. S., Chap, R., Green, B. S., Sela, M., & Eshhar, Z. (1995). Unexpected high occurence of catalytic antibodies in MRL/lpr and SJL mice immunized with a transition-state analog: Is there a linkage to autoimmunity?, *Proc Natl Acad Sci U S A.* **92**, 2145–9.

Tellier, C. (2002). Exploiting antibodies as catalysts: potential therapeutic applications. *Transfus Clin Biol.* **9**, 1-8.

Thiagarajan, P. Dannenbring, R. Matsuura, K. Tramontano, A. Gololobov, G. Paul, S. (2000). Monoclonal antibody light chain with prothrombinase activity. *Biochemistry.* **39**, 6459-65.

Tomanicek, S.J. Blakeley, M.P. Cooper, J. Chen, Y. Afonine, P.V. Coates, L. (2010). Neutron diffraction studies of a class A beta-lactamase Toho-1 E166A/R274N/R276N triple mutant. *J Mol Biol.* **396**, 1070-80.

Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature*. **302**, 575-81.

Tramontano, A. Janda, K.D. Lerner, R. (1986). Chemical reactivity at an antibody binding site elicited by a mechanistic design of a synthetic antigen. *Proc Natl Acad Sci U S A.* **89**, 7114-8.

Uda, T. Hifumi, E. (2004). Super catalytic antibody and antigenase. *J Biosci Bioeng.* **97**, 143-52.

Ulrich, H.D. Mundroff, E. Santarsiero, B.D. Driggers, E.M. Stevens, R.C. Schultz, P.G. (1997). The interplay between binding energy and catalysis in the evolution of a catalytic antibody. *Nature.* **389**, 271-5.

Vanwetswinkel, S. Avalle, B. Fastrez, J. (2000). Selection of beta-lactamases and penicillin binding mutants from a library of phage displayed TEM-1 beta-lactamase randomly mutated in the active site omega-loop. *J Mol Biol.* **295**, 527-40.

Vanwetswinkel, S. Touillaux, R. Fastrez, J. Marchand-Brynaert, J. (1995). Bifunctional activity labels for selection of flamentous bateriophages displaying enzymes. *Bioorg Med Chem*. **3**, 907-15.

Venkateshrao, S. Yin, J. Jarzecki, A.A. Schultz, P.G. Spiro, T.G. (2004). Porphyrin distortion during affinity maturation of a ferrochelatase antibody, monitored by resonance raman spectroscopy. *J Am Chem Soc.* **126**, 16361-7.

Vivier, E. Tomasello, E. Baratin, M. Walzer, T. Ugolini, S. (2008). Functions of natural killer cells. *Nat Immunol.* **9**, 503-10.

Vlassov, A. Florentz, C. Helm, M. Naumov, V. Buneva, V. Nevinsky, G. Geigé, R. (1998) Characterization and selectivity of catalytic antibodies from human serum with RNase activity. *Nucleic Acids Res.* **26**, 5243-50.

Vollmer, W. Blanot, D. de Pedro, M.A. (2008). Peptidoglycan structure and architecture. *FEMS Microbiol Rev.* **32**, 149-67.

Von Boehmer, H. (1994). Positive selection of lymphocytes. *Cell.* **76**, 219-28.

Wagner, J. Lerner, R.A. Barbas C.F. III. (1995). Efficient aldolase catalytic antibodies that use the enamine mechanism of natural enzymes. *Science.* **270**,1797-800.

Wang, C. Liu, Y. Cavanagh, M.M. Le Saux, S. Qi, Q. Roskin, K.M. Looney, T.J. Lee, J. Dixit, V. Dekker, C.L. Swan, G.E. Goronzy, J.J. Boyd, S.D. (2015). B-cell repertoire responses to varicella-zoster vaccination in human identical twins. *Proc Natl Acad Sci U S A.* **112**, 500-5.

Wang, C. Liu, Y. Xu L.Y. Jackson, K.J. Roskin, K.M. Pham, T.D. Laserson, J. Marshall, E.L. Seo, K. Lee, J.Y. Furman, D. Koller, D. Dekker, C.L. Davis, M.M. Fire A.Z. Boyd, S.D. (2014). Effects of aging, cytomegalovirus infection, and EBV infection on human b cell repertoires. *J Immunol.* **192**, 603-11.

Weber, M. Bujak, E. Putelli, A. Villa, A. Matasci, M. Gualandi, L. Hemmerie, T. Wulhfard, S. Neri, D. (2014). A highly functional synthetic phage display library containing over 40 billion human antibody clones. *PLoSOne.* **9**, e100000.

Weill, J.C. Reynaud, C.A. (1996). Rearrangement/hypermutation/gene conversion: when, where and why? *Immunol Today*. **17**, 92-7.

Wentworth, P. Datta, A. Blakey, D. Boyle, T. Partridge, I.J. Blackburn, G.M. (1996). Toward antibody-directed "abzyme" prodrug therapy, ADAPT: carbamate prodrug activation by a catalytic antibody and its *in vitro* application to human tumor cell killing. *Proc Natl Acad Sci U S A.* **93**, 799-803.

Wentworth, P. Janda, K. (2001). Catalytic antibodies, structure and function. *Cell Biochem Biophys.* **35**, 63-87.

Wentworth, P. Liu, Y.Q. Wentworth, A.D. Fan, P. Foley, M.J. Janda, K.D. (1998). A bait and switch hapten strategy generates catalytic antibodies for phosphodiester hydrolysis. *Proc Natl Acad Sci U S A.* **95**, 5971-5.

Wentworth, P. McDunn J.E. Wentworth, A.D. Takeuchi, C. Nieva, J. Jones, T. Bautista, C. Ruedi, J.M. Gutierrez, A. Janda, K.D. Babior, B.M. Eschenmoser, A. Lerner R.A. (2002). Evidence for antibody-catalyzed ozone formation in bacterial killing and inflammation. *Science.* **298**, 2195-9.

Wentzell, L.M. Nobbs, T.J. Halford, S.E. (1995). The SfiI restriction endonuclease makes a four-strand DNA break at two copies of its recognition sequence. *J Mol Biol.* **248**, 581-95.

Wilke, C.M. Bishop, K. Fox, D. Zou W. (2011). Deciphering the role of Th17 cells in human disease. *Trends Immunol*. **32**, 603-11.

Wirsching, P. Ashley, J.A. Lo C.H. Janda, K.D. Lerner, R.A. (1995). Reactive immunization. *Science.* **270**, 1775-82.

Wootla, B. Christophe, O.D. Mahendra, A. Dimitrov, J.D. Repressé, Y. Ollivier, V. Friboulet, A. Borel-Delon, A. Levesque, H. Borg, J.Y. Andre, S. Bayry, J. Calvez, T. Kaveri, S.V. Lacroix-Desmazes, S. (2011a). Proteolytic antibodies activate factor IX in patients with acquired hemophilia. *Blood*. **117**, 2257-64.

Wootla, B. Dasgupta, S. Dimitrov, J.D. Bayry, J. Lévesque, H. Borg, J.Y. Borel-Delon, A. Rao, D.N. Friboulet, A. Kaveri, S.V. Lacroix-Desmazes, S. (2008a). Factor VIII hydrolysis mediated by anti-factor VIII autoantibodies in acquired hemophilia. *J Immunol.* **180**, 7714-20.

Wootla, B. Dasgupta, S. Mallet, V. Kazatchkine, M.D. Nagaraja, V. Friboulet, A. Kaveri, S.V. Lacroix-Desmazes, S. (2006). Physiopathology of catalytic antibodies: the case for factor VIII-hydrolyzing immunoglobulin G. *Blood Coagul Fibrinolysis.* **17**, 229-34.

Wootla, B. Lacroix-Desmazes, S. Warrington, A.E. Bieber, A.J. Kaveri S.V. Rodriguez, M. (2011b). Autoantibodies with enzymatic properties in human autoimmune diseases. *J Autoimmun.* **37**, 144-50.

Wootla, B. Nicoletti, A. Patey, N. Dimitrov, J.D. Legendre, C. Christophe, O.D. Friboulet, A. Kaveri, S.V. Lacroix-Desmazes, S. Thaunat, O. (2008b). Hydrolysis of coagulation factors by circulating IgG is associated with a reduced risk of chronic allograft nephropathy in renal transplanted patients. *J Immunol.* **180**, 8455-60.

Wu, T.T. Kabat, E.A. (1970). An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J Exp Med*. **132**, 211-50.

Wykes, M. (2003). Why do B cells produce CD40 ligand? *Immunol Cell Biol.* **81**, 328-31.

Xu, J. Deng, Q. Chenm J. Houk, K.N. Bartek, J. Hilvert, D. Wilson, I.A. (1999). Evolution of shape complementary and catalytic efficiency from a primordial antibody template. *Science.* **286**, 2345-8.

Xu, J. Song, J. Su, J. Wei, J. Yu, Y. Lv, S. Li, W. Nie, G. (2010). A new human catalytic antibody Se-scFv-2D8 and its selenium-containing single domains with high GPX activity. *J Mol Recognit.* **23**, 352-9.

Xu, Y. Yamamoto, N. Janda, K. (2004). Catalytic antibodies: hapten design strategies and screening methods. *Bioorg Med Chem.* **12**, 5247-68.

Yli-Kauhaluoma, J.T. Ashley, J.A. Lo, C.H. Tucker, L. Wolfe, M.M. Janda, K.D. (1995). Anti-metallocene antibodies: A new approach to enantioselective catalysis of the Diels-Alder reaction. *J Am Chem Soc.* **117**, 7041-2.

Yribarren A.S. Thomas, D. Friboulet, A. Avalle, B. (2003). Selection of peptides inhibiting β-lactamase-like-activity. *Eur J Biochem.* **270**, 2789-95.

Zein, H.S. Teixeira da Silva, J.A. Miyatake, K. (2010). Structure-finction analysis and molecular modeling of DNase catalytic antibodies. *Immunol Lett.* **129**, 13-22.

Zhang, D.Y. Brandwein, M. Hsuih, T.C. Li, H. (1998) Detection of nucleic acid targets using ramified rolling circle DNA amplification: a single nucleotide polymorphism assay model. *Gene*. **211**, 277-85.

Zhao, W. Ali, M.M. Brook, M. Li, Y. (2008). Rolling circle amplification: applications in nanotechnology and biodetection with functional nucleic acids. *Angewandte Chem.* **47**, 6330-7.

Zhong, G. Lerner, R.A. Barbas, C.F. III. (1999). Broadening the aldolase catalytic antibody repertoire by combining reactive immunization and transition state theory: New enantio- and diastereoselectivities. *Angew Chem Int Ed.* **38**, 3738-41.

Zhou, Z.H. Tzioufas, A.G. Notkins, A.L. (2007) Properties and function of polyreactive antibodies and polyreactive antigen-binding B-cells. *J Autoimmun.* **29**, 219-28.

Zocher, I. Roschenthaler, F. Kirschbaum, T. Schable, K.F. Horlein, R. Fleischmann, B. Kofler, R. Geley, S. Hameister, H. Zachau, H.G. (1995). Clustered and interspersed gene families in the mouse immunoglobulin kappa locus. *Eur J Immunol*. **25**, 3326-31.

Zuckerman, N.S. McCann, K.J. Ottensmeier, C.H. Barak, M. Shahag, G. Edelman, H. Dunn-Walters, D. Abraham, R.S. Stevenson, F.K. Mehr, R. (2010). Ig gene diversification and selection in follicular lymphoma, diffuse large B cell lymphoma and primary central nervous system lymphoma revealed by lineage tree and mutation analysis. *Int Immunol*. **22**, 875-87.

ANNEXES

ANNEXES

I. Materials and Methods ..... **ii**
I.1. Mice and Immunization ii
I.2. Library Construction ii
I.2.A. cDNA synthesis ii
I.2.B. V_L and V_H amplification by nested PCR ii
I.2.C. Assembly of scFv iii
I.2.D. Rolling Circle Amplification, cloning and transformation iii
I.3. Construction of host vectors iv
I.4. Phage Purification iv
I.4.A. Preparation of the Amplification Culture iv
I.4.B. Phage Precipitation v
I.5. Library Characterization v
I.5.A. Western Blot v
I.5.B. PCR screening v
I.6. Sequence Analysis vi
I.7. Bioinformatics vi
I.8. Statistical Analysis vi
I.9. Selection vii
I.10. Phage titration vii
I.11. Phage-ELISA vii

II. Buffers and Reagents ..... **ix**
II.1. SDS-PAGE solutions ix
II.2. Western Blot and Phage-ELISA buffers x

III. Primers ..... **xi**
Table A.1. Primers targeting the L-PART1 regions of immunoglobulin light chain κ and λ. xi
Table A.2. Primers targeting the L-PART1 regions of immunoglobulin heavy chain γ sequences. xii
Table A.3. Primers targeting FR1 regions of immunoglobulin light chain κ and λ sequences... xiii
Table A.4. Primers targeting FR1 regions of immunoglobulin heavy chain γ sequences xiv
Table A.5. Primers targeting FR4 regions of immunoglobulin light chain κ and λ sequences.... xv
Table A.6. Primers targeting FR4 regions of immunoglobulin heavy chain γ sequences. xv

IV. Pearl diagrams of 9G4H9 and selected antibody fragments. ..... **xvi**
IV.1. 9G4H9 xvi
IV.2. P90C1 xvi
IV.3. P90C2 xvii
IV.4. P90C3 xvii
IV.5. PSC1 xviii
IV.6. PSC2 xviii

V. Graphic map of phagemid vector pAK100 ..... **xix**

VI. Evolution of IMGT data between 2014 and 2015. ..... **xx**

I. Materials and Methods

I.1. Mice and Immunization

The protocol for animal work was performed according to the guidelines of the Charles Darwin ethical committee for animal experimentation (UPMC Paris) at the pathogen-free animal facility of the Cordeliers Research Center, Paris, France. Immunization of mice was performed by the society In-Cell-Art (Nantes, France). The antigen used was a biotinylated penam sulfone hapten conjugated to a KLH carrier, synthesized by the company AlphaChimica (Chatenay Malabry, France). The protocol used for the immunization was that detailed by Tawfik's group (Tawfik *et al.*, 1995). Briefly, Balb/C and SJL/J mice at 8 weeks of age were given a sub-cutanous injection of 50 µg of the antigen formulated with Freund's complete adjuvant on day 0. A second dose of 50 µg of antigen formulated with Freund's incomplete adjuvant was given on day 17. Both immunized and naive mice were sacrificed by cervical dislocation on day 21 and the spleens were immediately removed.

I.2. Library Construction

I.2.A. cDNA synthesis

Total RNA was extracted from 30 mg of isolated and frozen mouse spleen tissue (~5×10^6 cells) using the NucleoSpin® RNA II kit (Cat. 740955, Machery-Nagel) as per the manufacturer's protocol. The RNA purity and concentrations were determined by Experion™ RNA StdSens Analysis Kit (Cat. 7007103, Biorad). About 5 µg of RNA were reverse transcribed by the Phusion RT-PCR kit (Car. F5465, ThermoScientific) using 10 pmol of a specific primer mix targeting segments in the constant region of the heavy chain gamma (γ), light chains kappa (κ) and lambda (λ) (Table 4, Chapter 1). The reaction mix and thermal cycling were followed as per the manufacturer's protocol.

I.2.B. V_L and V_H amplification by nested PCR

Synthesized cDNA (5 µg) was used in a nested PCR strategy for the amplification of variable light VL and variable heavy VH sequences. For the amplification of VL in the first PCR reaction (PCR1), 5 µL of cDNA were mixed with 200 µM of dNTP, 4.5 µM of $MgSO_4$, 2 µM forward and 2 µM of back primers, 1 U of Vent® DNA polymerase (Cat. M0254S, New England Biolabs), and 5 µL of 10× buffer in a 50 µL total reaction mix. The forward primers target regions in the signal sequence L-PART1 of light chains κ and λ sequences (Table A.1 in Annexes III.). The back primers target segments in the constant regions of these sequences, same as those used in RT-PCR (Table 4 in Chapter 1: IgK and IgL). For the amplification of V_H in PCR1, the similar parameters were used as above with only the

$MgSO_4$ concentration decreased to 3.5 µM. The salt concentrations were optimized separately for each PCR reaction. The V_H forward primers targeting the L-PART1 sequences of heavy chain γ are listed in Table A.2 in Annexes III.) and the back primers in Table 3, Chapter 1 (IgG1, 2a, 2b, and 3). Both reactions were heated for 3 minutes at 95 °C before performing 25 cycles (1 min at 95 °C, 30 sec at a temperature gradient of 68-58 °C, 1 min at 75 °C). At the end of the PCR, the reaction was heated at 75 °C for 7 min. The temperature gradient for the annealing step was used due to the presence of a large number of different primers in each mix, in order to ensure the annealing of all to their corresponding sequences.

A second PCR reaction (PCR2) was followed using 1 µL of a 10 times diluted sample of the PCR1 results for V_L and 1 µL of the undiluted PCR1 results for V_H. The dilution of the VL PCR1 results is necessary prior to use in PCR2, due to the high concentration of amplified genes after PCR1. The composition of these 2 reactions is similar to PCR1 except for the salt concentrations: 6 µM $MgSO_4$ used for VL and 4.5 µM for V_H amplification. The forward primers, targeting framework 1 (FR1), for the amplification of V_H and V_L are inspired by those described in a previous publication (Essono *et al.*, 2003). The back primers targeting framework 4 (FR4) regions of IGVH, IGVK, and IGVL are also those previously described have also been modified. These primers are included in Tables A.3 to A.6 (Annexes III.). The thermal cycling was the same as that in PCR1. The final PCR products of VL and VH were then run on a 1.5% agarose gel, cut out and purified by the QIAquick Gel Extraction Kit (Cat. 28704, QIAGEN) as per the manufacturer's protocol.

I.2.C. Assembly of scFv

The purified VL and VH (approximately 5 ng/µL) were then assembled by Single Overlap Extension Polymerase Chain Reaction (SOE-PCR) in a 50 µL reaction volume using PWO Master mix (Cat. 03789403001, Roche Applied Sciences) (Shahsavarian *et al.*, 2014). Briefly, VL and VH were associated via the overlapping linker sequences, introduced by the PCR2 primers, in a first step without any additional primers. This step involved an initial heating at 94 °C, followed by 10 cycles (45 sec at 94 °C, 45 sec at 50 °C, and 1 min at 72 °C) and finished by heating at 72 °C for 7 min. In the second step, 1 µM of primers scFor (5' GGA ATT CGG CCC CCG AG 3') and scBack (5' TTA TTA CTC GCG GCC CAG CCG GCC 3') were added to the products of the first reaction and heated for 3 min at 94 °C, followed by 30 cycles (45 sec at 94 °C, 50 sec at 61 °C, 30 sec at 58 °C, and 1 min at 72 °C), and a final heating step of 7 min at 72 °C. The assembled scFv fragments were then purified by QIAquick PCR Purification Kit (Cat. 28104, QIAGEN) as per the manufacturer's protocol.

I.2.D. Rolling Circle Amplification, cloning and transformation

Rolling Circle Amplification (RCA) was performed following a protocol previously published[15]. Briefly 300 ng of assembled scFv was ligated end to end by the T4 Ligase (Cat. M1804, Promega) in order to yield circularized scFv fragments. These fragments were then

subjected to amplification by the Phi29 polymerase using the illustra TempliPhi amplification kit (Cat. 25640010, GE Healthcare). Phagemid host vectors were subjected to a multi-digestion step prior to the digestion of both phagemids and RCA amplified scFv fragments by the SfiI restriction enzyme. The scFv fragments were finally ligated into their host vectors by the T4 DNA Ligase and electroporated into 50 µL of freshly prepared electro-competent *E coli* TG1 bacteria. The transformation results were grown overnight on solid media at 37°C, then scraped off and stored in 10 mL of liquid media with 20% glycerol at -80 °C.

I.3. Construction of host vectors

The host vector for the construction libraries is the phagemid pAK100 engineered by Krebber's team (Krebber *et al.*, 1997). In order to have 4 uniquely bar-coded vectors, three different versions of the phagemid pAK100 were constructed by performing point mutations using the QuikChangeII XL Site-Directed Mutagenesis Kit (Cat. 200521, Agilent Technologies), as per the manufacturer's protocol. For each modified vector, primers were designed in order to remove a single restriction site by a nucleotide deletion. For the phagemid pAK101, the restriction site KpnI is removed by using the primers KpnFor (5' GCT TTT ATC GGG GAC CGC TCA CTG CCC 3') and KpnBack (5' GCG GGC AGT GAG CGG TCC CCG ATA AAA GC 3'). For the phagemid pAK102, the restriction site EcoRI is removed by using the primers EcoFor (5' CTT CTG CTC GAC TTC GGC CCC CGA GGC 3') and EcoBack (5' GCC TCG GGG GCC GAA GTC GAG CAG AAG 3'). Finally, for the phagemid pAK103, the restriction site BstEII was removed by using the primers BstFor (5' CGC GAT TTG CTG GTG CCC CAA TGC GAC C 3') and BstBack (5' GGT CGC ATT GGG GCA CCA GCA AAT CGC G 3').

I.4. Phage Purification

I.4.A. Preparation of the Amplification Culture

A small sample of *E coli* bacteria from each stored library was pooled together in 2×YT liquid medium with 25 µg/mL chloramphenicol and 1% glucose and grown at 37°C under agitation until the culture reached the optical density (OD) at 600 nm of 0.5. A volume of 20 mL of this culture was then infected with M13K07 helper phage (New England Biolabs) at 1:20 (phage:bacterial cell) ratio and incubated at 37°C for 30 min without agitation. The infected cells were then centrifuged at 3300 ×g for 10 min and the pellet was resuspended with 30 mL of 2×YT-25 µg/mL chloramphenicol- 25 µg/mL kanamyecin. This resuspension was completed to 300 mL with the addition of 270 mL of the same medium previously heated and grown over-night at 30°C under agitation.

I.4.B. Phage Precipitation

The amplification culture was centrifuged at 3300 ×g for 30 min. The supernatant was collected and 1/5 (v/v) of a PEG 8000 solution with 2.5 mM NaCl was added. The solution was thoroughly mixed and incubated at 4°C on ice for 1 hour. The phage were then pelleted by a centrifugation at 10800 ×g at 4°C for 30 min. The supernatant was discarded and the phage resuspended in 40 mL of H_2O. To precipitate the phage 1/5 (v/v) of the PEG solution were once again added and the suspension incubated on ice for 30 min. The phage were once again pelleted by centrifugation at 3300 ×g for 30 min and resuspended in 1-2 mL of H_2O.

I.5. Library Characterization

I.5.A. Western Blot

Purified phages from a pooled sample of the 4 libraries (~5×10^{12} phage particles) were denatured by incubation at 100 °C for 10 min in non-reducing SDS-PAGE buffer. The wild-type phage pAK100 was used as a control following the same procedure. The solutions containing denatured phage particles were mixed with Laemmli buffer and run on a 10% Tris-glycine HCl SDS-PAGE polyacrylamide gel and subsequently electrotransferred to a BioTrace™ NT Pure Nitrocellulose Blotting Membrane (Cat. 66485, PALL Life Sciences). The membrane was then blocked with a 2% milk-Phosphate buffered saline solution containing 0.05% (v/v) Tween 20 (PBST-M) for 1 hour, washed 3 times with PBS containing 0.05% (v/v) Tween 20 (PBST), and incubated with an anti-pIII M13 antibody (Cat. E8033S, New England Biolabs) in a 1/5000 dilution from stock in PBS for 1 hour. Next, the membrane was washed 3 times in PBST, incubated with a secondary anti-mouse IgG/horseradish peroxidase (HRP) conjugate antibody (Cat. BI2413C, Abliance) in a 1/5000 dilution from stock in PBS for 1 hour, and again washed 3 times. The membrane was finally developed with Amersham™ ECL™ Prime Western Blotting Detection Reagent (Cat. RPN2106, GE Healthcare) and visualized on a ChemiDoc™ XRS System (BioRad).

I.5.B. PCR screening

The percentage of successful cloning was evaluated by PCR on randomly selected 24 colonies for each of the 4 libraries according to the protocol previously described.[15] Briefly, after plating an appropriate dilution of each library on solid media in order to have isolated colonies, a pinch of each bacterial clone was added to a reaction mix for amplification with OneTaq® Quick-Load 2X Master Mix (Cat. M0486L, New England Biolabs) as per the manufacturer's protocol using primers pAK100For (5' `GGA TAA CAA TTT CAC ACA GG` 3') and pAK100Back (5' `TGT AGC GCG TTT TCA TCG GC` 3'). The results were analyzed on a 0.8% agarose gel.

I.6. Sequence Analysis

For each library, about 100 individual clones were randomly selected and separately subcultured in 3 ml of Luria-Bertani (LB) medium in the presence of 25 µg/mL chloramphenicol. After incubation overnight at 37 °C under agitation, the phagemids were purified by the QIAprep Spin Miniprep Kit (Cat. 27106, Qiagen) as per the manufacturer's protocol. The purified plasmids were sent for sequencing to Eurofins MWG operon (Ebersbarg, Germany), using the same pAK100 primers listed above.

I.7. Bioinformatics

The immunoglobulin sequences were obtained from the IMGT® database (IMGT®, the international ImMunoGeneTics information system® http://www.imgt.org) (Lefranc *et al.*, 2009; Giudicelli *et al.*, 2011) using the IMGT/LIGM-DB tool (Giudicelli *et al.*, 2006). The determination of immunoglobulin gene subgroups was performed by the IMGT/V-QUEST online tool (Brochet *et al.*, 2008). The alignment of the sequences retrieved from IMGT® was performed by the Multalin software (Corpet, 1988). The determination of the percentage of homology between sequences as well as the analysis of the phylogenetic trees of the sequences available in each library were performed by the ClustalW2 (EMBL-EBI) tool (Larkin *et al.*, 2007).

I.8. Statistical Analysis

The significance of each fragment of genes issued by the 4 libraries was analyzed by a statistical test with null hypothesis H0 : $p_i = p_j$ (for *i* and *j* libraries) against H1 : $p_i \neq p_j$, where p denotes the pooled estimated solution for both samples. The test random variable Z, for equal dimension samples, denoted by n, is

$$Z = \frac{(p_i - p_j)}{\sqrt{2p(1-p)/n}}$$

Here $p = (p_i + p_j)/2$. This random variable is asymptotically (for large samples) a standard normal under H0 hypothesis. The sample realization of Z is denoted by z^*, and for a significance level α, we reject the H0 hypothesis if $z^* > z_{\alpha/2}$ or $z^* < -z_{\alpha/2}$. The P-value of the test is : $P \approx 2P(Z \geq z^* \mid H0)$. Here z_γ is the γ quantile of the standard normal distribution.

I.9. Selection

The selection procedure is performed on straptividin-conjugated magnetic beads (Cat. S1420S, New England Biolabs). The beads (200 µg) are washed three times with PBS. The wash buffer supernatant is collected after each wash step by placing the beads in contact with a magnet. An adequate quantity of each target, determined for saturation (Chapter 2, Section 2.IV.1 & 2), is immobilized on the beads by incubation at room temperature (RT) for two hours under rotation. The Pep90 peptide stock solution is at 1.4 mM in PBS-2% Dimethyl sulfoxide (DMSO). For the immobilization of beads, 588 µL of a 2000× dilution in PBS of this stock solution is used. The penam sulfone molecule stock solution is at 20 mM in 100% DMSO. For the immobilization of beads, 400 µL of a 10000× dilution in PBS of this stock solution is used. After the 2-hour incubation, the target-immobilized beads are washed three times with PBS-2% milk (PBS-M). Phages (5×10^{12} particles for rounds 1-4; 5×10^{11} particles for the following rounds) are added to the beads and the volume is completed to 1 mL with PBS-M. The suspension is incubated for three hours at RT under rotation. The beads are then washed ten times, with a 5-minute incubation time under rotation for each washing step, with PBS-Tween 0.1% (PBS-T) to eliminate unbound phage. The bound phage are finally eluted by a pH drop with an acidic solution (0.1 M HCl-glycine pH 2.2) for 30 minutes at RT under rotation. The eluted fraction is recovered from the beads and neutralized immediately by the addition of 11% (v/v) of a 2M Tris-base solution. The recovered phage is then submitted to a titration procedure for quantification and amplified for the following selection round.

I.10. Phage titration

TG1 bacteria are grown in LB medium until they reach OD (600 nm) of 0.5. Recovered phage after each selection round is diluted in H_2O by a serial dilution of factors of 10. One-hundred µL of each dilution is used to infect 900 µL of the bacterial culture. The infected cells are incubated at 37°C for 30 min. One-hundred µL of each infected culture is plated on LB-25 µg/mL chloramphenicol solid medium and incubated over-night at 37°C. The number of colonies are counted the following day to determine the quantity of infectious phage.

I.11. Phage-ELISA

Nunc™ 96-well immuno plates (Cat. 3455, Thermo Scientific) are coated with 100 µL anti-biotin antibody (Cat. MA5-11251, Thermo Scientific) at 2 µg/mL in carbonate buffer. The plate is incubated at 37°C for 2 hours. The wells are washed 3 times with PBS-T. The well are blocked with 200 µL of PBS-BSA 1% incubated at 37°C for 2 hours. One-hundred µL of phages (5×10^{12} particles) are incubated with 100 µL of 1 µM target for 2 hours at 37°C under rotation. After washing the blocked plates 3 times with PBS-T, each phage

suspension is added to the wells in duplicates. The plate is incubated at 4°C over-night. The wells are washed the following day 15 times with PBS-T. The secondary HRP-conjugated anti-pVIII M13 antibody (Cat. 27942101, GE Healthcare) 100 µL at 2 µg/mL in PBS is added to the wells and incubated at 37°C for 2 hours. The wells are washed 3 times with PBS-T. The revelation is performed by adding 100 µL of ABTS (Cat. 37615, Thermo Scientific) is added to each well, incubated for 10 minutes, and the optical density is finally read at 405 nm.

II. Buffers and Reagents

II.1. SDS-PAGE solutions

Migration gel (10%): H_2O ---------- 8 mL
Tris-HCl, 1.5 M, pH 8.8 ---------- 5 mL
Acrylamide.Bis-acrylamide (30%) -------- 6.6 mL
SDS,10% (w/v) ---------- 200 µL
Ammonium persulfate, 10%(w/v) -------- 200 µL
TEMED ---------- 20 µL

Stacking gel (5%): H_2O ---------- 6.8 mL
Tris-HCl, 1.5 M, pH 8.8 ---------- 3 mL
Acrylamide.Bis-acrylamide (30%) ---------- 3 mL
SDS,10% (w/v) ---------- 120 µL
Ammonium persulfate, 10%(w/v) ---------- 120 µL
TEMED ---------- 12 µL

Loading buffer (Laemmli 4×): Tris, 1 M, pH 6.8 ---------- 2.4 mL
SDS ---------- 0.8 g
Bromophenol blue ---------- 0.01%
β-mercaptoethanol ---------- 1 mL
H_2O ---------- 2.8 mL

Migration buffer: Tris ---------- 25 mM
Glycine ---------- 192 mM
SDS ---------- 0.1%

Coloration solution: Coomassie blue R-250 ---------- 2 g
Ethanol ---------- 200 mL
Acetic acid ---------- 100 mL
H_2O ---------- 700 mL

De-coloration solution: Ethanol ---------- 200 mL
Acetic acid ---------- 100 mL
H_2O ---------- 700 mL

II.2. Western Blot and Phage-ELISA buffers

PBS: NaCl ---------- 137 mM
KCl ---------- 2.7 mM
Na_2HPO_4 ---------- 10 mM
KH_2PO_4 ---------- 1.8 mM

PBS-T: PBS ---------- 1X
Tween-20 ---------- 0.1%

PBS-M: PBS ----------1X
Milk powder ---------- 2%

Migration buffer: Tris ---------- 25 mM
Glycine ---------- 192 mM
SDS ---------- 0.1%

Transfer buffer: Tris ---------- 25 mM, pH 8.3
Glycine ---------- 192 mM
Methanol ---------- 20% (v/v)

Carbonate solution: Carbonate ---------- 0.5 M, pH 9.6
Na_2CO_3 ---------- 1.5 g
$NaHCO_3$ ---------- 2.93 g
H_2O ---------- upto 1L

III. Primers.

Table A.1. Primers targeting the L-PART1 regions of immunoglobulin light chain κ and λ.

Primer	Sequence	d	v (μL)
LPK1a	GYT ART GCT CTG GAT TCR GG	8	1.6
LPK1b	GCC TGT TMG GYT GTT GGT GC	4	0.8
LPK1c	GAG TCC TGT CCA CTC C	1	0.2
LPK1d	GCG GCG CCA TTT TCC CGG TTC GGG	1	0.2
LPK1e	GGG AAT TTG GGT GAG GCC	1	0.2
LPK2a	GAG GTG YTC TCY TCA KTT CYT GGR G	32	6.4
LPK2b	CTT GTG CTC TGG ATC CCT G	1	0.2
LPK2c	CAG CTY CTG RGG ATK CTT RTG C	16	3.2
LPK3	GGG TVC TGC TKC TCT GGG TTC CAG	6	1.2
LPK4a	GTG CAG ATT WTC AGC TTC CTG C	2	0.4
LPK4b	CAG STT CMT GCT WAT CAG TKT CAC AG	16	3.2
LPK4c	CCT GTT AAA CAG TGT CTC AG	1	0.2
LPK5	GGT DTY CAC ACC TCA GWT CCT TG	12	2.4
LPK6a	CRT GTT KCT STG GTT GTC TG	8	1.6
LPK6b	GRA RTC ACA GAC YCY GG	16	3.2
LPK7	CCC AGG TAC TCA TGT CCC TGC TGC	1	0.2
LPK8	GCT GYT HTG GGT RTC TG	12	2.4
LPK9	GGS WTC TTG TTG CTC TKG	8	1.6
LPK10	CCT CTG CTC AGT TCC TTG GTC TCC	1	0.2
LPK11	GGG YTS CTG TTG YTC TG	8	1.6
LPK12	TGC TGC TGT GGC TTA CAG	1	0.2
LPK13a	CTC CCT CYC AGC KTC TGM	8	1.6
LPK13b	CCT TYT CAA CTT CTG CTS	4	0.8
LPK14	GGA CAT GAG GRM YYC TSC TC	32	6.4
LPK15a	CTS AGY TYC TKG GGC TSC TGC	32	6.4
LPK15b	GTG TCC TTC CTG ACC TCC TAG CTC	1	0.2
LPK16	CCA GGT TCA GGT TCT GGG GCT CC	1	0.2
LPK17	CTC CTC AGY YTT CTT CTC C	4	0.8
LPK18	GCC TCC TTT TAC TGT GGA CCA CGG	1	0.2
LPK19	CAG TTC CTG GGG CTC	1	0.2
LPL1	GGC CTG GAC TCC TCT CTT C	1	0.2
LPL2	CTC TCT CTC CTG GCT CTC WGC	2	0.4
		243	

These primers target the 5' extremity and are used for the amplification of light chain sequences in PCR1.

The sequences are given using the IUPAC nomenclature of mixed bases (R = A or G; Y = C or T; M = A or C; K = G or T; S = C or G; W = A or T; H = A or C or T; B = C or G or T; V = A or C or G; D = A or G or T). The orientation of the primers is written from 5' to 3' end. Column d lists the d-fold degeneration encoded in each primer. The volumes of each primer used to set up the PCR mix are listed in column v (μL).

Table A.2. Primers targeting the L-PART1 regions of immunoglobulin heavy chain γ sequences.

Primer	Sequence	d	v (μL)
LPH1a	TGG TAK CAR CAG CWA CAG	8	1.6
LPH1b	TGK CAG CAA SAG CTA CAG	4	0.8
LPH1c	TSS TGT CAG GAA CTG CAS	8	1.6
LPH1d	AWG GGA KGG AGC TGT RTC	8	1.6
LPH1e	ATG GAA TGG AGC TGG RTC	2	0.4
LPH1f	TCC TST CAD TAA CTG CAK	12	2.4
LPH2	GCC TRG TRA CAT TCC CAA GCT	4	0.8
LPH3	GTR SCT GTT SWC AGC C	16	3.2
LPH4	ATG GAT TTT GGG CTG	1	0.2
LPH5	GRA CTT CGG GCT CAG CTT G	2	0.4
LPH7a	GTT GTG GBT RAA CTG GR	12	2.4
LPH7b	CTT CAC AGT AAC ACC	1	0.2
LPH8	CTC ATT CCT GCT GCT G	1	0.2
LPH9	CTG ATG GCA GCT GCC CAA	1	0.2
LPH10	GST GTT GGG GCT KAA GTG GG	4	0.8
LPH11	ATG GAG TGG GAA CTG RGC	2	0.2
LPH12	CTC CTG TGC TTG GCA GCA GC	1	0.2
LPH14a	TTC CTG ATG GCA GTG	1	0.2
LPH14b	AAG GGG CTC CCA AGT CCC	1	0.2
		89	

These primers target the 5' extremity and are used for the amplification of heavy chain sequences in PCR1.

The sequences are given using the IUPAC nomenclature of mixed bases (R = A or G; Y = C or T; M = A or C; K = G or T; S = C or G; W = A or T; H = A or C or T; B = C or G or T; V = A or C or G; D = A or G or T). The orientation of the primers is written from 5' to 3' end. Column d lists the d-fold degeneration encoded in each primer. The volumes of each primer used to set up the PCR mix are listed in column v (μL).

Table A.3. Primers targeting FR1 regions of immunoglobulin light chain κ and λ sequences.

Primer	Sequence	d	v(µl)
OVK1	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GATGYTKTKVTGACCCAAACTCC**	24	4.8
OVK2a	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GATATTGTGATRACBCAGGYTGA**	12	2.4
OVK2b	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GATATTGTGATRACBCAGGYTGC**	12	2.4
OVK3	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**RACATTGTGCTGACMCAATCTCC**	4	0.8
OVK4a	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**SAAAWTGTKCTCWCCCAGTCTCC**	16	3.2
OVK4b	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**SAAAWTCTKCTCWCCCAGTCTCC**	16	3.2
OVK4c	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**SAAAWTTTKCTCWCCCAGTCTCC**	16	3.2
OVK5a	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GAYATYCTGATRACYCAGTCTCC**	16	3.2
OVK5b	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GAYATYGTGATRACYCAGTCTCC**	16	3.2
OVK5c	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GATATYTTGATRACYCAGTCTCC**	16	3.2
OVK5d	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GAYATYCTGCTRACYCAGTCTCC**	16	3.2
OVK5e	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GAYATYGTGCTRACYCAGTCTCC**	16	3.2
OVK5f	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GAYATYTTGCTRACYCAGTCTCC**	16	3.2
OVK6a	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**ARCATTGTGATGACCCAGWCTCA**	4	0.8
OVK6b	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**ARCATTGTGATGACCCAGWCTCC**	4	0.8
OVK6c	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GRCATTGTGATGACCCAGWCTCA**	4	0.8
OVK6d	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GRCATTGTGATGACCCAGWCTCC**	4	0.8
OVK7	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GACATTGTGATGACTCAGTCTCC**	1	0.2
OVK8b	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**RAMATTATGWTGWCACAGTCTAT**	16	3.2
OVK8d	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**RAMATTGTGWTGWCACAGTCTAT**	16	3.2
OVK8f	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**RAMATTTTGWTGWCACAGTCTAT**	16	3.2
OVK8g	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**RAMATTATGWTGWCACAGTCTCC**	16	3.2
OVK8h	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**RAMATTATGWTGWCACAGTCTCT**	16	3.2
OVK8i	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**RAMATTGTGWTGWCACAGTCTCC**	16	3.2
OVK8j	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**RAMATTGTGWTGWCACAGTCTCT**	16	3.2
OVK8k	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**RAMATTTTGWTGWCACAGTCTCC**	16	3.2
OVK8l	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**RAMATTTTGWTGWCACAGTCTCT**	16	3.2
OVK9	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GACATCCAGATGAYYCAGTCTCC**	4	0.8
OVK10	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GATATCCAGATGACACAGACTAC**	1	0.2
OVK11-16	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GATRTYCARATRAYYCAGTCTCC**	12	2.4
OVK12-13	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GACATCCAGATGACWCARTCTYC**	8	1.6
OVK14	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GAMATCMWGATGACCCARTCTCC**	16	3.2
OVK15,19	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GAYATCCAGATGAMMCAGTCTCC**	6	1.2
OVK17	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GAAAYAACTGTGACCCAGTCTCC**	2	0.4
OVK18	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**ACTGGAGAAACAACACAGGCTCC**	1	0.2
OVL	TTATTACTCGC*GGCCCAGCCGGCC*ATGGCGCTGGCC**GATGCTGTTGTGACTCAGGAATC**	1	0.2
		460	

These primers target the 5' extremity and are used for the amplification of light chain sequences PCR2.

The sequences are given using the IUPAC nomenclature of mixed bases (R = A or G; Y = C or T; M = A or C; K = G or T; S = C or G; W = A or T; H = A or C or T; B = C or G or T; V = A or C or G; D = A or G or T). The orientation of the primers is written from 5' to 3' end. Column d lists the d-fold degeneration encoded in each primer. The volumes of each primer used to set up the PCR mix are listed in column v (µL).

The nucleic acids listed in *italic* are the SfiI restriction sites and those in bold are homologous to part of the FR1 sequences. Legend is similar to table 2.

Table A.4. Primers targeting FR1 regions of immunoglobulin heavy chain γ sequences.

Primers	Sequence	d	v (µl)
OVH1	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**SAGGTCCARCTGCAGCAGYYTGG**	16	3.2
OVH2	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**CAGGTRCAGCTGAAGSAGTCAGG**	4	0.8
OVH3	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**GAKGTGCAGCTTCAGGAGTCRGG**	4	0.8
OVH4	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**GAGGTGAAGCTTCTCGAGTCTGG**	1	0.2
OVH5	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**GAVGTGAWGCTGGTGGAGTCTGR**	12	2.4
OVH6a	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**GATGTGAACTTGGAAGTGTCTGG**	1	0.2
OVH7	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**GAGGTGAAGCTGRTGGARTCTGR**	8	1.6
OVH8	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**CAGGTTACTCTGAAAGAGTCTGG**	1	0.2
OVH9	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**CAGATCCAGTTGGYGCAGTCTGG**	2	0.4
OVH10	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**GAGGTGCAGCTTGTTGAGWCTGG**	2	0.4
OVH11	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**GAAGTGCAGCTGTTGGAGACTGG**	1	0.2
OVH12	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**CAGATGCAGCTTCAGGAGTCAGG**	1	0.2
OVH13	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**GAAGTGAAGCTTGAGGAGTCTGG**	1	0.2
OVH14	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**GAGGTTCAGCTGCAGCAGTCTGG**	1	0.2
OVH15	GGCGGCGGCGGCTCCGGTGGTGGTGGATCC**CAGGTTCACCTACAACAGTCTGG**	1	0.2
		56	

These primers target the 5' extremity and are used for the amplification of heavy chain sequences PCR2.

The sequences are given using the IUPAC nomenclature of mixed bases (R = A or G; Y = C or T; M = A or C; K = G or T; S = C or G; W = A or T; H = A or C or T; B = C or G or T; V = A or C or G; D = A or G or T). The orientation of the primers is written from 5' to 3' end. Column d lists the d-fold degeneration encoded in each primer. The volumes of each primer used to set up the PCR mix are listed in column v (µL).

The nucleic acids listed in bold are homologous to part of the FR1 sequences and those underlined correspond to the G_4S_4 linker sequence. Legend is similar to table 2.

Table A.5. Primers targeting FR4 regions of immunoglobulin light chain κ and λ sequences.

Primer	Sequence	d	v(µl)
VL3'a	GGAGCCGCCGCCGCCAGAACCACCACCACCAGAACCACCACCACCACGTTTGATTTCCAGCTTGG	1	1
VL3'b	GGAGCCGCCGCCGCCAGAACCACCACCACCAGAACCACCACCACCACGTTTTATTTCCAGCTTGG	1	1
VL3'c	GGAGCCGCCGCCGCCAGAACCACCACCACCAGAACCACCACCACCACGTTTTATTTCCAACTTTG	1	1
VL3'd	GGAGCCGCCGCCGCCAGAACCACCACCACCAGAACCACCACCACCACGTTTCAGCTCCAGCTTGG	1	1
VL3'e	GGAGCCGCCGCCGCCAGAACCACCACCACCAGAACCACCACCACCACCTAGGACAGTCAGTTTGG	1	0.25
		5	

These primers target the 3' extremity and are used for the amplification of light chain sequences PCR2.

The nucleic acids that are highlighted are homologous to part of the FR4 sequences, and those underlined correspond to the G_4S_4 linker sequence. The orientation of the primers is written from 5' to 3' end.

Table A.6. Primers targeting FR4 regions of immunoglobulin heavy chain γ sequences.

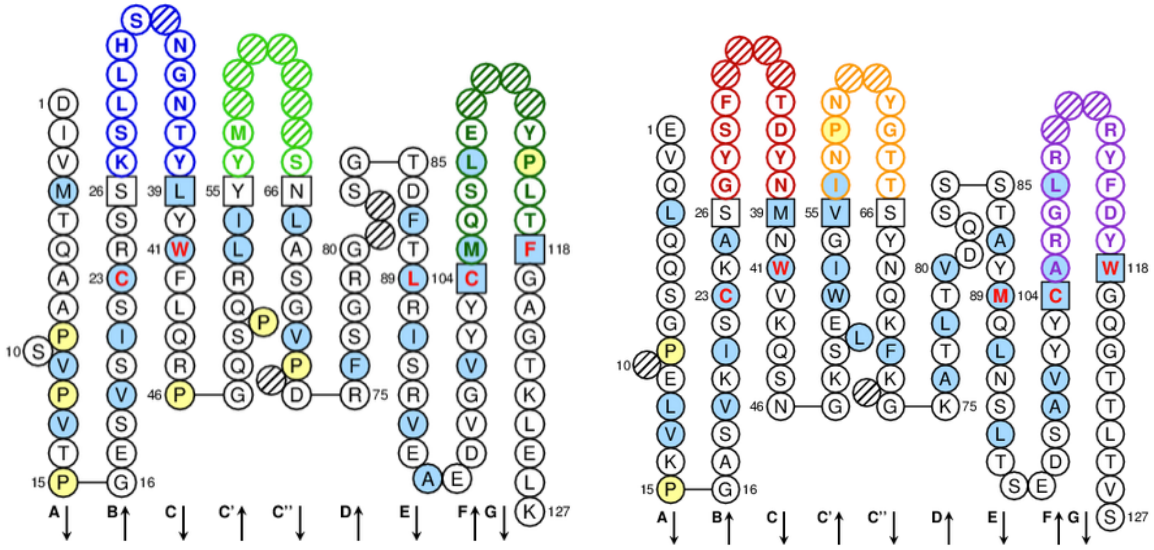
Name	Sequence	d	v(ul)
VH3'a	GGAATTC*GGCCCCCGAGGCC*GAGGAAACGGTGACCGTGGT	1	1
VH3'b	GGAATTC*GGCCCCCGAGGCC*GAGGAGACTGTGAGAGTGGT	1	1
VH3'c	GGAATT*CGGCCCCCGAGGCC*GCAGAGACAGTGACCAGAGT	1	1
VH3'd	GGAATT*CGGCCCCCGAGGCC*GAGGAGACGGTGACTGAGGT	1	1
		4	

These primers target the 3' extremity and are used for the amplification of heavy chain sequences PCR2.

The nucleic acids that are highlighted are homologous to part of the FR4 sequences, and those in *italic* are the SfiI restriction sites. The orientation of the primers is written from 5' to 3' end.

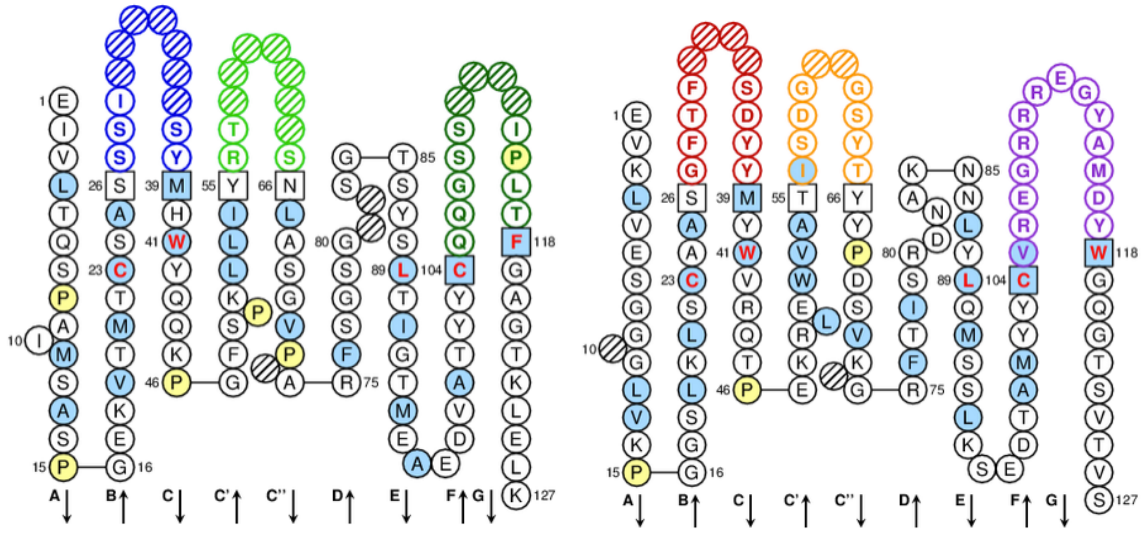
IV. Pearl diagrams of 9G4H9 and selected antibody fragments.

IV.1. 9G4H9

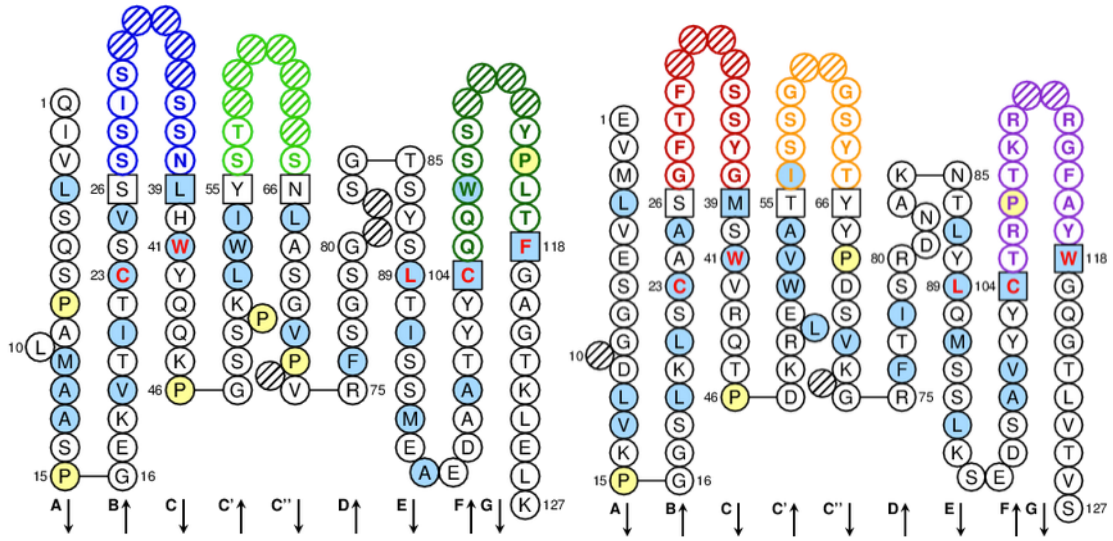


N.B. The presented numbering in this diagram is according to the criteria set by IMGT, which is different from the numbering scheme used in previous publications (Phichith *et al.*, 2009).

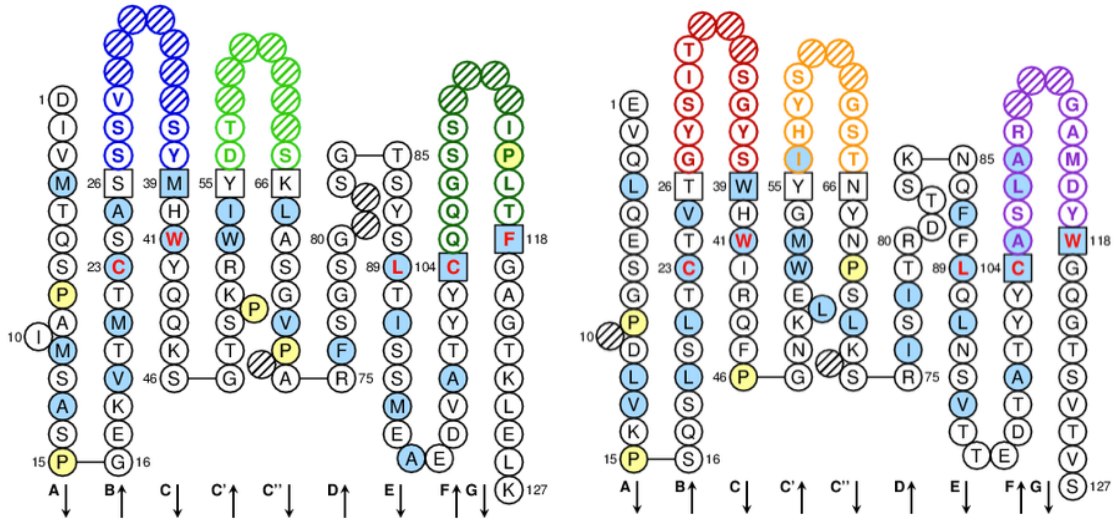
IV.2. P90C1



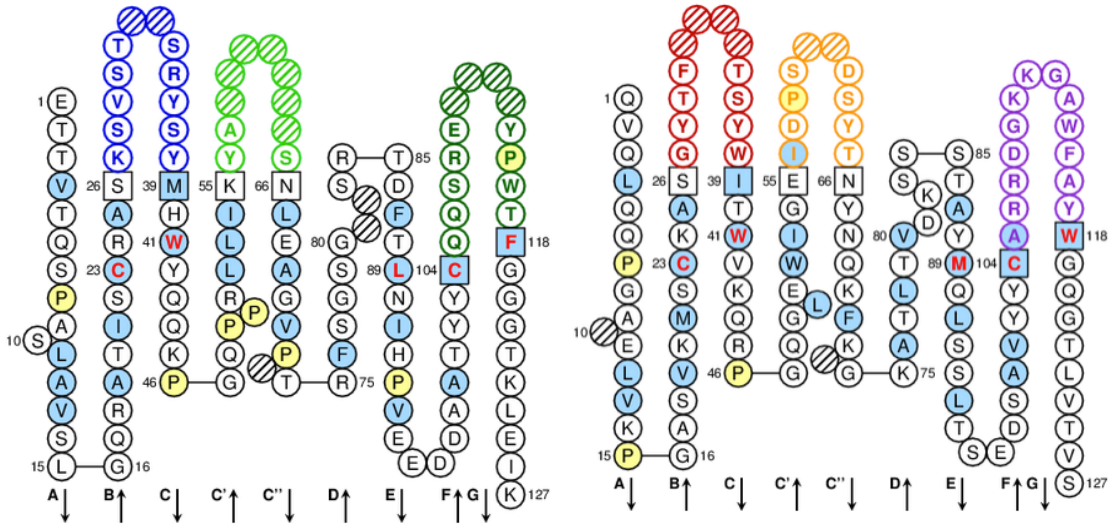
IV.3. P90C2



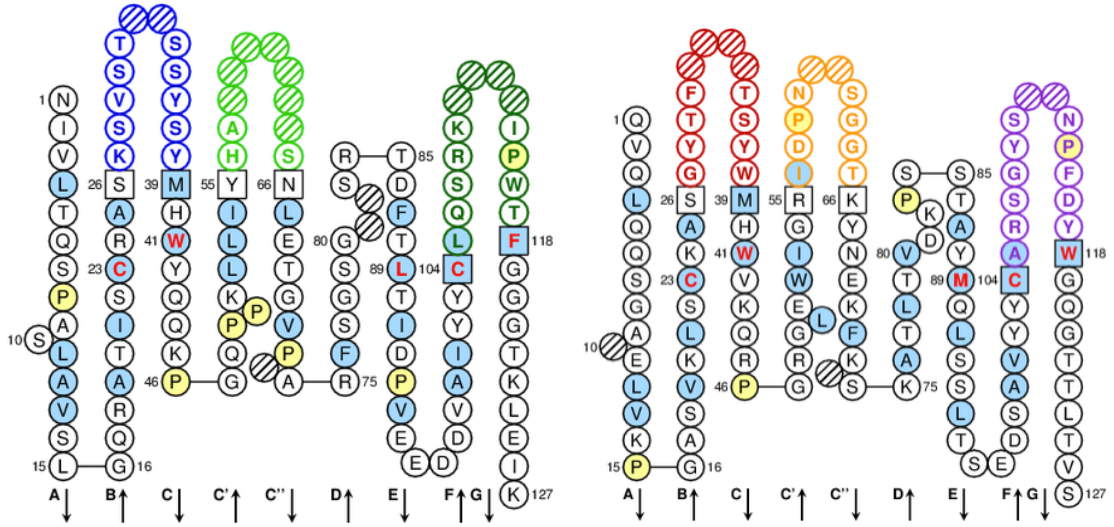
IV.4. P90C3



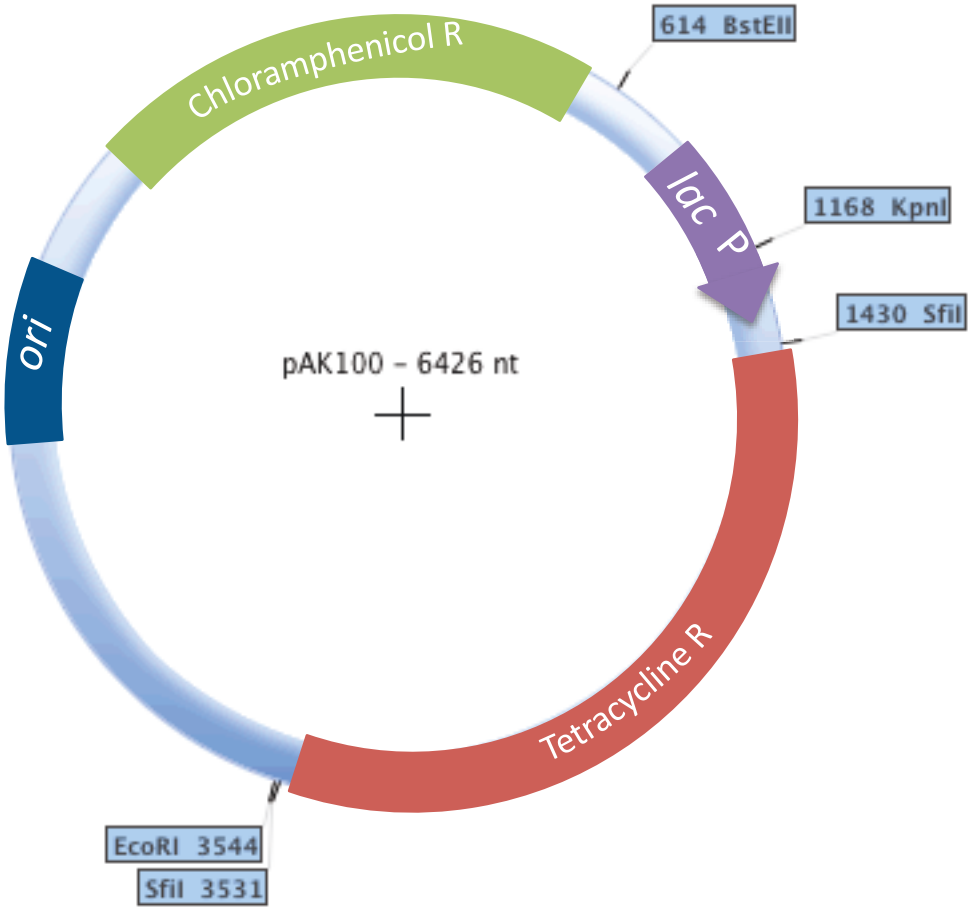
IV.5. PSC1



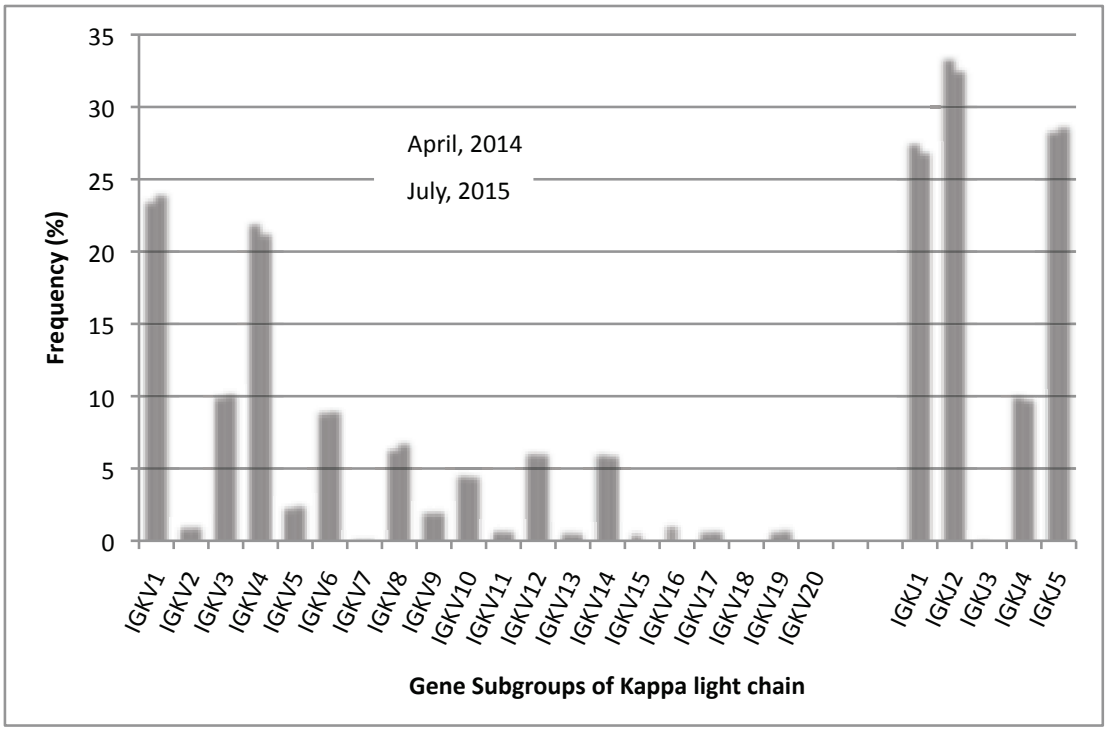
IV.6. PSC2



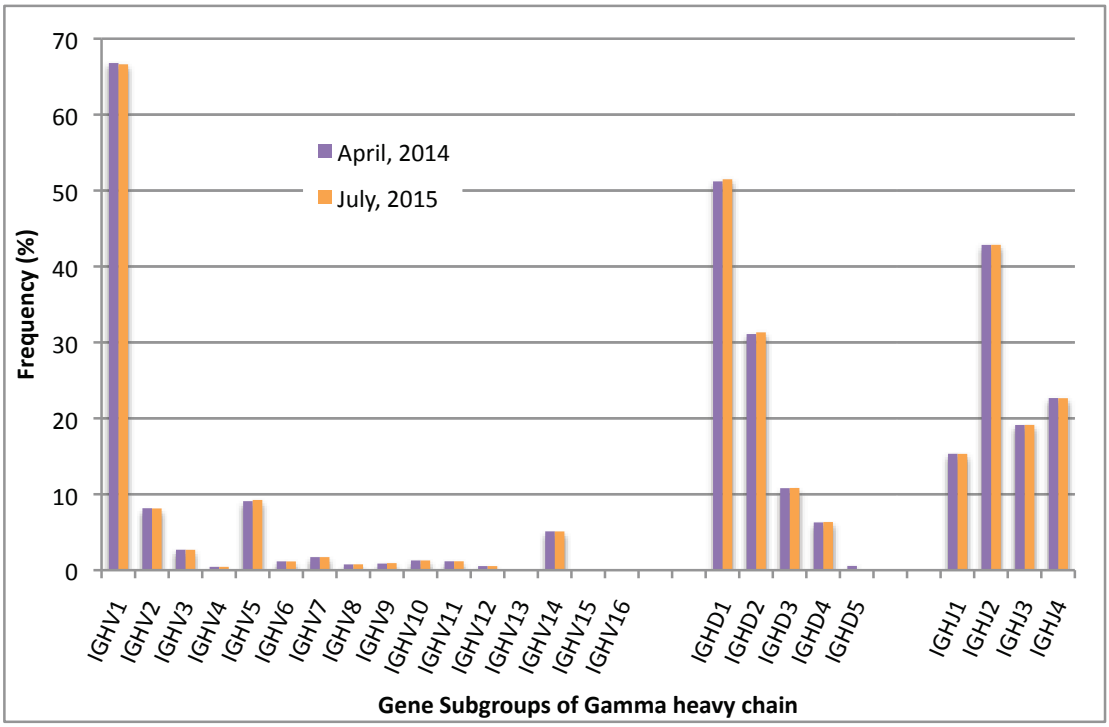
V. Graphic map of phagemid vector pAK100



VI. Evolution of IMGT data between 2014 and 2015.



The number of sequences available in IMGT in 2014 were the following: IGKV: n=1571 and IGKJ: n=1397. The number of sequences in 2015 were IGKV: n=1587 and IGKJ: n=1344.



The number of sequences available in IMGT in 2014 were the following: IGHV: n=8144, IGHD: n=6725, and IGHJ: n=8207. The number of sequences in 2015 were IGHV: n=8167, IGHD: n=6691, and IGHJ: n=8212.