

**Newcastle**  
University

**QSAR model development for  
early stage screening of  
monoclonal antibody therapeutics  
to facilitate rapid developability.**

**Arathi Kizhedath**

**(B.Tech, MSc.ir)**

**Doctor of Philosophy**

**Faculty of Science, Agriculture and Engineering**

**School of Engineering**

**February 2019**



---

# QSAR model development for early stage screening of monoclonal antibody therapeutics to facilitate rapid developability.

---

**Arathi Kizhedath**



**Supervisor: Prof. Jarka Glassey**

**Co-supervisor: Dr Simon Wilkinson**

**School of Engineering, Newcastle University  
Institute of Cellular Medicine, Newcastle University**

**October 2018**





## Abstract

Monoclonal antibodies (mAbs) and related therapeutics are highly desirable from a biopharmaceutical perspective as they are highly target specific and well tolerated within the human system. Nevertheless, several mAbs have been discontinued or withdrawn based either on their inability to demonstrate efficacy and/or due to adverse effects. With nearly 80% of drugs failing in clinical development mainly due to lack of efficacy and safety there arises an urgent need for better understanding of biological activity, affinity, pharmacology, toxicity, immunogenicity etc. thus leading to early prediction of success/failure. In this study a hybrid modelling framework was developed that enabled early stage screening of mAbs. The applicability of the experimental methods was first tested on chemical compounds to assess the assay quality following which they were used to assess potential off target adverse effects of mAbs. Furthermore, hypersensitivity reactions were assessed using Skimune™, a non-artificial human skin explants based assay for safety and efficacy assessment of novel compounds and drugs, developed by Alcyomics Ltd. The suitability of Skimune™ for assessing the immune related adverse effects of aggregated mAbs was studied where aggregation was induced using a heat stress protocol. The aggregates were characterised by protein analysis techniques such as analytical ultra-centrifugation following which the immunogenicity tested using Skimune™ assay. Numerical features (descriptors) of mAbs were identified and generated using ProtDCal, EMBOSS Pepstat software as well as amino acid scales for different. Five independent and novel X block datasets consisting of these descriptors were generated based on the physicochemical, electronic, thermodynamic, electronic and topological properties of amino acids: Domain, Window, Substructure, Single Amino Acid, and Running Sum. This study describes the development of a hybrid QSAR based model with a structured workflow and clear evaluation metrics, with several optimisation steps, that could be beneficial for broader and more generic PLS modelling. Based on the results and observation from this study, it was demonstrated incremental improvement via selection of datasets and variables help in further optimisation of these hybrid models. Furthermore, using hypersensitivity and cross reactivity as responses and physicochemical characteristics of mAbs as descriptors, the QSAR models generated for different applicability domains allow for rapid early stage screening and developability. These models were validated with external test set comprising of proprietary compounds from industrial partners, thus paving way for enhanced developability that tackles manufacturing failures as well as attrition rates.



## List of Publications

Kizhedath, Arathi, Simon Wilkinson, and Jarka Glassey. "Applicability of predictive toxicology methods for monoclonal antibody therapeutics: status Quo and scope." *Archives of toxicology* 91.4 (2017): 1595-1612.

Kizhedath, Arathi, Simon Wilkinson, and Jarka Glassey. "Applicability of Traditional In Vitro Toxicity Tests for Assessing Adverse Effects of Monoclonal Antibodies: A Case Study of Rituximab and Trastuzumab." *Antibodies* 7.3 (2018): 30.

Kizhedath, Arathi, and Jarka Glassey. "The importance of critical quality attributes in Quality by Design for rapid bioprocess development strategies." *European Pharmaceutical Review* 22.6 (2017): 48-50.

Udugama, Isuru A., Hannah Feldman, Simoneta Caño de las Heras, Arathi Kizhedath, Jesper Bryde-Jacobsen, Frans van den Berg, Seyed Soheil Mansouri, and Krist V. Gernaey. "BIOPRO World Talent Campus: A week of real world challenge for biotechnology post-graduate students." *Education for Chemical Engineers* 25 (2018): 1-8.

Kizhedath, Arathi, Simon Wilkinson, and Jarka Glassey. "Assessment of hepatotoxicity and dermal toxicity of butyl paraben and methyl paraben using HepG2 and HDFn in vitro models." *Toxicology in Vitro* 55 (2018) 108-115.

Kizhedath, Arathi, Micael Karlberg, and Jarka Glassey "Cross interaction chromatography based QSAR model for early stage screening to facilitate enhanced developability of monoclonal antibody therapeutics." *Biotechnology Journal* (2018) *Accepted*.

## List of major presentations

Kizhedath, Arathi, Simon Wilkinson, and Jarka Glassey. "Hybrid modelling approaches for early stage screening of monoclonal antibodies based on safety for rapid mAb developability" ESBES 2018, Lisbon, Portugal.

Kizhedath, Arathi, Simon Wilkinson, and Jarka Glassey. "Assessment of hepatotoxic and dermal toxicity effects of Parabens using HepG2 and HDFN in vitro models." British Toxicological Society Annual Congress 2018, Newcastle Upon Tyne, United Kingdom.

Kizhedath, Arathi, Simon Wilkinson, and Jarka Glassey. "Early toxicity prediction based on hybrid modelling." BESIG Young Researchers Meeting, IChemE, 2016, Dublin, Ireland.



## Acknowledgements

The past three years have been a true learning experience for me, professionally and personally. This project was truly interesting and challenging. I also had the opportunity to broaden my horizons and get exposed to new perspectives in pharmaceutical safety testing strategies via conferences, training and outreach activities.

Firstly, I would like to thank my extremely supportive supervisor, Prof Jarka Glassy without whom this dissertation would not be possible. I would also like to extend my gratitude to Dr Simon Wilkinson who co-supervised my work. Their patient guidance, enthusiastic encouragement and useful critiques helped shape up this dissertation. I would also like to acknowledge my colleague, Micael Karlbeg for the collaborative work on descriptor generation. I would also like to thank him for his professional and personal support throughout my PhD.

In Chapter 4, Skimune™ assay was used to assess immunogenicity of aggregated mAbs and this was done by Alcyomics Ltd. I would like to thank Prof Anne Dickinson, Dr. Shaheda Ahmed and Ana Ribeiro for their support. In Chapter 7, MAb Fab structures were generated by João Victor de Souza Cunha and Dr Agnieszka Bronowska, School of Natural and Environmental Sciences, Newcastle University. I would like to thank them for them as well as Dr Milan Mijajlovic for their time and valuable academic discussions. I would like to thank all the Early stage researchers and Scientist in charge of the BIORAPID consortium. This project is a part of the BIORAPID consortium which has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie actions grant agreement No. 643056. I gratefully acknowledge this funding provided by the European Commission. I would like to thank Mr. Paul Sterling for his assistance in the Biolab. I would also like to thank Dr Achim Treumann and Dr Alexandra Solovyova from the Newcastle University Protein and Proteome Analysis (NUPPA) unit for their services on the Analytical Ultracentrifugation used in Chapter 4. I would also like to acknowledge Dr Andrew Filby from the Newcastle University Flow Cytometry Core Facility for the Cell Cycle analysis services used in Chapter 2.

Last but not the least, I extend my heartfelt gratitude to my parents, brother, husband and friends whose continued support kept me motivated throughout my research period.



## Table of contents

Chapter 1. Applicability of predictive toxicology methods for monoclonal antibody therapeutics: Status Quo and Scope	1
1.1. Aims and objectives	3
1.2. Mabs: Safety pharmacology and side effects	4
1.3. Quality by Design in mAb developability	14
1.4. In vitro systems for toxicity testing	16
1.5. In silico tools for predictive toxicology	18
1.6. Predictive Model Development	21
1.5.1. Databases	22
1.5.2. Descriptor Generation and model development:	23
1.5.3. Models	23
1.5.4. Validation	25
1.7. Scope of the study	25
Chapter 2. Assessment of hepatotoxicity and dermal toxicity of butyl paraben and methyl paraben using HepG2 and HDFn in vitro models	26
2.1. Materials and Methods	28
2.1.1. Cell lines and reagents	28
2.1.2. Cell culture and maintenance	28
2.1.3. Cell seeding and treatment	28
2.1.4. WST-1 cell proliferation Assay	29
2.1.5. CellTiter-Glo® Luminescent Cell Viability Assay	29
2.1.6. GSH-Glo™ Glutathione Assay	29
2.1.7. Detection of Reactive oxygen species	30
2.1.8. Cell cycle Analysis	30
2.1.9. Mitochondrial Membrane Potential Assay	30
2.1.10. Statistical analysis	31
2.2. Results	31
2.2.1. Paraben induced cytotoxicity	31
2.2.2. Paraben induced reduction in ATP levels	31
2.2.3. Paraben induced reduction in glutathione levels	32
2.2.4. Paraben induced time dependent decrease in ATP levels	35
2.2.5. Cell cycle analysis to assess genotoxicity of parabens	36
2.2.6. Paraben induced Oxidative stress and mitochondrial dysfunction	38
2.3. Discussion	41
2.4. Chapter summary	44
Chapter 3 Applicability of traditional and novel in vitro toxicity tests for assessing adverse effects of monoclonal antibodies: A case study of rituximab and trastuzumab.	45
3.1. Materials and Methods	48
3.1.1. Materials and reagents	48
3.1.2. Cell culture and maintenance	48
3.1.3. Cell seeding and exposure to mAbs	49

3.1.4. Complement dependent cytotoxicity (CDC)	49
3.1.5. Antibody dependent cellular cytotoxicity (ADCC)	49
3.1.6. Complement dependent cellular cytotoxicity (CDCC)	49
3.1.7. WST-1 cell proliferation Assay	49
3.1.8. CellTiter-Glo® Luminescent Cell Viability Assay	50
3.1.9. Skimune analysis	50
3.1.10. Statistical analysis	50
3.2. Results	50
3.2.1. MAb induced effect on in cell viability	50
3.2.2. MAb induced effect in ATP levels	51
3.2.3. Effect of donor variability and intrinsic variation	54
3.2.4. Immunogenicity of mAbs assessed by Skimune™ assay	55
3.3. Discussion	55
3.4. Chapter summary	58
Chapter 4: Aggregation of mAbs at physiological pH and temperature and detection of immunogenicity using Skimune™ assay	59
4.1. Materials and Methods	60
4.1.1. Thermal stress of antibodies	60
4.1.2. AUC analysis	60
4.1.3. Immunogenicity of aggregated mAbs	61
4.1.4. Analysis of protein sequence and aggregation propensities	61
4.2. Results	62
4.2.1. AUC analysis of mAb aggregation	62
4.2.2. Immunogenicity of mAb aggregates	65
4.2.3. Prediction of aggregation propensities based on primary sequence.	67
4.2.4. Accessibility of aggregation prone regions in globular structures	71
4.3. Discussion	73
4.4. Chapter summary	75
Chapter 5: QSAR development using novel descriptors sets for mAbs: Influence of intrinsic properties.	76
5.1. Materials and methods	79
5.1.1. Data collection	79
5.1.2. Multiple sequence alignment	79
5.1.3. Descriptor generation	79
5.1.4. Dataset generation	80
5.1.5. Data pre-treatment and variable reduction	82
5.1.6. Principal component analysis (PCA)	83
5.2. Results	83
5.2.1. Multiple sequence alignment (MSA)	83
5.2.2. Domain specific influence of intrinsic properties	87
5.2.3. Influence of intrinsic properties in hypervariable regions	89
5.3. Discussion	94
5.4. Chapter summary	96

Chapter 6: Cross interaction chromatography based QSAR model for early stage screening to facilitate enhanced developability of monoclonal antibody therapeutics.	97
6.1. Materials and Methods	100
6.1.1. Data collection	100
6.1.2. Descriptor generation	100
6.1.3. Data curation and Variable reduction	100
6.1.4. Exploratory Analysis	100
6.1.5. Cross Interaction chromatography data	101
6.1.6. Model development	101
6.1.7. Model Performance metrics	102
6.1.8. Y randomization	102
6.2. Results	103
6.2.1. Exploratory analysis	103
6.2.2. Variable Reduction and variable Selection	107
6.2.3. Model performance evaluation	107
6.2.4. Benchmarking of model performance metrics	117
6.2.5. Y randomisation	120
6.3. Discussion	121
6.4. Chapter summary	124
Chapter 7: QSAR model development using 3D descriptors of mAbs generated via homology modelling and molecular dynamics simulation.	125
7.1. Materials and Methods	126
7.1.1. Data collection	126
7.1.2. Generation of mAb structures	127
7.1.3. Descriptor generation	129
7.1.4. Data curation and variable reduction	130
7.1.5. Exploratory Analysis	131
7.1.6. Cross Interaction chromatography data	131
7.1.7. Model development	131
7.1.8. Model Performance metrics	132
7.1.9. Y randomization	132
7.2. Results	132
7.2.1. Exploratory Analysis	132
7.2.2. Variable reduction and selection	134
7.2.3. Model performance	135
7.2.4. Benchmarking of datasets and models.	139
7.2.5. Y randomization	140
7.3. Discussion	141
7.4. Chapter summary	143
Chapter 8: QSAR model development based on a combination of primary and 3D descriptors	144
8.1. Materials and Methods	144
8.1.1. Data distribution	144
8.1.2. Data augmentation methodology	145

8.1.3. Combination Method I	145
8.1.4. Combination method II	146
8.2. Results	146
8.2.1. Combination Method I - Model performance	146
8.2.2. Combination Method II - Model Performance	148
8.3. Discussion	150
8.4. Chapter summary	151
Chapter 9: A human skin explant (Skimune™) based QSAR model for early adverse effect prediction of monoclonal antibody therapeutics.	152
9.1. Materials and methods	154
9.1.1. Data collection and descriptor generation	154
9.1.2. Data curation and Variable reduction	155
9.1.3. Biological response data	155
9.1.4. Model development	155
9.1.5. Model Performance metrics	156
9.1.6. Y randomization	156
9.2. Results	156
9.2.1. Variable reduction and variable selection	156
9.2.2. Model performance evaluation	157
9.2.3. Benchmarking of model performance metrics	158
9.2.4. Model Robustness	160
9.2.5. Correlation with Clinical incidence rates	160
9.3. Discussion	162
9.4. Chapter summary	163
Chapter 10: Conclusions and recommendation	165
Appendix A	170
Appendix B	172
Appendix C	176
Appendix D	183
References	190

## List of Figures

- Figure 1.1** Generic monoclonal antibody derived therapeutic structures as adapted from IMGT (World Health, 2006; Lefranc *et al.*, 2009). Fc: Constant region which contributes to effector function, immune response and increased half-life, Fv: Variable region that contains Complementarity Determining Regions (CDRs) facilitating antigen binding, Fab: Antigen Binding Fragment which lack Fc Region, scFv: Single chain Fragment variable, FP: Fc Fusion Proteins that contain Fc region for effector functionality (e.g. Abatacept), CP: Composite protein that contain Fc region for increasing half-life and not for effector functionality (e.g. Strensiq™)(World Health, 2006). 3
- Figure 1.2 (a)** Monoclonal antibody structure with binding site for antigen, FcγR and FcRn receptor as well as glycosylation sites (Glycan); Ag: Antigen; CDC: Complement Dependent Cytotoxicity; ADCC: Antibody Dependent Cell Cytotoxicity; ADCP: Antibody Dependent Cell Phagocytosis **(b)** Glycosylation profile at N297 residue of the Fc region of antibodies. The bold line indicates core structures and dotted line indicates variable structures. Gal: Galactose; SA: Sialic Acid; Man: Mannose; GlcNAc: N-Acetylglucosamine Fuc: Fucose; Asn: Asparagine (N297). 13
- Figure 1.3** A schematic for Quality by design in biopharmaceutical industry; RA: Risk assessment, PC: Process characterisation (Rathore and Winkle, 2009). 15
- Figure 1.4 a)** Computational toxicology model development workflow **b)** Techniques involved in different types of predictive models. 22
- Figure 2.1** Chemical structures of methyl paraben(CSID:7176) butyl paraben(CSID:6916) and 4-Hydroxybenzoic acid (CSID:132) (Source ChemSpider, Royal Society of Chemistry). 26
- Figure 2.2** Results of the different in vitro tests performed on HepG2 cell after treatment with Butyl paraben and Methyl paraben for 24 h. Cell viability assay using WST-1 reagent (a) Dose response curve (b) concentration vs percentage of the control response. ATP reduction assay using CellTiter-Glo® (c) Dose response curve (d) concentration vs percentage of the control response. GSH reduction using GSH-Glo™ (e) Dose response curve (f) concentration vs percentage of the control response. All values are expressed as a percentage of the control and mean ± standard error of the mean (SEM) for n=3. Statistical differences were assessed with one way ANOVA. Dunnett's post hoc test was used for multiple comparisons to control. \*\* indicates p<0.05, \*\*\* indicates p<0.001 33
- Figure 2.3** Results of the different in vitro tests performed on HDFN cell after treatment with Butyl paraben and Methyl paraben. Cell viability assay using WST-1 reagent (a) Dose response curve (b) concentration vs percentage of control 34

response. ATP reduction assay using CellTiter-Glo® (c) Dose response curve (d) concentration vs percentage of control response. GSH reduction using GSH-Glo™ (e) Dose response curve (f) concentration vs percentage of control response. All values are expressed as percentage of control and mean ± standard error of the mean (SEM) for n=3. Statistical differences were assessed with one way ANOVA. Dunnett's post hoc test was used for multiple comparisons to control. \*\* indicates p<0.05, \*\*\* indicates p<0.001.

**Figure 2.4** Dose response curves of the reduction in ATP levels measured over time after the exposure of cells to different concentrations of parabens for 1, 4, 8, 12 and 24 h using CellTiter-Glo®. Dose response curves for HDFN after treatment with (a) Butyl Paraben (b) Methyl Paraben. Dose response curves for HepG2 after treatment with (c) Butyl Paraben (d) Methyl Paraben All values are expressed as percentage of control and mean ± standard error of the mean (SEM) for n=3. 36

**Figure 2.5** Graphical representation of cell cycle phase proportions following cell fixing, Propidium Iodide staining and analysis by flow cytometry for (a) HepG2 cells treated with butyl paraben (b) HepG2 cells treated with methyl paraben. All values are expressed as a percentage of the control and mean ± standard error of the mean (SEM) for n=3. Statistical differences were assessed with one way ANOVA. Dunnett's post hoc test was used for multiple comparisons to control. \*\* indicates p<0.05. 37

**Figure 2.6** Graphical representation of cell cycle phase proportions following cell fixing, Propidium Iodide staining and analysis by flow cytometry for (a) HDFn cells treated with butyl paraben (b) HDFn cells treated with methyl paraben. 38

**Figure 2.7** Results of the DCFDA assay to measure ROS production in HepG2 cells after treatments with (a) butyl paraben and (b) methyl paraben, All values are expressed as percentage of control and mean ± standard error of the mean (SEM) for n=3. 39

**Figure 2.8** Results of the DHR 123 assay to measure oxidative stress in (a) HepG2 and (b) HDFN cell lines after treatment with Butyl paraben and Methyl paraben. Results of the TMRE assay to measure mitochondrial membrane potential in (c) HepG2 cells after treatments with parabens, All values are expressed as percentage of control and mean ± standard error of the mean (SEM) for n=3. 40

**Figure 3.1** General Methodology for an in vitro assay to detect toxicity of mAb based therapeutics. 46

**Figure 3.2.** Skimune™ ® assay. This autologous model consists in priming of immune cells and co-culture with skin biopsy from the same donor to induce tissue 47

damage. Histopathological assessment of damage is done according to a scoring grade (I-IV).

**Figure 3.3** Grading the reactivity of a test therapeutic. Assessment of the histopathological damage caused by a test therapeutic according to Lerner score grading (I to IV). 48

**Figure 3.4** (a) Control conditions without effector cells/serum (b) CDC (c) CDCC and (d) ADCC assay results of rituximab and Trastuzumab based on the WST assay. Results represent pooled responses from four donors (n=4). All values are expressed as Relative Absorbance Units (RAU) of control (mean±SE). R: Rituximab and H: Trastuzumab. Positive control is 5% (v/v) of absolute ethanol. 51

**Figure 3.5** (a) Control conditions without effector cells/serum, (b) CDC (b) CDCC and (d) ADCC assay results of rituximab and trastuzumab based on the ATP content compared to control in HepG2 cells exposed to mAbs. Results represent pooled responses from four donors (n=4). All values expressed as Relative fluorescence units (RFU) compared to control (mean±SE). R: Rituximab and H: Trastuzumab. Positive control is 5% (v/v) of absolute ethanol. 52

**Figure 3.6** (a) Control conditions without effector cells/serum, (b) CDC (b) CDCC and (d) ADCC assay results of rituximab and Trastuzumab based on the ATP content compared to control in HDFn cells exposed to mAbs. Results represent pooled responses from four donors (n=4). All values expressed as Relative fluorescence units (RFU) compared to control (mean±SE). R: Rituximab and H: Trastuzumab. Positive control is 5% (v/v) of absolute ethanol. 53

**Figure 3.7** Intrinsic variation in responses owing to donor variability in ADCC and CDCC assay (a) results of rituximab and Trastuzumab based on the ATP content compared to control in HepG2 cells exposed to mAbs. Results represent pooled responses from four donors (n=4). All values expressed as Relative fluorescence units (RFU) compared to control (mean±SE). R: Rituximab and H: Trastuzumab. Positive control is 5% (v/v) of absolute ethanol. 54

**Figure 3.8** Skimune™ assay results with the percentage response based on the different histopathological damage grades induced by rituximab and trastuzumab based on 10 and 5 donors respectively. 55

**Figure 4.1** Overview of experimental methodology to assess thermal-stress induced aggregation and associated immunogenicity. 60

**Figure 4.2** Sedimentation coefficient distribution of rituximab (Mabthera) species for different durations and storage times at temperatures of (a) 4°C, (b)37°C, and (c)40°C and (d)65°C (black line). The results fit the Bayesian prediction model (red dotted line) based on optimally low RMSD. Monomer species are seen around 6S with elongated dimers appearing at 8S and globular dimers appearing at 10s. 64

<b>Figure 4.3</b> Sedimentation coefficient distribution of trastuzumab (Herceptin) species for different duration at temperatures of 4°C, 37°C, 40°C, 65 (black line). The results fit the Bayesian prediction model (red dotted line) based on optimally low RMSD. Monomer species are seen around 6S with elongated dimers appearing at 8S and globular dimers appearing at 10s.	65
<b>Figure 4.4</b> Percentage Skimune® responses for heat-stressed samples of Herceptin and Rituximab (1 and 10µg/mL) based on responses from 5 donors. The red dotted line indicates the 40 % threshold. Samples have been given a 3 index code name for easier naming purposes, with the first index representing the mAb tested (Rituximab, R or Herceptin, H), the second index representing the temperature conditions (4, 37 or 40°C) and the third index representing the duration of exposure (0, 3, 6, 12, 24 or 48 hours). OKT3 : Muromonab-CD3. Med+Skin: Only medium and skin , Med+Skin+PBMCs: Medium, skin and PBMCs, ctrl IgG1: Isotype IgG1 control, ctrl pH3: Control of mAb treated at pH 3.	66
<b>Figure 4.5</b> Multiple Sequence alignment of (a) heavy chain of rituximab, herceptin and human Igg1 and (b) light chain of rituximab and Herceptin performed by T-coffee web server (Notredame et al., 2000).	68
<b>Figure 4.6</b> Zyggs residue scores reflective of potential intrinsic aggregation propensities of dissimilar amino acid residues of rituximab and Herceptin (a) heavy chain and (b) light chain.	70
<b>Figure 4.7</b> Normalised overall aggregation propensity descriptors for rituximab and Herceptin (a) heavy chain and (b) light chain.	71
<b>Figure 4.8</b> Relative surface accessibility of exposed dissimilar amino acid residues with a Zyggs score>1 in rituximab and Herceptin (a) heavy and (b) light chain.	72
<b>Figure 4.9</b> 3D structures rituximab and Herceptin VH and VL regions generated by SWISS-Pdb viewer (Guex and Peitsch, 1997). The blue are shows the accessible residues (>30%). The amino acids coloured in yellow are those accessible residues which have a high intrinsic aggregation propensity.	73
<b>Figure 5.1</b> Overview of methodology for generating descriptors from primary sequence of antibodies.	78
<b>Figure 5.2</b> Distribution of 285 mAbs based on intrinsic properties (a) heavy chain isotype, (b) light chain isotype and (c) species.	78
<b>Figure 5.3</b> Percentage identity matrix of complete heavy chain sequence and respective domains of 285 mAbs for (a) heavy chain (HC), (b) VH, (c) CH1, (d)	85

CH2, (e) Hinge, (f) CH3. Colour gradient indicates percentage similarity between sequences.

**Figure 5.4** Percentage identity matrix of complete light chain sequence and respective domains of 285 mAbs for (a) light chain (LC), (b) VL and (c) CL, Color gradient indicates percentage similarity between sequences. 86

**Figure 5.5** (a) Scores plots and (b) loadings plot following PCA analysis of Domain datablock generated from light chain of mAbs. VL: variable region of the Light Chain; CL: constant region of the Light Chain. 87

**Figure 5.6** (a) Scores plots and (b) loadings plot following PCA analysis of Domain datablock generated from heavy chain of mAbs. VH: variable region of the heavy chain; CH1: first constant region of Heavy chain; Hinge: hinge region of Heavy Chain; CH2: second constant region of Heavy Chain and CH3: third constant region of Heavy Chain. 88

**Figure 5.7** Scores plot generated following PCA analysis of Domain datablock descriptors generated from (a) light chain and (b) heavy chain of mAbs. 88

**Figure 5.8** Scores plots of (A) Domain, (C) Window, (E) Substructure, (G) Single Amino and (I) Running Sum datablocks and loadings plot of (B) Domain, (D) Window, (F) Substructure, (H) Single Amino and (J) Running Sum generated after performing PCA on data blocks containing only hypervariable region sequences of 285 mAbs. X axis represents Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captured by the respective PC. 91

**Figure 5.9** Scores plots of (A) Domain, (C) Window, (E) Substructure, (G) Single Amino and (I) Running Sum datablocks and loadings plot of (B) Domain, (D) Window, (F) Substructure, (H) Single Amino and (J) Running Sum generated after performing PCA on data blocks containing only hypervariable region sequences of 253 mAbs. Colored ellipses indicate the 95% confidence limits of the corresponding class. X axis represents Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captured by the respective PC. 92

**Figure 5.10** Scores plots of (a) Domain, (b) Window, (c) Substructure, (d) Single Amino and (e) Running Sum generated after performing PCA on data blocks containing only hypervariable region sequences of 253 mAbs. Coloured ellipses indicate the 95% confidence limits of the corresponding class. X axis represents Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captured by the respective PC. 93

<b>Figure 6.1</b> Hybrid model development workflow outlining the different steps involved in descriptor generation, pre-treatment and variable reduction; model development followed by model evaluation and optimisation. MSA: Multiple Sequence Alignment.	99
<b>Figure 6.2</b> Score plots generated after performing PCA on (a) Domain, (b) Window, (c) Substructure, (d) Single Amino acid and (e) Running Sum datasets for 137 mAbs. X axis represent Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captures by the respective PC. Coloured ellipses indicate the 95% Confidence limits of the corresponding class.	104
<b>Figure 6.3</b> Score plots generated after performing PCA on (a) Domain, (b) Window, (c) Substructure, (d) Single Amino acid and (e) Running Sum datasets for 67 mAbs. Coloured ellipses indicate the 95% Confidence limits of the corresponding class. X axis represent Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captures by the respective PC.	106
<b>Figure 6.4</b> Model summary of Domain dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.	108
<b>Figure 6.5</b> Regression coefficients of GA selected variables for domain dataset model (a) VH domain and (b) VL domain	109
<b>Figure 6.6</b> Model summary of Window dataset model developed following GA-PLS selection	109
<b>Figure 6.7</b> Regression coefficients of GA selected variables for Window dataset model (a) VH domain and (b) VL domain.	111
<b>Figure 6.8</b> Model summary of Substructure dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.	112
<b>Figure 6.9</b> Regression coefficients of GA selected variables for Substructure dataset model (a) VH domain and (b) VL domain.	113
<b>Figure 6.10</b> Model summary of Single Amino dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.	114
<b>Figure 6.11</b> Regression coefficients of GA selected variables for Single amino dataset model (a) VH domain and (b) VL domain	115

<b>Figure 6.12</b> Model summary of Running Sum dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.	116
<b>Figure 6.13</b> Regression coefficients of GA selected variables for Running sum dataset model (a) VH domain and (b) VL domain.	117
<b>Figure 6.14</b> Performance based on (a) $R^2$ and (b) for PLS Calibration models generated from the four data blocks: Domain, Window, Single amino, and Running sum modelled against CIC. The models were generated using the original dataset, V-WSP reduced dataset as well as after variable selection (GA selected). Red dashed lines indicate values at 0.6 for $R^2$ and 0.5 for RMSE. Black dashed line indicates RMSE=0.3.	118
<b>Figure 6.15</b> Performance based on (a) $R^2$ (b) $Q^2$ and (c) RMSE for PLS cross validation models generated from the four data blocks: Domain, Window, Single amino, and running sum modelled against CIC. The models were generated using the original dataset, V-WSP reduced dataset as well as after variable selection (GA selected). Red dashed lines indicate values at 0.5 for $R^2$ , $Q^2$ and 0.5 for RMSE. Black dashed line indicates RMSE=0.3.	119
<b>Figure 6.16</b> Performance based on (a) $R^2$ and (b) RMSE for PLS prediction models generated from the four data blocks: Domain, Window, Single amino, and Running sum modelled against CIC. The models were generated using the original dataset, V-WSP reduced dataset as well as after variable selection (GA selected). Red dashed lines indicate values at 0.6 for $R^2$ and 0.5 for RMSE. Black dashed line indicates RMSE=0.3.	120
<b>Figure 7.1:</b> Overview of methodology for generating (a) 3D structures from primary sequence of antibodies followed by descriptor generation (b) Workflow for generating 3D structures from mAbs useful for generating structural descriptors as well as docking studies.	126
<b>Figure 7.2</b> 3D structure of mAb 2fgw with the coloured spheres representing the disulphide bridges between the fixed Cysteine residues.	128
<b>Figure 7.3</b> Score plots generated after performing PCA on (a) Chain, (b) Domain and (c) Substructure datasets for 134 mAbs. X axis represents Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captures by the respective PC. Coloured ellipses indicate the 95% Confidence limits of the corresponding class.	133
<b>Figure 7.4</b> Score plots generated after performing PCA on (a) Chain, (b) Domain and (c) Substructure datasets for 121 mAbs. X axis represents Principal Component	133

(PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captures by the respective PC. Coloured ellipses indicate the 95% confidence limits of the corresponding class.

**Figure 7.5** Model summary of Chain dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions. 135

**Figure 7.6** Model summary of Domain dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions. 136

**Figure 7.7** Model summary of Substructure dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions. 137

**Figure 7.8** PLS model-based Y residuals of test set samples of mAbs based on GA-selected variables substructure datasets. 137

**Figure 7.9** Regression coefficients of GA selected variables for Substructure dataset model. 138

**Figure 8.1** Overall methodology describing the two data augmentation methods. 145

**Figure 8.2** Model summary of model developed using combination method II. (a) Measured vs predicted values for training and test set model. (b) Test set predictions. 147

**Figure 8.3** Regression coefficients of variables present in model from running sum and substructure dataset for all Fab domains. 147

**Figure 8.4** Model summary of model developed using combination method I. (a) Measured vs predicted values for training and test set model. (b) Test set predictions. 148

**Figure 8.5** Regression coefficients of variables present in model from running sum and substructure dataset for all Fab domains. 149

**Figure 9.1** Hybrid model development workflow outlining the different steps involved in pre-treatment and variable reduction; model development followed by model evaluation and optimisation. 154

<b>Figure 9.2</b> Model summary of Running Sum dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions	157
<b>Figure 9.3</b> Regression coefficients of GA selected variables for Running Sum dataset model	158
<b>Figure 9.4</b> Score plot of GA selected variables based PLS model built on Running Sum dataset.	161

## List of Tables

<b>Table 1.1</b> List of approved monoclonal antibody derived therapeutics and associated toxicity.	5
<b>Table 1.2</b> IgG receptors and effector functions.	12
<b>Table 1.3</b> A comprehensive overview of in silico prediction tools for assessing toxicology.	19
<b>Table 2.1</b> Time dependent butyl paraben induced decrease in ATP levels expressed in term of logIC <sub>50</sub> , IC <sub>50</sub> and standard error (SE) in HepG2 and HDFn cells.	35
<b>Table 4.1</b> Quantification of aggregated content of heat-stressed mAb samples by analytical ultra-centrifugation.	62
<b>Table 4.2</b> Results of amino acid composition following EMBOSS pepstats analysis of rituximab and Herceptin primary sequences.	67
<b>Table 5.1</b> Protein descriptor sets generated using different software for predicting different mAb isotypes.	80
<b>Table 6.1</b> Dimensionality reduction by variable reduction and selection	107
<b>Table 6.2</b> Relative positions of CDR and Fr regions for heavy and light chain following sequence alignment.	110
<b>Table 6.3</b> Results of $\gamma$ randomization and permutations tests for GA selected datasets.	121
<b>Table 7.1</b> List of descriptors generated from 3D descriptors of mAbs used subsequently for model development.	130
<b>Table 7.2</b> Variable reduction based on the V-WSP reduction algorithm indicating the selected thresholds and corresponding Procrustes index as well as the final number of descriptors in the reduced dataset	134
<b>Table 7.3</b> Results of GA-PLS based variable selection method	135
<b>Table 7.4</b> Benchmarking of datasets based on Calibration, Cross validation and external test set prediction metrics. Models that have passed QSAR validation criteria are shown in green, those with moderate performance are indicated in Yellow and those that have failed the QSAR validation criteria are indicated in red.	140

<b>Table 7.5</b> Results of permutations tests for GA selected substructure datasets.	141
<b>Table 8.1</b> Data distribution of mAb samples into training and test set based on response dataset.	145
<b>Table 8.2</b> Results of permutations tests for Combination method I dataset.	148
<b>Table 8.3</b> Results of permutations tests for Combination method II dataset.	149
<b>Table 9.1</b> Details of mAbs used and sample distribution for model development.	153
<b>Table 9.2</b> Relative positions of CDR and FR regions for heavy and light chain following sequence alignment.	157
<b>Table 9.3</b> Model performance metrics of Calibration, Cross validation and external test set prediction. Models that have passes QSAR validation criteria are shown in green, those with moderate performance are indicated in Yellow and those that have failed the QSAR validation criteria are indicated in red	159
<b>Table 9.4</b> Benchmarking of PLS model performance metrics for the original, V-WSP reduced and GA selected Cross validated models. CI: Confidence limits.	159
<b>Table 9.5</b> Results of y randomization and permutations tests for GA selected datasets.	160

## List of Abbreviations

ADCC	Antibody Dependent Cytotoxicity
AICC	Antibody independent cellular cytotoxicity
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
AUC	Analytical UltraCentrifugation
CDC	Complement Dependent Cytotoxicity
CDCC	Complement dependent cellular cytotoxicity
CDR	Complementary Determining Region
CIC	Cross interaction chromatography
CP	Composite Protein
DCFDA	2',7' –dichlorofluorescin diacetate
DOE	Design of experiment
Fab	Antigen Binding Fragment
FDA	Food and Drugs Administration,
FR/FW	Framework region
GA	Genetic Algorithm
HM	Homology modelling
IC	Isotype control
ICH	International Conference on Harmonisation
IgG	Immunoglobulin G
mAb	Monoclonal Antibody
MD	Molecular dynamics
MLR	Multiple Linear Regression
MSA	Multiple sequence alignment
MVDA	Multivariate data analysis
NC	Negative control
OECD	Organisation for Economic Co-operation and Development
PBMCs	Peripheral blood mononuclear cells
PC	Positive control
PCA	Principal Component Analysis
PLS	Partial Least Squares
QSAR	Quantitative Structure Activity Relationship

QSPR	Quantitative sequence property relationship
RAU	Relative absorbance units
RLU	Relative luminescence units
RMSE	Root Mean Square Error
ROS	Reactive oxygen species
SC	Solvent control
TMRE	tetramethylrhodamine, ethyl ester
UV-B	Ultraviolet-B
V-WSP	Variable - Wootton, Sergent, Phan-Tan-Luu algorithm
WST	2-(2-methoxy-4-nitrophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H-tetrazolium)

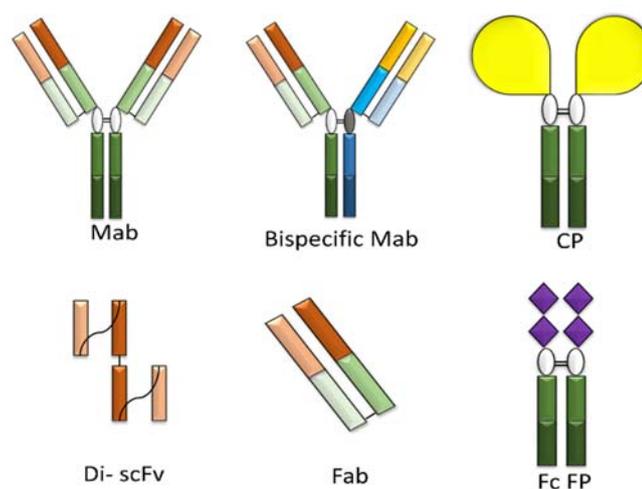


## Chapter 1. Applicability of predictive toxicology methods for monoclonal antibody therapeutics: Status Quo and Scope

The pharmaceutical market is currently valued at \$830 billion and is expected to reach \$1.2 trillion by 2024 (EvaluatePharma®, 2018). Biopharmaceuticals play a very important role in this increase and the contribution will increase from 49% to 52% of the world pharmaceutical sales and will dominate the oncology and anti-rheumatics market till 2024 both in sales and R&D (EvaluatePharma®, 2018). Biological drugs are associated to living entities (cells and tissues) and/or their product such as recombinant therapeutic proteins and vaccines to name a few. Based on historical data, a shift towards biologics seems imminent owing to increasing profits and lower attrition rates when compared to small molecule drugs. Biological drugs comprised 70% of the top ten selling products of the world in 2017 and the percentage sales of biotechnology products within the top 100 was 49% (EvaluatePharma®, 2018). Twenty new biologics were approved by FDA in 2014 compared to the 11 that were approved in 2009. Monoclonal antibodies have higher approval rates of 26% in the biopharmaceutical sector than that of conventional small molecule drugs (10%) (Hay *et al.*, 2014; Sewell *et al.*, 2017). Based on the area of therapy, the largest segments of oncology and anti-rheumatoid drugs, which contribute to a combined compound annual growth rate of 13%, continue to be dominated by biological drugs. Although the trends seem to be in favour of biopharmaceutical development, the growth rates have not yet reached their full potential due to financial and technical complexities involved in early stages of research and development, bioprocess development and preclinical testing. Compared to the 21 biologics approved by FDA in 2015 only 19 were approved in 2017 indicating a scarcity in innovativeness, lack of progress in bioprocess development strategies mainly in early stage screen and manufacturability. This coupled with the second patent expiration cliff around the corner leading to the biosimilar boom, there is a pressing need for rapid bioprocess development strategies to be put into place as well as push for faster commercialisation. The aforementioned challenges are estimated to put around \$251 billion of sales at risk (EvaluatePharma®, 2018). Tackling these challenges will allow for redirection of resources into more innovation within the industry as well as focus on drugs for rare and neglected diseases.

Even though the therapeutic efficiency of immunoglobulin molecules was demonstrated in 1890, it was only after Kohler and Milstein elucidated the murine hybridoma technology for *in vitro* production of mAbs (see Figure 1.1 for generic mAb structures) that the market for mAbs grew and expanded to different therapy areas, such as haematology, oncology, immunology,

cardiology, infectiology and ophthalmology as well as diagnostics and imaging (Köhler and Milstein, 1975). The shift from murine mAbs to chimeric (human Fc region with murine Fv region) was mainly to increase titres as well as decrease immunogenic effects (Zhu, 2012). To further decrease the murine composition and enhance Fc functionality, humanized mAbs were first developed in 1986 (Jones *et al.*, 1985). The production systems routinely used for chimeric and humanized mAbs are Chinese hamster ovary (CHO) cells, NS0 and Sp2/0 myeloma cell line. To fully eliminate the immunogenic potential of murine epitopes while maintaining optimal Fc region functionality, fully human mAbs were developed by phage display technology and commercially produced by CHO system (Lai *et al.*, 2013). Human Embryonic Kidney (HEK) and human retinal cell derived (Per.C6) cell lines are the new potential candidates for biopharmaceutical production (Zhu, 2012). In addition to being stable and producing high titres, the fully human cell lines offer the advantage of proper post translation modification and glycosylation as they incorporate human biosynthetic pathways. Plant expression systems, such as recombinant *Agrobacterium tumefaciens*, and microbial systems, such as *Escherichia coli*, are gaining popularity for production of monoclonal antibodies against viruses (Ma *et al.*, 2003; Berlec and Štrukelj, 2013; Rosenberg *et al.*, 2013). Transfected HEK cells have already been used to produce recombinant coagulation factors which have been approved by FDA (Food and Drug Administration), however full length mAbs produced by them are still awaiting approval (Berlec and Štrukelj, 2013; Lai *et al.*, 2013). Furthermore proprietary technologies, such as VelocImmune®, POTELLIGENT™, UltiMAb® and Xenomouse®, are used for production of monoclonal antibodies (Jakobovits *et al.*, 2007; Murphy, 2009; Shitara, 2009; Sheridan, 2010; Nelson and Paulos, 2015; Hurrell, 2018; Kennedy *et al.*, 2018). The mAb derived products include fusion proteins, antigen binding fragments as well as composite proteins (Povey *et al.*, 2001; Lefranc *et al.*, 2009; Li and Zhu, 2010; Ecker *et al.*, 2015; Hurrell, 2018; Kennedy *et al.*, 2018).



**Figure 1.1** Generic monoclonal antibody derived therapeutic structures as adapted from IMGT (World Health, 2006; Lefranc *et al.*, 2009). Fc: Constant region which contributes to effector function, immune response and increased half-life, Fv: Variable region that contains Complementarity Determining Regions (CDRs) facilitating antigen binding, Fab: Antigen Binding Fragment which lack Fc Region, scFv: Single chain Fragment variable, FP: Fc Fusion Proteins that contain Fc region for effector functionality (e.g. Abatacept), CP: Composite protein that contain Fc region for increasing half-life and not for effector functionality (e.g Strensiq™)(World Health, 2006).

### 1.1. Aims and objectives of the research

The main aim of this research was to facilitate QSAR model development framework for early stage screening of mAbs candidates that would allow for their rapid developability. Advancements in early stage screening of mAb candidates would reduce the number of lead candidates entering the bioprocess pipeline that are associated with adverse effects thereby reducing attrition rates. The research objectives of this study are as follows:

- a) To assess the utility of traditional toxicity assays for assessing the different toxicity endpoints of chemical compounds.
- b) To investigate the applicability of traditional toxicity assays for assessing the adverse effects associated with mAb therapeutics.
- c) To assess the utility of novel assays for assessing immune related adverse effects of mAbs and/or their aggregates.
- d) To identify and generate numerical features from the primary sequence and 3D structure of mAbs.
- e) To use multivariate data analysis techniques to identify any potential correlation between the numerical features and biological responses (adverse effects) elicited by mAbs.

- f) To develop a hybrid QSAR based model with a structured workflow and clear evaluation metrics, with several optimisation steps, that could be beneficial for broader and more generic modelling.

## **1.2. MAbs: Safety pharmacology and side effects**

MAbs and related therapeutics are highly desirable from a biopharmaceutical perspective as they are highly target specific and well tolerated within the human system (Shepard *et al.*, 2017). Nevertheless several mAbs have been discontinued or withdrawn based either on their inability to demonstrate efficacy and/or due to adverse effect, for example efalizumab, biciromab and fanelesomab while others were discontinued due to high manufacturing costs, for example imciromab and arcitumomab (Lefranc *et al.*, 2009). Approved monoclonal antibodies as well as derived products have been associated with adverse effects and these effects have been classified into categories of specialised toxicity as indicated in Table 1.1. The adverse effects associated with these toxicities are outline in Box 1.1 (Hansel *et al.*, 2010; Peluso *et al.*, 2013).The reporting of these adverse effects is to be treated with caution as there are several factors that influence them, such as underlying conditions, drug combinations, reporting practices and clinical practice involved in the clinical trials.

**Table 1.1** List of approved monoclonal antibody derived therapeutics and associated toxicity.

Generic name Trade name	Type	Antigen <sup>a</sup>	Species	Therapy area	Production	Therapy associated adverse effect
<b>Abatacept</b> ORENCIA®	FP	CD80, CD86	<i>Homo sapiens</i>	Immunology	CHO	Occular toxicity, Immunotoxicity, Dermal toxicity, Infection
<b>Abciximab</b> REOPRO®	Fab IgG1κ	ITGA2B_ ITGB3	Chimeric	Cardiology	Sp2/0	Immunotoxicity, Haemotoxicity
<b>Adalimumab</b> HUMIRA®	IgG1κ	TNF	<i>Homo sapiens</i>	Immunology	CHO	Immunotoxicity Cardiotoxicity, Infection, Hepatotoxicity Haemotoxicity, Others
<b>Aflibercept</b> ZALTRAP® EYLEA®	FP	VEGFA	<i>Homo sapiens</i>	Ophthalmolog y, oncology	CHO K-1	Occular toxicity Haemotoxcity Cardiotoxicity
<b>Alemtuzamab</b> CAMPATH- 1H®, LEMTRADA®	IgG1κ	CD52	Humanized	Haematology, oncology, immunology	CHO	Immunotoxicity, Haemotoxicity Cardiotoxicity Others
<b>Alirocumab</b> PRALUENT®	IgG1κ	PCSK9	<i>Homo sapiens</i>	Cardiology	VelocImmu ne®	Neurotoxicity, Dermal toxicity Occular toxicity Cardiotoxicity
<b>Asfotase alpha</b> STRENSIQ™	CP		<i>Homo sapiens</i>	Hypophosphat as-ia	CHO	Immunotoxicity, Dermal toxicity, Renal toxicity, Occular toxicity Others
<b>Basiliximab</b> SIMULECT®	IgG1κ	IL2RA	Chimeric	Immunology	Sp2/0	Immunotoxicity Dermal toxicity
<b>Belatacept</b> NULOJIX®	FP	CD80, CD86	<i>Homo sapiens</i>	Immunology	CHO	Renal toxicity, Infection, Others
<b>Belimumab</b> BENLYSTA®	IgG1λ	TNFSF13 B	<i>Homo sapiens</i>	Immunology	NS0 (serum free)	Immunotoxicity Infection, Others
<b>Besilesomab</b> SCINTIMUN®	IgG1κ	CEACAM 8	<i>Mus musculus</i>	Osteology (diagnostic)	Hybridoma technology*	Cardiotoxicity Immunotoxicity
<b>Bevacizumab</b> AVASTIN®	IgG1κ	VEGFA	Humanized	Oncology	CHO	Cardiotoxicity, Infection, Haemotoxicity, Gastrointestinal, Others
<b>Blinatumomab</b> BLINCYTO®	scFv κH - scFv κH	CD19, CD3E	<i>Mus musculus</i>	Haematology, oncology	CHO	Immunotoxicity, Neurotoxicity
<b>Brentuximab</b> ADCETRIS™	IgG1κ	TNFRSF8	Chimeric	Oncology	CHO	Cardiotoxicity, Infection, Pulmonary toxicity
<b>Canakinumab</b> ILARIS®	IgG1κ	IL1B	<i>Homo sapiens</i>	Hereditary inflammatory diseases; Immunology	UltiMab®	Infection, Others
<b>Capromab</b> PROSTASCIN T®	IgG1κ	FOLH1	<i>Mus musculus</i>	Oncology	Hybridoma technology* *	NR
<b>Catumaxomab</b> REMOVABT®	IgG2a κ/G2b λ	CD3E, EPCAM	<i>Mus musculus</i> <i>Rattus sp.</i> Hybrid	Oncology	Quadroma technology <sup>+</sup>	Haemotoxicity, Immunotoxicity, Others
<b>Certolizumab</b> CIMZIA®	Fab'- G1κ	TNF	Humanized	Immunology	Escherichia coli	Immunotoxicity Cardiotoxicity, Infection, Hepatotoxicity Haemotoxicity
<b>Cetuximab</b> ERBITUX®	IgG1κ	EGFR	Chimeric	Oncology	Sp2/0	Immunotoxicity, Dermal toxicity, pulmonary Toxicity

<b>Daclizumab</b> <sup>***</sup> ZENAPAX®	IgG1κ	IL2RA	Humanized	Immunology	NS0	Immunotoxicity, Dermal toxicity
<b>Daratumumab</b> DARZALEX™	IgG1κ	CD38	<i>Homo sapiens</i>	Haematology, oncology, immunology	UltiMAb®	Haemotoxicity, Immunotoxicity, Pulmonary toxicity
<b>Denosumab</b> PROLIAS® XGEVAS®	IgG2	TNSF11	<i>Homo sapiens</i>	Osteology	XenoMouse®	Haemotoxicity, Infection
<b>Eculizumab</b> SOLIRIS™	IgG2/ G4κ	C5	Humanized	Haematology	NS0	Haemotoxicity, Infection
<b>Edrecolomab</b> PANOREX®	IgG2a κ	EPCAM	<i>Mus musculus</i>	Oncology	Sp2/0	Immunotoxicity, Others
<b>Elotuzumab</b> EMPLICITI™	IgG1κ	SLAMF7	Humanized	Haematology, oncology, immunology	NS0(Varma <i>et al.</i> , 2014)	Haemotoxicity, Gastrointestinal, Others
<b>Etanercept</b> ENBREL®	FP	TNF	<i>Homo sapiens</i>	Immunology	CHO	Infection, Cardiotoxicity, Hepatotoxicity, Immunotoxicity
<b>Evolocumab</b> REPATHA™	IgG2λ	PCSK9	<i>Homo sapiens</i>	Cardiovascular diseases	XenoMouse®	Immunotoxicity, Haemotoxicity, Infection, Others
<b>Factor IX Fc FP</b> ALPROLIX®	CP	NA	<i>Homo sapiens</i>	Haematology	Transfected HEK cell line	NR
<b>Factor VIII Fc FP</b> ELOCTATE®	CP	NA	<i>Homo sapiens</i>	Haematology	Transfected HEK cell line.	NR
<b>Golimumab</b> SIMPONI® ARIA®	IgG1κ	TNF	<i>Homo sapiens</i>	Immunology	UltiMAb®	Dermal toxicity
<b>Ibritumomab</b> ZEVALIN®	IgG1κ	MS4A1	<i>Mus musculus</i>	Oncology	CHO	Haemotoxicity, Dermal toxicity, Others
<b>Idarucizumab</b> PRAXBIND®	Fab- G1κ	Pradaxa®: Dabigatran etexilate mesylate	Humanized	Reversal of drug overdose	CHO	Dermal toxicity, Gastrointestinal, Infection, Others
<b>Infliximab</b> REMICADE®	IgG1κ	TNF	Chimeric	Immunology	Sp2/0	Immunotoxicity Cardiotoxicity, Infection, Hepatotoxicity Haemotoxicity, Others
<b>Ipilimumab</b> YERVOY®	IgG1κ	CTLA4	<i>Homo sapiens</i>	Oncology	UltiMAb®	Hepatotoxicity, Neurotoxicity, Pulmonary toxicity, Gastrointestinal toxicity
<b>Mepolizumab</b> NUCALA®	IgG1κ	IL5	Humanized	Immunology	CHO	Infection, Cardiotoxicity, Others
<b>Mogamulizumab</b> POTELIGEO®	IgG1κ	CCR4	Humanized	Haematology, oncology	POTELLIGENT®	Immunotoxicity, Dermal toxicity
<b>Muromonab-CD3</b> ORTHOCLON E OKT3®	IgG2a κ	CD3E	<i>Mus musculus</i>	Immunology	Hybridoma murine ascites	Immunotoxicity, Hepatotoxicity, Cardiotoxicity
<b>Natalizumab</b> TYSABRI®	IgG4	ITGA4	Humanized	Immunology	NS0	Immunotoxicity, Hepatotoxicity, Infection
<b>Necitumumab</b> PORTRAZZA™	IgG1κ	EGFR	<i>Homo sapiens</i>	Oncology	UltiMAb®	Haemotoxicity, Immunotoxicity, Pulmonary toxicity, Hepatotoxicity
<b>Nimotuzumab</b>	IgG1κ	EGFR	Humanized	Oncology	NS0	Dermal Toxicity

THERACIM®						
<b>Nivolumab</b> OPDIVO®	IgG4κ	PDCD1	<i>Homo sapiens</i>	Oncology	UltiMAb®	Immunotoxicity, Hepatotoxicity, Gastrointestinal Toxicity, Pulmonary toxicity, Renal toxicity
<b>Obinutuzumab</b> GAZYVA®	IgG1κ	MS4A1	Humanized	Haematology, oncology	GlycoMAb®	Infection
<b>Ofatumumab</b> ARZERRA®	IgG1κ	MS4A1	<i>Homo sapiens</i>	Haematology, oncology	UltiMAb®, NS0	Infection, Gastrointestinal Toxicity
<b>Omalizumab</b> XOLAIR®	IgG1κ	IGHE	Humanized	Immunology	CHO	Immunotoxicity, Dermal Toxicity, Infection
<b>Palivizumab</b> SYNAGIS	IgG1κ	RSV glycoprotein in F	Humanized	Infectiology	NS0	Immunotoxicity, Others
<b>Panitumumab</b> VECTIBIX®	IgG2κ	EGFR	<i>Homo sapiens</i>	Oncology	XenoMouse® CHO	Immunotoxicity, Pulmonary Toxicity, Dermal Toxicity
<b>Pembrolizumab</b> KEYTRUDA®	IgG4κ	PDCD1	Humanized	Oncology	CHO	Immunotoxicity, Pulmonary, Others
<b>Pertuzumab</b> PERJETA®	IgG1κ	ERBB2	Humanized	Oncology	CHO <sup>++</sup>	Reproductive and developmental toxicity, Dermal toxicity, Haemotoxicity, Immunotoxicity, Cardiotoxicity
<b>Ramucirumab</b> CYRAMZA®	IgG1κ	KDR	<i>Homo sapiens</i>	Oncology	NS0	Haemotoxicity, Cardiotoxicity, Gastrointestinal, Others
<b>Ranibizumab</b> LUCENTISO®	Fab G1κ	VEGFA	Humanized	Ophthalmology, immunology	<i>Escherichia coli</i>	Cardiotoxicity, Haemotoxicity, Ocular toxicity
<b>Raxibacumab</b> ABTHRAX®	IgG1λ	anthrax protective antigen	<i>Homo sapiens</i>	Infectiology	CHO	Haemotoxicity, Infection, Dermal toxicity, Others
<b>Rilonacept</b> ARCALYSTF®	FP	IL1A	<i>Homo sapiens</i>	Immunology	CHO	Dermal toxicity, Immunotoxicity
<b>Rituximab</b> MABTHERA®, RITUXAN®	IgG1κ	MS4A1	Chimeric	Haematology, oncology, immunology	CHO-MR	Immunotoxicity, Cardiotoxicity, Infection, Others
<b>Romiplostim</b> NPLATE®	CP	MPL	<i>Homo sapiens</i>	Immunology	<i>Escherichia coli</i>	Haemotoxicity, Infection, Others
<b>Secukinumab</b> COSENTYX®	IgG1κ	IL17A	<i>Homo sapiens</i>	Immunology	XenoMouse®	Infection, Haemotoxicity, Cardiotoxicity
<b>Siltuximab</b> SYLVANT®	IgG1κ	IL6	Chimeric	Haematology, oncology, immunology	CHO	Immunotoxicity, Gastrointestinal toxicity, Infection
<b>Tocilizumab</b> ACTEMRA®, RoACTEMRA®	IgG1κ	IL6R	Humanized	Oncology, immunology	CHO-DR	Immunotoxicity, Infection, Hepatotoxicity, Others
<b>Trastuzumab</b> TRASTUZUMAB®	IgG1κ	ERBB2	Humanized	Oncology	CHO-MR	Immunotoxicity, Hepatotoxicity, Cardiotoxicity, Pulmonary Toxicity, Dermal toxicity
<b>Ado-trastuzumab (emantsine)</b> KADCYLAN®	IgG1κ	ERBB2	Humanized	Oncology	CHO	Reproductive and Developmental Toxicity, Dermal Toxicity,

<b>Ustekinumab</b> STELARA®	IgG1κ	IL12B	<i>Homo sapiens</i>	Immunology	UltiMAb®	Hepatotoxicity, Cardiotoxicity Neurotoxicity, Cardiotoxicity Others
<b>Vedolizumab</b> ENTYVIO®	IgG1κ	ITGA4 ITGB7	Humanized	Immunology	CHO	Infection, Pulmonary Toxicity, Other

<sup>a</sup>Nomenclature derived from HUGO Gene nomenclature Committee resources (Povey *et al.*, 2001) \* X63Ag8.653 and spleen cells from Balb/c mice previously immunised with CEA antigen (from human liver metastasis) \*\* fusing P3x63Ag8.653 myeloma cells with spleen cells from BALB/c mice immunized with whole cells and membrane extracts of the human prostate adenocarcinoma cell line LNCaP. + Consists of mouse IgG2a and rat IgG2b; ++ Fed-batch process using a suspension-adapted CHO cell line \*\*\* EC withdrawal; FP: Fusion Protein, CP: Composite Protein, Fab: Antigen binding Fragment, IgG: Immunoglobulin G, CHO: Chinese Hamster ovary cells, CHO-DR: Chinese Hamster ovary cells dihydrofolate reductase; CHO-MR: Chinese Hamster ovary cells methotrexate resistant; NS0+NS1: Nonsecreting Murine myeloma cells, Sp2/0: Hybridoma B lymphocyte, NA: Not Applicable, HEK: Human embryonic Kidney Cell line HEK293.

**Box 1.1** Adverse effects of mAbs categorised into specialised toxicities.

***Immunotoxicity:***

Infusion reactions (acute, severe) hypersensitivity, Immunogenicity, anaphylaxis (0.1%), Churg-Strauss syndrome, acute infusion reactions, cytokine release syndrome, immunosuppression, IgE against oligosaccharide and HAMA, immune haemolytic anaemia, Immune thrombocytopenia, Serum Sickness.

***Infection:***

Upper respiratory tract infections Progressive multifocal leukoencephalopathy Hepatitis B reactivation, serious opportunistic viral and/or bacterial infection, Meningococcal and Neisseria infection, tuberculosis reactivation, osteomyelitis.

***Haemotoxicity:***

Intravascular haemolysis, haemolytic anaemia, thrombocytopenia, haemorrhage, arterial and venous thromboembolic events, pancytopenia, lymphopenia, leukopenia, neutropenia, increased risk of bleeding, osteonecrosis, arterial thromboembolism, epistaxis.

***Cardiotoxicity:***

Transient hypotension, cardiac arrhythmias, severe hypertension, congestive heart failure, cardiomyopathy, pericarditis, myocardial infarction, ischaemic attack.

***Hepatotoxicity:***

Elevated liver transamines, abnormal liver function, lipid deregulation, neutropenia.

***Gastrointestinal Toxicity:***

Diarrhoea nausea vomiting, gastrointestinal perforation, bowel obstruction, enterocolitis.

***Pulmonary toxicity:***

Bronchospasm, interstitial lung disease, pulmonary fibrosis, pneumonitis, pulmonary embolism, pneumothorax.

***Ocular Toxicity:***

Conjunctival haemorrhage, intraocular inflammation, increased intraocular pressure, retinal detachment, endophthalmitis, uveitis.

***Dermal toxicity:***

Injection site reaction, severe mucocutaneous reactions, skin rashes and reactions, urticarial, angioedema.

***Renal Toxicity:***

Nephritis, proteinuria, tubular damage, hypophosphatemia, hypomagnesemia, renal failure, pyelonephritis.

***Neurotoxicity:***

Guillain–Barré syndrome, encephalitis, meningitis, neuropathy, neurocognitive disorders.

***Reproductive/Developmental Toxicity:***

Birth defects, embryo/foetal mortality.

***Others:***

Fever, headache, cough, secondary malignancies, wound dehiscence, thyroid disorders,

The catastrophic TGN1412 clinical trial that resulted in multiple organ failure of six healthy volunteers reiterated the need for more appropriate preclinical safety testing. The underlying problems that were subsequently identified in this trial were mainly the lack of appropriate preclinical testing and model organisms chosen for study of adverse effects. The standard *in vitro* assays failed to capture the *in vivo* adverse effects in humans (Stebbing *et al.*, 2013). *In vivo* toxicity studies using rodent or primate models are not always representative of the human system. Human therapeutics such as monoclonal antibodies are highly specific and targeted and there is therefore a higher likelihood of false positive efficacy or false negative toxicity if such entities are tested in non-human models, both outcomes being highly undesirable (Sewell *et al.*, 2017).

Eloctate showed haemotoxicity and hepatotoxicity in animal studies (mice and monkeys) but none have been reported in human clinical trials (Lower, 2015). TGN1412 did not show the pro-inflammatory cytokine storm in *in vivo* tests (cynomolgus macaques) due to the absence of CD28 on its CD4<sup>+</sup> effector memory T cells as well as in *in vitro* tests (human lymphocytes) due to the lack of localisation of cell receptor (Stebbing *et al.*, 2013). There are different factors which can influence the safety and efficacy of mAbs. Binding affinity, glycoforms, valency and density of antigens as well as antibodies, cell surface receptor and binding interface are some of the factors that contribute to the biological activity of mAbs and, if suboptimal, could lead to reduction of efficacy or an increase in toxicity (Stebbing *et al.*, 2013; Jefferis, 2014). nimotuzumab exhibits lower dermal toxicity due to optimal binding affinity to EGFR that ensures its binding below toxic levels (Boland and Bebb, 2009)

Effector functions of mAbs and related products, such as antibody dependent cell phagocytosis (ADCP), antibody dependent cytotoxicity (ADCC), complement dependent cytotoxicity (CDC), complement dependent cellular cytotoxicity (CDCC) as well as evoking other cell mediated immune responses, are modulated via the Fc region by interaction with FcγR receptors on different immune responsive cells (Figure 1.2 (a)) (Carter, 2006) . This also regulates the pharmacokinetics, transcytosis, catabolism and placental transfer of antibodies via the FcRn (neonatal Fc Receptor) as summarised in Table 1.2 (Roopenian and Akilesh, 2007). Glycosylation at the Fc region occurs at N297 and consists of a core heptasaccharide region comprising mostly N-Acetylglucosamine and mannose residues as well as the variable region as seen in Figure 1.2 (b) (Carter, 2006). Modifying the Fc region either via amino acid substitution or by a change in glycosylation pattern has shown to change effector functionality. IgG1 based therapeutic antibodies have shown increased ADCC and ADCP activity with

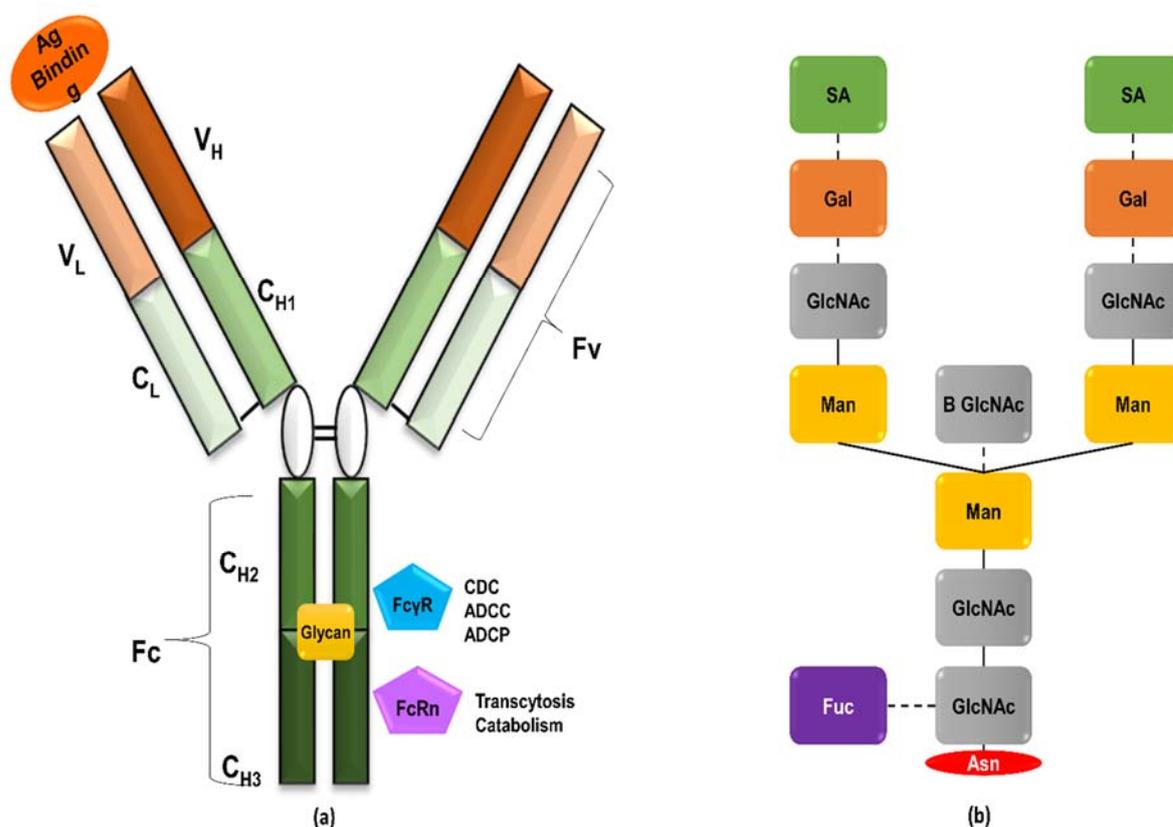
substitution at amino acid positions 298,333 and 334 whereas oteelixizumab has shown reduced ADCP and ADCC activity with an N297A substitutions (Bolt *et al.*, 1993; Shields *et al.*, 2001). The mammalian cell production systems could alter the glycoform and this could either change the effector function mediated therapeutic activity or induce immunogenic effects of mAbs (Jefferis, 2009). Afucosylation and bisecting N acetylglucosamine were reported for antibodies produced in CHO cells and they were associated with increased ADCC activity (Umaña *et al.*, 1999; Shields *et al.*, 2002). Galactosylation levels are important for different functions, such as transport of IgG molecules across placenta and complement activation. Mammalian cell lines generally produce hypogalactosylated products however if this hypogalactosylation is unintended, it could impact effector function. This has been demonstrated with alemtuzumab and rituximab where the removal of galactose residues reduced complement activation (Boyd *et al.*, 1995; Raju and Jordan, 2012). Mammalian production systems can also add oligosaccharides not present in human system, such as addition of N-glycolylneuraminic acid by CHO, NS0 and Sp2/0 systems, which can be immunogenic (Jefferis, 2014). Differing glycation patterns can also alter antigen binding and this has been quantitatively determined in a recent study (Mo *et al.*, 2018). These modification on engineered mAbs, when intentional, can alter effector functionality to enhance the therapeutic profile of mAbs (Wang *et al.*, 2018). However, when accidental it could lead to potential adverse effects.

From Table 1.1 it can be observed that following immunotoxicity, hepatotoxicity and dermal toxicity constitute majority of the adverse effects elicited by mAbs. These could be either a consequence of the exaggerated immune response elicited by mAbs or a potential off target effect wherein the mAb bind to an off-target tissue and elicits a response. In vitro systems traditionally used for assessing hepatotoxicity, such as HepG2 cells, and dermal toxicity, such as HDFn cells, of small molecules and chemical compounds may provide an insight into these potential off target effects of mAbs. These aspects have been explored in Chapter 3.

**Table 1.2** IgG receptors and effector functions.

	Function	Binding affinity		Expression	Important AA residues	Impact of Glycosylation <sup>b</sup>
		IgG subclass	Ka(10 <sup>6</sup> M <sup>-1</sup> )			
<b>C1q</b>	CDC	**	NA	Present in serum	L235,D265,D270,K322,P329,P331,H433	Galactose: ↑ CDC; Mannose: ↓ CDC
		*	NA			
		***	NA			
		-	NA			
<b>FcγRI</b>	Activation	***	65	Monocytes, macrophages Dendritic Cells Neutrophils <sup>1</sup> Mast Cells <sup>1</sup>	E233, L235, G236	Unclear
		-	-			
		****	61			
		**	34			
<b>FcγRIIA (H131)</b>	Activation	***	5.2	Monocytes, macrophages Dendritic Cells Neutrophils Mast Cells	L234,L235,G236, A327	Unclear
		**	0.45			
		****	0.89			
		**	0.17			
<b>FcγRIIA (R131)</b>		***	3.5	Neutrophils Basophils Eosinophils		
		*	0.10			
		****	0.91			
		**	0.21			
<b>FcγRIIB/C</b>	Inhibition	*	0.12	B cells Dendritic cells Basophils Monocytes <sup>a</sup> Macrophages <sup>a</sup> Neutrophils <sup>a</sup>	Unclear	Unclear
		-	0.02			
		**	0.17			
		*	0.20			
<b>FcγRIIA (F158)</b>	Activation	**	1.2	Natural Killer Cells Monocytes Macrophages	E233,L234,L235G236	Mannose, Bisecting GlcNAc: ↑ ADCC; Sialic acid, fucose: ↓ADCC
		-	0.03			
		****	7.7			
		-	0.20			
<b>FcγRIIA (V158)</b>		***	2.0			
		*	0.07			
		****	9.8			
		**	0.25			
<b>FcγRIIB</b>	Unclear	***	0.2	Neutrophils Basophils	L234,L235G236,G237,P238	Unclear
		-	-			
		****	1.1			
		-	-			
<b>FcRn</b>	Transcytosis Catabolism Antigen uptake	***	80	Monocytes, macrophages, Dendritic Cells Neutrophils Endothelium Syncytiotrophoblast	H433,N434,H435,Y436	Galactose, Mannose, GlcNAc: ↑ Clearance
		***	NA			
		**/****	NA			
		***	NA			

● IgG1, ● IgG2, ● IgG3, ● IgG4; \*\*\*\* very high affinity; \*\*\*high affinity; \*\* moderate affinity; \* low affinity; - no binding; <sup>1</sup>Inducible expression <sup>a</sup> low percentages; <sup>b</sup> (Liu, 2015); NA: Not applicable; AA: Amino acid;



**Figure 1.2 (a)** Monoclonal antibody structure with binding site for antigen, Fc $\gamma$ R and FcRn receptor as well as glycosylation sites (Glycan); Ag: Antigen; CDC: Complement Dependent Cytotoxicity; ADCC: Antibody Dependent Cell Cytotoxicity; ADCP: Antibody Dependent Cell Phagocytosis **(b)** Glycosylation profile at N297 residue of the Fc region of antibodies. The bold line indicates core structures and dotted line indicates variable structures. Gal: Galactose; SA: Sialic Acid; Man: Mannose; GlcNAc: N-Acetylglucosamine Fuc: Fucose; Asn: Asparagine (N297).

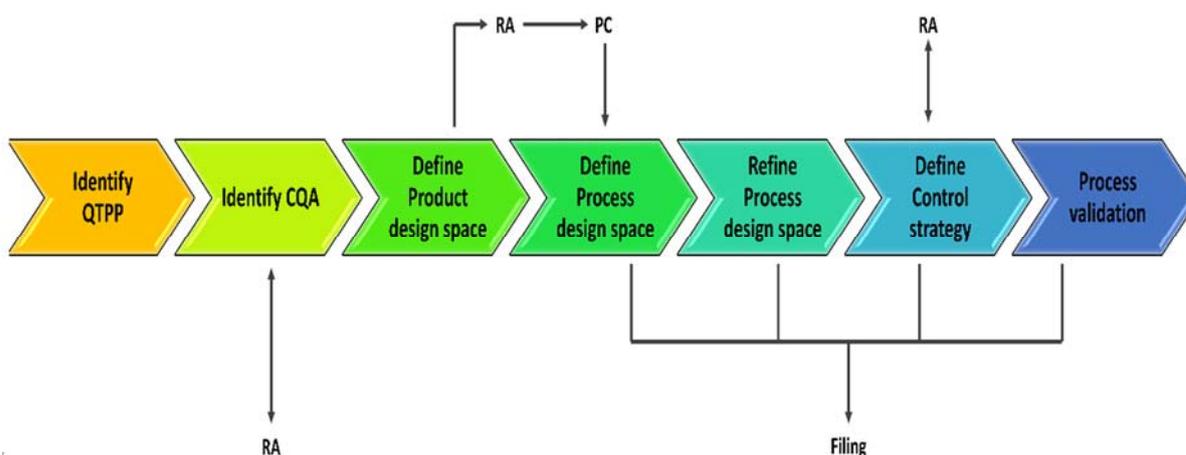
Although the trends seem to be in favour of biopharmaceutical development, the growth rates have not yet reached their full potential due to financial and technical complexities involved in early stages of research and development and preclinical testing as described in the following sections. The comprehensive costs of developing a new drug amount to \$2.8 billion (Pharma, 2014). Studies done over the past decade show that nearly 90% of drugs failed in clinical development (30% in Phase I and 33% in Phase II) and this high attrition rate is the major contributing factor to the exorbitant cost of new drug development (Hay et al., 2014; Kola and Landis, 2004; Paul et al., 2010). Thus, it is more beneficial to address attrition, as a 10-15% decrease in attrition rate could reduce the cost of drug development by nearly 35% (Paul et al., 2010). Recent studies reported that toxicity and lack of efficacy were the most important factors for high attrition rates in small molecule drug development (Waring et al., 2015)(Pellicciari, 2017). Unlike conventional drugs which mainly revolve around small molecule chemistry, biological drugs are far more complex to produce and characterise as they are 200-1000 times

larger, structurally more complex and highly sensitive to their manufacturing conditions. The costs involved in development and production of biopharmaceutical entities is 1.5-2.5 times higher than that of small molecule drugs (Blackstone and Fuhr, 2007). With nearly 80% of biological drugs failing in clinical development mainly due to lack of efficacy and safety there arises an urgent need for smarter preclinical development. This requires better product understanding i.e. examining characteristics which contribute to product quality such as biological activity, affinity, pharmacology, toxicity, immunogenicity etc. thus leading to early prediction of success/failure. Improved product understanding and rapid screening of potential drug candidates by utilising different in vitro and in silico methods to predict efficacy and safety techniques would lead to better preclinical design.

### **1.3. Quality by Design in mAb developability**

The need for Quality by Design (QbD) arose in early 2000s as the traditional approach to biopharmaceutical drug development was primarily empirical. This, combined with limited emphasis being placed on process understanding, led to manufacturing failures and consequently increase in product wastage. Furthermore, limited attention was being given to the analysis of the root causes of these failures as well as to the prediction of scale up effects on the final product quality and yield.

The concept of QbD was incorporated into the pharmaceutical manufacturing control review process in 2004 following the pharmaceutical cGMP for the 21st century initiative (Figure 1.3). The main aim of the initiative was to harmonise and modernise product, process and regulatory aspects of pharmaceutical manufacturing. The QbD paradigm urges industries put in place the following principles into their own process: identify and have a better understanding of their product profile in terms of safety and efficacy; identify physical, chemical or microbiological properties that can affect safety and efficacy and thus should be controlled within a predetermined range; design the process so as to deliver the required quality product consistently and controlling it by putting in place a robust control strategy; validate, document and continuously monitor the process to make sure the system/bioprocess performs robustly over the product life cycle (Rathore, 2009; Rathore and Winkle, 2009).



**Figure 1.3** A schematic for Quality by design in biopharmaceutical industry; RA: Risk assessment, PC: Process characterisation (Rathore and Winkle, 2009).

Recent studies reported that toxicity and lack of efficacy were the most important factors for high attrition rates in small molecule drug development (Waring *et al.*, 2015). With nearly 80% of biological drugs failing in clinical development mainly due to lack of efficacy and safety there arises an urgent need for smarter preclinical development using quality by design-based approaches. This requires better product understanding i.e. examining characteristics which contribute to product quality such as biological activity, affinity, pharmacology, toxicity, immunogenicity etc. thus leading to early prediction of success/failure. From a QbD point of view, this encompasses the principles of Quality Target Product Profile (QTTP) and Critical Quality Attributes (CQA). QTTP includes dosage form and strength, route of administration, intended use, release/delivery of therapeutic moiety, pharmacokinetics as well as drug product criteria required for intended market (stability and purity). Following the identification of QTTP, CQAs are then defined based on risk assessments. The rationale for identifying CQAs and relating them to product safety and efficacy arises from QTTP, prior product/process knowledge, literature, clinical data, non-clinical data including platform based, *in vivo* as well as *in vitro* assays and relevant data from similar products. The main CQAs for biological drugs would fall into the categories of size, charge, primary structure, post translation modifications and higher order structure. For monoclonal antibodies, examples of product related CQAs are fragmentation, aggregation, inadvertent glycosylation and conformation whereas process related CQAs are host cell proteins, formulation buffer components, culture media residues to name a few (Group, 2009).

QbD offers a plethora of advantages to the biopharmaceutical industry as well as regulatory agencies. It allows for easy implementation of innovative technologies, technical and scientific

harmonisation of regulations, minimisation of manufacturing failures, reduction of attrition rates and it ensures consistent production of high quality products by continuously monitoring and updating bioprocesses. However, for the full potential of QbD to be exploited, there are still many challenges to be overcome like filing terminology, training, fitting existing legacy products into the QbD frameworks, conflict of opinions amongst regulatory agencies across the world and most importantly, a limited understanding of how CQAs impact the quality and efficacy of a product. A possible solution could be to increase the use of non-clinical studies as well as *in silico* tools to augment knowledge about the impact of CQAs on the clinical potency and safety of biological drugs (Kizhedath *et al.*, 2016). These rapid screening methods can be seamlessly integrated into a Quality by design approach when linked with the target product profile as well.

Implementation of QbD and PAT could indeed revolutionise the way biopharma industries operate by cutting down their attrition rates and this could drastically change their business model enabling them to save time, money and resources. Historical data in this context from approved products can be used to power pattern recognition models that would allow for early stage screening of lead candidates thereby reducing the load on the bioprocess pipeline and allowing for better product and consequently process design i.e. by narrowing the design space. This could lead to reduction of prices for biologic drugs as well as more focus on medicines for rare diseases thus proving beneficial to society.

#### **1.4. In vitro systems for toxicity testing**

The general *in vitro* toxicity testing panel includes cellular, biochemical and molecular assays to study cytotoxicity, reactive oxygen species production as well as specialised toxicity effects including genotoxicity, hepatotoxicity, immunotoxicity to name a few. They are assessed via standard, specialised or target organ cell-based assays. Techniques such as WST, MTT, MTS, BrDu and Alamar blue are commonly used to assess basal cytotoxic or direct effect on cell proliferation, whereas Annexin V/Propidium iodide staining can help distinguish between necrotic and apoptotic events. Mitochondrial damage can be assessed by mitochondrial membrane potential assays and luminescent cell viability assays that quantify ATP. Protein marker-based techniques, such as assessing caspase cleavage via flow cytometry or western blotting techniques, can also be used to understand the mode of action of particular compounds. Reactive oxygen species production leads to oxidative stress and this can also lead to cellular damage. There are different dyes, such as fluorescent and bioluminescent dyes, that can be utilised to study this effect. For gauging specialised toxicity effects, different types of

biochemical, molecular and mode of action-based endpoints can be utilised. *In vitro* experimental data when combined with physicochemical properties and Absorption, Distribution, Metabolism and Elimination (ADME) characteristics help establish Physiologically Based Pharmacokinetic (PBPK) and partitioning models (based on fundamental thermodynamic principles). Metabolism of parent compound, toxicity and likelihood of metabolites also allow for a more robust model to be developed as they help to take into account biotransformation and bioavailability. The above information helps to identify the doses and the class of compounds that have to be further tested in *in vivo* tests as specified by OECD guidelines for toxicity testing.

Monoclonal antibodies evoke an effector response mainly via antibody dependent cytotoxicity, phagocytosis and complement dependent cytotoxicity for eliminating tumour target cells (Kindt *et al.*, 2007). For testing the biological activity of mAb based therapeutics *in vitro*, the target cell line is cocultured with the molecule as well as effector cells derived either from PBMCs in human blood or cultured effector cells in a defined target to effector ratio (Golay *et al.*, 2013). These effects can be studied by techniques which involve loading target cells with fluorescent membrane permeable dyes that are released upon target cell lysis. To assess mast cell degranulation, *in vitro* systems are incubated with drug of interest and endpoints like histamine are then measured via spectroscopy or flow cytometry (Demo *et al.*, 1999). Alternatively specific biomarkers like complement fragments can be used to detect specific events such as complement activation (Golay and Introna, 2012). Cytokine release assays provide information about the extent and the kind of pro inflammatory cytokine release. This is often assessed by introducing the monoclonal antibody to human lymphocytes and then assessing the supernatant for different types of cytokines and this assay can often be performed in a multiplex format with flow cytometer analysis (Lash *et al.*, 2006). A cytokine storm is a life threatening adverse effect induced by monoclonal antibodies such as in the case of TGN1412 (Suntharalingam *et al.*, 2006). Animal models utilised for assessing immunotoxicity involve lymph node proliferation assay, local lymph node assay and more recently the mouse drug allergy model though the predictive ability of these *in vivo* models haven't been well characterised or validated (Whritenour *et al.*, 2016). For assessing specialised toxicity assays specific endpoints or biomarkers can be studied. Drug induced liver injury, liver enzyme inhibition or induction (particularly cytochromes 450, flavin monooxygenases and numerous others), change in human pregnane X Receptor activity as well as drug transporter activities for hepatotoxicity; Ames test for mutagenicity, *in vitro* single cell electrophoresis (comet) assay and DNA based dyes for genotoxicity; human ether-related à-gogo gene related (hERG) assays, prolongation of QT

interval, patch clamp assay, embryonic stem cell differentiation assay for cardiotoxicity and so on are examples used in small molecule drug development (Ekins, 2014).

These issues regarding pharmacodynamics, selection of model organism, route of administration, dose, metabolism, toxicity studies have been addressed by the ICH Safety Pharmacology guideline *S6 (R1) Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals*. The safety pharmacology of mAbs however cannot be optimally assessed by standard toxicological assays alone (Guideline, 1997; Cavagnaro, 2002).

### **1.5. In silico tools for predictive toxicology**

Computational toxicology tools could substantially aid in safety pharmacology testing of monoclonal antibody derived therapeutics as they impart elements of automation, consistency and reliability to standard toxicological assays. There are a multitude of advantages offered by computational toxicology methods. They help to realise the 3R principle i.e. Replacement, Reduction and Refinement, by reducing the number of experimental animals used in drug safety testing. They also address the practical and economical concern of industries by providing a rapid and cost effective way for safety testing of novel drug molecules. This in turn helps to cut down attrition rates and thus reduce the financial burden on the discovery and the development of new drugs. Furthermore, computational toxicology methods help to prioritise testing of those compounds which could be associated with toxic hazards by virtue of a problematic chemical space. This could be by means of structural similarity, indiscriminate interaction with closely related pharmacological targets and/or off target effect or other molecular events which are adaptable to *in silico* methods. Computational toxicology methods also prove useful when animal studies do not adequately represent the fate of drugs in humans (Cronin and Madden, 2010; Wilson, 2011; Ekins, 2014; Greene and Pennie, 2015). Though these *in vitro* and *in silico* methods, such as physiologically based pharmacokinetic (PBPK) modelling and qualitative/quantitative structure activity relationships (QSAR), are extensively used for predicting biological activity as well as toxicity during small molecule drug development (Table 1.3), their full potential has not been utilised for biologic drug development.

**Table 1.3** A comprehensive overview of in silico prediction tools for assessing toxicology.

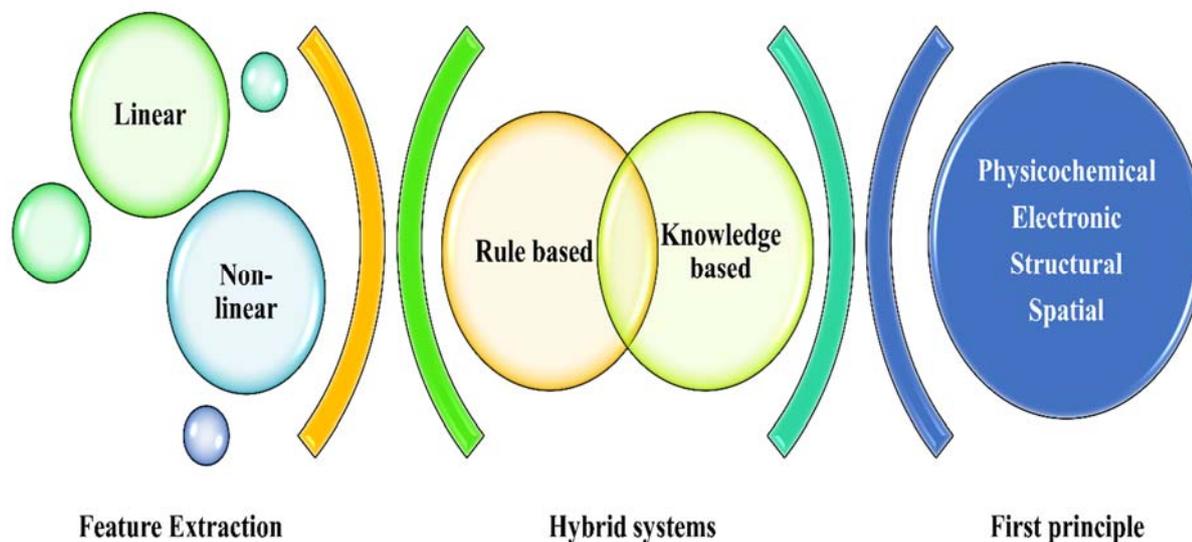
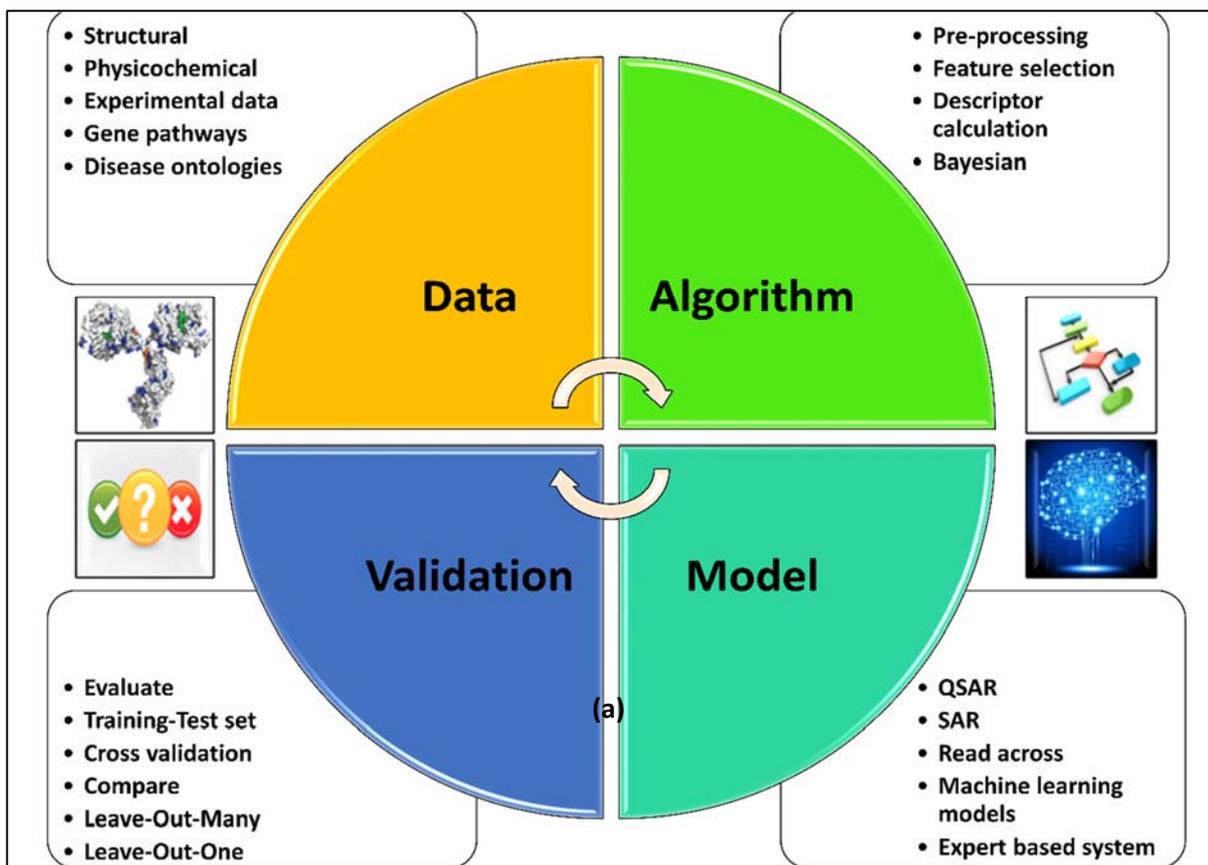
Name	Particulars	Accessibility	Owned by
<b>ACD ToxSuite</b>	Molecular Fragment QSAR and knowledge expert system, (Perceptra platform) employing machine learning a,,h,i,j,k,l,m,r,s,	Commercial	ACD Labs, Pharma Algorithms
<b>Admensa interactive™</b>	QSAR based system <sup>h,k,l</sup>	Commercial	Inpharmatica Ltd
<b>ADMET™ predictor</b>	QSAR based expert system and machine learning <sup>b,c,d,e,f,j,k</sup>	Commercial	Stimulation Plus Inc.
<b>ADMEWORKS Predictor</b>	QSAR,QSPR based expert system <sup>a,b,l</sup>	Commercial	Fujitsu, Poland
<b>AIM</b>	Category formation and read across	Free	US EPA
<b>BfR decision support system</b>	SAR and physicochemical exclusion rule-based system. Employs concordance decision tree approach <sup>d,i,o</sup>	Free	German Federal Institute for Risk Assessment
<b>BioEpisteme</b>	Molecular descriptor QSAR <sup>b,h,k,n</sup>	Commercial	Prous Institute for Biomedical Research, Spain
<b>Bio-loom</b>	QSAR database CLOGP,CMR <sup>h,j</sup>	Commercial	Biobyte
<b>CAESAR</b>	QSAR based expert systems based on Dragon descriptors and Multivariate approaches <sup>a,b,d,e</sup>	Free	EU
<b>CaseUltra (MC4PC)</b>	Molecular fragment QSAR based expert System using machine learning <sup>a,b,c,d,l,j,k</sup>	Commercial	MultiCASE Inc.
<b>Cerius<sup>2</sup>/Material Studio</b>	Molecular modelling software <sup>k,l</sup>	Commercial	Accelrys Inc.
<b>COMPACT</b>	SAR and knowledge based system Employs Molecular orbital descriptors. a,b,c,k	Free	US NTP
<b>CSgenoTOX</b>	QSAR based system and machine learning(ANN) <sup>a</sup>	Commercial	ChemSilico
<b>DEREK NEXUS</b>	SAR Knowledge based expert system a,b,c,d,e,	Commercial	Lhasa Ltd
<b>HazardExpert (ToxAlert)</b>	QSAR Knowledge based expert system a,b,d,e,n,o,p	Commercial	Compudrug Inc.
<b>Insilicofirst</b>	Common User Interface Expert System	Commercial	Lhasa Ltd., Leadscope, Multicase, MN GmbH
<b>KNIME®</b>	QSAR workflow tool	Open	KNIME.com
<b>LAZAR</b>	KNN approach (machine learning) <sup>a,b,k</sup>	Open source	<i>In silico</i> toxicology GmbH
<b>Leadscope Model Applier</b>	QSAR and Expert rule based knowledge system <sup>b,c,e,g,h,k,n</sup>	Commercial	Leadscope Inc.
<b>MDL QSAR</b>	Molecular descriptor QSAR, QSPR, multivariate approaches <sup>a,b,h,j</sup>	Commercial	Symyx - MDL, Inc.
<b>Molcode Toolbox</b>	QSAR based prediction tool <sup>a,b,d,i,j</sup>	Commercial	Molcode Ltd.
<b>OECD QSAR Toolbox</b>	Category formation and read across, QSAR for multiple endpoints	Free	OECD



Relationships, QSPR: Quantitative Structure Property Relationship, QSTR: quantitative Structure toxicity relationship, TOPKAT: Toxicity Prediction by Komputer Assisted Technology, PASS: Prediction of biological Activity Spectra for Substances, CAESAR: computer Assisted, Evaluation, of industrial chemical Substances According to Regulations, T.E.S.T: Toxicity Estimation Software Tool, COMPACT: Computer-optimised Parametric Analysis of Chemical Toxicity, LAZAR: Lazy Structure Activity Relationships, TIMES: Tissue Metabolism Simulator, ADMET: Absorption Distribution Metabolism Excretion Toxicity, MNA: Multilevel Neighbourhood of Atoms, COREPA: Common Reactivity Pattern Approach, ANN: Artificial neural networks

## 1.6. Predictive Model Development

From the different *in silico* tools listed in Table 1.3, a summarised workflow for predictive toxicology model development is depicted in Figure 1.4a. The main question to consider while developing a computational model is what can be modelled that is of use to facilitate existing or new processes i.e. value addition? The starting point of model development is data which can be of different types such as numeric, categorical, discrete or continuous and can be acquired from different sources like experiments, structures, physicochemical properties and so on. Algorithms are then required to pre-process this data as well as for feature extraction. This is mainly for selecting the inputs and outputs of models as well as to convert raw data into parameters that can be modelled mathematically i.e. profilers or descriptors. Different linear and nonlinear mathematical techniques can be used for associating these descriptors to an adverse effect or toxicity by means of statistics, rules, multivariate data analysis and/or expert knowledge thus leading to development of a predictive model as shown in Figure 3b. The resulting model must be validated to ensure non-discriminatory comparison with other existing models. Several factors would have to be taken into consideration while selecting a software platform/tool such as availability, accessibility, user expertise levels, transparency of algorithm and knowledge base, choice and complexity of methodology and inclusion of mechanistic elucidation. Performance would depend on choice of measures for robustness and goodness of fit as well as validation parameters and methods chosen. Some of these aspects are described in detail in the following sections keeping in mind the proteinaceous nature of mAbs.



**Figure 1.4** a) Computational toxicology model development workflow b) Techniques involved in different types of predictive models.

### 1.5.1. Databases

A number of databases have been utilised for developing predictive toxicology models during small molecule development such as Open TG GATEs, Pharmapendium, Drugmatrix® and

ToxFX® (Greene and Pennie, 2015). Databases containing information about mAbs and derived therapeutics are being developed extensively and the IMGT mAb database is particularly noteworthy in this regard as it provides comprehensive information on structure, primary sequences, developmental status, targets as well as documents relating to approval for more than 589 entities (Lefranc *et al.*, 2015). Sources like Drug Bank, patents, FDA documents and UniProt could yield useful information regarding sequences of mAbs whereas Protein Data Bank (PDB) could provide structural information. The choice of a dataset for training model impacts its performance as studies have frequently indicated the discrepancies between public and proprietary datasets i.e. performance of a model developed on public datasets is lower when applied on a proprietary dataset (Greene and Pennie, 2015).

### ***1.5.2. Descriptor Generation and model development:***

Multivariate and statistical data analysis techniques have further allowed for rapid and easier descriptor calculation and model development. For proteins, the primary amino acid sequence and in some cases the 3D structure form the basis of generating different physicochemical, thermodynamic and topographic indices where the physicochemical and structural characteristics of amino acids are utilised to derive descriptors. These include Principal Component Analysis derived descriptors such as z scales and T-scales; 3D structure based ones such as Isotropic Surface Area and Electronic Charge Index; Atomic Charge density derived ones such as Transferable Atomic Equivalent, to name a few (van Westen *et al.*, 2013b). Several machine learning and statistical methodologies, such as support vector machines (SVM), artificial neural networks (ANNs), k-nearest neighbour approach (kNN), decision forest approach, Naïve Bayes, C4.5 decision tree, Bayesian models, random forest approaches, recursive partitioning, multiple linear regression (MLR), discriminant analysis (DA) and self-organising maps (SOM), have been used to predict hepatotoxicity, genotoxicity, cardiotoxicity, and renal toxicity of small molecules (Hardy *et al.*, 2010; Wilson, 2011; Ekins, 2014; Greene and Pennie, 2015). They can be used to build standalone inference-based models or combined with Quantitative Structure activity relationship modelling.

### ***1.5.3. Models***

Quantitative structure activity relationships (QSAR) approach is based on connecting an activity, in particular toxicity (QSTR) or any other property (QSPR), to descriptors which can be derived from physicochemical, structural, electronic or steric parameters (Hansch *et al.*, 1995). QSAR methodology works best when the biological activity in question is based on a single endpoint or a simplistic mechanism of action. The development of QSAR models has

been supported extensively by workflow tools, QSAR databases as well as uniform reporting and summarising formats. Expert/Hybrid systems are extension of QSAR models and they can be based on rules, knowledge or statistics as well as a combination of two or more approaches. The multivariate techniques used can either be linear, such as principal component analysis (PCA) or partial least square regression (PLS) used in TOPKAT, or non-linear techniques, such as ANNs, used in CSgenoTox (Cronin and Madden, 2010). Knowledge based expert systems have incorporated a more mechanistic basis to their predictive tools (Cronin and Madden, 2010). QSAM (Quantitative Sequence Activity Modelling) is another paradigm of QSAR modelling which is being used extensively for protein based predictive models. Angiotensin-converting enzyme- (ACE-) inhibitory peptides were screened based on models generated using PLS, MLR and most recently ANN (Zhou *et al.*, 2008). PLS, SVM and HM based models have been used with smaller peptides (9 amino acids residues) for predicting binding affinity with Class I Major Histocompatibility Complex (Zhao *et al.*, 2007). Proteochemometric modeling is an extension of QSAR that uses Multiplication of Ligand and Protein Descriptors (MLPD) to include interaction space information in addition to protein and ligand descriptors (Qiu *et al.*, 2016).

The advantages of QSAR based expert systems are that they are rapid, well developed and regularly updated. The disadvantages are that the datasets, algorithms and knowledge base are usually not transparent. Most of the tools are commercial and use proprietary datasets. Due to the high level of automation there is a possibility of losing the mechanistic understanding of action.

In addition to the models mentioned above, significant advances have been made with regard to ADME models as understanding the ADME characteristics of molecules is very important in assessing their bioavailability. A target mediate drug disposition based pharmacokinetic model has been developed from preclinical data for predicting pharmacokinetics of mAbs within the human system which could aid in clinical designs (Luu *et al.*, 2012). There have been several machine learning techniques that have been employed in skin absorption and metabolising studies which enable to predict the extent of toxicity caused by compounds (Moore *et al.*, 2014; Ashrafi *et al.*, 2015). It is also worthwhile to mention that the latest techniques seem to revolve around consensus modelling where the outputs from different predictive models are averaged or inferred by several approaches, for example leverage-weighted means (Cronin and Madden, 2010). The success of these models however has been

debatable as some report better predictivity while others report no significant benefits when compared to single models (Hewitt *et al.*, 2007).

#### **1.5.4. Validation**

Models are assessed for specificity, sensitivity and concordance based on either a different dataset typically referred to as the test set or by other appropriate means of validation. Internal validation procedures implemented include cross validation (leave out one and/or leave out many) and bootstrapping. External and independent validation strategies can also be used such as testing the model with new experimental data. The predictive ability can be quantified using different parameters like root mean square error (RMSE), determination coefficient ( $R^2$ ) and predictive squared correlation coefficient ( $Q^2$ ) for QSAR model and these have been evaluated in previous studies (Abshear *et al.*, 2006; Consonni *et al.*, 2009).

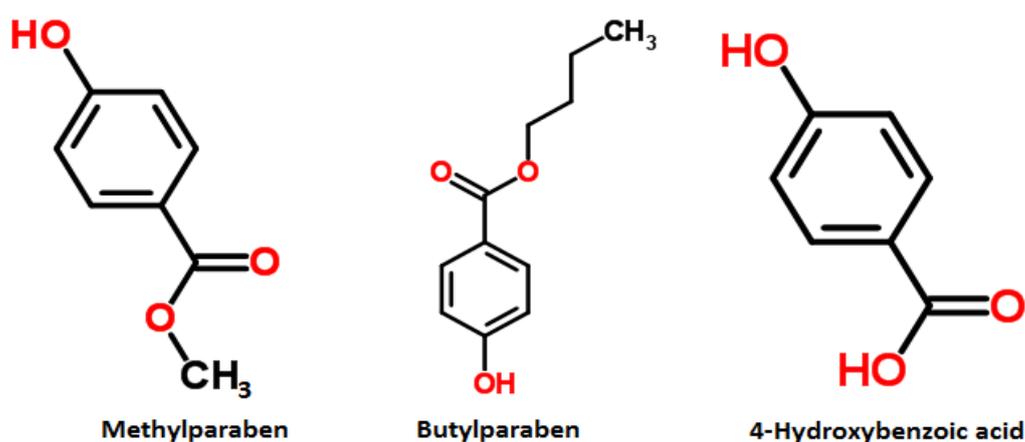
#### **1.6. Scope of the study**

This study focused on developing a QSAR model for early stage screening of monoclonal antibody therapeutics to facilitate rapid developability. Considering that hepatotoxicity and dermal toxicity constitute a major portion of adverse effects elicited by mAbs, the applicability of the experimental methods was first tested on chemical compounds to assess the assay quality following which they were used to assess potential off target adverse effects of mAbs (Chapter 2 and Chapter 3). Furthermore, hypersensitivity reactions were assessed using Skimune™, a non-artificial human skin explants-based assay for safety and efficacy assessment of novel compounds and drugs, developed by Alcyomics Ltd. The suitability of Skimune for assessing the immunogenic adverse effects of aggregated mAbs was studied where aggregation was induced using a heat stress protocol following which the aggregates were characterised and immunogenicity tested using Skimune™ assay (Chapter 4). Identification and creation of data blocks from the primary sequence of mAbs is described in Chapter 5 wherein the influence of intrinsic properties on these descriptors have been described. Multivariate regression models were developed to assess correlations between the biological activities of mAbs, which was expressed in terms of cross reactivity, with structural and primary sequence-based descriptors individually (Chapter 6 and Chapter 7). The models, datasets and responses were benchmarked using various performance metrics. The effect of using a combination of sequence based and structural features was assessed in Chapter 8. Finally, the applicability of the best descriptor set was assessed for developing a QSAR model with Skimune™. Unless otherwise stated, the work reported below was personally carried out by the author of this thesis.

## Chapter 2. Assessment of hepatotoxicity and dermal toxicity of butyl paraben and methyl paraben using HepG2 and HDFn *in vitro* models

As discussed in the previous chapter, *in vitro* toxicity testing panels are prevalent for assessing the different toxicity endpoints of chemical compounds. In this chapter these toxicity tests have been used to assess various toxic endpoints elicited by parabens in hepatocarcinoma (HepG2) and dermal fibroblast (HDFn) cell line (Kizhedath *et al.*, 2018b).

Parabens, esters of parahydroxybenzoic acid, are widely used in cosmetic, food and pharmaceutical industries mainly for their antibacterial and fungicidal properties. Stability over a wide pH range, low cost, broad spectrum activity and low toxicity were some of the reasons for the popularity of paraben as preservatives (Soni *et al.*, 2005). p-Hydroxybenzoic acid is esterified at C-4 position and the paraben series mainly include methyl-, ethyl-, propyl-, butyl-, heptyl- and benzyl-paraben (Figure 2.1). Of these, methyl paraben has shown least toxicity in a wide range of *in vitro* and animal tests (acute and chronic studies) and is the most widely used paraben. Butyl paraben is routinely used as a preservative in some foods, cosmetics, drug formulations and baby products. Humans are also exposed to parabens from the environment (soil, air, biota and water) via inhalation, skin contact and ingestion (Ma *et al.*, 2014). Even though parabens are rapidly hydrolysed by carboxylesterases, biomonitoring of butyl paraben based on their concentrations in urine (free/conjugated) showed presence of the parent compound in 50 – 70% (US), 80% (Denmark) and 36% (Germany) of the test population (Ye *et al.*, 2006; Calafat *et al.*, 2010; Moos *et al.*, 2014).



**Figure 2.1** Chemical structures of methyl paraben(CSID:7176) butyl paraben(CSID:6916) and 4-Hydroxybenzoic acid (CSID:132) (Source ChemSpider, Royal Society of Chemistry).

In vitro studies have shown that the effects of parabens on a number of endpoints increases with an increase in the alkyl chain length (Tavares *et al.*, 2009; Błędzka *et al.*, 2014; Uramaru *et al.*, 2014). The mode of action of parabens is thought to be the disruption of membrane transport and inhibition of mitochondrial function (Porceddu *et al.*, 2012). Propyl paraben associated cytotoxicity was observed in rat primary hepatocytes as well as in *in vitro* studies using HepG2 cells (Nakagawa and Moldéus, 1998; Szeląg *et al.*, 2016). Parabens are deemed to be mildly estrogenic and reproductive toxicity has been reported in several studies (Soni *et al.*, 2005; Tavares *et al.*, 2009). Several *in vitro* and *in vivo* studies have been conducted to ascertain the endocrine disrupting activity of parabens with special emphasis on the presence of methyl paraben being reported in breast cancer tissues (Golden *et al.*, 2005; Darbre and Harvey, 2008). Butyl paraben and isobutyl paraben, are classified as allergens and have been shown to induce male reproductive disorders, male sexual developmental toxicity as well as multiple endocrine disrupting effects (Kang *et al.*, 2002; Oishi, 2002; Uramaru *et al.*, 2014; Boberg *et al.*, 2016; Zhang *et al.*, 2016; Guerra *et al.*, 2017). Methyl paraben was shown to accelerate cellular aging in NHEK cells as well as produce oxidative stress in HaCat cells upon exposure to ultraviolet light-B (Handa *et al.*, 2006; Ishiwatari *et al.*, 2007).

As parabens are approved preservatives in cosmetics there has been considerable interest in the permeation and metabolism of parabens in skin. Multiple studies have been conducted in different *in vitro*, *in vivo* and *ex vivo* systems to assess these factors wherein the permeation ability as well as the permeation flux appeared to be related to the lipophilicity and molecular weight of parabens (Soni *et al.*, 2005; Pedersen *et al.*, 2007; Caon *et al.*, 2010; Moos *et al.*, 2016; Hatami *et al.*, 2017).

The concentration of butyl paraben considered safe, as stated by the European Union, is below 0.19% (w/v) (SSC/1514/13). However, these limits are still subject to scientific scrutiny due to the following factors: lack of human studies; metabolic differences in rats and humans for parabens leading to difficulties in extrapolation; limited information on the systemic availability of free parabens and their metabolites as well as differential dermal absorption and metabolism in *in vivo* studies when compared to humans (Harville *et al.*, 2007; Ulrike Bernauer *et al.*, 2013). The permissible levels of parabens correspond to 5mM in solution. Taking into consideration the increasing exposure of humans to butyl parabens via multiple routes, including the environment, the goal of this study was to measure the harmful effects of exposure to methyl or butyl paraben in an immortalised hepatocyte cell line and human neonatal dermal fibroblasts using an extended *in vitro* toxicity assay panel. These cells lines were considered as their

usefulness in assessing hepatotoxicity and dermal toxicity has been well documented for small molecules. Furthermore, they do not present the antigens required for mAbs binding and therefore could potentially aid in detecting off target effect of mAbs as shown in Chapter 3. The concentrations used in this study have been formulated based on a range below this (highest concentration is 1mM). In vitro studies on human cell lines can contribute to filling in the gap between animal studies and deriving limits for human exposure. The results obtained in this study could supplement existing in vivo toxicity data for defining more robust limits for human exposure.

## **2.1. Materials and Methods**

### **2.1.1. Cell lines and reagents**

HepG2 (ATCC® HB-8065™) and HDFn (ATCC® PCS-201-010™) cells were obtained from Newcastle Central Biobank maintained on the ATCC cultures, Dulbecco's Modified Eagle Medium (DMEM, high glucose, with bicarbonates), Fetal Bovine Serum (FBS), Penicillin-Streptomycin (10,000 units penicillin and 10 mg streptomycin/mL), Phosphate Buffered Saline (PBS), Trypsin EDTA solution, MEM non-essential amino acids, L glutamine solutions 200mM, Sodium Pyruvate Solution, Methyl paraben (CAS No: 99-76-3), butyl paraben (CAS No: 94-26-8), WST-1 (Cat. No: 05015944001), 2', 7'-dichlorofluorescein diacetate (DCFDA; CAS No: 4091-99-0), Dihydrorhodamine 123 (DHR123 CAS No: 109244-58-8), Hydrogen peroxide (CAS No: 7722-84-1) were purchased from Sigma Aldrich, UK. CellTiter-Glo® and GSH-Glo™ Reagent was purchased from Promega, UK and TMRE-Mitochondrial Membrane Potential Assay Kit was purchased from abcam, UK. The 96 well F bottom plates were purchased from GreinerBioOne.

### **2.1.2. Cell culture and maintenance**

HepG2 and HDFn cells were grown as an adherent culture in complete growth media (Dulbecco's Modified Eagle's Media supplemented with 10% Fetal Bovine Serum, 1% Penicillin/Streptomycin, 1% Non-Essential Amino Acids, 1% L-Glutamic acid and 1% Sodium Pyruvate) in T75 tissue culture flasks. The cells were subcultured 3 times a week using the following procedure: the spent media was removed, and the cells were given a Phosphate Buffered Saline wash following which 1x diluted Trypsin was added to gently lift the cells. The cells were then re-suspended in 1:15 dilution in T75 tissue culture flasks. The cells were maintained up to passage 20. Two vials of cells were maintained so as to take into account batch variability.

### ***2.1.3. Cell seeding and treatment***

The HepG2 and HDFn cells were seeded at a density of 10000cells/well onto a Greiner 96-well F bottom tissue culture plate. The cells were then exposed to concentrations ranging between 5-1000 $\mu$ M of the compounds: Methyl paraben or Butyl paraben and further incubated at 37°C 5% CO<sub>2</sub> for 24 hours(h). Each assay contained a control, positive control and solvent control. The control were cells in media, solvent control contained the volume of solvent used for the highest concentrations used and the positive control was dependent on the assay and endpoint measured.

### ***2.1.4. WST-1 cell proliferation Assay***

Following the exposure of cells to the test compounds for 24 h, 10 $\mu$ l of WST-1 reagent was added per well and the plates were incubated for an additional 4 h. Endpoint measurements of absorbance were taken at 480nm and 600nm (background) on FLUOstar® Omega multimode microplate reader. Cell viability was expressed as a percentage of the control. Absorbance was also measured for solvent control.

### ***2.1.5. CellTiter-Glo® Luminescent Cell Viability Assay***

Following the exposure of cells to the test compounds for 24 h, the plate and its contents were equilibrated at room temperature for approximately 30 minutes. Volume of Cell Titer-Glo® Reagent equal to the volume of cell culture medium present in each well (e.g., 100 $\mu$ l of reagent to 100 $\mu$ l of medium containing cells for a 96-well plate) was added. Contents were mixed for 2 minutes on an orbital shaker to induce cell lysis. The plate was allowed to incubate at room temperature for 10 minutes to stabilise luminescent signal. Luminescence was recorded in FLUOStar Omega multiplate reader.

### ***2.1.6. GSH-Glo™ Glutathione Assay***

Following the exposure of cells, seeded in 96 well plates, to the test compounds for 24 h, culture media was removed from the plates. GSH-Glo™ reagent (100 $\mu$ l Luciferin-NT, 100 $\mu$ l of Glutathione S-Transferase and 10 ml of GSH-Glo™ reaction buffer) was added (100 $\mu$ l per well) and mixed briefly on a plate shaker followed by further incubation for 30 minutes at room temperature. Reconstituted Luciferin Detection reagent (100 $\mu$ l) was added to each well of the 96 well plate. The contents of each plate were then mixed and incubated for 15 minutes at room temperature to stabilise luminescent signal. Luminescence was recorded in FLUOStar Omega multiplate reader.

### ***2.1.7. Detection of Reactive oxygen species***

Following the exposure of cells to the test compounds for 24 h, DCFDA was added to all wells at a final concentration of 10 $\mu$ M in well and the plates were incubated for an additional 30 minutes. The media was removed from plates and the wells were washed once with PBS. The cells were resuspended in 200 $\mu$ l of PBS. Endpoint measurements of fluorescence were taken at 480nm (excitation maximum) and 520nm (emission maximum) on FLUOstar  $\text{\textcircled{R}}$  Omega multimode microplate reader. Reactive oxygen species (ROS) levels were expressed as a percentage of the control. Absorbance measurements were also read for solvent control as well as positive control (20 $\mu$ M-5mM final concentration of H<sub>2</sub>O<sub>2</sub>). A similar protocol was used for the Dihydrorhodamine 123 assay reagent.

### ***2.1.8. Cell cycle Analysis***

HepG2 cells were seeded in 24 well plates at a density of 2 x10<sup>6</sup> cells/ml. Following the exposure to the test compounds for 24h, the cells were gently lifted with 1X trypsin and cells washed once with Phosphate buffered saline (PBS). The cells were then fixed in 70% ethanol, added dropwise to avoid clumping. The samples were then washed twice in PBS and treated with 50 $\mu$ l of 100 $\mu$ g/ml ribonuclease followed by 50 $\mu$ l of 50 $\mu$ g/ml Propidium iodide (PI). The samples were then analysed by flow cytometry. Forward and side scatter was measured to identify various cell populations. Pulse processing was used to exclude cell doublets from analysis. PI has an emission maximum of 636 nm and it was measured in the FL2 channel (585/42bp filter) on Caliburs (pulse area vs pulse width). Typical voltage for PI in mammalian cells is around 400V on the Calibur. The data was analysed using FlowJo software where Gaussian curves were fit to obtain G1, S and G2/M % for each cell cycle distribution. The analysis was performed by the Flow Cytometry Core Facility, at Newcastle University.

### ***2.1.9. Mitochondrial Membrane Potential Assay***

Following the exposure of cells to the test compounds for 24 h, the plate and its contents was equilibrated at room temperature for approximately 30 minutes. After removing the media, to eliminate background fluorescence, working solution of tetramethylrhodamine, ethyl ester (TMRE) solution prepared in media (final concentration 200nM) was added to cells. The plates were incubated for 30 minutes with subsequent media removal and replacement with 100 $\mu$ l of PBS. Endpoint measurements of fluorescence were taken at 549nm (excitation maximum) and 575nm (emission maximum) on FLUOstar  $\text{\textcircled{R}}$  Omega multimode microplate reader. Final

concentration of 20 $\mu$ M (carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone) was used as a positive control.

#### **2.1.10. Statistical analysis**

Statistical analysis was carried out using Minitab 17 software. Statistically significant results were reported based on one way Analysis of variance (ANOVA) test followed by post hoc tests (Tukey's/Fishers/Dunnett's) (Ruxton and Beauchamp, 2008) Graphpad Prism V.6.0 (Prism, 2014) was used for curve fitting and computing IC50 values.

## **2.2. Results**

### **2.2.1. Paraben induced cytotoxicity**

HepG2 and HDFn cells were treated with varying concentrations of butyl and methyl parabens. WST-1 is a tetrazolium salt that is converted by mitochondrial dehydrogenase enzymes into a water-soluble coloured formazan compound which is a measure of the metabolic activity of cells (Riss *et al.*, 2016). Cell proliferation was expressed as a percentage of the control. A concentration dependent decrease in cell viability can be observed for HepG2 cells exposed to butyl paraben with a logIC50 of  $2.81 \pm 0.04$  (IC50 643.7 $\mu$ M) (Figure 2.2a). For HDFn cells the logIC50 for butyl paraben was  $2.70 \pm 0.04$  (IC50 502.5 $\mu$ M) (Figure 2.3a). No statistically significant decrease in cell viability was observed in either of the cell lines for the entire concentration range of methyl paraben tested (Figure 2.2a, 2.3a). A significant decrease in cell viability was measured for butyl paraben concentrations >300 $\mu$ M in both cell lines (Figure 2.2b, 2.3b) ( $p < 0.05$ , one-way ANOVA).

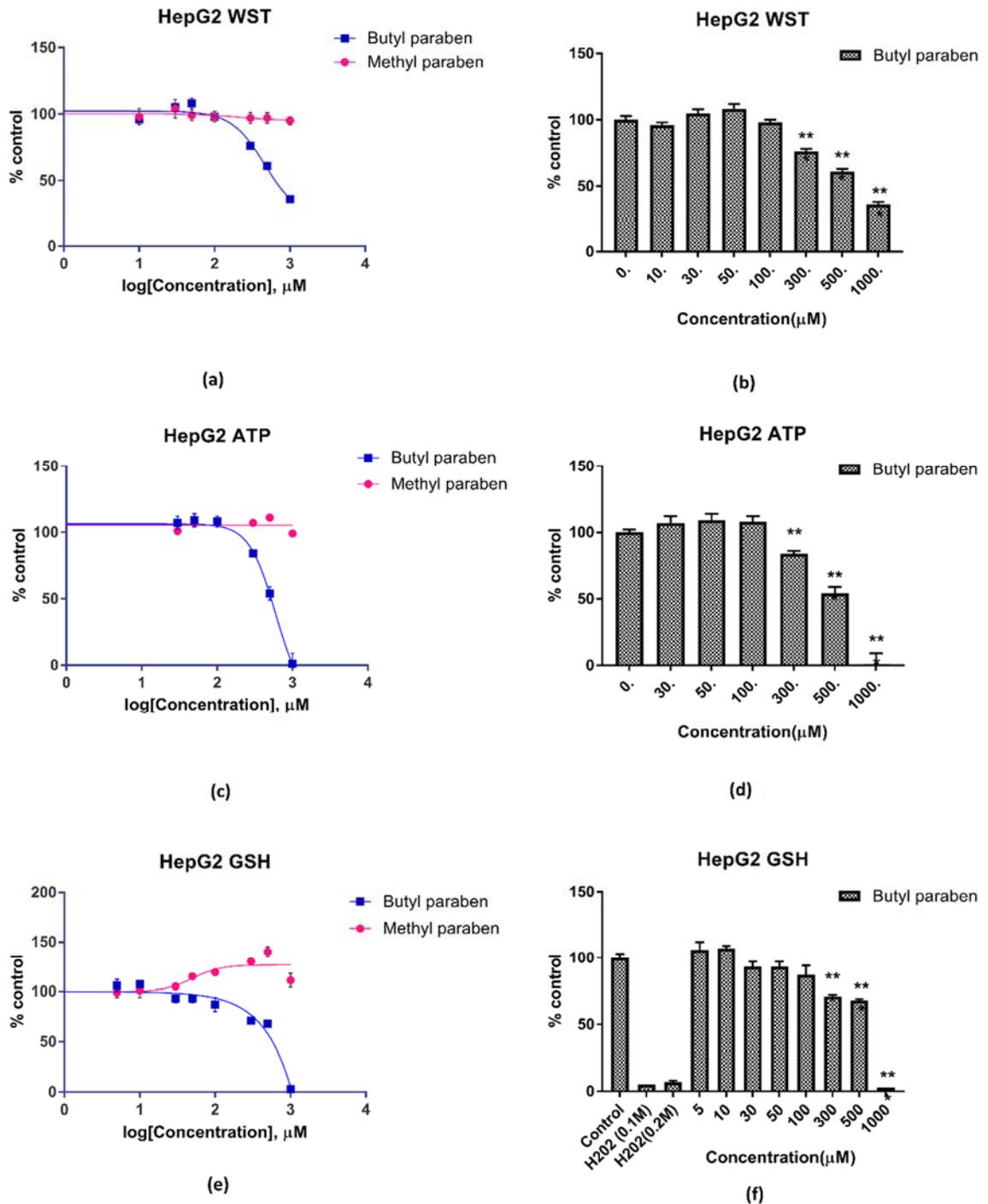
### **2.2.2. Paraben induced reduction in ATP levels**

HepG2 and HDFn cells were treated with varying concentrations of butyl and methyl parabens. Cell Titer-Glo® Luminescent Cell Viability Assay allows for detection of metabolically active cells through the quantification of Adenosine triphosphate (ATP). Luciferin emits light upon interaction with ATP in a reaction catalysed by firefly luciferase and this can be measured by recording the luminescence. For HepG2 cells a concentration dependent decrease in luminescence was measured for butyl paraben with a logIC50 of  $2.69 \pm 0.03$  (IC50 483.7 $\mu$ M) (Figure 2.2c). For HDFn a similar concentration dependent decrease was observed for butyl paraben with a logIC50 of  $2.63 \pm 0.03$  (IC50 425.2 $\mu$ M) (Figure 2.3c). No reduction in ATP levels was measured in either of the cell lines for the entire concentration range of methyl paraben tested (Figure 2.2c, 2.3c). A significant reduction ( $p < 0.05$ , one-way ANOVA) in ATP

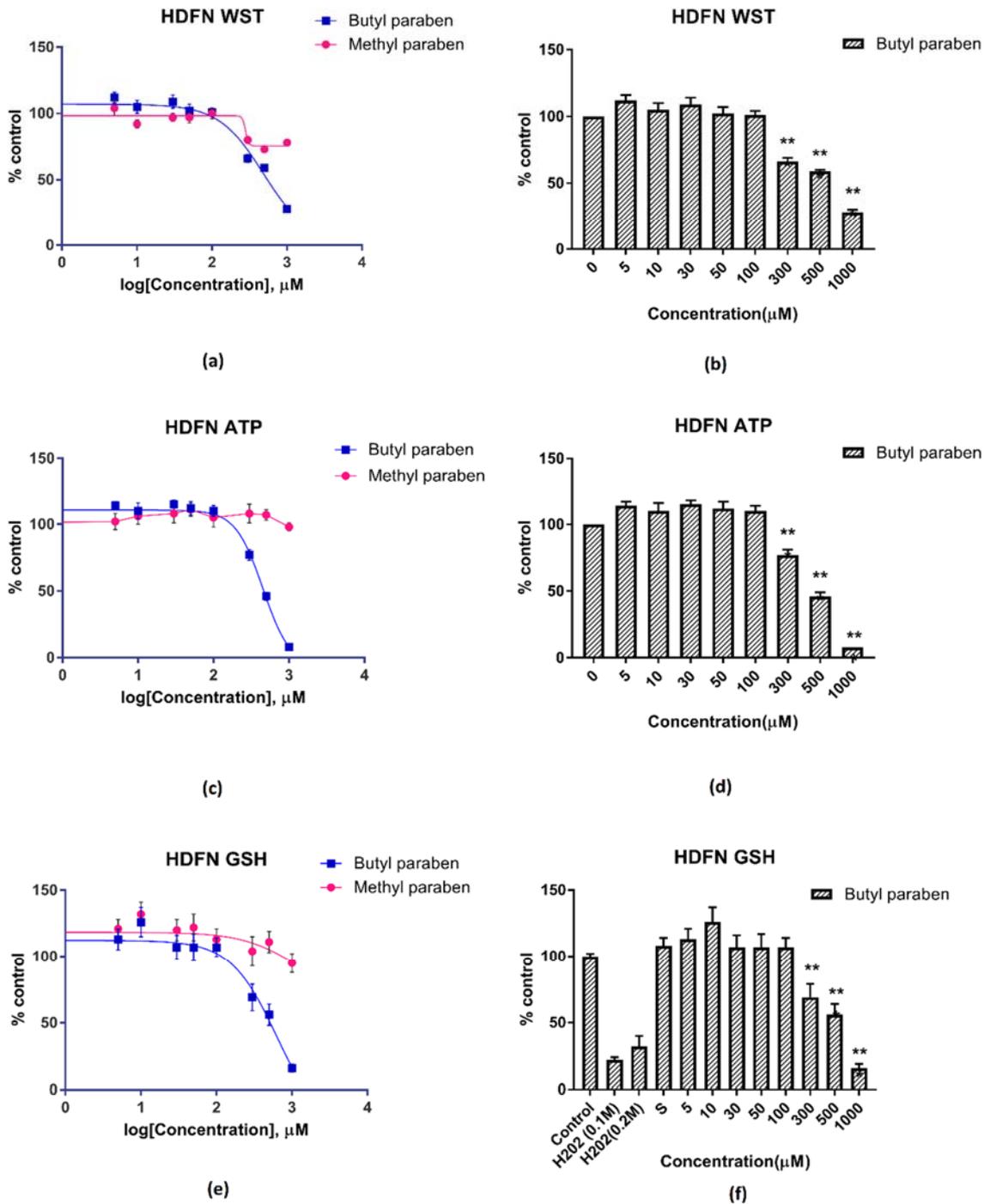
levels was measured for butyl paraben concentrations  $>300\mu\text{M}$  in both cell lines (Figure 2.2d, 2.3d).

### ***2.2.3. Paraben induced reduction in glutathione levels***

HepG2 and HDFn cells were treated with varying concentrations of butyl and methyl parabens. GSH-Glo™ assay allows for luminescence-based detection and quantification of glutathione in cells. Glutathione transferase catalyses the conversion of a luciferin derivative into luciferin when GSH is present and this luciferin is detected using Ultra-Glo™ recombinant luciferase. The luminescence generated is proportional to the glutathione levels present in cells. As in the previous tests, for HepG2 cells a concentration dependent decrease in GSH levels was observed for butyl paraben with a  $\log\text{IC}_{50}$  of  $2.65 \pm 0.15$  ( $\text{IC}_{50}$   $448.9\mu\text{M}$ ) (Figure 2.2e). For HDFn a similar concentration dependent decrease in GSH levels was observed for butyl paraben with a  $\log\text{IC}_{50}$  of  $2.64 \pm 0.06$  ( $\text{IC}_{50}$   $438.1\mu\text{M}$ ) (Figure 2.3e). Again, no reduction in GSH levels was observed in either of the cell lines for the concentration range of methyl paraben tested (Figure 2.2e, 2.3e). Results were considered to be significant at  $p < 0.05$  based on one-way ANOVA test. A significant decrease in GSH levels can be observed for butyl paraben concentrations  $>300\mu\text{M}$  in both cell lines (Figure 2.2f, 2.3f).



**Figure 2.2** Results of the different in vitro tests performed on HepG2 cell after treatment with Butyl paraben and Methyl paraben for 24 h. Cell viability assay using WST-1 reagent (a) Dose response curve (b) concentration vs percentage of the control response. ATP reduction assay using CellTiter-Glo® (c) Dose response curve (d) concentration vs percentage of the control response. GSH reduction using GSH-Glo™ (e) Dose response curve (f) concentration vs percentage of the control response. All values are expressed as a percentage of the control and mean  $\pm$  standard error of the mean (SEM) for  $n=3$ . Statistical differences were assessed with one way ANOVA. Dunnett's post hoc test was used for multiple comparisons to control. \*\* indicates  $p<0.05$ , \*\*\* indicates  $p<0.001$



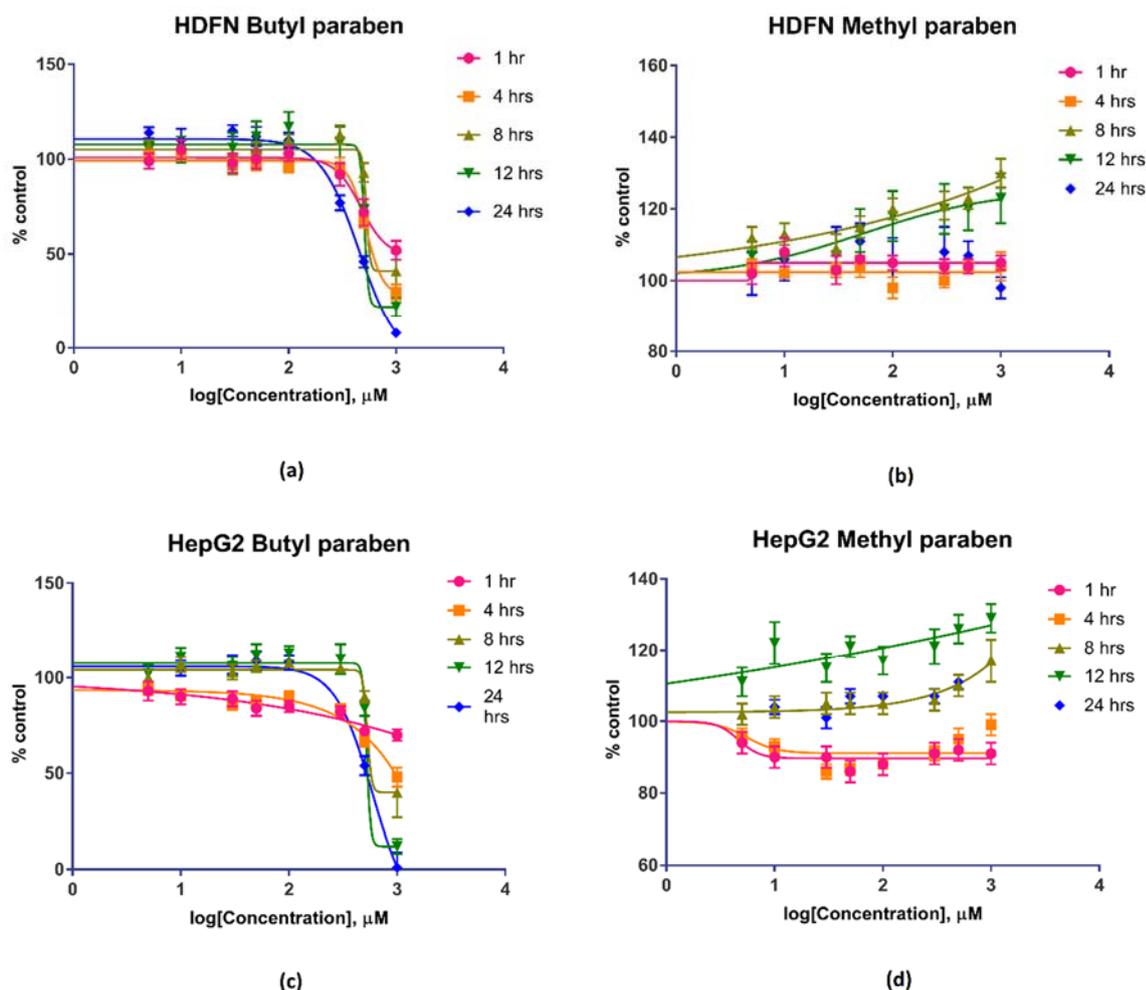
**Figure 2.3** Results of the different in vitro tests performed on HDFN cell after treatment with Butyl paraben and Methyl paraben. Cell viability assay using WST-1 reagent (a) Dose response curve (b) concentration vs percentage of control response. ATP reduction assay using CellTiter-Glo® (c) Dose response curve (d) concentration vs percentage of control response. GSH reduction using GSH-Glo™ (e) Dose response curve (f) concentration vs percentage of control response. All values are expressed as percentage of control and mean  $\pm$  standard error of the mean (SEM) for  $n=3$ . Statistical differences were assessed with one way ANOVA. Dunnett's post hoc test was used for multiple comparisons to control. \*\* indicates  $p<0.05$ , \*\*\* indicates  $p<0.001$ .

#### 2.2.4. Paraben induced time dependent decrease in ATP levels

Concentration and time dependent decrease of ATP was observed for butyl paraben in both cell lines (Figure 2.4). The percentage of ATP was measured at time intervals of 1, 4, 8, 12 and 24 h and this was used to derive logIC<sub>50</sub> and IC<sub>50</sub> values as shown in Table 2.1. In HepG2 cells a 50% in ATP levels and in HDFn cells approximately 97% decrease in ATP levels was observed based on the IC<sub>50</sub> values calculated at 1h and 24h (Figure 2.4c, 2.4a). Additionally, the time dependent decrease in ATP levels was faster for HDFn when compared to HepG2. As seen in Figure 4a, ATP levels fall to about 50% of control at 1000  $\mu$ M with HDFn after 1 h incubation whereas the decrease in HepG2 cells after 1 h is much lower (only to about 70%) as seen in Figure 4c. Similarly, ATP levels fall to 30% of control at 4 h in HDFn cells compared to the 50% of control in HepG2 cells (Figure 2.4a and 2.4c). This indicates that the HDFn cell line is more sensitive to butyl paraben. Methyl paraben did not show any decrease in ATP levels in response to concentration or time in either cell line. Indeed, ATP levels increased with methyl paraben concentrations with respect to the control for exposure durations of 4 h and 8 h (Figure 2.4b, 2.4d).

**Table 2.1** Time dependent butyl paraben induced decrease in ATP levels expressed in term of logIC<sub>50</sub>, IC<sub>50</sub> and standard error (SE) in HepG2 and HDFn cells.

Duration of exposure (hrs)	HepG2			HDFn		
	logIC <sub>50</sub>	SE	IC <sub>50</sub> ( $\mu$ M)	logIC <sub>50</sub>	SE	IC <sub>50</sub> ( $\mu$ M)
1	2.99	$\pm$ 0.04	991.5	4.16	$\pm$ 0.26	14517
4	2.85	$\pm$ 0.02	708.6	3.02	$\pm$ 0.06	1044
8	2.95	$\pm$ 0.03	891.1	2.94	$\pm$ 0.02	870.2
12	2.81	$\pm$ 0.03	647.7	2.81	$\pm$ 0.02	649.9
24	2.63	$\pm$ 0.03	425.2	2.68	$\pm$ 0.02	484.1

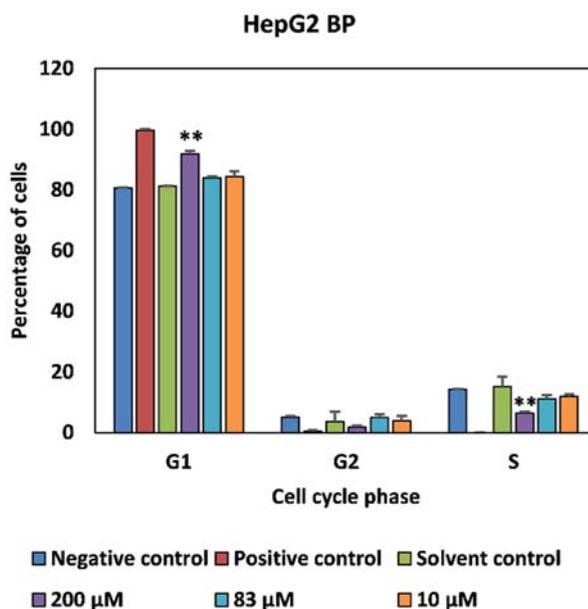


**Figure 2.4** Dose response curves of the reduction in ATP levels measured over time after the exposure of cells to different concentrations of parabens for 1, 4, 8, 12 and 24 h using CellTiter-Glo®. Dose response curves for HDFN after treatment with (a) Butyl Paraben (b) Methyl Paraben. Dose response curves for HepG2 after treatment with (c) Butyl Paraben (d) Methyl Paraben. All values are expressed as percentage of control and mean  $\pm$  standard error of the mean (SEM) for n=3.

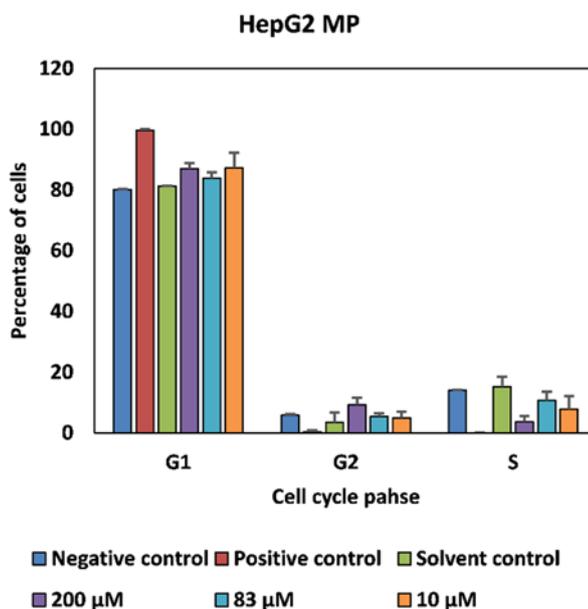
### 2.2.5. Cell cycle analysis to assess genotoxicity of parabens

Propidium iodide, an intercalating DNA dye, is widely used for quantification of total DNA content by flow cytometry analysis (Krishan, 1975). Cell cycle arrest compared to the control could be indicative of DNA damage and thus potential genotoxicity of compounds. Compared to the control there was a small increase in HepG2 G1 cell population as well as a decrease in G2 and S for the highest concentration of Butyl paraben i.e. 200 $\mu$ M (Figure 2.5a). An increase in the cell population in G1 and G2 phases, as well as a concomitant decrease in the proportion of cells in S phase, was measured for methyl paraben when compared to the control cells for the highest concentration of methyl paraben i.e. 200 $\mu$ M (Figure 2.5b). This cell cycle arrest in G1 phase was also measured with the positive control (50mM H<sub>2</sub>O<sub>2</sub>). For HDFn, due to the

sensitivity of these cells the protocol used here was perhaps not best suited (freeze thaw cycle) and there the results are not of the highest quality owing to methodological issues.

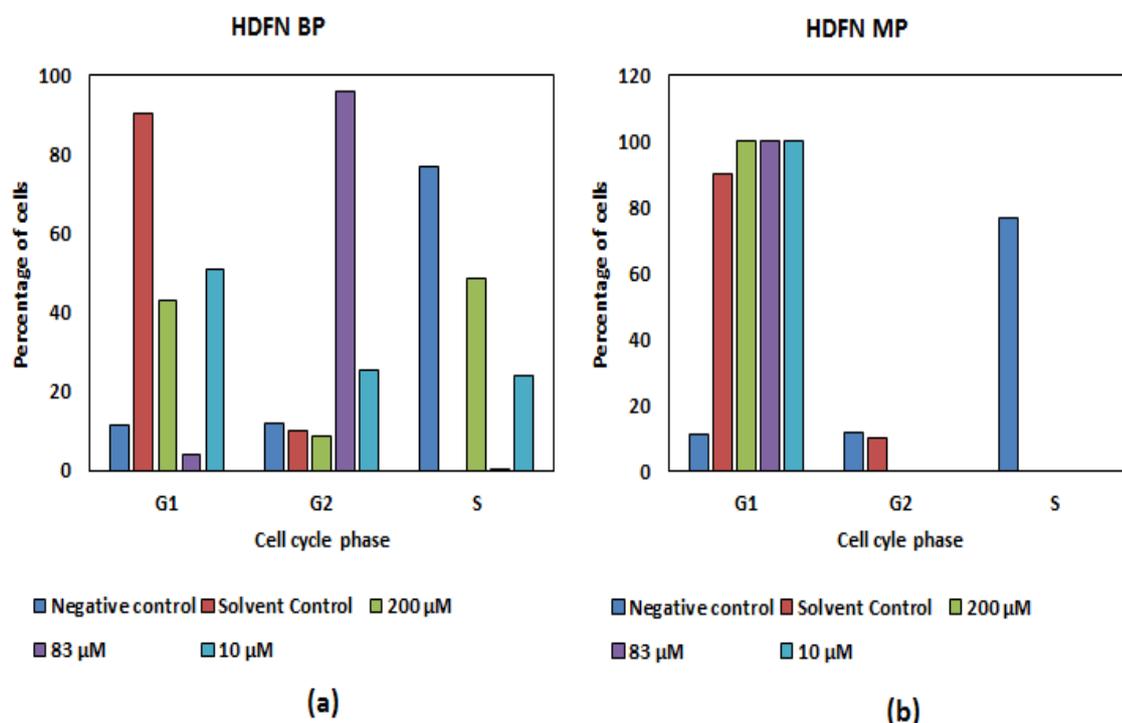


(a)



(b)

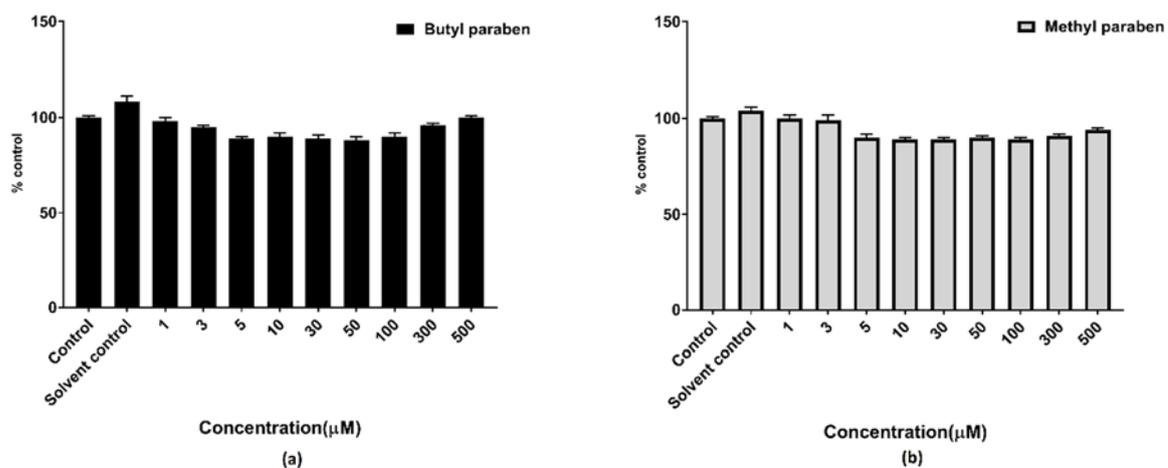
**Figure 2.5** Graphical representation of cell cycle phase proportions following cell fixing, Propidium Iodide staining and analysis by flow cytometry for (a) HepG2 cells treated with butyl paraben (BP) (b) HepG2 cells treated with methyl paraben (MP). All values are expressed as a percentage of the control and mean  $\pm$  standard error of the mean (SEM) for  $n=3$ . Statistical differences were assessed with one way ANOVA. Dunnett's post hoc test was used for multiple comparisons to control. \*\* indicates  $p<0.05$ .



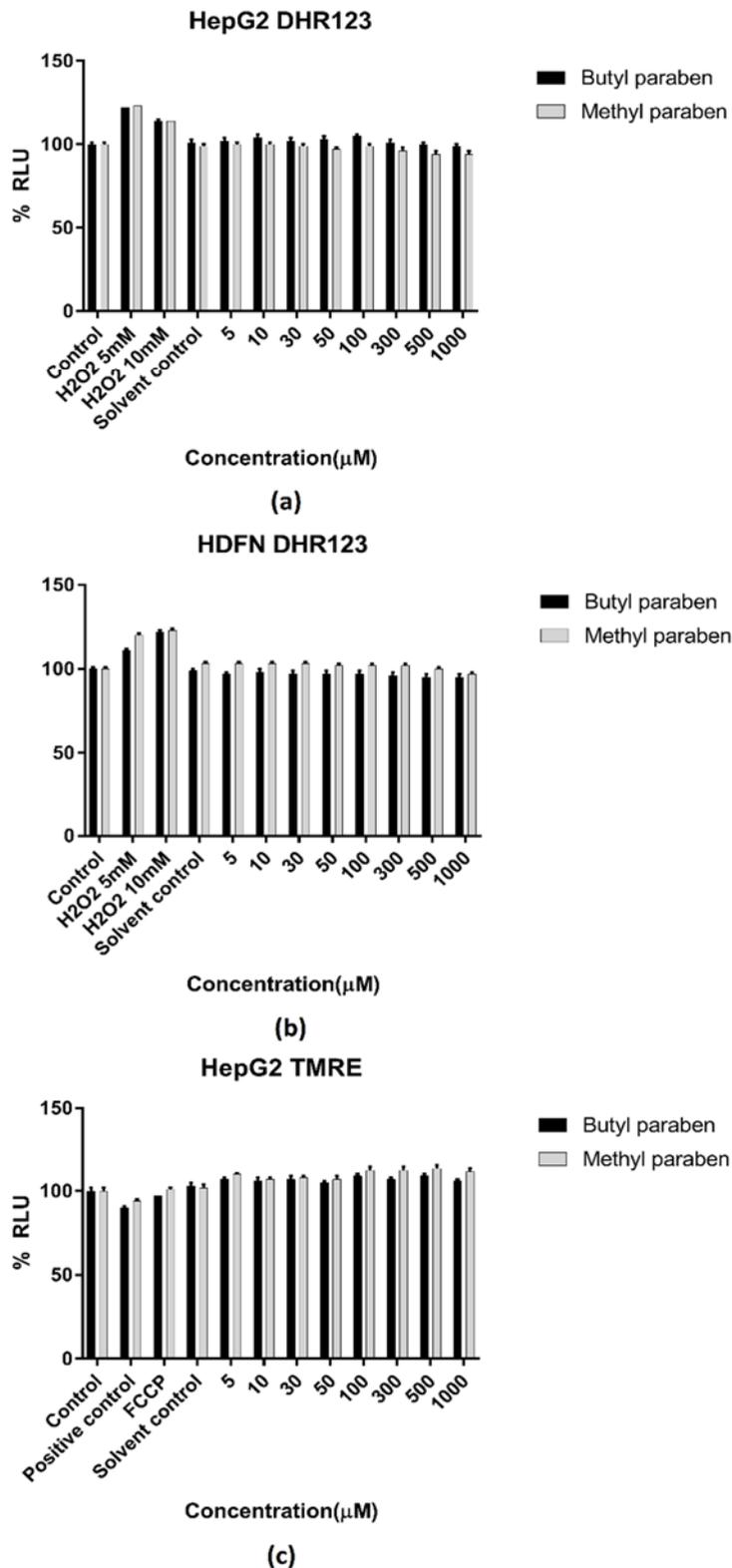
**Figure 2.6** Graphical representation of cell cycle phase proportions following cell fixing, Propidium Iodide staining and analysis by flow cytometry for (a) HDFn cells treated with butyl paraben (b) HDFn cells treated with methyl paraben.

#### 2.2.6. Paraben induced Oxidative stress and mitochondrial dysfunction

HepG2 cells and HDFn cells were treated with varying concentrations of butyl or methyl paraben. DCFDA is a cell permeant fluorogenic dye that, upon deacetylation by cellular esterases, can be converted to a fluorescent compound in the presence of reactive oxygen species (ROS). The higher the fluorescence detected, the greater the amount of free radicals and thus higher oxidative stress (Liu *et al.*, 2014a). The positive control (hydrogen peroxide, 5mM at 24h) showed a high level of fluorescence compared to the control, however no concentration dependent increase in ROS was observed for either of the parabens (Figure 2.7a,2.7b). Similar results were obtained using DHR123, a fluorogenic ROS indicator, for both HepG2 and HDFn cells exposed to varying concentrations of parabens (Possel *et al.*, 1997). No concentration dependent increase in fluorescence signal was observed for either of the parabens but a high level was observed for the positive control (Figure 2.8a, 2.8b). No changes in mitochondrial membrane potential were observed in HepG2 cells for either of the parabens over the concentration range 0 to 1000 μM. (Figure 2.8c).



**Figure 2.7** Results of the DCFDA assay to measure ROS production in HepG2 cells after treatments with (a) butyl paraben and (b) methyl paraben. All values are expressed as percentage of control and mean  $\pm$  standard error of the mean (SEM) for n=3.



**Figure 2.8** Results of the DHR 123 assay to measure oxidative stress in (a) HepG2 and (b) HDFN cell lines after treatment with Butyl paraben and Methyl paraben. Results of the TMRE assay to measure mitochondrial membrane potential in (c) HepG2 cells after treatments with parabens. All values are expressed as percentage of control and mean  $\pm$  standard error of the mean (SEM) for n=3.

### 2.3. Discussion

Parabens are widely used in pharmaceutical, cosmetic and food sector applications as preservatives. However, they have been shown to induce hepatotoxic and dermal toxic effects in several studies (Nakagawa and Moore, 1999; Harville *et al.*, 2007; Shah and Verma, 2011; Porceddu *et al.*, 2012).

In our studies, butyl paraben was observed to cause a significant concentration dependent decrease in cell viability for HepG2 cells (IC<sub>50</sub> 643.7 $\mu$ M when incubated with cells for 24 h. Besides being in agreement with previous *in vivo* studies, we have also demonstrated cytotoxicity at concentrations many times lower than those reported in primary rat hepatocytes and mice (Nakagawa and Moldéus, 1998; Shah and Verma, 2011). The toxicity of butyl paraben observed appears to be accompanied by a significant dose dependent reduction in ATP (IC<sub>50</sub> 483.7 $\mu$ M). ATP is a sensitive marker for cell viability since as the cells lose membrane integrity, they fail to synthesise ATP and any remaining ATP in the cytoplasm is rapidly depleted by ATPases (Riss *et al.*, 2016). This can also be reflective of mitochondrial dysfunction. This has been previously reported in HepG2 cells on exposure to propyl paraben as well as mice and rat primary hepatocytes upon exposure to butyl paraben (Nakagawa and Moldéus, 1998; Shah and Verma, 2011; Szeląg *et al.*, 2016). A significant dose dependent depletion of glutathione (GSH) levels was also observed for higher concentrations (IC<sub>50</sub> 448.9 $\mu$ M) of butyl paraben which may indicate an oxidative stress mechanism via GSH depletion. Dose dependent oxidative stress mediated hepatotoxicity has been observed in *in vivo* studies upon oral exposure to butyl paraben where significant dose dependent increase in lipid peroxidation (a consequence of oxidative stress) and decrease in glutathione levels were observed for higher concentrations (1000mM) (Shah and Verma, 2011).

The decrease in ATP content of cells in response to butyl paraben was found to be time-sensitive as well as concentration dependent. A significant time and dose dependent decrease in ATP was also observed in HepG2 cells with 50% decrease in IC<sub>50</sub> values based on ATP levels over a 24 h period. Butyl paraben was observed to cause a significant dose dependent cytotoxicity (IC<sub>50</sub> 502.5 $\mu$ M) in HDFn cells which was accompanied by a corresponding significant dose dependent decrease in ATP levels (IC<sub>50</sub> 425.2 $\mu$ M) as well as GSH levels (IC<sub>50</sub> 438.1 $\mu$ M). A significant time and dose dependent decrease in ATP was also observed for HDFn cells with 97% decrease in cell viability over a 24-hr period. At shorter exposure times, HDFn cells appeared to be more sensitive to butyl paraben than HepG2 cells. After 1 h exposure, HDFn cells showed almost complete depletion of cellular ATP compared to controls, whilst HepG2

cells showed only a 50% reduction in ATP content. This may be clinically significant, as liver cells are more likely to be exposed to the metabolite than the parent compound (either as a result of hydrolysis during absorption in the GI tract, in the plasma or by hepatic metabolism, whilst fibroblasts are more likely to be exposed to the parent compound. Cutaneous tissues are well understood to possess lower activities of hydrolytic enzymes than hepatic tissues.

Methyl paraben did not show any significant decrease in cell viability, reduction in ATP or glutathione levels in HepG2 and HDFn cell lines at the concentrations tested. Cell proliferation was observed for higher concentration of methyl paraben in both cell lines. IC50 values could not be derived for methyl paraben as there was no apparent concentration dependent toxicity at the concentrations tested. Low concentrations of methyl paraben (200µM) showed accelerated cell aging in skin keratinocytes, however the incubation time was 32 days (Ishiwatari *et al.*, 2007).

There appeared to be no ROS mediated effect of either methyl or butyl paraben on cell viability as demonstrated by the DCFDA assay and DHR123 assay, a finding supported by research in previous studies (Soni *et al.*, 2005). Some previous studies have reported ROS production as well as mitochondrial damage in response to paraben exposure, although the concentration range tested was far higher in those reports (Nakagawa and Moldéus, 1998; Shah and Verma, 2011) than those tested in the present study. The potential to induce hepatotoxicity in isolated mouse liver mitochondria via damage to mitochondrial respiration has been reported for parabens (Porceddu *et al.*, 2012). Butyl paraben induced mitochondrial permeation transition-mediated mitochondrial swelling was reported in the presence of 50µM calcium ions in isolated rat hepatocytes at a concentration of 0.25mM (Nakagawa and Moore, 1999). In the present study the TMRE assay was not able to detect any changes in mitochondrial membrane permeability for the highest concentration of parabens tested. This could be due to a lack of ions required to facilitate detection of increased membrane permeation via increased ion influx.

Cell cycle analysis using propidium iodide staining methods has been used to study cell cycle arrest at different phases to provide an indirect measure of genotoxicity (Krishan, 1975; Esmaelian *et al.*, 2013). A dose dependent cell cycle arrest in the G1 phase and decrease in S phase was observed for HepG2 cells exposed to the highest non-cytotoxic concentration of butyl paraben (200µM). These results are comparable to a previous study which showed a significant dose dependent decrease in mitotic cells as well as cycle arrest in G1 phase observed in Vero cells exposed to 500µM propyl paraben (Martín *et al.*, 2010). Chromosomal aberrations and an increase in polyploid cells were observed in an *in vitro* genotoxicity assays of butyl

paraben at a dose of 308  $\mu\text{M}$  (Ishidate Jr *et al.*, 1978). Further analysis such as gene mutation in mammalian cells systems,  $\gamma\text{-H2AX}$  assay (test for DNA damage) as well as chromosomal aberrations tests would be required to provide clarity on the mechanism of potential genotoxicity (Knight *et al.*, 2009; Ivashkevich *et al.*, 2012).

Considering the two main routes for paraben exposure, oral and/or topical, two main aspects must be taken into consideration while interpreting and setting toxicity limits: hydrolysis and permeation, which is in turn influence by the alkyl chain length of parabens. Permeation of parabens into skin layers upon topical application has been demonstrated by different studies and this is influenced by the lipophilicity and molecular weight of parabens in different *in vivo* and *ex vivo* models (Pedersen *et al.*, 2007; Caon *et al.*, 2010). Previous reports have shown that parabens with increased chain length exhibit increased potency (Prusakiewicz *et al.*, 2007) and this could be the reason for the toxicity of butyl paraben in this study while methyl paraben showed no effects. Elimination rate via urine decreases with increasing length of alkyl chain as this increases the lipophilicity of paraben as reflected in previous studies (Moos *et al.*, 2016). Butyl parabens have shown rapid biotransformation in terms of glucuronidation and hydrolysis when compared to methyl paraben in human liver microsomes and plasma respectively, further studies have to be conducted to distinguish the effect of the butyl or methyl parabens from its primary metabolite which is p-hydroxybenzoic acid (Abbas *et al.*, 2010).

Human exposure to parabens has become extensive due to their presence in consumer products as well as via environmental routes. In conclusion our results have demonstrated the potential hepatotoxic and dermal toxic effect of butyl paraben which appears to be accompanied by ATP and GSH depletion reflective of mitochondrial dysfunction and oxidative stress respectively using HepG2 and HDFn as *in vitro* models. Our findings also indicate a time dependent decrease in ATP levels which provides a useful basis for further toxicokinetic studies. An insight has also been provided into the potential cell cycle arresting tendency of butyl paraben. Due to a lack of data from human studies as well as difference in pharmacokinetic/pharmacodynamics behaviour of animal model studies when compared to humans, a scientifically sound conclusion cannot be reached by regulatory authorities on the safe usable limits of parabens. Whilst additional factors such as absorption and metabolism must be taken into account, it has been established that methyl paraben and butyl paraben are extensively and rapidly absorbed through the skin and that metabolism of parabens in cutaneous tissues is not extensive. Whilst the resemblance of HepG2 cells to human liver cells may be limited, the dermal fibroblasts used in this study were indeed primary human cells. Given the

accumulated evidence for the effects of parabens (and the potential interaction with other environmental stressors) such as UV radiation, the case could be made to re-examine permissible limits. These *in vitro* studies could form the preliminary step at bridging the gap between *in vivo* data and extrapolation to set safe use limits in pharmaceutical, cosmetics and food products.

#### 2.4. Chapter Summary

In the present study the effects of exposure to methyl or butyl paraben (5-1000 $\mu$ M) on cytotoxicity, oxidative stress, mitochondrial dysfunction and genotoxicity were measured in a hepatocarcinoma cell line (HepG2) and human dermal fibroblasts neonatal (HDFn). Butyl paraben caused a concentration dependent decrease (above 400 $\mu$ M) in cell viability for both cell lines. Toxicity of butyl paraben observed appeared to be mediated via ATP depletion as seen from luminescence assays. Depletion of glutathione was also observed for higher concentrations of butyl paraben, which may indicate the involvement of oxidative stress. Methyl paraben, however, did not show any significant decrease in cell viability, reduction in ATP or glutathione levels in HepG2 and HDFn cell lines at the concentrations tested. *In vitro* studies based on human cell lines can provide information in the early stages of multitier paraben toxicity studies and can be combined with *in vivo* and *ex vivo* studies to build more comprehensive, scientifically sound strategies for paraben safety testing.

As evident from the results obtained from the above study, *in vitro* toxicity assays were able to assess the potential hepatotoxicity and dermal toxicity of parabens. Complexity of mAbs compared to small molecules, co incubation with effectors cells (PBMCs), exposure times and donor variability are factors that could influence the ability of traditional toxicity assays to assess adverse effects of mAbs. The next chapter discusses the applicability of some of these assays to detect any potential off target effects of mAbs.

### **Chapter 3 Applicability of traditional and novel in vitro toxicity tests for assessing adverse effects of monoclonal antibodies: A case study of rituximab and trastuzumab.**

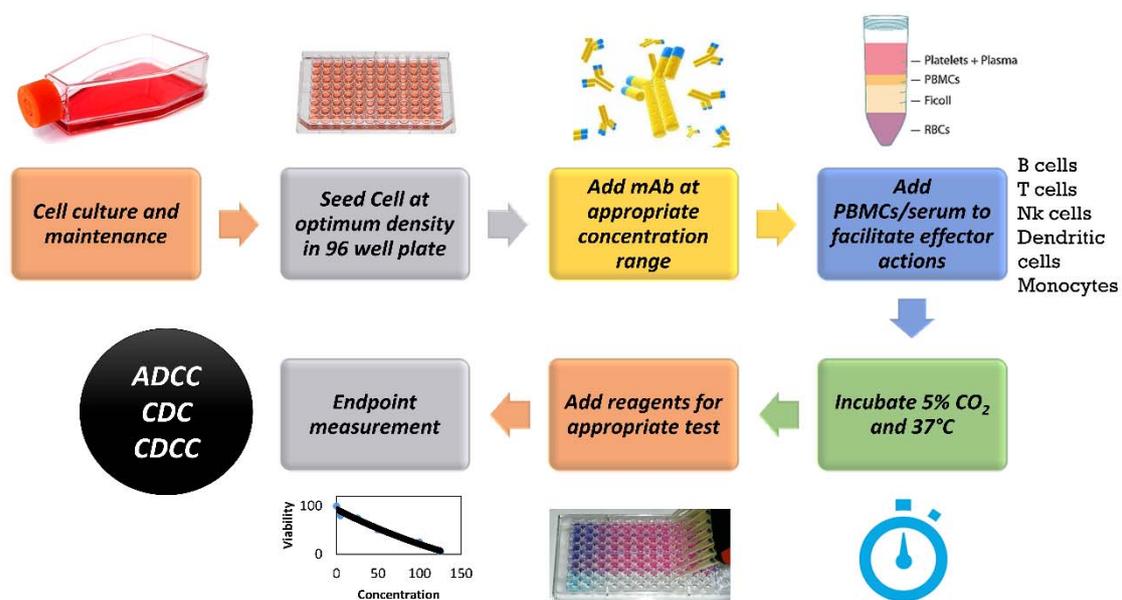
Previous studies indicate that following immunogenicity, hepatotoxicity and dermal toxicity are the two main adverse effect categories associated with mAbs and this has been outlined in Chapter 1 Table 1.1. As outlined in the previous chapter, different in vitro tests were used to assess the hepatotoxicity and dermal toxicity of parabens wherein the methodology and endpoints were discussed in detail. This chapter aims to assess the suitability of those traditional toxicity assays to investigate potential organ/system related adverse effects of mAbs that could lead to hepatotoxicity and dermal toxicity (as outlined in Table 1.1 and Box 1.1. in Chapter 1) using hepatocarcinoma cell line (HepG2) and human dermal fibroblasts neonatal (HDFn), respectively. These cells lines were chosen also based on them not expressing the antigens necessary for either of the two mAbs chosen in this study. This would thus allow for detection of any potential of target effect. Furthermore, the potential immunogenicity of mAbs was assessed using the novel Skimune™ skin explant assay.

The mAbs used in this case study are rituximab and trastuzumab both of which were commercially purchased. Rituximab is an AntiCD20 chimeric monoclonal antibody, with an IgG1 heavy chain and kappa light chain, used for therapeutic indications such as Non-Hodgkin's lymphoma, chronic lymphocytic leukaemia, rheumatoid arthritis, granulomatosis with polyangiitis and microscopic polyangiitis. It has been shown to elicit hepatobiliary and skin related adverse effects in addition to immunogenicity and hypersensitivity (*MabThera, INN-rituximab - European Medicines Agency - Europa EU*). Trastuzumab is an AntiHER2 humanised monoclonal antibody, with an IgG1 heavy chain and kappa lights chain, indicated for breast cancer and gastric cancer. It has shown to cause infusion related reactions, hypersensitivity as well as hepatobiliary and skin related adverse effects (*Herceptin - European Medicines Agency - Europa EU*).

Effector functions of mAbs such as phagocytosis, antibody dependent cytotoxicity (ADCC), complement dependent cytotoxicity (CDC) via complement activation, complement dependent cellular cytotoxicity (CDCC) as well as evoking other cell mediated immune responses, are modulated via the Fc region by interaction with FcγR receptors on different immune responsive cells (Table 1.2)(Carter, 2006; Meyer *et al.*, 2014). This requires the co culture of target cells with immune responsive cells, such as Peripheral Blood Mononuclear Cells(PBMCs) , to elicit the immune response pre requisite for causing adverse effects that could lead to off-target

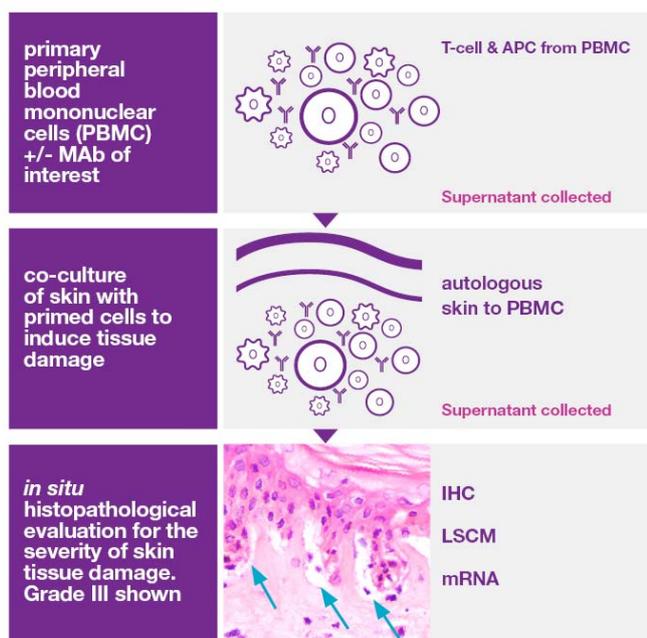
toxicity. PBMCs comprise of B cells, T cells, monocytes, dendritic cells and natural killer cells and these cells express various Fc $\gamma$ R receptors which bind to the Fc region of IgG mAbs and induce effector functions like ADCC, phagocytosis, transport and catabolism (Vidarsson *et al.*, 2014b). For complement activation and CDC, exposure to human serum containing complement proteins is required (Bologna *et al.*, 2011). A combination of PBMCs and complement protein is used to assess potential CDCC where a primary binding to complement protein is followed by engaging with complement receptors on natural killer cells or macrophages (Meyer *et al.*, 2014).

Different in vitro toxicity assays are in place for assessing toxicity endpoints as described in Chapter 1 Section 1.1.3. Of these assays, WST1 assay is routinely used for assessing cytotoxicity of compounds. WST-1 is a tetrazolium salt that is converted by mitochondrial dehydrogenase enzymes into a soluble coloured formazan compound which can be quantified using absorbance endpoint measured using a spectrophotometer. The absorbance values are reflective of mitochondrial enzyme activity which is a measure of the metabolic activity of cells. Another sensitive marker for cell viability is by measuring the adenosine tri phosphate (ATP). As the cells lose membrane integrity, they fail to synthesise ATP and any remaining ATP in the cytoplasm is rapidly depleted by ATPases (Riss *et al.*, 2016). Cell Titer-Glo<sup>®</sup> Luminescent Cell Viability Assay allows for detection of metabolically active cells through quantification of Adenosine triphosphate (ATP). Luciferin upon interaction with ATP emits light in a reaction catalysed by firefly luciferase and this can be measured by recording the luminescence. Potential cytotoxicity and reduction in ATP levels of HepG2 and HDFn cells following exposure to mAbs were investigated using the methodology described in Figure 3.1.



**Figure 3.1** General Methodology for an *in vitro* assay to detect toxicity of mAb based therapeutics.

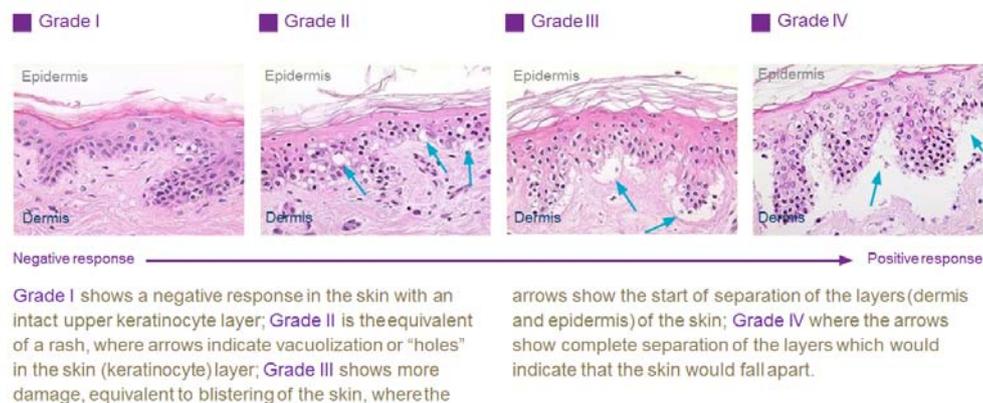
To assess the potential hypersensitivity and immunogenicity of mAbs, a novel human Skin explant assay, Skimune™, was used. Skimune™ ® is an autologous system for preclinical *in vitro* screening of drugs and other chemical compounds developed by Alcyomics (patent EP2524227). It provides a predictive readout of skin damage which also correlates with inflammatory cytokine release and T cell proliferation responses. It can be used in risk assessment to predict the hypersensitivity responses of novel drugs. It is a first-line tool for the assessment of adverse immune reactions. Originally developed for graft-versus-host disease (Vogelsang *et al.*, 1985; Dickinson *et al.*, 1998; Sviland and Dickinson, 1999), a systemic post-transplant complication, it was afterwards modified to become an autologous skin explant model (Ahmed *et al.*, 2016). Skimune™ ® relies, in an initial step, in priming of dendritic cells with consequent T cell activation. Afterwards, these activated cells are co-cultured with a skin biopsy from the same donor (hence an autologous system) to induce tissue damage (Figure 3.2).



**Figure 3.2.** Skimune™ ® assay. This autologous model consists in priming of immune cells and co-culture with skin biopsy from the same donor to induce tissue damage. Histopathological assessment of damage is done according to a scoring grade (I-IV).

This system allows for an *in situ* histopathological characterisation according to Lerner's grading system (Lerner *et al.*, 1974) (Figure 3.3). While Skimune™ ® is not a 3D model, it can be used as a first-line tool to safely predict adverse immune reactions (hypersensitivity and immunogenicity), screening of dose-response compounds or even to comparative studies with other biomolecules. Moreover, Skimune™ ® can be used during the different stages of a

product development pipeline, either in the initial stages of product screening, during product optimisation, as well as the final stage of approval for human testing. Overall, Skimune™ ® is an improvement since it reduces development costs and time, by providing a safe *in vitro* screening tool for assessment of sensitisation.



**Figure 3.3** Grading the reactivity of a test therapeutic. Assessment of the histopathological damage caused by a test therapeutic according to Lerner score grading (I to IV).

### 3.1. Materials and Methods

#### 3.1.1. Materials and reagents

Freshly isolated peripheral blood mononuclear cells (PBMC), human universal AB serum, rituximab (stock 100mg/ml) and trastuzumab (600mg/ml) were kindly provided by Alcyomics Ltd. Dulbecco's Modified Eagle Medium (DMEM, high glucose, with bicarbonates), Fetal Bovine Serum (FBS), Pencillin-Streptomycin (10,000 units penicillin and 10 mg streptomycin/mL), Phosphate Buffered Saline (PBS), Trypsin EDTA solution, MEM non-essential Amino acids, L glutamine solutions 200mM and Sodium Pyruvate Solution, tissue culture flask, 96 well F-bottom plates and WST-1 (Cat. No: 05015944001) were purchased from Sigma Aldrich. CellTiter-Glo® Luminescent Cell Viability Assay was purchased from Promega. All kits will be used as per manufacturers' instructions.

#### 3.1.2. Cell culture and maintenance

HepG2 and HDFn cells were grown as an adherent culture in complete growth media (Dulbecco's Modified Eagle's Media supplemented with 10% Fetal Bovine Serum, 1% Penicillin/Streptomycin, 1% Non-Essential Amino Acids, 1% L-Glutamic acid and 1% Sodium Pyruvate) in T75 tissue culture flasks at 37°C with 5% CO<sub>2</sub>. The cells were subcultured 3 times a week using the following procedure: The spent medium was removed, and the cells were given a Phosphate Buffered Saline wash following which 1x diluted Trypsin was added to

gently lift the cells. The cells were then re-suspended in 1:15 dilution in T75 tissue culture flasks.

### **3.1.3. *Cell seeding and exposure to mAbs***

The HepG2 and HDFn cells were seeded at a density of 5000cells/well onto a Greiner 96-well F bottom tissue culture plate. 50µl of rituximab (R) and trastuzumab (H) were added to the test wells at final concentration of 0.1, 1 and 10µg/ml upon which the plates were incubated for a further 3 hours at 37°C with 5% CO<sub>2</sub>. Details of the final mAb concentration and volumes associated are shown in Appendix A Table A.1. Appropriate volumes of media were added to the control wells to compensate for volume differences arising due to addition of mAbs and effector cells/serum. A final concentration of 10µg/ml of human IgG was used as isotype control. 5% (v/v) of ethanol was used as the positive. Target control refers to the target cells (HDFn or HepG2) in media.

### **3.1.4. *Complement dependent cytotoxicity (CDC)***

Following incubation with varying concentrations of rituximab and trastuzumab, 50µl of human universal AB serum were added to the test wells. The plate was incubated overnight at 37°C with 5% CO<sub>2</sub> (Harjunpaa *et al.*, 2000; Wang *et al.*, 2004). The layout of the 96-well plate for CDC experiment is outlined in Appendix A Figure A.1c

### **3.1.5. *Antibody dependent cellular cytotoxicity (ADCC)***

Following incubation with varying concentrations of rituximab and trastuzumab, 50µl of PBMCS at a density of 50000 cells per well were added to the test wells to achieve an effector to target ratio of 10:1. The plate was incubated overnight at 37°C with 5% CO<sub>2</sub> (Harjunpaa *et al.*, 2000; Wang *et al.*, 2004). The layout of the 96-well plate for ADCC experiment is outlined in Appendix A Figure A.1a

### **3.1.6. *Complement dependent cellular cytotoxicity (CDCC)***

Following incubation with varying concentrations of rituximab and trastuzumab, 50µl of human universal AB serum and 50µl of PBMCS at a density of 50000 cells per well were added to the test wells to achieve an effector to target ratio of 10:1. The plate was incubated overnight at 37°C with 5% CO<sub>2</sub> (Harjunpaa *et al.*, 2000; Wang *et al.*, 2004). The layout of the 96-well plate for CDCC experiment is outlined in Appendix A Figure A.1b

### **3.1.7. *WST-1 cell proliferation assay***

Following exposure of cells to the test compounds for 24 hours, 10µl of WST-1 reagent were added per well and the plates were incubated for an additional 4 hours at 37°C with 5% CO<sub>2</sub>. Endpoint measurements of absorbance were taken at 480nm and 600nm (background) on FLUOstar® Omega multimode microplate reader. Cell viability was expressed as a percentage of the target control.

### **3.1.8. *CellTiter-Glo® Luminescent Cell Viability Assay***

Following exposure of cells to test compounds for 24 hours, the plate and its contents were equilibrated at room temperature for approximately 30 minutes. Volume of Cell Titer-Glo® Reagent equal to the volume of cell culture medium present in each well (e.g. 100µl of reagent to 100µl of medium containing cells for a 96-well plate) was added. Contents were mixed for 2 minutes on an orbital shaker to induce cell lysis. The plate was allowed to incubate at room temperature for 10 minutes to stabilise luminescent signal. Luminescence was recorded in FLUOStar® Omega multiplate reader.

### **3.1.9. *Skimune™ analysis***

The mAbs, at a final concentration of 1µg/ml were incubated with (PBMCs) and matching skin explant from healthy donors (n=10 for rituximab and n=5 for trastuzumab) for the Skimune® mAb assay. The endpoint is a grading score of histopathological damage, ranging from 1 to 4. The analysis was performed in collaboration with Alcyomics Ltd (Ahmed *et al.*, 2016).

### **3.1.10. *Statistical analysis***

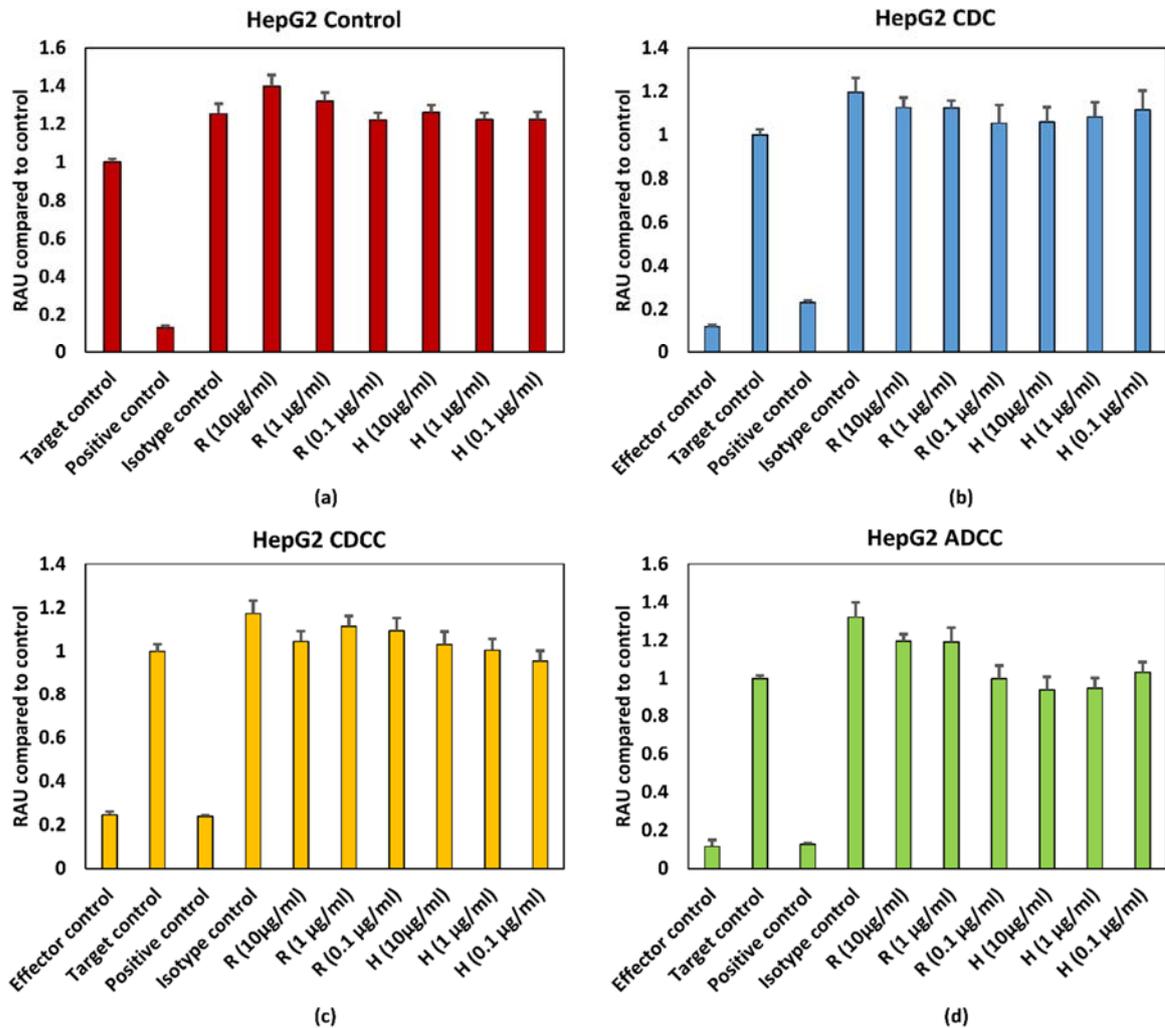
Statistical analysis was carried out using Minitab 17 software (Minitab, 2014). Statistically significant results were reported based on a one way analysis of variance (ANOVA) test followed by post hoc tests (Tukey's/Fishers/Dunnett's) (Keppel and Wickens, 2004; Ruxton and Beauchamp, 2008). All values are expressed as percentage of target control with mean ± standard error (SE).

## **3.2. Results**

### **3.2.1. *MAb induced effect on in cell viability***

HepG2 cells were treated with varying concentrations of rituximab and trastuzumab. As stated in section 3.1, cell viability was expressed as percentage of Target control which are wells containing only HepG2 cells. Effector cells/serum blank refer to those wells which contain only PBMCs for ADCC, serum for CDC and PBMCs+serum for CDCC assays to measure background absorbance for the effector cells and/or serum for the respective endpoint. Neither

a concentration dependent effect on cell viability nor an effector/serum dependent response were observed for either of the tested mAbs as shown in Figure 3.4. 5% (v/v) of ethanol (final concentration) was used as the positive control. Figure 3.4 represents the pooled responses from 4 donors for ADCC, CDC and CDCC tests.

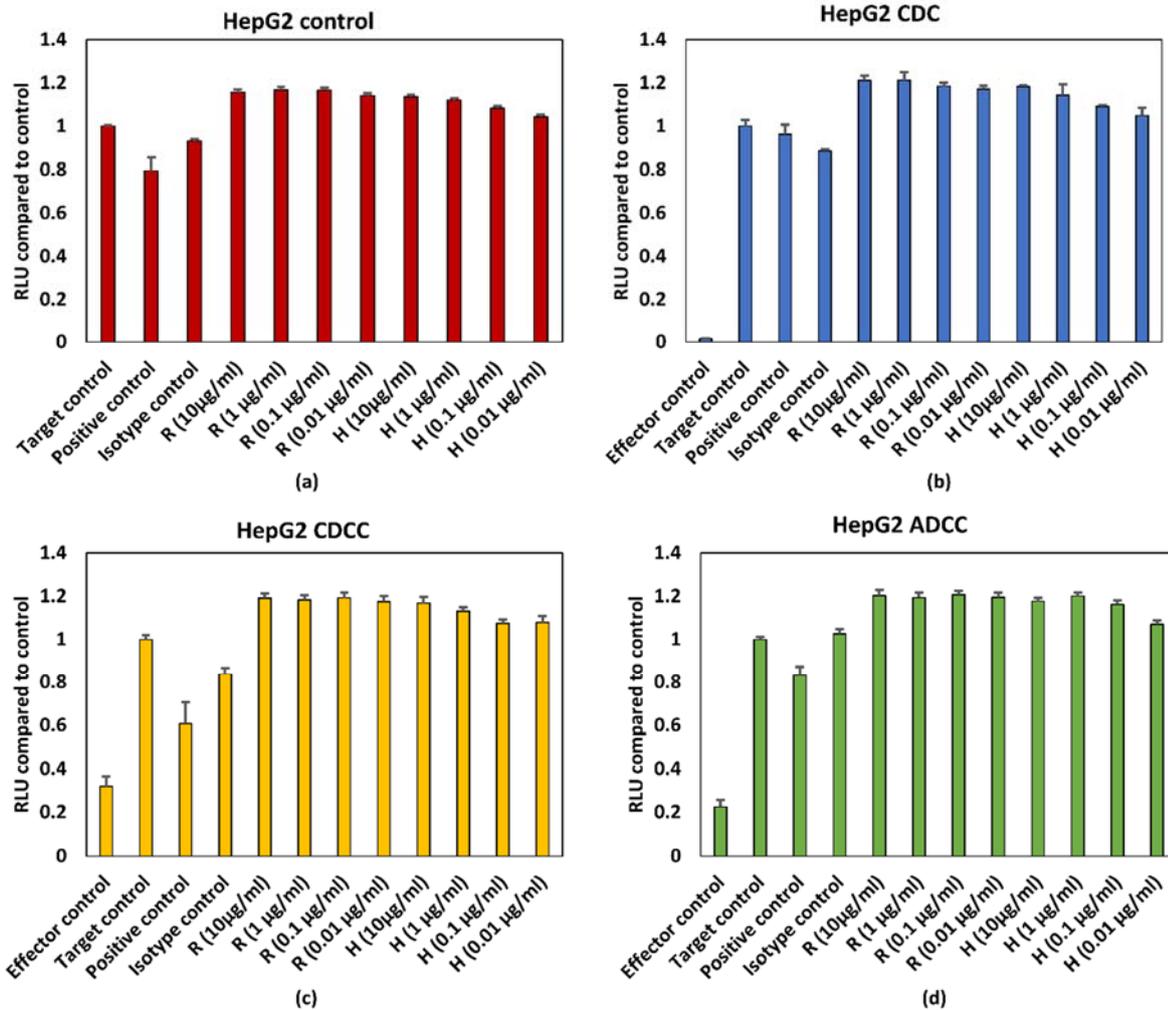


**Figure 3.4** (a) Control conditions without effector cells/serum (b) CDC (c) CDCC and (d) ADCC assay results of rituximab and trastuzumab based on the WST assay. Results represent pooled responses from four donors (n=4). All values are expressed as Relative Absorbance Units (RAU) of control (mean±SEM). R: Rituximab and H: Trastuzumab. Positive control is 5% (v/v) of absolute ethanol.

### 3.2.2. MAb induced effect in ATP levels

HepG2 and HDFn cells were treated with varying concentrations of rituximab and trastuzumab. ATP content was expressed as percentage of target control which are wells containing only HepG2 cells or HDFn cells. 5% (v/v) of ethanol (final concentration) was used as the positive control. Effector cells/serum blank refer to those wells which contain only PBMCs for ADCC, serum for CDC and PBMCs+serum for CDCC assays to measure background absorbance for

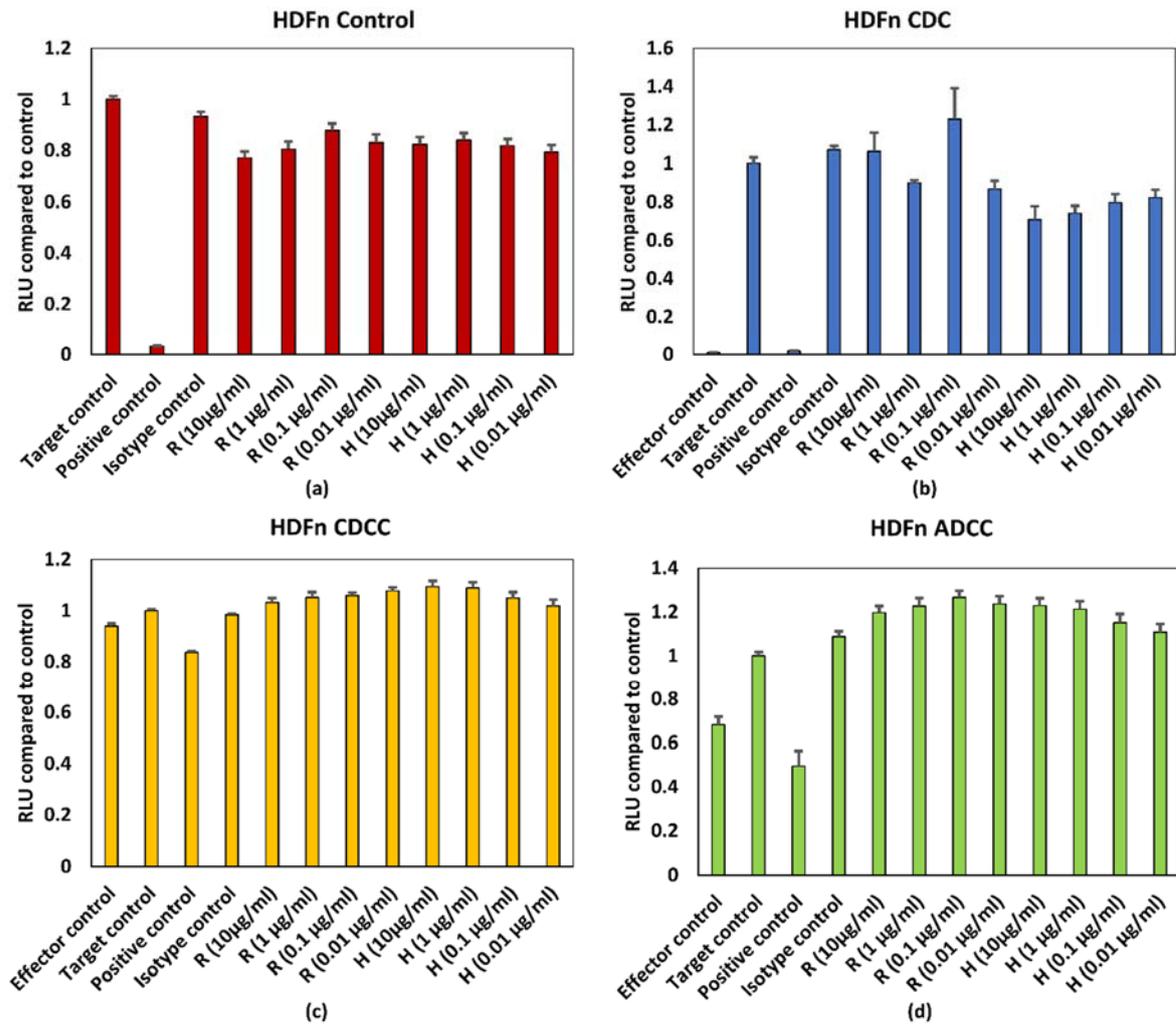
the effector cells and/or serum for the respective endpoint. Figure 3.5 and 3.6 represent pooled responses from 4 donors for ADCC, CDC and CDCC tests. Neither a concentration dependent effect on ATP content nor an effector/serum dependent response were observed for either of the tested mAbs on HepG2 cells (Figure 3.5).



**Figure 3.5** (a) Control conditions without effector cells/serum, (b) CDC (b) CDCC and (d) ADCC assay results of rituximab and trastuzumab based on the ATP content compared to control in HepG2 cells exposed to mAbs. Results represent pooled responses from four donors (n=4). All values expressed as Relative fluorescence units (RFU) compared to control (mean±SEM). R: Rituximab and H: Trastuzumab. Positive control is 5% (v/v) of absolute ethanol.

The HDFN cells seem to be slightly more sensitive to the responses evoked by mAbs when compared to HepG2 cells. Figure 3.6a depicts the response elicited by mAbs without the influence of PBMCs and/or serum. The response is generally lower than the target control, which contains only HDFn cells. When assessing responses resulting from CDC (Figure 3.6b), all concentrations of trastuzumab tested have lower responses when compared to the control. A similar trend can be observed for rituximab; however, the variation is higher when compared

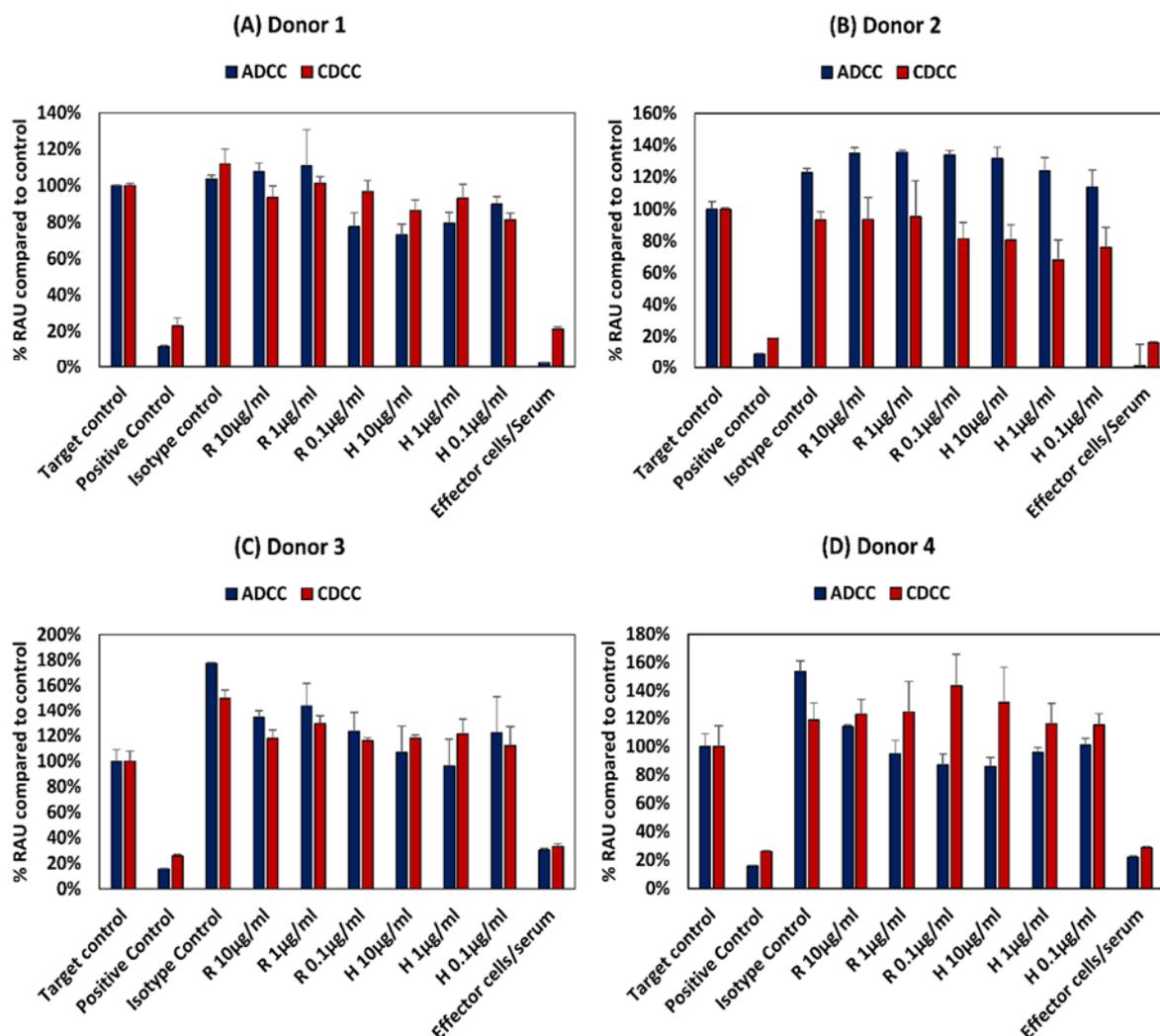
to trastuzumab (Figure 3.6b). Neither a concentration dependent effect on ATP content nor an effector/serum dependent response were observed for either of the tested mAbs resulting from ADCC and CDCC (Figure 3.6c and 3.6d). HDFn cells seem to be more sensitive to responses elicited by mAbs and this was observed in the previous chapter where the toxicity of parabens was assessed.



**Figure 3.6** (a) Control conditions without effector cells/serum, (b) CDC (b) CDCC and (d) ADCC assay results of rituximab and trastuzumab based on the ATP content compared to control in HDFn cells exposed to mAbs. Results represent pooled responses from four donors (n=4). All values expressed as Relative fluorescence units (RFU) compared to control (mean±SEM). R: Rituximab and H: Trastuzumab. Positive control is 5% (v/v) of absolute ethanol.

### 3.2.3. Effect of donor variability and intrinsic variation

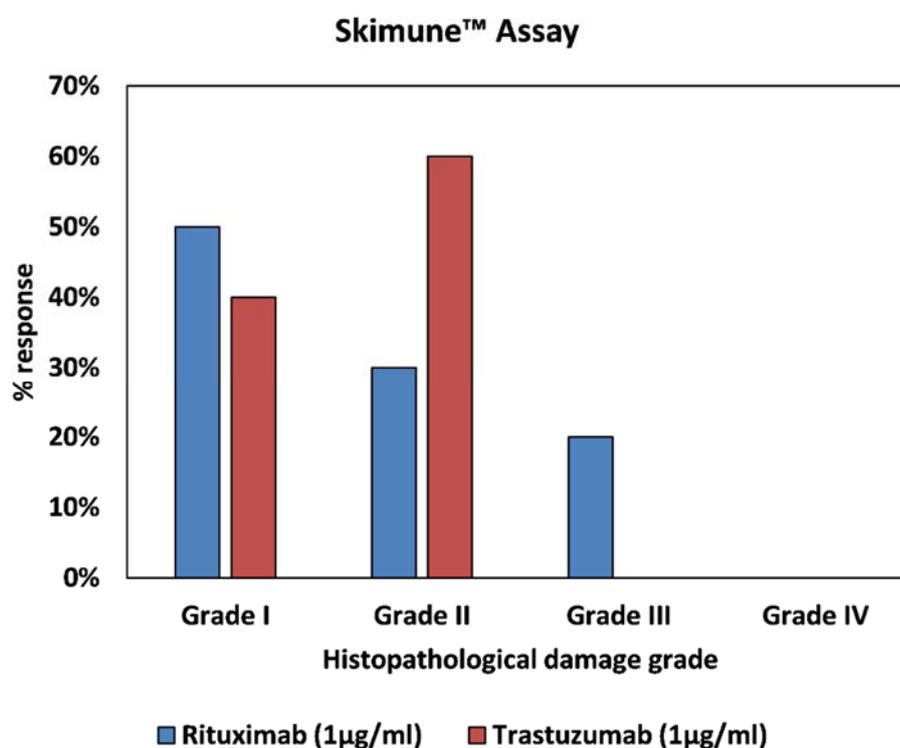
The PBMCs used in the assay were obtained from four different donors and the corresponding responses were varied and non-specific both in terms of the mAb used as well as dose. Figure 3.7 shows the intrinsic variability in the ADCC and CDCC assays owing to donor variability. This variability could be due the specificity of the immune response evoked by each individual which depends on many factors such as genetic make-up and environmental exposure. This intrinsic variation in the assay could potentially confound the outcome of any adverse effect elicited by mAbs.



**Figure 3.7** Intrinsic variation in responses owing to donor variability in ADCC and CDCC assay (a) results of rituximab and trastuzumab based on the ATP content compared to control in HepG2 cells exposed to mAbs. Results represent pooled responses from four donors (n=4). All values expressed as Relative fluorescence units (RFU) compared to control (mean±SE). R: Rituximab and H: Trastuzumab. Positive control is 5% (v/v) of absolute ethanol.

### 3.2.4. Immunogenicity of mAbs assessed by Skimune™ assay

Immunogenicity of rituximab and trastuzumab were assessed via a novel skin explant assay Skimune™. The results of the Skimune tests are shown in Figure 3.8 where the number of Grade I, II and III responses are expressed as percentage of total number of donors. Both rituximab and trastuzumab did not elicit any Grade IV response in any of the donors. The clinical immunogenicity associated with rituximab and trastuzumab are between 1-23% and 1% respectively as reported from clinical trials. (Joubert et al., 2016). Based on the Skimune test classification both trastuzumab and rituximab are deemed weak positives i.e. the testing compound, while known to be a sensitizer, fails to induce an adverse immune reaction in all tested donors. This means that the degree of variability present in clinical trials can be captured up to a certain extent by novel assays such as Skimune™.



**Figure 3.8** Skimune™ assay results with the percentage response based on the different histopathological damage grades induced by rituximab and trastuzumab based on 10 and 5 donors respectively.

### 3.3. Discussion

The *in vitro* systems selected in this study were based on the two main adverse effects associated with mAb based therapeutics: hepatotoxicity and dermal toxicity. Rituximab and trastuzumab elicit an immune mediated reaction to neutralize tumour cells via ADCC, CDC and/or CDCC (Harjunpaa *et al.*, 2000; Bologna *et al.*, 2011). The traditional toxicity tests used here are

routinely used for assessing safety and toxicity of compounds in multitier toxicological assessment studies (Eisenbrand *et al.*, 2002). The objective of the assay used here was to observe any adverse effects of mAbs on HepG2 and HDFn cells upon exposure to naïve PBMCs i.e. detection of any off-target toxicity elicited by mAbs. Both Rituximab and trastuzumab have shown hepatobiliary and skin/infusion related adverse effects in clinical trials (*MabThera, INN-rituximab - European Medicines Agency - Europa EU; Herceptin - European Medicines Agency - Europa EU*).

However as seen from the results shown in Figure 3.4 and 3.5, no dose dependent effect on cell viability or ATP levels were observed for either of the mAbs for HepG2 cells. The antigens for rituximab and trastuzumab are CD20 and HER2 respectively. As the potential off target effects were investigated, both cell lines were chosen such that they do not possess these antigens as surface markers. As HepG2 and HDFn cells do not express the antigen for either rituximab or trastuzumab ADCC and CDCC modes of decrease in cell viability were not observed owing to lack of direct cross target binding associated toxicity. While rituximab has shown to elicit higher CDC mediated responses, the CDC mediated effect of trastuzumab is comparatively lower (Harjunpaa *et al.*, 2000). This has shown to be due to the influence of membrane-bound complement regulatory proteins such as CD46, CD55 and CD59 which are overexpressed in tumour cells (Liu *et al.*, 2014b). CD46 is indeed overexpressed in HepG2 cells and this could be an additional reason why CDC mediated effect was not observed in HepG2 as compared to HDFn cells (Lu *et al.*, 2014). For HDFn cells a decrease in ATP compared to control was observed for all concentrations of trastuzumab tested and a similar trend was observed for rituximab but with higher variability, for CDC mediated response (Figure 3.6b).

Variability in donor responses will be a confounding factor affecting the potential to detect any off target adverse effects, as intra donor variability is quite high owing to the specificity of immune responses elicited (Figure 3.7). The clinical immunogenicity associated with rituximab and trastuzumab are between 1-23% and 1% respectively and this variation is reflective of clinical trials. (Joubert *et al.*, 2016).

Exposure time is another possible reason for the lack of response for these assays. Immune specific reaction could take 5-7 days to develop and this could lead to depletion in the target cell number and hence wouldn't be feasible in this context without a continuous culture in place. This has been shown in T cell proliferation assays for monoclonal antibodies wherein the early phase effects were identified at 20h and late phase effects at 7 days (Joubert *et al.*, 2016).

However, this requires additional measures in place for continuous maintenance of the PBMCs to maintain them at at least a minimum 90% viability. In the case of traditional toxicity tests used to assess off target effects, a continuous maintenance of PBMCs would confound the ability of the test to detect any decrease in cell viability. The percentage viability of PBMCs compared to control at the end of the 72 hours testing period was below 40% as seen from Figure 3.4c,d and 3.5c,d for HepG2 based testing and this distorts the effector target ratio essential for achieving a response. The percentage viability of PBMCs compared to control was around 70% for ADCC and around 85% for CDCC as seen in Figure 3.6c and 3.6d for HDFn cells.

Different approaches have to be adopted for safety evaluation of monoclonal antibody derived therapeutics when compared to small molecule drugs owing to innate differences, such as species specificity, degradation, increased half-life, complex dose response relationship, interaction, lack of generic testing material, pleotropic and synergistic mechanisms to name a few (Cavagnaro, 2002). The main bottleneck in using *in vitro* systems for assessing the toxicity of mAbs is that the effector cells must be co-incubated or co cultured with the cell line of interest. The sensitivity and specificity of these assays depends on several factors which must be optimised, such as cell density, incubation times as well as the choice of system and assay endpoint. The innate complexity, diversity and size of mAbs based therapeutic as well as their diverse mechanisms of actions that involve many pathways exacerbate the need for carefully designed *in vitro* systems that consider all the above factors.

New generation preclinical safety testing tools would have to be high throughput, rapid and cost effective to meet the accelerated growth of the biopharmaceutical market. They also need to be highly reproducible and be sufficiently predictive to allow for rapid screening facilitating reliable selection of new compounds at initial stages thus saving time and money. They would also provide an alternative to animal testing considering the various drawbacks of *in vivo* systems as seen in the case of TGN1412 (Stebbing *et al.*, 2007). *In vitro* systems have now evolved from 2D co-cultures to 3D spheroidal co-cultures, organs on chips as well as whole blood systems to better mimic the responses that could be produced in a human system (Whritenour *et al.*, 2016). Receptor binding studies are also considered to be indicative of biological activity of mAbs as binding to different FcγR receptors elicit different effector functions (Gillis *et al.*, 2014). These studies can either be cell based or conjugated beads based such as αscreen™ technology (Wojtal *et al.*, 2012; Kim *et al.*, 2017). Immunogenicity testing of mAb based therapeutics using T cell proliferation and cytokine assay have been reported

previously for rituximab and trastuzumab (Joubert *et al.*, 2016). Hypersensitivity reactions have been assessed using Skimune™, a non-artificial human skin explants based assay for safety and efficacy assessment of novel compounds and drugs, developed by Alcyomics Ltd (Ahmed *et al.*, 2016). Immunotoxicogenomics and expression profiling of both *in vivo* and *in vitro* systems are being used to identify pathways, mechanism of action as well as biomarkers for study of delayed hypersensitivity reactions (Shao *et al.*, 2014). These advancements may contribute to better designed pre-clinical testing strategies for monoclonal antibody derived therapeutics

### 3.4. Chapter summary

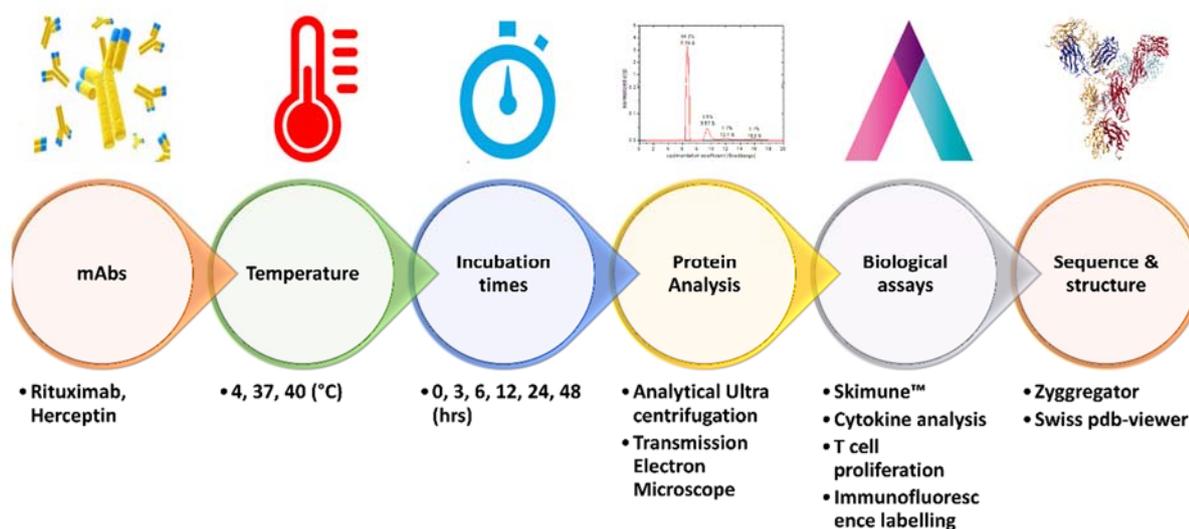
This study explores the applicability of traditional *in vitro* toxicity tests for detecting any off-target adverse effect elicited by mAbs on specific organ systems using hepatocarcinoma cell line (HepG2) and human dermal fibroblasts neonatal (HDFn), respectively. The mechanism of antibody dependent cytotoxicity (ADCC), complement dependent cytotoxicity (CDC) via complement activation, complement dependent cellular cytotoxicity (CDCC) were assessed. The rationale behind use of these particular cells lines were that they do not possess the antigens required by either of the mAbs and would therefore allow for observing off target effects if any. Major results: No apparent ADCC, CDCC or CDC mediated decrease in cell viability was measured for HepG2 cells. For HDFn cells, though ADCC or CDCC mediated decrease in cell viability wasn't detected, a CDC mediated decrease in cell viability was observed. Several considerations have been elucidated for development of *in vitro* assays better suited to detect off target toxicity of mAbs (Kizhedath *et al.*, 2018a). The immunogenicity of rituximab and trastuzumab were assessed using the novel Skimune™ skin explant assay which showed a weak positive test classification for both mAbs tested. Therefore, novel assays such as Skimune™ deemed to be more suitable for detecting potential immune related adverse effect elicited by mAbs. The different Skimune™ ® classes reflect the likelihood of an adverse immune reaction to occur during a stage IV clinical trial. The subsequent chapter will assess the suitability of the Skimune assay to assess the immunogenicity of mAbs aggregates formed at physiological conditions (temperature and pH) during administration or in storage.

## Chapter 4: Aggregation of mAbs at physiological pH and temperature and detection of immunogenicity using Skimune™ assay

The previous chapter explored the applicability of traditional toxicity assays for detecting adverse effects elicited by mAbs. It also described the novel Skimune™ skin explant assay and its potential to detect immunogenicity of mAbs. This chapter describes the heat stress induced aggregation of mAbs under physiological conditions and their detection using analytical techniques such as Analytical ultracentrifugation (AUC) followed by detection of immunogenicity using Skimune™ assay. Although multiple studies describe aggregation potentials at longer periods of time and elevated temperatures, this study was designed to assess the potential of commercial mAbs to form aggregates and induce an immunological response at physiological conditions. Increased aggregation potential can prove detrimental to the monoclonal antibody production process and lead to manufacturing failures (Rathore *et al.*, 2013). It is also a critical quality parameter during patient administration as aggregated forms of mAbs have been shown to elicit immunogenic responses (Ahmadi *et al.*, 2015). Therefore, during the development of mAb based therapeutics suitable analytical techniques are required to characterise these aggregates. Analytical ultracentrifugation (AUC) is an analytical technique that allows for investigation of molecular weight of biological molecules and is regaining popularity in biopharmaceutical development (Berkowitz and Philo, 2015; Liu *et al.*, 2015). AUC allows for the quantitative description of protein aggregation and the formation of large supramolecular complexes based on their sedimentation properties (Berkowitz, 2006). AUC experiments can be carried out in two modes which are highly complementary (Cole *et al.*, 2008). Sedimentation velocity analytical ultracentrifugation is a hydrodynamic approach which provides details of particle mass and shape and is particularly useful for studying multicomponent irreversible and reversible mixtures of species. Sedimentation equilibrium analytical ultracentrifugation on the other hand is a thermodynamic approach which provides insights into equilibrium constants of the process and stoichiometry and is useful for studying reversible self- and hetero- association.

In this study a sedimentation velocity experiment for mAbs in buffered solution was performed to explore the formation of stable and metastable oligomers, irreversible aggregates and degradation products appearing as a result of storage and thermal stress. The appearance of mAb aggregates and their breakdown products in solution as result of time (storage) and temperature (thermal) stress was explored. The potential of these mAbs to cause hypersensitivity was subsequently investigated using Skimune® (Skimune) as described in

Chapter 3. (Ahmed *et al.*, 2016). Furthermore, the potential aggregation propensities of mAbs based on their primary sequence and 3D structure were explored. The overall methodology is outlined in Figure 4.1.



**Figure 4.1** Overview of experimental methodology to assess thermal-stress induced aggregation and associated immunogenicity.

## 4.1. Materials and Methods

### 4.1.1. Thermal stress of antibodies

Rituximab (chimeric IgG1) and trastuzumab (humanized IgG1) were subjected to thermal stress for different durations at physiological pH based on a full factorial design of experiment (DOE) as shown in Appendix B (Table B.1). The samples of rituximab and trastuzumab were prepared in Dulbecco buffer at concentrations of 1 mg/ml and 10 mg/ml treated initially at 4°C, 37°C, 40°C for time periods of 0, 3, 6, 12, 24, 48 hrs. Heat treated rituximab samples were also stored for several weeks at 4°C. At the end of the storage period, samples were heated at 65°C for 2 hours. A positive control for aggregation was also prepared for each mAb, by leaving the samples at 65°C for one hour in an acidic buffer (pH=3). For the purposes of this study trastuzumab and Herceptin has been used interchangeably.

### 4.1.2. AUC analysis

The treated 1mg/ml samples were subjected to Analytical ultracentrifugation (AUC) to further characterise the aggregates. Sedimentation velocity experiments were carried out using ProteomeLab XL-I analytical ultracentrifuge (Beckman Coulter, Palo Alto, USA) at the temperature of 20°C and rotation speed of 40,000 rpm. Scanning protocol started simultaneously as the rotation began. The density and viscosity of the buffer at experimental

temperature (20°C) was calculated using program SEDNTERP (Laue *et al.*, 1992). Sedimentation velocity profiles (65 scans taken 1 second apart) were treated using size-distribution  $c(s)$  model implemented in the program SEDFIT with consecutive refining using Bayesian statistics to obtain  $cP(s)$  (Schuck, 1998; Brown *et al.*, 2008). In order to determine the mass of each species, the  $c(s)$  distribution was converted to  $c(M)$  distribution. Each peak on the distribution plot was integrated in order to obtain the weight-averaged values for sedimentation coefficient and molecular mass. Integrated values of sedimentation coefficient ( $s$ ) obtained at experimental conditions were converted to the standard conditions ( $s_{20, w}$ ) (which is the value of sedimentation coefficient in water at 20°C). The analysis was performed by the Newcastle University Protein Purification and Analysis Unit (NUPPA).

#### ***4.1.3. Immunogenicity of aggregated mAbs***

Heat-stressed samples of trastuzumab and Rituximab (1 and 10µg/mL) were incubated with Peripheral blood mononuclear cells (PBMCs) and matching skin explant from healthy donors (n=5) for the Skimune® mAb assay. The endpoint is a grading score of histopathological damage, ranging from 1 to 4. The analysis was performed in collaboration with Alcyomics Ltd. Samples were stored immediately after the heat shock protocol at 4°C until further analysis. They were further analysed in T cell proliferation assay, cytokine analysis as well as immunofluorescence labelling for cell death protein such as heat shock Protein 70 and caspase 3 (data has not been presented here as it is in preparation for publication by Alcyomics Ltd.).

#### ***4.1.4. Analysis of protein sequence and aggregation propensities***

The primary sequence of rituximab and trastuzumab were obtained from IMGT/mAb-DB (Lefranc *et al.*, 2009). Multiple sequence alignment was performed using MSA tool of T Coffee server (Notredame *et al.*, 2000). Physicochemical properties as well as composition of amino acids present in the primary sequence of mAbs were assessed using EMBOSS Pepstat (Rice *et al.*, 2000). To predict the potential aggregation propensity the primary sequence was input into an online platform called Zyggregator which analyses the tendency of protein aggregation based on parameters such as hydrophobicity,  $\beta$  sheet -  $\alpha$  helical propensities as well as charge of amino acids present in the primary sequence (Tartaglia and Vendruscolo, 2008). The surface activity of dissimilar residues were assessed using BepiPred 2.0 (Jespersen *et al.*, 2017). Additionally the 3D structures of the fab regions of both rituximab and trastuzumab were studied using Swiss PDB viewer to further explore the position of these residues within the mAbs structure (Guex and Peitsch, 1997; Schwede *et al.*, 2003).

## 4.2. Results

### 4.2.1. AUC analysis of mAb aggregation

Interference data show presence of small species, which is not detected by absorbance at 280 nm; this implies that these species have non-proteinaceous nature. The presence of very large aggregates can be detected by distortion of early scan above the plateau region for monomer. Summarised results of the integrated sedimentation parameters were used to represent percentage of species measured at different temperature and time points. Overall, the loss of monomers by heat stress was found to be very low, with less than 5% aggregation of the total protein content (Table 4.1).

**Table 4.1** Quantification of aggregated content of heat-stressed mAb samples by analytical ultra-centrifugation.

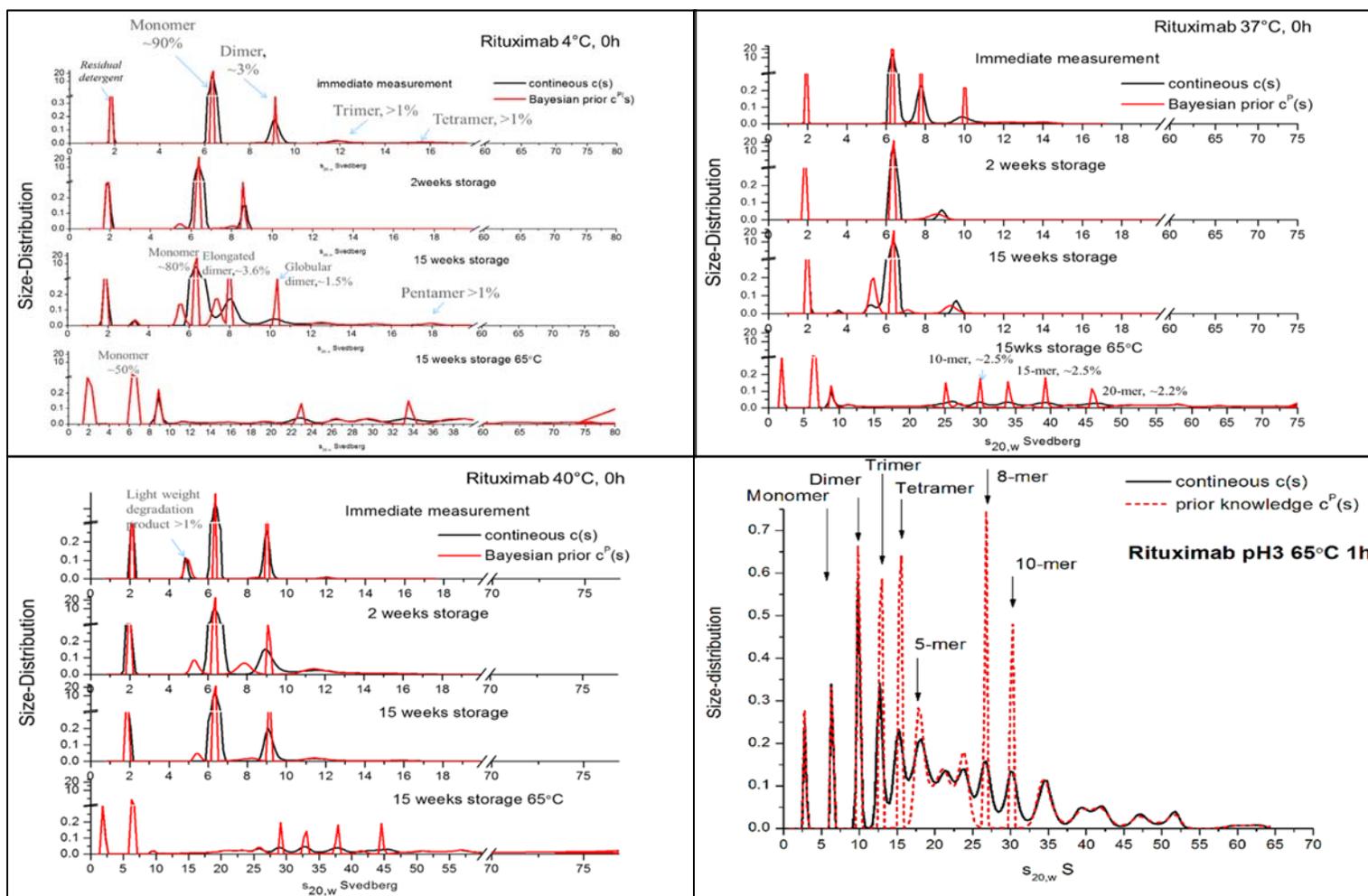
		Absorbance data								RMSD	
		monomer			dimer			Heavier species			
	Temp (°C)	time (h)	sedimentation (S)	Mass (kDa)	%	sedimentation (S)	Mass (kDa)	%	%		
Rituximab	4	0	6.412	152.1	97.92	9.250	263.653	1.14	0.93	0.015	
		48	6.415	138.3	97.72	9.715	257.707	1.70	0.58	0.017	
	37	0	6.411	153.4	96.16	9.089	258.781	2.69	1.15	0.014	
		3	6.411	149.9	97.41	9.270	274.930	1.68	0.91	0.013	
		6	6.408	152.5	96.88	9.190	261.938	2.07	1.05	0.011	
		12	6.413	149.3	96.69	9.522	270.064	2.28	1.11	0.010	
		24	6.426	146.8	99.00	-	-	-	1.00	0.018	
		48	6.424	148.3	97.38	8.340	229.272	1.61	1.48	0.014	
	40	0	6.415	153.0	97.26	9.042	262.435	2.23	1.24	0.016	
		3	6.407	147.4	98.34	8.997	262.004	1.18	1.30	0.017	
		6	6.399	149.5	98.16	8.768	249.582	1.15	1.04	0.017	
		12	6.397	155.3	98.40	-	-	-	1.60	0.017	
		24	6.425	150.3	97.38	9.461	281.061	1.32	1.42	0.018	
		48	6.420	147.6	97.73	9.649	281.750	1.15	1.44	0.016	
	65	1	1.587	13.1	0.71	6.503	108.750	2.54	73.01	-	
	Trastuzuma	4	0	6.441	149.0	97.96	9.432	263.977	1.70	0.47	0.010
		37	48	6.427	148.1	94.80	9.109	249.839	2.85	2.65	0.010
		40	48	6.421	148.1	95.51	9.235	255.364	2.61	1.93	0.011

	65	1	6.346	132.6	2.66	9.862	256.928	6.11	69.25	0.010
--	----	---	-------	-------	------	-------	---------	------	-------	-------

Quantification of monomers, dimers, and heavier species in Rituximab and Trastuzumab monoclonal antibodies (at 1 mg/mL) after exposure to a heat-stress protocol (4, 37 and 40°C for 0, 3, 6, 12, 24 and 48 hours). Quantification by sedimentation velocity (S), mass (kDa) and percentage in overall mAb sample (%). Missing values (-). RMSD : Root mean square deviation determines the model fit.

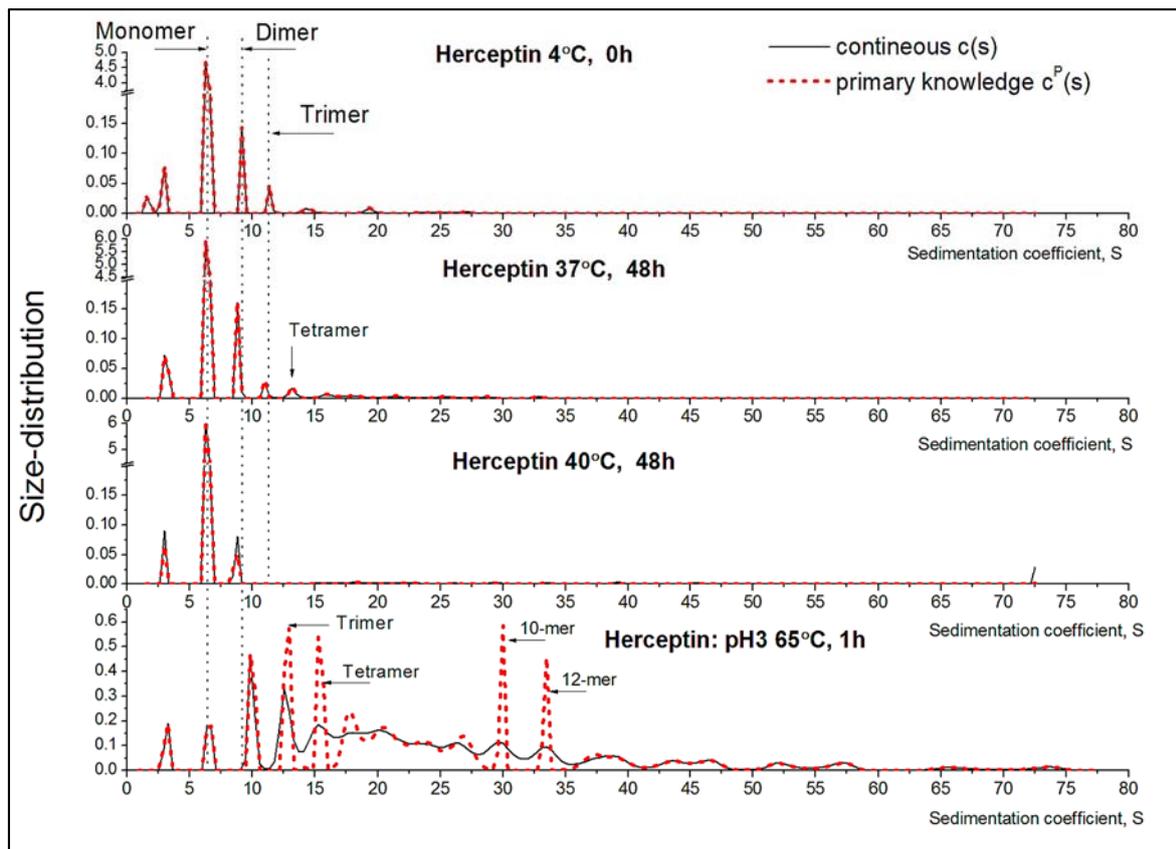
For rituximab, despite the significant temperature stress (65°C for 2 hours), more than 50% of the monomer remains in solution. AUC obtained size-distribution shows the majority of monomeric species (sedimenting at 6.3 S, with the molecular weight of 152 kDa) present in solution and accompanied by the appearance of some dimeric species. These dimeric species could have different stoichiometry: more elongated (sedimenting close to 8 S) and more globular configuration (sedimenting just below 10 S) (Schuck *et al.*, 2002; Aziz *et al.*, 2007; Chou *et al.*, 2011). The presence of heavier oligomers was rather insignificant even despite of prolonged storage time. Moderate temperature stress (40°C) result in appearance of some light-weight degradation products such as isolated heavy chain (sedimenting at about 5S). A substantial amount of aggregates in MDa range was observed only after treatment at 65°C where monomeric species were reduced to 50%. Sedimentation coefficient distributions of rituximab (Mabthera) species for all experimental conditions are shown in Figure 4.2. Furthermore, TEM analysis, the heat stressed samples of both rituximab and trastuzumab showed presence of small aggregates. The amount of detectable aggregates was higher for the samples at 65°C.

As observed in Table 4.1, the total mass percentages do not add up to 100%. This could be owing to the presence of fragments and could not be detected by AUC. Methods such as LC MS would have been more suited to the detection of fragments and this analysis have not been performed in the current study. The missing values were due to non-availability of data owing to non-suitability of sample.



**Figure 4.2** Sedimentation coefficient distribution of rituximab (Mabthera) species for different durations and storage times at temperatures of (a) 4°C, (b) 37°C, and (c) 40°C and (d) 65°C (black line). The results fit the Bayesian prediction model (red dotted line) based on optimally low RMSD. Monomer species are seen around 6S with elongated dimers appearing at 8S and globular dimers appearing at 10s.

Sedimentation coefficient distributions of trastuzumab species generated based on treatment are shown in Figure 4.3 where different peaks can be identified representing monomeric and dimeric species. It is possible to see different oligomer structures for the three temperature conditions tested, with trimers and tetramers reported at 37°C. In heavily aggregated samples, at pH3, the amount of the monomer drops dramatically, the main species in solution are dimers, trimers, tetramers and so on up to 10-15mers (Figure 4.3). The aggregated species are hard globular particles (frictional ratio+hydration is 1.3) compared with elongated monomeric form (frictional ratio+hydration is 1.6) (Siegel and Monty, 1966; Lebowitz *et al.*, 2002).

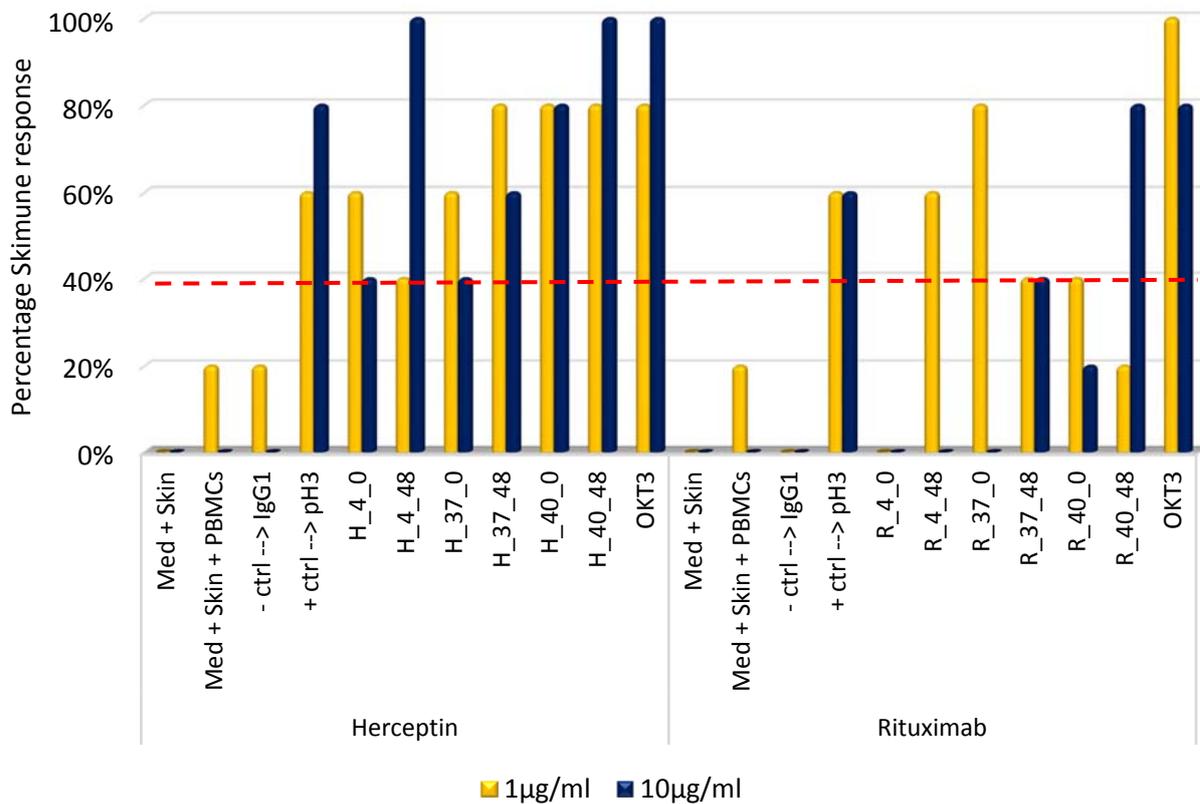


**Figure 4.3** Sedimentation coefficient distribution of trastuzumab (Herceptin) species for different duration at temperatures of 4°C, 37°C, 40°C, 65 (black line). The results fit the Bayesian prediction model (red dotted line) based on optimally low RMSD. Monomer species are seen around 6S with elongated dimers appearing at 8S and globular dimers appearing at 10s.

#### 4.2.2. Immunogenicity of mAb aggregates

The mAb samples subjected to heat stress were tested in the Skimune assay as describe in section 4.2.3. Skimune responses were calculated on number of grade 2 responses and above from 5 donors and are presented in Figure 4.4. The threshold was set at 40% above which the potential to cause an adverse effect is likely. Human IgG1 was used as the control and

Muromona-CD3 (OKT3) as the positive control. For trastuzumab, even at normal storage temperatures (4°C) there seems to be a response greater than that of the control (human IgG1) for both concentrations. There seems to be an increased response for the highest temperature of 40°C over time for higher concentrations of trastuzumab. Most combinations of the concentrations, temperature and storage times appear to elicit a Skimune response of 40% and above, indicating the likelihood of an adverse effect. The highest Skimune responses were comparable to those elicited by a known positive mAb - OKT3 (Muromonab-CD3). For rituximab, responses were observed to increase for longer storage times. Only three conditions of concentration, temperature and storage were above the 40% Skimune threshold. trastuzumab appears to show a slightly higher tendency to cause hypersensitivity when compared to rituximab for both heat treated and control samples.



**Figure 4.4** Percentage Skimune® responses for heat-stressed samples of trastuzumab and Rituximab (1 and 10µg/mL) based on responses from 5 donors. The red dotted line indicates the 40 % threshold. Samples have been given a 3 index code name for easier naming purposes, with the first index representing the mAb tested (Rituximab, R or Trastuzumab, H), the second index representing the temperature conditions (4, 37 or 40°C) and the third index representing the duration of exposure (0, 3, 6, 12, 24 or 48 hours). OKT3 : Muromonab-CD3. Med+Skin: Only medium and skin , Med+Skin+PBMCs: Medium, skin and PBMCs, ctrl IgG1: Isotype IgG1 control, ctrl pH3: Control of mAb treated at pH 3.

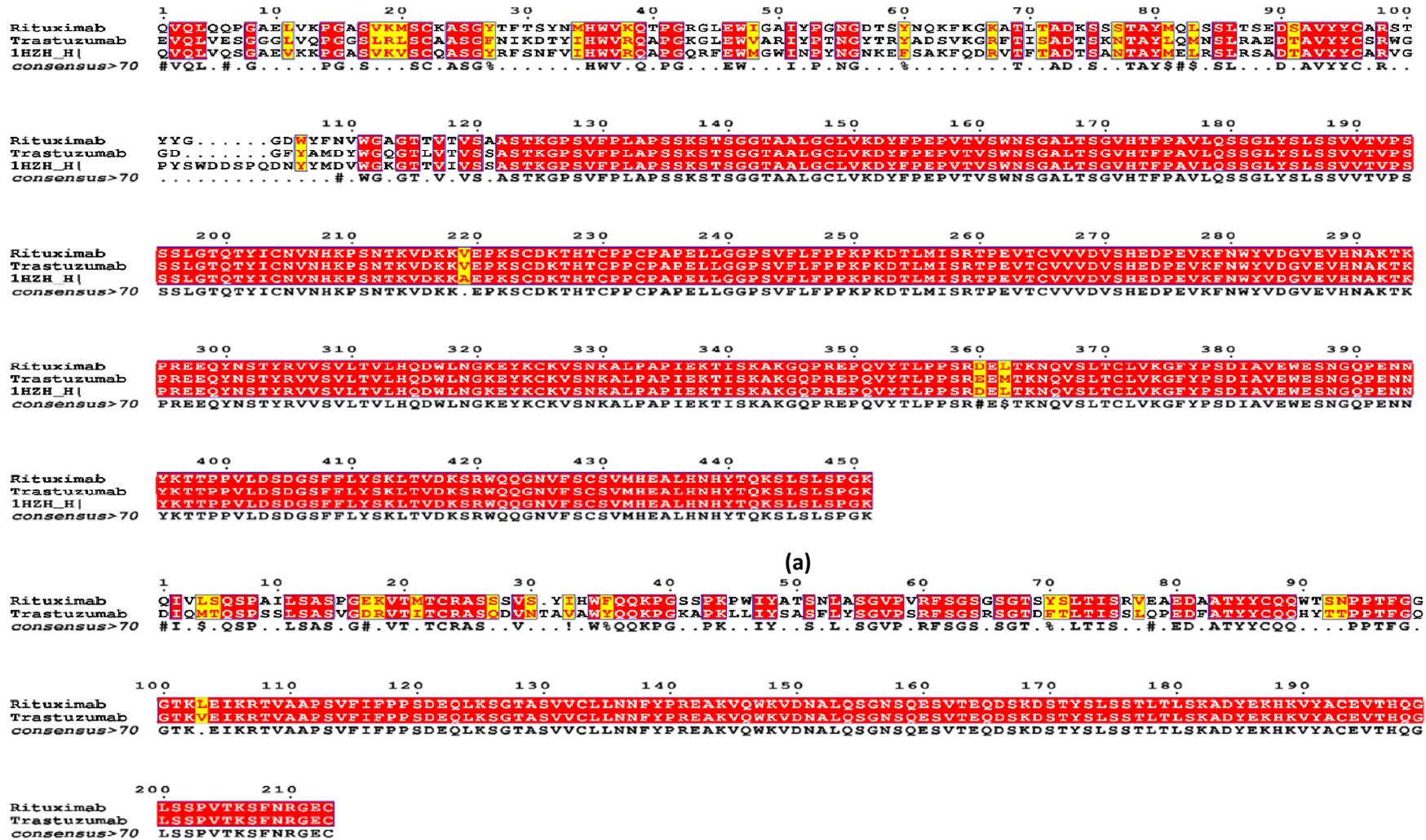
Additional tests for immunogenicity such as T cell proliferation, multiplex cytokine analysis and immunofluorescence labelling of cell death protein such as Hsp70 and Csp3 were performed to further understand the mechanism of immunogenicity. No significant increase of T cell proliferation was observed for any of the temperature and storage conditions tested. An increase in IL-10 and IFN- $\gamma$  was observed for 10 $\mu$ g/ml concentrations for the 37°C and 40°C conditions in both rituximab and trastuzumab. Positive staining Hsp70 and negative staining of Csp3 was observed for the 37°C and 40°C conditions in both rituximab and trastuzumab which indicated mild apoptotic damage that wasn't strong enough to cause cell death (data has not been presented here as it is in preparation for publication by Alcyomics Ltd.).

#### *4.2.3. Prediction of aggregation propensities based on primary sequence.*

Multiple sequence alignment (MSA) of rituximab, trastuzumab (Herceptin) and human IgG1 heavy chain shows that the CH1, CH2 and CH3 region are mostly identical whereas the VH region has considerable differences in amino acid compositions for residue indices 1-120 (Figure 4.5). MSA of light chain regions of rituximab and Trastuzumab shows difference in amino acid composition in the VL region whereas the CL regions remain fairly similar (Figure 5b). This could explain the increased response of trastuzumab in Skimune™ assay when compared to rituximab owing to differences in physicochemical, electronic and topological properties due to difference in primary sequences (Figure 4.4). These properties have been calculated using EMBOSS Pepstats and are shown in Appendix B (Table B.3 and B.4). The number of hydrophobic and positive residues is lower for rituximab than trastuzumab and this could explain the increased response of trastuzumab (Table 4.2). The amino acid composition of the heavy and light chain of rituximab and trastuzumab are shown in Appendix B (Table B.2 and B.4).

**Table 4.2** Results of amino acid composition following EMBOSS Pepstats analysis of rituximab and trastuzumab primary sequences.

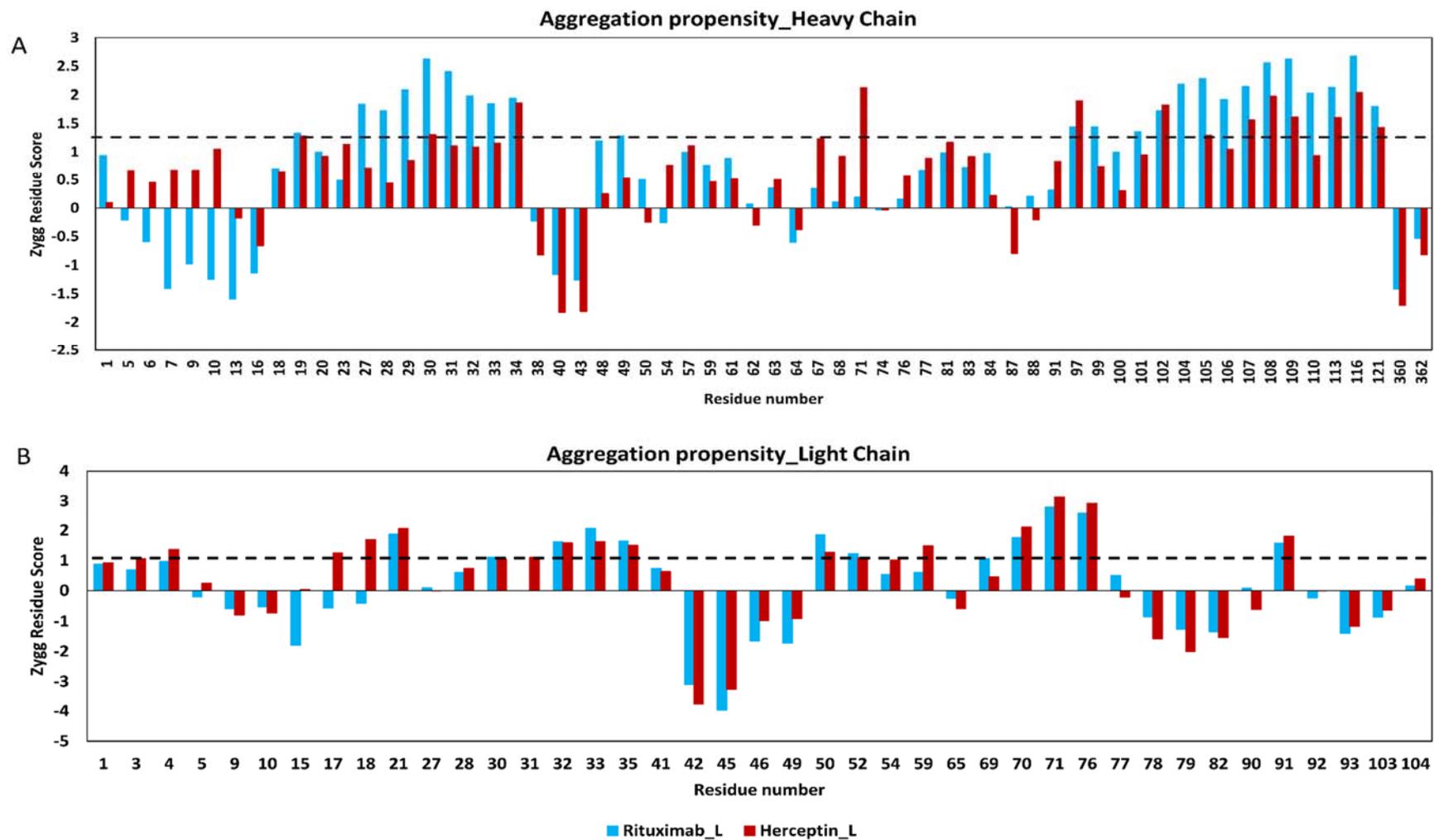
Residue type	Chain	Number of residues	
		Rituximab	Trastuzumab
Hydrophobic residues	Heavy Chain	232	235
	Light Chain	103	100
Charged residues	Heavy chain	91	95
	Light Chain	39	42



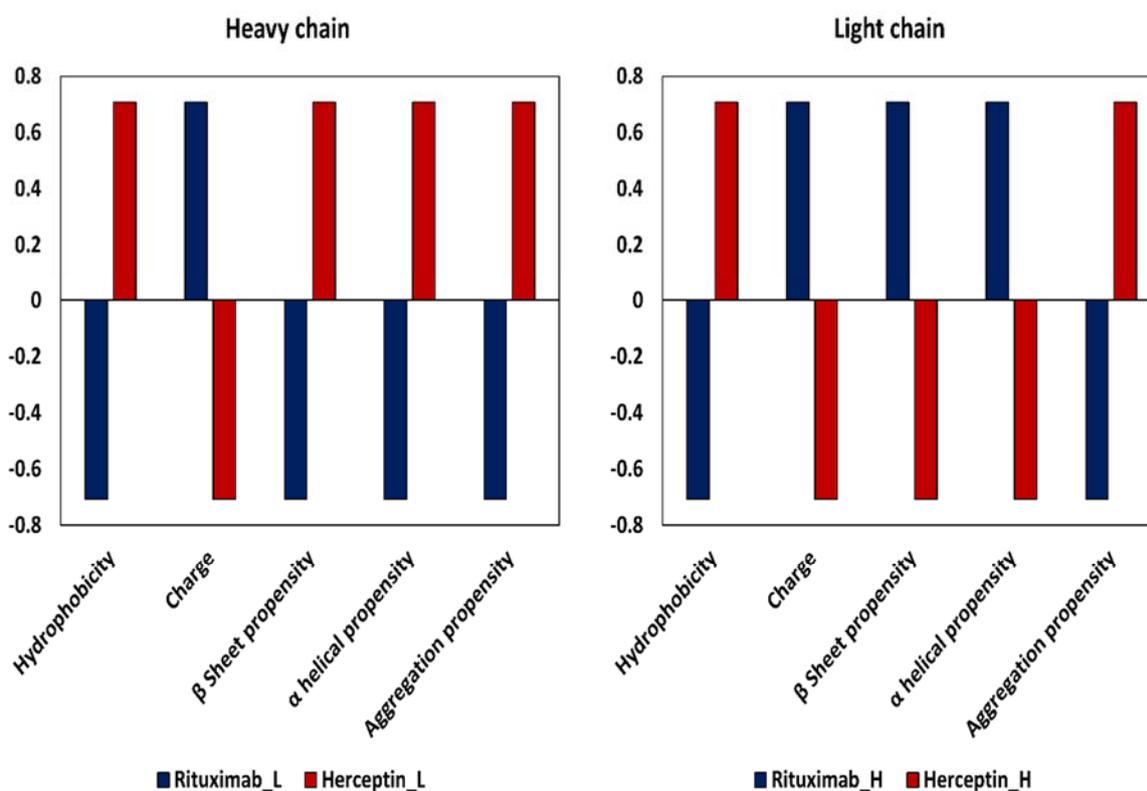
**Figure 4.5** Multiple Sequence alignment of (a) heavy chain of rituximab, herceptin and human Igg1 and (b) light chain of rituximab and Herceptin performed by T-coffee web server (Notredame *et al.*, 2000).

The primary sequences of rituximab and trastuzumab were assessed for aggregation propensities by using the Zyggregator software developed at the Vendruscolo Lab, University of Cambridge (Tartaglia and Vendruscolo, 2008). Individual aggregation propensities of amino acids were calculated based on amino acid scales for hydrophobicity,  $\alpha$  helix/ $\beta$ sheet formation and charge, at physiological pH. Additional terms for capturing information about hydrophobic patterns were also included. These individual scores were then normalised by considering the mean and standard deviation of the aggregation propensities of random sequences generated from the SWISS PROT database. The aggregation propensity at a particular residue is represented as the Zygg score at a particular residue index is 0 if the aggregation propensities is equivalent to that of a random sequence and 1 if it is one standard deviation higher than those of random sequences (DuBay *et al.*, 2004; Pawar *et al.*, 2005; Tartaglia *et al.*, 2008; Tartaglia and Vendruscolo, 2008).

Figure 4.6 show Zygg scores of those amino acids which are dissimilar at a particular residue index following the MSA of rituximab and trastuzumab primary sequences. Residue indices with a Zygg score  $>1$  have higher aggregation propensities. Figure 4.7 depicts the normalised overall values over the different parameters such as hydrophobicity, charge,  $\beta$  sheet propensity,  $\alpha$  helical propensity as well as aggregation propensity for the whole sequence. Both the light chain and heavy chain of trastuzumab shows higher hydrophobicity as well as overall aggregation propensity. For the heavy chain, trastuzumab has higher values for all parameters expect charge when compared to rituximab. For the light chain, even though rituximab has higher overall charge,  $\beta$  sheet and  $\alpha$  helical propensity, trastuzumab shows higher hydrophobicity which may contribute to the higher aggregation propensity.



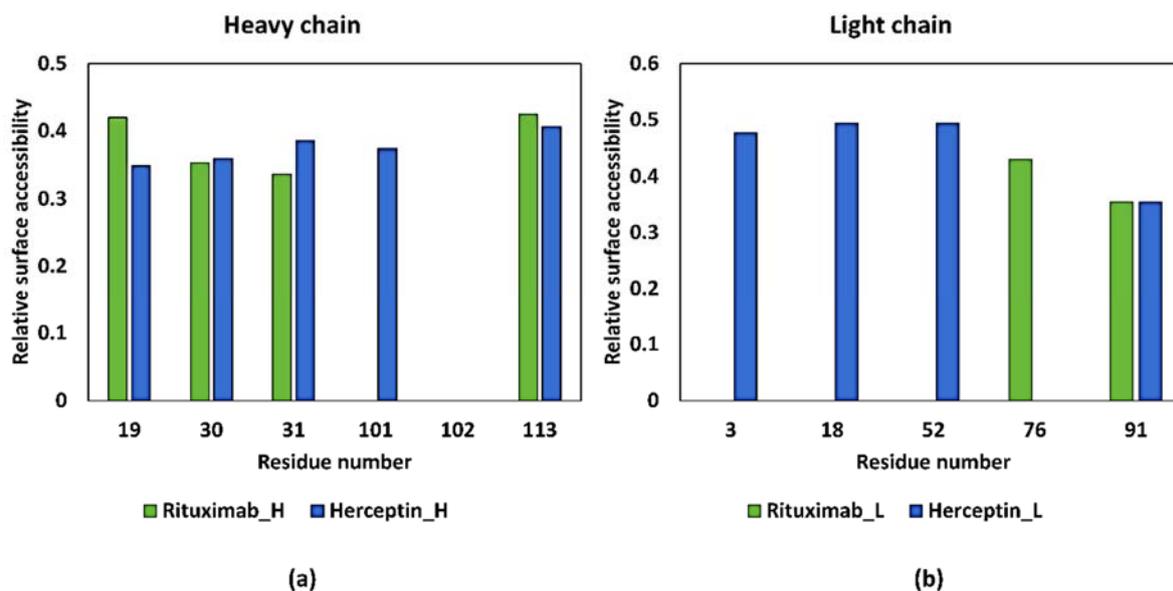
**Figure 4.6** Zygg residue scores reflective of potential intrinsic aggregation propensities of dissimilar amino acid residues of rituximab and trastuzumab (a) heavy chain and (b) light chain.



**Figure 4.7** Normalised overall aggregation propensity descriptors for rituximab and trastuzumab (Herceptin) (a) heavy chain and (b) light chain.

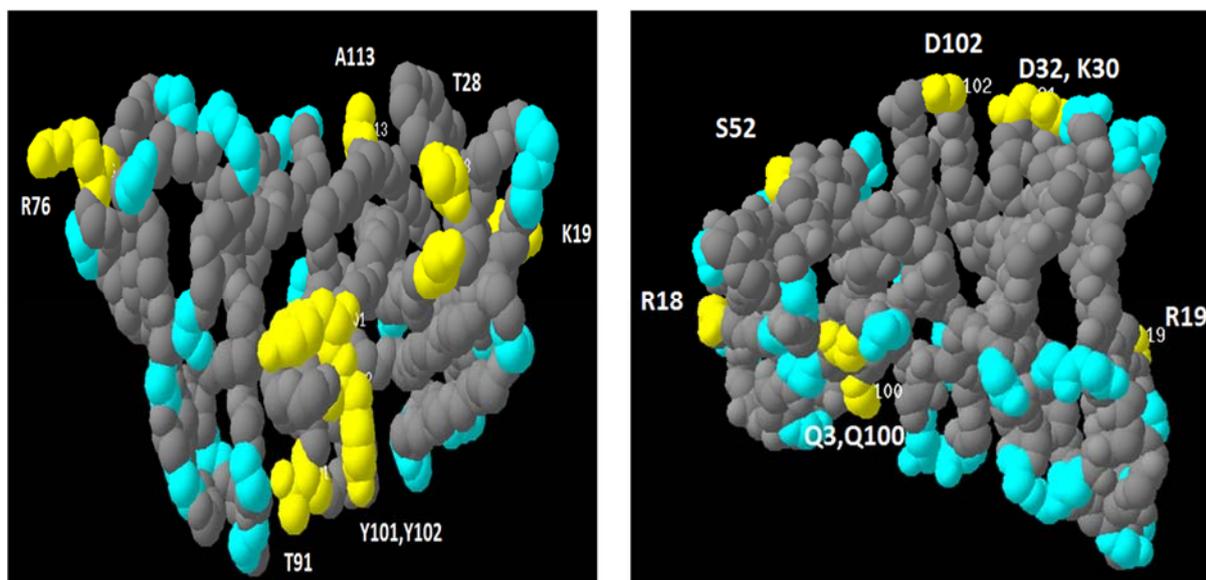
#### 4.2.4. Accessibility of aggregation prone regions in globular structures

For globular proteins like mAbs, accessibility is a prerequisite for increased aggregation propensity for those polypeptide regions which shows a higher intrinsic aggregation propensity. This stems from the fact that if the aggregation prone regions are deep in the core of the globular protein there is much less chance that they would be unstructured or unstable to promote intermolecular interactions that could cause aggregations (Tartaglia and Vendruscolo, 2008). Surface accessibility is thus an important feature that could contribute to potential aggregation tendencies. The relative surface accessibility of exposed dissimilar amino acids with a Zyg score >1 is shown for both heavy and light chain of rituximab and trastuzumab in Figure 4.8. trastuzumab has relatively more residues which have a higher surface accessibility when compared to rituximab especially for the light chain. This could contribute to explaining the higher aggregation tendency of trastuzumab when compared to rituximab.



**Figure 4.8** Relative surface accessibility of exposed dissimilar amino acid residues with a Zyg score >1 in rituximab and trastuzumab (Herceptin) (a) heavy and (b) light chain.

Three dimensional structures for both trastuzumab (PDB ID: 1N8Z) and rituximab (PDB ID: 2OSL) were sourced from RCSB PDB database (Cho *et al.*, 2003; Du *et al.*, 2007). These were then input into Swiss-Pdb viewer for further exploration as shown in Figure 4.8 (Guex and Peitsch, 1997; Schwede *et al.*, 2003). Once the relevant regions were visualised for the dimer containing both the heavy and light chain (VH and VL domains) (1-120 residue index). The structures were assessed based on residues having an accessible surface of greater >30% which were then cross checked with amino acids which showed a Zyg score of >1. ‘E’ refers to exposed residues whereas ‘B’ refers to buried residues. For trastuzumab, the following amino acids were identified in the VL domain: Gln 3 (E), Arg 18 (E), Ser 52(B), His 91 (E) and in the VH domain: Arg 19 (E), Lys 30 (E), ASP 31 (E), Asp 102 (E) (coloured in yellow) as shown in Figure 4.9. For rituximab the following amino acids were identified in the VL domain: Arg 76 (E), Thr 91 (E) in VL and in the VH domain: Lys 19 (E), Thr 28 (E), Ser 31 (E), Tyr 101 (B), Tyr 102 (B), Ala 113 (E).



**Figure 4.9** 3D structures rituximab and trastuzumab VH and VL regions generated by SWISS-Pdb viewer (Guex and Peitsch, 1997). The blue are shows the accessible residues (>30%). The amino acids coloured in yellow are those accessible residues which have a high intrinsic aggregation propensity.

### 4.3. Discussion

Aggregated forms of mAbs have been shown to elicit immunogenic reactions (Gaza-Bulseco and Liu, 2008; Joubert *et al.*, 2012). In this study, temperatures for mAb aggregation analysis were chosen based on stipulated storage temperature (4°C), normal physiological temperature (37°C) and elevated body temperature during an immune response (40°C). The results reported here indicate some level of protein unfolding and consequent rearrangement of the mAb structural conformation by evidence of some heavier species being present in the sample. Though overall aggregated species contributed to only 5% of the total protein content at these temperatures or storage conditions (see Figure 4.2 and 4.3) there appears to be more extensive aggregation at 65°C with a 50% reduction in monomeric species (see Figure 4.2 and 4.3). This is in concordance with earlier studies, where a low level (<3%) of total protein content was reported as aggregated (Ahmadi *et al.*, 2015). Aggregation of mAbs is dependent on factors such as type of monoclonal antibody, type of stress (sheer vs temperature induced, pH induced) and the duration of stress (Rathore *et al.*, 2013; Joubert *et al.*, 2016). Rituximab and trastuzumab have been shown to aggregate when exposed to stirring induced sheer stress for 20 hours (Joubert *et al.*, 2011; Joubert *et al.*, 2012). Elevated temperatures, sheer stress and longer storage times however are factors more prevalent during processing, formulation and shipping rather than during drug administration (Joubert *et al.*, 2011; Joubert *et al.*, 2012; Joubert *et al.*, 2016).

Taking into consideration the size of protein aggregates, their characterization becomes quite complex. Even though AUC is a powerful and useful method for detection of aggregates in the biopharmaceutical industry it has some drawbacks. AUC does not provide insight into the mechanism of aggregation and has reduced accuracy for low level aggregate detection (Liu *et al.*, 2006; Gómez de la Cuesta *et al.*, 2014). Furthermore, the performance and detection limits of SEDFIT, AUCs data analysis software, are affected by random noise as well as non-ideal thermodynamic conditions of the sample (Liu *et al.*, 2006). Therefore, a multi-technique analysis or a combination of a number of approaches may provide more information. Techniques such as size exclusion chromatography-multi angle light scattering (SEC-MALS); asymmetric flow field-flow fractionation multi-angle light scattering (AF4-MALS) and Raman spectroscopy have shown to very useful in monitoring antibody aggregation profiles and can thus be used to supplement AUC data (Fraunhofer and Winter, 2004; Ye, 2006; Gómez de la Cuesta *et al.*, 2014)

Different in vitro assays were used to study the potential immunogenicity of aggregated mAbs. Previous reports indicate that mechanically stressed trastuzumab samples showed higher number of aggregates and higher immunogenic response when compared to mechanically stressed rituximab (Joubert *et al.*, 2016). In vitro human explant-based assays are also an important tool for assessing immunogenic potential of mAb aggregates. Skimune® is one such platform which was used to assess the potential immunogenic profile of the above aggregated mAbs and it was able to detect the immunogenicity in a small population (Ahmed *et al.*, 2016). Skimune® results showed that aggregated trastuzumab and rituximab caused a mild apoptotic damage, by evident histopathological damage (scores 2 and 3 of Lerner's score system) and positive staining for HSP70, but not for casp3 (data not shown). These results indicate that, at these specific temperatures, aggregated mAb samples cause a mild apoptotic damage that is repairable. No increased proliferation of T cells or release of IFN- $\gamma$  was reported (data has not been presented here as it is in preparation for publication by Alcyomics Ltd.).

The differences observed in the responses of heat stressed rituximab and trastuzumab samples may be due to the differences in amino acid composition. Trastuzumab is a humanized IgG1 kappa mAb whereas rituximab is a chimeric IgG1 kappa mAb. The difference in amino acid composition owing to difference in species as well as those in the heavy and light chain may lead to differences in physicochemical properties such as hydrophobicity,  $\alpha$  helical and  $\beta$  sheet propensities as well as charge could lead to varying aggregation tendencies (Obrezanova *et al.*, 2015). This has been shown in potential aggregation propensity prediction analysis where

trastuzumab has shown higher hydrophobicity and corresponding overall aggregation propensities for both light chain and heavy chain over rituximab (Figure 4.7). Apart from intrinsic aggregation propensities of amino acid residues in protein, environmental parameters such as solvents, pH and temperature conditions along with physiological, cellular and molecular interaction in *in vivo* and *in vitro* experimental conditions influence aggregation of mAbs (Joubert *et al.*, 2011; Joubert *et al.*, 2012; Gómez de la Cuesta *et al.*, 2014; Joubert *et al.*, 2016). Additionally, the location of aggregation prone regions within the globular structure of proteins, their accessibility and stability also play an important role in determining the overall aggregation of proteins (DuBay *et al.*, 2004; Pawar *et al.*, 2005; Tartaglia *et al.*, 2008).

#### 4.4. Chapter summary

To summarise, the heat stress induced aggregation of mAbs at physiological conditions was assessed using AUC analysis revealing a small percentage of dimeric species in the heat stressed trastuzumab samples. Furthermore, TEM analysis showed presence of small aggregated species in both heat-treated trastuzumab and rituximab. The immune related adverse effects elicited by these heat stressed samples were assessed using Skimune™ assay which showed an elevated response for heat treated trastuzumab samples. Although the T cell proliferation assay was inconclusive, the cytokine analysis showed a general increase in pro-inflammatory cytokines such as IL10 and IFN  $\gamma$ . The cell death marker immunofluorescence labelling showed a positive staining for Hsp70 and not Csp 3 indicating mild reversible apoptotic damage (data has not been presented here as it is in preparation for publication by Alcyomics Ltd.). The primary sequence and 3D structure of rituximab and trastuzumab were assessed for aggregation propensities based on physicochemical, structural and surface properties wherein trastuzumab was shown to have slightly higher aggregation tendencies. In conclusion this chapter outlines the potential aggregation of mAbs that may occur during administration or storage at physiologically and storage relevant temperatures and how Skimune™ as an *in vitro* assay could be particularly useful to detect immune related adverse effects of aggregated mAbs. Although multiple studies describe aggregation potentials at longer periods of time and elevated temperatures, this study was designed to assess the potential of commercial mAbs to form aggregates and induce an immunological response at physiological conditions.

The following chapters describe the identification and generation of numerical features from mAbs that were utilised for developing a modelling framework that would allow for predicting responses elicited by mAbs such as immunogenicity and/or cross reactivity.

## Chapter 5: QSAR development using novel descriptors sets for mAbs: Influence of intrinsic properties.

Based on the QSAR development workflow mentioned in Chapter 1 both responses elicited by mAbs as well as numerical representation of their properties, descriptors, are critical for model development. While the previous chapter focused on the different experimental methodologies for assessing the adverse effects of mAbs, this chapter focusses on the generation of descriptors from mAbs based on their primary sequences.

Monoclonal antibody therapeutics represent a rapidly growing market which dominates the biopharmaceutical sector with combined EU prescription sales of more than €9 billion (EvaluatePharma®, 2017). This creates a need for rapid bioprocess development strategies to be put in place to support this expansion as well as to further reduce attrition and processing failures. This has been facilitated by the Quality by design paradigm where quality is built in by design by not only designing products to meet their intended performance, clinically, but also designing bioprocesses to deliver the required product quality consistently (Rathore *et al.*, 2014). Thus, better understanding of the monoclonal antibody therapeutic profile would not only allow for early stage screening of potential candidates, but also allow for better design of processes for their production.

Predictive models in developability assessment are proving to be of increasing importance in early risk assessment studies for lead selections in the area of mAb based therapeutics. Generating sensible descriptors is the first step of building protein predictive models and they have been used for proteins in general but not for mAb based therapeutics. These descriptors are based on the structural, physicochemical, topological and electrostatic properties of the constituent amino acids in the primary sequence of proteins. Furthermore, descriptor sets called amino acid scales have been generated by principal component analysis (PCA) of these properties (van Westen *et al.*, 2013a; van Westen *et al.*, 2013b). Additionally there are many software packages which can be utilised for generating these descriptors such as ProtDCal, eMBOSS Pepstat and ProFEAT (Ruiz-Blanco *et al.*, 2015). Primary sequence based descriptors have been used extensively for developing predictive models using multivariate data analysis techniques as well as machine learning methods such as protein function family prediction using SVM, k-NN and probabilistic neural networks (pNN) (Ong *et al.*, 2007; Li *et al.*, 2016); quantitative structure activity relationship (QSAR) models for prediction of peptide binding affinity to the human major histocompatibility complex using SVM and GA PLS; QSAR models for bioactive peptide derived from food proteins (Nongonierma and FitzGerald,

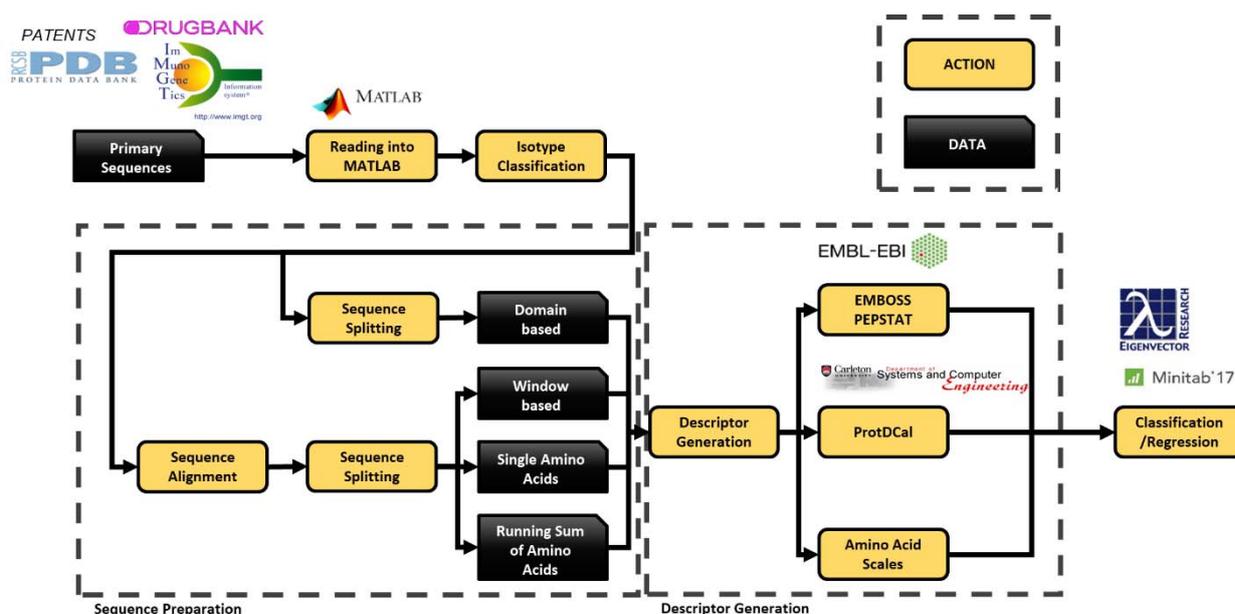
2016); proteochemometric models for antibody-antigen interactions using SVM regression to name a few(Qiu *et al.*, 2015).

As mAb based therapeutics are primarily glycosylated proteins, their primary sequence contains a wealth of information. This information could be translated into descriptors which can essentially be used for developing predictive models for predicting safety or behaviour at a particular process step e.g. the heavy chain and light chain substructure of monoclonal antibodies can have an impact on mode of action as well as their behaviour during production (Rathore *et al.*, 2014; Alt *et al.*, 2016). The heavy chain isotype influences the effector function of monoclonal antibodies as the amino acid residues on the Fc region dictate the binding affinity to the different Fcγ Receptors. This has been utilised for engineering monoclonal antibodies for increased/decreased effector function (Shields *et al.*, 2001; Strohl, 2009). For the light chain isotype, differences in the physicochemical properties can have effects in production and processing as well as in antigen binding (Townsend *et al.*, 2016). Finally with regard to species, the evolution of mAbs from mouse to chimeric to humanized and finally fully human occurred in response to the high risk of hypersensitivity, reduced efficacy, rapid clearance and production of human anti-mouse antibodies elicited by mouse antibodies owing to the species dependent difference of the immune system (Yamashita *et al.*, 2007; Catapano and Papadopoulos, 2013).

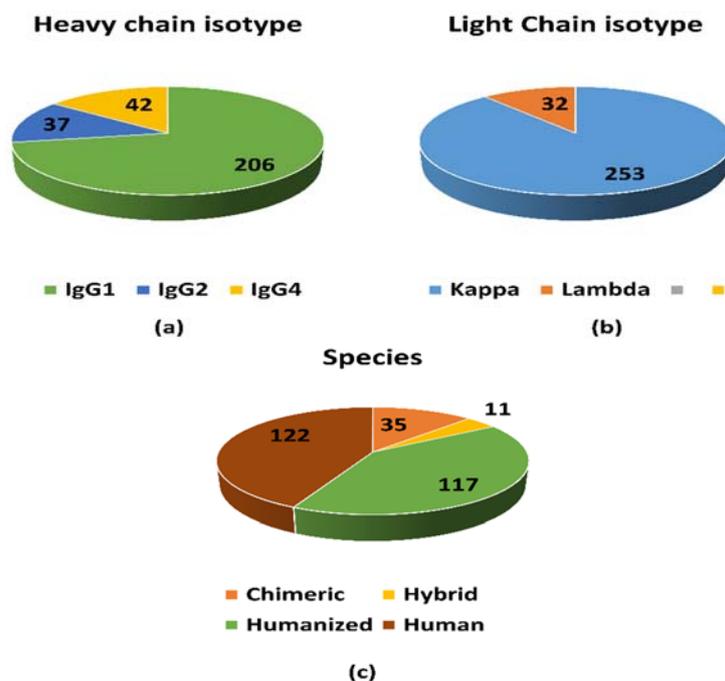
In this study, the identification and creation of data blocks from the primary sequence of mAbs is described. These were used for descriptor generation using eMBOSS Pepstat, ProtDcal and amino acid scales. Five data blocks namely, Domain based, Window based, Substructure based, Single Amino Acid based and Running sum based were generated from the primary sequences of 285 mAbs. Descriptor sets containing 5, 120 and 11 features were generated by using eMBOSS, ProtDcal and amino acid scales respectively. These descriptors capture the physicochemical properties of mAbs with varying degrees of resolution (local to global; singular to cluster). These descriptors from different data blocks were subjected to an exploratory analysis to study any separation based on intrinsic properties such as light and heavy chain isotypes as well as species type of mAbs that would then allow for selection of descriptor sets that could be used for QSAR model development. Exploratory analysis of the descriptor data was carried out using unsupervised pattern recognition methods such as Principal Component Analysis (PCA). PCA is a dimensionality reduction method which generates new variables (principal components) which best capture the variations within a

dataset (Bro and Smilde, 2014). Preliminary groupings within the data can also be visualised using PCA scores.

The overall methodology is illustrated in Figure 5.1. The distribution of 285 mAbs used in this study based on heavy chain isotype (IgG1, IgG2 and IgG4), light chain isotype (kappa and lambda) as well as species (Chimeric, human, humanized and hybrid) have been shown in Figure 5.2.



**Figure 5.1** Overview of methodology for generating descriptors from primary sequence of antibodies.



**Figure 5.2** Distribution of 285 mAbs based on intrinsic properties (a) heavy chain isotype, (b) light chain isotype and (c) species.

## **5.1. Materials and methods**

### **5.1.1. Data collection**

Primary sequences of mAb based therapeutics were collected from different databases such as IMGT, patents, Drug Bank and PDB as well as literature sources. These sequences were stored in a database along with key information pertaining to the mAbs such as heavy and light chain substructure, target antigen, and species as shown in Appendix C Table C.1.

### **5.1.2. Multiple sequence alignment**

The primary sequence of 285 mAbs was subject to Multiple Sequence Analysis (MSA) using the Clustal Omega Version 1.2.4 tool from EMBL to assess the degree of sequence similarity (Sievers *et al.*, 2011). Sequences were input in FASTA format and MSA was performed using the clustalo program. This tool is a part of the EMBL-EBI bioinformatics web and programmatic tools framework (McWilliam *et al.*, 2013; Li *et al.*, 2015).

### **5.1.3. Descriptor generation**

Descriptor generating software such as eMBOSS Pepstat and ProtDcal was used calculate descriptors for the Domain, Window and Substructure data blocks (Rice *et al.*, 2000; Ruiz-Blanco *et al.*, 2015). Local version of the Pepstat and ProtDcal could be run directly in MATLAB by using Java runtime environment 8. Thus, automatic generation of descriptors could be performed. 5 descriptors were generated from Pepstat and 120 from ProtDcal. These descriptors were calculated to be averages based on all contributing residues in the provided sequence. Three amino acid scales; Z-scale (Sandberg *et al.*, 1998), T-scale (Tian *et al.*, 2007b) and MS-WHIM (Bravi *et al.*, 1997) were used to convert each individual residue in the sequences into 11 numerical descriptors; 3 from the Z-scale, 5 from the T-scale and 3 from the MS-WHIM scale. In the Domain, Window and Substructure data blocks each amino acid scale component was average based on all residues, which generated 11 descriptors per sequence domain, split and substructure. In the Single amino and Running sum data blocks, only the amino acid scales were used to generate the descriptors were each residue was generated into 11 numerical descriptors. Details for the individual descriptors are listed in Table 5.1. The automation of descriptor and dataset generation was done by MATLAB codes written by Micael Karlberg, Newcastle University.

**Table 5.1** Protein descriptor sets generated using different software for predicting different mAb isotypes.

Source	Descriptor information
eMBOSS Pepstat	<ol style="list-style-type: none"> <li>1. Molecular weight (Mw)</li> <li>2. Average Residue Weight (ARw)</li> <li>3. Isoelectric point (Ip)</li> <li>4. The net charge of the sequence (Charge)</li> <li>5. Number of residues in sequence (Residues)</li> </ol>
ProtDcal	<ol style="list-style-type: none"> <li>1. Index of the contribution to the free energy from the entropy of the first shell of water molecules in an unfolded state (Gw(U))</li> <li>2. Index of the interfacial free energy of an unfolded state (Gs(U))</li> <li>3. Number of water molecules close to a residue in an unfolded state (W(U))</li> <li>4. Molecular weight (Mw)</li> <li>5. Hydrophobicity by the Kyte-Doolittle Scale (HP)</li> <li>6. Isoelectric point (Ip)</li> <li>7. Electronic charge index (ESI)</li> <li>8. Heat of formation (DHf)</li> <li>9. Isotropic surface area (ISA)</li> <li>10. Polar area for each amino acid in unfolded state (AP)</li> </ol> <p><b>ProtDcal also allows to generate descriptors/variables based on amino acid groups. 12 such groups were used to generate this data set:</b></p> <ol style="list-style-type: none"> <li>11. Group of the most common residues in reverse turn structure: ASN, ASP, GLY, PRO and SER. (RTR)</li> <li>12. Group of the most common residues in Beta Sheet structure: ILE, PHE, THR, TRP, TYR and VAL. (BSR)</li> <li>13. Group of the most common residues in Alfa Helix structures: ALA, CYS, GLN, GLU, HIS, LEU, LYS and MET. (AHR)</li> <li>14. Group of all aliphatic residues contained in the protein. (ALR)</li> <li>15. Group of all non-polar residues: ALA, GLY, ILE, LEU, MET, PHE, PRO, TRP and VAL. (NPR)</li> <li>16. Group of all aromatic residues: HIS, PHE, TRP and TYR. (ARM)</li> <li>17. Group of all polar residues: ARG, ASN, ASP, CYS, GLN, GLU, HIS, LYS, SER, THR and TYR. (PLR)</li> <li>18. Group of all positive charged residues: ARG, HIS and LYS. (PCR)</li> <li>19. Group of all negative charged residues: ASP and GLU. (NCR)</li> <li>20. Group of all uncharged (neutral) residues: ASN, CYS, GLN, SER, THR, TYR. (UCR)</li> <li>21. Group of all unfolding residues: GLY and PRO. (UFR)</li> <li>22. All residues in the sequence (protein). (PRT)</li> </ol>
Amino acid scales	<ol style="list-style-type: none"> <li>1. Z-scale (Physiochemical) <ol style="list-style-type: none"> <li>a. Z1 – Captures the hydrophilicity of the residue</li> <li>b. Z2 – Captures the steric properties of the residue</li> <li>c. Z3 – Captures the polarity and charge properties of the residue</li> </ol> </li> <li>2. T-Scale (Topological) <ol style="list-style-type: none"> <li>a. T1,T2,T3,T4,T5</li> </ol> </li> <li>3. MSWHIM (Electrostatic) <ol style="list-style-type: none"> <li>a. C1 – Captures the molecular electrostatic potential and size of residue</li> <li>b. C2 – Captures differences of positive molecular electrostatic potential between residues</li> <li>c. C3 – Captures differences between positively charged residues</li> </ol> </li> </ol>

#### ***5.1.4. Dataset generation***

The primary sequences of the antibodies were imported into MATLAB 2015a. A custom script was used to split the full sequences into either domain sized sequences or into smaller sequences based on a specified window size. For the latter, multiple sequence alignment (MSA) of all sequences was first performed to align the conserved regions. This was done using the MATLAB function *multialign* with BLOSUM80 as scoring matrix before splitting the sequences. While the domain data block used the fully sized domain sequences, the substructure data block used the CDRs and Fr regions. The Window, Single amino and Running sum data blocks were generated with the MSA step and split into smaller sequence according to a specified window size (see details on each individual data block below).

##### ***Domain based***

Descriptors/variables generated from sequences belonging to the following domains: variable region of the heavy chain, around 110 residues (VH) and around 110 residues (VL). Each domain sequence was used to generate descriptors as described above where each descriptor was calculated as an average value based on all the residues present in each domain.

##### ***Window based***

A multiple alignment of all sequences was performed to align the conserved region of the mAbs. From the sequence alignment, the length of the longest consecutive gap present between all sequences was determined to be 23 residues. One residue was added to each side of the gap length and was used as a final window size to split the full sequence into equally sized sequence segments of 25 residues each. Descriptors were generated in the same fashion for these segments as was done for the different domains in the domain-based data block described above.

##### ***Substructure based***

These datasets are based on the Complementarity-determining regions (CDRs) and Framework regions (FRs) of the hypervariable regions of mAbs. The CDRs and FR regions were identified based on the unique IMGT numbering system (Lefranc, 1997; Lefranc *et al.*, 2003) which relies on anchor residues i.e. highly conserved such as Cys\_23 Trp\_41, a hydrophobic amino acid at position 89, Cys\_104), Phe/Trp\_118 and so on. A standardised delimitation of the framework regions and of the complementarity determining regions based on sequence lengths is also provided by IMGT numbering systems. The sequences were input into Matlab following which algorithms were used to identify the CDR and FR regions based on the rules specified. Further

alignment for substructure sequences was carried out based on additional rules formulated on conserved sequences.

### *Single Amino Acid based*

A multiple alignment of all sequences was performed to align the conserved region of the mAbs. Position in the alignment where amino acids were not conserved and showed high amino acid variability was selected and numerically converted using the Z-scale, T-scale and MS-WHIM. Similar to the window approach, an average of each component in the amino acid scales was calculated based on the residues in a window when gaps were present. To avoid null values, the window was extended with +1 residues to each side of the gap.

### *Running Sum based*

Similar to the single amino acid approach a multiple alignment of all sequences was performed to align the conserved region of the mAbs. A window of 13 residues was selected and all residues present in this window were numerically converted using the amino acid scales and each scale component was summed together. All descriptors were generated by sliding the window in the forward direction in the sequence one residue at the time. As the longest consecutive gap in the multiple sequence alignment of mAbs contained 23 residues, the window size was chosen in a way keeping 2 cases in mind. In the first case, for mAbs that have no residues in that RD window containing the longest gap, descriptors would be encoded as -999 in the Running sum data block due to no numerical conversion being possible. In the second case, for those mAbs that do contain residues in long gap regions, loss of information was minimised as considering the following possibilities; multiple representation of the same residues within the long gap regions are avoided as well as ensuring they are represented at least once in the windows that start at the outer boundaries of the longest gap. For example, let the longest gap be numbered from 2-24. This means that the RD window 1-13 and 12-25 would contain all the information contained in that region. The other windows starting between 2 till 12 (+13) would return null values and be discarded in the data analysis.

#### *5.1.5. Data pre-treatment and variable reduction*

All columns containing null values, coded as -999 in the data blocks were removed. Columns with a standard deviation below 0.0001 were also removed with the assumption of not containing sufficient deviation for the model. All data blocks were auto-scaled before being used in modelling in order to centre the data on zero as well as to scale all descriptors to unit variance. Following descriptor generation, the datasets were curated by removing variables

with null values as well as zero variance variables i.e. descriptors with standard deviation < 0.0001. This is referred to as the ‘Original dataset’. The original datasets were then subjected to V-WSP reduction, an unsupervised variable reduction method which allows for the elimination of a variable based on multicollinearity (Ballabio *et al.*, 2014). This was performed via a grid search to find the correlation coefficient threshold for the descriptors for which the Procrustes index was lower than 0.2. Procrustes index is a statistical measure that allows for the assessment of the degree of comparability between the original and reduced datasets based on informational content. A Procrustes value of 1 indicates complete dissimilarity and 0 indicates that both datasets are identical (Ballabio *et al.*, 2014). The grid search was carried out separately for the descriptors of different domains to avoid removing chance correlations due to the presence of similar amino acids in different domains of the variable regions i.e. VH and VL domains. This dataset is referred to as ‘V-WSP reduced’.

#### **5.1.6. Principal component analysis (PCA)**

PCA is a decomposition technique for dimensionality reduction for multivariate data. PCA identifies linear combinations of the variables or so called principal components which align with the directions of the highest variance within the data block. All components are calculated perpendicular to each other (Bro and Smilde, 2014). The analysis was performed using the PLS Toolbox from Eigenvector Research.

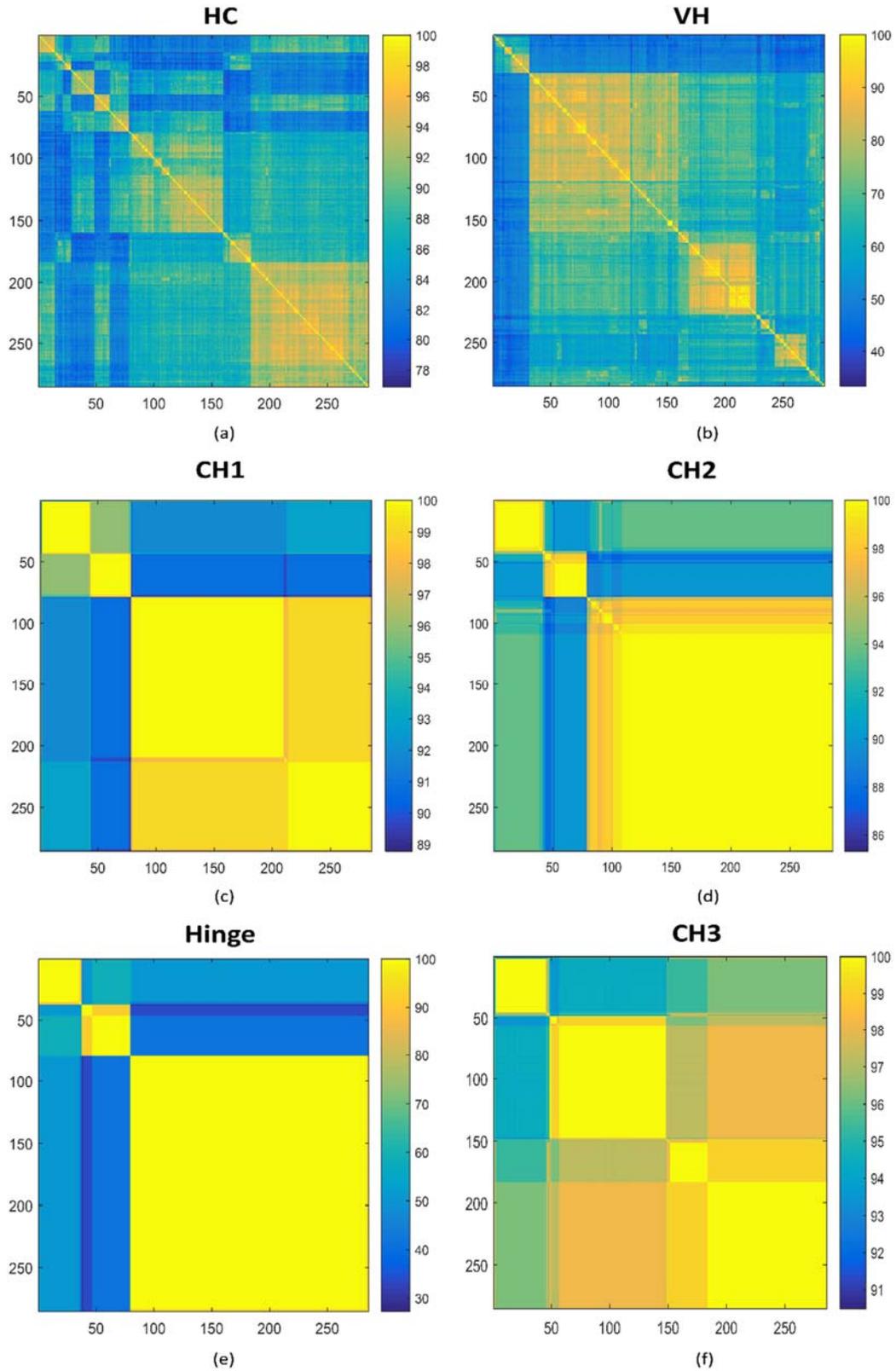
## **5.2. Results**

### **5.2.1. Multiple sequence alignment (MSA)**

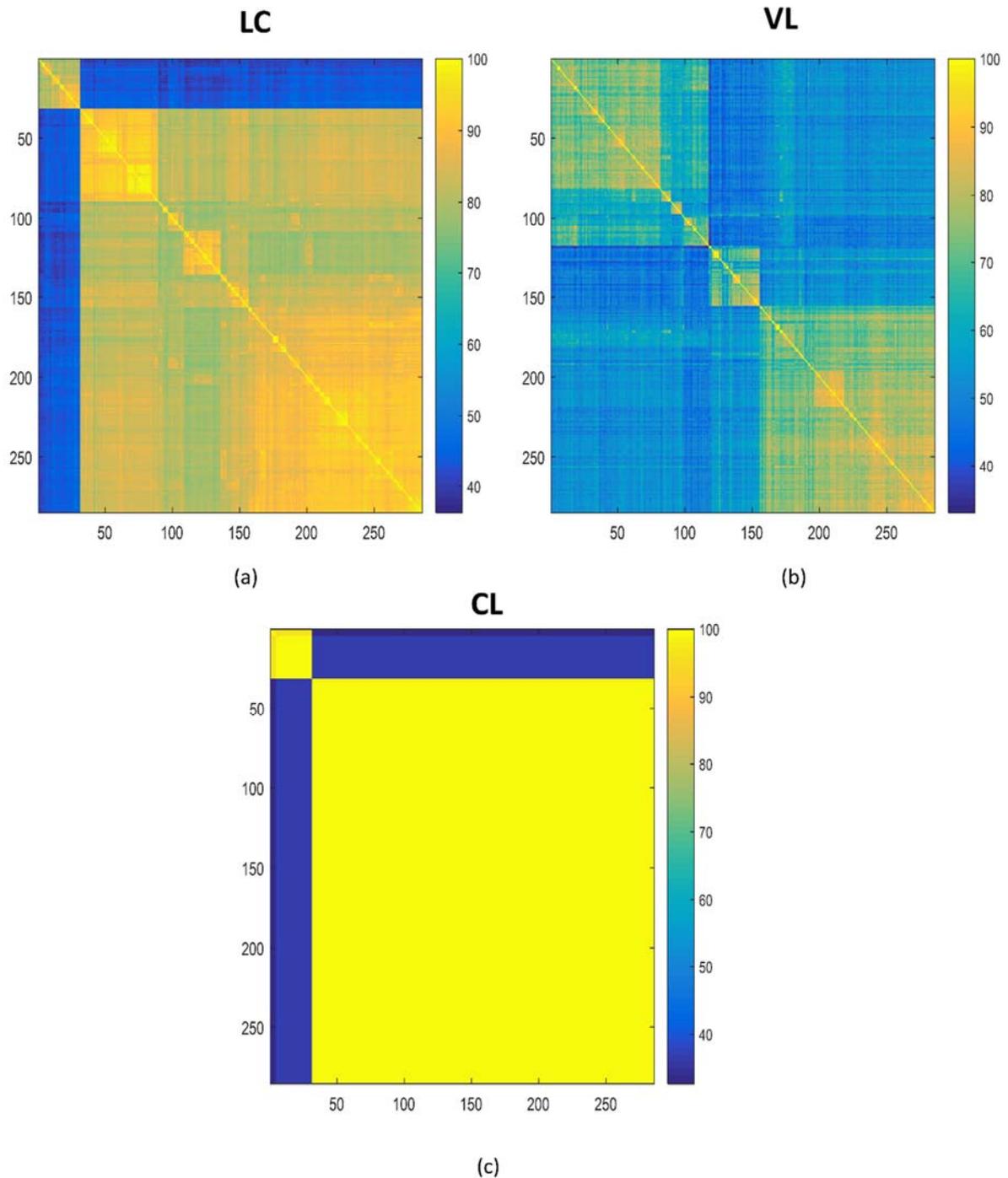
The whole heavy and light chain primary sequence of mAbs were subject to MSA. The percentage identity matrix is derived from percentage sequence similarity of all mAbs sequence. The whole heavy chain sequence was subject to MSA followed by its respective domains: VH, CH1, CH2, Hinge and CH3 (Figure 5.3). The CH1, CH2 and CH3 domains are fairly similar with 89, 86, 91% minimum sequence similarity respectively (Figure 5.3c, d and f). The Hinge region on the other hand has some differences and this is due to the differing lengths between isotypes (Vidarsson *et al.*, 2014b). The major points of dissimilarity between the constant domains for mAbs arises from FC engineering for enhancing their effector function as well as clearance rates based on the therapeutic indication (Kim *et al.*, 2005; Strohl, 2009). Despite Fc engineering, as seen from Figure 5.3c, d and f), the constant domain remains fairly similar. The VH region has a fair degree of dissimilarity as these sequences contain the

hypervariable region for antigen binding i.e. the complementarity determining regions (CDR) and framework regions (FR).

For the light chain sequence, similarly sequence analysis was performed on the complete light chain sequence as well as its constituent domains: VL and CL. The CL region of mAbs seem to have two distinct bands (Figure 5.4c) and this could mainly be from the light chain isotypes, kappa and lambda (Vidarsson *et al.*, 2014b). The VL region is fairly dissimilar as it contains the FR and CDR regions and is involved in antigen binding.



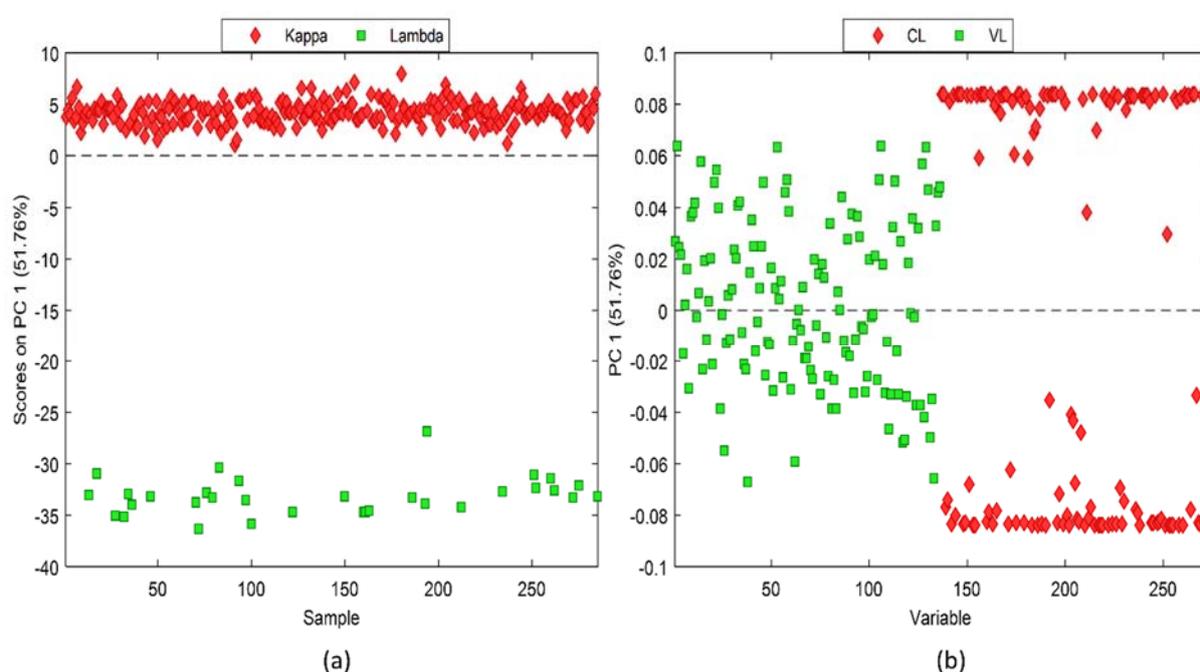
**Figure 5.3** Percentage identity matrix of complete heavy chain sequence and respective domains of 285 mAbs for (a) heavy chain (HC), (b) VH, (c) CH1, (d) CH2, (e) Hinge, (f) CH3. Colour gradient indicates percentage similarity between sequences.



**Figure 5.4** Percentage identity matrix of complete light chain sequence and respective domains of 285 mAbs for (a) light chain (LC), (b) VL and (c) CL, Colour gradient indicates percentage similarity between sequences.

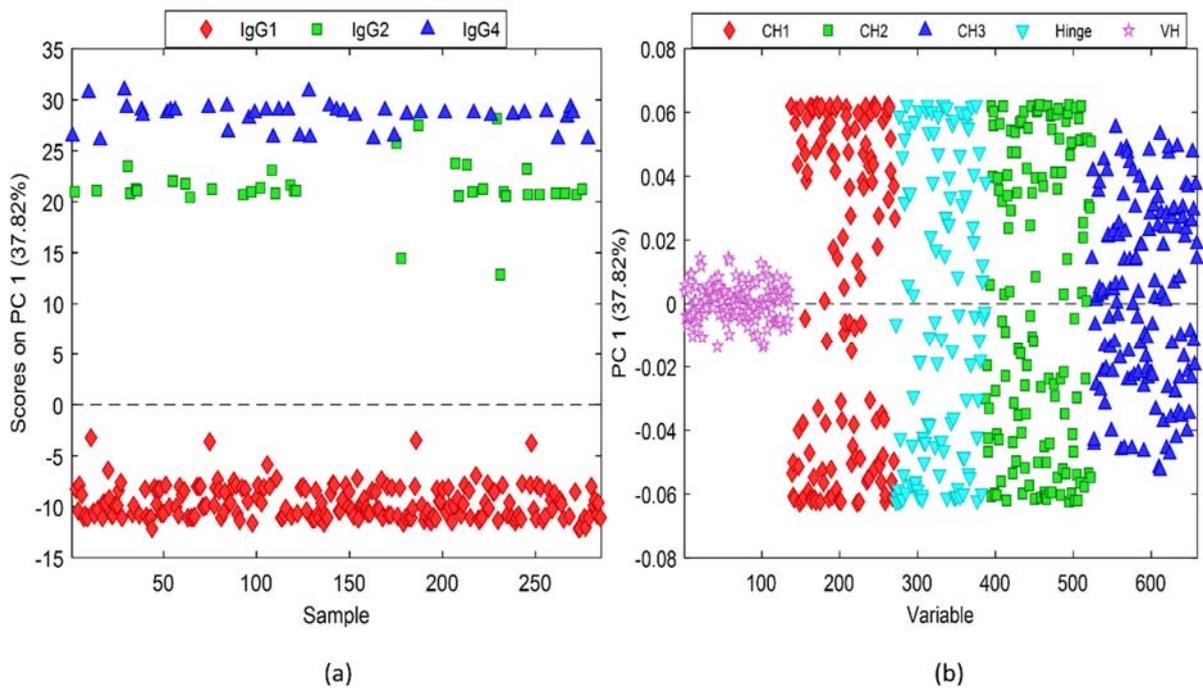
### 5.2.2. Domain specific influence of intrinsic properties

PCA was performed initially using descriptors generated from the primary sequence of the full light chain of mAbs (Figure 5.5). As shown in Figure 5.5a, there is a clear separation of mAbs based on light chain isotypes along PC1 capturing 51.75% of variance depicted in the scores plot. This separation seems to be influenced mostly by CL domain as can be seen from the loading plot (Figure 5.5b). The descriptors from the VL domain influences the separation as well but to a lesser extent when compared to CL. However, from the second PC onwards the loadings of CL descriptors were mostly centred around zero wherein these PCs captured between 2-6% of variance. (Figure 5.5b).



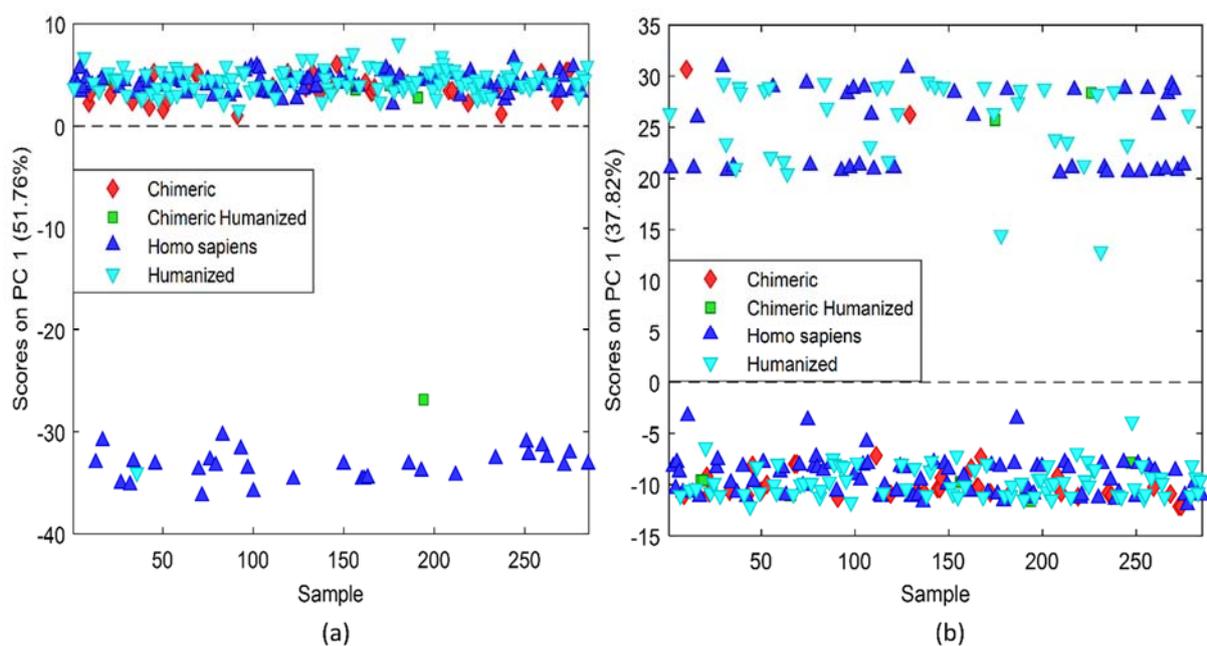
**Figure 5.5** (a) Scores plots and (b) loadings plot following PCA analysis of Domain datablock generated from light chain of mAbs. VL: variable region of the Light Chain; CL: constant region of the Light Chain.

PCA of the heavy chain descriptors of mAbs revealed that there is a clear separation of mAbs based on heavy chain isotype (Figure 5.6a). This separation along PC1, capturing 37.82% variance, is most pronounced for IgG1 from the others. IgG2 and IgG4 are separated as well but to a lesser extent compared to their separation from IgG1 (Figure 5.6a). The separation is most influenced by descriptors of the CH1, Hinge and CH2 domains followed by CH3 domain (Figure 5.6b). VH domain does not appear to influence the separation of mAbs based on heavy chain as the loadings were mainly centred on zero for the PC1 and PC2 (Figure 5.6d). However, from PC3 onwards, while the loadings of all other domains centre around zero, loadings of the VH domain influenced the separation wherein the PC captured only 7% of the variance.



**Figure 5.6** (a) Scores plots and (b) loadings plot following PCA analysis of Domain datablock generated from heavy chain of mAbs. VH: variable region of the heavy chain; CH1: first constant region of Heavy chain; Hinge: hinge region of Heavy Chain; CH2: second constant region of Heavy Chain and CH3: third constant region of Heavy Chain.

No apparent separation between different species of mAbs was observed when descriptors of the heavy or light chain were subjected to PCA (Figure 5.7). As the maximum variance for light chain-based descriptors leads to separation based on light chain isotype and for heavy chain descriptors maximum variance leads to separation based on heavy chain isotype, any species based difference is overshadowed.



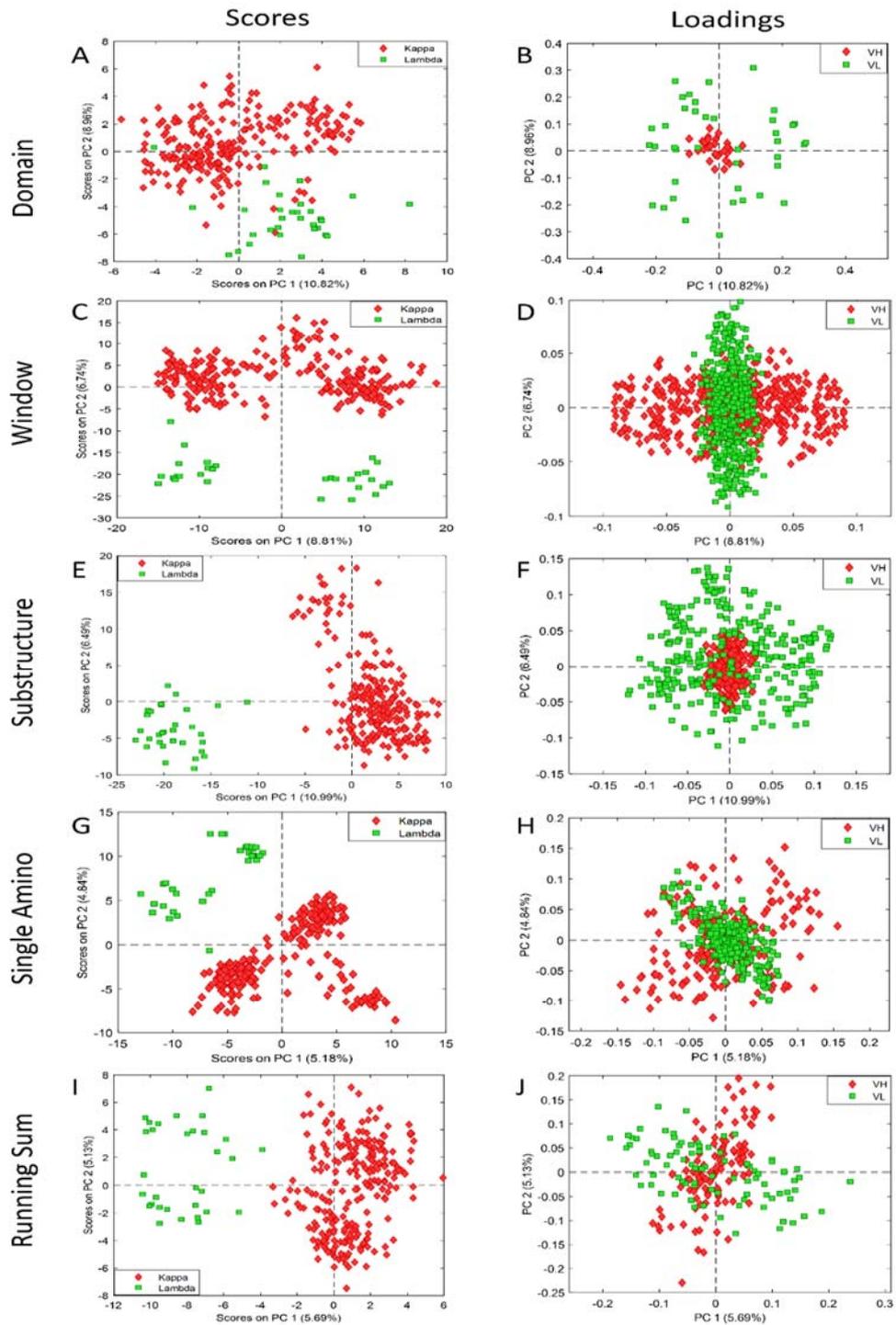
**Figure 5.7** Scores plot generated following PCA analysis of Domain datablock descriptors generated from (a) light chain and (b) heavy chain of mAbs.

### 5.2.3. *Influence of intrinsic properties in hypervariable regions*

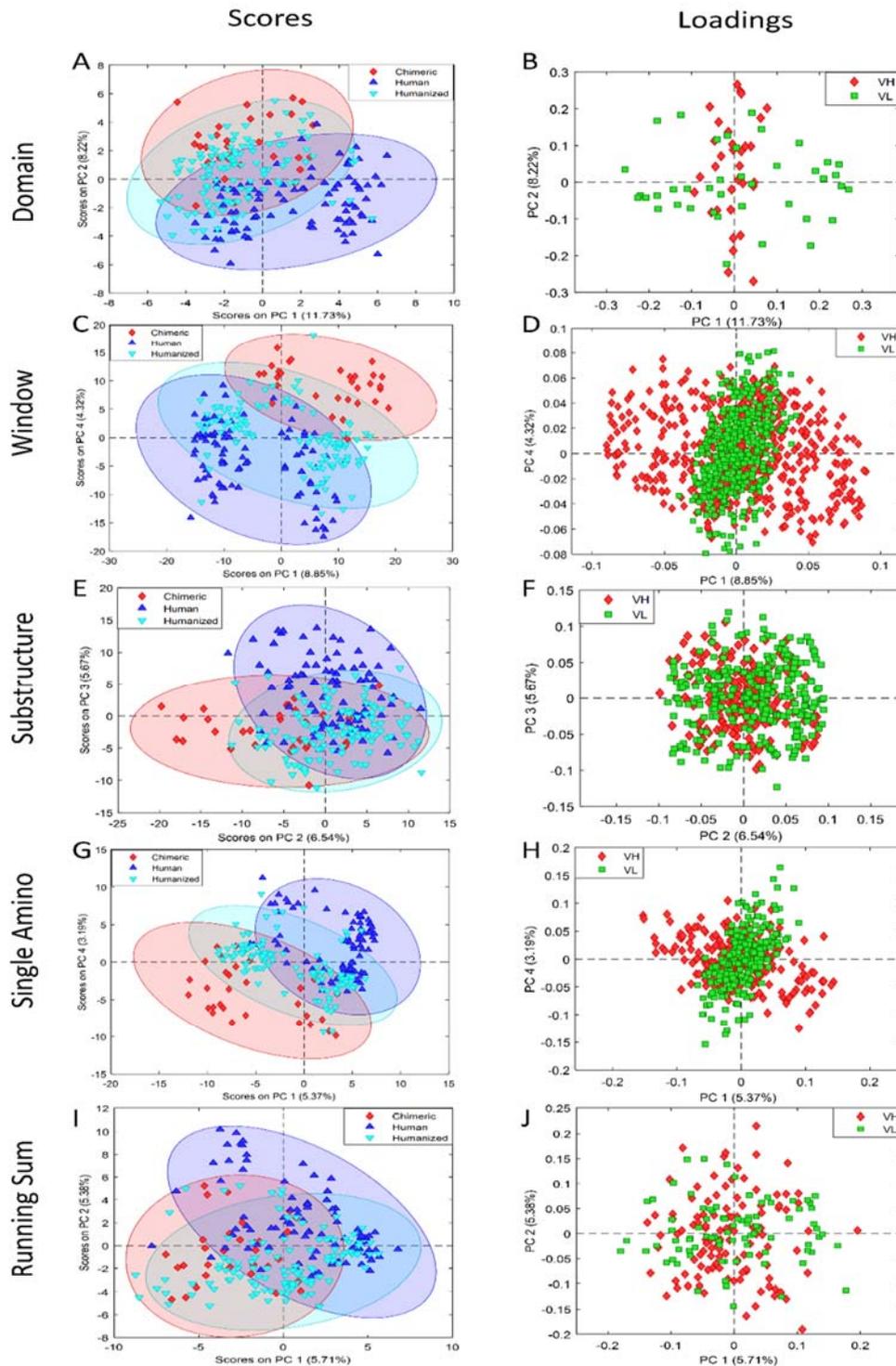
An important aspect of developing QSAR models is defining the applicability domain that will allow for the prediction of new compounds. It is therefore important to identify any potential separation or clustering of samples that could later influence model development. In the case of mAbs, clustering based on intrinsic properties such as light chain substructure, heavy chain substructure and species were observed. Principal component analysis was performed to visualise the separation of mAbs based on their intrinsic properties. As an unsupervised pattern recognition method, PCA allows for the decomposition of the X-block i.e. the primary sequence-based descriptors without the influence of the dependent variables or responses on such a separation. This helps visualise any potential clustering/grouping that might influence the QSAR model at a later stage.

As shown in Figure 5.8a separation based on light chain isotype can be seen for all the data blocks with resolution of the separation increasing with the resolution of the data. The Domain data block is the most global of all the data block and shows the least separation with a dispersion on both PC1 and PC2 (Figure 5.8A). For the Window datablock, the separation of samples based on light chain isotypes appears to be more on PC 2 (Figure 5.8C). Separation based on light chain substructures can also be seen on other PCs however to a lesser extent although these PCs captured <10% variance. All the other data blocks, which are based on either locally clustered amino acids or singular amino acid differences, show more resolution in the separation. The separation is most pronounced for the Substructure datablock with the first two principal component (PC) capturing 19% variance and contribution is mainly from PC1 (Figure 5.8E). For the Single amino acid datablock the separation can be seen on PC2 (Figure 5.8G) whereas for Running Sum it is on PC1 (Figure 5.8I). As seen from Figure 5.8 there is an influence of light chain substructure on the descriptors generated from the primary sequence. As shown in the loadings of the corresponding data block, the variables that influence this separation arises from the VL region. The number samples corresponding to the respective light chain substructure can also influence the analysis as there are greater number of kappa samples than lambda. Based on these observation, the samples were separated and only the Kappa samples (N=253) were retained for further analysis to check for any influence of other intrinsic properties such as species. As all of the Lambda samples were humanized mAbs they were left out of further analysis.

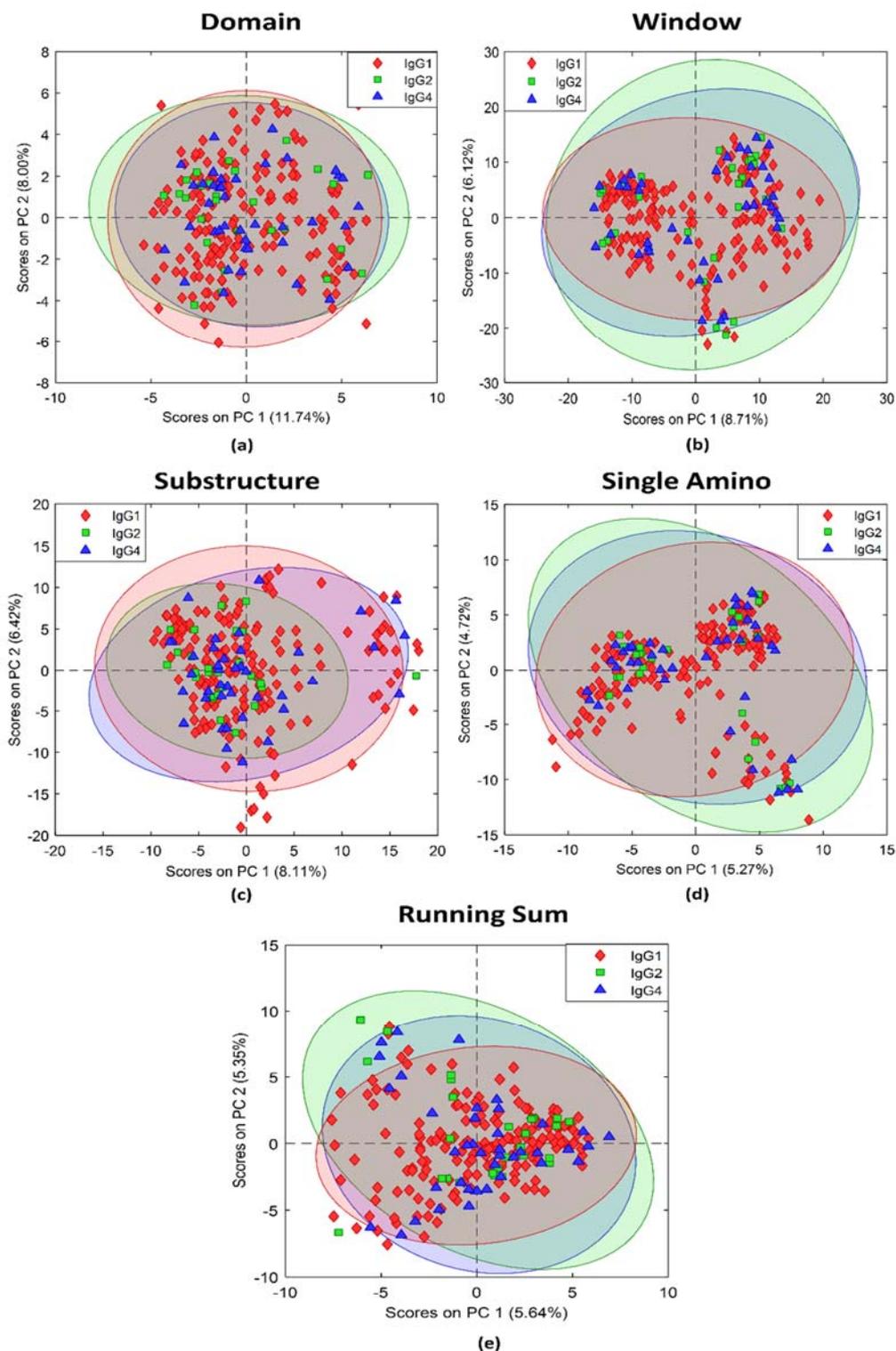
A slightly less apparent separation can be observed based on species of mAbs (Figure 5.9). The coloured ellipses indicate the 95% confidence limits of the corresponding species class and helps in better visualisation of the class behaviour. For the Domain datablock, there is a separation of the chimeric, human and humanized mAbs along PC2. The dispersion is higher for the human from the humanized and chimeric (Figure 5.9A). There seems to be a more mixed influence from the descriptors of both VH and VL domain (Figure 5.9B). The chimeric mAbs are separated farther from the human and humanized mAbs with the separation along PC4 for the Window datablock (Figure 5.9C). This separation is mainly due to the contribution from the descriptors of the VL domain (Figure 5.9D). In the substructure datablock the chimeric is separated from the rest on PC2 and the human mAbs on PC3 (Figure 5.9E). The contribution for this dispersion in the substructure datablock is mixed from descriptors from both VH and VL domain (Figure 5.9F). There is clearer separation of the three species classes in the Single Amino datablock wherein the chimeric, humanized and human mAbs are separated on PC4 (Figure 5.9G). The separation mainly arises from contributions of the VL descriptors (Figure 5.9H). For the Running Sum datablock the resolution of the separation is a bit lower with the human mAbs separated on PC2 (Figure 5.9I). The contributions from the VL and VH descriptors are quite mixed for the Running Sum datablock (Figure 5.9J). No apparent separation of mAbs based on heavy chain isotype was observed from descriptors generated from the hypervariable regions of mAbs (Figure 5.10). This can be seen from the scores plots generated following PCA of VH and VL region descriptors of mAbs (Figure 5.10), the confidence ellipses set at 95% confidence limits mostly overlap among the different isotypes.



**Figure 5.8** Scores plots of (A) Domain, (C) Window, (E) Substructure, (G) Single Amino and (I) Running Sum datablocks and loadings plot of (B) Domain, (D) Window, (F) Substructure, (H) Single Amino and (J) Running Sum generated after performing PCA on data blocks containing only hypervariable region sequences of 285 mAbs. X axis represents Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captured by the respective PC.



**Figure 5.9** Scores plots of (A) Domain, (C) Window, (E) Substructure, (G) Single Amino and (I) Running Sum datablocks and loadings plot of (B) Domain, (D) Window, (F) Substructure, (H) Single Amino and (J) Running Sum generated after performing PCA on data blocks containing only hypervariable region sequences of 253 mAbs. Coloured ellipses indicate the 95% confidence limits of the corresponding class. X axis represents Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captured by the respective PC.



**Figure 5.10** Scores plots of (a) Domain, (b) Window, (c) Substructure, (d) Single Amino and (e) Running Sum generated after performing PCA on data blocks containing only hypervariable region sequences of 253 mAbs. Coloured ellipses indicate the 95% confidence limits of the corresponding class. X axis represents Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captured by the respective PC.

### 5.3. Discussion

This study mainly focused on the generation of descriptors derived from the primary sequence of monoclonal antibodies. These descriptors were then analysed using unsupervised pattern recognition method, PCA, to assess the influence of intrinsic properties such as light and heavy chain isotypes as well as species on these descriptors. This could affect QSAR model development as any apparent grouping or clustering descriptors based on intrinsic properties could confound the model. Additionally this could also affect defining the applicability domain of QSAR models i.e. the scope of models to make reliable predictions (Roy *et al.*, 2015). A well-defined applicability domain has an optimum level of generalisation as well as restriction allowing for better understanding of the compound space and better predictions (Netzeva *et al.*, 2005).

The MSA of primary sequences of mAbs reveals a high degree of similarity between constant domains of both heavy and light chain (Figure 5.3, 5.4). The main differences arise from the different heavy chain isotype (CH1, CH2 Hinge and CH3) as well as the light chain isotype (CL) of mAb samples. Thus, the descriptors generated from these domains would capture information that contributes to the difference in isotype and confound any correlation with the actual output during predictive model development. This was further examined by assessing descriptors generated from the Domain datablock for the whole heavy and light chain as well as their constituent domains (Figure 5.5 and 5.6). For the light chain the first PC always captured the variance between the light chain isotype and the main contribution for this arose from the CL domain followed by VL (Figure 5.5). A similar trend was observed for the heavy chain descriptors wherein the first PC captured the difference between different heavy chain isotypes and this was influence mainly form descriptors in the CH1, CH2, hinge and Ch3 domains (Figure 5.6). The VH region had very little influence in the separation based on heavy chain isotype. The separation was highest for IgG1 mAbs when compared to IgG2/IgG4 (Figure 5.6). Due to the high structural similarity between IgG2 and IgG4 monoclonal antibodies, only a few residues contribute to the structural difference when comparing the isotypes (Vidarsson *et al.*, 2014a). Based on these preliminary investigations it can be concluded that the descriptors generated from the constant domain will always be influenced by intrinsic properties and this becomes an issue in multivariate data driven regression models such as Partial least squares which works on maximising the covariance between descriptors and response. This has led to many QSAR studies in the area of mAb development to utilise

only the hypervariable regions for further model development (Yadav *et al.*, 2011; Hötzel *et al.*, 2012; Sharma *et al.*, 2014; Robinson *et al.*, 2017).

The descriptors generated from the hypervariable regions of mAbs for all datablocks (Domain, Window, Substructure, Single amino and Running sum) were further explored to assess the influence of any intrinsic properties. The VL region descriptors indeed influenced the separation in light chain isotypes for all datablocks with separation increasing with data resolution (Figure 5.8). The contribution from the VH region had minimal influence on light isotype based separation. The hypervariable descriptors were further assessed for influence of species following the removal of Lambda mAbs. The hypervariable region descriptors were influenced by the species type as well (Figure 5.9) to a slightly lesser extent than light chain isotype. This makes sense as the constant domains in all the antibodies that were investigated were of human origin, while the actual species conformation/properties is only present in the variable domains (Kim *et al.*, 2005). The heavy chain isotype did not influence the descriptors generated from the VH and VL region (Figure 5.10). Thus, it would be prudent to separate mAbs based on light chain isotype as well as species before predictive model development as they have a clear influence on the descriptors generated. This has two advantages as it removes any influence of intrinsic properties as well as allowing for a development of a predictive model with a well-defined applicability domain that facilitates better prediction and interpretability.

Aspects which contribute to isotype separation might be different from those that contribute to bioactivity. Thus, selection of domains as well as descriptor sets has to be performed in such cases to improve the performance of the regression/classification models which will allow for better prediction by removing the influence of redundant or noisy data. Different machine learning based predictive models have already been developed for predicting aggregation tendencies of mAbs in manufacturing as well as for epitope mapping studies (Singh *et al.*, 2013; Obrezanova *et al.*, 2015).

Evaluations of safety, manufacturability and efficacy allow for early risk assessment for developing and improving mAb based therapeutics. The main advantages of this methodology are that it is high throughput, requires only the primary sequence; translatable to all mAbs based therapeutics. The main challenges for further development of such predictive tools would be to capture the structural information such as hydrophobic and hydrophilic patches, solvent accessible surface area, ligand binding pockets which could greatly aid not only in better selection of lead candidates but also aid in better design of the process in accordance with the

QbD paradigm. 3D structures of approved mAbs are not readily available due to the highly competitive nature of the biopharmaceutical industry. Homology modelling and molecular dynamics techniques can be utilised to overcome this challenge. This has been previously used for the development of QSAR models for chromatographic techniques using the surface properties of fab variants (Robinson *et al.*, 2017). This methodology has also been used to predict the critical product properties of mAb based therapeutics such as clearance, viscosity and chemical stability (Agrawal *et al.*, 2016). These techniques can be further extended to predict the safety or adverse effect profile of mAbs using docking studies. This has been previously demonstrated for smaller peptide-based QSAR models (Lin *et al.*, 2017).

Development of such *in silico* tools will have a significant impact on all stages of the mAb production including ascertaining the critical quality attributes of product and process which leads to better product and process design. These methods also allow for seamless integration into existing processes for better profiling of the product, thereby reducing attrition, as well as reducing manufacturing failures owing to the complexities of the mAb based therapeutics.

#### **5.4. Chapter summary**

In conclusion, this chapter elucidates the development of five novel descriptor datablocks: Domain, Window, Substructure, Single Amino and Running sum datasets, generated from the primary sequence of mAbs that could be used for QSAR model development of mAbs. Unsupervised pattern recognition via PCA was performed on the descriptor sets to visualise any intrinsic property-based separation or clustering. As the method is unsupervised it facilitates detection of any intrinsic groupings or clusters. The descriptor sets were assessed for the influence of intrinsic properties that could hamper with predictive model development that facilitates mAb developability. Based on the results, the descriptors of the hypervariable region were chosen for further model development. Furthermore, the samples were divided into appropriate heavy chain, light chain and species type for developing models that have a better-defined applicability domain.

## **Chapter 6: Cross interaction chromatography based QSAR model for early stage screening to facilitate enhanced developability of monoclonal antibody therapeutics.**

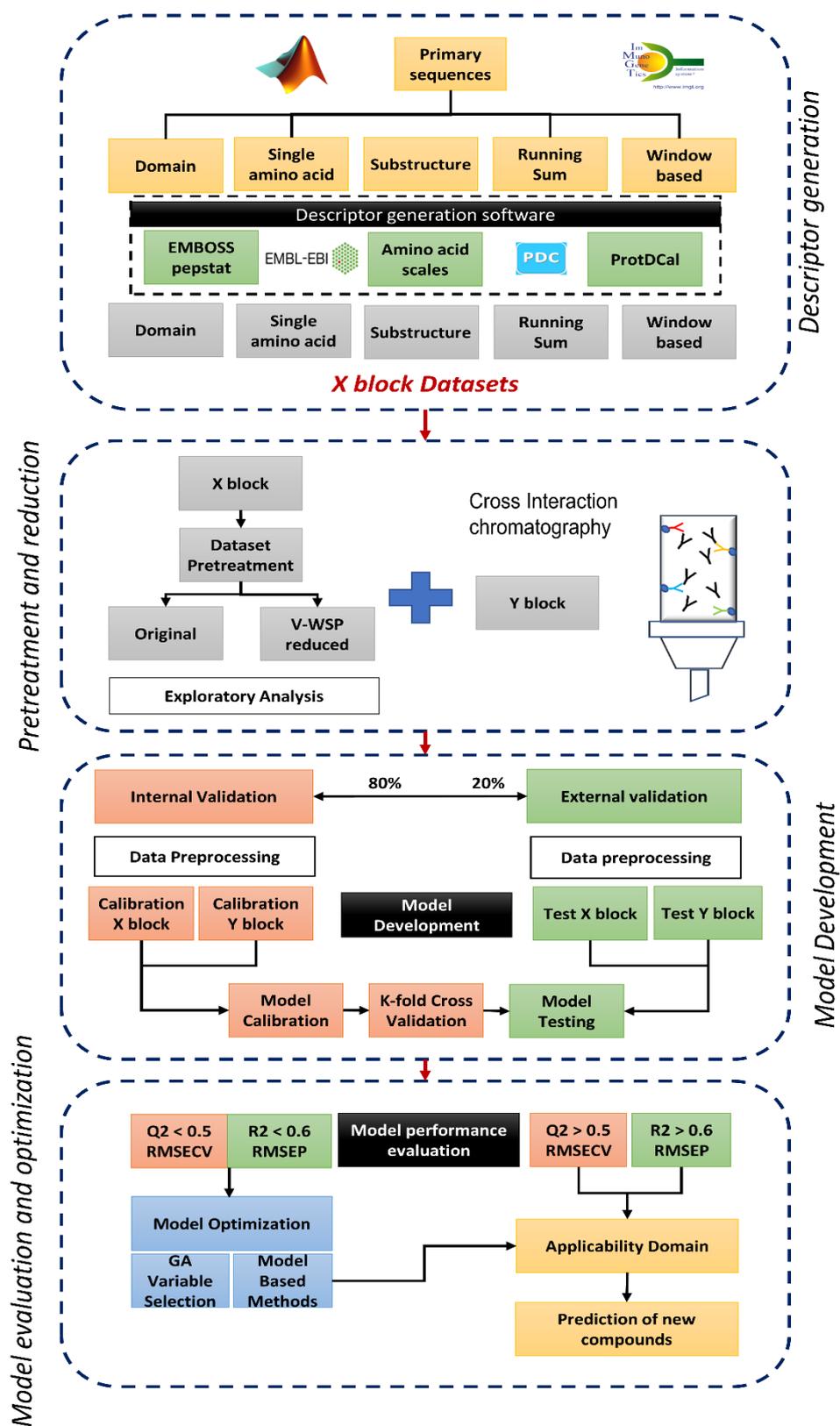
The growth in the area of monoclonal antibody-based therapeutics, is impeded by developability issues regarding the biophysical properties of mAbs that cause them to fail either at the manufacturing stage or in clinical trials. With more mAbs being administered subcutaneously, higher concentration formulations of mAbs are required so as to administer higher amounts in smaller volumes (Pindrus *et al.*, 2015). The two main problems associated with high concentration mAb formulations are 1) increased aggregation propensities that could lead to immune related adverse effects and 2) poor solubility causing problems during administration (Yadav *et al.*, 2011; Pindrus *et al.*, 2015; Joubert *et al.*, 2016).

Thus, it is beneficial to screen mAbs for developability attributes such as aggregation propensity and solubility issues at the discovery stage. There exists a battery of assays to measure these biophysical properties which usually assess the potential of mAb interactions with itself (self-associate) and/or with other antibodies (cross interaction) (Xu *et al.*, 2013; Jain *et al.*, 2017). One such assay is Cross interaction chromatography (CIC), an early stage screening technique utilised during affinity and activity characterisation of mAbs. It assesses weak cross interaction of mAbs with polyclonal antibodies when eluting through a column coupled with human serum polyclonal antibodies (Jacobs *et al.*, 2010; Kelly *et al.*, 2015). Longer retention times indicate increased tendencies of nonspecific interactions. This could lead to issues in developability such as polyspecificity, low solubility and higher clearance rates (Jacobs *et al.*, 2010; Kelly *et al.*, 2015). CIC has multiple advantages over other techniques that estimate protein solubility and polyspecificity (Jacobs *et al.*, 2010). As demonstrated in previous studies it requires very low amount of protein for analysis and allows for rapid high throughput screening of mAbs at their discovery stage which could help in minimising cost and time of analysis (Jacobs *et al.*, 2010; Jain *et al.*, 2017). However, considering the scale and number of mAbs in development and the costs involved, there is an urgent need for *in silico* tools that could allow for the prediction of protein behaviour that would allow for prioritisation of lead candidates to be tested in predictive assays.

Quantitative sequence activity/property relationship (QSAR/QSPR) models are powerful *in silico* tools that allow for the prediction of mAb behaviour where the descriptors (based on structure, sequence and/or interaction) are modelled against a biological activity (response) using multivariate data analysis approach for approved mAbs. This is an extension of

Quantitative structure activity/property relationship models which are routinely used for the development of predictive models in biopharmaceutical development and toxicological applications (Zhao *et al.*, 2007; Zhou *et al.*, 2008; Matthews *et al.*, 2009; Nongonierma and FitzGerald, 2016). QSAR models have been developed for a number of chromatographic applications for a myriad of proteins at different developmental stages (Song *et al.*, 2002; Woo *et al.*, 2015). Descriptors for protein molecules can be generated by different software, such as EMBOSS and ProtDcal, both of which are freely available with a friendly graphical user interface and the capacity to generate a higher number of molecular descriptors for proteins from FASTA or PDB files (Ruiz-Blanco *et al.*, 2015). Different sets of protein descriptors have been utilised in studies to distinguish between protein functional families where models were developed based on machine learning methods such as Support vector machines and artificial neural networks (Ong *et al.*, 2007; Li *et al.*, 2016). Among these methods is Partial least square (PLS) regression, a combination of principal component analysis and multiple linear regression. In this method a linear combination of new predictor variables is created from the original variable dataset, which has a large covariance with the response variable. It takes into consideration the correlations and variances of the responses and predictors (Geladi and Kowalski, 1986; Ng, 2013).

In this study, development of a robust QSAR model was extended using five novel datasets generated from the primary sequence of mAbs and they were modelled against corresponding CIC retention time based on the data recently published by Jain *et al.*, (Jain *et al.*, 2017). As the mAbs tested in their study had different variable domains but the same IgG1 constant domain descriptors were generated only for the variable domains of mAbs i.e. for the VH and VL regions (Veerasamy *et al.*, 2011). In this study descriptors were generated using ProtDcal, EMBOSS Pepstat software as well as amino acid scales for variable regions of mAbs, the details of which are outlined in Table S1. Five independent and novel X block datasets consisting of these descriptors were generated based on the physicochemical, electronic, thermodynamic, electronic and topological properties of amino acids: Domain based, Window based, Substructure based, Single Amino Acid based, and Running Sum based. The relationship between the descriptors and the responses was captured by Partial Least Squares based models and the datasets as well as models were benchmarked based on metrics as outlined in QSAR validation guidelines (Roy *et al.*, 2015). Furthermore, the models were validated with an external test of mAbs and the corresponding model performance metrics were benchmarked for the different datasets. The overall methodology is outlined in Figure 6.1.



**Figure 6.1** Hybrid model development workflow outlining the different steps involved in descriptor generation, pre-treatment and variable reduction; model development followed by model evaluation and optimisation. MSA: Multiple Sequence Alignment.

## **6.1. Materials and Methods**

### **6.1.1. Data collection**

Sequence information, substructure, species and phase of development of 137 mAbs have been collected from IMGT, literature, patents as well as from industrial partners (Apweiler *et al.*, 2004; Lefranc *et al.*, 2009). An overview of these 137 mAbs have been provided in Appendix D Table D.1. The FASTA format sequences of mAbs were read into Matlab and were subjected to multiple sequence alignment. The sequences were then split into domains and the variable region domains VH and VL were selected for further descriptor generation.

### **6.1.2. Descriptor generation**

The primary sequences of the antibodies were imported into MATLAB 2015a. The details of descriptor generation have been outlined in Chapter 5, Section 5.1.3 and the data block generation has been described in Chapter 5, Section 5.1.4. The automation of descriptor and dataset generation was done by MATLAB codes written by Micael Karlberg, Newcastle University.

### **6.1.3. Data curation and Variable reduction**

Following descriptor generation, the datasets were curated by removing variables with null values as well as zero variance variables i.e. descriptors with standard deviation  $< 0.0001$ . This is referred to as the ‘Original dataset’. The original datasets were then subjected to V-WSP reduction, an unsupervised variable reduction method which allows for the elimination of variable based on multicollinearity (Ballabio *et al.*, 2014). This was performed via a grid search to find the correlation coefficient threshold for the descriptors for which the Procrustes index was lower than 0.2. Procrustes index is a statistical measure that allows for the assessment of the degree of comparability between the original and reduced datasets based on informational content. A Procrustes value of 1 indicates complete dissimilarity and 0 indicates that both datasets are identical (Ballabio *et al.*, 2014). The grid search was carried out separately for the descriptors of different domains to avoid removing chance correlations due to the presence of similar amino acids in different domains of the variable regions i.e. VH and VL domains. This dataset is referred to as ‘V-WSP reduced’.

### **6.1.4. Exploratory Analysis**

Principal component analysis, an unsupervised pattern recognition method, was performed on the reduced datasets: Domain, Window, Single amino and Running Sum using the PLS

Toolbox from Eigenvector Research (Bro and Smilde, 2014). The data was auto scaled (each descriptor was mean centred and scaled with its individual standard deviation) prior to analysis. Corresponding score plots were then assessed to visualise the separation, if any, of mAbs based on their intrinsic properties, such as heavy chain substructure, light chain substructure as well as species.

#### ***6.1.5. Cross Interaction chromatography data***

The cross interaction chromatography data for 137 mAbs were obtained from a previous study performed by Jain et al., where 30 mg of human serum polyclonal antibodies was coupled to a column followed by testing approximately 5 µg of each antibody at a flow rate of 0.1 mL/min using PBS as a mobile phase on an HPLC system (Jain *et al.*, 2017). This data was auto scaled prior to modelling.

#### ***6.1.6. Model development***

The data was divided into a training and test set via Kennard Stone algorithm, maintaining an 80%-20% split of data (Galvao *et al.*, 2005). This division was based on the small size of the dataset i.e. to retain sufficient samples in the training set whilst having a small dataset. The mAbs with the most dissimilarity in Euclidean space are placed into the training set. CIC retention times were modelled against the descriptors generated using the PLS Toolbox from Eigenvector Research. Both the descriptors and responses of the training set were entered into Matlab, auto scaled separately i.e. for the training set the mean and standard deviation of the training set was used and for the test set the mean and standard deviation of the test set was used. The datasets were consequently subjected to PLS analysis. The cross-validation method used was a random subset cross validation with 5 splits and 20 iterations (Minitab, 2014; Hahn and Valentine, 2016). A maximum of 10 latent variables were tested for each developed model and used to investigate the model error to choose an appropriate model complexity, i.e. number of latent variables. Models were developed first for the original datasets followed by the V-WSP reduced datasets. Models were further optimised by Genetic Algorithm-Partial least squares (GA-PLS) based supervised variable selection method. The parameters were set as follows: population size of 100; maximum generations of 100; mutation rate of 0.005, window width of 1; convergence rate of 80%; 30 initial terms, cross over of 2; random subset 5 fold cross validation with 10 iterations and data retreatment set for autoscaling (Hasegawa *et al.*, 1997; Andersen and Bro, 2010). For each stage of model development and improvement i.e. Original, V-WSP and GA-Selected, the model was assessed with the test set.

### 6.1.7. Model Performance metrics

The datasets, models and outputs were benchmarked based on the following metrics for the multivariate regression analysis:  $R^2$  values based on Pearson correlation coefficient between observed and predicted values (Equation 6.1);  $Q^2$  values based on goodness of fit (Equation 6.2); Root mean square error of cross validation (RMSECV) (Equation 6.3) and cross validation bias (Equation 6.4) (Alexander *et al.*, 2015). The model metrics were further assessed against the OECD guidelines for the Calibration, internally validated (Cross validation) and externally validate models (test set predictions) (Veerasamy *et al.*, 2011; Organisation for Economic and Development, 2014). For a dataset with  $n$  samples containing observations  $y_1$  to  $y_n$ , each associated with a predicted /modelled value  $f_1$  to  $f_n$  wherein  $y_i$  is a sample within the dataset and  $f_i$  is its associated predicted/modelled value. SS refers to Sum of squares for the estimated ( $SS_E$ ) and true values ( $SS_T$ ) and summation is denoted by  $\Sigma$ .

$$R^2 = \left( \frac{n \Sigma_{i=1}^n f_i y_i - \Sigma_{i=1}^n f_i \Sigma_{i=1}^n y_i}{\sqrt{n \Sigma_{i=1}^n f_i^2 - (\Sigma_{i=1}^n f_i)^2} \sqrt{n \Sigma_{i=1}^n y_i^2 - (\Sigma_{i=1}^n y_i)^2}} \right)^2 \quad (\text{Equation 6.1})$$

$$Q^2 = 1 - \frac{SS_E}{SS_T} = 1 - \frac{\Sigma_{i=1}^n (f_i - y_i)^2}{\Sigma_{i=1}^n (y_i - \bar{y})^2} \quad (\text{Equation 6.2})$$

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^n (f_i - y_i)^2}{n}} \quad (\text{Equation 6.3})$$

$$bias = \frac{\Sigma_{i=1}^n f(x)_i - y_i}{n} \quad (\text{Equation 6.4})$$

### 6.1.8. Y randomisation

The developed models were further evaluated by performing Y randomisation (Y scrambling) to investigate whether the models did indeed capture the underlying relationships in the data rather than capturing chance correlation between the independent descriptors and dependent response data (Rücker *et al.*, 2007). This was done by randomly rearranging the order of the response data while keeping the original order in the descriptor set, thus creating a mismatch between the input and output data. A model was then developed according to specifications given above in the Model Development section. The procedure was repeated 50 times to test different random permutations of the response values and averages of  $R^2$  and  $Q^2$  from the cross-validation results from all repetitions were calculated. In addition, statistical tests based on Wilcoxon Signed Rank test, Sign test and Student t-test were performed by pairwise comparison of the predictions of the randomised model to that of the non-randomised model

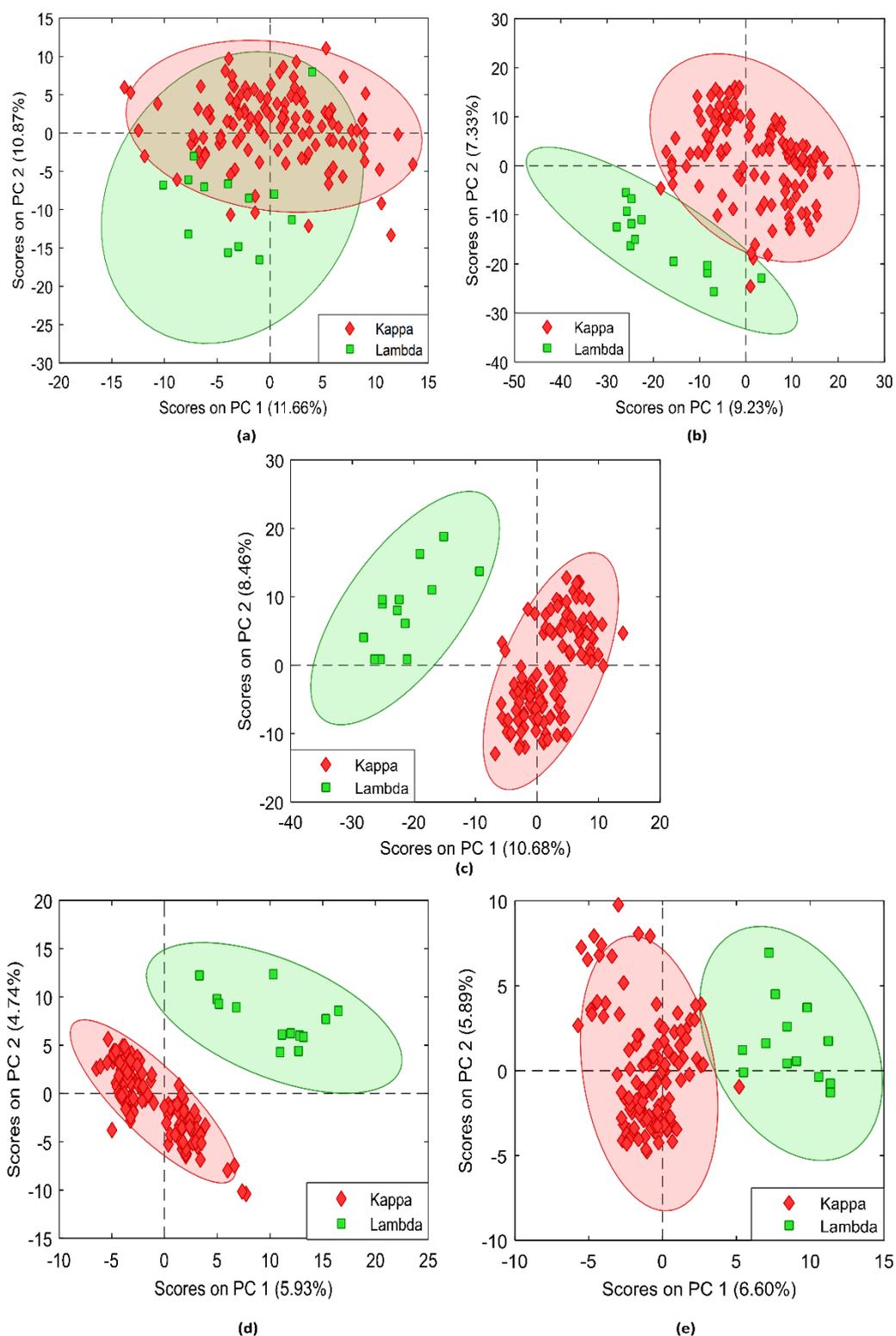
(Thomas, 2003). A significance level of 0.05 was used to reject or accept the null hypothesis of predictions from the randomised model being significantly different from the predictions of the non-randomised model (below 0.05).

## **6.2. Results**

### **6.2.1. Exploratory analysis**

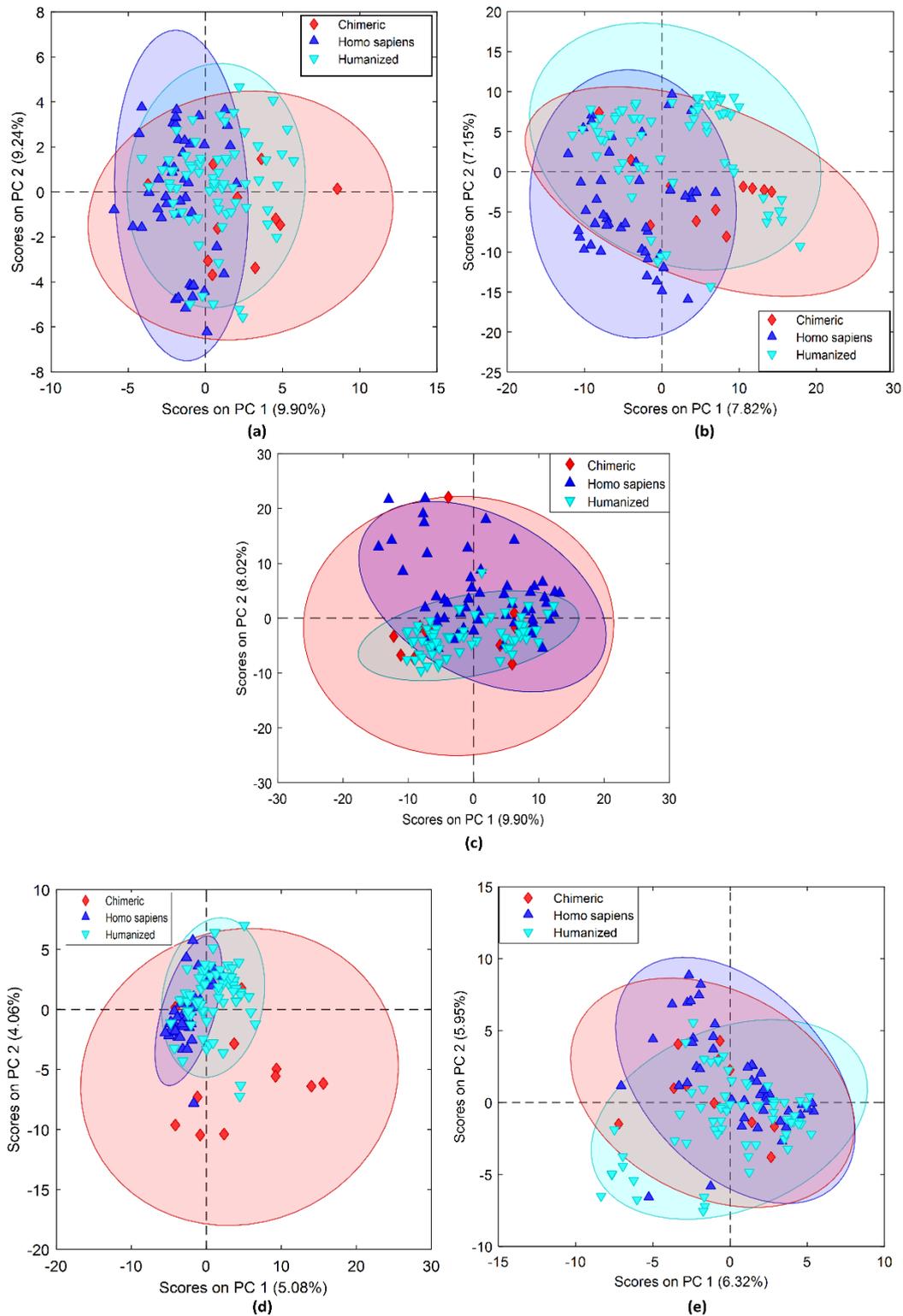
An important aspect of developing QSAR models is defining the applicability domain that will allow for the prediction of new compounds. It is therefore important to identify any potential separation or clustering of samples that could later influence model development. In the case of mAbs, clustering based on intrinsic properties such as light chain substructure, heavy chain substructure and species were observed. Principal component analysis was performed to visualise the separation of mAbs based on their intrinsic properties. Being an unsupervised pattern recognition method, PCA allows for the decomposition of the X-block i.e. the primary sequence based descriptors without the influence of the dependent variable or response.

A clear separation between samples containing kappa and lambda chain based on primary sequence descriptors can be observed in Figure 6.2. This is evident in all datasets: Domain, Window, Single and Running Sum. The separation is in the first PC as the highest amount of variance is between kappa and lambda light chain substructures. At this stage only those mAbs with a kappa light chain were selected (n=124) and lambda light chain mAbs were removed from the datasets (n=13).



**Figure 6.2** Score plots generated after performing PCA on (a) Domain, (b) Window, (c) Substructure, (d) Single Amino acid and (e) Running Sum datasets for 137 mAbs. X axis represent Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captures by the respective PC. Coloured ellipses indicate the 95% Confidence limits of the corresponding class.

A slightly less apparent separation can be observed based on species of mAbs (Figure 6.3). The coloured ellipses indicate the 95% Confidence limits of the corresponding species class and helps in better visualisation of the class behaviour. This is particularly pronounced for chimeric mAbs when compared to *Homo sapiens* (human) and humanized. As reflected in the influence plot (Appendix D Figure D.1) the chimeric species has higher Q residuals, a lack-of-fit statistic, than the rest. Another factor influencing this could be the sample sizes which are 67, 46 and 11 samples for humanized, *Homo sapiens* and chimeric species respectively. As PLS models are based on covariance matrix, the variance in the X block which attributes to the difference in intrinsic properties, influences the model development. Only the humanized mAbs were chosen at this stage (n=67).



**Figure 6.3** Score plots generated after performing PCA on (a) Domain, (b) Window, (c) Substructure, (d) Single Amino acid and (e) Running Sum datasets for 67 mAbs. Coloured ellipses indicate the 95% Confidence limits of the corresponding class. X axis represent Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captures by the respective PC.

Most of the residues attributing for the difference of IgG heavy chain substructures are present in the constant region. The number of samples for IgG2, IgG4 and IgG1 are 7, 13 and 46 samples respectively and this would influence the spread as sample descriptor spaces are sparsely and varying populated based on species. Based on the above results, IgG1-Kappa-Humanized mAbs were chosen for further model development (n=46) such that the QSAR model developed is for a homologous set of mAbs which has been seen in previous studies as well (Robinson *et al.*, 2017). Furthermore, the other species of mAbs have too little samples in them to develop an applicability domain around chimeric or human mAb. The details of the mAbs chosen are shown in Appendix D Table D.2.

### 6.2.2. Variable Reduction and variable Selection

Considering the dimensionality of the data with many variables and the limited number of observations, variable reduction and variable selection methods were used prior to model building. There was a 78%, 31%, 87% and 95% reduction in variables of the domain, window, single amino acid and running sum datasets, respectively, following V-WSP reduction. The associated correlation coefficient threshold as well as Procrustes indices are reported in Appendix D Table D.3. The V-WSP reduced datasets were then divided into training (80%) and test set (20%). Furthermore, GA-PLS based variable selection led to a further reduction of variables resulting in approximately 4%, 14%, 25 and 1% of the variables from the original dataset being retained in the GA selected dataset for domain, window, single amino acid and running sum datasets, respectively. The number of unique models generated, the average RMSE values of these models as well as the number of generations required to reach convergence are reported in Appendix D Table D.4. The total number of variables retained following variable reduction and variable selection is shown in Table 6.1.

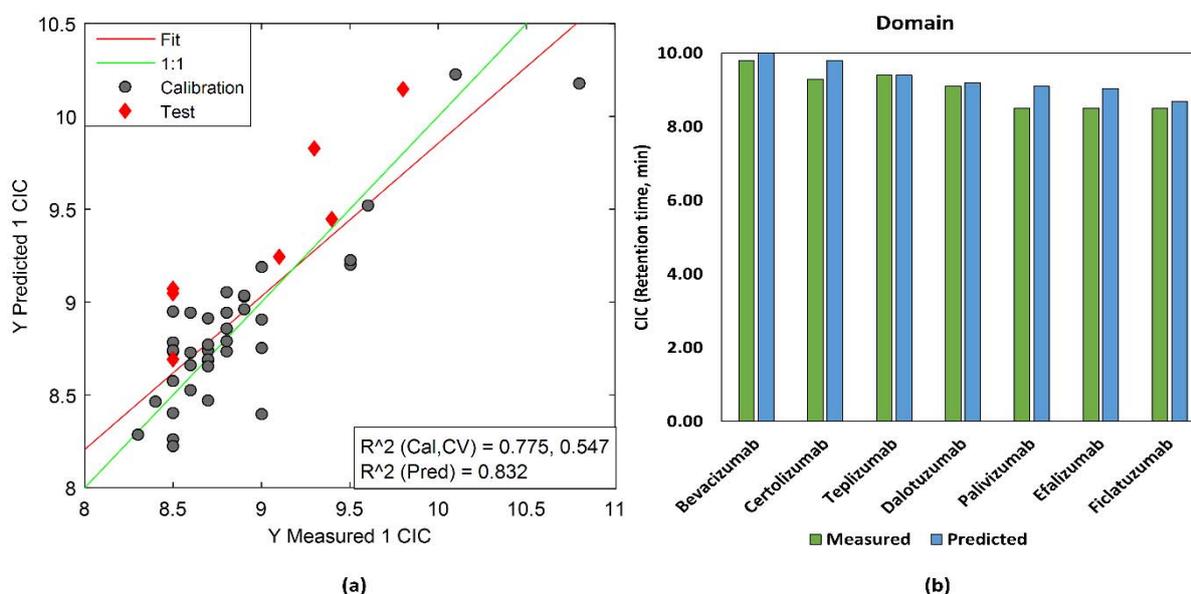
**Table 6.1** Dimensionality reduction by variable reduction and selection

<b>Dataset</b>	<b>Original</b>	<b>V-WSP reduced</b>	<b>GA selected</b>
Domain	272	61	11
Window	1336	921	191
Substructure	639	193	43
Single amino	1474	191	36
Running sum	2596	140	29

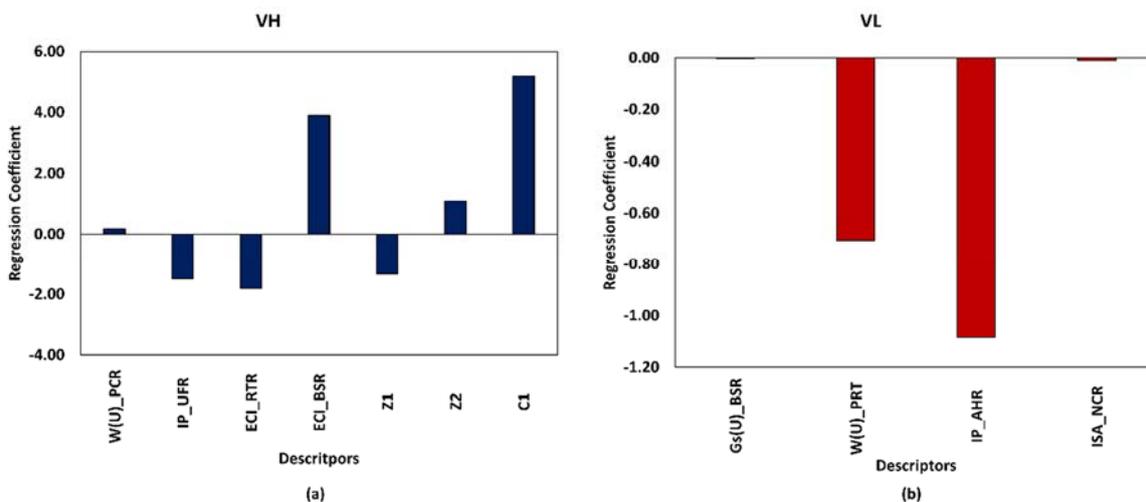
### 6.2.3. Model performance evaluation

#### Domain dataset

Figure 6.4 shows the overall model summary for the GA variable selected PLS model (5 LVs) for the domain dataset. There is more dispersion of the samples from the parity line in the lower extreme of the data i.e. for lower retention times and this is where most samples lie due to the skewed nature of the data distribution (Figure 6.4a). Duligotuzumab was removed from the test set as it had a high Y residual (Appendix D Figure D.2a). This indicates that the Y value for this sample is different for the group of samples for which duligotuzumab shares a high degree of sequence similarity and thus possessing similar values for descriptors. Most of the test set predictions fall within 30 seconds of the experimental value i.e. within a 5% range of the measure values (Figure 6.4b). Palivizumab and efalizumab however have been overestimated by 36 seconds and 32 seconds respectively (Figure 6.4b). Figure 6.5 shows the regression coefficients for the variables chosen in the model following GA based variable selection. In VH region the highest absolute values for regression coefficients are those of electronic properties of the residues in  $\beta$  sheet structure and the size associated electrostatic potential of VH domain (Figure 6.5a). In the VL region the number of water molecules close to residues in unfolded state and the isoelectric point of alpha helix structures have the highest absolute regression coefficients (Figure 6.5b).



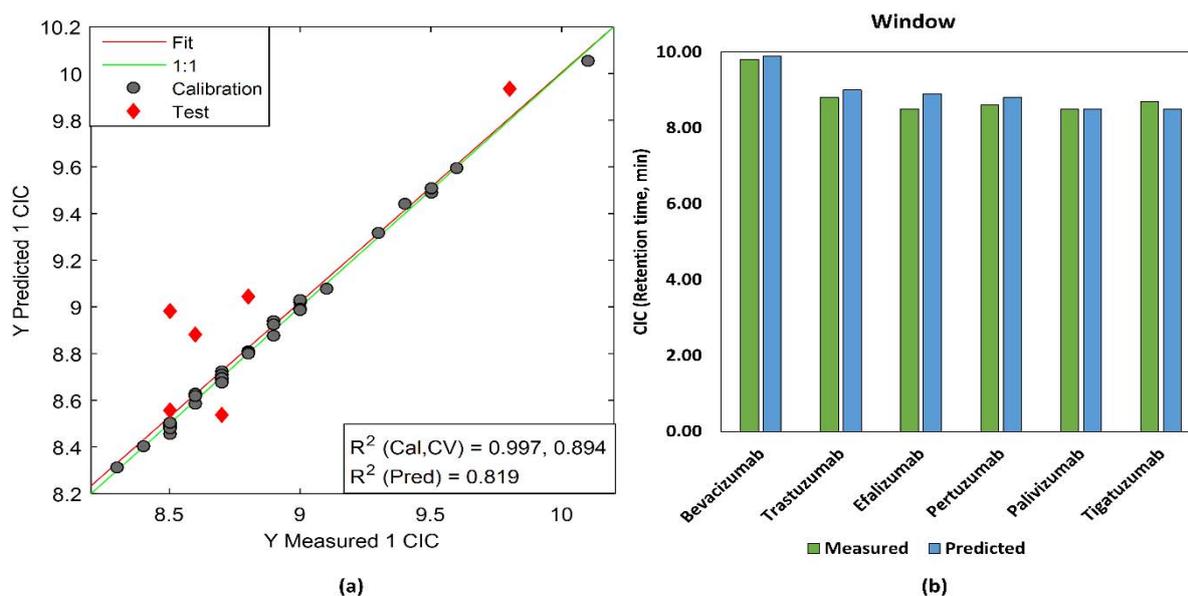
**Figure 6.4** Model summary of Domain dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.



**Figure 6.5** Regression coefficients of GA selected variables for domain dataset model (a) VH domain and (b) VL domain

### Window Dataset

Figure 6.6 provides an overview of the GA-selected variables based PLS model (5 LVs) for the window dataset. A good model fit can be observed throughout the data distribution for the calibration model (Figure 6.6a). For the test set prediction, as seen from Figure 6.6b, all the predicted values are within 30 seconds of the experimentally determined retention times. Atezolizumab and duligotuzumab have been removed from the test set due to abnormally high Y residuals (Appendix D Figure D.2b). Alternately the descriptors were not able to capture this information.



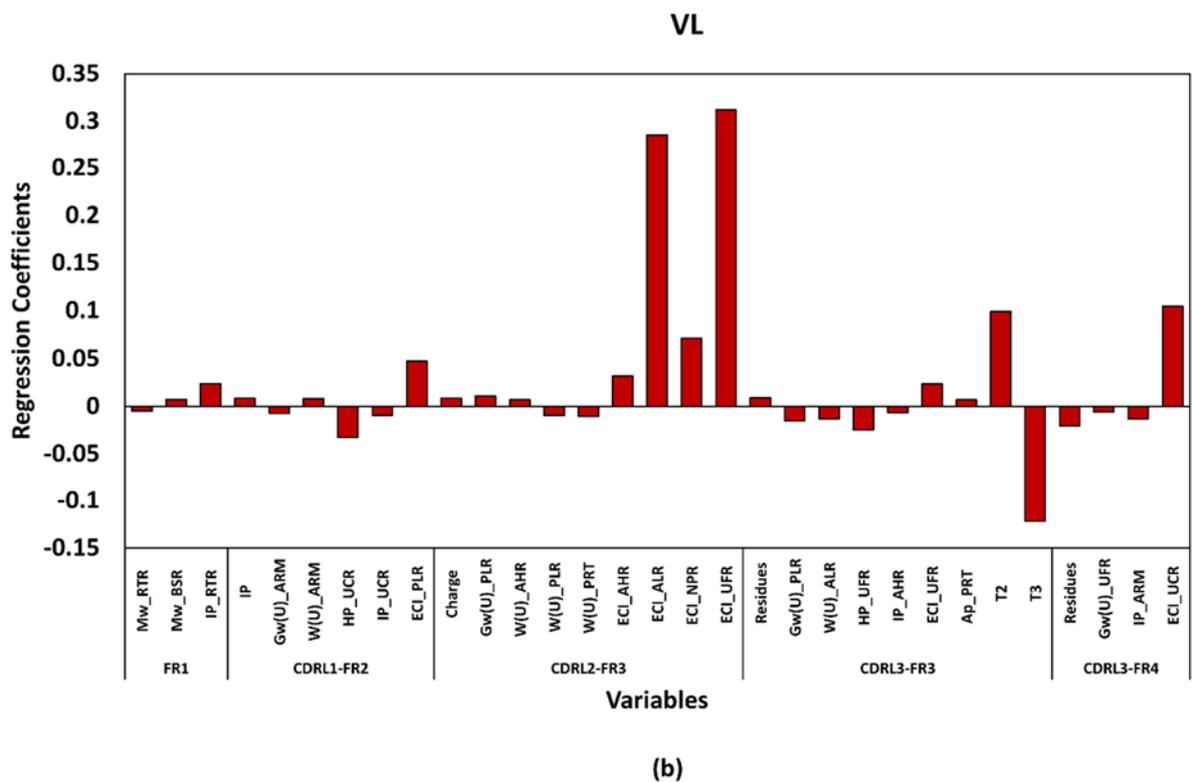
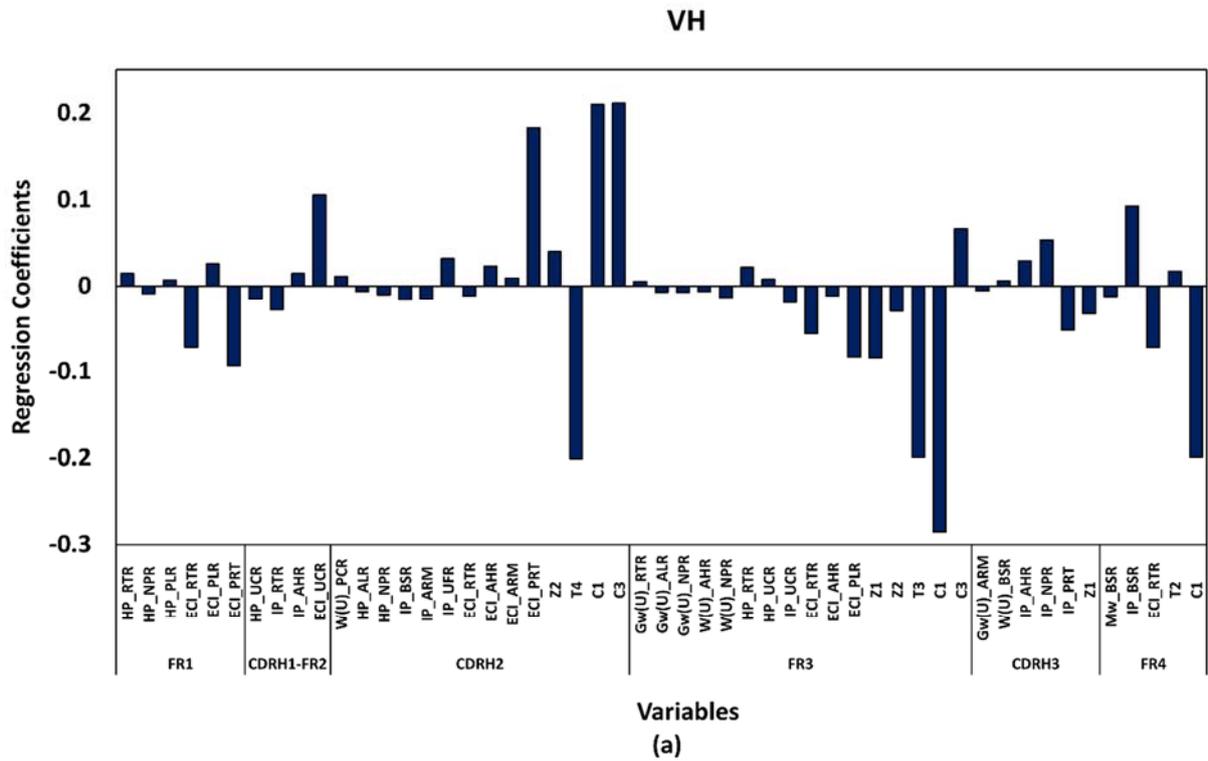
**Figure 6.6** Model summary of Window dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.

As the Window, Single Amino and Running Sum datasets were generated following multiple sequence alignment, the relative positions of residues of the complementary determine region (CDR) and framework region (FR) that constitute the hypervariable region of mAbs needed to be identified. This has been shown in Table 6.2.

**Table 6.2** Relative positions of CDR and Fr regions for heavy and light chain following sequence alignment.

<b>Heavy</b>	<b>Indices</b>	<b>Light</b>	<b>Indices</b>
FR1	1-21	FR1	1-23
Start	22-26	Start	24
CDRH1	27-38	CDRL1	25-40
End	39-40	End	41-43
FR2	41-47	FR2	44-53
Start	48-52	Start	54-55
CDRH2	53-71	CDRL2	56-62
End	72-74	End	-
FR3	75-100	FR3	63-95
Start	101-103	Start	96
CDRH3	104-124	CDRL3	97-111
End	125-128	End	112-115
FR4	129-135	FR4	116-121

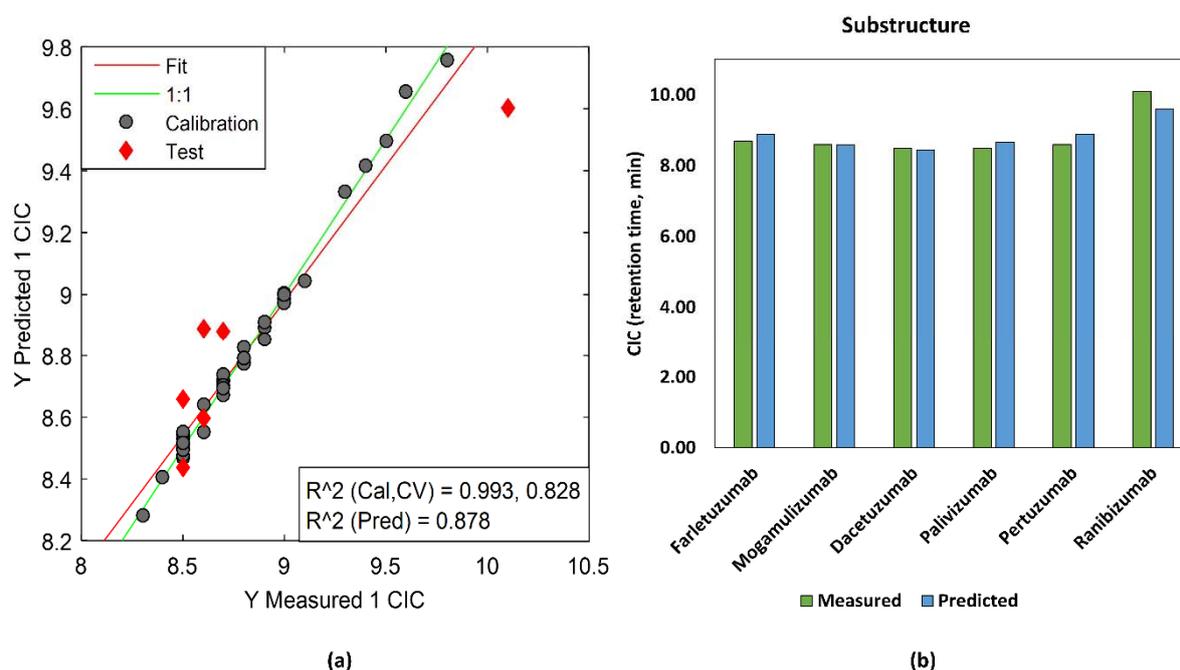
GA-Selected variables with regression coefficients  $>0.001$  are shown in Figure 6.7. In the VH domain, descriptors pertaining to electronic charge index, topological properties and the electrostatic potential based properties of residues in CDRH2 as well as electrostatic potential based and topological properties of residues in the FR3 showcase the highest contributions (Figure 6.7a). Additionally, size based molecular electrostatic potential of residues in the FR4 contribute to the model. For the VL region, clear contribution can be seen from the residues in the CDRL2-FR3 regions, arising from the electronic charge indices of the aliphatic residues and unfolding residues as well as electronic charge indices of aliphatic residues as well as that of unfolding residues (Figure 6.7b). Furthermore, topological properties of residues in the CDRL3-FR3 region as well as electronic charge index of the residues in the CDRL3-FR4 region contribute to the model.



**Figure 6.7** Regression coefficients of GA selected variables for Window dataset model (a)VH domain and (b) VL domain.

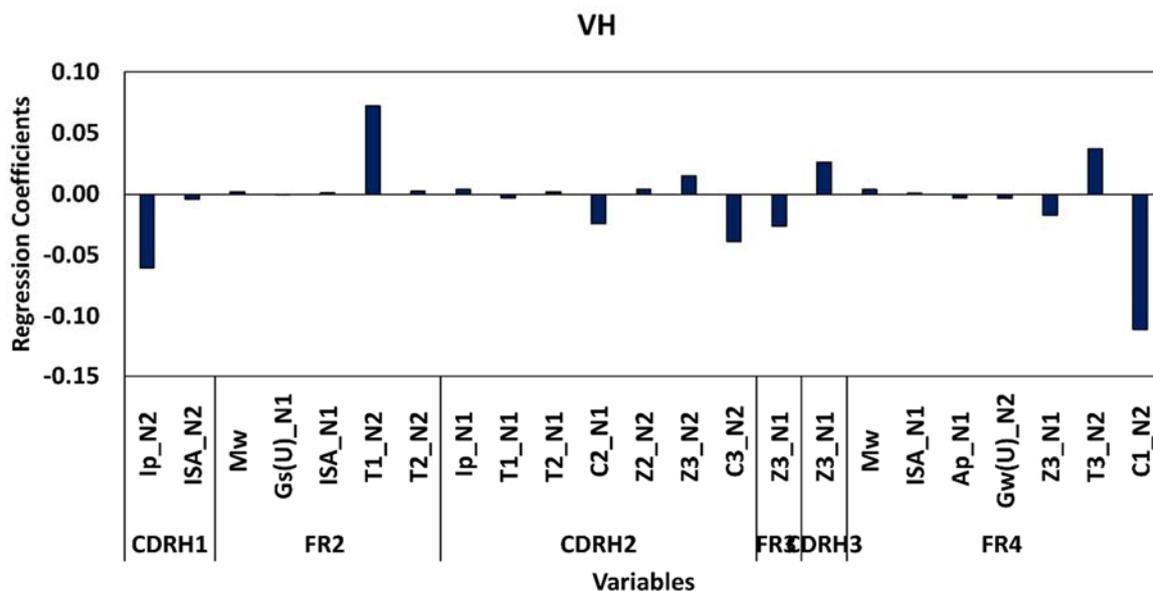
### Substructure dataset

Figure 6.8 provides an overview of the GA-selected variables based PLS model (5 LVs) for the substructure dataset. A good model fit can be observed throughout the data distribution for the calibration model (Figure 6.8a). For the test set prediction, as seen from Figure 6.8b, all the predicted values are within 30 seconds of the experimentally determined retention times. Atezolizumab and duligotuzumab have been removed from the test set due to abnormally high Y residuals (Appendix D Figure D.2c). Alternately the descriptors were not able to capture this information.

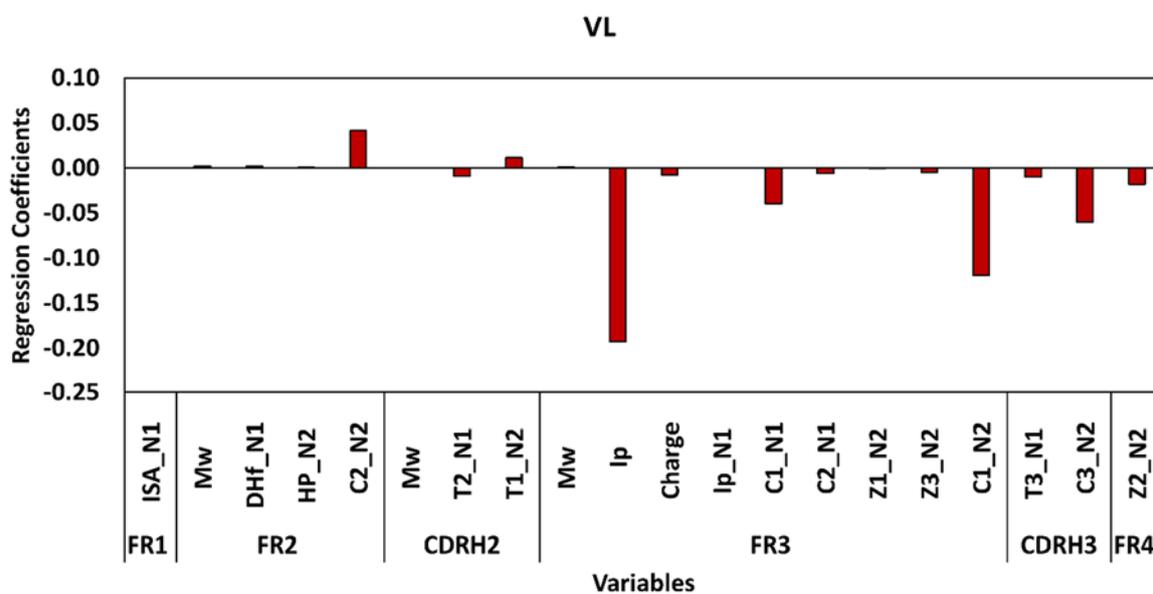


**Figure 6.8** Model summary of Substructure dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.

GA-Selected variables with regression coefficients  $>0.001$  are shown in Figure 6.9. In the VH domain, descriptors pertaining to isoelectric point in CDRH1, topological properties of residues in FR2 as well as molecular electrostatic potential based of residues in the FR4 showcase the highest contributions (Figure 6.9a). Additionally, size based molecular electrostatic potential of residues in the FR4 contribute to the model. For the VL region, clear contribution can be seen from the isoelectric potential and molecular electrostatic potential of residues in the FR3 region (Figure 6.9b).



(a)



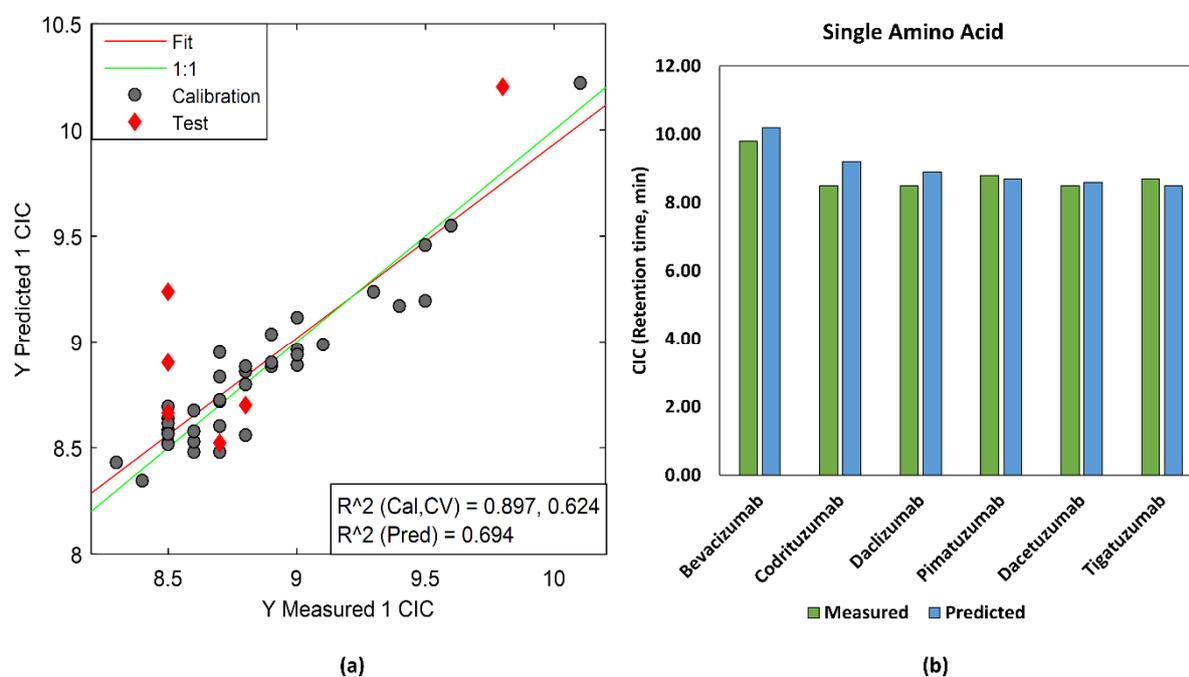
(b)

**Figure 6.9** Regression coefficients of GA selected variables for Substructure dataset model (a)VH domain and (b) VL domain.

*Single Amino Acid Dataset*

Figure 6.10 provides an overview of the GA-selected variables based PLS model (2 LVs) for the single amino dataset. Figure 6.10a depicts the model fit where dispersion can be observed for the samples on the lower end of the retention times from the parity line. Atezolizumab and duligotuzumab have been removed from the test set due to abnormally high Y residuals

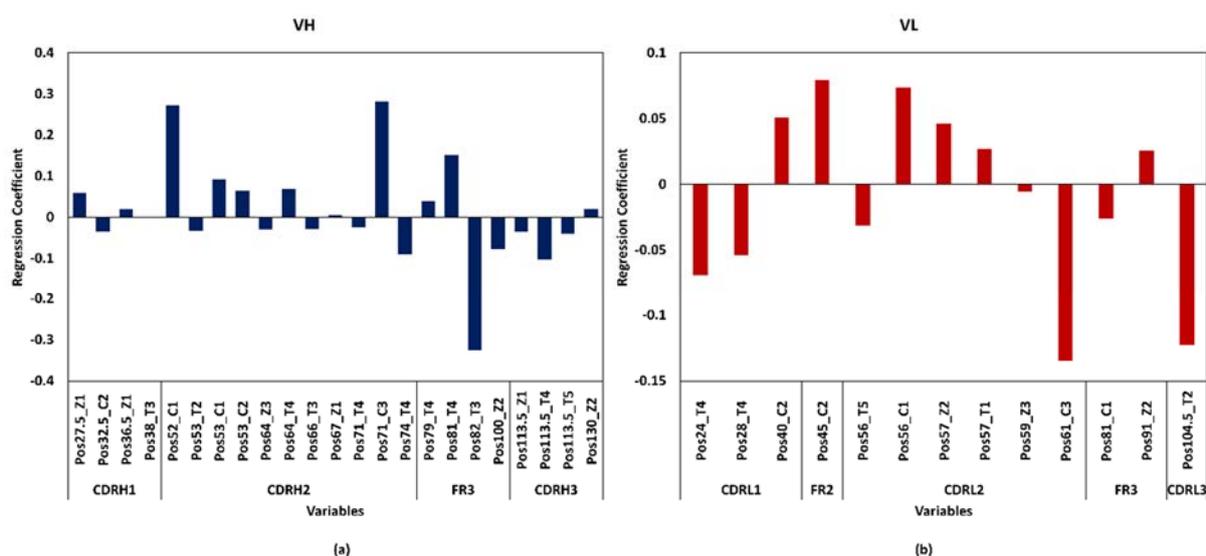
(Appendix D Figure D.2d) For the test set prediction, as seen from Figure 6.10b, all the predicted values are within 30 seconds of the experimentally determined retention time except for codrituzumab which is overestimated by 42 seconds.



**Figure 6.10** Model summary of Single Amino dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.

Position numbers in the Single amino acid dataset refer to the residue indices following MSA as shown in Table 6.2. The properties of residues in the CDRH2 region as well as FR3 have shown to influence the model. In the VH domain the size based molecular electrostatic potential of Pos 52 and the difference between positively charged residues at Pos 71 of the CDRH2 region show most contribution to the model (Figure 6.11a). For Pos 52 the residues are either Alanine or Glycine, wherein samples with Alanine, the retention times are towards the lower end of the range with the exception of atezolizumab, which had the highest retention time of all samples. This could explain why atezolizumab has a higher Y residual and its unusual behaviour in the model. Pos 71 is Serine for all samples except ozanezumab, where it is replaced by Glycine. This could be indicative of the model trying to fit the sample (overfitting) rather than capturing information that contributes to the correlation between descriptors and response (pattern recognition). Additionally, the topological properties of Pos 81 and 82 in the FR3 region contributed to the model (Figure 6.11a). Asparagine, Threonine or Serine are possible amino acids at Pos 82 and their topological properties influence the model, with Asparagine and Threonine being bulkier than Serine and this could indeed influence the local

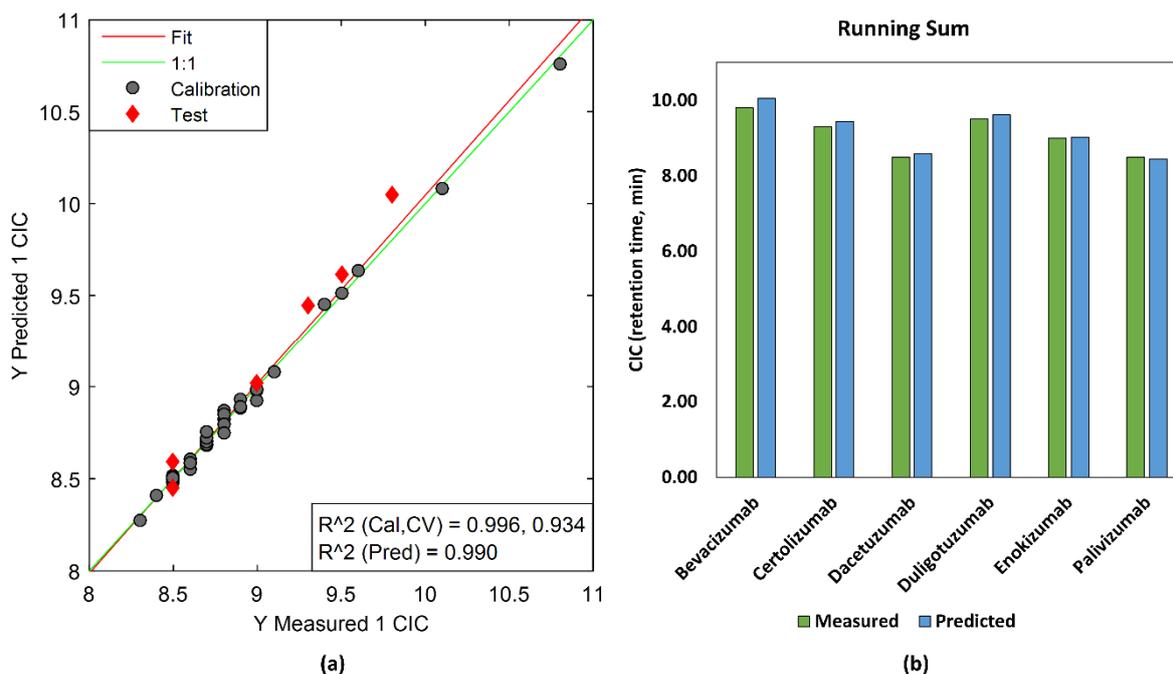
conformation of the protein (Zimmerman *et al.*, 1968; Cho *et al.*, 2007). For the VL domain, the electrostatic potential base properties of Pos 61 and Pos 56 have shown to have most contribution. Topological properties of residues in CRL1 and CDRL3 (Pos 104.5) regions as well as charge base properties of FR2 region contribute to the model the model (Figure 6.11b). Since most of these residues are in the CDR region or adjacent to it, the amino acid composition at each residue will vary from sample to sample.



**Figure 6.11** Regression coefficients of GA selected variables for Single amino dataset model (a)VH domain and (b) VL domain

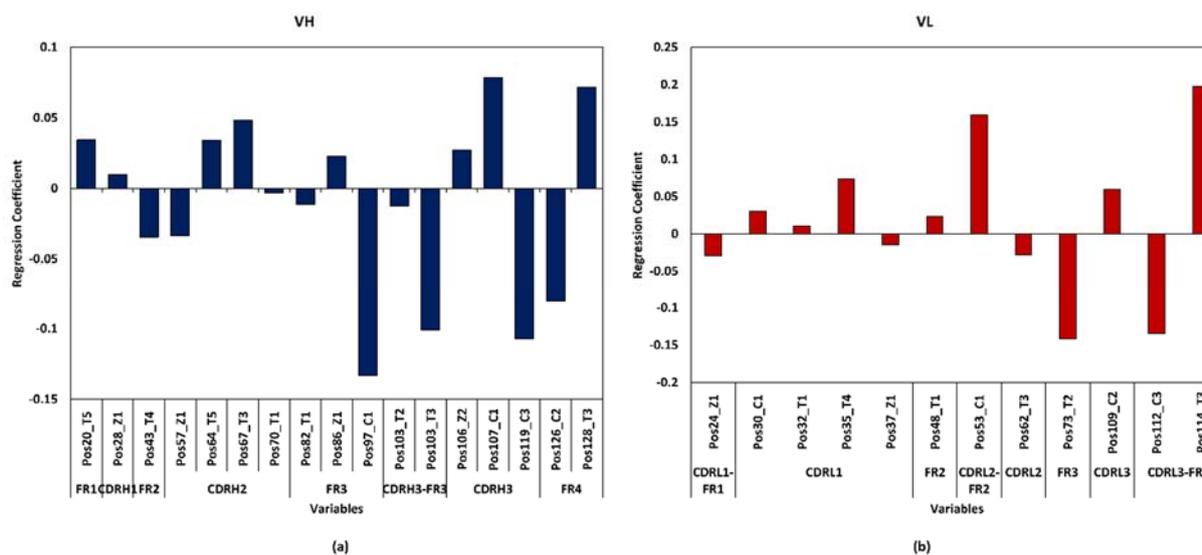
### *Running Sum dataset*

The GA selected variable based PLS model (7 LVs) summary for the running sum dataset is shown in Figure 6.12. Figure 6.12a depicts that the model fit is best throughout the distribution, especially the higher extreme. For the test set predictions, the predicted values of retention times are within 20 seconds of experimental values (Figure 6.12b). Ocrelizumab and elotuzumab were removed from the test set as it had unusually high Y residuals (Appendix D Figure D.2e). Even when these samples were in the training set the model performance was quite poor. This could be due to the fact they have very similar descriptors to other mAbs in the dataset but very dissimilar responses. A larger dataset could provide more information regarding the discrepancy of these data points.



**Figure 6.12** Model summary of Running Sum dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.

Running sum descriptor calculations consider 13 residues around a central/ pivotal residue (indicated by Pos numbers) and the values are summed over this 13-residue segment. Position numbers in the Running Sum acid dataset refers to the residue indices following MSA as shown in Table 6.2. For the VH region, size based Electrostatic potential of residues in FR3 region showed maximum contribution to the model followed by topological properties of residues in CDRH3-FR3 region. Electrostatic potential based properties of residue in CDRH3 as well as FR4 have shown to influence the model (Figure 6.13a). The topological properties of the CDRL3-FR4 and FR3 regions as well as the electrostatic potential based properties at the CDRL2-FR2 and CDRL3-FR4 regions contribute to the model from the VL region (Figure 6.13b).



**Figure 6.13** Regression coefficients of GA selected variables for Running sum dataset model (a) VH domain and (b) VL domain.

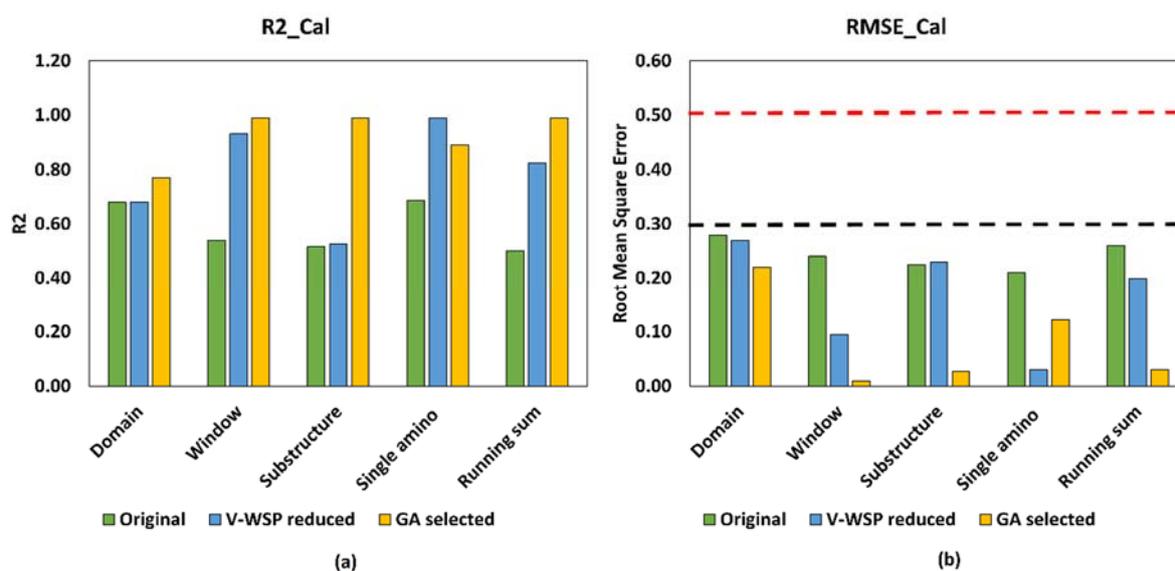
#### 6.2.4. Benchmarking of model performance metrics

All the models generated for the different datasets were benchmarked based on the OECD guidelines for QSAR model Validation (Veerasamy *et al.*, 2011; Organisation for Economic and Development, 2014). The models were considered acceptable for the model fit if Calibration  $R^2 > 0.6$ , Cross validation  $R^2 > 0.6$  and  $Q^2 > 0.5$  and finally External test prediction  $R^2 > 0.6$ . The models were further benchmarked for numerical accuracy wherein a  $RMSE < 0.3$  is usually recommended. However previous studies have shown that an  $RMSE < 10\%$  of the output range is considered acceptable if accompanied by external validation especially for QSAR models developed for early stage screening (Alexander *et al.*, 2015). The above values are more of an indicative measure about the model performance rather than hard and fast rules especially for early stage screening applications. The performance metrics for the original, V-WSP reduced and GA selected Cross validated PLS models are reported in Appendix D Table D.5.

#### Calibration

The performance metrics of the Calibration models are shown in Figure 6.14. For Domain dataset the models after variable selection perform better than the original and reduced dataset. For Window, Substructure and Running Sum there is improvement in model performance after each step of model optimisation, with GA-Selected models performing the best. Window, Substructure and Running sum have the high calibration  $R^2$  following variable selection (Figure 6.14a). Among the best performing datasets, the window dataset has a higher number

of variables compared to the Substructure and Running sum (191, 43 and 29 respectively) making model interpretability an issue. For the Single amino acid dataset, the reduced dataset has better performance than the original dataset and GA selected variable model. In general, the original dataset performs worst, and this is since this large dataset contains redundant and noisy variables as confirmed by the superior performance of the reduced dataset that address the problem of collinearity and redundancies whereas the GA-PLS selection further addresses the issue of noisy variables. The Root mean square error is not only below the 10% output range for all the calibration models but adheres to the OECD guidelines of  $RMSE < 0.3$  as well (Figure 6.14b). For Domain based dataset the RMSE is reflective of model performance with GA selected models having only slightly lower error than the original and reduced dataset. For both Window, Substructure and Running Sum dataset with model optimisation and variable selection the RMSE values show marked decrease with nearly 90% reduction. The RMSE values for the Single amino acid have a slightly different trend which is reflective of the  $R^2$  coefficients with reduced dataset having lower RMSE than GA selected.

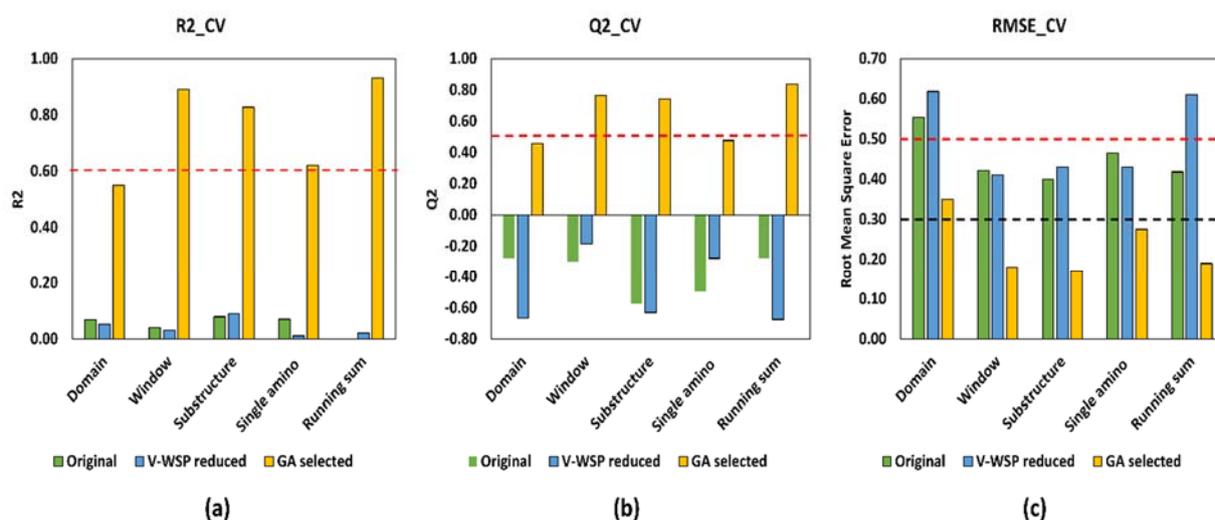


**Figure 6.14** Performance based on (a)  $R^2$  and (b) for PLS Calibration models generated from the four data blocks: Domain, Window, Single amino, and Running sum modelled against CIC. The models were generated using the original dataset, V-WSP reduced dataset as well as after variable selection (GA selected). Red dashed lines indicate values at 0.6 for  $R^2$  and 0.5 for RMSE. Black dashed line indicates  $RMSE=0.3$ .

### Cross validation

Internal validation of the models was performed via 5-fold random subset cross validation with 10 iterations and the corresponding performance metrics are shown in Figure 6.15. The GA

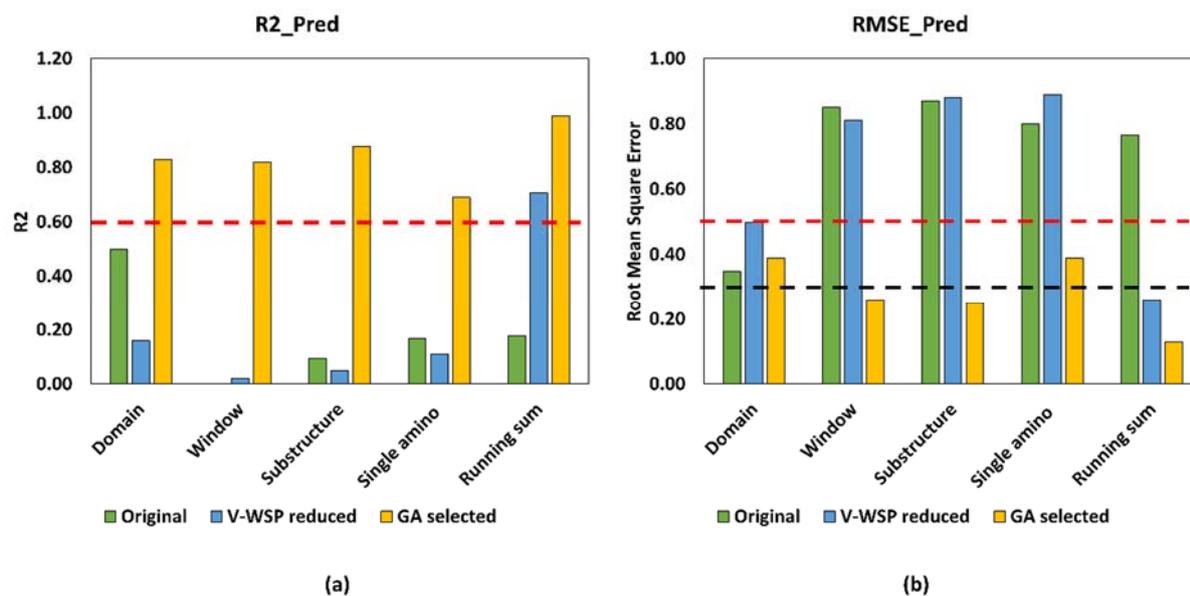
Selected models have superior performance with respect to both  $R^2$  and  $Q^2$  for all datasets when compared with their respective original and reduced datasets (Figure 6.15a, 6.15b). The RMSE values follow the same trend as the GA models of all datasets with RMSE less than 0.5 (Figure 6.15c). The original and reduced models for the domain dataset along with the reduced model for running sum dataset also reported an  $RMSE > 0.5$ . For the RMSE of the GA-selected datasets were ranked as followed: Domain>Single>window>Running>Substructure. The GA selected models pass the criteria with regard to both  $R^2$  as well as for RMSE with Domain have the lowest performance and the Running Sum performing the best.



**Figure 6.15** Performance based on (a)  $R^2$  (b)  $Q^2$  and (c) RMSE for PLS cross validation models generated from the four data blocks: Domain, Window, Single amino, and running sum modelled against CIC. The models were generated using the original dataset, V-WSP reduced dataset as well as after variable selection (GA selected). Red dashed lines indicate values at 0.5 for  $R^2$ ,  $Q^2$  and 0.5 for RMSE. Black dashed line indicates  $RMSE=0.3$ .

### *External test set prediction*

Distance based method of Kennard Stone was used to split the dataset after variable reduction (80% training and 20% test). All GA-selected models performed well for prediction of the external dataset with the following ranking for prediction  $R^2$ : Running Sum>Substructure>Domain>Window>Single amino (Figure 6.16a). The reduced dataset for running sum also passed the OECD criteria for prediction. RMSE values were lower than 0.5 for all GA-selected models with Running Sum<Substructure<Window<Single Amino<Domain (Figure 6.16b). The RMSE values for Original as well as V-WSP reduced models for Running Sum were below 0.5 as well. Furthermore, the RMSE of Window, Substructure and Running sum datasets were below 0.3 making them acceptable as per OECD guidelines.



**Figure 6.16** Performance based on (a)  $R^2$  and (b) RMSE for PLS prediction models generated from the four data blocks: Domain, Window, Single amino, and Running sum modelled against CIC. The models were generated using the original dataset, V-WSP reduced dataset as well as after variable selection (GA selected). Red dashed lines indicate values at 0.6 for  $R^2$  and 0.5 for RMSE. Black dashed line indicates  $RMSE=0.3$ .

### 6.2.5. *Y* randomisation

When dealing with a smaller dataset it is essential to check for overfitting of the models. When a model is overfitting the data, it means that the current model works well for the data at hand but may have poor predictive performance for new samples. (Roy, 2007). Table 6.3 indicates the  $R^2$  and  $Q^2$  values of the five datasets following *y* randomisation. As seen from the calibration and cross validation metrics of the *Y* scrambled model, they perform poorly. This reinforces the fact that the model performance was not just based on chance correlation but due to an underlying causal relationship between descriptors and response as all the  $R^2$  and  $Q^2$  values are below 0.5 (Guha and Jurs, 2005). Apart from the Domain dataset all the GA models for Window, Substructure, Single amino and Running sum datasets show a significant difference between the original models and those generated followed by *Y* scrambling (within the same range of output values). The null hypothesis of the test is that there is no significant difference between the GA selected dataset and the *Y* randomised dataset, both built under the same conditions. The permutation analysis reports the results of non-parametric statistical hypothesis tests as shown in Table 6.2 for both the calibration (self-prediction) as well as cross validation model comparison. Values lower than 0.05 indicate that the models based on GA variable selection are significantly different from the *Y* randomised models at 95% confidence

level. This indicates that the models generated in this study are capturing an underlying relationship between the descriptors and the response and not a chance/random correlation.

**Table 6.3** Results of y randomisation and permutations tests for GA selected datasets.

Datasets	Self-prediction			Cross Validation			CV	
	Wilcoxon	Sign Test	Rand t-test	Wilcoxon	Sign Test	Rand t-test	R <sup>2</sup>	Q <sup>2</sup>
Domain	0.01	0.058	0.019	0.010	0.051	0.014	0.04	-0.30
Window	0.000	0.001	0.005	0.000	0.005	0.005	0.05	-0.38
Substructure	0.000	0.001	0.005	0.000	0.000	0.005	0.04	-0.37
Single Amino	0.003	0.020	0.006	0.005	0.03	0.017	0.05	-0.49
Running sum	0.000	0.000	0.005	0.000	0.000	0.005	0.04	-0.40

### 6.3. Discussion

Early stage screening of mAbs during the discovery phase, based on the potential to cause developability issues, such as aggregation propensity and solubility problems that arise from non-specific interactions, would aid in reducing manufacturing failures and attrition rates. QSAR models can serve as important *in silico* tools that allow for prediction of mAb behaviour with respect to non-specific interaction. When the main objective of the model is to correctly identify the trends and rank molecules based on their activity/property in question, the metrics that capture the correlation between observed and predicted values of the test set i.e. prediction R<sup>2</sup>, maybe more reflective and relevant when compared to RMSE, accuracy and error of the models (Alexander *et al.*, 2015). With respect to mAbs, models that allow for early stage screening based on safety and efficacy may lead to effective reduction of the number of lead candidates that progress through process development into manufacturing and thus reduce costs associated to process development failures and attrition. The results of these studies could contribute to early stage screening and better design of mAbs for increased efficacy and decreased polyspecificity.

This study presents a hybrid modelling framework employing five novel descriptor sets: Domain, Window, Substructure, Single amino and Running Sum, extracted from the primary sequence of mAbs with increasing resolution of information and decreasing level of confounding of information. The datasets also look at the difference between the global and local perspectives with the Domain set consisting of global descriptors and single amino consisting of more local descriptors. Overall the Window, Substructure and Running sum datasets perform the best, based on the Calibration, Cross validation and Prediction metrics (Figure 13,14 and 15) and TableS5 (Supplementary File SF1). The reason behind it could be

that these datasets consider segment-based properties of the amino acid sequence which highlights the influence of adjacent amino acids as well. This could be particularly important with charge based and hydrophobicity based properties at the CDR and FR region wherein the physicochemical, electronic and topological properties are better represented by cluster and/or lower volume localised descriptors rather than averaged values (Sharma *et al.*, 2014; Robinson *et al.*, 2017). This is also supported by the slightly superior performance of the Running sum dataset wherein the descriptors are summed over smaller segments whereas the Window data and Substructure are averaged over segments. The Domain dataset is highly global and thus specific information could be confounded whereas the Single amino acid descriptor is highly local, and this could lead to higher variation owing to specific residue difference. Descriptors that have been highlighted by the various datasets revolve around hydrophobicity and charge-based properties which is reflected by the regression vectors of the descriptors selected. This is in consensus with previously shown experimental as well as QSAR based studies (Sharma *et al.*, 2014; Robinson *et al.*, 2017). Primary descriptors based multivariate models have been previously described for biophysical properties of mAbs wherein the electrostatic interactions and charge asymmetry of VH and VL regions play an import role in viscosity, hydrophobicity and charge for in vivo clearance (Sharma *et al.*, 2014).

Biophysical properties such as viscosity and in vivo clearance are vastly different for mAbs that possess the same Fc region but different variables domain, indicative of the fact that the hypervariable regions consisting of the CDR and FR and their corresponding physicochemical and charge based properties are responsible for these differences (Hötzel *et al.*, 2012). This is reflective of the current dataset used as well, since the variable regions have been grafted on to the same IgG1 Fc region for all mAbs and still display different CIC retention profiles. The descriptors generated in this study thus involve only those for the VH and VL regions. The exposed charged residues in the VH and VL regions have been identified to influence the extent of nonspecific interactions of mAbs (Yadav *et al.*, 2011). It also reflects the involvement of the hypervariable region, consisting of CDR and FR regions, in nonspecific interactions that could lead to interactions and aggregation propensities (Bethea *et al.*, 2012). Weak electrostatic interactions between mAb domains have also been identified to induce mAb interactions (Kanai *et al.*, 2008; Nishi *et al.*, 2011).

Previous studies show correlation of CIC with accelerated stability slope and increased clearance rates in mice which is indicative of increased aggregation propensities as well as poor solubility respectively, both of which are potent developability issues (Kelly *et al.*, 2015).

Higher polyspecificity profile, both cross reactivity or self-interaction, leads to faster clearance rates and this could lead to inadequate dosing making it difficult to achieve the requisite therapeutic response (Tessier *et al.*, 2014). Furthermore, increased non-specific interaction could alter the pharmacokinetics of the mAbs. Negative correlation between IgG-FcRn interaction and half-life has been observed in previous studies for mAbs with identical Fc region and different variable regions leading to differing pharmacokinetic outcomes (Wang *et al.*, 2011; Kelly *et al.*, 2016).

Increased number of variables in the dataset could lead to issues of noise, redundancy and chance correlation which can ultimately affect the results of QSAR model development which relies on finding the correlation between the descriptors and response. Unsupervised variable reduction techniques would be the first step to tackle the problem wherein V-WSP has shown superior performance to other data reduction techniques, especially in QSAR model development and was thus used in this study (Ballabio *et al.*, 2014). V-WSP data reduction also facilitates subsequent supervised variable selection method as the latter suffers if there is high degree of correlation present between variables, leading to overfitting (Hawkins, 2004). The second step in tackling the above problem and to further optimise the model is supervised variable selection using methods such as which have been detailed in previous studies (Andersen and Bro, 2010). GA-PLS was chosen here as there is no requirement for the initial model to be perform well, works with many noisy/irrelevant variables and allows for repeated application (Hasegawa *et al.*, 1997; Andersen and Bro, 2010). However, the main caveat of GA is overfitting and this can be addressed by strict cross validation regimes such as random subsets with multiple iterations, as performed in this study (Gao *et al.*, 2002). This can also be checked for using a test set for external validation of the model. For the modelling techniques used here, PLS is quite common for protein and peptide based QSAR studies (Zhao *et al.*, 2007; Nongonierma and FitzGerald, 2016). Other studies have been carried out using techniques such Support Vector Regression (SVR), Artificial Neural Networks (ANN), Principal Component Regression (PCR) and Decision Tree based techniques which could be a potential next step in this study (Zhou *et al.*, 2008; Sharma *et al.*, 2014; Robinson *et al.*, 2017). However, PLS models have shown comparable performance to SVR in a previous QSAR model development for predicting mAb retention behaviour (Robinson *et al.*, 2017). Further advances would be developing 3D descriptors based on homology modelling and molecular dynamic simulations which would allow for better understanding of surface properties that would allow for better mechanistic understanding of nonspecific interactions. Primary sequence based descriptors

have been successfully used in previous studies to develop QSAR models for mAb developability and the key advantages are that they are easier to use, rapid, less computationally intensive and provide a reasonable performance for predicting mAb behaviour (Sharma *et al.*, 2014). Primary sequence based descriptors have also been used in bioprocess route selection via machine learning models (O'Malley *et al.*, 2012). Extension to this study would be to involve higher sample sizes as well as modelling other mAb isotypes provided there are sufficient samples to allow development of an applicability domain. Furthermore, this framework can also be extended for application in bispecific mAbs.

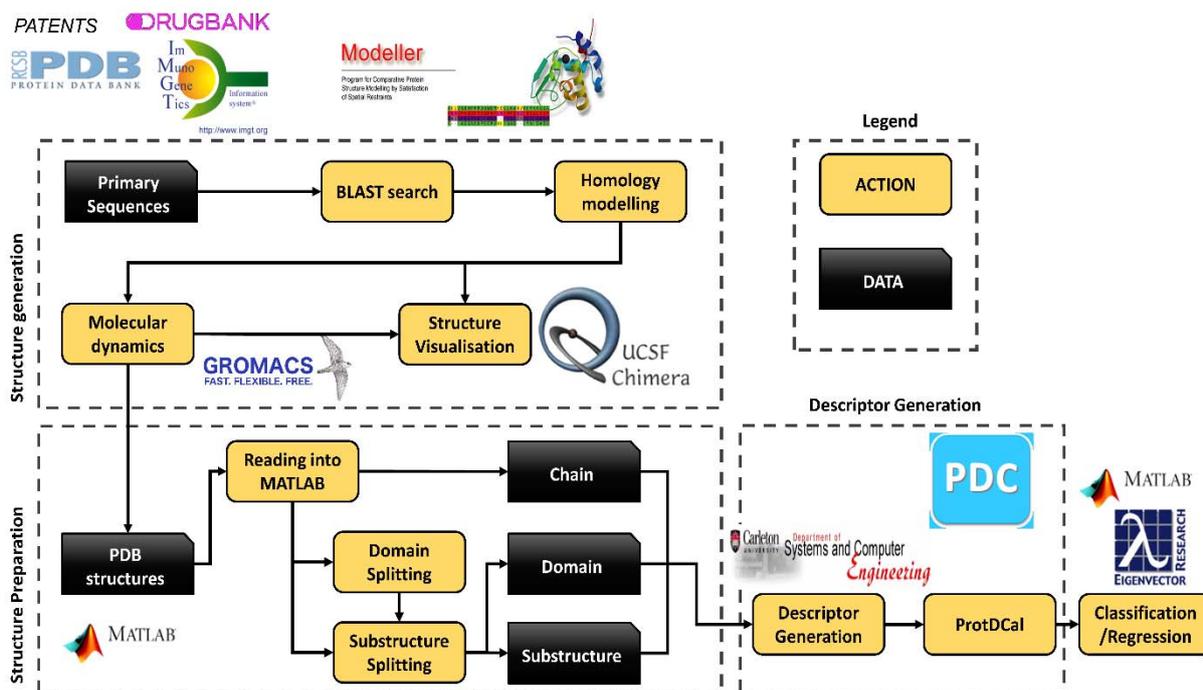
#### 6.4. Chapter summary

This chapter discussed the development of a predictive model based on the CIC retention times as an output which could thus contribute to and facilitate the early stage screening and characterisation of mAbs during the developability phase. This framework uses four novel datasets: Domain, Window, Substructure, Single Amino and Running Sum, derived from the primary sequences of homologous mAbs (IgG1-kappa-Humanized) with varying degrees of resolutions. Unsupervised pattern recognition was first performed on the descriptor sets to visualise any intrinsic property-based separation or clustering following which the descriptors were regressed against reported Cross Interaction chromatography retention times. Model optimisation was performed via unsupervised variable reduction followed by supervised variable selection. Finally, the models and datasets were benchmarked based on regression model performance metrics such as  $R^2$ ,  $Q^2$  and RMSE for the calibration, cross validation as well as external validation using a test set. To check whether the model is capturing only chance correlations, output (Y) randomisation and permutation test was performed. Datasets that contain information pertaining to a cluster of amino acids represented by localised descriptors rather than averaged value over the entire protein had better predictive performance of CIC retention behaviour with  $R^2 > 0.8$  and  $RMSE < 0.3$ . Furthermore, the results indicate the physicochemical, electronic and topological properties of hypervariable regions of antibodies contribute most to the CIC retention times.

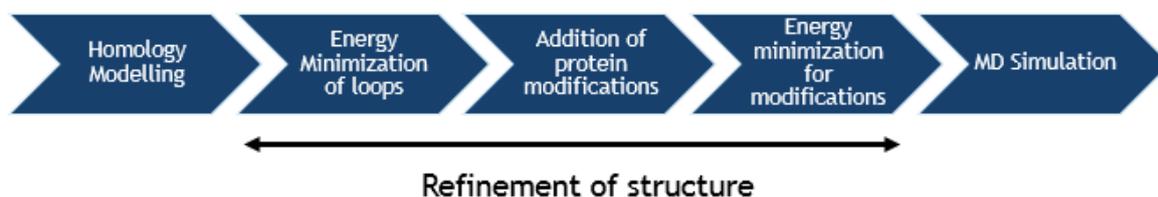
The next chapter focusses on 3D structure-based descriptors of mAbs for QSAR model development with CIC as an output so as to capture the functional characterisation aspect provided by structural information.

## **Chapter 7: QSAR model development using 3D descriptors of mAbs generated via homology modelling and molecular dynamics simulation.**

The previous chapter described the application of primary sequence-based descriptor datasets in QSAR model development. Even though models developed using all the datasets performed satisfactorily based on the performance evaluation metrics, the aspect of functional characterisation is slightly diminished due to the absence of the 3D information. This chapter describes the generation of descriptors from 3D structures of mAbs based on homology modelling and molecular dynamic simulations which allows for better understanding of surface properties that would facilitate better mechanistic understanding of nonspecific interactions. As the mAbs tested in their study had different variable domains but the same IgG1 constant domain, only the Fab region of mAbs were utilised in this study. Furthermore, it was computationally advantageous as it speeds up the process since pruning of whole molecules is not required. The first step was to generate the 3D structures from the primary sequence of mAbs which was achieved via homology modelling and molecular dynamics simulation. Homology modelling is a technique that enables the prediction of 3D structure via comparative protein modelling of target protein sequences with a template protein structure that shares a high sequence similarity with the target protein (Eswar *et al.*, 2006). These structures developed were then subjected to energy minimisation and loop refinement so as to achieve a stable 3D structure (Fiser and Šali, 2003). Once homology models have been developed they were then subjected to molecular dynamic simulations that further simulate the atoms and their interactions within a dynamically evolving system; in this case mAbs in aqueous solutions (Hospital *et al.*, 2015). Following the generation of Fab structures of mAbs, the development of a robust QSAR model was extended using three novel datasets generated from the 3D structure of mAbs and they were modelled against corresponding CIC retention time based on the data recently published by Jain *et al.*, (Jain *et al.*, 2017).. Three independent and novel X block datasets consisting of these descriptors were generated based on the physicochemical, electronic, thermodynamic, electronic and topological properties of amino acids: Chain based, Domain based, and Substructure based. The relationship between the descriptors and the responses was captured by Partial Least Squares based models and the datasets as well as models were benchmarked based on metrics as outlined in QSAR validation guidelines (Roy *et al.*, 2015). Furthermore, the models were validated with an external set of mAbs and the corresponding model performance metrics were benchmarked for the different datasets. The overall methodology is outlined in Figure 7.1.



(a)



(b)

**Figure 7.1:** Overview of methodology for generating (a) 3D structures from primary sequence of antibodies followed by descriptor generation (b) Workflow for generating 3D structures from mAbs useful for generating structural descriptors as well as docking studies.

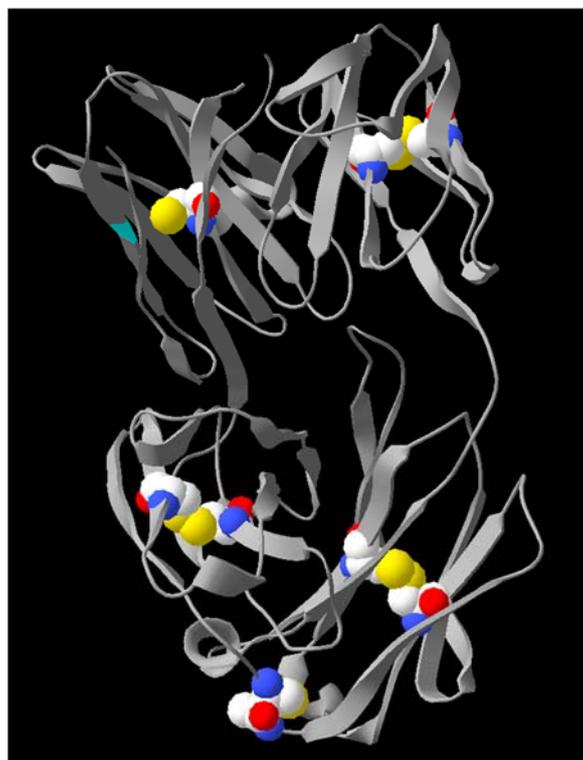
## 7.1. Materials and Methods

### 7.1.1. Data collection

Sequence information, substructure, species and phase of development of 134 mAbs have been collected from IMGT, literature, patents as well as from industrial partners (Apweiler *et al.*, 2004; Lefranc *et al.*, 2009). The FASTA format sequences of mAbs were used for building 3D structures as described in Section 7.1.2. Only the Fab regions of mAbs were utilised in this study and for further model development as explained above.

### 7.1.2. Generation of mAb structures

**Homology Modelling:** Fab sequences of mAbs were first put through a BLAST search (NCBI) so as to identify sequences that share the highest similarity and have an experimentally verified 3D structure within the database. These structures can then be used as templates. In the case of the mAbs used in this study a single template molecule was chosen with 70 % similarity to the starting set of sequences. An advantage of using single template is that automation is possible via means of BASH scripting once the initial parameters are set for alignment of mAbs to the template structure. Parameters initially set were identifying cysteine residues, specifying the Sulphur-Sulphur bond distance of 2Å and to identify the disulphide bond residues. Furthermore, alignment of residues was also checked. and the endpoint is to model against non-specific binding. 2fgw was chosen based on correct positioning of C terminal and all disulphide bonds being in place (Figure 7.2). Modeller 9.19 was used with Python library-based scripts for generating the homology models (Šali *et al.*, 1995; Fiser and Šali, 2003; Eswar *et al.*, 2006). Once the parameters were set all 133 mAbs were modelled using a BASH script wherein five disulphide bonds were set as restraints (4 intra domain and one inter chain). For mAbs with more than 10 Cysteine residues the homology models were developed and refined individually (as in the case of TGN1412). Once the homology models were generated, the selection of good models were based on Modeller assigned scores based on energy wherein the criteria of acceptance were values <10000 for good models. Apart from being computational advantageous, a single template approach was used in this case taking into consideration the preliminary and explorative nature of the study where the endpoint was to develop a model to predict nonspecific interaction. MAb Fab structures were generated by João Victor de Souza Cunha and Dr Agnieszka, School of Natural and Environmental Sciences, Newcastle University.



**Figure 7.2** 3D structure of mAb 2fgw with the coloured spheres representing the disulphide bridges between the fixed Cysteine residues.

**Molecular Dynamic Simulation:** Molecular dynamics simulation was carried out using GROMACS 5-37 for a run of 50 ns using AMBER99SB-ILDN forcefield. The forcefield was reassigned along with water molecules (TIP3P). A box (system) was assigned around the protein with the edge of the box  $9\text{\AA}$  away from protein.  $9\text{-}10\text{\AA}$  distance is needed to surround the molecule with three layers of water molecules (Hospital *et al.*, 2015). Ions were added to mimic biological physiological conditions of protein ( $0.1\text{M}$  NaCl). This starting set up was followed by energy minimisation wherein the potential energy of this system i.e. structural/system energy was minimised. For each energy minimisation step the atom was allowed to move  $0.002\text{\AA}$  and this was run for 10000 steps. The temperature and pressure equilibration steps followed energy minimisation. The first equilibration step was NVT (Number of atoms, Volume inside box, Temperature). Keeping NV constant, the temperature was increased from  $0\text{-}300\text{K}$  and the simulation was run for  $100\text{ps}$  and this added Kinetic energy to the system. During this step pressure of the system was allowed to change whereas the temperature stayed constant at  $300\text{K}$ . Positional constraints were placed on the backbone during this step. The next equilibration step was NPT (Number of atoms, Pressure, Temperature) where pressure was kept constant at  $1\text{ bar}$ , but volume was allowed to change. The density of water was also inputted. Finally, the simulation proceeded to the production run i.e.  $50\text{ns}$  in NPT configuration with no structural restraints. MAb Fab structures were generated

by João Victor de Souza Cunha and Dr Agnieszka, School of Natural and Environmental Sciences, Newcastle University.

### 7.1.3. Descriptor generation

The PDB structures of mAbs fab regions were inputted into ProtDCal software for descriptor generation. The details of descriptors generated from the 3D structures are outlined in Table 7.1. Three different datasets were generated from the PDB structures of mAb Fab regions with varying degrees of resolution as well as localisation. **Chain dataset:** The chain dataset consisted of descriptors summed over the entire chain i.e. Heavy chain (VH+CH1) and light chain (VL+CL) with 53 descriptors generated for each chain resulting in a total of 106 descriptors. The chain dataset was the most global of all datasets. **Domain Dataset:** The Domain dataset consisted of descriptors summed over each domain i.e. VH, CH1, VL and CL with 53 descriptors generated for each domain resulting in a total of 212 descriptors. **Substructure dataset:** For the generation of substructure dataset, the positions of the CDR and FR for the variable region as well as strands for the constant domains were identified. This resulted in 14 substructures per chain of the mAb Fab region. Thus 53 descriptors were generated for each of the 28 substructures defined resulting in the total of 1484 descriptors. The automation of descriptor and dataset generation was done by MATLAB codes written by Micael Karlberg, Newcastle University.

**Table 7.1** List of descriptors generated from 3D descriptors of mAbs used subsequently for model development.

Descriptor Type	Descriptors	Description
<b>Thermodynamic properties</b>	Gc(F)	Free energy from the conformational entropy in a folded state
	Gw(F)	Free energy from the entropy of the first shell of water molecules in a folded state
	Gs(F)	Interfacial free energy of a folded state
	W (F)	Number of water molecules close to a residue in a folded state
	Hbd	Number of hydrogen bond in the backbone of the protein
	DGw(F)	Folding free energy of the first shell off water molecules
	DGs	Variation of the interfacial free energy between folded and unfolded states
	DGel	Free energy contribution of the charge distribution within the protein

	DGLJ	Van der Wals interaction to the folding free energy
	DGtor	Dihedral torsion potential to the folding free energy
<b>Topological properties</b>	A	Solvent accessible surface area
	DA	Difference of the Solvent accessible surface area
	DAnp	Difference of the Solvent accessible surface area
	Psi & Phi	Torsion angles Psi and Phi
	lnFD	Logarithms of the Folding Degree
<b>Physicochemical properties</b>	Mw	Molecular weight
	HP	Hydrophobicity by the Kyte-Doolittle Scale
	Ip	Isoelectric point
	AP	Polar area for each amino acid in unfolded state
	ECI	Electronic charge index
	ISA	Isotropic surface area
<b>Transferable Atom Equivalent (TAE)</b>	Energy	Energy based
	VOLTAE	Volume based
	Surface	Surface based
	Population	Population based
	Rho derived	SIDel(Rho)N, Del(Rho)Min, Del(Rho)Max, Del(Rho)NIA
	K derived	SIDel(K)N, Del(K)Min, Del(K)Max, Del(K)IA SIK, SIKMin, SIKMax, SIKIA
	G derived	SIDel(G)N, Del(G)NMin, Del(G)NMax, Del(G)NIA, SIG, SIGMin, SIGMax, SIGIA
	EP derived	SIEP, SIEPMin, SIEPMax, SIEPIA
	PIP	PIPMin, PIPMax, PIPAvg

#### 7.1.4. Data curation and variable reduction

Following descriptor generation, the datasets were curated by removing variables with null values as well as zero variance variables i.e. descriptors with standard deviation  $< 0.0001$ . This set is referred to throughout the Chapter as the ‘Original dataset’. The original datasets were then subjected to V-WSP reduction, an unsupervised variable reduction method which allows for the elimination of variable based on multicollinearity (Ballabio *et al.*, 2014). This was performed via a grid search to find the correlation coefficient threshold for the descriptors for which the Procrustes index was lower than 0.2. Procrustes index is a statistical measure that allows for the assessment of the degree of comparability between the original and reduced datasets based on informational content. A Procrustes value of 1 indicates complete dissimilarity and 0 indicates that both datasets are identical (Ballabio *et al.*, 2014). For the Chain and Domain datasets, a grid search was carried out separately for the descriptors of

different domains to avoid removing chance correlations due to the presence of similar amino acids in different domains of the variable regions i.e. VH and VL domains. For the Substructure dataset, a similar grid search was performed on separate substructures however. This dataset is referred to as ‘V-WSP reduced’.

#### **7.1.5. Exploratory Analysis**

Principal component analysis, an unsupervised pattern recognition method, was performed on the reduced datasets: Domain, Window, Single amino and Running Sum using the PLS Toolbox from Eigenvector Research. The data was auto scaled (each descriptor was mean centred and scaled with its individual standard deviation) prior to analysis. Corresponding score plots were then assessed to visualise the separation, if any, of mAbs based on their intrinsic properties such as heavy chain substructure, light chain substructure as well as species.

#### **7.1.6. Cross Interaction chromatography data**

The cross interaction chromatography data for 137 mAbs were obtained from a previous study performed by Jain et al., where 30 mg of human serum polyclonal antibodies was coupled to a column followed by testing approximately 5 µg of each antibody at a flow rate of 0.1 mL/min using PBS as a mobile phase on an HPLC system(Jain *et al.*, 2017). This data was auto scaled prior to modelling.

#### **7.1.7. Model development**

The data was divided into a training and test set via Kennard stone algorithm, maintaining an 80%-20% split of data(Galvao *et al.*, 2005). The mAbs with the most dissimilarity in Euclidean space are placed into the training set. CIC retention times were modelled against the descriptors generated using the PLS Toolbox from Eigenvector Research. Both the descriptors and responses of the training set were entered into Matlab, auto scaled separately i.e. for the training set the mean and standard deviation of the training set was used and for the test set the mean and standard deviation of the test set was used. The datasets were consequently subjected to PLS analysis. The cross-validation method used was a random subset cross validation with 5 splits and 20 iterations (Minitab, 2014; Hahn and Valentine, 2016). A maximum of 10 latent variables were tested for each developed model and used to investigate the model error in order to choose the appropriate model complexity, i.e. number of latent variables. Models were developed first for the original datasets followed by the V-WSP reduced datasets. Models were further optimised by Genetic Algorithm-Partial least squares (GA-PLS) based supervised

variable selection method. The parameters were set as follows: population size of 100; maximum generations of 100; mutation rate of 0.005, window width of 1; convergence rate of 80%; 30 initial terms, cross over of 2; random subset 5 fold cross validation with 10 iterations and data retreatment set for auto scaling (Hasegawa *et al.*, 1997; Andersen and Bro, 2010). For each stage of model development and improvement i.e. Original, V-WSP and GA-Selected, the model was assessed with the test set.

#### **7.1.8. Model Performance metrics**

The datasets, models and outputs were benchmarked based on the following metrics for the multivariate regression analysis:  $R^2$  values based on Pearson correlation coefficient between observed and predicted values;  $Q^2$  values based on goodness of fit; Root mean square error of cross validation (RMSECV) and cross validation bias (Alexander *et al.*, 2015). These have been described previously in Chapter 6, Section 6.1.7. The model metrics were further assessed against the OECD guidelines for the Calibration, internally validated (Cross validation) and externally validate models (test set predictions) (Veerasingam *et al.*, 2011; Organisation for Economic and Development, 2014).

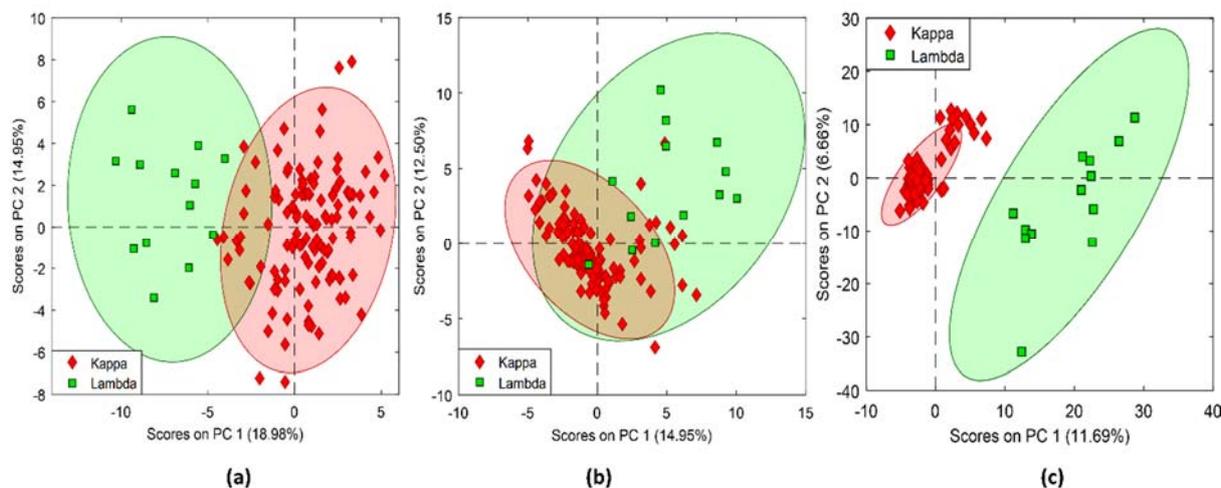
#### **7.1.9. Y randomisation**

The developed models were further evaluated by performing Y randomisation/Y scrambling to investigate whether the models did indeed capture the underlying relationships in the data rather than capturing chance correlation between the independent descriptors and dependent response data (Rücker *et al.*, 2007) as described in Chapter 6 Section 6.1.8.

### **7.2. Results**

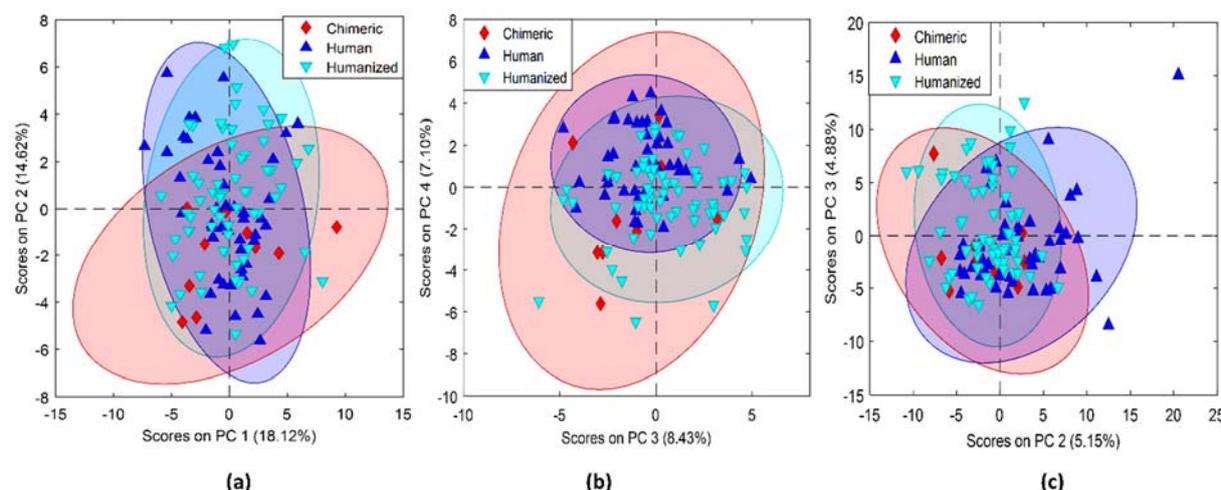
#### **7.2.1. Exploratory Analysis**

Similar to the exploratory analysis performed on the primary sequence dataset explained in the previous chapter, PCA was performed on the 3D structure-based datasets. As shown in Figure 7.3 an increased separation between mAbs based on light chain isotype can be observed with increasing resolution of the dataset, chain being most global and substructure being most local. . At this stage only those mAbs with a kappa light chain were selected (n=121) and lambda light chain mAbs were removed from the datasets (n=13).



**Figure 7.3** Score plots generated after performing PCA on (a) Chain, (b) Domain and (c) Substructure datasets for 134 mAbs. X axis represents Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captures by the respective PC. Coloured ellipses indicate the 95% Confidence limits of the corresponding class.

Grouping based on species was not as discernible for the 3D structure-based descriptors when compared to primary sequence descriptors (Chapter 6 Figure 6.3) as shown in Figure 7.4. As the output data was based on mAbs with their respective Fab regions grafted onto an IgG1 constant domains IgG1-Kappa-Humanized mAbs were chosen for further model development (n=44). This was also done for keeping the methodology consistent with that developed for primary sequence-based descriptors.



**Figure 7.4** Score plots generated after performing PCA on (a) Chain, (b) Domain and (c) Substructure datasets for 121 mAbs. X axis represents Principal Component (PC) 1 and Y axis represents PC2. The percentage values indicate the percentage variance captures by the respective PC. Coloured ellipses indicate the 95% confidence limits of the corresponding class.

### 7.2.2. Variable reduction and selection

An approximate 50% reduction in variables was observed for both Chain and Domain dataset following V-WSP reduction whereas for the substructure dataset there was a 40% reduction in the number of variables. The associated correlation coefficient threshold as well as overall Procrustes indices are reported in Table 7.2. The V-WSP reduced datasets were then divided into training (80%) and test set (20%) using Kennard Stone Algorithm. Furthermore, GA-PLS based variable selection led to a further reduction of variables resulting in approximately 20%, 13% and 9% of the variables from the original dataset being retained in the GA selected dataset for Chain, Domain and Substructure datasets respectively. The number of unique models generated, the average RMSE values of these models as well as the number of generations required to reach convergence are reported in Table 7.3.

**Table 7.2** Variable reduction based on the V-WSP reduction algorithm indicating the selected thresholds and corresponding Procrustes index as well as the final number of descriptors in the reduced dataset

			Threshold		
			Chain	Domain	Substructure
HC	VH	FW1	0.9	0.9	0.56
		CDR1			0.77
		FW2			0.51
		CDR2			0.76
		FW3			0.51
		CDR3			0.96
		FW4			0.63
	CH1	A		0.5	0.73
		B			0.63
		C			0.55
		D			0.59
		E			0.5
		F			0.54
		G			0.66
LC	VL	FW1	0.94	0.94	0.74
		CDR1			0.5
		FW2			0.6
		CDR2			0.55
		FW3			0.75
		CDR3			0.77
		FW4			0.51
	CL	A		0.53	0.57
		B			0.5
		C			0.61
		D			0.5
		E			0.58
		F			0.57
		G			0.62

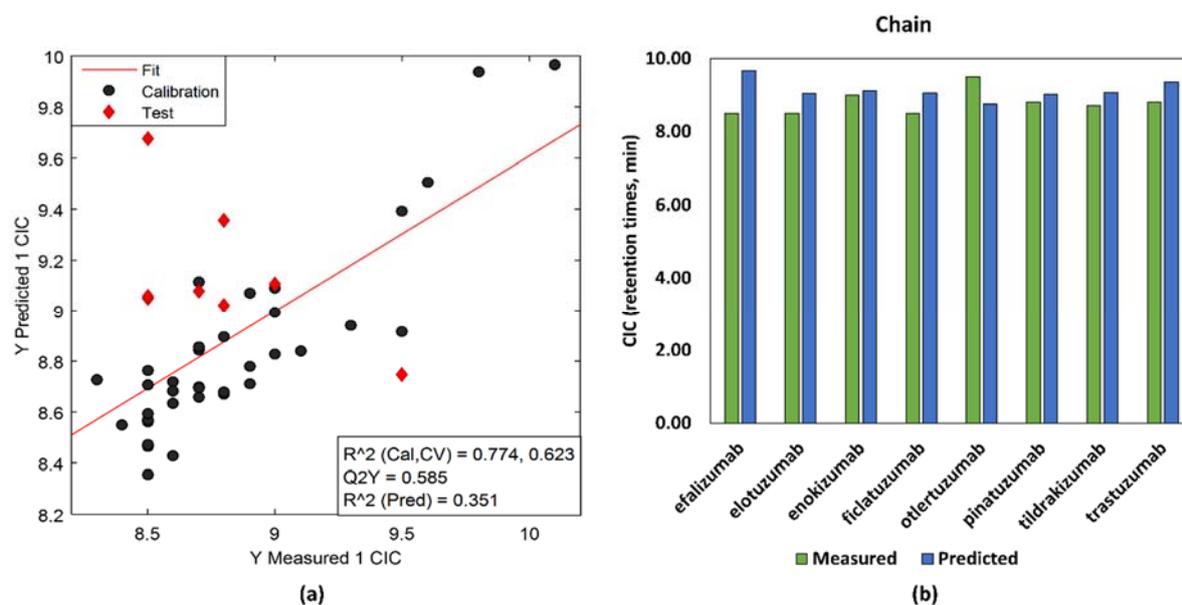
<b>Number of descriptors</b>	102	134	886
<b>Number of reduced descriptors</b>	51	66	348
<b>Overall Procrustes Index</b>	0.0935	0.0976	0.14

**Table 7.3** Results of GA-PLS based variable selection method.

<b>Datasets</b>	<b>No. of unique models</b>	<b>Best fit of models (RMSE CV)</b>	<b>No. of descriptors selected</b>	<b>No. of generations</b>
Chain	53	0.266098	14	21
Domain	70	0.273	18	21
Substructure	89	0.168	76	50

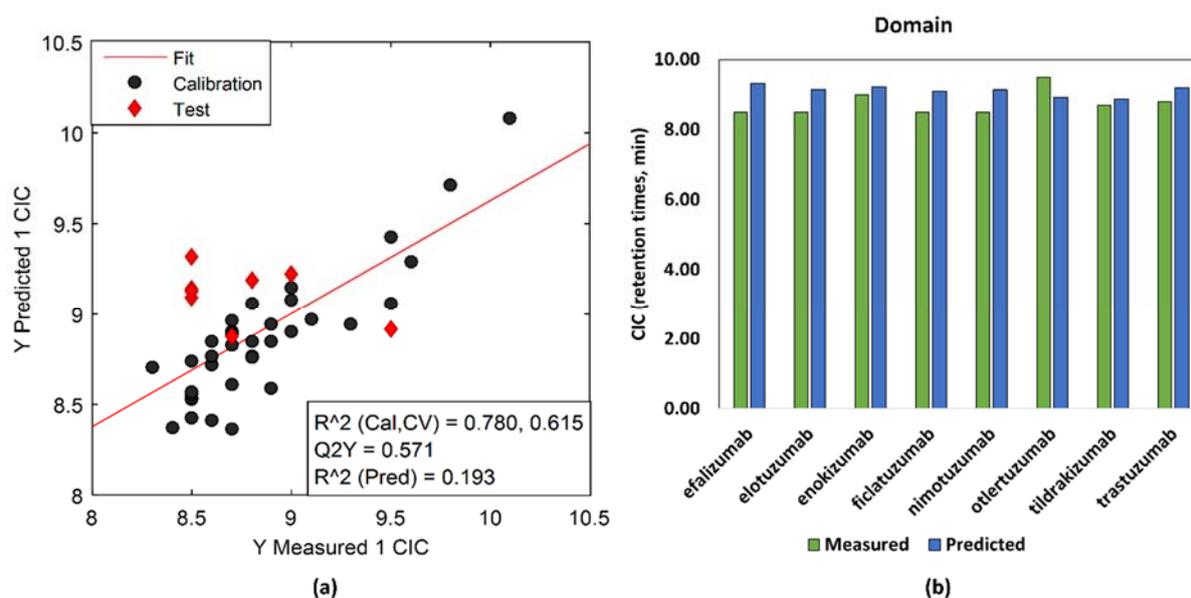
### 7.2.3. Model performance

Figure 7.5 shows the overall model summary for the GA variable selected PLS model (3 LVs) for the domain dataset. There is more dispersion of the samples from the parity line throughout the data distribution with slightly more dispersion in the lower extreme of the data i.e. for lower retention times and this is where most samples lie due to the skewed nature of the data distribution (Figure 7.5a). Efalizumab and olertuzumab in particular had larger differences between their measured and predicted values (Figure 7.5b). Removing these samples however did not result in improvement of the model based on the model performance criteria such as  $R^2$  and RMSE.



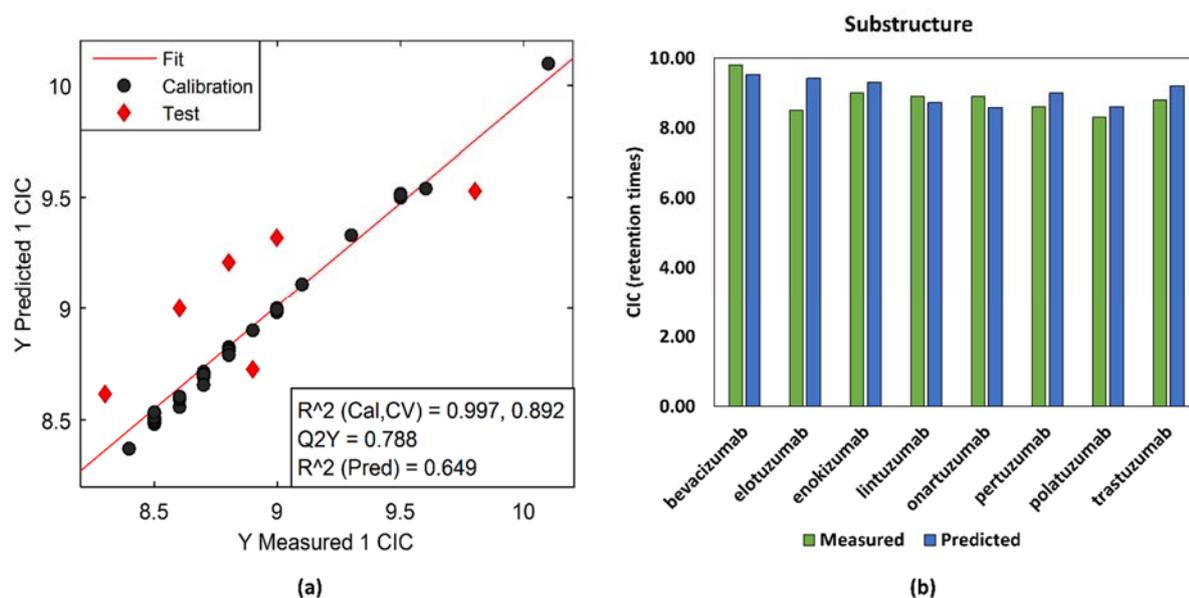
**Figure 7.5** Model summary of Chain dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.

A similar trend was observed for the domain dataset (PLS models with 3LVs) where the data was dispersed from the parity line, however to a slightly lesser extent than for the chain dataset (Figure 7.6a). Efalizumab and olertuzumab in particular had larger differences between their measured and predicted values in the domain dataset as well (Figure 7.6b). Removing these samples however did not result in improvement of the model.

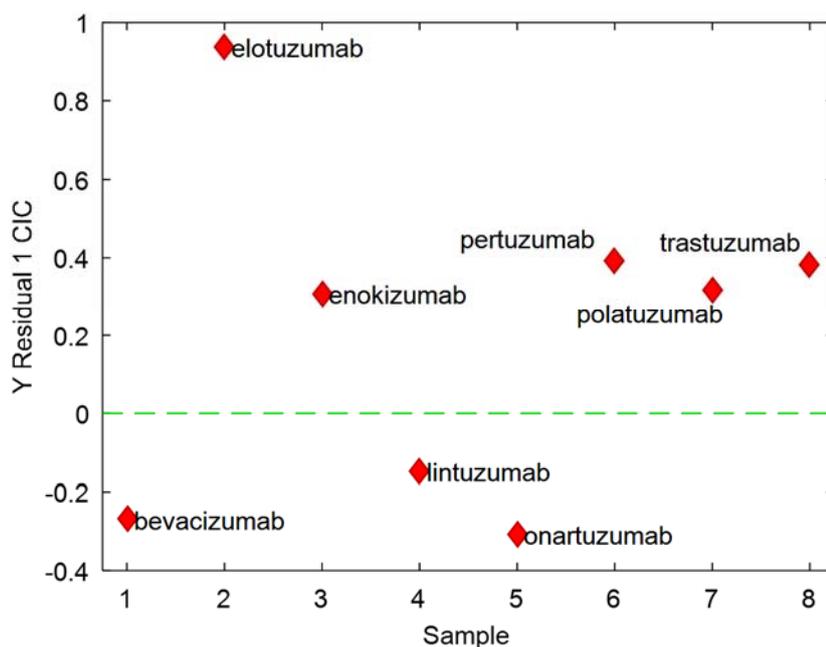


**Figure 7.6** Model summary of Domain dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.

The GA selected variable based PLS model (8 LVs) summary for the Substructure dataset is shown in Figure 7.7. Figure 7.7a depicts that the model fit was best throughout the distribution, especially the higher extreme. For the test set predictions, the predicted values of retention times were within 20 seconds of experimental values (7.7b). Elotuzumab was removed from the test set as it had unusually high Y residuals (Figure 7.8).



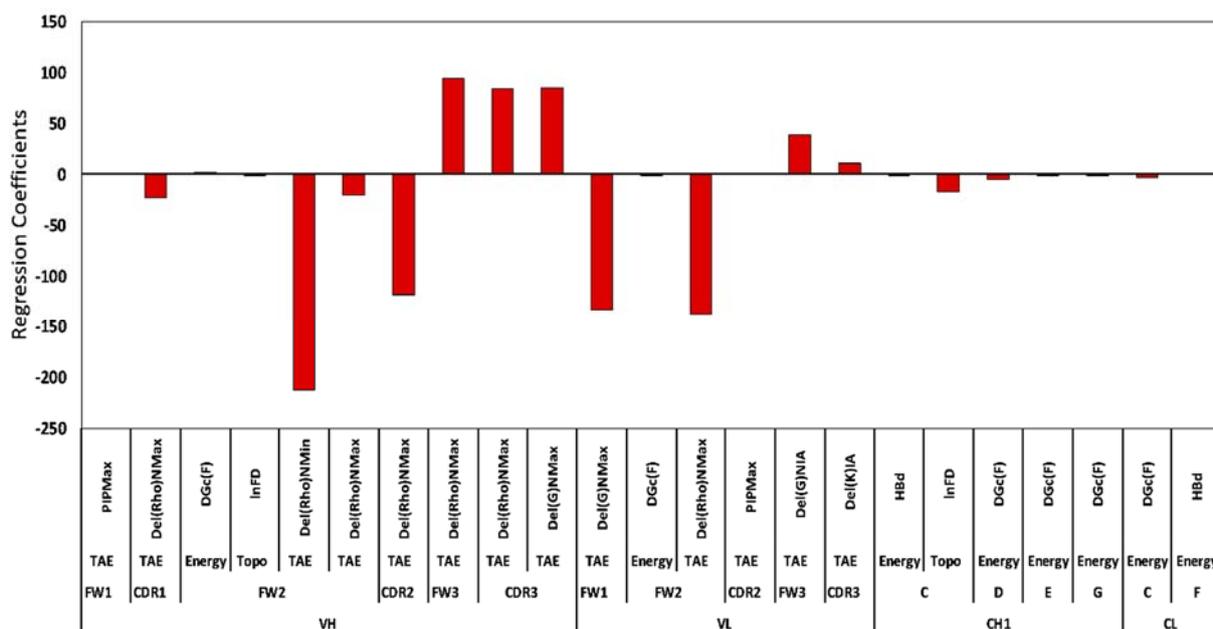
**Figure 7.7** Model summary of Substructure dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.



**Figure 7.8** PLS model-based Y residuals of test set samples of mAbs based on GA-selected variables substructure datasets.

The regression coefficients of variables  $>0.1$  are shown in Figure 7.9. The variables of the constant domains (CH1 and CL regions) do not have a major influence on the model. The maximum contribution arises from the TAE variables of the VH region followed by the VL region. Del(Rho)N descriptor fields distinguish between loose regions of polarizable electron density from more tightly held regions. These values are much smaller over electron-rich pi

systems and aromatic rings than over polarised or electron deficient alkyl carbons. The Del(K)N and Del(G)N descriptors describe the difference in polarizability and hydrophobicity of molecular regions i.e. more negative ranges indicate that the regions are more hydrophobic and less susceptible to electrophilic attack. In the VH domain the TAE properties FW2, CDR2, FW3 and CDR3 regions have the highest regression coefficients. In the VL region FW1 and FW2 have the highest contributions.



**Figure 7.9** Regression coefficients of GA selected variables for Substructure dataset model.

#### ***7.2.4. Benchmarking of datasets and models.***

All the models generated for the different datasets were benchmarked based on the OECD guidelines for QSAR model validation (Veerasamy *et al.*, 2011; Organisation for Economic and Development, 2014). The models were considered acceptable for the model fit if Calibration  $R^2 > 0.6$ , Cross validation  $R^2 > 0.6$  and  $Q^2 > 0.5$  and finally External test prediction  $R^2 > 0.6$ . The models were further benchmarked for numerical accuracy wherein a  $RMSE < 0.3$  is usually recommended. However previous studies have shown that an  $RMSE < 10\%$  of the output range is considered acceptable if accompanied by external validation especially for QSAR models developed for early stage screening (Alexander *et al.*, 2015). The performance metrics for the original, V-WSP reduced and GA selected Cross validated PLS models are reported in Table 7.4.

**Table 7.4** Benchmarking of datasets based on Calibration, Cross validation and external test set prediction metrics. Models that have passed QSAR validation criteria are shown in green, those with moderate performance are indicated in Yellow and those that have failed the QSAR validation criteria are indicated in red.

			Chain	Domain	Substructure
Calibration	Original	R <sup>2</sup>	0.42	0.46	0.99
		RMSE	0.27	0.26	0.01
		Bias	0.00	0.00	0.00
	V-WSP reduced	R <sup>2</sup>	0.72	0.41	0.91
		RMSE	0.22	0.31	0.11
		Bias	0.00	0.00	0.00
	GA selected	R <sup>2</sup>	0.77	0.78	0.99
		RMSE	0.19	0.19	0.02
		Bias	0.00	0.00	0.00
Cross validation	Original	R <sup>2</sup>	0.01	0.02	0.09
		Q <sup>2</sup>	-0.59	-0.68	-0.03
		RMSE	0.44	0.46	0.36
		Bias	0.01	0.01	0.02
	V-WSP reduced	R <sup>2</sup>	0.22	0.14	0.05
		Q <sup>2</sup>	0.02	0.06	-0.14
		RMSE	0.41	0.39	0.42
		Bias	-0.02	0.00	0.01
	GA selected	R <sup>2</sup>	0.62	0.61	0.89
		Q <sup>2</sup>	0.59	0.57	0.79
		RMSE	0.26	0.26	0.18
		Bias	-0.01	0.01	-0.01
Test	Original	R <sup>2</sup>	0.00	0.17	0.54
		RMSE	0.57	0.49	0.34
		Bias	0.11	0.01	0.03
	V-WSP reduced	R <sup>2</sup>	0.23	0.19	0.32
		RMSE	0.49	0.48	0.39
		Bias	0.27	0.22	0.15
	GA selected	R <sup>2</sup>	0.35	0.19	0.65
		RMSE	0.62	0.54	0.32
		Bias	0.35	0.36	0.16

### 7.2.5. Y randomisation

As the GA-PLS selected substructure dataset-based model passed the OECD criteria for model performance Y randomisation and permutation test was performed on this model. The R<sup>2</sup> and Q<sup>2</sup> values of the substructure dataset following y randomisation was 0.043 and -0.343. This

reinforces the fact that the model performance was not just based on chance correlation but due to an underlying causal relationship between descriptors and response as the  $R^2$  and  $Q^2$  values are below 0.5 (Guha and Jurs, 2005). There is a significant difference between the original models and those generated followed by Y scrambling (within the same range of output values) as well. The permutation analysis reports the results of non-parametric statistical hypothesis tests as shown in Table 7.5 for both the calibration (self-prediction) as well as cross validation model comparison. Values lower than 0.05 indicate that the models based on GA variable selection are significantly different from the Y randomised models at 95% confidence level. This indicates that the models generated in this study are capturing an underlying relationship between the descriptors and the response and not a chance/random correlation.

**Table 7.5** Results of permutations tests for GA selected substructure datasets.

	<b>Wilcoxon</b>	<b>Sign test</b>	<b>Rand t-test</b>
Self-Prediction	0.000	0.001	0.005
Cross-Validated	0.000	0.001	0.005

### 7.3. Discussion

In this particular study development of a QSAR model was attempted using 3D structure-based descriptors of Fab regions of mAbs with CIC as response. 3D structures were first generated for the mAbs using homology modelling followed by molecular dynamic simulation, the details for which are described in section 7.1.2. As elucidated in section 7.1.3, three different datasets with varying nature of localisation and resolution were generated from the PDB structures of the Fab fragments of 134 mAbs.

A similar trend to that seen from the behaviour of primary sequence-based descriptors was observed for the 3D structure-based ones with regards to resolution. The descriptors that were most global and had least resolution, chain dataset in this case, performed poorly when compared to ore local and clustered descriptor dataset such as the substructure-based ones (Table 7.5). This has been observed previously in QSAR studies involving Fab variants (Robinson *et al.*, 2017) wherein the local and cluster based descriptors were identified to be more important than global ones. Within the substructure-based descriptors, the electronic and charge-based properties of both framework and CDR regions were highlighted mainly for the VH and VL domain (Figure 7.8). This is consistent with results from the previous chapter

where descriptors describing similar electronic and charge-based properties were selected. This is also consistent with previous studies where QSAR model developed on mAb 3D structures based descriptors using chromatographic behaviour and retention times as responses (Chung *et al.*, 2010; Robinson *et al.*, 2017). There was a greater influence of TAE descriptors on the model, as reflected by their respective regression coefficients (Figure 7.5). This has also been shown in previous studies wherein machine learning based models were developed to predict chromatographic retention times (Song *et al.*, 2002; Ladiwala *et al.*, 2006). QSAR models have also been developing for understating how orientation affects behaviour of mAbs in ion exchange chromatography (Kittelmann *et al.*, 2017).

However, when comparing the model performance metric between the best performing primary sequence-based descriptors and 3D structure-based descriptors, the former had better performance compared to the latter. This could be due to the fact that the primary sequence-based descriptors offered more resolution than the 3D structure-based ones. There is increased complexity in developing 3D structure using homology modelling and molecular dynamics simulation. For homology modelling the framework chosen for assembly of the protein structures is important. While in the case of mAbs this is relatively straight forward due to the high degree of similarity between mAbs, further studies should be conducted for looking at the effect of different framework structures chosen for homology modelling, however, this would then hinder the high throughputness of structure development for further QSAR applications (Šali *et al.*, 1995; Fiser and Šali, 2003; Eswar *et al.*, 2006). For molecular dynamic simulations additional factors must be taken into consideration such as the properties of the aqueous solution including pH, ion concentration, temperature etc (Van Der Spoel *et al.*, 2005). These will have to be optimised on a case to case basis with respect to the conditions of the bioprocessing step in question and opens avenues for a multitude of predictive models to be developed per unit step. There is also the additional caveat of the time scales involved in the generation of these 3D structures using molecular dynamic solutions, however advancements such as coarse grain simulations as well as advanced Graphics Processing Unit (GPU) could speed up the process considerably (Hospital *et al.*, 2015).

Characterising antibody cross reactivity from docking or cavity mapping studies could be potential next steps in this study. Indeed antibody-antigen interaction characterisation using molecular dynamics simulation has been reported in previous studies (Heymann and Grubmüller, 2001; Castellanos *et al.*, 2017). This would also allow for development of

descriptors which are more cluster based and have higher resolution than the current substructure dataset (Robinson *et al.*, 2017).

#### **7.4. Chapter summary**

To summarise, this chapter described generation of mAb structural descriptors which were extracted from 3D Fab structures generated via homology modelling and molecular dynamics simulation. These structural descriptors were generated for three data blocks: Chain, Domain and Substructure and they were modelled against CIC retention times. As seen in the case of primary descriptors resolution and type of descriptors selected were similar i.e. local and cluster-based Substructure dataset with electronic and charge-based descriptors were identified to be important. Most of these descriptors of importance arise from the hypervariable regions of mAbs similar to primary sequence-based descriptors.

The next chapter looks at the effect of combination the best primary based descriptor dataset (running sum) and the best 3D structure-based dataset (substructure). Two different data augmentation methods will be explored, and the models developed on the each of these new combined datasets will be assessed with CIC as the response variable.

## Chapter 8: QSAR model development based on a combination of primary and 3D descriptors.

The previous chapters examined the QSAR model development for mAb developability using primary sequence-based descriptors and 3D structure-based descriptors separately. This chapter looks at the combination of descriptors from the best primary sequence-based descriptor i.e. Running sum dataset and the best 3D structure-based descriptor set: Substructure dataset. Furthermore, two different methods of descriptor combination were also examined. Model performance was assessed using the same benchmarking metrics used in previous methods.

A combination of sequence and structural based features have been used in studies for achieving better accuracy and predictability of prediction models particularly in protein-RNA binding studies, recognition of native-like protein structures, identifying coreceptor usage for HIV-1 cell entry as well as for in silico design of antimicrobial peptides (Sander *et al.*, 2007; Jenssen *et al.*, 2008; Liu *et al.*, 2010; Dybowski *et al.*, 2011; Bozek *et al.*, 2013). Within the bioprocessing sector combination of feature sets have been used for developing machine learning models that predict chromatographic behaviour as well as microbial production of secreted proteins (Ladiwala *et al.*, 2006; van den Berg *et al.*, 2012).

### 8.1. Materials and Methods

#### 8.1.1. Data distribution

As explained in Chapter 7, the 3D structure of some mAb samples could not be generated due to their structure being different from the reference framework. Thus, to make both running sum and substructure datasets similar, atelizumab and teplizumab were removed. Thus 42 IgG1-Kappa-Humanized samples were used in this study. The distribution of samples into training and test set was done using Kennard stone algorithm applied to the Y dataset instead i.e. distance based splitting based on the CIC response values. This is shown in Table 8.1.

**Table 8.1** Data distribution of mAb samples into training and test set based on response dataset.

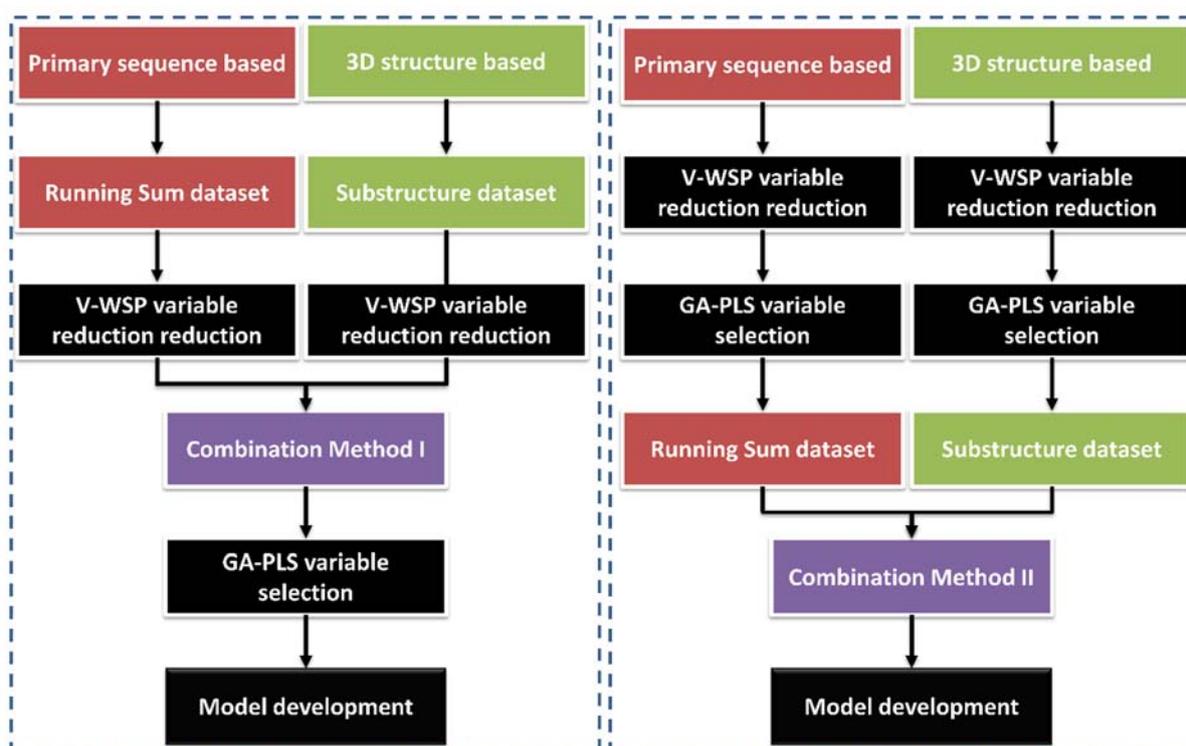
Training				Test
alemtuzumab	duligotuzumab	mepolizumab	palivizumab	dalotuzumab
bapineuzumab	efalizumab	mogamulizumab	parsatuzumab	etrolizumab
benralizumab	elotuzumab	motavizumab	polatuzumab	ficlatazumab
bevacizumab	enokizumab	nimotuzumab	ranibizumab	lintuzumab
certolizumab	epiratuzumab	obinutuzumab	tigatuzumab	onartuzumab

clazakizumab	farletuzumab	ocrelizumab	tildrakizumab	pertuzumab
codrituzumab	imgatuzumab	omalizumab	tocilizumab	pinatuzumab
dacetuzumab	lampalizumab	otlertuzumab	vedolizumab	trastuzumab
daclizumab	matuzumab	ozanezumab	veltuzumab	

Two different methods were subsequently utilised for combining the descriptors from the running sum and substructure dataset. These methods are explained in detail along with model performance assessment in the following sections.

### 8.1.2. Data augmentation methodology

Apart from the data combination method, explained in detail in their respective subsections below, the rest of the modelling methodology is the same as that described in previous chapters. The overall methodology of data augmentation is shown in Figure 8.1.



**Figure 8.1** Overall methodology describing the two data augmentation methods.

### 8.1.3. Combination Method I

In the first method of dataset augmentation the reduced primary sequence based running sum dataset and 3D structure-based substructure dataset were combined and then subjected to GA-PLS based variable selection as described in Chapter 6, Section 6.1.6. The cross-validation

method used was a random subset cross validation with 5 splits and 50 iterations (Minitab, 2014; Hahn and Valentine, 2016). A maximum of 10 latent variables were tested for each developed model and used to investigate the model error in order to choose an appropriate model complexity, i.e. number of latent variables. Models were further optimised by Genetic Algorithm-Partial least squares (GA-PLS) based supervised variable selection method. The parameters were set as follows: population size of 100; maximum generations of 100; mutation rate of 0.005, window width of 1; convergence rate of 80%; 30 initial terms, cross over of 2; random subset 5 fold cross validation with 10 iterations and data retreatment set for autoscaling (Hasegawa *et al.*, 1997; Andersen and Bro, 2010). The algorithm is terminated upon reaching convergence or reaching max generations as a starting point for variable selection. The model with the best fitness i.e. the lowest RMSE is chosen and subsequently the variables from that model is selected for further PLS analysis.

#### ***8.1.4. Combination method II***

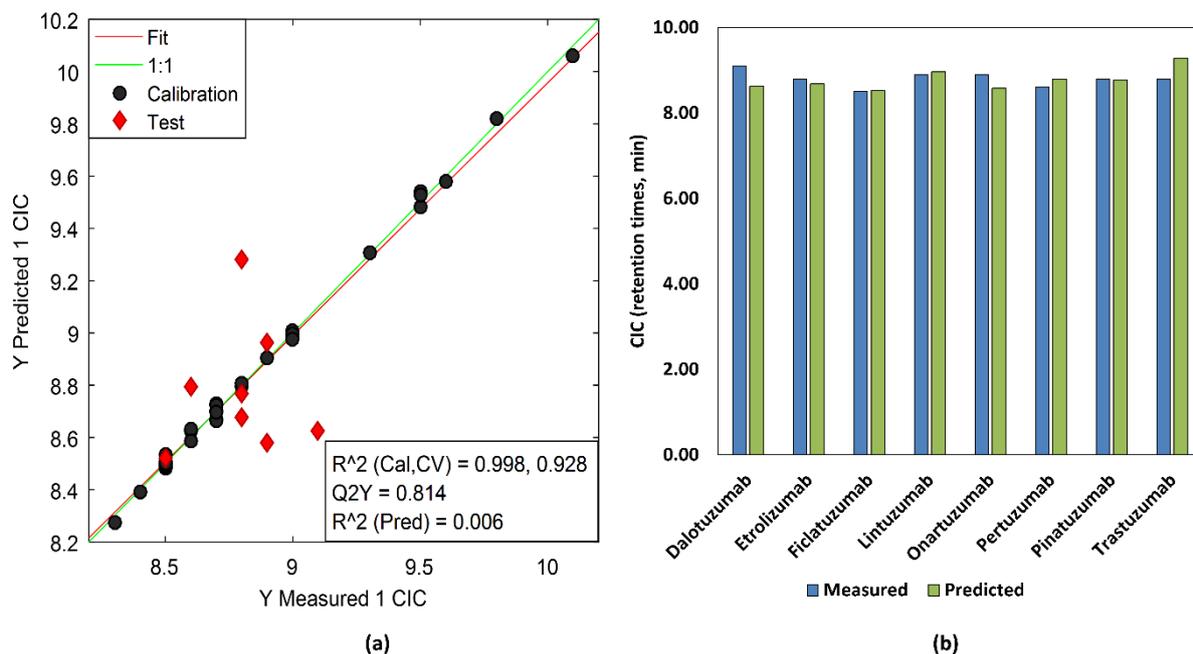
In the second method of data augmentation the GA-PLS selected descriptors of the primary sequence based running sum dataset and the 3 D structure-based substructure dataset were combined and later subjected to PLS modelling. The cross-validation method used was a random subset cross validation with 5 splits and 50 iterations (Minitab, 2014; Hahn and Valentine, 2016). A maximum of 10 latent variables were tested for each developed model and used to investigate the model error in order to choose an appropriate model complexity, i.e. number of latent variables.

## **8.2. Results**

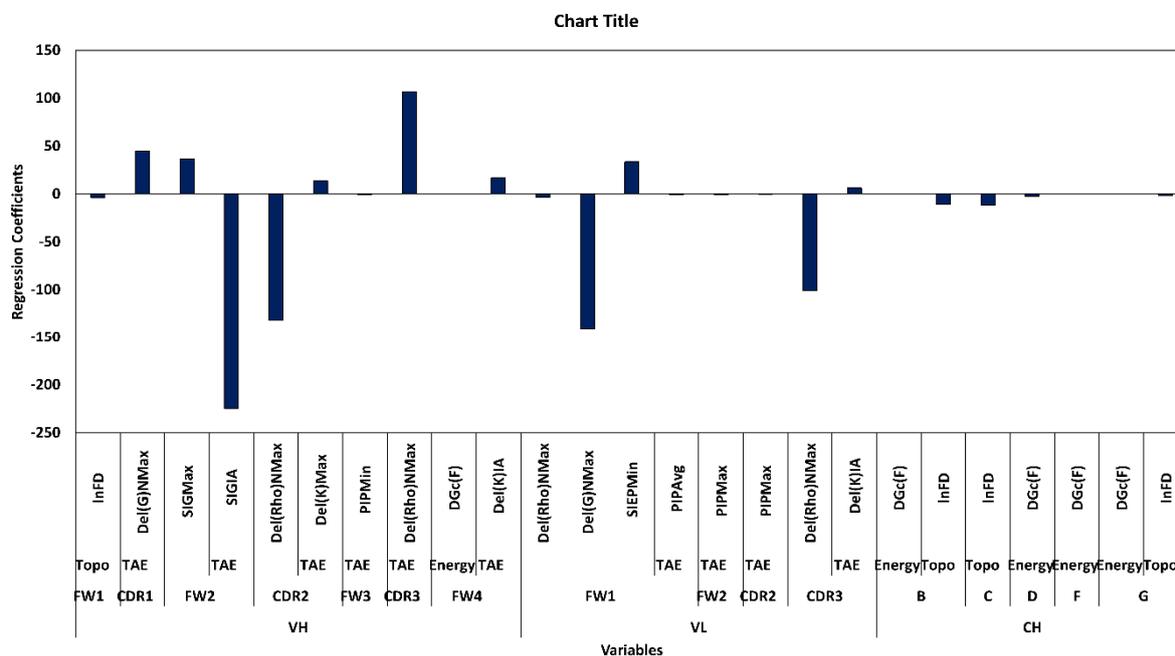
### ***8.2.1. Combination Method I - Model performance***

The GA-PLS based variable selection was terminated at generation 50. resulting in 94 unique models with the best fitness being 0.18 and the average fitness being 0.19. A total of 113 descriptors were selected following GA-PLS based variable selection. The summary of PLS model (3LVs) generated from the descriptors selected via the first data augmentation method is shown in Figure 8.4. A good model fit can be observed throughout the data distribution for the calibration model (Figure 8.4a). All the test set predictions fall within 30 seconds of the experimental values i.e. within a 5% range of the measure values (Figure 8.4b) however the prediction performance is relatively poor as indicated by prediction  $R^2$ . Figure 8.5 shows the regression coefficients for the variables chosen in the model. The major contributions for this

model arise from the 3D structure-based descriptors mainly from the Framework 2 (FW2), CDR2 and CDR3 substructures of the VH region as well as FW1 and CDR3 of the VL domain.



**Figure 8.2** Model summary of model developed using combination method I. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.



**Figure 8.3** Regression coefficients of variables present in model from running sum and substructure dataset for all Fab domains.

The results of the permutation test are given in Table 8.2. The augmented dataset in this case does not show a significant difference between the original models and those generated followed by Y scrambling (within the same range of output values). Values greater than 0.05

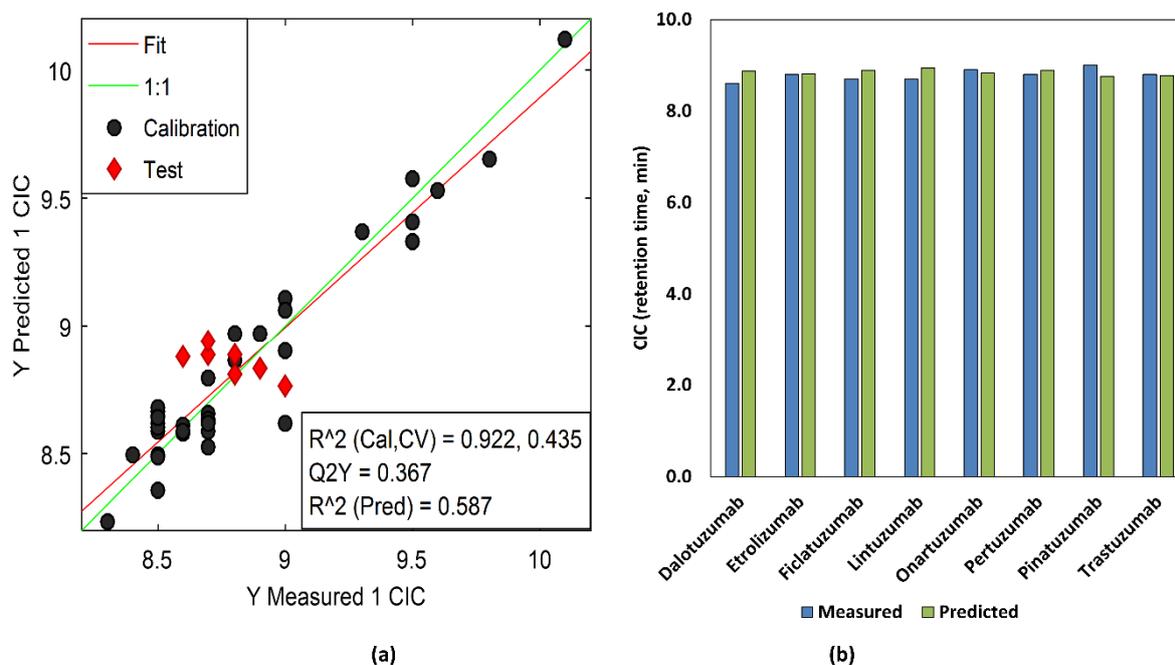
indicate that the models based on GA variable selection are not significantly different from the Y randomised models at 95% confidence level.

**Table 8.2** Results of permutations tests for Combination method I dataset.

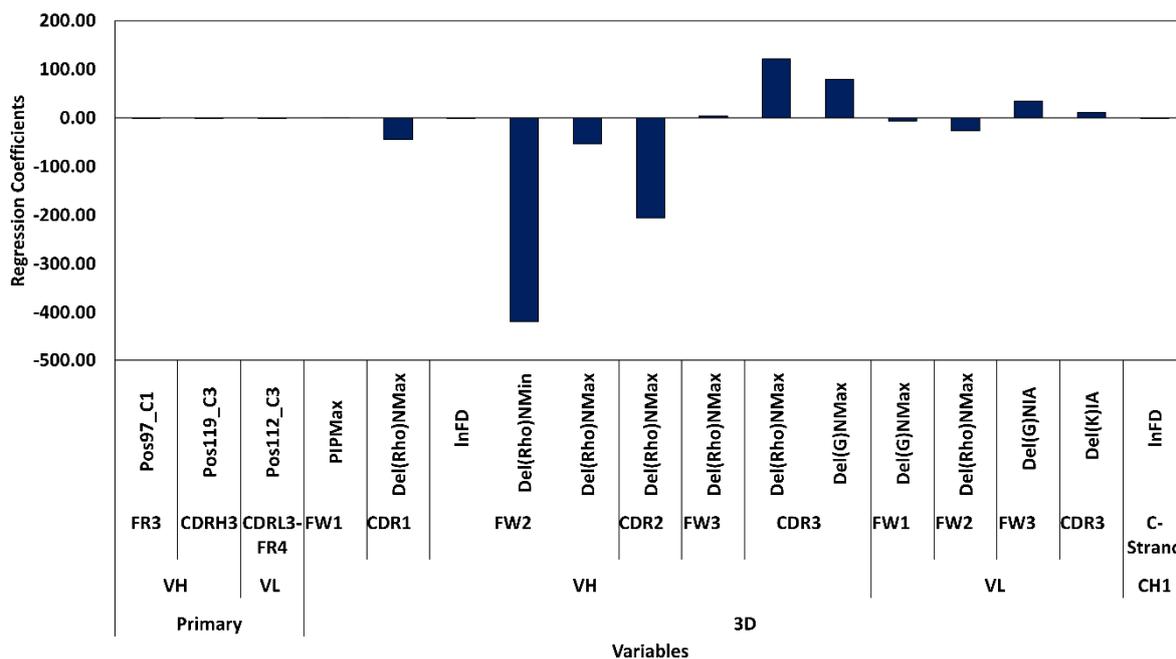
	Wilcoxon	Sign Test	Rand t-test
Self-Prediction:	0.061	0.172	0.076
Cross-Validated:	0.042	0.076	0.040

### 8.2.2. Combination Method II - Model Performance

A total of 216 descriptors were present in this combined dataset, 140 from Running sum and 76 from the substructure-based dataset. Figure 8.2 shows the overall model summary of the PLS model developed following combination method II. There is more dispersion of the samples from the parity line in the lower extreme of the data i.e. for lower retention times and this is where most of the samples lie due to the skewed nature of the data distribution (Figure 8.2a). All the test set predictions fall within 30 seconds of the experimental value i.e. within a 5% range of the measure values (Figure 2.3b). Figure 8.3 shows the regression coefficients for the variables chosen in the model. The major contributions for this model arise from the 3D structure-based descriptors mainly from the Framework 2 (FW2), CDR2 and CDR3 substructures of the VH region.



**Figure 8.4** Model summary of model developed using combination method II. (a) Measured vs predicted values for training and test set model. (b) Test set predictions.



**Figure 8.5** Regression coefficients of variables present in model from running sum and substructure dataset for all Fab domains.

The results of the permutation test are given in Table 8.3. The augmented dataset shows a significant difference between the original models and those generated followed by Y scrambling (within the same range of output values). The null hypothesis of the test is that there is no significant difference between the GA selected dataset and the Y randomised dataset, both built under the same conditions. The permutation analysis reports the results of non-parametric statistical hypothesis tests as shown in Table 8.3 for both the calibration (self-prediction) as well as cross validation model comparison. Values lower than 0.05 indicate that the models based on GA variable selection are significantly different from the Y randomised models at 95% confidence level.

**Table 8.3** Results of permutations tests for Combination method II dataset.

	Wilcoxon	Sign Test	Rand t-test
Self-Prediction:	0.000	0.002	0.005
Cross-Validated:	0.000	0.002	0.005

### 8.3. Discussion

The aim of the above study was to assess the effect of using a combined dataset that contains both primary sequence and 3D structure-based descriptors for developing a QSAR model with CIC retention time as response. As seen from the above results the combined datasets, irrespective of the method of combination, do not perform as well as the individual primary sequence-based or 3D Structure based descriptors sets (Figure 8.2 and 8.4). Of the two combination methods tested, combining the best descriptors from the GA selected running sum dataset and GA selected substructure dataset performed slightly better than the other (Figure 8.4). Neither of the datasets satisfies the criteria set down for QSAR model validation guidelines set by OECD (Organisation for Economic and Development, 2014). This has been shown in previous studies pertaining to other protein studies, wherein a combination of structural and sequence based descriptors did not yield an improvement in prediction performance when compared to the models developed separately (Bozek *et al.*, 2013).

Another interesting factor that emerged from the above data augmentation study was that the 3D descriptors appear to outweigh the primary sequence-based descriptors in terms of the contribution to the model predictive capacity (Figure 8.3 and 8.5). The TAE descriptors in particular have higher regression coefficients compared to other descriptors and this type of descriptors has been used extensively in 3D QSAR docking studies as they describe molecular electron densities and allow for assessing electrostatic interactions (Kharkar *et al.*, 2008; Xu *et al.*, 2012; Lorca *et al.*, 2018). Therefore, interpretability of these descriptors could be accentuated in molecular docking studies of mAbs. However larger datasets would be required to accentuate the above findings. 3D descriptors, even at substructure level, capture more global and holistic properties that differ between mAbs, but these may not correlate to a response. Furthermore, this could potentially lead to confounding of information that could be relevant for the response. Previous studies, wherein QSAR models have been successfully developed using the Fab 3D structure based descriptors, reflect the superior performance of localised and cluster descriptors over global descriptors (Robinson *et al.*, 2017). Another complexity arises from the nature of the environmental setting where the 3D structure was developed via molecular dynamics simulation as explained in the previous chapter (Castellanos *et al.*, 2017). Primary descriptors, such as Running Sum dataset, on the other hand are local and/or clustered, which allows the model to capture any potential correlation between the descriptors and response. This has been reflected in previous studies where QSAR models have

been successfully developed for assessing potential aggregation propensities of mAbs based on primary sequence-based descriptors (Obrezanova *et al.*, 2015). Studies that use a combination of 3D and primary descriptors are rare in the field of mAbs and therefore further preselection methods would have to be employed prior to data augmentation method. A larger sample size would be required to adequately model and understand the effect of different descriptors, sequence or structure based, with respect to a particular response.

#### **8.4. Chapter summary**

In this chapter a combination of primary and structural descriptors was used for developing a QSAR model with CIC retention time as response. Different data augmentation methods were used wherein the combination of best primary and structural descriptors from individual models performed better compared to combining all primary and structural descriptors and performing model optimisation steps. The descriptors selected were related to electronic and charge-based properties. The structural descriptors outweighed the primary sequence-based ones however, the overall model performance was less than that of the models developed individually using primary sequence-based descriptors and structural descriptors respectively as seen from previous chapters. This implies that careful selection of variables based on expert knowledge should be performed such that the descriptors selected capture both structural and sequence-based aspects of functional characterisation.

## **Chapter 9: A human skin explant (Skimune™) based QSAR model for early adverse effect prediction of monoclonal antibody therapeutics.**

Based on the results observed in the previous chapters the primary sequence-based descriptors showed better performance when compared to that of 3D structure-based descriptors. Additionally, 3D structures could not be generated for some samples resulting in further reduction of the sample size for generating the substructure-based descriptors. Thus, the best performing primary sequence-based dataset, running sum, was selected for developing a QSAR model with a human skin explant assay results as the response variable.

From a clinical perspective mAbs are highly desirable owing to their high target specificity and tolerance within the human system. The principles for pharmacological and safety testing of mAbs are derived from International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) safety guidelines S8, S9 and S6 which lay down the minimum requirements for immunotoxicology studies, non-clinical evaluation for anticancer pharmaceuticals and biotechnological product testing respectively (Guideline, 1997; Guideline, 2005; Ich, 2008). However, the current paradigm has limitations as seen in the case of adverse effects elicited following the administration of mAb based therapeutics such as a cytokine storm induced by TGN1412 as well as formation of anti-mAb antibodies during clinical trials (Stebbing *et al.*, 2007). Clinical toxicities associated with mAbs are usually immune related adverse effect such as cytokine storms, immunogenicity, anaphylactic release etc (Kizhedath *et al.*, 2016).

For nonclinical safety testing there are challenges, such as regulatory requirements, around in vitro testing and cross reactivity of mAbs only to non-human primates (Sewell *et al.*, 2017). This is again selective as seen in the case of the TGN412, which caused no undesired reaction in cynomolgus monkeys but elicited a cytokine storm in healthy volunteers. This reiterates the need for a fully functional immune system to be in places in safety testing models of mAbs. With the advent of minimum anticipated biological effect level (MABEL) dosing there is the advantage of reducing an adverse effect but a disadvantage of not reaching a pharmacologically effective dosage for humans (Muller *et al.*, 2009). Additionally, species cross reactivity, anti-mAb antibody formation, time frame of testing for eliciting these responses remain areas of concern (Baumann *et al.*, 2014).

Thus, there is a need for more complex, hybrid and comprehensive testing strategies that include in vitro, in vivo and in silico aspects to fill the gaps in non-clinical pharmacology and

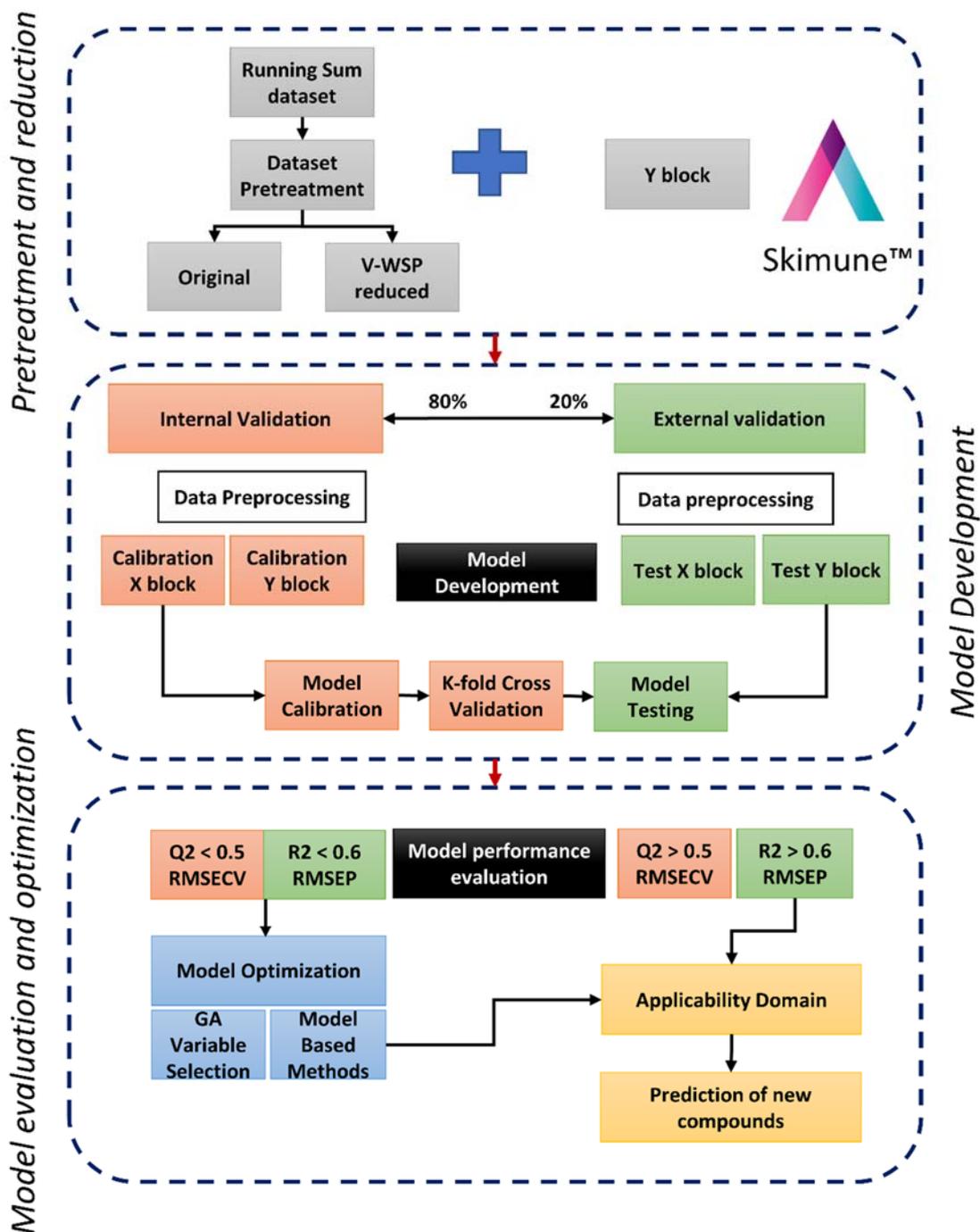
safety testing of mAbs especially since the age of biosimilars is just around the corner (Baumann *et al.*, 2014). Effective, high-throughput rapid screening of candidates based on adverse effect detection is required at an early stage to eliminate unsuitable candidates from proceeding to clinical trials. Appropriate and relevant experimental studies are of paramount importance in nonclinical safety testing as they also contribute to good datasets which can then be modelled. Computational toxicology tools like expert/hybrid systems provide a powerful complement to in vitro systems as they may allow for development of automated and reliable models for predicting toxicity or adverse effect of monoclonal antibody therapeutics. This study has focused on the development of Skimune™ based QSAR model development for early stage screening of mAbs based on descriptors derived from the primary sequence.

In this study the running sum dataset was generated using for variable regions of 15 mAbs the details of which are outlined in Table 9.1 (Ruiz-Blanco *et al.*, 2015). Skimune™ assay responses were used as a measure of hypersensitivity and this constituted the dependent variable or the y block (Ahmed *et al.*, 2016; Jain *et al.*, 2017). The relationship between the descriptors and the responses was captured by PLS and the datasets as well as models were benchmarked based on metrics relevant for the endpoint of the model i.e. early stage screening and identifying trends. The overall methodology is outlined in Figure 9.1.

**Table 9.1** Details of mAbs used and sample distribution for model development.

<i>Name</i>	<i>HC</i>	<i>LC</i>	<i>Species</i>	<i>Antigen</i>	<i>Skimune™ response (%)</i>	<i>Clinical incidence (%)<sup>+</sup></i>	<i>IMGT mAb* ID</i>
adalimumab	IgG1	κ	Homo sapiens	TNF	70	10	165
alemtuzumab	IgG1	κ	Humanized	CD52	70	10	11
basiliximab	IgG1	κ	Chimeric	IL2RA	54	1	148
bevacizumab	IgG1	κ	Humanized	VEGFA	50	10	24
certolizumab	IgG1	κ	Humanized	TNF	30	0.1	242
cetuximab	IgG1	κ	Chimeric	EGFR	50	10	151
golimumab	IgG1	κ	Homo sapiens	TNF	40	0.1	175
infliximab	IgG1	κ	Chimeric	TNF	30	0.1	156
muromonab	IgG2	κ	Chimeric	CD3E	90	10	132
natalizumab	IgG4	κ	Humanized	ITGA4	20	0.1	75
ofatumumab	IgG1	κ	Homo sapiens	MS4A1 CD20	50	10	194
panitumumab	IgG1	κ	Homo sapiens	EGFR	30	0.1	196
rituximab	IgG1	κ	Chimeric	MS4A1	50	10	161
trastuzumab	IgG1	κ	Humanized	ERBB2 CD340	60	10	97
TGN1412	IgG4	κ	Humanized	CD28	90	10	-b, **

<sup>b</sup>Sequence was sourced from patent., HC = Heavy chain isotype, LC = Light chain isotype <sup>+</sup> Upper limits taken from label information. \*(Lefranc *et al.*, 2009; Lefranc *et al.*, 2015), \*\*(Hanke *et al.*, 2014).



**Figure 9.1** Hybrid model development workflow outlining the different steps involved in pretreatment and variable reduction; model development followed by model evaluation and optimisation.

## 9.1. Materials and methods

### 9.1.1. Data collection and descriptor generation

Sequence information, substructure, species and phase of development of 15 mAbs have been collected from IMGT, literature, patents as well as from industrial partners (Apweiler *et al.*,

2004; Lefranc *et al.*, 2009). The FASTA format sequences of mAbs were read into Matlab and were subjected to multiple sequence alignment. The sequences were then split into domains and the variable region domains VH and VL were selected for further descriptor generation. 2530 features for Running Sum based dataset was generated as described in Chapter 5. The descriptors based on primary sequences provide information about the physicochemical properties of amino acids and these are outlined (Zaliani and Gancia, 1999; Tian *et al.*, 2007a). The automation of descriptor and dataset generation was done by MATLAB codes written by Micael Karlberg, Newcastle University.

#### ***9.1.2. Data curation and Variable reduction***

Following descriptor generation, the datasets were curated by and V-WSP reduction was performed as described in Chapter 6, Section 6.1.3. These datasets are referred to as ‘Original dataset’ and ‘V-WSP reduced’ dataset respectively.

#### ***9.1.3. Biological response data***

Hypersensitivity reactions were assessed using Skimune™, a non-artificial human skin explants based assay for safety and efficacy assessment of novel compounds and drugs, developed by Alcyomics Ltd (Ahmed *et al.*, 2016). For the Skimune™ assay, a percentage response was attributed based on number of grade 2 and above responses for all donors (n=10). The details of the responses used for regression analysis used is shown in Table 9.1.

#### ***9.1.4. Model development***

Responses ascertaining hypersensitivity were modelled against the descriptors generated using the PLS Toolbox from Eigenvector Research. Both the descriptors and responses were entered into Matlab, auto scaled (each descriptor was mean centred and scaled with its individual standard deviation) and subsequently subjected to PLS analysis. Both the descriptors and responses of the training set were inputted into Matlab and auto scaled separately i.e. for the training set the mean and standard deviation of the training set was used and for the test set the mean and standard deviation of the test set was used. The cross-validation method used was a random subset cross validation with 5 splits and 50 iterations (Minitab, 2014; Hahn and Valentine, 2016). A maximum of 10 latent variables were tested for each developed model and used to investigate the model error in order to choose an appropriate model complexity, i.e. number of latent variables. Models were further optimised by Genetic Algorithm-Partial least squares (GA-PLS) based variable selection method. The parameters were set as follows: population size of 100; maximum generations of 50; mutation rate of 0.005, window width of

1; convergence rate of 25; 30 initial terms, cross over of 2; random subset 5 fold cross validation with 50 iterations and data retreatment set for autoscaling (Hasegawa *et al.*, 1997; Andersen and Bro, 2010). The algorithm is terminated upon reaching convergence or reaching max generations. The model with the best fitness i.e. the lowest RMSE is chosen and subsequently the variables from that model is selected for further PLS analysis.

#### ***9.1.5. Model Performance metrics***

The datasets, models and outputs were benchmarked based on the following metrics for the multivariate regression analysis:  $R^2$  values based on Pearson correlation coefficient between observed and predicted values;  $Q^2$  values based on goodness of fit; Root mean square error of cross validation (RMSECV) and cross validation bias (Alexander *et al.*, 2015). These have been described previously in Chapter 6, Section 6.1.7. The model metrics were further assessed against the OECD guidelines for the Calibration, internally validated (Cross validation) and externally validate models (test set predictions) (Veerasamy *et al.*, 2011; Organisation for Economic and Development, 2014).

#### ***9.1.6. Y randomisation***

The developed models were further evaluated by performing Y randomisation and Permutation tests as described in Chapter 6, Section 6.1.8 to investigate whether the models did indeed capture the underlying relationships in the data rather than capturing chance correlation between the independent descriptors and dependent response data (Rücker *et al.*, 2007).

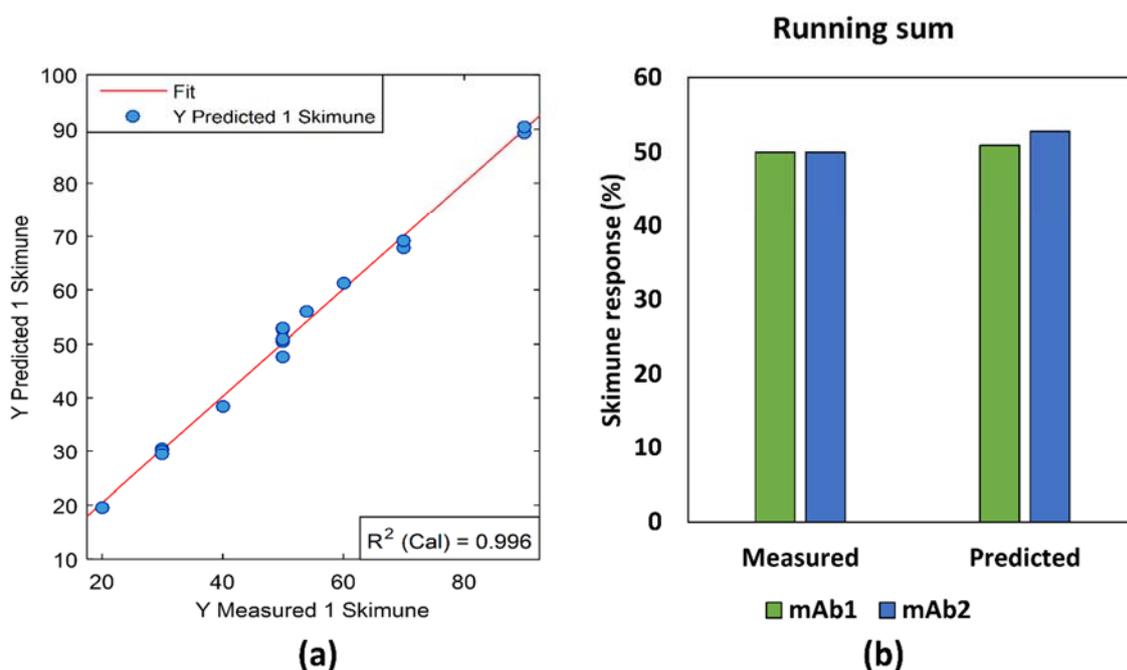
## **9.2. Results**

### ***9.2.1. Variable reduction and variable selection***

Considering the dimensionality of the data with a large number of variables and a limited number of observations, variable reduction and variable selection methods were used prior to model building. There was a 98% reduction in variables of the running sum dataset following V-WSP reduction. The correlation coefficient thresholds of VH and VL were both 0.50 and the overall Procrustes index was 0.09 resulting in 72 descriptors. Furthermore, GA-PLS based variable selection led to a further reduction of variables resulting in 1% of the variables from the original dataset being retained in the GA selected dataset for running sum dataset. The number of unique models generated was 61, the average RMSE values of these models was 0.82 and the number of generations required to reach convergence was 18. A final set of 21 descriptors was retained in the final GA selected running sum dataset.

### 9.2.2. Model performance evaluation

Figure 9.2 shows the predicted vs measured response plots for the calibration model of the running sum dataset for the GA selected set of variables. The model has a good fit as most points are close to the parity line. The test set predictions are within 3% of the measured Skimune™ responses.



**Figure 9.2** Model summary of Running Sum dataset model developed following GA-PLS selection of variables. (a) Measured vs predicted values for training and test set model. (b) Test set predictions

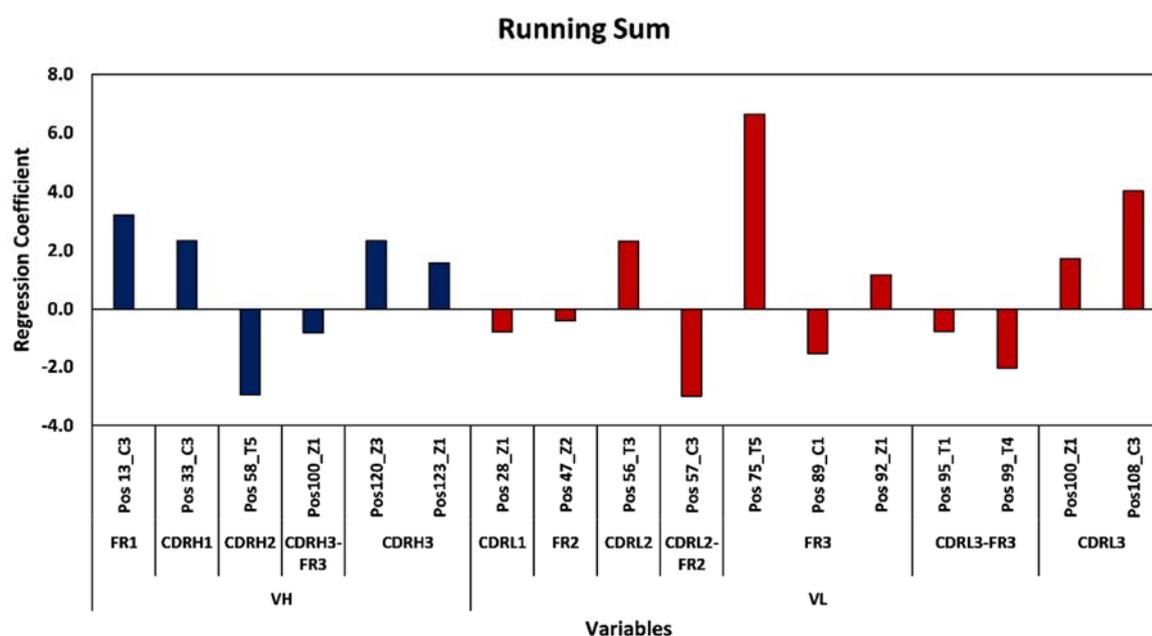
As the Running Sum datasets were generated following multiple sequence alignment, the relative positions of residues of the complementary determine region (CDR) and framework region (FR) that constitute the hypervariable region of mAbs needed to be identified. This has been shown in Table 9.2.

**Table 9.2** Relative positions of CDR and FR regions for heavy and light chain following sequence alignment.

Heavy	Indices	Light	Indices
FR1	1-21	FR1	1-22
Start	22-25	Start	23
CDRH1	27-37	CDRL1	24-34
End	38-39	End	35-37
FR2	40-46	FR2	38-47

Start	47-51	Start	48-49
CDRH2	52-70	CDRL2	50-56
End	71-73	End	-
FR3	74-99	FR3	57-87
Start	100-102	Start	88
CDRH3	103-119	CDRL3	89-98
End	120-123	End	99-102
FR4	124-130	FR4	103-110

The highest correlation coefficients from the VH region arise from the topological properties of CDRH2 as well as electrostatic potential based properties of FR1 and CDRH1. In the VL region the highest correlation coefficients arise from the topological properties of the FR3 region followed by electrostatic potential properties residues in CDR2-FR2 and CDRL3 regions.



**Figure 9.3** Regression coefficients of GA selected variables for Running Sum dataset model

### 9.2.3. Benchmarking of model performance metrics

The performance of calibration models was evaluated based on  $R^2$  and RMSE as shown in Table 9.3. The model performance acceptance criteria have been discussed in the previous chapters. The range for Skimune™ responses is 70 which means for the model to have high accuracy an RMSE of less than 7 would be preferred. The  $R^2$  and RMSE values of calibration model show marked improvement following V-WSP reduction and GA-PLS based variable selection (Table 9.3).

**Table 9.3** Model performance metrics of Calibration, Cross validation and external test set prediction. Models that have passes QSAR validation criteria are shown in green, those with

moderate performance are indicated in Yellow and those that have failed the QSAR validation criteria are indicated in red.

			Running Sum
Original	Calibration	R2	0.58
		RMSE	13.21
		Bias	0.00
	Cross validation	R2	0.15
		Q2	-0.99
		RMSE	28.70
		Bias	-0.30
	Test	RMSE	10.90
		Bias	8.89
V-WSP reduced	Calibration	R2	0.98
		RMSE	2.65
		Bias	0.00
	Cross validation	R2	0.04
		Q2	-0.36
		RMSE	23.76
		Bias	-1.81
	Test	RMSE	0.99
		Bias	-0.98
GA selected	Calibration	R2	0.99
		RMSE	1.36
		Bias	0.00
	Cross validation	R2	0.84
		Q2	0.80
		RMSE	8.93
		Bias	-1.50
	Test	RMSE	2.11
		Bias	1.88

The performance of the models generated was evaluated using R<sup>2</sup>, Q<sup>2</sup>, RMSE and Bias for the cross-validation models as shown in Table 9.3. Confidence intervals were calculated over the 50 cross validation iterations to give a measure of the model uncertainty i.e. the probability of the measured metric falling between the upper and lower confidence limits (97.5% and 2.5%) respectively).

**Table 9.4** Benchmarking of PLS model performance metrics for the original, V-WSP reduced and GA selected Cross validated models. CI: Confidence limits.

Metrics	GA Selected	Lower CI	Upper CI	V-WSP Reduced	Lower CI	Upper CI	Original	Lower CI	Upper CI
R <sup>2</sup>	0.84	0.83	0.86	0.03	0.02	0.04	0.15	0.12	0.18
Q <sup>2</sup>	0.80	0.78	0.82	-0.31	-0.36	-0.27	-0.99	-1.05	-0.92
RMSE	8.93	8.46	9.39	23.34	22.95	23.73	28.7	28.22	29.18
Bias	-1.40	-1.87	-1.14	-1.20	-1.51	-0.90	-0.3	-0.77	0.17

Both  $R^2$  and  $Q^2$  show a marked increase following GA selection of variables Table 9.4. Following variable selection using GA the model performance improves. As shown in Table 9.4 the models have a cross validation  $R^2$  greater than 0.6 which shows that the models are generally able to predict the trend of the responses. The RMSE values are close to the 10% cut off for the GA selected variable based model which indicates that the numerical accuracy of this model is considerably higher compared to the original and V-WSP reduced dataset. Correspondingly, the absolute cross validation bias is very low compared the original and V-WSP reduced datasets. This indicates that the model has relatively good cross validation  $R^2$ ,  $Q^2$  values and low RMSE, bias making them useful in predicting the trends and for ranking mAbs according to hypersensitivity. Owing to the objective of the modelling framework, i.e. early stage screening, the trend capturing  $R^2$  metric would be more relevant than RMSE which reflects the accurate numerical representation. This was also tested by PLSDA models for the datasets above (data not shown) wherein the models had very poor performance as the data was imbalanced (more negatives than positives). Therefore, PLS models were better at predicting the trends of responses and placing the extreme data points correctly.

#### 9.2.4. Model Robustness

Table 9.5 indicates the  $R^2$  and  $Q^2$  values of the Running Sum dataset following  $y$  randomisation. As seen from the metrics the  $y$  randomised models perform poorly. Additionally, permutation test was performed where the prediction residuals of the unperturbed and  $y$ -block shuffled data are compared (van der Voet, 1994; Thomas, 2003). The permutation analysis reports the results of non-parametric statistical hypothesis tests as shown in Table 9.5 for both the calibration (self-prediction) as well as cross validation model comparison. This indicates that the model generated in this study is capturing an underlying relationship between the descriptors and the response and not a chance/random correlation.

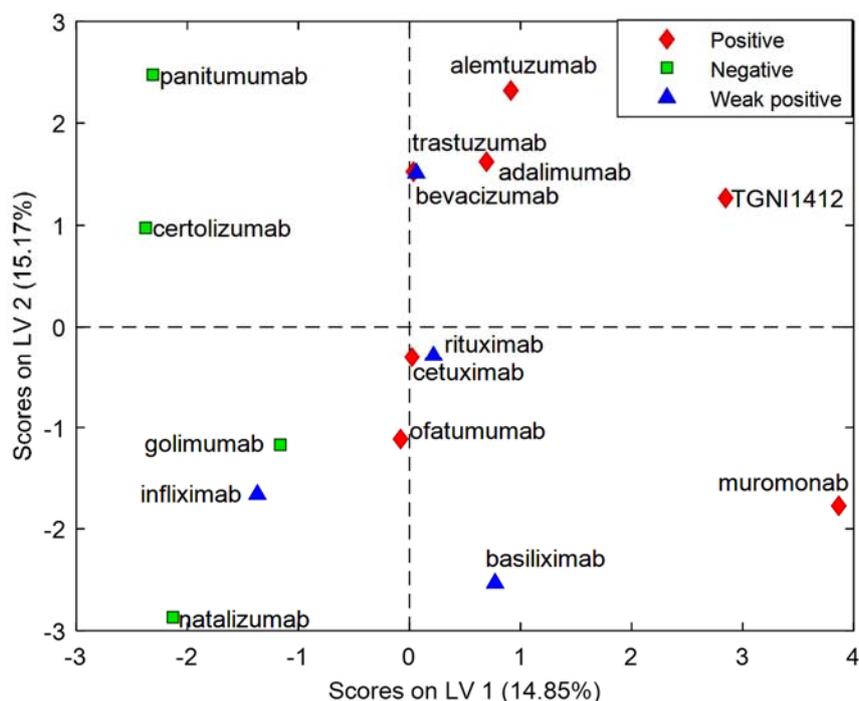
**Table 9.5** results of  $y$  randomisation and permutations tests for GA selected datasets.

Datasets	Self-prediction			Cross Validation			$R^2$ CV	$Q^2$ CV
	Wilcoxon	Sign Test	Rand t-test	Wilcoxon	Sign Test	Rand t-test	Mean	Mean
<b>Running sum</b>	0.009	0.035	0.011	0.003	0.021	0.006	0.1003	-0.535

#### 9.2.5. Correlation with Clinical incidence rates

The advantage of Skimune over other immunogenicity assay is the autologous nature of the assay as the results would be representative of clinical trials assay. This can already be seen

from the results of TGN1412 assays which passes in vivo testing and causes a cytokine storm in clinical trial patients (healthy volunteers). The clinical incidence of hypersensitivity associated adverse effects elicited by the mAbs in this study were retrieved from labels as accessed from the European Medicine Agency (EMA). Correlation analysis of Skimmune response with upper limit of clinical incidence rates were analysed using both Pearson and Spearman rho analyses. This yielded an  $R^2$  of 0.74 and 0.79 as well as p-value of 0.002 and 0.005 respectively. This indicates that responses used in this assay are reflective of clinical incidence rates and the QSAR model can thus be extended to include more mAbs. This can also be seen from the score plot of the PLS model (4 LVs) generated for Running Sum dataset wherein the classes correspond to the clinical classes, as defined by the incidence rates within the EMA labels, and separation between positives and negatives can be observed on the first latent variable of the model (Figure 9.4). The only exception to this is infliximab which is classified by the model as a weak positive. However, these immune related adverse events have been categorised as rare (<1%) based on clinical trial information (EMA). This separation or clustering wasn't observed on the other LVs.



**Figure 9.4** Score plot of GA selected variables based PLS model built on Running Sum dataset.

### 9.3. Discussion

Predictive models could substantially aid in safety pharmacology testing of mAb monoclonal antibody derived therapeutics as they impart elements of automation, consistency and reliability to standard toxicological assays. There are a multitude of advantages offered by computational toxicology methods. They help to realise the 3R principle i.e. Replacement, Reduction and Refinement, by reducing the number of experimental animals used in drug safety testing. However, as with most predictive model strategies, the availability of data pertaining to mAbs is relatively limited by the cost of testing and the restrictions on access to information owing to the highly competitive nature of the biopharmaceutical industry.

Criteria for characterising the predictive ability of a model are based on the objective of the model. When the main objective of the model is to correctly identify the trends and rank molecules based on their activity/property in question, the metrics that capture the correlation between observed and predicted values of the test set i.e. prediction  $R^2$ , maybe more reflective and relevant when compared to RMSE, accuracy and error of the models(Alexander *et al.*, 2015). With respect to mAbs, models that allow for early stage screening based on safety and efficacy may lead to effective reduction of the number of lead candidates that progress through process development into manufacturing and thus reduce costs associated to process development failures and attrition. To that end these models based on the tendency of mAbs to cause hypersensitivity show relatively good prediction  $R^2$  values and thus will be useful for identifying trends in mAb therapeutics developability.

Due to the limited sample size caution must be exercised while interpreting the predictive ability of the models. Even though the models show the general trend of responses, more samples have to be included to increase the robustness and the predictive ability of all the models as overfitting becomes an important problem. Model performance can also be increased by bootstrapping and iterative resampling methods, however it would still present the dangers of an overly optimistic model due to a limited sample size as well as the influence of the training set on predictions (Kohavi, 1995; Consonni *et al.*, 2009). Model optimisation can also be carried out by reducing the number of variables used in a model. To this end unsupervised variable reduction method, V-WSP, was used to address the challenge of the number variables being greater than the number of samples. Subsequently, GA-PLS based supervised variable selection was used to further optimise the model.

The rationale for the Skimune™ ® grading is based on its correlation with the clinical occurrence of adverse events during clinical trials. The different Skimune™ ® classes reflect the likelihood of an adverse immune reaction to occur during a stage IV clinical trial. Furthermore, donor variability is also accounted for, especially for the “weak positive” class of Skimune™ ® grading score. This refers to cases where the testing compound, while known to be a sensitizer, fails to induce an adverse immune reaction in all tested donors. The same logic can be applied in the opposite situation, where a negative sensitizer can induce an adverse immune response in a small percentage of donors. Thus, as an in vitro assay system for safety testing of mAb based therapeutics, the Skimune™ assay captures the variability associated with clinical testing in terms of patient variability. This has also been reflected in the high correlation of Skimune™ responses with clinical immunogenicity incidence rates (Figure 9.4). However, this variability also increases the intrinsic variance when used as a biological response in QSAR studies.

As shown in Figure 9.3 the electronic and topological properties of VH and VL substructures, both framework and CDR, have shown to influence the model. As the framework regions contain sequences pertaining to species information this could be one of the potential reasons for the influence of these regions on Skimune™ response prediction since immunogenic responses can be elicited by presence of non-human sequences in mAbs (Harding *et al.*, 2010). For the regression models, inclusion of constant region descriptors as well as glycoform information together with increased number of samples would allow for better prediction of the response range (Umaña *et al.*, 1999). Primary sequence-based descriptors do not consider either the interactions between amino acid residues or the antibody-antigen and antibody-receptor interaction space. This could be addressed by using the 3D structures of mAbs generated via homology modelling and molecular dynamics simulations so as to extract useful 3D structural descriptors provide energy minimised 3D structures which can be used for in silico receptor binding studies resulting in effector function related descriptors (using docking) (Pettersen *et al.*, 2004; Phillips *et al.*, 2005).

#### 9.4. Chapter summary

To summarise this chapter looked at the applicability of the modelling framework developed with Skimune™ as response. There is plenty of scope for application of this methodology for other studies involving different biological activities as output. Extension could possibly include reducing number of lead candidates and providing more information to in vitro studies

when requisite in vivo species is not available. Supplementing existing safety strategies by providing a hybrid approach to safety and efficacy testing of biotherapeutics would be particularly useful while exploring the combined mAb chemical space for combination therapies. In conclusion this study could be translated to developing QSAR models for mAbs using larger developability or adverse effects datasets to allow for early stage screening, thus reducing manufacturing failures and attrition rates.

## Chapter 10: Conclusions and recommendations

Monoclonal antibodies (mAbs) and related therapeutics are highly desirable from a biopharmaceutical perspective as they are highly target specific and well tolerated within the human system. Advancements in early stage screening of mAb candidates would reduce the number of lead candidates entering the bioprocess pipeline that are associated with adverse effects thereby reducing attrition rates. As defined in the objectives, the toxicity of parabens was assessed via traditional toxicity assays wherein the hepatotoxicity and dermal toxicity of butyl paraben was observed. These traditional toxicity assays were then used to assess potential off target adverse effects of mAbs. No apparent ADCC, CDCC or CDC mediated decrease in cell viability was measured for either of the two cells lines tested except for a CDC mediated decrease in cell viability in HDFn cells. Several considerations have been elucidated for development of *in vitro* assays better suited to detect off target toxicity of mAbs. Furthermore, hypersensitivity reactions of mAbs and their aggregates were assessed using the novel Skimune™ assay indicating the utility of these assays for detection of immune related adverse effects of mAbs. The development of a hybrid QSAR based model with a structured workflow and clear evaluation metrics, with several optimisation steps, was also described that could be beneficial for broader and more generic PLS modelling. Based on the results and observation from this study, it was demonstrated incremental improvement via selection of datasets and variables help in further optimisation of these hybrid models. Furthermore, using hypersensitivity and cross reactivity as responses and physicochemical characteristics of mAbs as descriptors, the QSAR models generated for different applicability domains allow for rapid early stage screening and developability.

Whether it is for assessing preclinical safety or for rapid screening, *in vivo* systems are not the most suitable models for studying the effects of monoclonal antibody-based therapeutics. The rationale behind using *in vivo* studies in preclinical safety testing is that the indirect immune mediated response induced by the antibody as well as the magnitude of the effect cannot be gauged via standard *in vitro* tests. However, species specificity remains the main obstacle (Brennan *et al.*, 2018). Studying the effector function becomes difficult due to differences in the Fc $\gamma$ R receptors structure and affinity, complement system response and absence of target antigen (Golay and Introna, 2012). Presence, number, interactions as well as distribution of target antigen also plays an important role in assessing the biological activity of monoclonal antibodies (Golay *et al.*, 2001). Attempts have been made to solve this problem by different strategies, such as knockout mice that lack mouse Fc $\gamma$ R, transgenic mice expressing human

FcγR, generating xenografts with human antigen in mouse cell lines, using completely mouse systems and using primate models such as rhesus monkey (Golay and Introna, 2012; Barouch *et al.*, 2013; Strasser *et al.*, 2013; Bournazos *et al.*, 2014). Animal testing is also expensive, sample size dependent and resource intensive.

*In vitro* systems have been routinely used for assessing the adverse effect of chemical compounds. This has been illustrated during the assessment of hepatotoxicity and dermal toxicity of butyl paraben and methyl paraben using HepG2 and HDFn *in vitro* models as shown in **Chapter 2**. However, applying traditional toxicity tests for assessing adverse effects of mAbs was not as straightforward as evident from the results in **Chapter 3**. The innate complexity, diversity and size of mAbs based therapeutic as well as their diverse mechanisms of actions that involve many pathways exacerbate the need for carefully designed *in vitro* systems that consider all of the above factors (Brennan and Kiessling, 2017). In standard cytokine release assays, the mAbs bind to receptors all over the cell which is not an accurate representation of the human systems where cytokine release is sometimes dependent on localised receptor interaction (Stebbins *et al.*, 2007). Sophisticated analytical techniques used in studying the endpoints of these assays have to be carefully assessed for resolution as well as sensitivity in detecting events as they can be prone to artefacts owing to nature of assay in question as well as the size of biological molecules. Artefacts can arise while using flow cytometry techniques due to homotypic adhesion as demonstrated with anti CD20 antibodies monoclonal antibodies (Golay *et al.*, 2010). New generation preclinical safety testing tools would have to be high throughput, rapid and cost effective to meet the accelerated growth of the biopharmaceutical market. They also need to be highly reproducible and be fairly predictive to allow for rapid screening facilitating reliable selection of new compounds at initial stages thus saving time and money to allow more focus on drug development for rare diseases. They would also provide an alternative to animal testing considering the various drawbacks of *in vivo* systems as seen in the case of TGN1412. Hypersensitivity reactions have been assessed using Skimune™, a non-artificial human skin explants based assay for safety and efficacy assessment of novel compounds and drugs, developed by Alcyomics Ltd (Ahmed *et al.*, 2016). The applicability of the Skimune™ assay have been demonstrated in **Chapter 4** for assessing the immunogenicity of mAb aggregates.

Appropriate and relevant experimental studies are of paramount importance in nonclinical safety testing as they also contribute to good datasets which can then be modelled. Most of the models are developed based on public data sets and fail to perform adequately when tested with

proprietary datasets. The highly competitive nature of the biopharmaceutical industry makes information access very difficult. There are also difficulties in feature extraction for biological molecules owing to their complexity and size. The applicability of such modelling techniques in rapid screening depends on the experimental set up as well as on identifying and forming sensible profilers and descriptors.

Identification and generation of descriptors from the primary sequence of mAbs as well as the influence intrinsic properties exert on the descriptors was shown in **Chapter 5**. These descriptors were then used for developing a QSAR model for prediction of cross interaction chromatographic retention times, the results of which was elucidated in **Chapter 6**. Primary sequence-based descriptors do not take into account interactions between amino acid residues or the antibody-antigen and antibody-receptor interaction space. To address this, 3D structures of mAbs were first generated via homology modelling and molecular dynamics simulation upon which structural descriptors were generated for QSAR model development (**Chapter 7**). A combination of sequence based, and structural descriptors were also utilised for model development (**Chapter 8**) followed by testing their applicability to predict Skimune™ responses elicited by mAbs (**Chapter 9**).

Advances made in PCM techniques include a new descriptor for antigen-antibody interaction called Epitope-Paratope Interaction Fingerprint (EPIF) which tries to address the higher time-complexity of MLPD, thus allowing for simplification the antigen-antibody interaction term (Qiu *et al.*, 2015; Qiu *et al.*, 2016). Platforms like proABC, ABangle and LYRA allow for modelling antigen-antibody interactions, orientation of variable chain and lymphocyte receptor respectively (Dunbar *et al.*, 2013; Olimpieri *et al.*, 2013; Klausen *et al.*, 2015). Physicochemical characteristics of mAbs will influence PK/PD properties (increased binding to serum proteins and increased half-life) which affects ADME characteristics thus impacting bioavailability and biological activity (Dostalek *et al.*, 2017). Glycosylation is another aspect that has to be taken into consideration as change in glycosylation pattern could affect functionality as well as impact PK/PD characteristics of mAbs (Liu, 2015). Successful attempts have been made from a bioengineering point of view to investigate the effects of the production process on glycosylation profiles of monoclonal antibodies by using multivariate techniques, such as principal component analysis, partial least squares and parallel factor analysis (Green and Glassey, 2015), (Glassey, 2012). Glycoengineered antibodies were produced by CHO cells with higher glycosyltransferase which enabled the production of engineered antibodies with the N acetylglucosamine profiles required to achieve higher neutrophil mediated phagocytosis

activity and thus greater efficacy in killing tumour cells (Umaña *et al.*, 1999; Golay *et al.*, 2013). Indeed, engineered glycoforms of anti CD20 antibodies, such as obinutuzumab and Rituximab, have sevenfold higher binding affinity to neutrophils and thus an increased neutrophil mediated phagocytosis based killing of tumour cells (Golay *et al.*, 2013). The challenge would then be to associate these attributes to potential adverse effects which will then allow for development of predictive toxicology models. Intricate algorithms would also be required for associating profilers and descriptors with synergistic endpoints of toxicity. Along with carefully designed experimental procedures, extensive expert knowledge would be required for such model development.

Future work pertaining to the studies done in this thesis would be to use nonlinear modelling techniques, applying the modelling framework on larger and/or more industrial datasets as well as newer types of mAbs such as bispecifics, developing QSAR models around other substructural and species related applicability domains of MAbs, including glycoform conformation as a feature for model development to name a few. Modelling techniques such as those that use categorical responses (Partial Least Squares Discriminant Analysis, PLS-DA) as well as nonlinear modelling methods such as Support Vector Machines (SVM) were also briefly investigated in this project (data not shown). The performance of PLS-DA models was poor mainly due to class imbalance. Furthermore, the influence of expert knowledge is greater for these models as defining the classes is arbitrary. With regard to the SVM based models, the increase in model complexity only led to similar model performance as that of the PLS models. Interpretability is lower for complex models such as those generated by SVM for similar model performance. Thus, further investigation is required to ascertain the applicability of complex modelling techniques such as Artificial neural networks, Random forests, Bayesian models etc for early stage screening of mAb therapeutics. Model predictability and utility can be further improved with the inclusion of better mAb features such as glycoform conformation and distribution which are linked to mAb efficacy and safety. Increasing the sample sizes within each substructure and specifies applicability domain could expand the applicability of this QSAR modelling framework. Furthermore, this methodology could then extend to different therapeutic types such as fusion proteins, bispecifics, single chain fragment variables and other novel mAb based therapeutics.

Biopharmaceuticals have positively impacted the lives of millions. They have paved the way for personalised medicines, improve prognosis of cancer, genetic and immune disorders as well as breakthroughs in rare disease management (Dostalek *et al.*, 2017; Shepard *et al.*, 2017;

Kennedy *et al.*, 2018). The advances made are however impeded by a lack of progress in bioprocess development strategies as well as increasing costs owing to attrition, wherein the lack of efficacy and safety accounts for nearly 60% of all factors contributing to attrition (Kola and Landis, 2004). This reiterates the need for carefully designed predictive models to assess the efficacy as well as toxicity of potential drug candidates at an early stage. A more effective, high-throughput rapid screening of candidates based on adverse effects is required at an early stage to filter out the number of candidates proceeding to clinical trials. A choice of appropriate *in vivo* systems should be in place along with better proof of concept studies as animal models are not representative of human systems for assessing the efficacy and safety of biopharmaceuticals in specialised therapy areas like oncology and immunology. Alternative approaches such as specialised *in vitro* toxicology tests, better biomarkers and *omics* approaches can be utilised for this purpose. In this regard, computational toxicology tools like expert/hybrid systems provide a powerful complement to *in vitro* systems as they will allow for development of automated and reliable models for predicting toxicity or adverse effect of monoclonal antibody therapeutics. To make these predictive platforms more robust, descriptor calculation, feature extraction, inclusion of pharmacokinetics and bioavailability characteristics, mechanistic understanding and multidisciplinary expert knowledge will be of paramount importance. This could aid in reducing the number of lead candidates that could go forward into the bioprocessing /manufacturing pipeline. Thus, this would tackle two of the main setbacks biopharmaceutical industries face today: manufacturing failure and attrition. This will pave way for the development of rapid bioprocess development strategies for faster development of effective and safe biopharmaceuticals and may in fact change the face of biopharmaceutical manufacturing as we see today.

## Appendix A

**Table A.1** Details of the final mAb concentration and associated dilutions.

Working concentrations	Concentration in well (µg/ml)	Volume of mAb (µl)	Volume of DMEM (µl)	Source	Final Volume in well (µl)	Volume added/well (µl)
40	10	48	1152	1mg/ml	200	50
4	1	120	1080	40	200	50
10.4	0.1	120	1080	4	200	50

Working concentration= Concentration in well x dilution factor (4 as total volume in cell is 200µl) Rituximab stock is 100mg/ml; Working stock of 10mg/ml : 100µl in 900µl of PBS): Stock solution of 1mg/ml. Trastuzumab stock is 600mg/ml; working stock of 10mg/ml; 16.5 µl in 984 µl of PBS): Stock Solution of 1mg/ml

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	NC	PC	IC	mAb1 10µg/ml	mAb1 1µg/ml	mAb1 0.1µg/ml	mAb2 10µg/ml	mAb2 1µg/ml	mAb2 0.1µg/ml	M		
C	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	M		
D	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	M		
E	NC	PC	IC	mAb1 10µg/ml	mAb1 1µg/ml	mAb1 0.1µg/ml	mAb2 10µg/ml	mAb2 1µg/ml	mAb2 0.1µg/ml	EC	PBMCs	
F	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	EC		
G	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	EC		
H	AICC			ADCC	ADCC	ADCC	ADCC	ADCC	ADCC	ADCC		

(a)

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	NC	PC	IC	mAb1 10µg/ml	mAb1 1µg/ml	mAb1 0.1µg/ml	mAb2 10µg/ml	mAb2 1µg/ml	mAb2 0.1µg/ml	M		
C	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	M		
D	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	M		
E	NC	PC	IC	mAb1 10µg/ml	mAb1 1µg/ml	mAb1 0.1µg/ml	mAb2 10µg/ml	mAb2 1µg/ml	mAb2 0.1µg/ml	EC	PBMCs + Serum	
F	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	EC		
G	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	EC		
H	AICC			CDCC	CDCC	CDCC	CDCC	CDCC	CDCC	CDCC		

(b)

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	NC	PC	IC	mAb1 10µg/ml	mAb1 1µg/ml	mAb1 0.1µg/ml	mAb2 10µg/ml	mAb2 1µg/ml	mAb2 0.1µg/ml	M	Serum	
C	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	M		
D	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	M		
E	NC	PC	IC	mAb1 10µg/ml	mAb1 1µg/ml	mAb1 0.1µg/ml	mAb2 10µg/ml	mAb2 1µg/ml	mAb2 0.1µg/ml	EC		
F	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	EC		
G	NC	PC	IC	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	mAb 10µg/ml	mAb 1µg/ml	mAb 0.1µg/ml	EC		
H	AICC			CDC	CDC	CDC	CDC	CDC	CDC			

(c)

**Figure A.1** Layout of (a) ADCC, (b) CDCC and (c) CDC experiment. ADCC: Antibody Dependent Cellular Cytotoxicity, CDCC: Complement Dependent Cellular Cytotoxicity, CDC: Complement Dependent Cytotoxicity, AICC: Antibody independent cytotoxicity, NC: Negative/target Control, PC: Positive Control, IC: Isotype Control, EC: Effector Control. The coloured cells indicate wells that contain PBMCs, PBMCs+Serum and Serum respectively.

## Appendix B

**Table B.1** Full factorial Design of Experiments for mAb aggregation

StdOrder	RunOrder	Blocks	Temperature (°C)	TIME (Hours)
1	1	1	4	0
2	2	1	4	3
3	3	1	4	6
4	4	1	4	12
5	5	1	4	24
6	6	1	4	48
7	7	1	37	0
8	8	1	37	3
9	9	1	37	6
10	10	1	37	12
11	11	1	37	24
12	12	1	37	48
13	13	1	40	0
14	14	1	40	3
15	15	1	40	6
16	16	1	40	12
17	17	1	40	24
18	18	1	40	48
19	19	1	4	0
20	20	1	4	3
21	21	1	4	6
22	22	1	4	12
23	23	1	4	24
24	24	1	4	48
25	25	1	37	0
26	26	1	37	3
27	27	1	37	6
28	28	1	37	12
29	29	1	37	24
30	30	1	37	48
31	31	1	40	0
32	32	1	40	3
33	33	1	40	6
34	34	1	40	12
35	35	1	40	24
36	36	1	40	48
37	37	1	4	0
38	38	1	4	3
39	39	1	4	6
40	40	1	4	12
41	41	1	4	24
42	42	1	4	48
43	43	1	37	0

44	44	1	37	3
45	45	1	37	6
46	46	1	37	12
47	47	1	37	24
48	48	1	37	48
49	49	1	40	0
50	50	1	40	3
51	51	1	40	6
52	52	1	40	12
53	53	1	40	24
54	54	1	40	48

**Table B.2** Heavy chain amino acid composition of rituximab and Trastuzumab

Residue	Number		Mole%		DayhoffStat	
	Rituxima b	Trastuzuma b	Rituxima b	Trastuzuma b	Rituxima b	Trastuzumab
A = Ala	24	23	5.322	5.111	0.619	0.594
B = Asx	0	0	0	0	0	0
C = Cys	11	11	2.439	2.444	0.841	0.843
D = Asp	17	18	3.769	4	0.685	0.727
E = Glu	20	22	4.435	4.889	0.739	0.815
F = Phe	13	13	2.882	2.889	0.801	0.802
G = Gly	31	33	6.874	7.333	0.818	0.873
H = His	10	10	2.217	2.222	1.109	1.111
I = Ile	7	9	1.552	2	0.345	0.444
K = Lys	36	32	7.982	7.111	1.209	1.077
L = Leu	32	33	7.095	7.333	0.959	0.991
M = Met	5	5	1.109	1.111	0.652	0.654
N = Asn	19	19	4.213	4.222	0.98	0.982
P = Pro	35	34	7.761	7.556	1.492	1.453
Q = Gln	18	16	3.991	3.556	1.023	0.912
R = Arg	8	13	1.774	2.889	0.362	0.59
S = Ser	53	50	11.752	11.111	1.679	1.587
T = Thr	38	35	8.426	7.778	1.381	1.275
V = Val	43	44	9.534	9.778	1.445	1.481

W = Trp	9	9	1.996	2	1.535	1.538
X = Xaa	0	0	0	0	0	0
Y = Tyr	22	21	4.878	4.667	1.435	1.373

**Table B.3** Properties of amino acid composition of rituximab and Trastuzumab (heavy chain).

Property	Residues	Number		Mole%	
		Rituximab	Trastuzumab	Rituximab	Trastuzumab
		ab	ab	ab	ab
Tiny	(A+C+G+S+T)	157	152	34.812	33.778
Small	(A+B+C+D+G+N+P+S+T+V)	271	267	60.089	59.333
Aliphatic	(A+I+L+V)	106	109	23.503	24.222
Aromatic	(F+H+W+Y)	54	53	11.973	11.778
Non-polar	(A+C+F+G+I+L+M+P+V+W+Y)	232	235	51.441	52.222
Polar	(D+E+H+K+N+Q+R+S+T+Z)	219	215	48.559	47.778
Charged	(B+D+E+H+K+R+Z)	91	95	20.177	21.111
Basic	(H+K+R)	54	55	11.973	12.222
Acidic	(B+D+E+Z)	37	40	8.204	8.889

**Table B.4** Light chain amino acid composition of rituximab and trastuzumab

Residue	Number		Mole%		DayhoffStat	
	Rituximab	Trastuzumab	Rituximab	Trastuzumab	Rituximab	Trastuzumab
A = Ala	15	14	7.042	6.542	0.819	0.761
C = Cys	5	5	2.347	2.336	0.809	0.806
D = Asp	6	10	2.817	4.673	0.512	0.85
E = Glu	11	9	5.164	4.206	0.861	0.701
F = Phe	7	9	3.286	4.206	0.913	1.168
G = Gly	13	11	6.103	5.14	0.727	0.612
H = His	3	3	1.408	1.402	0.704	0.701
I = Ile	7	6	3.286	2.804	0.73	0.623

K = Lys	13	13	6.103	6.075	0.925	0.92
L = Leu	13	14	6.103	6.542	0.825	0.884
M = Met	1	1	0.469	0.467	0.276	0.275
N = Asn	7	6	3.286	2.804	0.764	0.652
P = Pro	13	12	6.103	5.607	1.174	1.078
Q = Gln	12	15	5.634	7.009	1.445	1.797
R = Arg	6	7	2.817	3.271	0.575	0.668
S = Ser	35	31	16.432	14.486	2.347	2.069
T = Thr	17	20	7.981	9.346	1.308	1.532
V = Val	16	16	7.512	7.477	1.138	1.133
W = Trp	4	2	1.878	0.935	1.445	0.719
Y = Tyr	9	10	4.225	4.673	1.243	1.374

**Table B.5** Properties of amino acid composition of rituximab and Trastuzumab (heavy chain).

Property	Residues	Number		Mole%	
		Rituximab	Trastuzumab	Rituximab	Trastuzumab
Tiny	(A+C+G+S+T)	85	81	39.906	37.85
Small	(A+B+C+D+G+N+P+S+T+V)	127	125	59.624	58.411
Aliphatic	(A+I+L+V)	51	50	23.944	23.364
Aromatic	(F+H+W+Y)	23	24	10.798	11.215
Non-polar	(A+C+F+G+I+L+M+P+V+W+Y)	103	100	48.357	46.729
Polar	(D+E+H+K+N+Q+R+S+T+Z)	110	114	51.643	53.271
Charged	(B+D+E+H+K+R+Z)	39	42	18.31	19.626
Basic	(H+K+R)	22	23	10.329	10.748
Acidic	(B+D+E+Z)	17	19	7.981	8.879

## Appendix C

**Table C.1** Information of the mAbs used in the study and the primary sequence sources.

Antibody	Heavy chain	Light Chain	Species	Target
TGN1412	IgG4	Kappa	Humanized	CD28
abrilumab	IgG2	Kappa	Homo sapiens	ITGA4_ITGB7 D49d
actoxumab	IgG1	Kappa	Homo sapiens	Toxin A
adalimumab	IgG1	Kappa	Homo sapiens	TNF
aducanumab	IgG1	Kappa	Homo sapiens	APP
afasevikumab	IgG1	Kappa	Homo sapiens	IL17A IL17F
alemtuzumab	IgG1	Kappa	Humanized	CD52
alirocumab	IgG1	Kappa	Homo sapiens	PCSK9
amatuximab	IgG1	Kappa	Chimeric	MSLN
andecaliximab	IgG4	Kappa	Chimeric	MMP9
anifrolumab	IgG1	Kappa	Homo sapiens	IFNAR1
anrukinzumab	IgG1	Kappa	Humanized	IL13
aprutumab	IgG1	Lambda	Homo sapiens	FGFR2
ascrinvacumab	IgG2	Kappa	Homo sapiens	ACVRL1
atezolizumab	IgG1	Kappa	Humanized	CD274
atinumab	IgG4	Kappa	Homo sapiens	RTN4
avelumab	IgG1	Lambda	Homo sapiens	CD274
azintuxizumab vedotin	IgG1	Kappa	Chimeric Humanized	SLAMF7
azintuxizumab	IgG1	Kappa	Chimeric Humanized	SLAMF7
bapineuzumab	IgG1	Kappa	Humanized	APP
basiliximab	IgG1	Kappa	Chimeric	IL2RA
bavituximab	IgG1	Kappa	Chimeric	phosphatidylserine
benralizumab	IgG1	Kappa	Humanized	IL5RA CD125
bevacizumab beta	IgG1	Kappa	Humanized	VEGFA
bevacizumab	IgG1	Kappa	Humanized	VEGFA
bezlotoxumab	IgG1	Kappa	Homo sapiens	Toxin B
bimagrumab	IgG1	Lambda	Homo sapiens	ACVR12B/A
bimekizumab	IgG1	Kappa	Humanized	IL17A IL17F
bleselumab	IgG4	Kappa	Homo sapiens	CD40
blosozumab	IgG4	Kappa	Humanized	SOST
bococizumab	IgG2	Kappa	Humanized	PCSK9
brazikumab	IgG2	Lambda	Homo sapiens	IL23A
brentuximab vedotin	IgG1	Kappa	Chimeric	TNFRSF8
briakinumab	IgG1	Lambda	Homo sapiens	IL12B
brodalumab	IgG2	Kappa	Homo sapiens	IL17RA
brontictuzumab	IgG2	Lambda	Humanized	NOTCH1
burosumab	IgG1	Kappa	Homo sapiens	FGF23
cabiralizumab	IgG4	Kappa	Humanized	CSF1R
camrelizumab	IgG4	Kappa	Humanized	PDCD1
canakinumab	IgG1	Kappa	Homo sapiens	IL1B

cantuzumab ravtansine	IgG1	Kappa	Humanized	MUC1
carlumab	IgG1	Kappa	Homo sapiens	CC12
carotuximab	IgG1	Kappa	Chimeric	ENG CD105
cergutuzumab amunaleukin	IgG1	Kappa	Humanized	CEACAM5 CD66e
cetuximab	IgG1	Kappa	Chimeric	EGFR
cixutumumab	IgG1	Lambda	lambda	IGF1R
clazakizumab	IgG1	Kappa	Humanized	IL6
clivatuzumab tetraxetan	IgG1	Kappa	Humanized	MUC1
codrituzumab	IgG1	Kappa	Humanized	GPC3
coltuximab ravtansine	IgG1	Kappa	Chimeric	CD19
conatumumab	IgG1	Kappa	Homo sapiens	TNFRSF10B
concizumab	IgG4	Kappa	Humanized	TFPI
cosfroviximab	IgG1	Kappa	Chimeric	EBOV
crenezumab	IgG4	Kappa	Humanized	APP
crizanlizumab	IgG2	Kappa	Humanized	SELP
crotedumab	IgG4	Kappa	Homo sapiens	GCGR
dacetuzumab	IgG1	Kappa	Humanized	CD40
daclizumab	IgG1	Kappa	Humanized	IL2RA
dalotuzumab	IgG1	Kappa	Humanized	IGF1R
daratumumab	IgG1	Kappa	Homo sapiens	CD38
dectrekumab	IgG1	Kappa	Homo sapiens	IL13
demcizumab	IgG2	Kappa	Humanized	DLL4
denintuzumab mafodotin	IgG1	Kappa	Homo sapiens	TNSF11 CD254
denosumab	IgG2	Kappa	Humanized	CD19
depatuxizumab mafodotin	IgG1	Kappa	Chimeric Humanized	EGFR
depatuxizumab	IgG1	Kappa	Chimeric Humanized	EGFR
dezamizumab	IgG1	Kappa	Humanized	APCS
dinutiximab beta	IgG1	Kappa	Chimeric	GD2
dinutuximab	IgG1	Kappa	Chimeric	GD2
diridavumab	IgG1	Lambda	Homo sapiens	Influenza A virus
domagrozumab	IgG1	Kappa	Humanized	MSTN
drozitumab	IgG1	Lambda	Homo sapiens	TNFRSF10B
duligotuzumab	IgG1	Kappa	Humanized	ERBB3
dupilumab	IgG4	Kappa	Homo sapiens	IL4R
durvalumab	IgG1	Kappa	Homo sapiens	CD274
dusigitumab	IgG2	Lambda	Homo sapiens	IGF2
efalizumab	IgG1	Kappa	Humanized	ITGAL
eldelumab	IgG1	Kappa	Homo sapiens	CXCL10
elezanumab	IgG1	Lambda	Homo sapiens	RGMA
elgemtumab	IgG1	Kappa	Homo sapiens	ERBB3
elotuzumab	IgG1	Kappa	Humanized	SLAMF7
emactuzumab	IgG1	Kappa	Humanized	CSF1R
emapalumab	IgG1	Lambda	Homo sapiens	IFNG
emibetuzumab	IgG4	Kappa	Humanized	MET
emicizumab	IgG4	Kappa	Humanized	F9
enavatuzumab	IgG1	Kappa	Humanized	TNFRSF12A

enfortumab vedotin	IgG1	Kappa	Homo sapiens	PVRL4
enoblituzumab	IgG1	Kappa	Humanized	CD276
enokizumab	IgG1	Kappa	Humanized	IL9
enoticumab	IgG1	Kappa	Homo sapiens	DLL4
ensituximab	IgG1	Kappa	Chimeric	MUC5AC
eptinezumab	IgG1	Kappa	Humanized	CALCA/B
erenumab	IgG2	Lambda	Homo sapiens	CRCP
etaracizumab	IgG1	Kappa	Humanized	ITGAV ITGB3
etrolizumab	IgG1	Kappa	Humanized	ITGA4 ITGB7
evinacumab	IgG4	Kappa	Homo sapiens	ANGPTL3
evolocumab	IgG2	Lambda	Homo sapiens	PCSK9
farletuzumab	IgG1	Kappa	Humanized	FOLR1
fasinumab	IgG4	Kappa	Homo sapiens	NGF
fezakinumab,	IgG1	Lambda	Homo sapiens	IL22
ficlatuzumab	IgG1	Kappa	Humanized	HGF
figitumumab	IgG2	Kappa	Homo sapiens	IGF1R
firivumab	IgG1	Kappa	Homo sapiens	Influenza A virus
flanvotumab	IgG1	Kappa	Homo sapiens	TYRP1
fletikumab	IgG4	Kappa	Homo sapiens	IL20
foralumab	IgG1	Kappa	Homo sapiens	CD3E
foravirumab	IgG1	Kappa	Homo sapiens	RV
fremanezumab	IgG2	Kappa	Humanized	CALCA/B
fresolimumab	IgG4	Kappa	Homo sapiens	TGFB
fulranumab	IgG2	Kappa	Homo sapiens	NGF
futuximab	IgG1	Kappa	Chimeric	EGFR
galcanezumab	IgG4	Kappa	Humanized	CALCA/B
ganitumab	IgG1	Kappa	Homo sapiens	IGF1R
gantenerumab	IgG1	Kappa	Homo sapiens	APP
gatipotuzumab	IgG1	Kappa	Humanized	MUC1
gedivumab	IgG1	Kappa	Homo sapiens	Influenza A virus
gemtuzumab ozogamicin	IgG4	Kappa	Humanized	CD33
gevokizumab	IgG2	Kappa	Humanized	1L1B
girentuximab	IgG1	Kappa	Chimeric	CA9
glebatumumab vedotin	IgG2	Kappa	Homo sapiens	GPNMB
glebatumumab	IgG2	Kappa	Homo sapiens	GPNMB
guselkumab	IgG1	Lambda	Homo sapiens	IL23A
ibalizumab	IgG4	Kappa	Humanized	CD4
icrucumab	IgG1	Kappa	Homo sapiens	FLT1
ifabotuzumab	IgG1	Kappa	Humanized	EPHA3
imalumab	IgG1	Kappa	Homo sapiens	MIF
imgatuzumab	IgG1	Kappa	Humanized	EGFR
inclacumab	IgG4	Kappa	Homo sapiens	SELP
indatuximab ravtansine	IgG4	Kappa	Chimeric	SDC1
indusatumab vedotin	IgG1	Kappa	Homo sapiens	GUCY2C
indusatumab	IgG1	Kappa	Homo sapiens	GUCY2C
inebilizumab	IgG1	Kappa	Humanized	CD19

infliximab	IgG1	Kappa	Chimeric	TNF
intetumumab	IgG1	Kappa	Homo sapiens	ITGAV ITGB3
ipilimumab	IgG1	Kappa	Homo sapiens	CTLA4
iratumumab	IgG1	Kappa	Homo sapiens	TNFR TNFRSF8 CD30
isatuximab	IgG1	Kappa	Chimeric	CD38
itolizumab	IgG1	Kappa	Humanized	CD6
ixekizumab	IgG4	Kappa	Humanized	IL17A
labetuzumab govitecan	IgG1	Kappa	Humanized	CEACAM5 CD66e
lacnotuzumab	IgG1	Kappa	Humanized	CSF1
lanadelumab	IgG1	Kappa	Homo sapiens	KLKB1
landogrozumab	IgG4	Kappa	Humanized	MSTN
laprituximab emtansine	IgG1	Kappa	Chimeric	EGFR
laprituximab	IgG1	Kappa	Chimeric	EGFR
larcaviximab	IgG1	Kappa	Chimeric	EBOV Zaire Ebola Virus
lebrikizumab	IgG4	Kappa	Humanized	IL13
lenzilumab	IgG1	Kappa	Homo sapiens	CSF2
lesofavumab	IgG1	Kappa	Homo sapiens	influenze B virus hemagglutinin
lexatumumab	IgG1	Lambda	Homo sapiens	TNFR TNFRSF10B CD262
lifastuzumab vedotin	IgG1	Kappa	Humanized	SLC34A2
ligelizumab	IgG1	Kappa	Humanized	IGHE
lirilumab	IgG4	Kappa	Homo sapiens	KIRD2 subgroup
lodelcizumab	IgG1	Kappa	Humanized	PCSK9
lorvotuzumab mertansine	IgG1	Kappa	Humanized	NCAM1
losatuxizumab vedotin	IgG1	Kappa	Chimeric Humanized	EGFR
losatuxizumab	IgG1	Kappa	Chimeric Humanized	EGFR
lucatumumab	IgG1	Kappa	Homo sapiens	CD40
lumretuzumab	IgG1	Kappa	Humanized	ERBB3
lupartumab amadotin	IgG1	Lambda	Homo sapiens	LYPD3
lupartumab	IgG1	Lambda	Homo sapiens	LYPD3
margetuximab	IgG1	Kappa	Chimeric	ERBB2 CD340
mavrilimumab	IgG4	Lambda	Homo sapiens	CSF2RA
milatuzumab doxorubicin	IgG1	Kappa	Humanized	CD74
mirvetuximab soravtansine	IgG1	Kappa	Chimeric	FOLR1
mirvetuximab	IgG1	Kappa	Chimeric	FOLR1
modotuximab	IgG1	Kappa	Chimeric	EGFR
mogamulizumab	IgG1	Kappa	Humanized	CCR4
monalizumab	IgG4	Kappa	Humanized	KLRC1
motavizumab	IgG1	Kappa	Humanized	RSV glycoprotein F
namilumab	IgG1	Kappa	Homo sapiens	CSF2
naratuximab	IgG1	Kappa	Chimeric	CD37
narnatumab	IgG1	Kappa	Homo sapiens	MST1R
natalizumab	IgG4	Kappa	Humanized	ITGA4
navicixizumab	IgG2	Kappa	Chimeric Humanized	VEGFA

navivumab	IgG1	Kappa	Homo sapiens	influenze A virus hemagglutinin HA
necitumumab	IgG1	Kappa	Homo sapiens	EGFR
nemolizumab	IgG2	Kappa	Humanized	IL31RA
nesvacumab	IgG1	Kappa	Homo sapiens	ANGPT2
nimotuzumab	IgG1	Kappa	Humanized	EGFR
nivolumab	IgG4	Kappa	Homo sapiens	PDCD1
obiltoxaximab	IgG1	Kappa	Chimeric	anthrax protective antigen
obinutuzumab	IgG1	Kappa	Humanized	MS4A1 CD20
ocaratumab	IgG1	Kappa	Humanized	MS4A1 CD20
olaratumab	IgG1	Kappa	Homo sapiens	PDGFRA
oleclumab	IgG1	Lambda	Homo sapiens	NT5E CD73
olendalizumab	IgG2	Kappa	Humanized	C5
olokizumab	IgG4	Kappa	Humanized	IL6
omalizumab	IgG1	Kappa	Humanized	IGHE
onartuzumab	IgG1	Kappa	Humanized	MET
ontuxizumab	IgG1	Kappa	Chimeric humanized	CD248
opicinumab	IgG1	Kappa	Homo sapiens	LINGO1
orticumab	IgG1	Lambda	Homo sapiens	oxLDL
otelixizumab	IgG1	Lambda	ChimericHumanizedb	CD3E
oxelumab	IgG1	Kappa	Homo sapiens	TNFSF4 CD252
ozanezumab	IgG1	Kappa	Humanized	RTN4
pamrevlumab	IgG1	Kappa	Homo sapiens	CTGF
parsatuzumab	IgG1	Kappa	Humanized	EGFL7
pateclizumab	IgG1	Kappa	Humanized	LTA TNFSF1
patritumab	IgG1	Kappa	Homo sapiens	ERBB3
pembrolizumab	IgG4	Kappa	Humanized	PDCD1
perakizumab	IgG1	Kappa	Humanized	IL17A
pidilizumab	IgG1	Kappa	Humanized	PDCD1 CD279
pinatuzumab vedotin	IgG1	Kappa	Humanized	CD22
plozalizumab	IgG1	Kappa	Humanized	CCR2
polatuzumab vedotin	IgG1	Kappa	Humanized	CD79B
ponezumab	IgG2	Kappa	Humanized	APP
porgaviximab	IgG1	Kappa	Chimeric	EBOV Zaire Ebola Virus
prezalumab	IgG2	Kappa	Homo sapiens	ICOSLG CD275
pritoxaximab	IgG1	Kappa	Chimeric	Shiga toxin type 1
quilizumab	IgG1	Kappa	Humanized	IGHE
rafivirumab	IgG1	Lambda	Homo sapiens	RV
ralpancizumab	IgG2	Kappa	Humanized	PCSK9
ramucirumab	IgG1	Kappa	Homo sapiens	KDR
refanezumab	IgG1	Kappa	Humanized	MAG
rilotumumab	IgG2	Kappa	Homo sapiens	HGF
rinucumab	IgG4	Kappa	Homo sapiens	PDGFRB CD140b
risankizumab	IgG1	Kappa	Humanized	IL23A
rituximab	IgG1	Kappa	Chimeric	MS4A1

robotumumab	IgG1	Kappa	Homo sapiens	IGF1R
roledumab	IgG1	Kappa	Homo sapiens	RHD CD240D
romosozumab	IgG2	Kappa	Humanized	SOST
rontalizumab	IgG1	Kappa	Humanized	IFNA1
rosmantuzumab	IgG1	Kappa	Humanized	RSPO3
rovalpituzumab	IgG1	Kappa	Humanized	DLL3
rozanolixizumab	IgG4	Kappa	Chimeric Humanized	FCGRT
sacituzumab govitecan	IgG1	Kappa	Humanized	TACSTD2
sacituzumab	IgG1	Kappa	Humanized	TACSTD2
samalizumab	IgG2	Kappa	Humanized	CD200
sarilumab	IgG1	Kappa	Homo sapiens	IL6R
satralizumab	IgG2	Kappa	Humanized	IL6R
secukinumab	IgG1	Kappa	Homo sapiens	IL17A
selicrelumab	IgG2	Kappa	Homo sapiens	TNFR TNFRSF5 CD40
seribantumab	IgG2	Lambda	Homo sapiens	ERBB3
setoxaximab	IgG1	Kappa	Chimeric	Shiga toxin type 2
sifalimumab	IgG1	Kappa	Homo sapiens	IFNA1
siltuximab	IgG1	Kappa	Chimeric	IL6
simtuzumab	IgG4	Kappa	Humanized	LOXL2
sirukumab	IgG1	Kappa	Homo sapiens	IL6
solanezumab	IgG1	Kappa	Humanized	APP
suptavumab	IgG1	Kappa	Homo sapiens	RSV glycoprotein F
suvizumab	IgG1	Kappa	Humanized	HIV-1
suvratoxumab	IgG1	Kappa	Homo sapiens	AT (alpha toxin)
tabalumab	IgG4	Kappa	Homo sapiens	TNFSF13B
tanezumab	IgG2	Kappa	Humanized	NGF
tarextumab	IgG2	Kappa	Homo sapiens	NOTCH2 NOTCH3
tavolixizumab	IgG1	Kappa	Chimeric Humanized	TNFSF4 CD252
telisotuzumab	IgG1	Kappa	Humanized	MET
teplizumab	IgG1	Kappa	Humanized	CD3E
teprotumumab	IgG1	Kappa	Homo sapiens	IGF1R
tesidolumab	IgG1	Lambda	Homo sapiens	C5
tezepelumab	IgG2	Lambda	Homo sapiens	TSLP
tigatuzumab	IgG1	Kappa	Humanized	TNFR TNFRSF10B CD262
tildrakizumab	IgG1	Kappa	Humanized	IL23A
timigutuzumab	IgG1	Kappa	Humanized	ERBB2 CD340
timolumab	IgG4	Kappa	Homo sapiens	AOC3
tisotumab	IgG1	Kappa	Homo sapiens	F3 CD142
tocilizumab	IgG1	Kappa	Humanized	IL6R
tomuzotuximab	IgG1	Kappa	Chimeric	EGFR
tosatoxumab	IgG1	Lambda	Homo sapiens	AT (alpha toxin)
tovetumab	IgG2	Kappa	Homo sapiens	PDGFRA CD140a
tralokinumab	IgG4	Lambda	Homo sapiens	IL13
trastuzumab emtansine	IgG1	Kappa	Humanized	ERBB2 CD340
trastuzumab	IgG1	Kappa	Humanized	ERBB2 CD340

tregalizumab	IgG1	Kappa	Humanized	CD4
tremelimumab	IgG2	Kappa	Homo sapiens	CTLA4 CD152
trevogrumab	IgG4	Kappa	Homo sapiens	MSTN
ublituximab	IgG1	Kappa	Chimeric	MS4A1 CD20
ulocuplumab	IgG4	Kappa	Homo sapiens	CXCR4 CD184
urelumab	IgG4	Kappa	Homo sapiens	TNFR TNFRSF9 CD137
ustekinumab	IgG1	Kappa	Homo sapiens	IL12B
utomilumab	IgG2	Lambda	Homo sapiens	TNFR TNFRSF9 CD137
vadastuximab talirine	IgG1	Kappa	Chimeric	CD33
vadastuximab	IgG1	Kappa	Chimeric	CD33
vantictumab	IgG2	Lambda	Homo sapiens	FZD
varisacumab	IgG1	Kappa	Homo sapiens	VEGFA
varlilumab	IgG1	Kappa	Homo sapiens	TNFR TNFRSF7 CD27
vatelizumab	IgG4	Kappa	Humanized	ITGA2 CD49b
vedolizumab	IgG1	Kappa	Humanized	ITGA4 ITGB7
veltuzumab	IgG1	Kappa	Humanized	MS4A1
vesencumab	IgG1	Kappa	Homo sapiens	NRP1 CD304
vonlerolizumab	IgG1	Kappa	Humanized	TNFR TNFRSF4 CD134
vorsetuzumab	IgG1	Kappa	Humanized	TNFSF7 CD70
vunakizumab	IgG1	Kappa	Humanized	IL7A
xentuzumab	IgG1	Lambda	Homo sapiens	IGF1 IGF2

## Appendix D

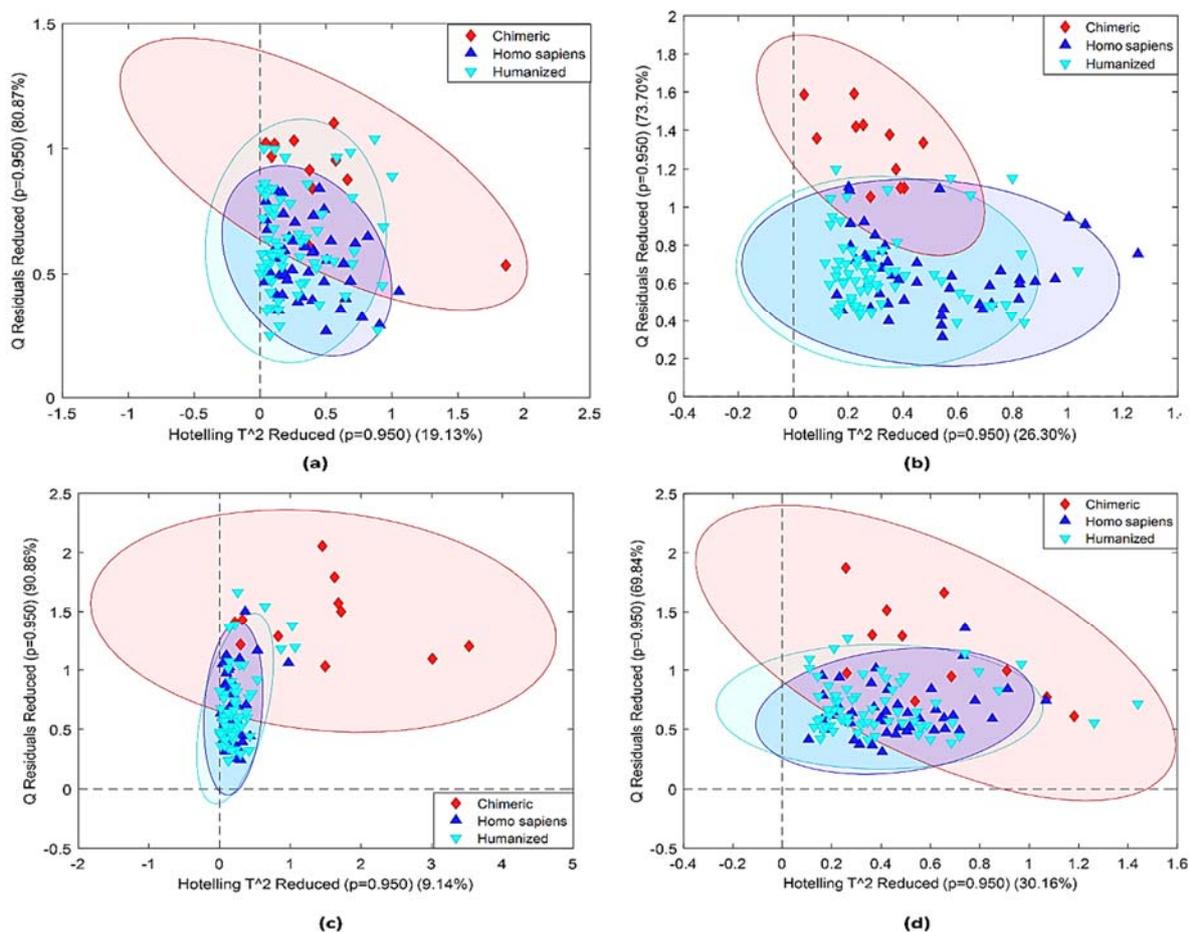
**Table D.1** Overview of 137 mAbs used in preliminary analysis

Name	HC	LC	Species	Antigen
abrituzumab	IgG2	Kappa	Humanized	ITGAV CD51
abrilumab	IgG2	Kappa	Homo sapiens	ITGA4_ITGB7 D49d
adalimumab	IgG1	Kappa	Homo sapiens	TNF
alemtuzumab	IgG1	Kappa	Humanized	CD52
alirocumab	IgG1	Kappa	Homo sapiens	PCSK9
anifrolumab	IgG1	Kappa	Homo sapiens	IFNAR1
atezolizumab	IgG1	Kappa	Humanized	CD274
bapineuzumab	IgG1	Kappa	Humanized	APP
basiliximab	IgG1	Kappa	Chimeric	IL2RA
bavituximab	IgG1	Kappa	Chimeric	phosphatidylserine
belimumab	IgG1	Lambda	Homo sapiens	TNFSF13B CD257
benralizumab	IgG1	Kappa	Humanized	IL5RA CD125
bevacizumab	IgG1	Kappa	Humanized	VEGFA
bimagrumab	IgG1	Lambda	Homo sapiens	ACVR12B/A
blosozumab	IgG4	Kappa	Humanized	SOST
bococizumab	IgG2	Kappa	Humanized	PCSK9
brentuximab	IgG1	Kappa	Chimeric	TNFRSF8
briakinumab	IgG1	Lambda	Homo sapiens	IL12B
brodalumab	IgG2	Kappa	Homo sapiens	1L17RA
canakinumab	IgG1	Kappa	Homo sapiens	IL1B
carlumab	IgG1	Kappa	Homo sapiens	CCl2
certolizumab	IgG1	Kappa	Humanized	TNF
cetuximab	IgG1	Kappa	Chimeric	EGFR
cixutumumab	IgG1	Lambda	Homo sapiens	IGF1R
clazakizumab	IgG1	Kappa	Humanized	IL6
codrituzumab	IgG1	Kappa	Humanized	GPC3
crenezumab	IgG4	Kappa	Humanized	APP
dacetuzumab	IgG1	Kappa	Humanized	CD40
daclizumab	IgG1	Kappa	Humanized	IL2RA
dalotuzumab	IgG1	Kappa	Humanized	IGF1R
daratumumab	IgG1	Kappa	Homo sapiens	CD38
denosumab	IgG2	Kappa	Homo sapiens	TNSF11 CD254
dinutuximab	IgG1	Kappa	Chimeric	GD2
drozitumab	IgG1	Lambda	Homo sapiens	TNFRSF10B
duligotuzumab	IgG1	Kappa	Humanized	ERBB3
dupilumab	IgG4	Kappa	Homo sapiens	IL4R
eculizumab	IgG2/G 4	Kappa	Humanized	C5
efalizumab	IgG1	Kappa	Humanized	ITGAL
eldelumab	IgG1	Kappa	Homo sapiens	CXCL10

elotuzumab	IgG1	Kappa	Humanized	SLAMF7
emibetuzumab	IgG4	Kappa	Humanized	MET
enokizumab	IgG1	Kappa	Humanized	IL9
epratuzumab	IgG1	Kappa	Humanized	CD22
etrolizumab	IgG1	Kappa	Humanized	ITGA4 ITGB7
evolocumab	IgG2	Lambda	Homo sapiens	PCSK9
farletuzumab	IgG1	Kappa	Humanized	FOLR1
fasinumab	IgG4	Kappa	Homo sapiens	NGF
fezakinumab	IgG1	Lambda	Homo sapiens	IL22
ficlatuzumab	IgG1	Kappa	Humanized	HGF
figitumumab	IgG2	Kappa	Homo sapiens	IGF1R
fletikumab	IgG4	Kappa	Homo sapiens	IL20
foralumab	IgG1	Kappa	Homo sapiens	CD3E
fresolimumab	IgG4	Kappa	Homo sapiens	TGFB
fulranumab	IgG2	Kappa	Homo sapiens	NGF
galiximab	IgG1	Lambda	Chimeric	CD80
ganitumab	IgG1	Kappa	Homo sapiens	IGF1R
gantenerumab	IgG1	Kappa	Homo sapiens	APP
gemtuzumab	IgG4	Kappa	Humanized	CD33
gevokizumab	IgG2	Kappa	Humanized	IL1B
girentuximab	IgG1	Kappa	Chimeric	CA9
glembatumumab	IgG2	Kappa	Homo sapiens	GPNMB
golimumab	IgG1	Kappa	Homo sapiens	TNF
guselkumab	IgG1	Lambda	Homo sapiens	IL23A
ibalizumab	IgG4	Kappa	Humanized	CD4
imgatuzumab	IgG1	Kappa	Humanized	EGFR
infliximab	IgG1	Kappa	Chimeric	TNF
inotuzumab	IgG4	Kappa	Humanized	CD22
ipilimumab	IgG1	Kappa	Homo sapiens	CTLA4
ixekizumab	IgG4	Kappa	Humanized	IL17A
lampalizumab	IgG1	Kappa	Humanized	CFD
lebrikizumab	IgG4	Kappa	Humanized	IL13
lenzilumab	IgG1	Kappa	Homo sapiens	CSF2
lintuzumab	IgG1	Kappa	Humanized	CD33
lirilumab	IgG4	Kappa	Homo sapiens	KIRD2 subgroup
lumiliximab	IgG1	Kappa	Chimeric	FCER2 CD23
matuzumab	IgG1	Kappa	Humanized	EGFR
mavrilimumab	IgG4	Lambda	Homo sapiens	CSF2RA
mepolizumab	IgG1	Kappa	Humanized	IL5
mogamulizumab	IgG1	Kappa	Humanized	CCR4
motavizumab	IgG1	Kappa	Humanized	RSV glycoprotein F
muromonab	IgG2	Kappa	Chimeric	CD3E
natalizumab	IgG4	Kappa	Humanized	ITGA4

necitumumab	IgG1	Kappa	Homo sapiens	EGFR
nimotuzumab	IgG1	Kappa	Humanized	EGFR
nivolumab	IgG4	Kappa	Homo sapiens	PDCD1
obinutuzumab	IgG1	Kappa	Humanized	MS4A1 CD20
ocrelizumab	IgG1	Kappa	Humanized	MS4A1 CD20
ofatumumab	IgG1	Kappa	Homo sapiens	MS4A1 CD20
olaratumab	IgG1	Kappa	Homo sapiens	PDGFRA
olokizumab	IgG4	Kappa	Humanized	IL6
omalizumab	IgG1	Kappa	Humanized	IGHE
onartuzumab	IgG1	Kappa	Humanized	MET
otelixizumab	IgG1	Lambda	Chimeric Humanized	CD3E
otlertuzumab	IgG1	Kappa	Humanized	CD37
ozanezumab	IgG1	Kappa	Humanized	RTN4
palivizumab	IgG1	Kappa	Humanized	RSV glycoprotein F
panitumumab	IgG2	Kappa	Homo sapiens	EGFR
panobacumab	IgM	Kappa	Homo sapiens	serotype IATS O11
parsatuzumab	IgG1	Kappa	Humanized	EGFL7
patritumab	IgG1	Kappa	Homo sapiens	ERBB3
pembrolizumab	IgG4	Kappa	Humanized	PDCD1
pertuzumab	IgG1	Kappa	Humanized	ERBB2 CD340
pinatuzumab	IgG1	Kappa	Humanized	CD22
polatuzumab	IgG1	Kappa	Humanized	CD79B
ponezumab	IgG2	Kappa	Humanized	APP
radretumab	IgE	Kappa	Homo sapiens	FN
ramucirumab	IgG1	Kappa	Homo sapiens	KDR
ranibizumab	IgG1	Kappa	Humanized	VEGFA
reslizumab	IgG4	Kappa	Humanized	IL5
rilotumumab	IgG2	Kappa	Homo sapiens	HGF
rituximab	IgG1	Kappa	Chimeric	MS4A1
robatumumab	IgG1	Kappa	Homo sapiens	IGF1R
romosozumab	IgG2	Kappa	Humanized	SOST
sarilumab	IgG1	Kappa	Homo sapiens	IL6R
secukinumab	IgG1	Kappa	Homo sapiens	IL17A
seribantumab	IgG2	Lambda	Homo sapiens	ERBB3
sifalimumab	IgG1	Kappa	Homo sapiens	IFNA1
siltuximab	IgG1	Kappa	Chimeric	IL6
simtuzumab	IgG4	Kappa	Humanized	LOXL2
sirukumab	IgG1	Kappa	Homo sapiens	IL6
tabalumab	IgG4	Kappa	Homo sapiens	TNFSF13B
tanezumab	IgG2	Kappa	Humanized	NGF
teplizumab	IgG1	Kappa	Humanized	CD3E
tigatuzumab	IgG1	Kappa	Humanized	TNFR TNFRSF10B CD262
tildrakizumab	IgG1	Kappa	Humanized	IL23A

tocilizumab	IgG1	Kappa	Humanized	IL6R
tovetumab	IgG2	Kappa	Homo sapiens	PDGFRA CD140a
tralokinumab	IgG4	Lambda	Homo sapiens	IL13
trastuzumab	IgG1	Kappa	Humanized	ERBB2 CD340
tremelimumab	IgG2	Kappa	Homo sapiens	CTLA4 CD152
urelumab	IgG4	Kappa	Homo sapiens	TNFR TNFRSF9 CD137
ustekinumab	IgG1	Kappa	Homo sapiens	IL12B
vedolizumab	IgG1	Kappa	Humanized	ITGA4 ITGB7
veltuzumab	IgG1	Kappa	Humanized	MS4A1
visilizumab	IgG2	Kappa	Humanized	CD3E
zalutumumab	IgG1	Kappa	Homo sapiens	EGFR
zanolimumab	IgG1	Kappa	Homo sapiens	CD4



**Figure D.1** Influence plots generated after performing PCA on (a) Domain, (b) Window, (c) Substructure, (d) Single Amino acid and (e) Running Sum datasets for 67 mAbs. Coloured ellipses indicate the 95% Confidence limits of the corresponding class. X axis represent Hotelling T<sup>2</sup> distribution and Y axis represents Q residuals. The percentage values indicate the percentage X variance captured.

**Table D.2** Details of mAbs selected for model development

<b>Name</b>	<b>Antigen3</b>	<b>CIC Retention Time (Min)</b>	<b>Clinical Status</b>	<b>Phage c</b>
atezolizumab	CD274	10.8	Approved	Yes
ranibizumab	VEGFA	10.1	Approved	Yes
bevacizumab	VEGFA	9.8	Approved	No
benralizumab	IL5RA CD125	9.6	Phase 3	No
duligotuzumab	ERBB3	9.5	Phase 2	Yes
ocrelizumab	MS4A1 CD20	9.5	Phase 3	No
otlertuzumab	CD37	9.5	Phase 2	No
teplizumab	CD3E	9.4	Phase 3	No
certolizumab	TNF	9.3	Approved	No
dalotuzumab	IGF1R	9.1	Phase 2	No
enokizumab	IL9	9.0	Phase 2	Yes
imgatuzumab	EGFR	9.0	Phase 2	No
ozanezumab	RTN4	9.0	Phase 2	No
vedolizumab	ITGA4 ITGB7	9.0	Approved	No
lintuzumab	CD33	8.9	Phase 3	No
onartuzumab	MET	8.9	Phase 3	No
tocilizumab	IL6R	8.9	Approved	No
etrolizumab	ITGA4 ITGB7	8.8	Phase 2	Yes
obinutuzumab	MS4A1 CD20	8.8	Approved	No
pinatuzumab	CD22	8.8	Phase 2	No
trastuzumab	ERBB2 CD340	8.8	Approved	No
veltuzumab	MS4A1	8.8	Phase 2	No
clazakizumab	IL6	8.7	Phase 2	No
epratuzumab	CD22	8.7	Phase 3	No
farletuzumab	FOLR1	8.7	Phase 3	No
motavizumab	RSV glycoprotein F	8.7	Phase 3	Yes
parsatuzumab	EGFL7	8.7	Phase 2	Yes
tigatuzumab	TNFR TNFRSF10B CD262	8.7	Phase 2	No
tildrakizumab	IL23A	8.7	Phase 3	No
bapineuzumab	APP	8.6	Phase 3	No
matuzumab	EGFR	8.6	Phase 2	No
mogamulizumab	CCR4	8.6	Approved	No
pertuzumab	ERBB2 CD340	8.6	Approved	No
alemtuzumab	CD52	8.5	Approved	No
codrituzumab	GPC3	8.5	Phase 2	No
dacetuzumab	CD40	8.5	Phase 2	No
daclizumab	IL2RA	8.5	Approved	No
efalizumab	ITGAL	8.5	Approved	No
elotuzumab	SLAMF7	8.5	Approved	No
ficlatuzumab	HGF	8.5	Phase 2	No
lampalizumab	CFD	8.5	Phase 3	No
nimotuzumab	EGFR	8.5	Approved	No

omalizumab	IGHE	8.5	Approved	No
palivizumab	RSV glycoprotein F	8.5	Approved	No
mepolizumab	IL5	8.4	Approved	No
polatuzumab	CD79B	8.3	Phase 2	Yes

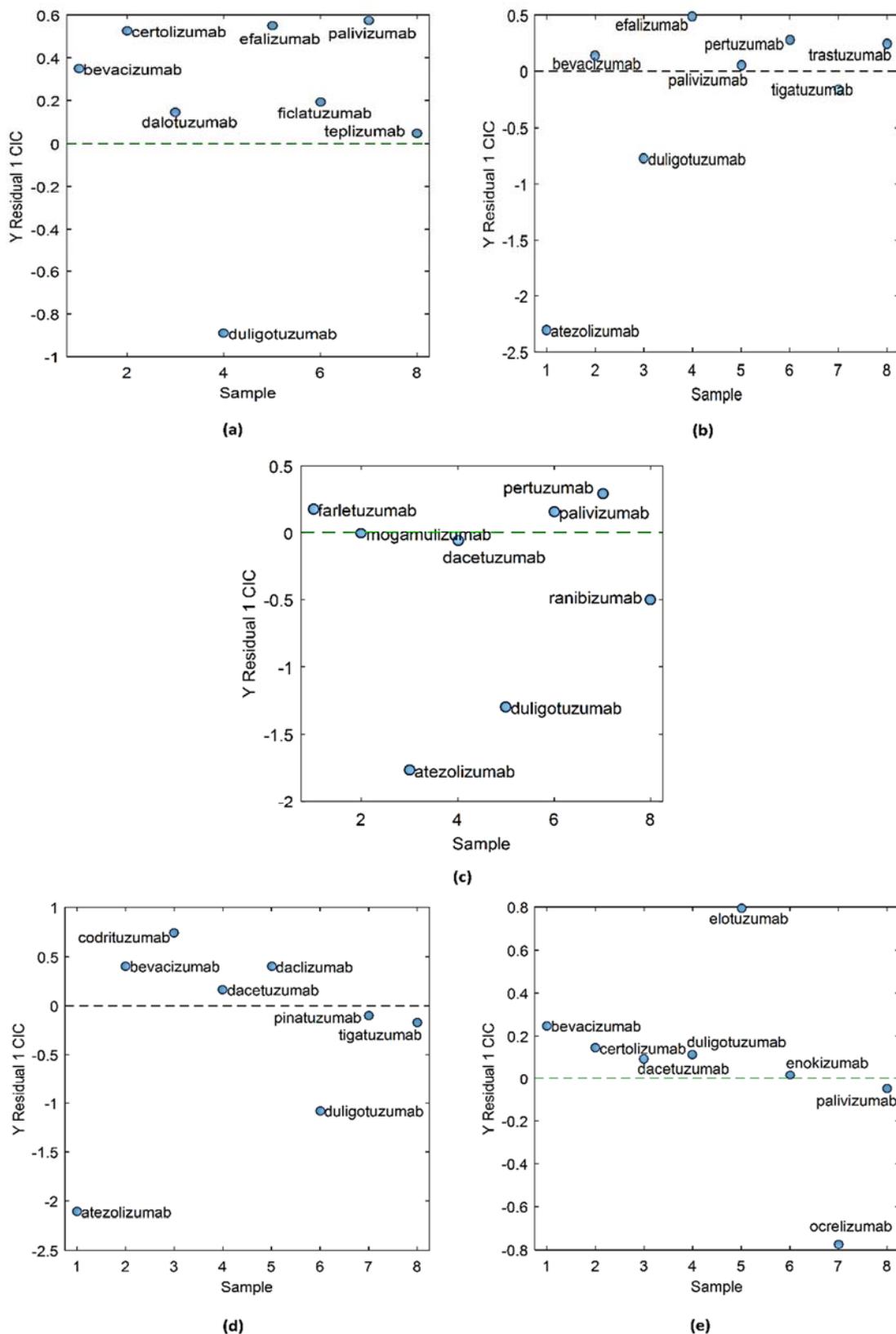
**Table D.3** Results of GA based variable selection method

Datasets	No. of unique models	Best fit of models (RMSE_CV)	No. of descriptors selected	No. of generations
Domain	14	0.35	11	25
Window	96	0.17	191	100
Substructure	73	0.11	43	100
Single amino	2	0.087619	36	98
Running sum	2	0.153	29	69

**Table D.4** Variable reduction based on the V-WSP reduction algorithm indicating the selected thresholds and corresponding Procrustes index as well as the final number of descriptors in the reduced dataset

Dataset	VH Threshold	VL Threshold	Procrustes index	Number of descriptors (original)	Number of descriptors (reduced)
Domain	0.64	0.53	0.175	272	61
Window	0.92	0.94	0.137	1336	921
Substructure	NA	NA	0.139	639	193
Single amino	0.62	0.56	0.145	1474	191
Running sum	0.5	0.51	0.137	2596	140

NA: not applicable as the thresholds were set for each substructure of each domain.



**Figure D.2** PLS model-based Y residuals of test set samples of mAbs based on GA-selected variables of (a) Domain, (b) Window, (c) Substructure, (d) Single Amino and (e) Running Sum datasets.

**Table D.5** Benchmarking of datasets based on Calibration, Cross validation and external test set prediction metrics. Models that have passes QSAR validation criteria are shown in green, those with moderate performance are indicated in Yellow and those that have failed the QSAR validation criteria are indicated in red.

			Domain	Window	Substructure	Single amino	Running sum
Calibration	Original	R2	0.68	0.54	0.52	0.69	0.50
		RMSE	0.28	0.24	0.23	0.21	0.26
		Bias	0.00	0.00	0.00	0.00	0.00
	V-WSP reduced	R2	0.68	0.93	0.53	0.99	0.83
		RMSE	0.27	0.10	0.23	0.03	0.20
		Bias	0.00	0.00	0.00	0.00	0.00
	GA selected	R2	0.77	0.99	0.99	0.89	0.99
		RMSE	0.22	0.01	0.03	0.12	0.03
		Bias	0.00	0.00	0.00	0.00	0.00
Coss validation	Original	R2	0.07	0.04	0.08	0.07	0.00
		Q2	-0.28	-0.30	-0.57	-0.50	-0.28
		RMSE	0.55	0.42	0.40	0.47	0.42
		Bias	0.01	0.02	0.02	0.00	0.01
	V-WSP reduced	R2	0.05	0.03	0.09	0.01	0.02
		Q2	-0.67	-0.18	-0.63	-0.28	-0.68
		RMSE	0.62	0.41	0.43	0.43	0.61
		Bias	0.01	0.00	0.03	0.02	-0.02
	GA selected	R2	0.55	0.89	0.83	0.62	0.93
		Q2	0.46	0.77	0.74	0.48	0.84
		RMSE	0.35	0.18	0.17	0.28	0.19
		Bias	0.00	0.00	0.02	0.00	0.01
Test	Original	R2	0.50	0.00	0.10	0.17	0.18
		RMSE	0.35	0.85	0.87	0.80	0.77
		Bias	0.05	-0.25	-0.51	-0.39	-0.13
	V-WSP reduced	R2	0.16	0.02	0.05	0.11	0.71
		RMSE	0.50	0.81	0.88	0.89	0.26
		Bias	0.10	-0.20	-0.37	-0.34	0.00
	GA selected	R2	0.83	0.82	0.88	0.69	0.99
		RMSE	0.39	0.26	0.25	0.39	0.13
		Bias	0.32	0.17	0.01	0.24	0.09

## References

- Abbas, S., Greige-Gerges, H., Karam, N., Piet, M.-H., Netter, P. and Magdalou, J. (2010) 'Metabolism of parabens (4-hydroxybenzoic acid esters) by hepatic esterases and UDP-glucuronosyltransferases in man', *Drug metabolism and pharmacokinetics*, 25(6), pp. 568-577.
- Abshear, T., Banik, G.M., D'Souza, M.L., Nedwed, K. and Peng, C. (2006) 'A model validation and consensus building environment', *SAR and QSAR in Environmental Research*, 17(3), pp. 311-321.
- Agrawal, N.J., Helk, B., Kumar, S., Mody, N., Sathish, H.A., Samra, H.S., Buck, P.M., Li, L. and Trout, B.L. (2016) *MABs*. Taylor & Francis.
- Ahmadi, M., Bryson, C.J., Cloake, E.A., Welch, K., Filipe, V., Romeijn, S., Hawe, A., Jiskoot, W., Baker, M.P. and Fogg, M.H. (2015) 'Small amounts of sub-visible aggregates enhance the immunogenic potential of monoclonal antibody therapeutics', *Pharmaceutical research*, 32(4), pp. 1383-1394.
- Ahmed, S.S., Wang, X.N., Fielding, M., Kerry, A., Dickinson, I., Munuswamy, R., Kimber, I. and Dickinson, A.M. (2016) 'An in vitro human skin test for assessing sensitization potential', *Journal of Applied Toxicology*, 36(5), pp. 669-684.
- Alexander, D.L.J., Tropsha, A. and Winkler, D.A. (2015) 'Beware of R 2: simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models', *Journal of chemical information and modeling*, 55(7), pp. 1316-1322.
- Alt, N., Zhang, T.Y., Motchnik, P., Taticek, R., Quarmby, V., Schlothauer, T., Beck, H., Emrich, T. and Harris, R.J. (2016) 'Determination of critical quality attributes for monoclonal antibodies using quality by design principles', *Biologicals*, 44(5), pp. 291-305.
- Andersen, C.M. and Bro, R. (2010) 'Variable selection in regression—a tutorial', *Journal of Chemometrics*, 24(11 - 12), pp. 728-737.
- Apweiler, R., Bairoch, A. and Wu, C.H. (2004) 'Protein sequence databases', *Current opinion in chemical biology*, 8(1), pp. 76-80.
- Ashrafi, P., Moss, G.P., Wilkinson, S.C., Davey, N. and Sun, Y. (2015) 'The application of machine learning to the modelling of percutaneous absorption: an overview and guide', *SAR and QSAR in Environmental Research*, 26(3), pp. 181-204.
- Aziz, Z., Behlke, J., Bernardi, G., Bourdillon, L., Butler, P.J.G., Carels, N., Clay, O., Colfen, H., Correia, J.J. and Daugherty, M.A. (2007) *Analytical ultracentrifugation: techniques and methods*. Royal Society of Chemistry.
- Ballabio, D., Consonni, V., Mauri, A., Claeys-Bruno, M., Sergent, M. and Todeschini, R. (2014) 'A novel variable reduction method adapted from space-filling designs', *Chemometrics and Intelligent Laboratory Systems*, 136, pp. 147-154.
- Barouch, D.H., Whitney, J.B., Moldt, B., Klein, F., Oliveira, T.Y., Liu, J., Stephenson, K.E., Chang, H.-W., Shekhar, K. and Gupta, S. (2013) 'Therapeutic efficacy of potent neutralizing

HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys', *Nature*, 503(7475), pp. 224-228.

Baumann, A., Flagella, K., Forster, R., de Haan, L., Kronenberg, S., Locher, M., Richter, W.F., Theil, F.-P. and Todd, M. (2014) 'New challenges and opportunities in nonclinical safety testing of biologics', *Regulatory Toxicology and Pharmacology*, 69(2), pp. 226-233.

Berkowitz, S.A. (2006) 'Role of analytical ultracentrifugation in assessing the aggregation of protein biopharmaceuticals', *The AAPS journal*, 8(3), pp. E590-E605.

Berkowitz, S.A. and Philo, J.S. (2015) 'Characterizing biopharmaceuticals using analytical ultracentrifugation', in *Biophysical characterization of proteins in developing biopharmaceuticals*. Elsevier, pp. 211-260.

Berlec, A. and Štrukelj, B. (2013) 'Current state and recent advances in biopharmaceutical production in Escherichia coli, yeasts and mammalian cells', *Journal of industrial microbiology & biotechnology*, 40(3-4), pp. 257-274.

Bethea, D., Wu, S.-J., Luo, J., Hyun, L., Lacy, E.R., Teplyakov, A., Jacobs, S.A., O'Neil, K.T., Gilliland, G.L. and Feng, Y. (2012) 'Mechanisms of self-association of a human monoclonal antibody CNTO607', *Protein Engineering, Design & Selection*, 25(10), pp. 531-538.

Błądzka, D., Gromadzińska, J. and Wąsowicz, W. (2014) 'Parabens. From environmental studies to human health', *Environment international*, 67, pp. 27-42.

Boberg, J., Axelstad, M., Svingen, T., Mandrup, K., Christiansen, S., Vinggaard, A.M. and Hass, U. (2016) 'Multiple endocrine disrupting effects in rats perinatally exposed to butylparaben', *Toxicological Sciences*, 152(1), pp. 244-256.

Boland, W.K. and Bebb, G. (2009) 'Nimotuzumab: a novel anti-EGFR monoclonal antibody that retains anti-EGFR activity while minimizing skin toxicity', *Expert opinion on biological therapy*, 9(9), pp. 1199-1206.

Bologna, L., Gotti, E., Manganini, M., Rambaldi, A., Intermesoli, T., Introna, M. and Golay, J. (2011) 'Mechanism of action of type II, glycoengineered, anti-CD20 monoclonal antibody GA101 in B-chronic lymphocytic leukemia whole blood assays in comparison with rituximab and alemtuzumab', *The Journal of Immunology*, 186(6), pp. 3762-3769.

Bolt, S., Routledge, E., Lloyd, I., Chatenoud, L., Pope, H., Gorman, S.D., Clark, M. and Waldmann, H. (1993) 'The generation of a humanized, non - mitogenic CD3 monoclonal antibody which retains in vitro immunosuppressive properties', *European journal of immunology*, 23(2), pp. 403-411.

Bournazos, S., Klein, F., Pietzsch, J., Seaman, M.S., Nussenzweig, M.C. and Ravetch, J.V. (2014) 'Broadly neutralizing anti-HIV-1 antibodies require Fc effector functions for in vivo activity', *Cell*, 158(6), pp. 1243-1253.

Boyd, P.N., Lines, A.C. and Patel, A.K. (1995) 'The effect of the removal of sialic acid, galactose and total carbohydrate on the functional activity of Campath-1H', *Molecular immunology*, 32(17), pp. 1311-1318.

- Bozek, K., Lengauer, T., Sierra, S., Kaiser, R. and Domingues, F.S. (2013) 'Analysis of physicochemical and structural properties determining HIV-1 coreceptor usage', *PLoS computational biology*, 9(3), p. e1002977.
- Bravi, G., Gancia, E., Mascagni, P., Pegna, M., Todeschini, R. and Zaliani, A. (1997) 'MS-WHIM, new 3D theoretical descriptors derived from molecular surface properties: A comparative 3D QSAR study in a series of steroids', *Journal of Computer-Aided Molecular Design*, 11(1), pp. 79-92.
- Brennan, F.R., Cavagnaro, J., McKeever, K., Ryan, P.C., Schutten, M.M., Vahle, J., Weinbauer, G.F., Marrer-Berger, E. and Black, L.E. (2018) *MAbs*. Taylor & Francis.
- Brennan, F.R. and Kiessling, A. (2017) 'In vitro assays supporting the safety assessment of immunomodulatory monoclonal antibodies', *Toxicology in Vitro*, 45, pp. 296-308.
- Bro, R. and Smilde, A.K. (2014) 'Principal component analysis', *Analytical Methods*, 6(9), pp. 2812-2831.
- Brown, P.H., Balbo, A. and Schuck, P. (2008) 'A bayesian approach for quantifying trace amounts of antibody aggregates by sedimentation velocity analytical ultracentrifugation', *The AAPS journal*, 10(3), p. 481.
- Calafat, A.M., Ye, X., Wong, L.-Y., Bishop, A.M. and Needham, L.L. (2010) 'Urinary concentrations of four parabens in the US population: NHANES 2005–2006', *Environmental health perspectives*, 118(5), p. 679.
- Caon, T., Costa, A.C.O., de Oliveira, M.A.L., Micke, G.A. and Simões, C.M.O. (2010) 'Evaluation of the transdermal permeation of different paraben combinations through a pig ear skin model', *International journal of pharmaceutics*, 391(1-2), pp. 1-6.
- Carter, P.J. (2006) 'Potent antibody therapeutics by design', *Nature Reviews Immunology*, 6(5), pp. 343-357.
- Castellanos, M.M., Snyder, J.A., Lee, M., Chakravarthy, S., Clark, N.J., McAuley, A. and Curtis, J.E. (2017) 'Characterization of Monoclonal Antibody–Protein Antigen Complexes Using Small-Angle Scattering and Molecular Modeling', *Antibodies*, 6(4), p. 25.
- Catapano, A.L. and Papadopoulos, N. (2013) 'The safety of therapeutic monoclonal antibodies: implications for cardiovascular disease and targeting the PCSK9 pathway', *Atherosclerosis*, 228(1), pp. 18-28.
- Cavagnaro, J.A. (2002) 'Preclinical safety evaluation of biotechnology-derived pharmaceuticals', *Nature Reviews Drug Discovery*, 1(6), pp. 469-475.
- Cho, H.-S., Mason, K., Ramyar, K.X., Stanley, A.M., Gabelli, S.B., Denney Jr, D.W. and Leahy, D.J. (2003) 'Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab', *Nature*, 421(6924), p. 756.
- Cho, M.-K., Kim, H.-Y., Bernado, P., Fernandez, C.O., Blackledge, M. and Zweckstetter, M. (2007) 'Amino acid bulkiness defines the local conformations and dynamics of natively unfolded  $\alpha$ -synuclein and tau', *Journal of the American Chemical Society*, 129(11), pp. 3032-3033.

Chou, C.-Y., Hsieh, Y.-H. and Chang, G.-G. (2011) 'Applications of analytical ultracentrifugation to protein size-and-shape distribution and structure-and-function analyses', *Methods*, 54(1), pp. 76-82.

Chung, W.K., Hou, Y., Holstein, M., Freed, A., Makhatadze, G.I. and Cramer, S.M. (2010) 'Investigation of protein binding affinity in multimodal chromatographic systems using a homologous protein library', *Journal of Chromatography A*, 1217(2), pp. 191-198.

Cole, J.L., Lary, J.W., Moody, T.P. and Laue, T.M. (2008) 'Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium', *Methods in cell biology*, 84, pp. 143-179.

Consonni, V., Ballabio, D. and Todeschini, R. (2009) 'Comments on the definition of the Q 2 parameter for QSAR validation', *Journal of chemical information and modeling*, 49(7), pp. 1669-1678.

Cronin, M.T.D. and Madden, J.C. (2010) *In silico toxicology: principles and applications*. Royal Society of Chemistry.

CSID:132 '<http://www.chemspider.com/Chemical-Structure.132.html>'.

CSID:6916 '<http://www.chemspider.com/Chemical-Structure.6916.html>'.

CSID:7176 '<http://www.chemspider.com/Chemical-Structure.7176.html>'.

Darbre, P.D. and Harvey, P.W. (2008) 'Paraben esters: review of recent studies of endocrine toxicity, absorption, esterase and human exposure, and discussion of potential human health risks', *Journal of applied toxicology*, 28(5), pp. 561-578.

Demo, S.D., Masuda, E., Rossi, A.B., Thronset, B.T., Gerard, A.L., Chan, E.H., Armstrong, R.J., Fox, B.P., Lorens, J.B. and Payan, D.G. (1999) 'Quantitative measurement of mast cell degranulation using a novel flow cytometric annexin-V binding assay', *Cytometry*, 36(4), pp. 340-348.

Dickinson, A.M., Sviland, L., Wang, X.N., Jackson, G., Taylor, P.R.A., Dunn, A. and Proctor, S.J. (1998) 'Predicting graft-versus-host disease in HLA-identical bone marrow transplants: A comparison of T-cell frequency analysis and a human skin explant model', *Transplantation*, 66(7), pp. 857-863.

Dostalek, M., Prueksaritanont, T. and Kelley, R.F. (2017) *MAbs*. Taylor & Francis.

Du, J., Wang, H., Zhong, C., Peng, B., Zhang, M., Li, B., Huo, S., Guo, Y. and Ding, J. (2007) 'Structural basis for recognition of CD20 by therapeutic antibody Rituximab', *Journal of Biological Chemistry*, 282(20), pp. 15073-15080.

DuBay, K.F., Pawar, A.P., Chiti, F., Zurdo, J., Dobson, C.M. and Vendruscolo, M. (2004) 'Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains', *Journal of molecular biology*, 341(5), pp. 1317-1326.

Dunbar, J., Fuchs, A., Shi, J. and Deane, C.M. (2013) 'ABangle: characterising the VH-VL orientation in antibodies', *Protein Engineering Design and Selection*, 26(10), pp. 611-620.

Dybowski, J.N., Riemenschneider, M., Hauke, S., Pyka, M., Verheyen, J., Hoffmann, D. and Heider, D. (2011) 'Improved Bevirimat resistance prediction by combination of structural and sequence-based classifiers', *BioData mining*, 4(1), p. 26.

Ecker, D.M., Jones, S.D. and Levine, H.L. (2015) *MAbs*. Taylor & Francis.

Eisenbrand, G., Pool-Zobel, B., Baker, V., Balls, M., Blaauboer, B.J., Boobis, A., Carere, A., Kevekordes, S., Lhuguenot, J.C. and Pieters, R. (2002) 'Methods of in vitro toxicology', *Food and Chemical Toxicology*, 40(2), pp. 193-236.

Ekins, S. (2014) 'Progress in computational toxicology', *Journal of pharmacological and toxicological methods*, 69(2), pp. 115-140.

EMA. Available at: <http://www.ema.europa.eu/ema/> (Accessed: 01-05-2018).

Esmaelian, B., Benkendorff, K., Johnston, M.R. and Abbott, C.A. (2013) 'Purified brominated indole derivatives from *Dicathais orbita* induce apoptosis and cell cycle arrest in colorectal cancer cell lines', *Marine drugs*, 11(10), pp. 3802-3822.

Eswar, N., Webb, B., Marti - Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.y., Pieper, U. and Sali, A. (2006) 'Comparative protein structure modeling using Modeller', *Current protocols in bioinformatics*, 15(1), pp. 5.6. 1-5.6. 30.

EvaluatePharma® (2017) 'World Preview (2017) Outlook to 2022', 10th edition.

EvaluatePharma® (2018) 'World Preview (2018) Outlook to 2024', 10th edition.

Fiser, A. and Šali, A. (2003) 'Modeller: generation and refinement of homology-based protein structure models', in *Methods in enzymology*. Elsevier, pp. 461-491.

Fraunhofer, W. and Winter, G. (2004) 'The use of asymmetrical flow field-flow fractionation in pharmaceuticals and biopharmaceuticals', *European Journal of Pharmaceutics and Biopharmaceutics*, 58(2), pp. 369-383.

Galvao, R.K.H., Araujo, M.C.U., Jose, G.E., Pontes, M.J.C., Silva, E.C. and Saldanha, T.C.B. (2005) 'A method for calibration and validation subset partitioning', *Talanta*, 67(4), pp. 736-740.

Gao, H., Lajiness, M.S. and Van Drie, J. (2002) 'Enhancement of binary QSAR analysis by a GA-based variable selection method', *Journal of Molecular Graphics and Modelling*, 20(4), pp. 259-268.

Gaza-Bulsecu, G. and Liu, H. (2008) 'Fragmentation of a recombinant monoclonal antibody at various pH', *Pharmaceutical research*, 25(8), pp. 1881-1890.

Geladi, P. and Kowalski, B.R. (1986) 'Partial least-squares regression: a tutorial', *Analytica chimica acta*, 185, pp. 1-17.

Gillis, C., Gouel-Chéron, A., Jönsson, F. and Bruhns, P. (2014) 'Contribution of human FcγRs to disease with evidence from human polymorphisms and transgenic animal studies', *Frontiers in immunology*, 5, p. 254.

Glasse, J. (2012) 'Multivariate data analysis for advancing the interpretation of bioprocess measurement and monitoring data', in *Measurement, Monitoring, Modelling and Control of Bioprocesses*. Springer, pp. 167-191.

Golay, J., Bologna, L., André, P.-A., Buchegger, F., Mach, J.P., Boumsell, L. and Introna, M. (2010) 'Possible misinterpretation of the mode of action of therapeutic antibodies in vitro: homotypic adhesion and flow cytometry result in artefactual direct cell death', *Blood*, 116(17), pp. 3372-3373.

Golay, J., Da Roit, F., Bologna, L., Ferrara, C., Leusen, J.H., Rambaldi, A., Klein, C. and Introna, M. (2013) 'Glycoengineered CD20 antibody obinutuzumab activates neutrophils and mediates phagocytosis through CD16B more efficiently than rituximab', *Blood*, 122(20), pp. 3482-3491.

Golay, J. and Introna, M. (2012) 'Mechanism of action of therapeutic monoclonal antibodies: Promises and pitfalls of in vitro and in vivo assays', *Archives of biochemistry and biophysics*, 526(2), pp. 146-153.

Golay, J., Lazzari, M., Facchinetti, V., Bernasconi, S., Borleri, G., Barbui, T., Rambaldi, A. and Introna, M. (2001) 'CD20 levels determine the in vitro susceptibility to rituximab and complement of B-cell chronic lymphocytic leukemia: further regulation by CD55 and CD59', *Blood*, 98(12), pp. 3383-3389.

Golden, R., Gandy, J. and Vollmer, G. (2005) 'A review of the endocrine activity of parabens and implications for potential risks to human health', *Critical reviews in toxicology*, 35(5), pp. 435-458.

Gómez de la Cuesta, R., Goodacre, R. and Ashton, L. (2014) 'Monitoring antibody aggregation in early drug development using raman spectroscopy and perturbation-correlation moving windows', *Analytical chemistry*, 86(22), pp. 11133-11140.

Green, A. and Glasse, J. (2015) 'Multivariate analysis of the effect of operating conditions on hybridoma cell metabolism and glycosylation of produced antibody', *Journal of Chemical Technology and Biotechnology*, 90(2), pp. 303-313.

Greene, N. and Pennie, W. (2015) 'Computational toxicology, friend or foe?', *Toxicology Research*, 4(5), pp. 1159-1172.

Group, C.M.C.B.W. (2009) 'A-Mab: a case study in bioprocess development', *Emeryville, CA: CASSS*.

Guerra, M.T., Sanabria, M., Leite, G.A.A., Borges, C.S., Cuciolo, M.S., Anselmo - Franci, J.A., Foster, W.G. and Kempinas, W.G. (2017) 'Maternal exposure to butyl paraben impairs testicular structure and sperm quality on male rats', *Environmental toxicology*, 32(4), pp. 1273-1289.

Guex, N. and Peitsch, M.C. (1997) 'SWISS - MODEL and the Swiss - Pdb Viewer: an environment for comparative protein modeling', *electrophoresis*, 18(15), pp. 2714-2723.

Guha, R. and Jurs, P.C. (2005) 'Determining the validity of a QSAR model— a classification approach', *Journal of chemical information and modeling*, 45(1), pp. 65-73.

Guideline, I.C.H.H.T. (1997) *International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use*.

Guideline, I.C.H.H.T. (2005) 'Immunotoxicity studies for human pharmaceuticals', *Guideline S8 Google Scholar*.

Hahn, B. and Valentine, D.T. (2016) *Essential MATLAB for engineers and scientists*. Academic Press.

Handa, O., Kokura, S., Adachi, S., Takagi, T., Naito, Y., Tanigawa, T., Yoshida, N. and Yoshikawa, T. (2006) 'Methylparaben potentiates UV-induced damage of skin keratinocytes', *Toxicology*, 227(1), pp. 62-72.

Hanke, T., Trischler, M. and Guntermann, C. (2014) 'Superagonistic anti-CD28 antibodies'. Google Patents.

Hansch, C., Hoekman, D., Leo, A., Zhang, L. and Li, P. (1995) 'The expanding role of quantitative structure-activity relationships (QSAR) in toxicology', *Toxicology letters*, 79(1), pp. 45-53.

Hansel, T.T., Kropshofer, H., Singer, T., Mitchell, J.A. and George, A.J.T. (2010) 'The safety and side effects of monoclonal antibodies', *Nature reviews Drug discovery*, 9(4), pp. 325-338.

Harding, F.A., Stickler, M.M., Razo, J. and DuBridge, R. (2010) *MAbs*. Taylor & Francis.

Hardy, B., Douglas, N., Helma, C., Rautenberg, M., Jeliaskova, N., Jeliaskov, V., Nikolova, I., Benigni, R., Tcheremenskaia, O. and Kramer, S. (2010) 'Collaborative development of predictive toxicology applications', *Journal of cheminformatics*, 2(1), pp. 1-29.

Harjunpaa, A., Junnikkala, S. and Meri, S. (2000) 'Rituximab (anti-CD20) therapy of B-cell lymphomas: direct complement killing is superior to cellular effector mechanisms', *Scandinavian journal of immunology*, 51(6), p. 634.

Harville, H.M., Voorman, R. and Prusakiewicz, J.J. (2007) 'Comparison of paraben stability in human and rat skin', *Drug metabolism letters*, 1(1), pp. 17-21.

Hasegawa, K., Miyashita, Y. and Funatsu, K. (1997) 'GA strategy for variable selection in QSAR studies: GA-based PLS analysis of calcium channel antagonists', *Journal of Chemical Information and Computer Sciences*, 37(2), pp. 306-310.

Hatami, A., Carr, K., Whiteley, P., Wilkinson, S. and Dodou, K. (2017) 'Ex vivo studies for the passive transdermal permeation and extent of metabolism of methyl and butyl paraben from a cream', *British Journal of Pharmacy*, 2(2).

Hawkins, D.M. (2004) 'The problem of overfitting', *Journal of chemical information and computer sciences*, 44(1), pp. 1-12.

Hay, M., Thomas, D.W., Craighead, J.L., Economides, C. and Rosenthal, J. (2014) 'Clinical development success rates for investigational drugs', *Nature biotechnology*, 32(1), pp. 40-51.

*Herceptin* - European Medicines Agency - Europa EU. Available at: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/EPAR\\_-\\_Product\\_Information/human/000278/WC500074922.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000278/WC500074922.pdf) (Accessed: 23 July).

Hewitt, M., Cronin, M.T.D., Madden, J.C., Rowe, P.H., Johnson, C., Obi, A. and Enoch, S.J. (2007) 'Consensus QSAR models: do the benefits outweigh the complexity?', *Journal of chemical information and modeling*, 47(4), pp. 1460-1468.

Heymann, B. and Grubmüller, H. (2001) 'Molecular dynamics force probe simulations of antibody/antigen unbinding: entropic control and nonadditivity of unbinding forces', *Biophysical journal*, 81(3), pp. 1295-1313.

Hospital, A., Goñi, J.R., Orozco, M. and Gelpi, J.L. (2015) 'Molecular dynamics simulations: advances and applications', *Advances and applications in bioinformatics and chemistry: AABC*, 8, p. 37.

Hötzel, I., Theil, F.-P., Bernstein, L.J., Prabhu, S., Deng, R., Quintana, L., Lutman, J., Sibia, R., Chan, P. and Bumbaca, D. (2012) *MAbs*. Taylor & Francis.

Hurrell, J.G. (2018) *Monoclonal hybridoma antibodies: techniques and applications*. CRC press.

Ich, S. (2008) 'Guideline: Nonclinical evaluation for anticancer pharmaceuticals'. November.

Ishidate Jr, M., Hayashi, M., Sawada, M., Matsuoka, A., Yoshikawa, K., Ono, M. and Nakadate, M. (1978) 'Cytotoxicity test on medical drugs--chromosome aberration tests with Chinese hamster cells in vitro (author's transl)', *Eisei Shikenjo hokoku. Bulletin of National Institute of Hygienic Sciences*, (96), p. 55.

Ishiwatari, S., Suzuki, T., Hitomi, T., Yoshino, T., Matsukuma, S. and Tsuji, T. (2007) 'Effects of methyl paraben on skin keratinocytes', *Journal of applied toxicology*, 27(1), pp. 1-9.

Ivashkevich, A., Redon, C.E., Nakamura, A.J., Martin, R.F. and Martin, O.A. (2012) 'Use of the  $\gamma$ -H2AX assay to monitor DNA damage and repair in translational cancer research', *Cancer letters*, 327(1), pp. 123-133.

Jacobs, S.A., Wu, S.-J., Feng, Y., Bethea, D. and O'Neil, K.T. (2010) 'Cross-interaction chromatography: a rapid method to identify highly soluble monoclonal antibody candidates', *Pharmaceutical research*, 27(1), p. 65.

Jain, T., Sun, T., Durand, S., Hall, A., Houston, N.R., Nett, J.H., Sharkey, B., Bobrowicz, B., Caffry, I. and Yu, Y. (2017) 'Biophysical properties of the clinical-stage antibody landscape', *Proceedings of the National Academy of Sciences*, 114(5), pp. 944-949.

Jakobovits, A., Amado, R.G., Yang, X., Roskos, L. and Schwab, G. (2007) 'From XenoMouse technology to panitumumab, the first fully human antibody product from transgenic mice', *Nature biotechnology*, 25(10), pp. 1134-1143.

Jefferis, R. (2009) 'Glycosylation as a strategy to improve antibody-based therapeutics', *Nature reviews Drug discovery*, 8(3), pp. 226-234.

Jefferis, R. (2014) 'Monoclonal Antibodies: Mechanisms of Action', in *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 1. Monoclonal Antibody Therapeutics: Structure, Function, and Regulatory Space*. American Chemical Society, pp. 35-68.

Jenssen, H., Fjell, C.D., Cherkasov, A. and Hancock, R.E. (2008) 'QSAR modeling and computer - aided design of antimicrobial peptides', *Journal of peptide science: an official publication of the European Peptide Society*, 14(1), pp. 110-114.

Jespersen, M.C., Peters, B., Nielsen, M. and Marcatili, P. (2017) 'BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes', *Nucleic acids research*, 45(W1), pp. W24-W29.

Jones, P.T., Dear, P.H., Foote, J., Neuberger, M.S. and Winter, G. (1985) 'Replacing the complementarity-determining regions in a human antibody with those from a mouse', *Nature*, 321(6069), pp. 522-525.

Joubert, M.K., Deshpande, M., Yang, J., Reynolds, H., Bryson, C., Fogg, M., Baker, M.P., Herskovitz, J., Goletz, T.J. and Zhou, L. (2016) 'Use of In Vitro Assays to Assess Immunogenicity Risk of Antibody-Based Biotherapeutics', *PloS one*, 11(8), p. e0159328.

Joubert, M.K., Hokom, M., Eakin, C., Zhou, L., Deshpande, M., Baker, M.P., Goletz, T.J., Kerwin, B.A., Chirmule, N. and Narhi, L.O. (2012) 'Highly aggregated antibody therapeutics can enhance the in vitro innate and late-stage T-cell immune responses', *Journal of Biological Chemistry*, 287(30), pp. 25266-25279.

Joubert, M.K., Luo, Q., Nashed-Samuel, Y., Wypych, J. and Narhi, L.O. (2011) 'Classification and characterization of therapeutic antibody aggregates', *Journal of Biological Chemistry*, 286(28), pp. 25118-25133.

Kanai, S., Liu, J.U.N., Patapoff, T.W. and Shire, S.J. (2008) 'Reversible self-association of a concentrated monoclonal antibody solution mediated by Fab–Fab interaction that impacts solution viscosity', *Journal of pharmaceutical sciences*, 97(10), pp. 4219-4227.

Kang, K.-S., Che, J.-H., Ryu, D.-Y., Kim, T.-W., Li, G.-X. and Lee, Y.-S. (2002) 'Decreased sperm number and motile activity on the F1 offspring maternally exposed to butyl p-hydroxybenzoic acid (butyl paraben)', *Journal of Veterinary Medical Science*, 64(3), pp. 227-235.

Kelly, R.L., Sun, T., Jain, T., Caffry, I., Yu, Y., Cao, Y., Lynaugh, H., Brown, M., Vásquez, M. and Wittrup, K.D. (2015) *MABs*. Taylor & Francis.

Kelly, R.L., Yu, Y., Sun, T., Caffry, I., Lynaugh, H., Brown, M., Jain, T., Xu, Y. and Wittrup, K.D. (2016) *MABs*. Taylor & Francis.

Kennedy, P.J., Oliveira, C., Granja, P.L. and Sarmento, B. (2018) 'Monoclonal antibodies: Technologies for early discovery and engineering', *Critical reviews in biotechnology*, 38(3), pp. 394-408.

- Keppel, G. and Wickens, T.D. (2004) 'Simultaneous comparisons and the control of type I errors', *Design and analysis: A researcher's handbook. 4th ed. Upper Saddle River (NJ): Pearson Prentice Hall. p*, pp. 111-130.
- Kharkar, P.S., Reith, M.E. and Dutta, A.K. (2008) 'Three-dimensional quantitative structure-activity relationship (3D QSAR) and pharmacophore elucidation of tetrahydropyran derivatives as serotonin and norepinephrine transporter inhibitors', *Journal of computer-aided molecular design*, 22(1), pp. 1-17.
- Kim, S., Song, J., Park, S., Ham, S., Paek, K., Kang, M., Chae, Y., Seo, H., Kim, H.-C. and Flores, M. (2017) *mAbs*. Taylor & Francis.
- Kim, S.J., Park, Y. and Hong, H.J. (2005) 'Antibody engineering for the development of therapeutic antibodies', *Molecules and Cells*, 20(1), pp. 17-29.
- Kindt, T.J., Goldsby, R.A., Osborne, B.A. and Kuby, J. (2007) *Kuby immunology*. Macmillan.
- Kittelmann, J., Lang, K.M., Ottens, M. and Hubbuch, J. (2017) 'Orientation of monoclonal antibodies in ion-exchange chromatography: A predictive quantitative structure–activity relationship modeling approach', *Journal of Chromatography A*, 1510, pp. 33-39.
- Kizhedath, A., Wilkinson, S. and Glassey, J. (2016) 'Applicability of predictive toxicology methods for monoclonal antibody therapeutics: status Quo and scope', *Archives of Toxicology*, pp. 1-18.
- Kizhedath, A., Wilkinson, S. and Glassey, J. (2018a) 'Applicability of Traditional In Vitro Toxicity Tests for Assessing Adverse Effects of Monoclonal Antibodies: A Case Study of Rituximab and Trastuzumab', *Antibodies*, 7(3), p. 30.
- Kizhedath, A., Wilkinson, S. and Glassey, J. (2018b) 'Assessment of hepatotoxicity and dermal toxicity of butyl paraben and methyl paraben using HepG2 and HDFn in vitro models', *Toxicology in Vitro*.
- Klausen, M.S., Anderson, M.V., Jespersen, M.C., Nielsen, M. and Marcatili, P. (2015) 'LYRA, a webserver for lymphocyte receptor structural modeling', *Nucleic acids research*, 43(W1), pp. W349-W355.
- Knight, A.W., Little, S., Houck, K., Dix, D., Judson, R., Richard, A., McCarroll, N., Akerman, G., Yang, C. and Birrell, L. (2009) 'Evaluation of high-throughput genotoxicity assays used in profiling the US EPA ToxCast™ chemicals', *Regulatory Toxicology and Pharmacology*, 55(2), pp. 188-199.
- Kohavi, R. (1995) *Ijcai*. Montreal, Canada.
- Köhler, G. and Milstein, C. (1975) 'Continuous cultures of fused cells secreting antibody of predefined specificity', *nature*, 256, pp. 495-497.
- Kola, I. and Landis, J. (2004) 'Can the pharmaceutical industry reduce attrition rates?', *Nature reviews Drug discovery*, 3(8), pp. 711-716.
- Krishan, A. (1975) 'Rapid flow cytofluorometric analysis of mammalian cell cycle by propidium iodide staining', *The Journal of cell biology*, 66(1), pp. 188-193.

- Ladiwala, A., Xia, F., Luo, Q., Breneman, C.M. and Cramer, S.M. (2006) 'Investigation of protein retention and selectivity in HIC systems using quantitative structure retention relationship models', *Biotechnology and bioengineering*, 93(5), pp. 836-850.
- Lai, T., Yang, Y. and Ng, S.K. (2013) 'Advances in mammalian cell line development technologies for recombinant protein production', *Pharmaceuticals*, 6(5), pp. 579-603.
- Lash, G.E., Scaife, P.J., Innes, B.A., Otun, H.A., Robson, S.C., Searle, R.F. and Bulmer, J.N. (2006) 'Comparison of three multiplex cytokine analysis systems: Luminex, SearchLight™ and FAST Quant®', *Journal of immunological methods*, 309(1), pp. 205-208.
- Laue, T.M., Shah, B.D., Ridgeway, T.M. and Pelletier, S.L. (1992) 'Analytical Ultracentrifugation in Biochemistry and Polymer Science, edited by SE Harding, AJ Rowe & JC Horton'. Cambridge: Royal Society of Chemistry.
- Lebowitz, J., Lewis, M.S. and Schuck, P. (2002) 'Modern analytical ultracentrifugation in protein science: a tutorial review', *Protein Science*, 11(9), pp. 2067-2079.
- Lefranc, M.-P., Giudicelli, V., Duroux, P., Jabado-Michaloud, J., Folch, G., Aouinti, S., Carillon, E., Duvergey, H., Houles, A. and Paysan-Lafosse, T. (2015) 'IMGT®, the international ImMunoGeneTics information system® 25 years on', *Nucleic acids research*, 43(D1), pp. D413-D422.
- Lefranc, M.-P., Giudicelli, V., Ginestoux, C., Jabado-Michaloud, J., Folch, G., Bellahcene, F., Wu, Y., Gemrot, E., Brochet, X. and Lane, J.m. (2009) 'IMGT®, the international ImMunoGeneTics information system®', *Nucleic acids research*, 37(suppl 1), pp. D1006-D1012.
- Lefranc, M.-P., Pommié, C., Ruiz, M., Giudicelli, V., Foulquier, E., Truong, L., Thouvenin-Contet, V. and Lefranc, G. (2003) 'IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains', *Developmental & Comparative Immunology*, 27(1), pp. 55-77.
- Lefranc, M.P. (1997) 'Unique database numbering system for immunogenetic analysis', *Immunology today*, 18(11), pp. 509-509.
- Lerner, K.G., Kao, G.F., Storb, R., Buckner, C.D., Clift, R.A. and Thomas, E.D. (1974) *Transplantation proceedings*. Elsevier USA.
- Li, J. and Zhu, Z. (2010) 'Research and development of next generation of antibody-based therapeutics', *Acta Pharmacologica Sinica*, 31(9), pp. 1198-1207.
- Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y.M., Buso, N. and Lopez, R. (2015) 'The EMBL-EBI bioinformatics web and programmatic tools framework', *Nucleic acids research*, 43(W1), pp. W580-W584.
- Li, Y.H., Xu, J.Y., Tao, L., Li, X.F., Li, S., Zeng, X., Chen, S.Y., Zhang, P., Qin, C. and Zhang, C. (2016) 'SVM-Prot 2016: A Web-Server for Machine Learning Prediction of Protein Functional Families from Sequence Irrespective of Similarity', *PloS one*, 11(8), p. e0155290.
- Lin, K., Zhang, L.-w., Han, X. and Cheng, D.-y. (2017) 'Novel angiotensin I-converting enzyme inhibitory peptides from protease hydrolysates of Qula casein: Quantitative structure-

activity relationship modeling and molecular docking study', *Journal of Functional Foods*, 32, pp. 266-277.

Liu, B., Tan, X., Liang, J., Wu, S., Liu, J., Zhang, Q. and Zhu, R. (2014a) 'A reduction in reactive oxygen species contributes to dihydromyricetin-induced apoptosis in human hepatocellular carcinoma cells', *Scientific reports*, 4, p. 7041.

Liu, J., Andya, J.D. and Shire, S.J. (2006) 'A critical review of analytical ultracentrifugation and field flow fractionation methods for measuring protein aggregation', *The AAPS journal*, 8(3), pp. E580-E589.

Liu, J., Yadav, S., Andya, J., Demeule, B. and Shire, S.J. (2015) 'Chapter Nineteen-Analytical Ultracentrifugation and Its Role in Development and Research of Therapeutical Proteins', *Methods in enzymology*, 562, pp. 441-476.

Liu, L. (2015) 'Antibody Glycosylation and Its Impact on the Pharmacokinetics and Pharmacodynamics of Monoclonal Antibodies and Fc - Fusion Proteins', *Journal of pharmaceutical sciences*, 104(6), pp. 1866-1884.

Liu, M., Yang, Y.-J., Zheng, H., Zhong, X.-R., Wang, Y., Wang, Z., Wang, Y.-G. and Wang, Y.-P. (2014b) 'Membrane-bound complement regulatory proteins are prognostic factors of operable breast cancer treated with adjuvant trastuzumab: A retrospective study', *Oncology reports*, 32(6), pp. 2619-2627.

Liu, Z.-P., Wu, L.-Y., Wang, Y., Zhang, X.-S. and Chen, L. (2010) 'Prediction of protein-RNA binding sites by a random forest method with combined features', *Bioinformatics*, 26(13), pp. 1616-1622.

Lorca, M., Morales-Verdejo, C., Vásquez-Velásquez, D., Andrades-Lagos, J., Campanini-Salinas, J., Soto-Delgado, J., Recabarren-Gajardo, G. and Mella, J. (2018) 'Structure-Activity Relationships Based on 3D-QSAR CoMFA/CoMSIA and Design of Aryloxypropanol-Amine Agonists with Selectivity for the Human  $\beta$ 3-Adrenergic Receptor and Anti-Obesity and Anti-Diabetic Profiles', *Molecules*, 23(5), p. 1191.

Lower, P.A.T.A. (2015) 'Mechanism of Action', *Drugs*.

Lu, Z., Zhang, C., Cui, J., Song, Q., Wang, L., Kang, J., Li, P., Hu, X., Song, H. and Yang, J. (2014) 'Bioinformatic analysis of the membrane cofactor protein CD46 and microRNA expression in hepatocellular carcinoma', *Oncology reports*, 31(2), pp. 557-564.

Luu, K.T., Bergqvist, S., Chen, E., Hu-Lowe, D. and Kraynov, E. (2012) 'A model-based approach to predicting the human pharmacokinetics of a monoclonal antibody exhibiting target-mediated drug disposition', *Journal of Pharmacology and Experimental Therapeutics*, 341(3), pp. 702-708.

Ma, J.K.C., Drake, P.M.W. and Christou, P. (2003) 'The production of recombinant pharmaceutical proteins in plants', *Nature Reviews Genetics*, 4(10), pp. 794-805.

Ma, W.-L., Subedi, B. and Kannan, K. (2014) 'The occurrence of bisphenol A, phthalates, parabens and other environmental phenolic compounds in house dust: a review', *Current Organic Chemistry*, 18(17), pp. 2182-2199.

*MabThera*, INN-rituximab - European Medicines Agency - Europa EU. Available at: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/EPAR\\_-\\_Product\\_Information/human/000165/WC500025821.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000165/WC500025821.pdf) (Accessed: 23 July).

Martín, J.M.P., Peropadre, A., Herrero, Ó., Freire, P.F., Labrador, V. and Hazen, M.J. (2010) 'Oxidative DNA damage contributes to the toxic activity of propylparaben in mammalian cells', *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 702(1), pp. 86-91.

Matthews, E.J., Kruhlak, N.L., Benz, R.D., Sabaté, D.A., Marchant, C.A. and Contrera, J.F. (2009) 'Identification of structure–activity relationships for adverse effects of pharmaceuticals in humans: Part C: Use of QSAR and an expert system for the estimation of the mechanism of action of drug-induced hepatobiliary and urinary tract toxicities', *Regulatory toxicology and pharmacology*, 54(1), pp. 43-65.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P. and Lopez, R. (2013) 'Analysis tool web services from the EMBL-EBI', *Nucleic acids research*, 41(W1), pp. W597-W600.

Meyer, S., Leusen, J.H.W. and Boross, P. (2014) *MABs*. Taylor & Francis.

Minitab, I. (2014) 'MINITAB release 17: statistical software for windows', *Minitab Inc, USA*.

Mo, J., Jin, R., Yan, Q., Sokolowska, I., Lewis, M.J. and Hu, P. (2018) *mAbs*. Taylor & Francis.

Moore, C.A., Wilkinson, S.C., Blain, P.G., Dunn, M., Aust, G.A. and Williams, F.M. (2014) 'Percutaneous absorption and distribution of organophosphates (chlorpyrifos and dichlorvos) following dermal exposure and decontamination scenarios using in vitro human skin model', *Toxicology letters*, 229(1), pp. 66-72.

Moos, R.K., Angerer, J., Dierkes, G., Brüning, T. and Koch, H.M. (2016) 'Metabolism and elimination of methyl, iso- and n-butyl paraben in human urine after single oral dosage', *Archives of toxicology*, 90(11), pp. 2699-2709.

Moos, R.K., Angerer, J., Wittsiepe, J., Wilhelm, M., Brüning, T. and Koch, H.M. (2014) 'Rapid determination of nine parabens and seven other environmental phenols in urine samples of German children and adults', *International journal of hygiene and environmental health*, 217(8), pp. 845-853.

Muller, P.Y., Milton, M., Lloyd, P., Sims, J. and Brennan, F.R. (2009) 'The minimum anticipated biological effect level (MABEL) for selection of first human dose in clinical trials with monoclonal antibodies', *Current opinion in biotechnology*, 20(6), pp. 722-729.

Murphy, A. (2009) 'VelocImmune: immunoglobulin variable region humanized mice', *Recombinant antibodies for immunotherapy*, p. 100.

Nakagawa, Y. and Moldéus, P. (1998) 'Mechanism of p-hydroxybenzoate ester-induced mitochondrial dysfunction and cytotoxicity in isolated rat hepatocytes', *Biochemical pharmacology*, 55(11), pp. 1907-1914.

Nakagawa, Y. and Moore, G. (1999) 'Role of mitochondrial membrane permeability transition in p-hydroxybenzoate ester-induced cytotoxicity in rat hepatocytes', *Biochemical pharmacology*, 58(5), pp. 811-816.

- Nelson, M.H. and Paulos, C.M. (2015) 'Novel immunotherapies for hematologic malignancies', *Immunological reviews*, 263(1), pp. 90-105.
- Netzeva, T.I., Worth, A.P., Aldenberg, T., Benigni, R., Cronin, M.T.D., Gramatica, P., Jaworska, J.S., Kahn, S., Klopman, G. and Marchant, C.A. (2005) 'Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships', *ATLA*, 33, pp. 155-173.
- Ng, K.S. (2013) 'A simple explanation of partial least squares'.
- Nishi, H., Miyajima, M., Wakiyama, N., Kubota, K., Hasegawa, J., Uchiyama, S. and Fukui, K. (2011) 'Fc domain mediated self-association of an IgG1 monoclonal antibody under a low ionic strength condition', *Journal of bioscience and bioengineering*, 112(4), pp. 326-332.
- Nongonierma, A.B. and FitzGerald, R.J. (2016) 'Learnings from quantitative structure-activity relationship (QSAR) studies with respect to food protein-derived bioactive peptides: a review', *RSC Advances*, 6(79), pp. 75400-75413.
- Notredame, C., Higgins, D.G. and Heringa, J. (2000) 'T-Coffee: A novel method for fast and accurate multiple sequence alignment', *Journal of molecular biology*, 302(1), pp. 205-217.
- O'Malley, C.J., Montague, G.A., Martin, E.B., Liddell, J.M., Kara, B. and Titchener-Hooker, N.J. (2012) 'Utilisation of key descriptors from protein sequence data to aid bioprocess route selection', *Food and bioproducts processing*, 90(4), pp. 755-761.
- Obrezanova, O., Arnell, A., de la Cuesta, R.G., Berthelot, M.E., Gallagher, T.R.A., Zurdo, J. and Stallwood, Y. (2015) *MABs*. Taylor & Francis.
- Oishi, S. (2002) 'Effects of butyl paraben on the male reproductive system in mice', *Archives of toxicology*, 76(7), pp. 423-429.
- Olimpieri, P.P., Chailyan, A., Tramontano, A. and Marcatili, P. (2013) 'Prediction of site-specific interactions in antibody-antigen complexes: the proABC method and server', *Bioinformatics*, p. btt369.
- Ong, S.A.K., Lin, H.H., Chen, Y.Z., Li, Z.R. and Cao, Z. (2007) 'Efficacy of different protein descriptors in predicting protein functional families', *Bmc Bioinformatics*, 8(1), p. 300.
- Organisation for Economic, C.-o. and Development (2014) *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q) SAR] Models*. OECD Publishing.
- Pawar, A.P., DuBay, K.F., Zurdo, J., Chiti, F., Vendruscolo, M. and Dobson, C.M. (2005) 'Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases', *Journal of molecular biology*, 350(2), pp. 379-392.
- Pedersen, S., Marra, F., Nicoli, S. and Santi, P. (2007) 'In vitro skin permeation and retention of parabens from cosmetic formulations', *International journal of cosmetic science*, 29(5), pp. 361-367.

- Pellicciari, R. (2017) 'Attrition in the Pharmaceutical Industry—Reasons, Implications, and Pathways Forward. Edited by Alexander Alex, C. John Harris, Dennis A. Smith', *ChemMedChem*, 12(13), pp. 1097-1098.
- Peluso, R., Cafaro, G., Di Minno, A., Iervolino, S., Ambrosino, P., Lupoli, G. and Di Minno, M.N.D. (2013) 'Side effects of TNF- $\alpha$  blockers in patients with psoriatic arthritis: evidences from literature studies', *Clinical rheumatology*, 32(6), pp. 743-753.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) 'UCSF Chimera—a visualization system for exploratory research and analysis', *Journal of computational chemistry*, 25(13), pp. 1605-1612.
- Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L. and Schulten, K. (2005) 'Scalable molecular dynamics with NAMD', *Journal of computational chemistry*, 26(16), pp. 1781-1802.
- Pindrus, M., Shire, S.J., Kelley, R.F., Demeule, B.I., Wong, R., Xu, Y. and Yadav, S. (2015) 'Solubility challenges in high concentration monoclonal antibody formulations: Relationship with amino acid sequence and intermolecular interactions', *Molecular pharmaceuticals*, 12(11), pp. 3896-3907.
- Porceddu, M., Buron, N., Roussel, C., Labbe, G., Fromenty, B. and Borgne-Sanchez, A. (2012) 'Prediction of liver injury induced by chemicals in human with a multiparametric assay on isolated mouse liver mitochondria', *Toxicological Sciences*, 129(2), pp. 332-345.
- Possel, H., Noack, H., Augustin, W., Keilhoff, G. and Wolf, G. (1997) '2, 7 - Dihydrodichlorofluorescein diacetate as a fluorescent marker for peroxynitrite formation', *FEBS letters*, 416(2), pp. 175-178.
- Povey, S., Lovering, R., Bruford, E., Wright, M., Lush, M. and Wain, H. (2001) 'The HUGO gene nomenclature committee (HGNC)', *Human genetics*, 109(6), pp. 678-680.
- Prism, G. (2014) 'version 6.0 e', *GraphPad Software*.
- Prusakiewicz, J.J., Harville, H.M., Zhang, Y., Ackermann, C. and Voorman, R.L. (2007) 'Parabens inhibit human skin estrogen sulfotransferase activity: possible link to paraben estrogenic effects', *Toxicology*, 232(3), pp. 248-256.
- Qiu, T., Qiu, J., Feng, J., Wu, D., Yang, Y., Tang, K., Cao, Z. and Zhu, R. (2016) 'The recent progress in proteochemometric modelling: focusing on target descriptors, cross-term descriptors and application scope', *Briefings in bioinformatics*, p. bbw004.
- Qiu, T., Xiao, H., Zhang, Q., Qiu, J., Yang, Y., Wu, D., Cao, Z. and Zhu, R. (2015) 'Proteochemometric Modeling of the Antigen-Antibody Interaction: New Fingerprints for Antigen, Antibody and Epitope-Paratope Interaction', *PLoS one*, 10(4), p. e0122416.
- Raju, T.S. and Jordan, R.E. (2012) *MAbs*. Taylor & Francis.
- Rathore, A.S. (2009) 'Roadmap for implementation of quality by design (QbD) for biotechnology products', *Trends in biotechnology*, 27(9), pp. 546-553.

- Rathore, A.S., Joshi, V. and Yadav, N. (2013) 'Aggregation of monoclonal antibody products: formation and removal', *BioPharm International*, 26(3), pp. 40-45.
- Rathore, A.S., Weiskopf, A. and Reason, A.J. (2014) 'Defining critical quality attributes for monoclonal antibody therapeutic products', *BioPharm International*, 27(7), pp. 34-43.
- Rathore, A.S. and Winkle, H. (2009) 'Quality by design for biopharmaceuticals', *Nature biotechnology*, 27(1), p. 26.
- Rice, P., Longden, I. and Bleasby, A. (2000) 'EMBOSS: the European molecular biology open software suite'. Elsevier Current Trends.
- Riss, T.L., Moravec, R.A., Niles, A.L., Duellman, S., Benink, H.A., Worzella, T.J. and Minor, L. (2016) 'Cell viability assays'.
- Robinson, J.R., Karkov, H.S., Woo, J.A., Krogh, B.O. and Cramer, S.M. (2017) 'QSAR models for prediction of chromatographic behavior of homologous Fab variants', *Biotechnology and bioengineering*, 114(6), pp. 1231-1240.
- Roopenian, D.C. and Akilesh, S. (2007) 'FcRn: the neonatal Fc receptor comes of age', *Nature Reviews Immunology*, 7(9), pp. 715-725.
- Rosenberg, Y., Sack, M., Montefiori, D., Forthal, D., Mao, L., Hernandez-Abanto, S., Urban, L., Landucci, G., Fischer, R. and Jiang, X. (2013) 'Rapid high-level production of functional HIV broadly neutralizing monoclonal antibodies in transient plant expression systems', *PLoS One*, 8(3), p. e58724.
- Roy, K. (2007) 'On some aspects of validation of predictive quantitative structure–activity relationship models', *Expert opinion on drug discovery*, 2(12), pp. 1567-1577.
- Roy, K., Kar, S. and Das, R.N. (2015) *Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment*. Academic press.
- Rücker, C., Rücker, G. and Meringer, M. (2007) 'y-Randomization and its variants in QSPR/QSAR', *Journal of chemical information and modeling*, 47(6), pp. 2345-2357.
- Ruiz-Blanco, Y.B., Paz, W., Green, J. and Marrero-Ponce, Y. (2015) 'ProtDCal: A program to compute general-purpose-numerical descriptors for sequences and 3D-structures of proteins', *BMC bioinformatics*, 16(1), p. 1.
- Ruxton, G.D. and Beauchamp, G. (2008) 'Time for some a priori thinking about post hoc testing', *Behavioral Ecology*, 19(3), pp. 690-693.
- Šali, A., Potterton, L., Yuan, F., van Vlijmen, H. and Karplus, M. (1995) 'Evaluation of comparative protein modeling by MODELLER', *Proteins: Structure, Function, and Bioinformatics*, 23(3), pp. 318-326.
- Sandberg, M., Eriksson, L., Jonsson, J., Sjostrom, M. and Wold, S. (1998) 'New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids', *Journal of Medicinal Chemistry*, 41(14), pp. 2481-2491.

Sander, O., Sing, T., Sommer, I., Low, A.J., Cheung, P.K., Harrigan, P.R., Lengauer, T. and Domingues, F.S. (2007) 'Structural descriptors of gp120 V3 loop for the prediction of HIV-1 coreceptor usage', *PLoS computational biology*, 3(3), p. e58.

Schuck, P. (1998) 'Sedimentation analysis of noninteracting and self-associating solutes using numerical solutions to the Lamm equation', *Biophysical Journal*, 75(3), pp. 1503-1512.

Schuck, P., Perugini, M.A., Gonzales, N.R., Howlett, G.J. and Schubert, D. (2002) 'Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems', *Biophysical journal*, 82(2), pp. 1096-1111.

Schwede, T., Kopp, J., Guex, N. and Peitsch, M.C. (2003) 'SWISS-MODEL: an automated protein homology-modeling server', *Nucleic acids research*, 31(13), pp. 3381-3385.

Sewell, F., Chapman, K., Couch, J., Dempster, M., Heidel, S., Loberg, L., Maier, C., Maclachlan, T.K., Todd, M. and van der Laan, J.W. (2017) *MAbs*. Taylor & Francis.

Shah, K.H. and Verma, R.J. (2011) 'Butyl p-hydroxybenzoic acid induces oxidative stress in mice liver—an in vivo study', *Acta Pol. Pharm*, 68(6), pp. 875-879.

Shao, J., Berger, L.F., Hendriksen, P.J.M., Peijnenburg, A.A.C.M., van Loveren, H. and Volger, O.L. (2014) 'Transcriptome-based functional classifiers for direct immunotoxicity', *Archives of toxicology*, 88(3), pp. 673-689.

Sharma, V.K., Patapoff, T.W., Kabakoff, B., Pai, S., Hilario, E., Zhang, B., Li, C., Borisov, O., Kelley, R.F. and Chorny, I. (2014) 'In silico selection of therapeutic antibodies for development: viscosity, clearance, and chemical stability', *Proceedings of the National Academy of Sciences*, 111(52), pp. 18601-18606.

Shepard, H.M., Phillips, G.L., Thanos, C.D. and Feldmann, M. (2017) 'Developments in therapy with monoclonal antibodies and related proteins', *Clinical medicine*, 17(3), pp. 220-232.

Sheridan, C. (2010) 'Fresh from the biologic pipeline [mdash] 2009', *Nature biotechnology*, 28(4), pp. 307-310.

Shields, R.L., Lai, J., Keck, R., O'Connell, L.Y., Hong, K., Meng, Y.G., Weikert, S.H.A. and Presta, L.G. (2002) 'Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fc $\gamma$ RIII and antibody-dependent cellular toxicity', *Journal of Biological Chemistry*, 277(30), pp. 26733-26740.

Shields, R.L., Namenuk, A.K., Hong, K., Meng, Y.G., Rae, J., Briggs, J., Xie, D., Lai, J., Stadlen, A. and Li, B. (2001) 'High resolution mapping of the binding site on human IgG1 for Fc $\gamma$ RI, Fc $\gamma$ RII, Fc $\gamma$ RIII, and FcRn and design of IgG1 variants with improved binding to the Fc $\gamma$ R', *Journal of Biological Chemistry*, 276(9), pp. 6591-6604.

Shitara, K. (2009) '[Potelligent antibodies as next generation therapeutic antibodies]', *Yakugaku zasshi: Journal of the Pharmaceutical Society of Japan*, 129(1), pp. 3-9.

Siegel, L.M. and Monty, K.J. (1966) 'Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation.

Application to crude preparations of sulfite and hydroxylamine reductases', *Biochimica Et Biophysica Acta (BBA)-Biophysics Including Photosynthesis*, 112(2), pp. 346-362.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M. and Söding, J. (2011) 'Fast, scalable generation of high - quality protein multiple sequence alignments using Clustal Omega', *Molecular systems biology*, 7(1), p. 539.

Singh, H., Ansari, H.R. and Raghava, G.P.S. (2013) 'Improved method for linear B-cell epitope prediction using antigen's primary sequence', *PloS one*, 8(5), p. e62216.

Song, M., Breneman, C.M., Bi, J., Sukumar, N., Bennett, K.P., Cramer, S. and Tugcu, N. (2002) 'Prediction of protein retention times in anion-exchange chromatography systems using support vector regression', *Journal of chemical information and computer sciences*, 42(6), pp. 1347-1357.

Soni, M.G., Carabin, I.G. and Burdock, G.A. (2005) 'Safety assessment of esters of p-hydroxybenzoic acid (parabens)', *Food and chemical toxicology*, 43(7), pp. 985-1015.

*StarDrop*. Available at: <http://www.optibrium.com/stardrop/stardrop-features.php> (Accessed: March).

Stebbing, R., Eastwood, D., Poole, S. and Thorpe, R. (2013) 'After TGN1412: Recent developments in cytokine release assays', *Journal of immunotoxicology*, 10(1), pp. 75-82.

Stebbing, R., Findlay, L., Edwards, C., Eastwood, D., Bird, C., North, D., Mistry, Y., Dilger, P., Liefoghe, E. and Cludts, I. (2007) "Cytokine storm" in the phase I trial of monoclonal antibody TGN1412: better understanding the causes to improve preclinical testing of immunotherapeutics', *The Journal of Immunology*, 179(5), pp. 3325-3331.

Strasser, A., Harris, A.W., Vaux, D.L., Webb, E., Bath, M.L., Adams, J.M. and Cory, S. (2013) 'Abnormalities of the immune system induced by dysregulated bcl-2 expression in transgenic mice', *Curr. Top. Microbiol. Immunol. 1990b*, 166, pp. 175-181.

Strohl, W.R. (2009) 'Optimization of Fc-mediated effector functions of monoclonal antibodies', *Current opinion in biotechnology*, 20(6), pp. 685-691.

Suntharalingam, G., Perry, M.R., Ward, S., Brett, S.J., Castello-Cortes, A., Brunner, M.D. and Panoskaltsis, N. (2006) 'Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412', *New England Journal of Medicine*, 355(10), pp. 1018-1028.

Sviland, L. and Dickinson, A.M. (1999) 'A human skin explant model for predicting graft-versus-host disease following bone marrow transplantation', *Journal of clinical pathology*, 52(12), pp. 910-913.

Szela, S., Zabłocka, A., Trzeciak, K., Drozd, A., Baranowska-Bosiacka, I., Kolasa, A., Goschorska, M., Chlubek, D. and Gutowska, I. (2016) 'Propylparaben-induced disruption of energy metabolism in human HepG2 cell line leads to increased synthesis of superoxide anions and apoptosis', *Toxicology in Vitro*, 31, pp. 30-34.

Tartaglia, G.G., Pawar, A.P., Campioni, S., Dobson, C.M., Chiti, F. and Vendruscolo, M. (2008) 'Prediction of aggregation-prone regions in structured proteins', *Journal of molecular biology*, 380(2), pp. 425-436.

Tartaglia, G.G. and Vendruscolo, M. (2008) 'The Zyggregator method for predicting protein aggregation propensities', *Chemical Society Reviews*, 37(7), pp. 1395-1401.

Tavares, R.S., Martins, F.C., Oliveira, P.J., Ramalho-Santos, J. and Peixoto, F.P. (2009) 'Parabens in male infertility—Is there a mitochondrial connection?', *Reproductive Toxicology*, 27(1), pp. 1-7.

Tessier, P.M., Wu, J. and Dickinson, C.D. (2014) 'Emerging methods for identifying monoclonal antibodies with low propensity to self-associate during the early discovery process'. Taylor & Francis.

Thomas, E.V. (2003) 'Non - parametric statistical methods for multivariate calibration model selection and comparison', *Journal of chemometrics*, 17(12), pp. 653-659.

Tian, F., Zhou, P. and Li, Z. (2007a) 'T-scale as a novel vector of topological descriptors for amino acids and its application in QSARs of peptides', *Journal of molecular structure*, 830(1), pp. 106-115.

Tian, F.F., Zhou, P. and Li, Z.L. (2007b) 'T-scale as a novel vector of topological descriptors for amino acids and its application in QSARs of peptides', *Journal of Molecular Structure*, 830(1-3), pp. 106-115.

Townsend, C.L., Laffy, J.M.J., Wu, Y.-C.B., O'Hare, J.S., Martin, V., Kipling, D., Fraternali, F. and Dunn-Walters, D.K. (2016) 'Significant differences in physicochemical properties of human immunoglobulin kappa and lambda CDR3 regions', *Frontiers in immunology*, 7.

ToxWiz. Available at: <http://camcellnet.com/products/toxwiz/> (Accessed: March).

Ulrike Bernauer, Q.C., Gisela Degen, Elsa Nielsen, Thomas Platzek, Suresh, Chandra Rastogi, C.R., Jan van Benthem, Pieter Coenraads, Maria, Dusinska, D.G., Werner Liliënblum, Andreas Luch, Manfred Metzler, Nancy and Monteiro-Rivière. (2013) 'Opinion on parabens, updated request on propyl- and butylparaben'.

Umaña, P., Jean-Mairet, J., Moudry, R., Amstutz, H. and Bailey, J.E. (1999) 'Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity', *Nature biotechnology*, 17(2), pp. 176-180.

Uramaru, N., Inoue, T., Watanabe, Y., Shigematsu, H., Ohta, S. and Kitamura, S. (2014) 'Structure-activity relationship of a series of 17 parabens and related compounds for histamine release in rat peritoneal mast cells and skin allergic reaction in guinea pigs', *The Journal of toxicological sciences*, 39(1), pp. 83-90.

van den Berg, B.A., Reinders, M.J., Hulsman, M., Wu, L., Pel, H.J., Roubos, J.A. and de Ridder, D. (2012) 'Exploring sequence characteristics related to high-level production of secreted proteins in *Aspergillus niger*', *PLoS One*, 7(10), p. e45869.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E. and Berendsen, H.J. (2005) 'GROMACS: fast, flexible, and free', *Journal of computational chemistry*, 26(16), pp. 1701-1718.

van der Voet, H. (1994) 'Comparing the predictive accuracy of models using a simple randomization test', *Chemometrics and intelligent laboratory systems*, 25(2), pp. 313-323.

van Westen, G.J.P., Swier, R.F., Cortes-Ciriano, I., Wegner, J.K., Overington, J.P., Ijzerman, A.P., van Vlijmen, H.W.T. and Bender, A. (2013a) 'Benchmarking of protein descriptor sets in proteochemometric modeling (part 2): modeling performance of 13 amino acid descriptor sets', *Journal of cheminformatics*, 5(1), p. 1.

van Westen, G.J.P., Swier, R.F., Wegner, J.K., Ijzerman, A.P., van Vlijmen, H.W.T. and Bender, A. (2013b) 'Benchmarking of protein descriptor sets in proteochemometric modeling (part 1): comparative study of 13 amino acid descriptor sets', *Journal of cheminformatics*, 5(1), p. 1.

Varma, A., Cuenca, J. and Zhu, Y. (2014) 'Compositions and methods for producing glycoproteins'. Google Patents. Available at: <https://www.google.com/patents/WO2014055370A1?cl=en>.

Veerasamy, R., Rajak, H., Jain, A., Sivadasan, S., Varghese, C.P. and Agrawal, R.K. (2011) 'Validation of QSAR models-strategies and importance', *International Journal of Drug Design & Discovery*, 3, pp. 511-519.

Vidarsson, G., Dekkers, G. and Rispens, T. (2014a) 'IgG subclasses and allotypes: from structure to effector functions', *Frontiers in Immunology*, 5.

Vidarsson, G., Dekkers, G. and Rispens, T. (2014b) 'IgG subclasses and allotypes: from structure to effector functions', *Front Immunol*, 5(520.10), p. 3389.

Vogelsang, G.B., Hess, A.D., Berkman, A.W., Tutschka, P.J., Farmer, E.R., Converse, P.J. and Santos, G.W. (1985) 'An in vitro predictive test for graft versus host disease in patients with genotypic HLA-identical bone marrow transplants', *New England Journal of Medicine*, 313(11), pp. 645-650.

Wang, W., Lu, P., Fang, Y., Hamuro, L., Pittman, T., Carr, B., Hochman, J. and Prueksaritanont, T. (2011) 'Monoclonal antibodies with identical Fc sequences can bind to FcRn differentially with pharmacokinetic consequences', *Drug Metabolism and Disposition*, 39(9), pp. 1469-1477.

Wang, X., Mathieu, M. and Brezski, R.J. (2018) 'IgG Fc engineering to modulate antibody effector functions', *Protein & cell*, 9(1), pp. 63-73.

Wang, Y., Fei, D., Vanderlaan, M. and Song, A. (2004) 'Biological activity of bevacizumab, a humanized anti-VEGF antibody in vitro', *Angiogenesis*, 7(4), pp. 335-345.

Waring, M.J., Arrowsmith, J., Leach, A.R., Leeson, P.D., Mandrell, S., Owen, R.M., Pairaudeau, G., Pennie, W.D., Pickett, S.D., Wang, J., Wallace, O. and Weir, A. (2015) 'An analysis of the attrition of drug candidates from four major pharmaceutical companies', *Nat Rev Drug Discov*, 14(7), pp. 475-486.

Whritenour, J., Casinghino, S., Collinge, M. and Zhu, X. (2016) 'Nonclinical Tools to Assess Risk of Drug Hypersensitivity Reactions', *Annual review of pharmacology and toxicology*, 56, pp. 561-576.

Wilson, A.G.E. (2011) *New Horizons in Predictive Toxicology: Current Status and Application*. Royal Society of Chemistry.

Wojtal, K.A., Rogler, G., Scharl, M., Biedermann, L., Frei, P., Fried, M., Weber, A., Eloranta, J.J., Kullak-Ublick, G.A. and Vavricka, S.R. (2012) 'Fc gamma receptor CD64 modulates the inhibitory activity of infliximab', *PloS one*, 7(8), p. e43361.

Woo, J., Parimal, S., Brown, M.R., Heden, R. and Cramer, S.M. (2015) 'The effect of geometrical presentation of multimodal cation-exchange ligands on selective recognition of hydrophobic regions on protein surfaces', *Journal of Chromatography A*, 1412, pp. 33-42.

World Health, O. (2006) 'International nonproprietary names (INN) for biological and biotechnological substances', *INN Working document*, 5, pp. 1-29.

Xu, G., Chu, Y., Jiang, N., Yang, J. and Li, F. (2012) 'The Three Dimensional Quantitative Structure Activity Relationships (3D-QSAR) and docking studies of curcumin derivatives as androgen receptor antagonists', *International journal of molecular sciences*, 13(5), pp. 6138-6155.

Xu, Y., Roach, W., Sun, T., Jain, T., Prinz, B., Yu, T.-Y., Torrey, J., Thomas, J., Bobrowicz, P. and Vásquez, M. (2013) 'Addressing polyspecificity of antibodies selected from an in vitro yeast presentation system: a FACS-based, high-throughput selection and analytical tool', *Protein Engineering, Design & Selection*, 26(10), pp. 663-670.

Yadav, S., Sreedhara, A., Kanai, S., Liu, J., Lien, S., Lowman, H., Kalonia, D.S. and Shire, S.J. (2011) 'Establishing a link between amino acid sequences and self-associating and viscoelastic behavior of two closely related monoclonal antibodies', *Pharmaceutical research*, 28(7), pp. 1750-1764.

Yamashita, M., Katakura, Y. and Shirahata, S. (2007) 'Recent advances in the generation of human monoclonal antibody', *Cytotechnology*, 55(2-3), pp. 55-60.

Ye, H. (2006) 'Simultaneous determination of protein aggregation, degradation, and absolute molecular weight by size exclusion chromatography–multiangle laser light scattering', *Analytical biochemistry*, 356(1), pp. 76-85.

Ye, X., Bishop, A.M., Reidy, J.A., Needham, L.L. and Calafat, A.M. (2006) 'Parabens as urinary biomarkers of exposure in humans', *Environmental health perspectives*, 114(12), p. 1843.

Zaliani, A. and Gancia, E. (1999) 'MS-WHIM scores for amino acids: a new 3D-description for peptide QSAR and QSPR studies', *Journal of chemical information and computer sciences*, 39(3), pp. 525-533.

Zhang, L., Ding, S., Qiao, P., Dong, L., Yu, M., Wang, C., Zhang, M., Zhang, L., Li, Y. and Tang, N. (2016) 'n - butylparaben induces male reproductive disorders via regulation of estradiol and estrogen receptors', *Journal of Applied Toxicology*, 36(9), pp. 1223-1234.

Zhao, C., Zhang, H., Luan, F., Zhang, R., Liu, M., Hu, Z. and Fan, B. (2007) 'QSAR method for prediction of protein-peptide binding affinity: application to MHC class I molecule HLA-A\* 0201', *Journal of Molecular Graphics and Modelling*, 26(1), pp. 246-254.

Zhou, P., Tian, F., Wu, Y., Li, Z. and Shang, Z. (2008) 'Quantitative sequence-activity model (QSAM): applying QSAR strategy to model and predict bioactivity and function of peptides, proteins and nucleic acids', *Current Computer-Aided Drug Design*, 4(4), pp. 311-321.

Zhu, J. (2012) 'Mammalian cell protein expression for biopharmaceutical production', *Biotechnology advances*, 30(5), pp. 1158-1170.

Zimmerman, J.M., Eliezer, N. and Simha, R. (1968) 'The characterization of amino acid sequences in proteins by statistical methods', *Journal of theoretical biology*, 21(2), pp. 170-201.