The dynamics of cholesterol metabolism and atherosclerosis across population subgroups

Andrew Parton (BSc)

Faculty of Life and Health Sciences of Ulster University

A thesis presented to Ulster University for the degree of Doctor of Philosophy

April 2018

I confirm that the word count of this thesis is less than 100,000 words

List of Figures

List of Tables

Acknowledgements

The completion of this thesis has required the guidance, assistance and support of the following individuals:

First, I would like to thank my supervisor Dr Steven Watterson for his support, dedication and enthusiasm towards this project. Steve has put an exceptional amount of work into this project, and has been insightful, encouraging and quite simply outstanding as a supervisor. Thank you for getting me here.

I would like to thank Dr Victoria McGilligan for her advice and guidance during this project. Simply knowing that there is a desk that I could knock on whenever I needed assistance was a great help. Thank you for keeping my feet in the real world when the theoretical side of this project threatened to swallow me whole.

I would like to thank my friends and colleagues at C-TRIC for their support, ideas and laughter over the last three years. In particular, I would not have been able to complete this project without the assistance of Ellie Yankova, who has given me obscene amounts of laughter, support, advice and kindness. Thank you for keeping my head in the game.

Thank you to my friends and family who have listened to me turn into a work obsessed bore during the writing of this thesis. There are too many of you to name, but thank you to each and every one of you. And mum, thank you for listening to me ramble.

To Nigel and Elna, thank you for believing in me. This is for you.

Abstract

Atherosclerosis is an inflammatory disorder characterized by the formation of plaque inside an artery wall. Despite the significance of atherosclerotic cardiovascular disease to healthcare, the pathophysiology of atherosclerosis is not fully understood. To allow us to examine the dynamical process of atherogenesis, a theoretical approach has the potential to increase knowledge of the interactions involved. A computational model of atherosclerosis has been built to study the process of atheroma formation and to suggest therapeutic hypotheses. Previously, computational models of disease pathways have aided in combinatorial drug discovery, and have led to the generation of therapeutic hypotheses. The model has been developed to conform to Systems Biology Markup Language (SBML) and Systems Biology Graphical Notation (SBGN) open standards. Collating parameters from multiple sources, the curated model displays atherosclerosis-like behaviour such as lipoprotein oxidation, cellular build-up, extra-cellular matrix formation and reverse cholesterol transport.

Publicly available genomic data has been utilised to evaluate the changes in pathway dynamics across population subgroups. Data taken from the 1000 Genome Project, a worldwide effort to create an expansive catalogue of human variation, has been used to predict a tertiary protein structure for all proteins contained within the mathematical model, and the variation in structure for a collection of mutations is studied.

A combination of molecular dynamics methods and electrostatic potential analysis are then used to estimate how the binding rates of these proteins are affected by individual mutations. These updated binding rates are subsequently used to reparameterise the mathematical model. With population data available from the 1000 Genomes Project, these new parameters can be used to study population specific dynamics of atherosclerosis, and subsequently suggest new therapeutic responses.

Abbreviations

ABCA	ATP-binding cassette transporter ABCA1	LDL	Low Density Lipoprotein
ABCG1	ATP-binding cassette sub-family G member 1	LDLC	Low Density Lipoprotein Cholesterol
ADAM17	ADAM metallopeptidase domain 17	LDLP	Low Density Lipoprotein Particle Number
AIDS	Acquired Immune Deficiency Syndrome	LDLR	Low Density Lipoprotein Receptor
AMBER	Assisted Model Building with Energy Refinement	LPL	Lipoprotein Lipase
APBS	Adaptive Poisson–Boltzmann Solver	MCP1	Monocyte Chemoattractant Protein 1
API	Application Programming Interface	MD	Molecular Dynamics
ApoB	Apolipoprotein B	MIASE	Minimum Information About a Simulation Experiment
ARVD	Atherosclerotic Renovascular Disease	MIRIAM	Minimum Information Required in the Annotation of Models
BD	Brownian Dynamics	MMP	Matrix Metalloproteinase
BRENDA	BRaunschweig ENzyme DAtabase	MMP1	Matrix Metalloproteinase 1
CAD	Coronary Artery Disease	MMP13	Matrix Metalloproteinase 13
CAMEO	Continuous Automated Model Evaluation	MMP2	Matrix Metalloproteinase 2
CAPRI	Critical Assessment of Prediction of Interactions	MMP3	Matrix Metalloproteinase 3
CASP	Critical Assessment of Protein Structure Prediction	MMP9	Matrix Metalloproteinase 9
CCL2	Chemokine (C-C motif) ligand 2	MRI	Magnetic Resonance Imaging
CCL5	Chemokine (C-C motif) ligand 5	MySQL	My Structured Query Language
CCR2	Chemokine (C-C motif) Receptor 2	NF-KB	Nuclear Factor Kappa B
CCR5	Chemokine (C-C motif) Receptor 5	NHS	National Health Service
CD34	Cluster of Differentiation 34	NMR	Nuclear Magnetic Resonance
CD36	Cluster of Differentiation 36	ODE	Ordinary Differential Equation
CD40	Cluster of Differentiation 40	OLR	Oxidized low-density lipoprotein receptor
CETP	Cholesterylester Transfer Protein	OxLDL	Oxidised Low Density Lipoprotein
CHARMM	Chemistry at Harvard Macromolecular Mechanics	PAD	Peripheral Artery Disease
CMA1	Gene for Chymase	PCSK9	Proprotein convertase subtilisin/kexin type 9
COPASI	Complex Pathway Simulator	PDB	Protein Data Bank
CPU	Central Processing Unit	PDE	Partial Differential Equation
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats	PDGF	Platelet Derived Growth Factor
CSF1	Colony Stimulating Factor 1	PDGFA	Platelet Derived Growth Factor A
CSF1R	Colony Stimulating Factor 1 Receptor	PDGFB	Platelet Derived Growth Factor B
CVD	Cardiovascular Disease	PDGFRA	Platelet Derived Growth Factor Receptor A
CXCL10	C-X-C motif chemokine 10	PDGFRB	Platelet Derived Growth Factor Receptor B
CXCL11	C-X-C motif chemokine 11	PEM	Potential Energy Minimisation

CXCL9	C-X-C motif chemokine 10	PMR2	Physiome Model Repository 2
DbSNP	Database of Single Nucleotide Polymorphisms	PPI	Protein Protein Interaction
EBI	European Bioinformatics Institute	PROCAM	Prospective Cardiovascular Munster
ECM	Extra Cellular Matrix	PRODIGY	PROtein binDIng enerGY prediction
EGF	Epidermal Growth Factor	PTM	Post-Translational Modification
EGFR	Epidermal Growth Factor Receptor	RAM	Random Access Memory
ERDF	European Regional Development Fund	RANTES	Regulated on Activation, Normal T Cell Expressed and Secreted
EU	European Union	RHD	Rheumatic Heart Disease
FFT	Fast Fourier Transform	RMSD	Root Mean Square Deviation
FRODOCK	Fast Rotational Docking	SASA	Surface Accessable Surface Area
GLG1	Golgi apparatus protein 1	SBGN	Systems Biology Graphical Notation
GO	Gene Ontology	SBGN-ML	Systems Biology Graphical Notation - Markup Language
GPU	Graphical Processing Unit	SBML	Systems Biology Markup Language
GWAS	Genome Wide Association Study	SCOP	Structural Classification of Proteins
HDL	High Density Lipoprotein	SCOPe	Structural Classification of Proteins Extended
HPF	High Power Field	SDA	Simulation of Diffusional Association
ICAM	intercellular adhesion molecules	SELP	Gene for P-Selectin
ICHEC	Irish Centre for High-End Computing	SEM	Standard Error of the Mean
IFNG	Interferon Gamma	SMC	Smooth Muscle Cells
IFNGR1	Interferon Gamma Receptor Subunit 1	SMPDB	Small Molecule Pathway Database
IFNGR2	Interferon Gamma Receptor Subunit 2	SNP	Single Nucleotide Polymorphism
IL10	Interleukin 10	SODE	Stochastic Ordinary Differential Equation
IL10RA	Interleukin 10 Receptor Subunit A	SPF	Spherical Polar Fourier Technique
IL12A	Interleukin 12 Subunit A	SPR	Surface Plasmon Resonance
IL12B	Interleukin 12 Subunit B	SPSS	Statistical Package for the Social Sciences
IL12RB1	Interleukin 12 Receptor Subunit B1	SREBP	Sterol Regulatory Element Binding Protein
IL12RB2	Interleukin 12 Receptor Subunit B2	TACE	Tumor necrosis factor-α-converting enzyme
IL17A	Interleukin 17 A	TCFA	Thin-cap Fibroatheroma
IL17RA	Interleukin 17 Receptor A	TGFB1	Transforming Growth Factor Beta
IL18	Interleukin 18	TGFBR1	Transforming Growth Factor Beta Receptor Subunit 1
IL18R1	Interleukin 18 Receptor 1	TGFBR2	Transforming Growth Factor Beta Receptor Subunit 2
IL18RAP	Interleukin 18 Receptor Accessory Protein	TIA	Transient Ischaemic Attack
IL1B	Interleukin 1B	TIMP1	Tissue Inhibitor of Matrix Metalloproteinases 1
IL1R1	Interleukin 1 Receptor 1	TIMP2	Tissue Inhibitor of Matrix Metalloproteinases 2
IL1RL1	Interleukin 1 Receptor-like 1	TIMP3	Tissue Inhibitor of Matrix Metalloproteinases 3
IL2	Interleukin 2	TIMP4	Tissue Inhibitor of Matrix

			Metalloproteinases 4
IL21R	Interleukin 21 Receptor	TNF	Tumor necrosis factor
IL21R	Interleukin 21 Receptor	TNFR	Tumor necrosis factor Receptor
IL2RA	Interleukin 2 Receptor Subunit A	TNFSFR1A	Tumor necrosis factor Superfamily Receptor 1 A
IL2RB	Interleukin 2 Receptor Subunit B	TNFSFR1B	Tumor necrosis factor Superfamily Receptor 1 A
IL2RG	Interleukin 2 Receptor Subunit G	TPSB2	Tryptase Beta 2
IL33	Interleukin 33	TPSG1	Tryptase Gamma 1
IL4	Interleukin 4	UK	United Kingdom
IL4R	Interleukin 4 Receptor	USA	United States of America
IL5	Interleukin 5	VAF	Variant Allele Frequency
IL5RA	Interleukin 5 Receptor Subunit Alpha	VANTED	Visualization and Analysis of Networks containing Experimental Data
IL6	Interleukin 6	VCAM	Vascular cell adhesion protein
IL6R	Interleukin 6 Receptor	VLDL	Very Low Density Lipoprotein
I-TASSER	Iterative Threading ASSEmbly Refinement	WES	Whole Exome Sequencing
KEGG	Kyoto Encyclopaedia of Genes and Genomes	WGS	Whole Genome Sequencing
LCP	Latex Clearing Protein	WSS	Wall Shear Stress

Declaration

I hereby declare that for 2 years following the date on which the thesis is deposited in Research Student Administration of Ulster University, the thesis shall remain confidential with access or copying prohibited. Following expiry of this period I permit:

1. The Librarian of the University to allow the thesis to be copied in whole or in part without reference to me on the understanding that such authority applies to the provision of single copies made for study purposes or for inclusion within the stock of another library
2. The thesis to be made available through the Ulster Institutional Repository and/or EThOS under the terms of the Ulster eTheses Deposit Agreement

Chapter 1:

General Introduction

1.1 Introduction

1.1.1 Stratified Medicine

Stratified medicine combines therapeutic strategies with a coinciding diagnosis to determine patient subgroups for treatment optimisation (Trusheim et al., 2011). Dividing patients into subgroups using indicative biomarkers can allow for more targeted therapeutics, creating efficiencies in drug response, medication cost, treatment length and patient safety. Stratified medicine, alongside personalised and precision medicine, could lead to a new era of healthcare customisation, utilising pharmacogenomics to maximise drug efficacy while minimising adverse effects (Mizzi et al., 2014). Patient stratification can be seen as a more practical and immediately feasible healthcare methodology than personalised medicine, due to the more simplistic nature of identifying groups with a particular drug response, rather than individual patient therapeutic optimisation (Lonergan et al., 2017). Discovery of biomarkers that indicate drug response efficacy and drug target identification are two of the main goals in stratified medicine. Increase in quantities of publicly available genome sequence data recently, due to cost reduction and technological advancements, have allowed for developments in pharmacogenomics and stratified medicine leading to a reduction in adverse drug reactions, increase in drug efficacy and reduced therapeutic cost (Trusheim et al., 2011).

1.1.2 Systems Biology

Understanding biological networks at the molecular level through biochemistry, mathematics, informatics, computer science and statistics is the primary goal of systems biology. The study of complex biological systems through mathematical modelling allows for predictions on how disruptions and perturbations affect disease pathophysiology. Biological networks, including protein-protein interaction (PPI), signalling and metabolic networks can be treated as mathematical graphs, allowing for the use of numerical approaches to biological problems. Disease typically leads to production of macromolecules not created under normal conditions, and identification and detection of these biomarkers could lead to improved therapeutic strategies. Systems Biology approaches can be used as part of a stratified medicine program, allowing for the development of therapeutic

hypotheses, discovery of disease biomarkers and the identification of drug targets (Hood, 2013).

1.1.3 Atherosclerosis

Atherosclerosis describes the build-up of fatty material embedded within the artery wall causing swelling and thickening. Damage to the endothelium, the layer of cells forming the interior surface within the vessel wall, leads to an influx of lipoproteins and immune cells (including monocytes, T cells, dendritic cells and mast cells) into the *tunica intima*, the innermost layer of the artery wall (Davignon and Ganz, 2004; Libby, 2002). Smooth muscle cells (SMCs) proliferate from the *tunica media* into the *tunica intima* and produce several matrix proteins, while long-term accumulation of these materials inside the vessel wall leads to intimal thickening and a reduction in blood flow through the artery due to the reduced size of the lumen.

1.1.4 Cardiovascular Disease

Cardiovascular disease (CVD) is responsible for more deaths worldwide than any other disorder (World Health Organisation, 2011), and has contributed to morbidity and mortality more than any other condition in the western world (Singh et al., 2002). CVD is an age-related collection of diseases of the circulatory system and heart that can lead to angina, stroke and heart failure (World Health Organisation, 2015). Types of CVD include coronary artery disease (CAD) (Hansson, 2005), rheumatic heart disease (RHD) (Sliwa and Zilla, 2012) and cerebrovascular diseases (Postiglione and Napoli, 1995). Atherosclerosis can lead to different types of CVD depending on the location and the stability of the atheroma. It is estimated that 73% of cardiovascular disease cases are due to underlying atherosclerosis (Nichols M Luengo-Fernandez R, Leal J, Gray A, Scarborough P, Rayner M, 2012). Examples of cardiovascular diseases are shown in Figure 1.1.

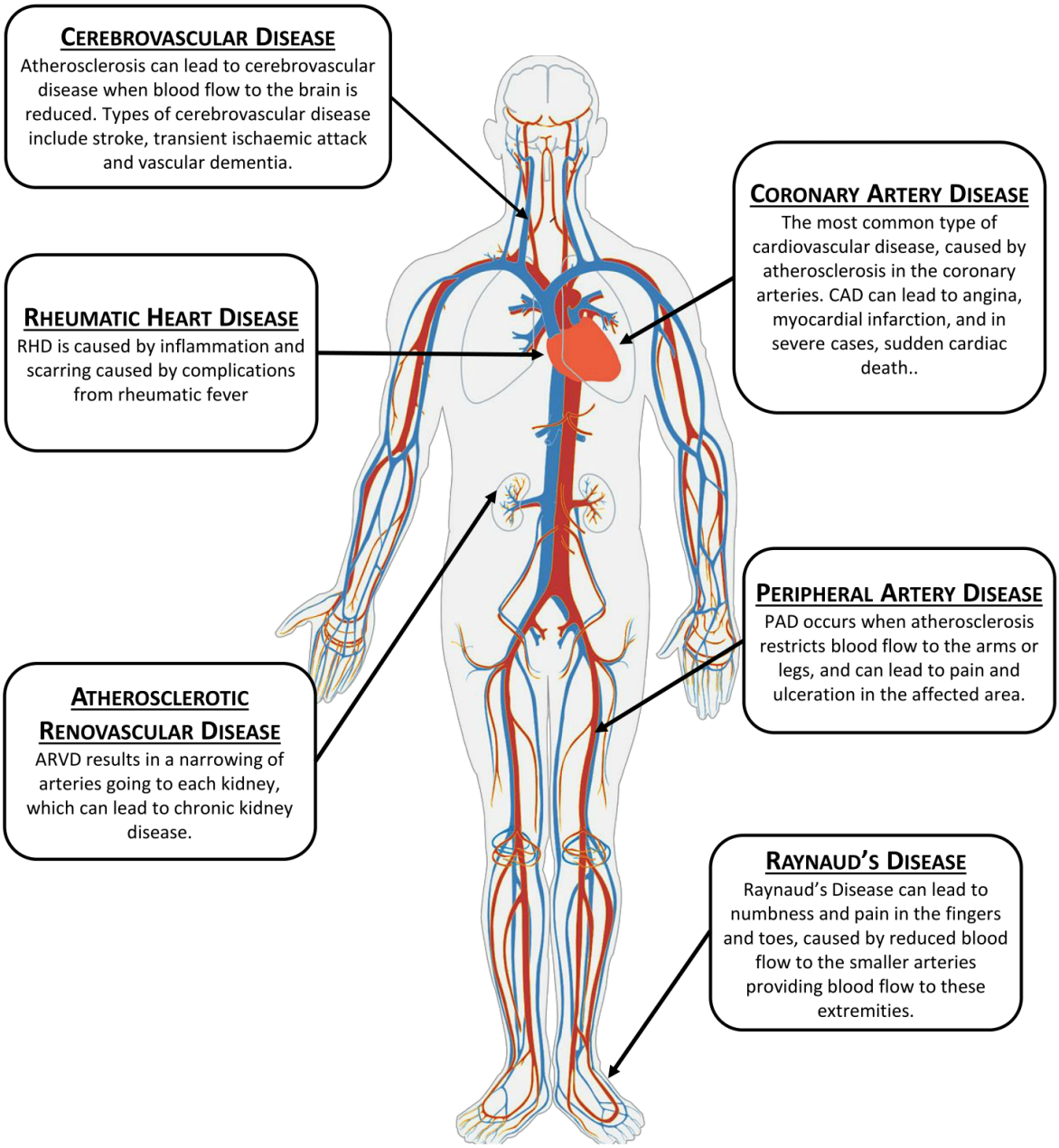


Figure 1.1: Examples of types of cardiovascular disease

1.1.5 Structure of the artery

The wall of an artery consists of three main layers, named *tunica intima*, *tunica media* and *tunica adventitia*. The *tunica intima* (also referred to as the intima) is the innermost layer of an artery wall. The intima contains the endothelium, the layer of cells within an artery, which are constantly in contact with blood flowing through the vessel. The subendothelial section of the *tunica intima* contains permeable connective tissue to provide strength and flexibility to the vessel wall. An internal elastic lamina is present in larger arteries which, similar to other intimal components, provide structure while allowing the artery to contract and expand to alter the rate of blood flow when required. Veins do not contain this internal elastic lamina (OpenStax CNX, 2013).

The *tunica media* is the middle layer of the three primary arterial layers. The predominant content of the *tunica media* is smooth muscle cells, supported by elastic fibres that function as connective tissue. This smooth muscle can control the size of the artery lumen through vasoconstriction and vasodilation. In atherosclerotic vessels, SMCs from the *tunica media* can proliferate into the *tunica intima* to help form a fibrous cap around the lipid core (Wagenseil and Mecham, 2009).

The outermost layer, the *tunica adventitia,* is mainly composed of collagen and, similar to the *tunica intima*, is supported by an external elastic lamina. Fibroblasts are the most common cell type found within the adventitia (Stenmark et al., 2013). The outer layers of the adventitia merge with the connective tissue that surrounds the artery, holding the vessel in place.

1.1.6 Atherosclerosis Incidence and Statistics

CVD is responsible for approximately 27% of all deaths in the UK, and is the second most prevalent disease after cancer (29%). In the UK, 25% of premature deaths in men (17% in women) were due to CVD in 2014. Coronary artery disease (CAD) was responsible for 45% of these deaths, and 25% were due to stroke, making these two the primary forms of CVD (Townsend et al., 2015).

It is estimated that 92.1 million adults in the USA have at least one type of CVD (Mozaffarian et al., 2014), and this is expected to rise to approximately 158 million by 2030 (Colby and Ortman, 2015; Mozaffarian et al., 2015).

While atherosclerosis is an age-related disorder, initial signs of atherosclerosis are present in most children who are older than three years old (Joseph et al., 1993). In addition, atherosclerotic lesions have been found in foetal arteries in cases of severe maternal hypercholesterolemia (Palinski and Napoli, 2002).

In addition to the human cost of cardiovascular disease, the economic costs of treatment and production loss are significant. More than 1.6 million episodes of CVD resulted in inpatient episodes in NHS hospitals in 2014. Additionally, twice as many of these inpatient episodes were for men than women. In 2014, approximately 369 million prescriptions were dispensed for CVD in the UK, with total CVD expenditure within the NHS reaching £4.292 billion. Economic costs, due to production losses, morbidity and mortality of those in working age and their relatives, contribute to the financial burden. Healthcare costs of CVD in the UK were concluded to be £11.3bn in 2014, with additional economic costs projected to be £151.6bn from reduced productivity (Townsend et al., 2015a). The economic costs of CVD in the USA, direct and indirect, were an estimated $316.1 billion in 2012/13. Including nursing home care costs in the prediction calculation gives a projection of $918 billion dollars of healthcare costs as a result of CVD between 2012 and 2030 in the USA. (Mozaffarian et al., 2015)

Patients with CVD have an increased risk of comorbidities and complications, which can lead to increases in morbidity and mortality and a reduction of quality of life. More myocardial infarctions in the US occur during a hospitalisation for another reason rather than being the initial reason for hospitalisation (Mozaffarian et al., 2015).

1.1.7 Risk Factors

In the 21st century, assessing the risk of cardiovascular disease in individuals has become easier due to the realisation that the combination of multiple risk factors can be used to predict atheroma-related risk in a patient more than any single risk factor can.

The Framingham Risk Score is a method used to estimate an individual's risk of developing cardiovascular disease (D'Agostino et al., 2008). The Framingham Risk Score combines the age, blood pressure, total cholesterol and high density lipoprotein cholesterol (HDL-C) of an individual, alongside their gender, whether they smoke and whether they have diabetes to estimate a 10-year risk percentage to aid in preventative therapy decision making.

Other scoring mechanisms for predicting CVD risk have been developed, such as Prospective Cardiovascular Munster (PROCAM), which focus on different risk factors including family history of myocardial infarction and triglyceride concentrations (Assmann et al., 2002). Focus on newly discovered and elucidated risk factors, alongside new combinations of previously known risk factors, could potentially allow for the development of a new method for assessing cardiovascular disease risk.

Major risk factors for atherosclerosis include cholesterol levels, sex, age, family history of CVD, blood pressure, smoking, obesity, diabetes, triglyceride levels, unhealthy diet, lack of physical activity, ethnicity and alcohol consumption (Go et al., 2014; Gotto, 1998; Grundy et al., 1999).

1.1.8 Clinical Outcomes

Atherosclerosis usually begins in childhood, and can remain asymptomatic for decades (Hong, 2010). Clinical outcomes are dependent on the location and the stability of an atheroma. Coronary artery disease, the most common form of CVD, is caused by presence of an atheroma within the coronary arteries, which can lead to angina, dizziness, nausea, sweating and arrhythmia. Reduced blood flow in the coronary arteries can lead to ischemia in myocardial cells, leading to myocardial infarction (MI, or heart attack). Coronary thrombosis can also lead to MI (Hansson, 2005). CAD is more common in men than women, and symptoms are more likely to present earlier in life in men (Townsend et al., 2015b).

Cerebrovascular disease, caused by atherosclerosis leading to reduced blood flow to the brain, can cause stroke, transient ischemic attack (TIA), subarachnoid haemorrhage and vascular dementia. Rupture of a plaque within the cerebral arterial network leads to thrombosis, and the subsequent blockage causes cerebral infarction, leading to stroke.

Bleeding within the cerebral circulation can also cause stroke (Postiglione and Napoli, 1995).

Peripheral artery disease, induced when atherosclerosis leads to reduced blood flow to the arms and legs, can lead to claudication, numbness, skin discolouration and muscle weakness. In extreme cases critical limb ischemia can lead to gangrene, requiring amputation (Abdulhannan et al., 2012).

Additionally, atherosclerosis can lead to fatigue, shortness of breath, muscle pain, confusion, vision problems, erectile dysfunction, hair loss, cold hands and feet, loss of kidney function and flash pulmonary oedema (Abdulhannan et al., 2012; Chiurlia et al., 2005; Chrysochou and Kalra, 2010; Muntner et al., 2005; Shahar et al., 2008; Vinkers DJ, Stek ML, van der Mast RC, de Craen AJM, Le Cessie S, Jolles J, Westendorp RGJ, 2005).

1.2 Progression of Atherosclerosis

Damage to the endothelium triggers a series of interconnected processes leading to atheroma formation, including cell signalling (Tall and Yvan-Charvet, 2015), lipoprotein transfer (Kaazempur-Mofrad and Ethier, 2001), haemodynamics (Glagov et al., 1988), lipid oxidation (Young and McEneny, 2001) and cell migration (Kraemer, 2000). Lipoproteins, including proatherogenic low density lipoproteins (LDL), enter the artery wall at sites of endothelial damage (Singh et al., 2002). Free oxygen radicals embedded within the *tunica intima* react with these lipoproteins, creating an oxidized form (Sato et al., 1990). Damaged endothelial cells release cytokines which recruit immune cells to the site of inflammation. Recruited monocytes enter the *tunica intima*, differentiate into macrophages and subsequently phagocytose the oxidized form of low density lipoprotein (oxLDL) (Zhang et al., 2013). These cholesterol filled macrophages, also known as foam cells, will go through apoptosis, leading to fatty streak formation within the arterial wall (Stoneman and Bennett, 2004). Chemotaxis factors released by immune cells within the lesion lead to smooth muscle cell proliferation from the *tunica media* into the *tunica intima*, leading to formation of a fibrous cap around the lesion in advanced atheroma (Bennett et al., 2016). Depending on plaque morphology, this fibrous cap can thin leading to an unstable plaque prone to rupture (Li et al., 2006). Plaque rupture can lead to debris being released from the vessel

wall into the blood stream, potentially leading to thrombosis, or other downstream complications (Green et al., 2002).

1.2.1 Lesion Categorisation

Structural characteristics of atherosclerotic lesions allow for classification based on histology and morphological description. Multiple categorisations of lesions have been developed over the years (Cai et al., 2002; Stary et al., 1995; Virmani et al., 2000), with the current standard described in (Yahagi et al., 2016). This classification is an update of previous efforts provided by the American Heart Association in the 1990s (Stary et al., 1994) and 2000s (Virmani et al., 2000). Four types of lesion have been split into ten subtypes to provide a histological characterisation of atherosclerotic lesions.

1.2.2 Nonatherosclerotic Intimal Lesions

1.2.2.1 Intimal thickening

Smooth muscle cells can accumulate naturally within the *tunica intima* without the presence of lipids, and this is often observed in arterial locations prone to atherosclerosis (Nakashima et al., 2002). Extra-cellular lipid and foam cells are not found within the artery wall at this point (Stary et al., 1994). These lesions do not necessarily have features that are required for advanced atheroma development, and are often known to regress, and as such are not considered to be a progressive atherosclerotic lesion.

1.2.2.2 Intimal Xanthoma

Intimal xanthomas, or 'fatty streaks,' contain layers of foam cells without a necrotic core. The majority of the lipid content within the lesion is contained within foam cells and smooth muscle cells. T lymphocytes and mast cells are present in the lesion at this time. Intimal xanthoma can be seen by the naked eye during autopsy (Stary et al., 1994).

1.2.3 Progressive Atherosclerotic Lesions

1.2.3.1 Pathological intimal thickening

Extracellular pools of lipid form beneath the layers of macrophages and foam cells contained within the *tunica intima*, disrupting the smooth muscle cells within the *tunica media* (Guyton and Klemp, 1996). These lesions are rich in SMCs and extra-cellular matrix

(ECM). As the disease state progresses, an accumulation of extracellular lipid occurs leading to a primary lipid core (Guyton and Klemp, 1996).

1.2.3.2 Fibroatheroma

Early and late fibroatheromas are characterised by the formation of a necrotic core and a significant volume of fibrous connective tissue, causing thickening of the *tunica intima*. Smooth muscle cells synthesise collagen and other ECM elements, leading to a narrowing of the lumen (Moore and Tabas, 2011). In early fibroatheroma, platelet derived growth factor (PDGF) is secreted by macrophages and foam cells, leading to smooth muscle cell migration and the formation of a thick fibrous cap (Kraemer, 2000).

1.2.3.3 Intraplaque haemorrhage or plaque fissure

Intraplaque haemorrhage can occur within a plaque when angiogenesis has occurred to form new blood vessels within the lesion. These new vessels are particularly prone to rupture (Kolodgie et al., 2003).

A plaque fissure can occur when a tear in the fibrous cap occurs without the formation of a thrombus. The necrotic core can become exposed to the bloodstream through this fissure (Lendon et al., 1992).

1.2.3.4 Thin-cap fibroatheroma

A lesion with a fibrous cap thinner than 65 µm is classified as a thin-cap fibroatheroma (TCFA) (Yahagi et al., 2016). These lesions, often called vulnerable plaques, have a high probability of rupture. TCFAs are characterised by a reduction of SMCs within the fibrous cap and the presence of a large necrotic core (Yahagi et al., 2016).

1.2.4 Lesions With Acute Thrombi

1.2.4.1 Plaque Rupture

Ruptured plaques show disruptions on the fibrous cap of the atheroma (Libby et al., 2013). Matrix metalloproteinases (MMPs) degrade the ECM and the fibrous connective tissue, leading to thinning of the fibrous cap and an increased risk of plaque rupture (Moore and Tabas, 2011).

1.2.4.2 Plaque Erosion

Plaque erosion can occur during pathological intimal thickening, or during the fibroatheroma stage of lesion development. Endothelial cells at the atheroma site are eroded and an abrasion is formed. Eroded plaques are more common in the young, smokers, and female patients (Lafont, 2003).

1.2.4.3 Calcified Nodule

A fibrocalcific plaque can lead to the jettison of a calcified nodule through unknown mechanisms. These nodules are more commonly found in older individuals (Yahagi et al., 2016).

1.2.5 Healed Lesions

1.2.5.1 Healed plaque rupture, erosion or calcified nodule

Plaques can heal after thrombus formation. A healed lesion can be calcified, contain a necrotic core, and cause significant stenosis (Yahagi et al., 2016). Lesion repair is driven by reverse cholesterol transport (Cuchel and Rader, 2006).

1.2.5.2 Plaque Rupture

The early stages of atherosclerosis are usually asymptomatic. Lesions can form over decades within an artery without knowledge and symptoms do not occur until blood flow is sufficiently reduced or thrombosis occurs (NHLBI, 2016). While knowledge of how atherogenic lipoproteins lead to plaque formation is significant, considerably less is known about the mechanisms behind plaque rupture. Bentzon *et al.* have written an excellent review considering the reasons behind this (Bentzon et al., 2014).

Acute coronary syndromes are almost always caused by the presence of thrombus or plaque haemorrhage (Davies, 2000). Plaque rupture is the most common cause of thrombosis (Badimon and Vilahur, 2014). Rupture occurs on the fibrous cap at points where thinning has occurred, potentially due to infiltration by foam cells. The thinnest point is often located in the shoulder region (Bentzon et al., 2014).

Two known mechanisms are involved in fibrous cap thinning — SMCs and collagen are slowly removed from the fibrous cap and macrophage infiltration leads to MMP production,

causing matrix degradation. Rupture can be a spontaneous event; however increases in stress can be the catalyst.

The ability to predict which plaques are vulnerable, and subsequently more likely to rupture and lead to thrombosis, would aid in reducing the burden of cardiovascular disease. A thin fibrous cap is often a sign of a vulnerable plaque, alongside a significant lipid core and inflammatory element (Fuster, 1995). Cap thickness, ECM remodelling, macrophage infiltration and wall shear stress are also involved in evaluating the vulnerable nature of an atheroma (Pinto, 2014). However, plaque vulnerability is not the same as occurrence of a coronary event – sometimes plaques can rupture and be asymptomatic in the short term.

Features of plaques prone to rupture are: large necrotic core, thin fibrous cap, high macrophage density of cap, low SMC density, significant ECM remodelling, formation of new blood vessels from the *vasa vasorum*, inflammation of the adventitia, and limited (or 'spotty') calcification (Moreno, 2010; Morrow, 2016).

1.3 Thesis Aims

1.3.1 Aim 1 – Develop a computational model of atherosclerosis

A mathematical model will be developed to study pathway dynamics of atherosclerosis. This model will become the basis of an *in silico* learning platform later in this thesis.

1.3.2 Aim 2 – Study the variation in structure for proteins related to atherosclerosis

Amino acid sequence data taken from the 1000 Genome Project will be used to predict a tertiary structure for the proteins involved in the aforementioned model of atherosclerosis, and a collection of mutations for each of these proteins.

1.3.3 Aim 3 – Predict how structural variance will change atherosclerosis dynamics

A collection of bioinformatics tools will be used to establish association rates for the protein structures previously derived from 1000 Genome Project data. These association rates will be used to predict how the structure of these proteins alter atherogenesis, and how these dynamics differ between population subgroups.

1.4 Thesis Outcomes

1.4.1 The Case For Computational Modelling

Modelling biological systems allows for the study of properties of complex systems. A system of differential equations can be developed to model biological processes, which can allow for the analysis and visualisation of these systems. Mathematical models have been built for biological systems such as a whole *Mycoplasma genitalium* cell (Karr et al., 2012a), cholesterol biosynthesis (Watterson et al., 2013) and circadian rhythms (Akman et al., 2012). Computational modelling can provide a mechanism and a framework to analyse pathway dynamics of a multi-scale system such as atherosclerosis. An approach like this can yield many benefits — there are examples of computational and mathematical modelling approaches assisting in the development of therapeutic hypotheses, and identifying potential drug targets (Berg, 2014). Development of a mathematical model of atherosclerosis would allow for the study of how model kinetics vary across multiple populations as part of a stratified medicine program. Computational biology methods have been successful within stratified medicine contexts previously (Velikova et al., 2014). This topic will be discussed further in Chapters 2 and 3.

1.4.2 Predicting Tertiary Structure

Atherosclerosis is a disorder heavily dependent on protein-protein binding interactions. Protein function is directly related to protein structure. Calculating the tertiary structure of a protein from its amino acid sequence is a problem that has been tackled by bioinformaticians and theoretical chemists for years (Dorn et al., 2014). Vast quantities of amino acid sequence data have been generated by sequencing groups, such as 1000 Genome Project (Abecasis et al., 2012) and UK10K (Muddyman et al., 2013), and while significantly more sequence data exists than experimentally derived structural data, refinement of structural prediction methods would be of great benefit to the biological community. Prediction of tertiary and quaternary structure for a collection of proteins involved in atherosclerosis would allow for the study of how these proteins vary across population subgroups. This topic will be discussed further in Chapter 4.

1.4.3 Estimating Kinetic Parameters

Protein-protein binding is a vital process throughout all phases of immune response. While surface plasmon resonance (SPR) techniques have allowed for the calculation of vast quantities of enzyme kinetics data, the ability to predict association constants through bioinformatics methods has been a challenge tackled by multiple groups with varying results (Pattnaik, 2005). Estimation of protein-protein association rate constants has been tackled using molecular dynamics methods by multiple groups (Qin et al., 2011; Yu et al., 2015b). Association constants can be used to establish how individual mutations alter model dynamics, and can subsequently be used to predict how the dynamics of atherosclerosis will differ between population subgroups. This topic will be discussed further in Chapters 5 and 6.

1.4.4 Thesis Summary

In Chapter 2, the current state of computational modelling in atherosclerosis is reviewed. Mathematical models of individual processes involved in atherosclerosis currently exist, and the work done in this field is reviewed and discussed. In addition, current knowledge of the pathogenesis of atherosclerosis is detailed.

In Chapter 3, the ordinary differential equation (ODE) model of atherosclerosis is developed. The literature mining procedure undertaken to construct the model is studied, and the differential equations used in model formation are discussed. Each biochemical reaction contained within the model is justified within the literature and sources for corresponding rate parameters are given. Experimental validity and multi-drug interventions are additionally examined.

In Chapter 4, protein structures are predicted for the proteins involved in the model of atherosclerosis. This is completed by utilising relevant genomic data from the 1000 Genome Project. The bioinformatics tools used to isolate data, formulate predictions, evaluate error and display data are discussed. The big data challenges, and the materials and methods used to overcome these are also reviewed.

In Chapter 5, protein structure data is used to predict association rates for the protein-protein interactions involved within the model. The tools used to isolate data, generate protein complexes and estimate association rates are examined.

In Chapter 6, calculated association rates are used to reparametrize the model of atherosclerosis. The generation of a series of models, representing different population subgroups, is discussed and the implications of mutations on the dynamics of our model are evaluated. A series of therapeutic hypotheses based on our *in silico* experimental system are developed.

In Chapter 7, the work undertaken in this thesis and PhD project as a whole is summarised and concluded, and potential avenues for further work within this nascent field are discussed.

Chapter 2:

Computational Modelling of Atherosclerosis

The following chapter has been published in:

Andrew Parton, Victoria McGilligan, Maurice O'Kane, Francina R. Baldrick, Steven Watterson; Computational modelling of atherosclerosis, *Briefings in Bioinformatics*, Volume 17, Issue 4, 1 July 2016, Pages 562-575, https://doi.org/10.1093/bib/bbv081

Permission has been obtained from the journal to include this article within this thesis. An additional section reviewing the literature in this field since publication has been included from section 2.7 onwards.

2.1 Introduction

Cardiovascular disease (CVD) is the primary cause of death globally (World Health Organisation, 2011) and contributes to morbidity and mortality more than any other disorder in the western world (Singh et al., 2002). In 2012, CVD was responsible for 31% of deaths worldwide, 47% of all deaths within Europe and 40% of all deaths within the European Union (Nichols M Luengo-Fernandez R, Leal J, Gray A, Scarborough P, Rayner M, 2012; Organisation, 2015). CVD covers a collection of disorders that can be split into atherosclerotic and non-atherosclerotic categories (World Health Organisation - Mendis S Norrving B editors, 2011). Atherosclerotic CVD includes cerebrovascular disease (Postiglione and Napoli, 1995), coronary artery disease (Hansson, 2005) and peripheral vascular disease (Hussein et al., 2011), and it is responsible for the majority of instances of CVD with a 2012 estimate attributing 71% of all CVD to atherosclerotic forms (Nichols M Luengo-Fernandez R, Leal J, Gray A, Scarborough P, Rayner M, 2012).

At least 75% of all CVD-related deaths occur in low and middle-income countries (World Health Organisation, 2015). In China, more than 4% of the gross national income is directly spent on the treatment of CVD (Federation, 2012) and in the EU, it is estimated that CVD costs the economy approximately €196 billion per year (Nichols M Luengo-Fernandez R, Leal J, Gray A, Scarborough P, Rayner M, 2012). Improvements in atherosclerosis and CVD treatment therefore have the potential to make a dramatic impact, not only on the quality of care, but also on the economics of healthcare.

CVD is predominantly an age related condition. Coronary heart disease in men occurs five times more frequently in 80+ year old patients than similar patients in the 40-59 age group (Mozaffarian et al., 2014). It is predicted that 22% of the global population will be 60+ years old in 2050, doubling from 11% in 2000 (Organisation, 2014). Comorbidities that drive CVD, such as diabetes (Grundy et al., 1999), are set to grow with a global increase of 55% in diabetes cases projected between 2013 and 2035 (International Diabetes Federation, 2013). The current and growing global risk of morbidity and mortality from atherosclerosis and the economic burden of treatment make atherosclerosis an important area of future research.

Despite the growing importance of atherosclerosis and its implications for public health, its pathogenesis is not fully understood (Weber and Noels, 2011). Traditionally, atherosclerosis

was viewed as a build-up of lipids (including cholesterol) within the innermost layer of the artery wall (Libby et al., 2013) (the *tunica intima*). However, our understanding has since developed and atherosclerotic CVD is now predominantly viewed as a chronic inflammatory condition, advanced by lipid build-up and triggered by innate immune responses (Joris and Majno, 1978; Weber and Noels, 2011).

Atherosclerosis emerges as the results of multiple dynamical cell processes. Damage to the endothelial cells (Viggers et al., 1986) will recruit monocytes to the site of inflammation via inter- and intra-cellular signalling (Boisvert, 2004). These monocytes will migrate into the artery wall (Kraemer, 2000), alongside lipoproteins, and phagocytose oxidised low density lipoproteins (oxLDL) (Schrijvers et al., 2007; Yoshida and Kisugi, 2010). The migration rate of these cells and particles is related to haemodynamics (Glagov et al., 1988) and vascular mechanical stress (Cunningham and Gotlieb, 2005). Cholesterol-laden macrophages within the artery wall will lead to plaque formation (Blum and Miller, 1996).

Studies aimed at understanding atherosclerosis need to be broad in scope and integrative in nature. The appropriate framework in which to consider emergent dynamical behaviour of this type is mathematical and computational modelling. A comprehensive programme of mathematical modelling and simulation can provide many benefits. Principally, it yields a framework for therapeutic hypothesis generation and for *in silico* drug target identification with the potential to streamline the drug development pipeline. This framework can be applied across populations or can be tuned to describe individual patients or patient groups as part of a programme of stratified, personalised and precision medicine (Auffray and Hood, 2012).

Mathematical and computational models can take a range of forms. Ordinary differential equations (ODEs) (Aldridge et al., 2006), partial differential equations (PDEs) (Aldridge et al., 2006) and stochastic ordinary differential equations (SODEs) (Meng et al., 2004), alongside binary (Watterson et al., 2008) and multivalued (Watterson and Ghazal, 2010) logic have all been used to model pathway dynamics. Process algebras such as pi (Guerriero et al., 2009) and kappa (Feret et al., 2009) calculus have been used to capture the structure of pathway systems, in particular addressing the exponential growth in possible network configurations to be considered as the number of pathway components increases (Kwiatkowska and Heath,

2009; Pedersen and Plotkin, 2008). Statistical models that infer pathway structure have been used to generate hypotheses from existing datasets (Grzegorczyk et al., 2008; Vert, 2010).

Computational biology approaches have previously been applied to a range of dynamical disease processes, examples include Alzheimer's disease (Lewis et al., 2010), diabetes (Ajmera et al., 2013) and breast cancer (Faratian et al., 2009). Furthermore, these computational models have been applied to pathway systems such as in nuclear factor kappa beta (NF-κB) signalling (Nelson et al., 2004), macrophage processing (Raza et al., 2010), human metabolism (Thiele et al., 2013) and iron metabolism (Mitchell and Mendes, 2013). In one of the more ambitious computational studies of recent years, the first computational model of whole cell activity has appeared describing *Mycoplasma genitalium* (Karr et al., 2012a).

Machine-readable standards for model representation have been developed to assist model development and model reuse. These standards have stimulated the creation of pathway informatics tools and have made models independent of the software tools used to create them. In particular, the Systems Biology Markup Language (SBML) (Finney and Hucka, 2003; Hucka et al., 2003) and CellML (Cuellar et al., 2003) file formats capture ODE models describing the kinetics of pathway interactions and the Systems Biology Graphical Notation Markup Language (SBGN-ML) (Van Iersel et al., 2012; Le Novere et al., 2009) encodes diagrams of pathway function in a biologically meaningful file format. The Minimum Information Requested in the Annotation of Biochemical Models (MIRIAM) (Le Novère et al., 2005) and Minimum Information About a Simulation Experiment (MIASE) (Waltemath et al., 2011) standards describe model annotation and use respectively, and online repositories of SBML files have been introduced to facilitate model reuse (Le Novere et al., 2006).

Previously, cholesterol biosynthesis and the impact of therapeutic interventions have been modelled in a series of computational studies (Bhattacharya et al., 2014; Lu et al., 2015; Mazein et al., 2013; Mc Auley et al., 2012; Watterson et al., 2013) and the role of lipid metabolism and CVD in aging has been reviewed (Mc Auley and Mooney, 2015). However, no review has yet brought together the significant volume of recent work completed on computational modelling of atherosclerosis. This paper reviews the current state of this

important nascent field, describing the work completed to date, discussing the approaches taken and highlighting the gaps in our understanding.

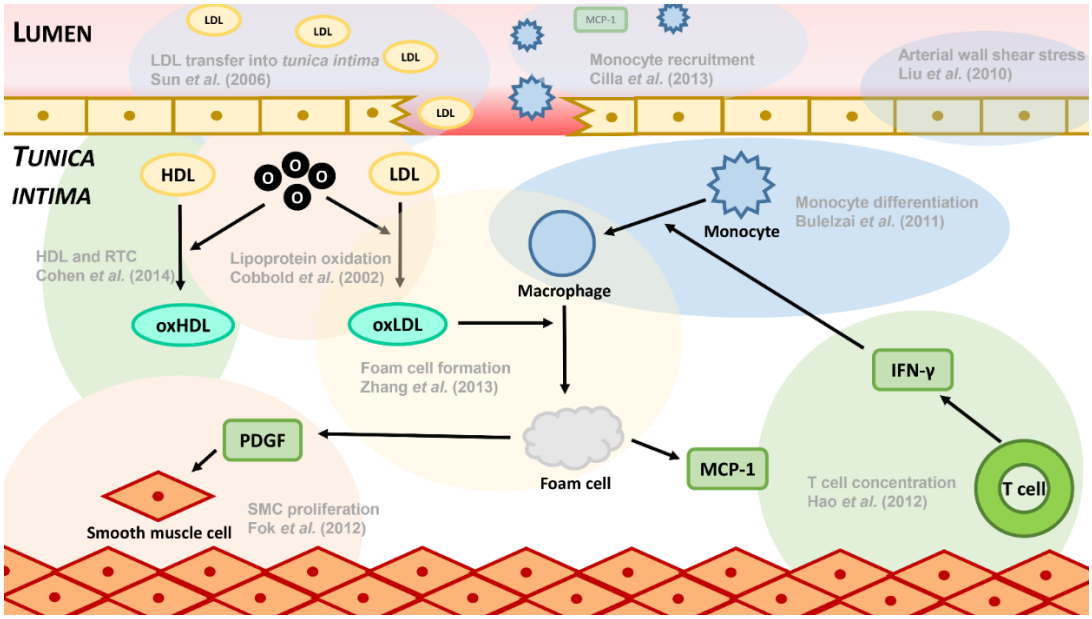


Figure 2.1: The Pathophysiology of Atherosclerosis.

Low density lipoproteins (LDL) will transfer into the artery wall at a site of endothelial damage. The role of arterial wall shear stress (WSS) and its relationship to lipoprotein transfer into the artery wall has been studied by Liu et al (Liu and Tang, 2010). Lipoproteins pass into the artery wall at a rate that depends on the WSS, lipoprotein diffusivity and lipoprotein concentration, as modeled by Sun et al (Sun et al., 2006). Once entering the intima, these lipoproteins become oxidized when they come into contact with free oxygen radicals, a process that has been modeled in more detail by Cobbold et al (Cobbold et al., 2002a). Monocytes are recruited to the site of inflammation via MCP-1 (modeled by Cilla et al. (Cilla et al., 2014)), which pass into the intima and differentiate into macrophages (Bulelzai et al. (Bulelzai and Dubbeldam, 2012)), catalyzed by T-Cell produced IFN-γ (Hao et al. (Hao and Friedman, 2014)). Macrophages will then phagocytose oxidized LDL within the artery wall, forming cholesterol-laden foam cells – a process included in Zhang et al.'s model of atherogenesis (Zhang et al., 2013). Foam cells secrete MCP-1, which recruits more monocytes to the lesion, and PDGF, which proliferates smooth muscle cells (SMCs) into the intima (Fok et al. (Fok, 2012)).

2.2 The pathophysiology of atherosclerosis

In Figure 2.1, we can see a representation of the processes that lead to atherosclerosis (Douglas Channon, KM., 2014; Lusis, 2000; Weber and Noels, 2011). Damage to the endothelial layer of the artery wall triggers an inflammatory response in which monocytes, T lymphocytes and other immune cells are recruited to the region of damage. These cells penetrate the endothelial layer, reaching the *tunica intima*, along with low density lipoprotein (LDL) and high-density lipoprotein (HDL) particles. Stimulated by the presence of interferon gamma (IFN-γ) and macrophage colony stimulating factor (M-CSF), monocytes differentiate into macrophages once they have entered the artery wall. While embedded within the *tunica intima*, both LDL and HDL become oxidized by free oxygen radicals. Macrophages will phagocytose oxidized LDL (oxLDL), but not oxidized HDL. Macrophages heavily loaded with oxLDL transform into foam cells that eventually undergo apoptosis. The resulting mass of debris embedded in the *tunica intima* is known as an atheroma. Foam cells, along with endothelial cells, secrete monocyte chemoattractant protein-1 (MCP-1) to recruit more monocytes to the site of inflammation. Naïve T cells contained within the artery wall differentiate into individual T cell types and secrete IFN-γ. Smooth muscle cells (SMCs) are also recruited into the *tunica intima* where they undergo apoptosis and contribute to the formation of a fibrous cap in the artery wall. This accumulation of cells and debris can cause a swelling of the artery wall that restricts blood flow, leading to stenosis. When the fibrous cap ruptures, the build-up in the *tunica intima* is released into the blood stream increasing the risk of blockages downstream. Further complications can occur including clotting at the site of the atheroma where a thrombus forms further impeding blood flow.

2.3 Computational modelling

2.3.1 Blood flow dynamics

Vascular damage is a key trigger for the onset of atherosclerosis that can be induced by factors such as hypertension (Alexander, 1995), smoking (Powell, 1998) and oxidative stress (Harrison et al., 2003). The elastic properties of arteries under hypertensive pressure have been modelled previously (Goriely and Vandiver, 2010). Obstructions to blood flow are known to be atherogenic (Nabel et al., 1988) and it has been shown that this is in part

attributable to the turbulent blood flow likely to be induced downstream (Cunningham and Gotlieb, 2005; Glagov et al., 1988; Li et al., 2006; Resnick et al., 2003; Stroud et al., 2002).

A number of computational studies have modelled the dynamics of blood flow (*haemodynamics*) and its relationship to vascular structure. Navier-Stokes equations are typically used to describe blood flow through arterial structures (Ai and Vafai, 2006; Calvez and Ebde, 2010; Cilla et al., 2014; Ethier, 2002; Filipovic and Kojic, 2004; Filipovic et al., 2011a; Goriely and Vandiver, 2010; Johnston et al., 2006; Li et al., 2006; Liu and Tang, 2010; Nabel et al., 1988; Olgac et al., 2008; Quarteroni et al., 2002; Rappitsch et al., 1997; Stroud et al., 2002; Sun et al., 2006; Tomaso et al., 2011; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010; Wada et al., 2002; Wang, 2001). Arterial wall shear stress (WSS) is widely used as a model output that serves as a marker for atherosclerotic prone regions within an artery (Bosnić et al., 2012; Bulelzai and Dubbeldam, 2012; Calvez and Ebde, 2010; Ethier, 2002; Filipovic and Kojic, 2004; Filipovic et al., 2011a; Gabriel et al., 2014; Glagov et al., 1988; Johnston et al., 2006; Liu and Tang, 2010; Quarteroni et al., 2002; Silva et al., 2013; Stroud et al., 2002; Sun et al., 2006; Tomaso et al., 2011; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010; Wada et al., 2002; Wang, 2001). Two-dimensional and three dimensional models of a Y-shaped arterial branch (Bosnić et al., 2012; Calvez and Ebde, 2010; Filipovic and Kojic, 2004; Filipovic et al., 2011a; Gabriel et al., 2014; Gessaghi et al., 2011; Silva et al., 2013; Stroud et al., 2002) have been created along with linear artery models (Cilla et al., 2014; Deepa et al., 2009; Filipovic and Kojic, 2004; Filipovic et al., 2011a; Green et al., 2002; Johnston et al., 2006; Li et al., 2006; Liu and Tang, 2010; Olgac et al., 2008; Rappitsch et al., 1997; Tomaso et al., 2011; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010; Wada et al., 2002).

Inflammation is driven by the penetration of the arterial wall by LDL, which in some cases is taken to be a function of the wall shear stress, demonstrating that an arterial branch can be a focal point for atheroma formation (Bosnić et al., 2012; Calvez and Ebde, 2010; Filipovic and Kojic, 2004; Filipovic et al., 2011a; Gabriel et al., 2014; Silva et al., 2013; Stroud et al., 2002; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010). As well as WSS, it has been shown that inflammation is related to blood viscosity, inlet flow rate and the diameter of the artery (Liu and Tang, 2010).

Simpler models of this process have been developed that enable atherosclerosis to be considered as a bistable system (Bulelzai and Dubbeldam, 2012). Haemodynamic models have also been developed to explore the turbulence downstream of an atherosclerotic constriction (Green et al., 2002). Haemodynamics and plaque formation have also been considered as a test case for novel numerical methods (Girke et al., 2014).

2.3.2 LDL concentration in the artery lumen

The turnover of LDL in the blood plays an important role as a primary factor that affects LDL penetration of the *tunica intima* in many models of atherosclerosis. Plasma LDL levels have been modelled as constant (Ai and Vafai, 2006; Calvez and Ebde, 2010; Filipovic et al., 2011a; Gessaghi et al., 2011), or as a variable (Cilla et al., 2014; Filipovic and Kojic, 2004; Filipovic et al., 2011a; Gabriel et al., 2014; Girke et al., 2014; McKay et al., 2005; Olgac et al., 2008; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010; Wada et al., 2002) where the system dynamics are typically governed by a series of convection-diffusion equations, or part of a combined mass flow (Quarteroni et al., 2002).

2.3.3 LDL penetration of the tunica intima

The process through which LDL passes into the *tunica intima* has been modelled at a range of levels. The simplest approaches consider this to be a mathematical function of arterial WSS (Vincent Calvez Nicolas Meunier, Annie Raoult, 2010) or constant (Ai and Vafai, 2006; Cobbold et al., 2002b; Gessaghi et al., 2011; Girke et al., 2014; Ibragimov et al., 2005; McKay et al., 2005). Some simply ignore LDL penetration, instead considering only LDL in the *tunica intima* (Bulelzai and Dubbeldam, 2012) or combining cells, proteins and other macromolecules into one mixed quantity (Prosi et al., 2005). More sophisticated approaches have considered diffusion (Cilla et al., 2014; Friedman and Hao, 2014; Gabriel et al., 2014; Hao and Friedman, 2014; Olgac et al., 2008; Quarteroni et al., 2002; Wada et al., 2002) and have modelled the artery wall as a semi-permeable membrane by utilising Kedem-Katchalsky equations (Filipovic et al., 2011a, 2011b, 2013; Tomaso et al., 2011; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010). LDL penetration appears to be considered as a boundary to many models and the description of its uptake reflects the scope of the model proposed.

2.3.4 LDL oxidation and the role of HDL

A range of approaches have been taken to describe LDL oxidation inside the *tunica intima* and they are coupled to LDL penetration to differing extents. Many studies consider the synthesis and turnover of oxLDL directly (Bulelzai and Dubbeldam, 2012; Calvez and Ebde, 2010; Cilla et al., 2014; Cobbold et al., 2002b; Cohen et al., 2014; Filipovic et al., 2011a, 2013; Friedman and Hao, 2014; Gessaghi et al., 2011; Girke et al., 2014; Hao and Friedman, 2014; Ibragimov et al., 2005; McKay et al., 2005; Ougrinovskaia et al., 2010; Silva et al., 2013; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010). In some, oxidation of LDL is a modelled reaction (Calvez and Ebde, 2010; Cilla et al., 2014; Cobbold et al., 2002b; Filipovic et al., 2011a; Friedman and Hao, 2014; Gessaghi et al., 2011; Girke et al., 2014; Hao and Friedman, 2014; Ibragimov et al., 2005; McKay et al., 2005; Silva et al., 2013; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010) whereas in others it is taken to be a process that is driven by other factors such as monocyte recruitment (Bulelzai and Dubbeldam, 2012) or a constant (Cohen et al., 2014; Ougrinovskaia et al., 2010). Intermediate stages of the oxidation process have been

considered by modelling the number of unoxidized antioxidant molecules attached to each LDL particle (Cobbold et al., 2002b).

The role of HDL has been incorporated into a portion of these studies. In particular, it has been modelled as competing for free radicals and suppressing inflammatory signalling in the *tunica intima* (Cobbold et al., 2002b; Hao and Friedman, 2014; McKay et al., 2005) and as having an atheroprotective effect on foam cells (Cohen et al., 2014; Friedman and Hao, 2014).

Elsewhere, the interplay between LDL, HDL, oxidising free radicals and antioxidant vitamins C and E have been studied (Cobbold et al., 2002b) with predictions of comparable atheroprotective power between HDL and vitamin C.

2.3.5 Monocyte recruitment and chemoattractants

Monocyte recruitment has been modelled as related to shear stress and the rate of LDL penetration (Bulelzai and Dubbeldam, 2012). The existence of monocytes in the lumen has rarely been considered (Cilla et al., 2014), but several studies have modelled the turnover of monocytes in the *tunica intima* (Bulelzai and Dubbeldam, 2012; Cilla et al., 2014; Little et al., 2009; McKay et al., 2005; Zohdi et al., 2004). Elsewhere, the process of monocyte recruitment and differentiation has also been simplified and incorporated into one step governing macrophage turnover, where this is linked to driving factors such as shear stress, diffusion and LDL penetration (Calvez and Ebde, 2010; Cohen et al., 2014; Filipovic et al., 2011b, 2013; Friedman and Hao, 2014; Girke et al., 2014; Hao and Friedman, 2014; Ibragimov et al., 2005; Ougrinovskaia et al., 2010; Silva et al., 2013; Tomaso et al., 2011; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010).

Similarly, the turnover of MCP-1 as a chemoattractant has been described explicitly in some studies (Friedman and Hao, 2014; Hao and Friedman, 2014) and grouped together with other chemoattractants including interleukin-1 (IL-1) and M-CSF in other studies (Cilla et al., 2014; Filipovic et al., 2011b; Girke et al., 2014; Ibragimov et al., 2005; El Khatib et al., 2007; McKay et al., 2005; Silva et al., 2013). One study has shown that exposure to radiation leads to enhanced levels of MCP-1 and is therefore atherogenic (Little et al., 2009). However, in many studies the role of chemoattractants has been ignored.

2.3.6 Monocytes to macrophage differentiation

The differentiation of monocytes to macrophages has been incorporated into a number of studies, although many simplify this step by considering both populations as one group on the grounds that differentiation occurs on a time scale too short to be significant (Filipovic et al., 2011a; Ougrinovskaia et al., 2010; Tomaso et al., 2011). Where differentiation has been modelled it is presented with mass action kinetics (Bulelzai and Dubbeldam, 2012; Cilla et al., 2014; Little et al., 2009; McKay et al., 2005).

2.3.7 Foam cell formation and the phagocytosis of oxidised LDL

The transformation of macrophages to foam cells due to the phagocytosis of oxLDL is a critical stage in the formation of atheroma that has been included in many studies. These are typically modelled as a combination of mass action and Michaelis-Menten terms (Bulelzai and Dubbeldam, 2012; Calvez and Ebde, 2010; Cilla et al., 2014; Cohen et al., 2014; Friedman and Hao, 2014; Hao and Friedman, 2014; Ougrinovskaia et al., 2010; Silva et al., 2013; Tomaso et al., 2011; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010), and in some cases reverse cholesterol efflux is included in the model (Cohen et al., 2014; Friedman and Hao, 2014; Zhang et al., 2013). Many studies, however, omit foam cell formation as a step, instead taking the volume of macrophages recruited to be representative of atheroma formation (Filipovic et al., 2013; Ibragimov et al., 2005; El Khatib et al., 2007; McKay et al., 2005; Zohdi et al., 2004).

2.3.8 T cell recruitment and the role of interferon-gamma (IFN-γ)

The role of T cells in coordinating the inflammatory response has rarely been included in computational studies. Where they have been included as a factor, T cells yield IFN-γ that modulates macrophage differentiation (Friedman and Hao, 2014; Hao and Friedman, 2014; Little et al., 2009; McKay et al., 2005) and are themselves modelled as being activated by interleukin 12 (IL-12) (Friedman and Hao, 2014; Hao and Friedman, 2014), although it has been shown experimentally that T cells can also be activated by IL-1 (Lichtman et al., 1988) and IFN-γ (McLaren and Ramji, 2009).

2.3.9 Proliferation of smooth muscle cells

Along with foam cells and cell debris, SMCs contribute to the formation of atheroma (Weber and Noels, 2011). However this factor has rarely been incorporated into models. Where it

has been incorporated, SMC recruitment occurs in response to MCP-1, platelet derived growth factor (PDGF) and extracellular matrix (ECM) either modelled explicitly as factors (Friedman and Hao, 2014; Hao and Friedman, 2014) or as a generic recruitment process (Cilla et al., 2014; Girke et al., 2014; Ibragimov et al., 2005; McKay et al., 2005). One study in particular has focused on the interplay between SMCs and PDGF identifying bistability in SMC-driven atheroma formation (Fok, 2012).

2.3.10 Plaque rupture and thrombosis

The rupture of atheroma has been modelled, establishing a criterion for atheroma instability that takes the form of a solution to a third order non-linear ODE (Green et al., 2002). Separate studies have established stability by evaluating the eigenvalues of a perturbed system (Ibragimov et al., 2005) and by calculating the mean *time-to-rupture* of atheroma formation (Zohdi et al., 2004). The WSS upon an atheroma has been calculated as a trigger for rupture and this model has been modified to incorporate the effects of abnormal axial G-forces (Deepa et al., 2009). Relevant models have been produced that describe thrombus formation in the absence (Xue et al., 2009) and presence of shear blood flow (Guy et al., 2007; Li et al., 2006).

2.4 Discussion

The majority of the work presented here has been published in the last 10 years, demonstrating that computational modelling of atherosclerosis is a developing field with growing support. These studies operate at a range of levels of abstraction and have variable scope. However, they have all been produced as separate bespoke computational models with little capacity for reuse by the wider modelling community. The introduction of community modelling standards such as SBML (Frank T. Bergmann, 2015; Hucka et al., 2003) and SBGN-ML (Van Iersel et al., 2012; Le Novere et al., 2009) would enable the community to progress together on the development of atherosclerosis modelling and it would be valuable to translate the most biologically detailed models (Cilla et al., 2014; Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010) into these community standards.

Online databases, such as BioModels (Chelliah et al., 2015; Li et al., 2010; Le Novere et al., 2006) and the Physiome Model Repository 2 (PMR2) (Yu et al., 2011), contain computational

models of biological processes. Such databases facilitate the codification of our understanding and, critically, enable models to be reused and built upon as our knowledge advances. However, no models of atherosclerosis currently exist within these repositories, although systems biology representations of the cardiovascular system (Shaw et al., 2007) and statin pharmacokinetics (Bucher et al., 2011) are available.

2.4.1 Factors not yet modelled

There are many components of atherosclerosis that to date have not been modelled. With accurate parameterisation each would increase the comprehensiveness and accuracy of our understanding of atherosclerosis as a dynamical process. Triglyceride rich lipoproteins contribute to plaque build-up with some studies showing that they trigger foam cell formation through mechanisms that bypass LDL oxidation (Gotto, 1998; Le and Walter, 2007; Talayero and Sacks, 2011). Elsewhere, it has been proposed that categories of HDL and their relative proportions may be more important than the absolute abundance of HDL (Douglas Channon, KM., 2014; Umaerus et al., 2012), suggesting that models could be adapted to incorporate a HDL profile that influences oxidation and reverse cholesterol efflux. It has also been shown that HDL can inhibit the recruitment of monocytes and subsequently reduce atherogenesis (Umemoto et al., 2013) suggesting further interactions to model. Clinically, it has been suggested that LDL particle number is a stronger risk factor for atherosclerosis than the abundance of LDL-bound cholesterol, implying that future models should include a description of the cholesterol load of lipoproteins as well as their abundance (Otvos et al., 2011). In addition, the role of neutrophils (Hartwig et al., 2014), nitrous oxide (Douglas Channon, KM., 2014), B cells (Perry et al., 2012), heat shock proteins (Kilic and Mandal, 2012; Xu et al., 2012), sterol regulatory element binding protein (SREBP) mediated regulation (Lu et al., 2015), various cell signalling proteins such as NLRP3 (Xiao et al., 2013), and miRNAs (Nazari-Jahantigh et al., 2014) have not been modelled in this context.

By far the majority of work to date has been on the buildup of atheroma. Some studies have addressed the mechanisms through which atheroma rupture, but they are in a significant minority. Very little work has been done on the consequences of rupture, such as thrombus formation. This presents a potential direction for the field. It is highly relevant to

patient treatment as most patients at risk of CVD are only identified after a cardiovascular event has occurred.

2.4.2 Computational modelling in therapy development

The application of computational modelling to therapy development in atherosclerosis has been historically poor. It is possible to predict both the efficacy of a drug and its potential side effects (Tatonetti et al., 2009; Wang et al., 2013; Yang et al., 2011) and there is growing interest in areas of combinatorial drug design (Sun et al., 2013) to optimise treatment. Such approaches have been demonstrated for the role of statins in the reduction of LDL levels in plasma along with dietary changes [49, 95]. Computational biology can also be used to identify potential molecular targets for drugs and has been used to reduce the high attrition rate of drug discovery (Chua and Roth, 2011). However, these technologies have yet to be exploited to their full potential. Finite element and analytical methods have been employed to model the interaction between a stent and artery wall when widening constricted arteries during angioplasty (Eftaxiopoulos and Atkinson, 2005; Holzapfel et al., 2000).

Creating more comprehensive models of atherosclerosis has the potential to improve the efficiency of therapy development with benefits for both the patient and the commercial vendor. However, obtaining accurate parameterisations for the models is a fundamental challenge. The lack of appropriate published experimental data is a critical obstacle to generating high confidence predictive models.

2.4.3 Difficulties in model generation

Developing a comprehensive predictive model of atherogenesis comes with many challenges. Our knowledge of the processes involved has increased significantly in recent years with the development of genomic technologies such as genome wide association studies (GWAS) (Schunkert et al., 2011). As atherosclerosis is a cardiovascular condition that affects critical circulatory systems, studying human atheroma poses logistical and ethical problems as access to live atherosclerotic tissue is limited and disturbances risk triggering plaque rupture. Consequently, data is limited. Animal studies of atherosclerosis do exist for mouse, rabbit and pig (Getz and Reardon, 2012) and profiling has been conducted for plaque material removed in carotid endarterectomy (Verhoeven et al., 2005).

The limited data available obstructs studies of atheroma at the macro scale and studies of the molecular biology involved. As a result, establishing biologically relevant kinetic parameters that can be used to simulate pathway dynamics will be challenging and comprehensive data for the pathogenesis of atherosclerosis is not available in the public domain. Subsequently, some studies have resorted to estimating parameters for models, based on expert opinion or inferred from other cell processes.

It is likely that approximate values can be obtained for a number of the parameters required by using recombinant proteins and *in vitro* studies. However, recreating the environment of the *tunica intima* and quantifying its impact on the parameterisation in order to obtain physiological values will be challenging(Rollins et al., 1991; Santoli et al., 1987; Yang et al., 1995).

2.4.4 Conclusion

Computational modelling of atherosclerosis presents an opportunity to contribute to the reduction of the global burden of CVD. By introducing accurate and quantitative models of atherosclerosis, we can create an *in silico* experimental system with the potential not only to displace *in vivo* experimentation but also to enable us to study details that cannot be measured *in vivo*. However this necessitates a physiologically accurate parameterisation and such data is not currently available in a comprehensive form.

Historically, little work has been completed developing computational modelling of atherosclerosis, although recent years have seen a clear growth of interest and the formation of a nascent field. Here we have gathered together and reviewed the recent results with a view to identifying where the gaps in our understanding lie and where progress can be made.

Most of the work completed in this area to date has focussed on the inflammatory response and shear stress of the artery wall and has involved modelling at a range of levels of abstraction.

The majority of work has focused on describing atheroma formation and few studies have addressed the mechanics of plaque rupture and its subsequent consequences. In most cases, models follow the canonical understanding of atherosclerosis: LDL penetration and

oxidation, monocyte recruitment and differentiation and foam cell formation. However, many additional factors remain outside this canonical picture that are known to contribute to atherosclerosis and there currently exist opportunities to explore their role in the dynamics of this disease through computational modelling.

2.5 Acknowledgements

We would like to acknowledge and thank Eliza Yankova (University of Ulster) for drawing Figure 1 (The pathophysiology of atherosclerosis) that we have included in this paper.

2.6 List of Computational Models

In Table 2.1, the mathematical models of atherosclerosis referenced within this review are summarized. These models are reproducible as their governing equations are explained unless otherwise stated.

Table 2.1: Mathematical Models of Atherosclerosis

First Author	Title	Year	Form	Parameters	Validation	Tools
Goriely (Goriely and Vandiver, 2010)	On the mechanical stability of growing arteries	2010	A coronary artery modelled as an incompressible 2-layer cylindrical structure was used to study the arterial response to stress	Related to experimental data	Compared to experimental data obtained by Schulze-Bauer et al. (Schulze-Bauer et al., 2003)	None mentioned
Li (Li et al., 2006)	How critical is fibrous cap thickness to carotid plaque stability? A flow-plaque interaction model	2006	A model of a stenotic carotid artery was used to relate fibrous cap thickness to WSS	Use a combination of estimated and experimentally validated parameters	The authors claim that the model fits well within the current literature, however no references are given to substantiate this claim	FEMLAB was used for model construction, SPSS was used to analyse this model
Stroud (Stroud et al., 2002)	Numerical analysis of flow through a severly stenotic carotid artery bifurcation	2002	A model of a carotid artery bifurcation is used to study pulsatile and steady blood flow	Related to experimental data	Compared to experimental data obtained by Ahmed and Giddens (Ahmed and Giddens, 1983)	None mentioned
Quarteroni (Quarteroni et	Mathematical and numerical modeling of	2002	Proposed two models of an arterial bifurcation to study	Parameter source unclear	None	None mentioned

First Author	Title	Year	Form	Parameters	Validation	Tools
al., 2002)	solute dynamics in blood flow and arterial walls		mass transfer			
Di Tomaso (Tomaso et al., 2011)	A Multiscale Model of Atherosclerotic Plaque Formation at Its Early Stage	2011	Built a multi-scale model of atherosclerosis to include mass transfer, LDL oxidation and foam cell formation	Use a combination of estimated and experimentally validated parameters	The model was compared with experimental data taken from Meyer *et al.* (Meyer et al., 1996)and against the model produced by Olgac *et al.* (Olgac et al., 2008)	None mentioned
Cilla (Cilla et al., 2014)	Mathematical modelling of atheroma plaque formation and development in coronary arteries	2013	Uses a standard left descending coronary artery model to study plaque growth	Taken from experimental data and other mathematical models	Parts of the model correspond with experimental data such as Meyer *et al.* (Meyer et al., 1996), however appropriate experimental data to cover the entire model is not currently available.	COMSOL Multiphysics
Filipovic (Filipovic and Kojic, 2004)	Computer simulations of blood flow with mass transport through the carotid artery bifurcation	2004	Proposed a simulation of mass transport to allow physicians to study individual patients	Parameter source unclear	The authors claim that the model fits well within the current literature, however no references are given to substantiate this	None mentioned

First Author	Title	Year	Form	Parameters	Validation	Tools
					claim	
Filipovic (Filipovic et al., 2011a)	ARTreat Project: Three-Dimensional Numerical Simulation of Plaque Formation and Development in the Arteries	2012	Presented a 3D model of plaque formation and development	Parameters were experimentally established through a rabbit animal model	Plaque progression within the model has been validated against experimental data taken from Boussel *et al.*(Boussel et al., 2009)	None mentioned
Johnston (Johnston et al., 2006)	Non-Newtonian blood flow in human right coronary arteries: Transient simulations	2005	Used right coronary artery models to study pulsatile blood flow	Experimentally observed	Findings were validated against experimental data taken from Kirpalani (Kirpalani et al., 1999) & Myers (Myers et al., 2001)	CFD-ACE
Liu (Liu and Tang, 2010)	Computer Simulations of Atherosclerotic Plaque Growth in Coronary Arteries	2010	Uses model of a stenosis-free curved human coronary artery to study plaque growth	Experimentally observed	None	COMSOL Multiphysics
Olgac (Olgac et al., 2008)	Computational modeling of coupled blood-wall mass transport of LDL: effects of local wall shear stress	2008	Developed a model of a stenosed coronary artery to study the effects of WSS on mass transport	Experimentally observed	Related to experimental data (Huang et al., 1997; Meyer et al., 1996; Yuan et al., 1991)	COMSOL Multiphysics
Rappitsch (Rappitsch et al., 1997)	Numerical Modelling of Shear-Dependent Mass Transfer in Large Arteries	1997	Used a curved-tube-artery model to study blood flow and lipoprotein transport processes	Use a combination of estimated and experimentally validated parameters	Validated against (Friedman and Ehrlich, 1975)	None mentioned

First Author	Title	Year	Form	Parameters	Validation	Tools
Sun (Sun et al., 2006)	Fluid-wall modelling of mass transfer in an axisymmetric Stenosis: Effects of shear-dependent transport properties	2006	Studies the influence of WSS on mass transport	Use a combination of estimated and experimentally validated parameters	Compared to experimental data taken from rabbit aortic walls (Meyer et al., 1996)	None mentioned
Ai (Ai and Vafai, 2006)	A coupling model for macromolecule transport in a stenosed arterial wall	2006	A model of a stenosed artery is used to study lipid transfer	Experimentally validated	Compared to other mathematical models, with arguments as to why their parameter set is more accurate	FIDAP
Wada (Wada et al., 2002)	Theoretical study of the effect of local flow disturbances on the concentration of low-density lipoproteins at the luminal surface of end-to-end anastomosed vessels.	2002	Femoral artery model is used to study the relationship between intimal thickness and the endothelial surface level of LDL	Parameters were taken from experimental data or estimated	Compared to experimental data taken from Ishibashi *et al.* (Ishibashi et al., 1995)	Star LT
Calvez (Calvez and Ebde, 2010)	Mathematical modelling of the atherosclerotic plaque formation.	2009	Developed a 2D geometry modelling the carotid artery to demonstrate plaque formation, based on the model of El Khatib *et al.* (El Khatib et al., 2007)	Parameters are taken from other mathematical models, relating to atherosclerosis (El Khatib et al., 2007) and hyperplasia (Budu-Grajdeanu et al., 2008)	None	FreeFem++ (Hecht, 2012)
Calvez (Vincent	Mathematical and numerical modelling of	2010	Expanded on their previous model (Calvez and Ebde, 2010)	Parameters were taken from	Experiments published by Cheng	FreeFem++ (Hecht, 2012)

First Author	Title	Year	Form	Parameters	Validation	Tools
Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010)	early atherosclerotic lesions		to include a model of lesion growth	experimental data or estimated	*et al.* (Cheng et al., 2005, 2006) were reproduced and were used to validate the model	
Bosnić (Bosnić et al., 2012)	Mining data from hemodynamic simulations for generating prediction and explanation models.	2012	Built a prototype of a system that could predict locations of increased WSS from artery models	Parameter source unclear	Presents a series of methods to estimate accuracy of the model, and relates these to experimental data	None mentioned
Bulelzai (Bulelzai and Dubbeldam, 2012)	Long time evolution of atherosclerotic plaques	2011	Present a series of ODEs for the concentrations of particular elements of atheromae.	Taken from experimental data and other mathematical models (McKay et al., 2005)	Compared to mathematical model of Zohdi *et al.* (Zohdi et al., 2004)	MATCONT (Dhooge et al., 2003)
Gabriel (Gabriel et al., 2014)	Deposition-driven Growth in Atherosclerosis Modelling.	2014	A simplified bifurcating artery is used to model LDL flux into the intima	Taken from experimental data	None	ANSYS Fluent
Silva (Silva et al., 2013)	Mathematical Modeling of Atherosclerotic Plaque Formation Coupled with a Non-Newtonian Model of blood Flow	2013	Built a 2D carotid artery bifurcation to study plaque formation with a non-Newtonian model of blood flow	Taken from other mathematical models (Chen and Lu, 2004)	None	COMSOL Multiphysics
Gessaghi (Gessaghi et al., 2011)	Growth model for cholesterol accumulation in the wall of a simplified 3D	2011	A 3D model of a carotid artery bifurcation is used to study the influx, efflux, oxidation and phagocytosis of LDL	Taken from experimental data	Compared with data obtained from Yang *et al.*(Yang and Vafai,	OpenFOAM (Weller and Tabor, 1998), Netgen

First Author	Title	Year	Form	Parameters	Validation	Tools
	geometry of the carotid bifurcation				2006).However, authors comment that not enough experimental data exists to fully validate the model.	
Green (Green et al., 2002)	Atherosclerotic plaque rupture	2002	A model of a straight, stenotic 2D artery is used to study atherosclerotic plaque rupture.	Parameter source unclear	None	AUTO
Deepa (Deepa et al., 2009)	Modelling Blood Flow and Analysis of Atherosclerotic Plaque Rupture under G-Force	2009	A 1D arterial model was used to study the rupture of plaques under g-force	Sources have not been cited for parameter values	None	MATLAB
Girke (Girke et al., 2014)	Efficient Parallel Simulation of Atherosclerotic Plaque Formation Using Higher Order Discontinuous Galerkin Schemes	2014	Girke *et al.* built a mathematical model based on the works of Ibragimov *et al.* (Ibragimov et al., 2005) and Calvez *et al.* (Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, 2010) to demonstrate the use of the compact discontinuous galerkin method (CDG2) in discretizing relevant equations	Taken from experimental data	None	DUNE-FEM (Dedner et al., 2010)
McKay (McKay et al., 2005)	Towards a Model of Atherosclerosis	2005	Proposed a mathematical mode to cover mass transfer, oxidation, immune cell activation and plaque growth	Taken from other mathematical models, or estimated by domain experts	None	None mentioned
Ibragimov (Ibragimov et al., 2005)	A mathematical model of atherogenesis as an inflammatory response	2005	Created a series of ODEs to study the concentrations of cell groups over time	Primarily estimated due to lack of relevant data	None	FEMLAB

First Author	Title	Year	Form	Parameters	Validation	Tools
Cobbold (Cobbold et al., 2002b)	Lipoprotein Oxidation and its Significance for Atherosclerosis: a Mathematical Approach	2002	Built a series of ODEs to study lipoprotein oxidation	Taken from experimental data	Compared to an experiment performed by Neužil *et al.* (Neužil et al., 1996)	None mentioned
Prosi (Prosi et al., 2005)	Mathematical and numerical models for transfer of low density lipoproteins through the arterial walls: a new methodology for the model set up with applications to the study of disturbed lumenal flow	2004	Built multiple models of lipoprotein transfer in order to maximse the accuracy of their prediction	Taken from experimental data	Experimentally validated against Meyer *et al.* (Meyer et al., 1996)	None mentioned
Friedman (Friedman and Hao, 2014)	A Mathematical Model of Atherosclerosis with Reverse Cholesterol Transport and Associated Risk Factors	2014	Expands on the previous model by the same group (Hao and Friedman, 2014) to include reverse cholesterol transport	Taken from experimental data and from other mathematical models (Hao and Friedman, 2014)	Validated qualitatively against experimental data (e.g. (Feig et al., 2011; Lovren et al., 2012; Schiopu et al., 2007))	None mentioned
Hao (Hao and Friedman, 2014)	The LDL-HDL profile determines the risk of atherosclerosis: a mathematical model	2014	Developed a series of PDEs to model the concentration of a series of cells and macromolecules contained within an atheroma, and related this information to plaque growth	Taken from experimental data or estimated	None	MATLAB
Filipovic (Filipovic et	Experimental and computer model of	2011	Built a model of plaque formation based on a pig left	Taken from experimental data,	Reproduced an experiment by	None explained

First Author	Title	Year	Form	Parameters	Validation	Tools
al., 2011b)	plaque formation in the artery		anterior descending coronary artery	or estimated where data was unavailable.	Cheng *et al.* (Cheng et al., 2006) and compared the results to their model of plaque formation	
Ougrinovskaia (Ougrinovskaia et al., 2010)	An ODE model of early stages of atherosclerosis: mechanisms of the inflammatory response	2010	Developed a series of ODEs to model mass transfer and foam cell formation	Estimated	Behaviour relates to qualitative data, but model has not been compared to quantitative data	MATLAB, XPPAUTO
Cohen (Cohen et al., 2014)	Athero-protective effects of High Density Lipoproteins (HDL): An ODE model of the early stages of atherosclerosis	2014	Expanded on their previous model (Ougrinovskaia et al., 2010) to include HDL and reverse cholesterol transport	Taken from experimental data	Noted that the behaviour of their model corresponds with an experiment performed by Feig *et al.* (Feig et al., 2011)	None mentioned
Zohdi (Zohdi et al., 2004)	A phenomenological model for atherosclerotic plaque growth and rupture	2004	Built a series of equations to study plaque growth and lesion rupture	Taken from experimental data, or estimated where data was unavailable.	None	None mentioned
Little (Little et al., 2009)	A model of cardiovascular disease giving a plausible mechanism for the effect of fractionated low-dose ionizing radiation exposure	2009	Built a series of equations to study the effect of small radiation doses to atherosclerosis and CVD	Taken from experimental data	Sections of this model are validated by matching with experimental data published by Cushing *et al.* (Cushing et al., 1990) and Shi *et al.*	None mentioned

First Author	Title	Year	Form	Parameters	Validation	Tools
					(Shi et al., 2005)	
El Khatib (El Khatib et al., 2007)	Atherosclerosis Initiation Modeled as an Inflammatory Process	2007	Built a series of reaction-diffusion equations by grouping together all cytokines and immune cells involved	Estimated	None	COMSOL Multiphysics
Zhang (Zhang et al., 2013)	Foam cell formation in atherosclerosis: HDL and macrophage reverse cholesterol transport	2013	Expanded on the model of Ibragimov *et al.* (Ibragimov et al., 2005) by focusing on the role of HDL and reverse cholesterol transport	Taken from experimental data and from other mathematical models (Cobbold et al., 2002b)	None	None mentioned
Fok (Fok, 2012)	Mathematical model of intimal thickening in atherosclerosis: vessel stenosis as a free boundary problem	2012	Focuses on SMC migration and the role of PDGF	Taken from experimental data	Compared to experimental data taken from New Zealand white rabbits (Stadius et al., 1992)	None mentioned
Xue (Xue et al., 2009)	A mathematical model of ischemic cutaneous wounds	2009	Xue *et al.* developed a series of PDEs to model ischemic dermal wounds	A combination of experimentally validated and estimated parameters are used	Compared to experimental data established by Roy *et al.* (Roy et al., 2009)	Livermore Solver
Guy (Guy et al., 2007)	Fibrin gel formation in a shear flow	2007	Presents a model of fibrin formation in a damaged blood vessel	A combination of experimentally validated and estimated parameters are used	None	None mentioned

2.7 Recent Updates

Since August 2015, the date of submission of the literature review 'Computational Modelling in Atherosclerosis', modelling within atherosclerosis has continued to improve. Multiple models detailing the effect of blood flow dynamics on atherosclerosis development, arterial remodelling, lesion size and thrombus growth have been developed (Cuomo et al., 2017; Filipovic et al., 2017; Mehrabadi et al., 2016). Models designed to predict lesion location and rate of thrombus growth from initial arterial geometry and arterial wall shear stress, validated using *in vitro* experiments across a variety of different blood vessels, providing insight into the consequences of arterial occlusion (Mehrabadi et al., 2016). Filipovic et al have provided updates to their earlier work utilising Navier-Stokes equations and Kadem-Katchalsky equations to represent fluid dynamics and the environmental changes between the artery wall and lumen, demonstrating methods which predict atheroma location and size from initial arterial geometry (Filipovic et al., 2017). A study into non-Newtonian oscillating flow in biomechanical analysis within a human aorta has been undertaken, concluding that non-Newtonian methods of modelling blood flow can be more accurate in capturing molecular viscosity (Soulis et al., 2016).

Age-related effects of cardiovascular disease on arterial stiffness have been modelled through a collection of blood-flow models on arterial segments considered as an incompressible elastic membrane where external tissue stability support is modelled as a boundary condition, limiting oscillations within the blood vessel wall (Cuomo et al., 2017). The influence of aging on cholesterol metabolism, including reverse cholesterol transport, has been modelled using SBGN-ML open standards, encouraging model adaptation and revision as knowledge into these processes improves (Morgan et al., 2016). LDL composition and lipoprotein metabolism has been modelled (including different lipoprotein forms, apolipoprotein concentrations and cholesteryl ester transfer protein (CETP) mechanisms), leading to the prediction of lipoprotein metabolism rate constants (Jansen et al., 2016).

Lipoprotein transfer into the *tunica intima* has been considered by multiple groups (Filipovic et al., 2017; Iasiello et al., 2016; Mel'nyk, 2017; Mpairaktaris et al., 2017). The role of adhesion molecules within LDL migration has been considered (Mel'nyk, 2017), improving upon previous models which solely considered diffusion when modelling this process (Hao

and Friedman, 2014). Lipoprotein transfer under hyperthermia conditions has been considered (Iasiello et al., 2016).

Different subtypes of macrophages, M_1 and M_2 (or classically and alternatively activated), have been considered within models which consider LDL transfer, monocyte recruitment and inflammatory mechanisms (Mel'nyk, 2017). Foam cell formation through phagocytosis has been considered in the context of oxLDL consumption by pro-inflammatory macrophages (Mel'nyk, 2017) and a function of shear forces within the blood (Filipovic et al., 2017).

T-Cell recruitment and their subsequent secretion of cytokines have been considered in an ODE model of abdominal aortic aneurysm (Hao et al., 2017), however all T-Cells are assumed to be Th1 cells - individual T-Cell subtypes have not been considered within this model. SMC density within the *tunica media* is also considered as part of this ODE model, however proliferation into the intima and fibrous cap formation are not considered here (Hao et al., 2017).

2.7.1 The future of atherosclerosis modelling

The complexity of atheroma formation and the sheer quantity of proteins, cell types, environments and biological processes involved in atherogenesis is a distinct challenge in computational modelling. As such, many aspects of atherosclerosis have not yet been modelled to a sufficient quality to allow for a clear and complete picture of atherosclerosis-related cardiovascular disease. Models exist to represent cholesterol metabolism, lipoprotein metabolism, diet, haemodynamics and plaque formation, and a combination of each of these processes is required to provide a complete representation of the underlying biology. HDL mechanisms are enigmatic and unclear, and computational modelling of known processes involving HDL could lead to clarification. Plaque calcification and its effects on stability and coronary events have not yet been modelled. Recent studies have been undertaken connecting both telomere length (Fernández-Alvira et al., 2016) and the ratio of platelets to lymphocytes (Akboga et al., 2015) to atherosclerosis.

Table 2.2: Recently Published Models of Atherosclerosis

First Author	Title	Year	Form	Parameters	Validation	Tools
Cuomo (Cuomo et al., 2017)	Effects of age-associated regional changes in aortic stiffness on human hemodynamics revealed by computational modeling	2017	Fluid-to-solid-interactions are modelled through haemodynamic studies to predict spatial changes in arterial structure	Parameterisation taken from literature, initial conditions taken from magnetic resonance images (MRI) from healthy volunteers.	Spatial distribution compared to experimental data	CRIMSON (Whiting and Jansen, 2001)
Filipovic (Filipovic et al., 2017)	Computational modeling of plaque development in the coronary arteries	2017	Develops methods to study plaque development in both left and right coronary arteries	Taken from experimental data, or estimated where data was unavailable.	None explained	None explained
Mehrabadi (Mehrabadi et al., 2016)	A Predictive Model of High Shear Thrombus Growth	2016	Predicts location and size of thrombus deposit from haemodynamic conditions and arterial geometry	Calculated form correlation data between thrombus growth and shear rate.	Results are compared to experimental data about in (Casa and Ku, 2014)	Matlab

First Author	Title	Year	Form	Parameters	Validation	Tools
Soulis (Soulis et al., 2016)	Low Density Lipoprotein and Non-Newtonian Oscillating Flow Biomechanical Parameters for Normal Human Aorta	2016	Four non-Newtonian models and a Newtonian model were created and tested for a typical human aorta exposed to oscillating blood flow	Taken from experimental data	LDL Transfer and spatial configurations were compared to experimental data	ANSYS Workbench 13
Morgan (Morgan et al., 2016)	Mathematically modelling the dynamics of cholesterol metabolism and ageing	2016	SBML model of cholesterol metabolism	Taken from experimental data	Model outputs were compared to clinical data taken from a feeding study	Copasi
Jansen (Jansen et al., 2016)	In silico modeling of the dynamics of low density lipoprotein composition via a single plasma sample	2016	LDL composition and lipoprotein metabolism are modelled	Some parameters are calculated from human plasma samples, others are inferred from the data	Validated against a second group of human plasma samples	SPSS
Mel'nyk (Mel'nyk,	A mathematical model of the atherosclerosis	2017	A model of atherogenesis within small blood	Taken from experimental studies and estimated	Boundary conditions are	None mentioned

First Author	Title	Year	Form	Parameters	Validation	Tools
2017)	development in thins blood vessels and its asymptotic approximation		vessels has been developed, including pro and anti-inflammatory macrophages		justified, however model results are not	
Hao (Hao et al., 2017)	A mathematical model of aortic aneurysm formation	2017	Hao *et al.* have built on their previous work and developed an ODE model of abdominal aortic aneurysm	Taken from experimental data and other mathematical models	None	None mentioned
Iasiello (Iasiello et al., 2016)	Low-density lipoprotein transport through an arterial wall under hyperthermia and hypertension conditions – An analytical solution	2016	Four layers of artery wall are considered and transfer of LDL and temperature are modelled under hypertension and hyperthermia conditions	Taken from experimental data where possible	Compared to previous analytical studies	None mentioned
Mpairaktaris (Mpairaktaris et al., 2017)	Low density lipoprotein transport through patient-specific	2017	Models patient specific LDL distribution within the artery wall	A combination of experimentally derived, mathematically justified	Results are compared to experimental data	Materialize Mimics

First Author	Title	Year	Form	Parameters	Validation	Tools
	thoracic arterial wall			and estimated parameters are used in this system	from (Nematollahi et al., 2012)	

Chapter 3:

A Computational Model of Atherosclerosis: Development of a Community Resource

The following chapter has been submitted to:

Circulation - http://circ.ahajournals.org/

Full Title

Computational modelling of atherosclerosis: developing a community resource.

Authors

Andrew Parton, Northern Ireland Centre for Stratified Medicine, Ulster University

Victoria McGilligan, Northern Ireland Centre for Stratified Medicine, Ulster University

Maurice O'Kane, Western Health and Social Care Trust, Altnagelvin Hospital; Northern Ireland Centre for Stratified Medicine, Ulster University

Steven Watterson, Northern Ireland Centre for Stratified Medicine, Ulster University

Author contribution

This work was conceived by AP and SW. The analysis and data compilation was undertaken by AP and SW. The manuscript was written by AP, SW, VM and MO'K.

3.1 Introduction

Cardiovascular disease (CVD) is the primary cause of global mortality. CVD is estimated to account for 17m deaths worldwide each year (Go et al., 2014), representing 31% of all-cause mortality worldwide and 47% of all-cause mortality within Europe (Nichols M Luengo-Fernandez R, Leal J, Gray A, Scarborough P, Rayner M, 2012). Such a prevalent condition incurs a significant financial burden, accounting for 17% of all healthcare expenditure in the USA (Heidenreich et al., 2011). Age is a significant risk factor for cardiovascular disease and with an aging population, the cost of CVD related therapies is predicted to almost triple in the USA from $273 billion in 2010 to $818 billion by 2030 (Heidenreich et al., 2011).

Atherosclerosis is estimated to account for 71% of CVD diagnoses (Nichols M Luengo-Fernandez R, Leal J, Gray A, Scarborough P, Rayner M, 2012). It is characterised by the hardening of an artery wall, and the formation of a fibrous-fatty lesion within the intimal layer. As the disorder progresses, thick extracellular cores of lipids occur within the artery wall, occluding the artery and subsequently reducing blood flow. Thrombosis can further occlude the artery either as a result of plaque rupture or turbulent blood flow induced around the site of the atheroma.

Atherosclerosis is understood to be a chronic inflammatory condition facilitated by a high blood lipid profile that involves arterial damage, low-density lipoprotein (LDL) and high-density lipoproteins (HDL) penetration and oxidation, monocyte and T-cell recruitment, monocyte differentiation and foam cell creation, smooth muscle cell (SMC) migration and fibrous cap formation and plaque rupture. Despite our increasing knowledge of the mechanisms driving this disorder, the pathogenesis of atherosclerosis is still not fully understood. In part, this is due to the significant challenge inherent in studying live, dynamic plaques. Accessing plaques *in vivo* is logistically difficult, necessitating catheterization, and ethically challenging as it can increase the risk of plaque rupture. As a result, alternative approaches to studying atherosclerosis dynamics are needed. Computational modelling has the potential to be especially valuable here due to its flexibility, low financial and ethical cost, consistency and ease of replication. However, as

yet there are no computational or mathematical models of atherosclerosis that are easily available to the research community for use in exploratory studies.

In previous modelling studies the majority of work has focused on plaque initiation and haemodynamics (Parton et al., 2015), where Navier-Stokes dynamics have described blood flow and wall shear stress has been calculated as an pro-atherogenic output (Tomaso et al., 2011), although plaque rupture and thrombosis have not gone unaddressed (Guy et al., 2007). We are interested in the molecular and cellular biology that mediate plaque formation and can furnish targets for therapeutic interventions. However, in previous studies these details have been routinely omitted or simplified for reasons of mathematical expediency (Bulelzai and Dubbeldam, 2012; Friedman and Hao, 2014). Furthermore, the resulting models have not been made publicly available. Reusing this work would necessitate reconstruction of the models in their entirety, a complex, time consuming and error-prone task. At the present time, the European Bioinformatics Institute (EBI) BioModels database (Novère, 2006), a publically available repository for mathematical models of biological processes, contains only one model pertaining to atheroma formation, focussing on lipoprotein action and B-cell signaling with little detail on the mechanisms of plaque formation (Gomez-Cabrero et al., 2011). KEGG (Kanehisa et al., 2017), Reactome (Fabregat et al., 2016), SMPDB (Frolkis et al., 2009), Wikipathways (Kutmon et al., 2016) contain no molecular biology maps of atherosclerosis.

Here we develop the first detailed, predictive dynamical computational model of atherogenesis using Systems Biology standards. The model is described as a map using the Systems Biology Graphical Notation (SBGN) (Le Novere et al., 2009) and is made available to the research community for reuse and refinement using the Systems Biology Graphical Notation Markup Language (SBGN-ML) (Van Iersel et al., 2012). This map is accompanied by a mathematical model describing the dynamics of the interactions in the map as a system of ordinary differential equations (ODEs) and is made available using the Systems Biology

Markup Language (SBML) (Hucka et al., 2003). Currently, there are many examples of SBGN[a] and SBML[b] compliant software.

The only clinically approved therapy for atherosclerosis is angioplasty (Cooper et al., 2014), although there is some evidence to suggest that intense statin treatment (Lima et al., 2004), combined statin-PCSK9 inhibitor treatment (SJ et al., 2016) or Cyclodextrin treatment (Zimmer et al., 2016) may yield a modest plaque reduction. New drug combinations that yield a substantial reduction in plaque size could have a dramatic impact on CVD morbidity and mortality and so their identification has high strategic importance. Here, we employ the model to develop effective therapeutic hypotheses comprising multi-drug combinations through a process of computational optimisation.

3.2 Methods

A list of the cell types involved in atherosclerosis was compiled from descriptions of atherosclerosis in the existing literature (see Appendix 3). Each article was also searched for references to proteins and small molecules with each entity found considered for the model. A protein or small molecule was incorporated into the model if its biological source, presence within a relevant compartment and its influence on atherogenesis (however minor) were all described. The model was assembled with CellDesigner (Funahashi et al., 2008) using the SBGN schema and with mass action and Michaelis-Menten equations primarily used to describe the dynamics. The resulting model was exported to SBGN-ML file format to disseminate the visual map and to SBML file format to disseminate the mathematical model describing the dynamics.

PubMed and Google Scholar searches were undertaken to find studies describing representative concentrations of the cells, proteins and small molecules. The BRENDA enzyme database was searched for relevant known rate parameters (Placzek et al., 2017). Values for unknown parameters were calculated by constraining the model to show dynamics in agreement with published CVD studies. We considered dynamics for three lipid profiles and, as a chronic condition, we considered plaque formation across a

[a] http://sbgn.github.io/sbgn/software_support
[b] http://sbml.org/SBML_Software_Guide

representative time scale of 80 years. High risk, medium-risk and low-risk lipid profiles comprised LDL concentrations of 190 mg/dl[c], 110 mg/dl[c] and 50mg/dl (O'Keefe et al., 2004), respectively and HDL concentrations of 40 mg/dl, 50 mg/dl and 50 mg/dl, respectively (Boden, 2000).

There are be between 5 and 800 cells within a plaque area per high powered field (HPF) at 400x magnification (Brandl et al., 1997), where one HPF displays approximately 0.2 mm^2 of plaque area (Bonanno et al., 2000). We estimate that a plaque contains between 25 and 4000 cells per mm^2. Average plaque area has been shown to be 15.2 mm^2 (von Birgelen et al., 1998), giving the number of cells in a plaque as being between 380 and 60800. With this, we identified the following constraints from the published literature.

3.2.1 Model Constraints

I) Smooth muscle cells comprise 35.10% of the cellular composition of plaques (Bonanno et al., 2000), corresponding to a range of 133 cells which we take to be representative of low LDL profiles to 21341 cells which we take to be representative of high LDL profiles.

II) Macrophages (including foam cells) comprise 34.07% of the cellular composition of plaques (Bonanno et al., 2000), corresponding to a range of 129 cells to 20715 cells.

III) T Cells comprise 30.82% of the cellular composition of plaques (Bonanno et al., 2000), corresponding to a range of 117 cells to 18739 cells.

IV) The ratio of T_h1 to non-T_h1 cells in a plaque is approximately 0.3 (van Dijk et al., 2015), corresponding to a range of 88 T_h1 cells to 14031 T_h1 cells.

V) Blood serum concentrations of MCP1/CCL2 were estimated from myocardial infarction and ischemic stroke patients, ranging from 100 pg/ml to 775 pg/ml (Arakelyan et al., 2005).

VI) Blood serum concentrations of CXCL9 were estimated from patients assessed for coronary artery calcium deposits, ranging from 17.4 pg/ml to 271.2 pg/ml (Yu et al., 2015a).

VII) Blood serum concentrations of CXCL10 were estimated from patients assessed for coronary artery disease, ranging from 127.6 pg/ml to 956.5 pg/ml (Tavakolian Ferdousie et al., 2017).

VIII) Blood serum concentrations of CXCL11 were estimated from control groups in transplantation studies, ranging from 420 pg/ml to 1062 pg/ml (Kao et al., 2003).

---

[c] https://www.nhlbi.nih.gov/health/resources/heart/heart-cholesterol-hbc-what-html

IX) Blood serum concentrations of triglycerides were estimated from control and hyperlipidemic patients, corresponding to 58 mg/dl and 1005 mg/dl, respectively (Sakai et al., 2003).

X) Blood serum concentrations of chylomicrons were estimated from control and hyperlipidemic patients, corresponding to 1.4 µg/ml and 52.6 µg/ml, respectively (Sakai et al., 2003).

XI) Blood serum concentrations of IL1β were estimated from congestive heart failure and control patients, ranging from 0.28 pg/ml to 2.12 pg/ml (Di Iorio et al., 2003).

XII) Blood serum concentrations of CCL5 were estimated from control and coronary event patients, ranging from 2.7 ng/ml to 176.0 ng/ml, respectively (Herder et al., 2012).

XIII) Plaque concentrations of MMP1 were estimated from carotid endarterectomy patients, ranging from 18 ng/g to 104 ng/g wet weight of plaque (Molloy et al., 2004).

XIV) Plaque concentrations of MMP9 were estimated from carotid endarterectomy patients, ranging from 121 ng/g to 722 ng/g wet weight of plaque (Molloy et al., 2004).

XV) Plaque concentrations of TIMP1 were estimated from carotid endarterectomy patients, ranging from 5.3 µg/g to 12.4 µg/g wet weight of plaque (Molloy et al., 2004).

XVI) Plaque concentrations of IL1β were estimated from carotid endarterectomy patients, ranging from 12 ng/g to 24 ng/g wet weight of plaque (Molloy et al., 2004).

XVII) Plaque concentrations of IL6 were estimated from carotid endarterectomy patients, ranging from 1.5 µg/g to 5.1 µg/g wet weight of plaque (Molloy et al., 2004).

XVIII) Plaque concentrations of TNFα were estimated from carotid endarterectomy patients, ranging from 15 ng/g to 27 ng/g wet weight of plaque (Molloy et al., 2004).

XIX) Plaque concentrations of IL2 were estimated from acute coronary syndrome patients, giving 24.0 pg/mg of protein (Ragino et al., 2012).

XX) Plaque concentrations of IL18 were estimated from acute coronary syndrome patients, giving 10.7 pg/mg of protein (Ragino et al., 2012).

XXI) Plaque concentrations of IL10 were estimated from arterial occlusion patients, ranging from 1.51 pg/mg to 2.29 pg/mg wet weight of plaque (Stein et al., 2008).

XXII) Plaque concentrations of IL12 were estimated from arterial occlusion patients, ranging from 3.6 pg/mg to 4.6 pg/mg wet weight of plaque (Stein et al., 2008).

XXIII) Plaque concentrations of IFNγ were estimated from carotid endarterectomy patients, ranging from 20 pg/g to 182 pg/g wet weight of plaque (Grufman et al., 2014).

XXIV) Plaque concentrations of TGFβ were estimated from control and coronary artery disease patients, ranging from 0.33 mg/g to 0.76 mg/g of protein (Herder et al., 2012).

XXV) Plaque density ratios of chymase to tryptase were recorded to be 107.8:135.1 in plaques (Ramalho et al., 2013).

XXVI) Plaque concentrations of elastin were estimated from acute coronary syndrome patients, giving 1.58 mg/g wet weight of plaque (Gonçalves et al., 2003).

XXVII) Plaque concentrations of collagen were estimated from acute coronary syndrome patients, giving 6.26 mg/g wet weight of plaque (Gonçalves et al., 2003).

XXVIII) Plaque concentrations of PDGF were estimated from carotid endarterectomy patients, ranging from 279 pg/g to 1381 pg/g wet weight of plaque (Grufman et al., 2014).

XXIX) The ratio of T_h1 to T_h2 cells has been shown to correlate with atherogenesis (Szodoray et al., 2006).

XXX) Animal models with advanced atherosclerosis have shown plaque reduction mediated by reverse cholesterol after a reduction in lipid profile (Trogan et al., 2006).

XXXI) The weight of oxidized LDL per weight of ApoB has been measured to be 19.6 ng/μg in macrophage rich plaques and 1.9 ng/μg in normal intimal tissue (Nishi et al., 2002). The plaque concentration of ApoB has been measured to range from 1.97 μg/mg to 0.13 μg/mg (Hoff et al., 1978), yielding upper and lower estimates for the oxidised LDL concentrations of 38.6 μg/g and 0.25 μg/g.

Once the model was constructed, the values of unknown parameters were optimised, so that the model recreated these results as far as possible.

3.2.2 Developing multi-drug plaque regression therapeutic hypotheses

In order to demonstrate the utility of the resulting model, we undertook to identify an optimal multi-drug intervention hypothesis that would reprogram the dynamics of the model leading to regression of advanced plaques. It has been demonstrated that multi-drug approaches have the potential to exploit compound effects to yield effective interventions at lower individual and collective dosages than in comparable single-drug interventions, reducing the risk from pleotropic effects (Benson et al., 2017). Critically, this

is an example of the type of investigation that would be extremely complex to undertake clinically and yet can be undertaken computationally with relative ease.

We identified the following 7 drugs with targets in the model (targets in brackets): Ustekinumab (IL12R), GSK1070806 (IL18R), Imatinib mesylate (PDGF), Bindarit (CCL2), cFMS Receptor Inhibitor III (MCSF), GW4869 (SMase) and SCH546738 (CXCR3).

We used the MATLAB software system (https://www.mathworks.com) and a genetic algorithm with a population size of 10000 for 70 generations to identify the optimal combination of drugs that would drive atherosclerosis regression. The genetic algorithm started from one instance of a set of drug concentrations and from this generated a further 19999 instances of sets of drug concentrations from the first by adding Gaussian noise to the concentration of each drug (with standard deviation 1). These 20000 instances comprised the first generation of candidate interventions. All instances were evaluated for their efficacy at plaque reduction and 10000 new instances were created as a second generation of candidate interventions from the two most effective instances of the first generation with the addition of Gaussian noise. The 10000 new instances were then themselves evaluated with the two most effective instances to generate a further 10000 new instances, the third generation. This process was iterated until we arrived at instances from which no improvement in efficacy could be found, at which point we interpreted the best performing instance as optimal. In order to evaluate the efficacy of a particular instance, we constructed a scoring function that allowed the model to develop using the high risk profile for the first forty years before introduction of the drug concentrations of the instance at forty years. The model then continued to run for a further forty years, and at eighty years, we calculated a score for the instance as $S = (C/C_{max} + T/T_{max})/2 + 0.01*(\text{sum of drug concentrations})$ where C is the sum of smooth muscle cells, macrophages, foam cells and T-cells observed and C_{max} is the sum of smooth muscle cells, macrophages, foam cells and T-cells that would occur at eighty years in the absence of any drugs. T is the collegen concentration observed and T_{max} is the collagen concentration that would occur at eighty years in the absence of any drugs. This score describes the efficacy of the instance of a set of drugs at driving plaque regression with effective interventions yielding lower numbers

and ineffective interventions yielding larger numbers. The score also included a sum of the concentrations of the drugs used. Low scores also ensure that the dosages are minimal, yielding therapeutic hypotheses with the least risks of off-target effects. At each generation, the genetic algorithm selected the two instances with the lowest score, using these to populate the next generation of instances. In this way, the algorithm converged on an optimal set of drug concentrations. Analyses were performed on an Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz (Octo-core) CPU with 64GB of RAM running CentOS 7.

3.3 Results

A visual map of the model obtained is shown in Figure 3.1 using the SBGN schema. The model covers five distinct organs and tissues: the liver, intestine, lumen, endothelium and tunica intima. It covers LDL retention, LDL oxidation, monocyte recruitment, monocyte differentiation, smooth muscle cell proliferation, phagocytosis, reverse cholesterol transport and T-cell proliferation. The cell types involved include monocytes, endothelial cells, T-cells, macrophages, foam cells, B-cells, smooth muscle cells, neutrophils, dendritic cells and mast cells. A legend describing the glyphs of the SBGN schema is shown in Figure 3.2. Each interaction represents a parameterized equation (see Appendix 3, Table 1 for the equations and Appendix 3, Table 2 for the parameters), enabling us to dynamically simulate the changing concentrations/abundances of the model as the plaque forms.

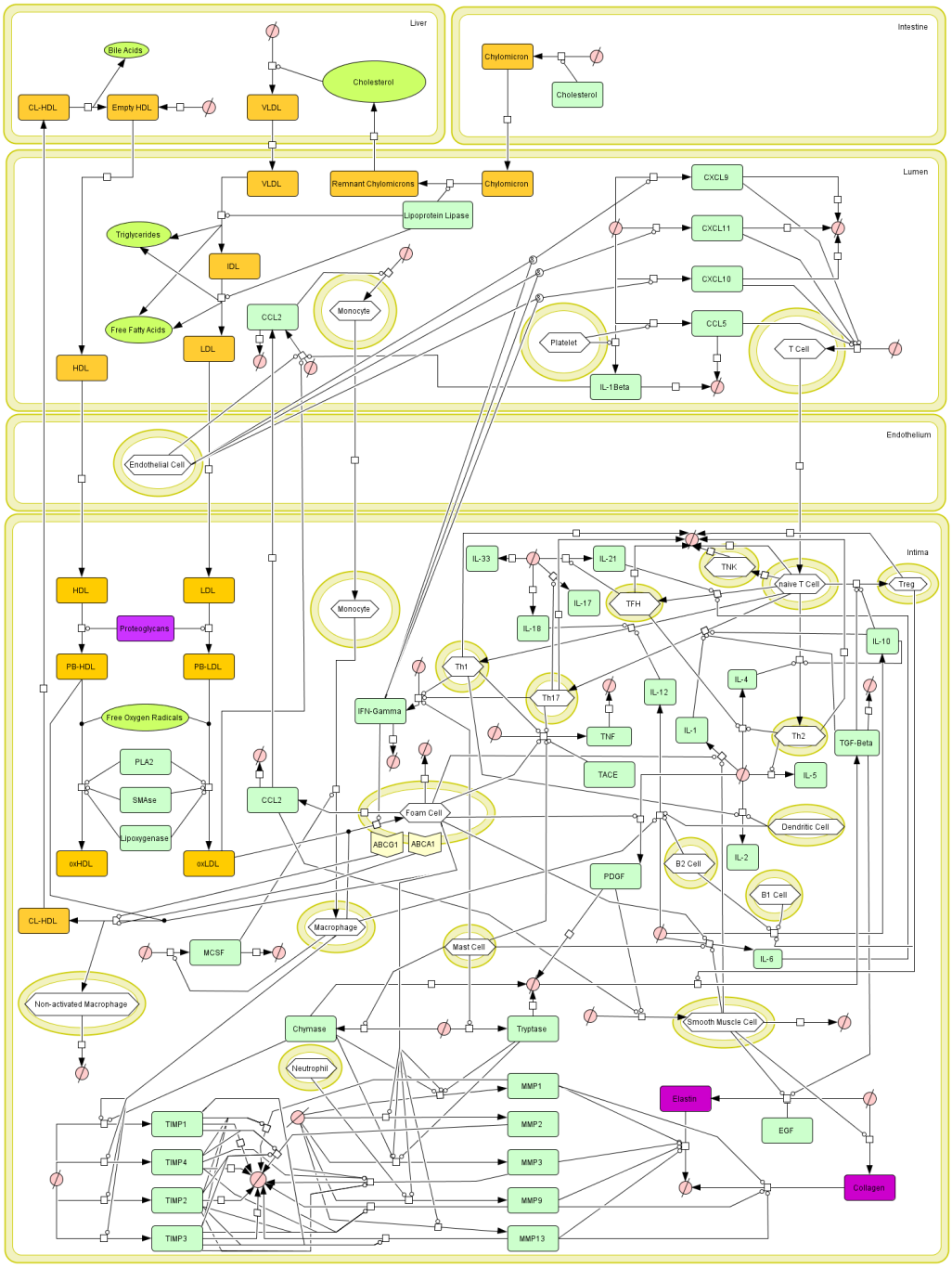


Figure 3.1: A map of atherosclerotic plaque dynamics shown using the Systems Biology Graphical Notation (SBGN)

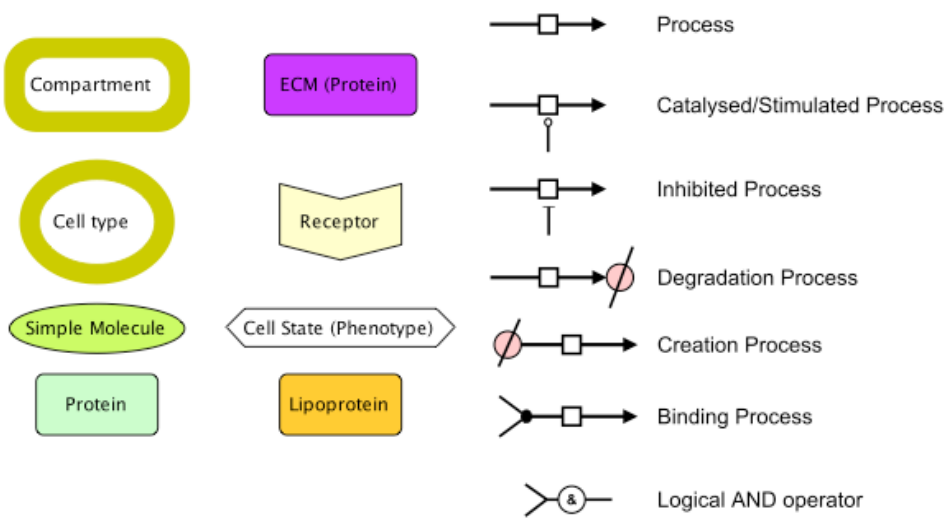


Figure 3.2: The legend for the SBGN schema used in Figure 3.1.

The initial conditions identified are described in Appendix 3, Table 3 and unknown parameters were optimised so that the model maximally satisfied all the constraints described above simultaneously. Key markers for plaque development include foam cell, macrophage and smooth muscle cell proliferation and their behaviour is shown in Figure 3.3 for the three risk profiles.

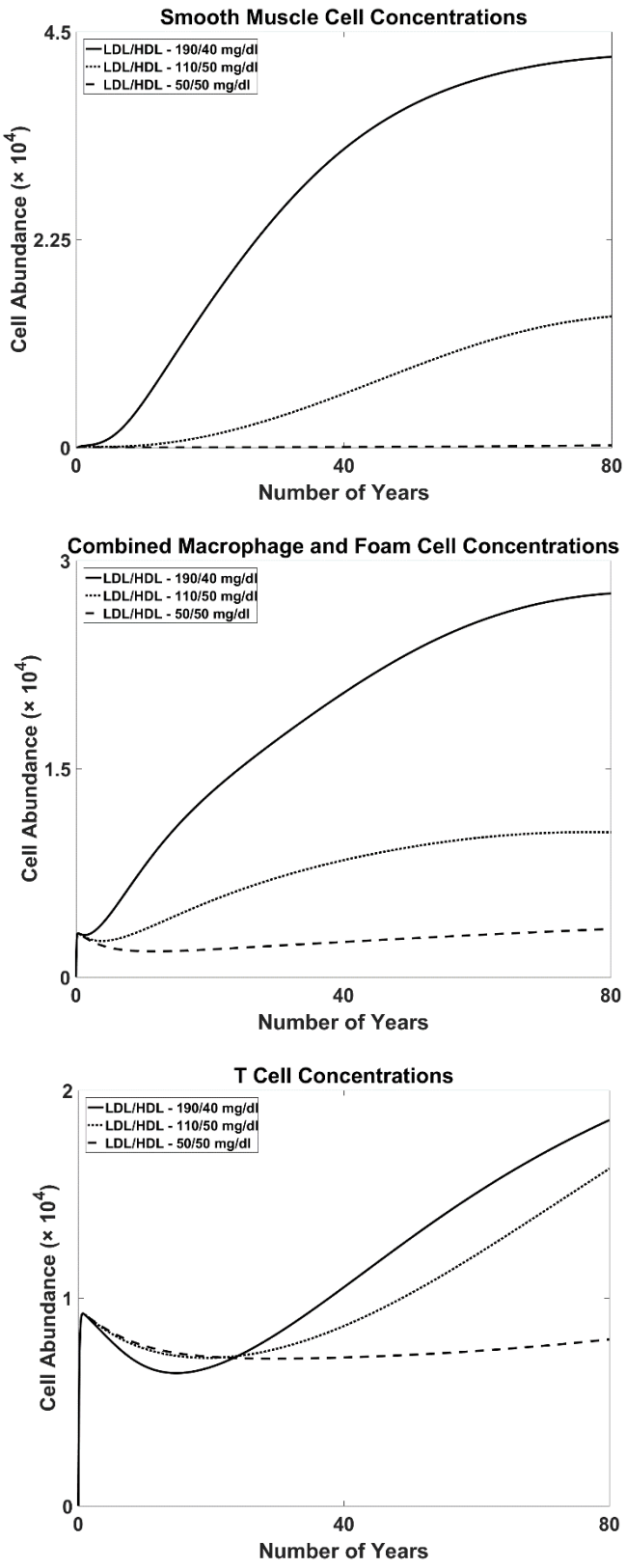


Figure 3.3: Smooth muscle cell, macrophage and foam cell proliferation during plaque development for three blood LDL profiles of 50 mg/dl, 120 mg/dl and 190 mg/dl.

The model satisfies the constraints as follows. Results are stated at 80 years with constraint values in brackets.

3.3.1 Model Constraints

I) Figure 3.3A shows smooth muscle cell abundance, yielding 42287 cells (21341) and 230 cells (133), for high and low LDL profiles, respectively.

II) Figure 3.3B shows combined macrophage and foam cell abundance, yielding 27630 cells (20715) and 3463 cells (129) for high and low LDL profiles, respectively.

III) Figure 3.3C shows total T cell abundance, yielding 18562 cells (18739) and 8012 cells (117) for high and low LDL profiles, respectively.

IV) Figure 3.4.1 shows T_h1 cell abundance, yielding 7186 cells (4324) and 223 cells (27) for high and low LDL profiles, respectively.

V) Figure 3.4.2 shows MCP1/CCL2 blood serum concentration, yielding 649.8 pg/ml (775 pg/ml) and 163.8 pg/ml (100 pg/ml) for high and low LDL profiles, respectively.

VI) Figure 3.4.3 shows CXCL9 blood serum concentration, yielding 283.9 pg/ml (271.2 pg/ml) and 23.8 pg/ml (17.4 pg/ml) for high and low LDL profiles, respectively.

VII) Figure 3.4.4 shows CXCL10 blood serum concentration, yielding 850.0 pg/ml (956.5 pg/ml) and 120.9 pg/ml (127.6 pg/ml) for high and low LDL profiles, respectively.

VIII) Figure 3.4.5 shows CXCL11 blood serum concentration, yielding 965 pg/ml (1062 pg/ml) and 355 pg/ml (420 pg/ml) for high and low LDL profiles, respectively.

IX) Figure 3.4.6 shows triglyceride blood serum concentration, yielding 754 mg/dl (1005 mg/dl) a value that does not change for low LDL profiles (58 mg/dl).

X) Figure 3.4.7 shows chylomicron blood serum concentration, yielding 49.1 μg/ml (52.6 μg/ml) a value that does not change for low LDL profiles (1.4 μg/ml).

XI) Figure 3.4.8 shows IL1β blood serum concentration, yielding 2.04 pg/ml (2.12 pg/ml) and 0.97 pg/ml (0.28 pg/ml) for high and low LDL profiles, respectively.

XII) Figure 3.4.9 shows CCL5 blood serum concentration, yielding 181.1 ng/ml (176.0 ng/ml) and 45.7 ng/ml (2.7 ng/ml) for high and low LDL profiles, respectively.

XIII) Figure 3.4.10 shows MMP1 plaque concentration, yielding 86.8ng/g (104 ng/g) and 0.2 ng/g (18 ng/g) for high and low LDL profiles, respectively.

XIV) Figure 3.4.11 shows MMP9 plaque concentration, yielding 609.6 ng/g (722 ng/g) and 1.6 ng/g (121 ng/g) for high and low LDL profiles, respectively.

XV) Figure 3.4.12 shows TIMP1 plaque concentration, yielding 11.5 μg/g (12.4 μg/g) and 3.6 μg/g (5.3 μg/g) for high and low LDL profiles, respectively.

XVI) Figure 3.4.13 shows IL1β plaque concentration, yielding 23.6 ng/g (24 ng/g) and 0.1 ng/g (12ng/g) for high and low LDL profiles, respectively.

XVII) Figure 3.4.14 shows IL6 plaque concentration, yielding 5.3 μg/g (5.1 μg/g) and 0.025 μg/g (1.5 μg/g) for high and low LDL profiles, respectively.

XVIII) Figure 3.4.15 shows TNFα plaque concentration, yielding 24 ng/g (27 ng/g) and 0.3 ng/g (15 ng/g) for high and low LDL profiles, respectively.

XIX) Figure 3.4.16 shows IL2 plaque concentration, yielding 27 ng/g (24 ng/g) for the high LDL profile.

XX) Figure 3.4.17 shows IL18 plaque concentration, yielding 10.9 ng/g (10.7 ng/g) for the high LDL profile.

XXI) Figure 3.4.18 shows IL10 plaque concentration, yielding 2.1 ng/g (2.3 ng/g) and 0.6 ng/g (3.6 ng/g) for high and low LDL profiles, respectively.

XXII) Figure 3.4.19 shows IL12 plaque concentration, yielding 5.2 ng/g (4.6 ng/g) and 0.7 ng/g (3.6 ng/g) for high and low LDL profiles, respectively.

XXIII) Figure 3.4.20 shows IFNγ plaque concentration, yielding 167 pg/g (182 pg/g) and 5 pg/g (20 pg/g) for high and low LDL profiles, respectively.

XXIV) Figure 3.4.21 shows TGFβ plaque concentration, yielding 0.80 mg/g (0.76 mg/g) and 0.05 mg/g (0.33 mg/g) for high and low LDL profiles, respectively.

XXV) Figure 3.4.22 shows the ratio of plaque density between chymase and tryptase, yielding 106.0:134.3 (107.8:135.1) for the high LDL profiles.

XXVI) Figure 3.4.23 shows the elastin plaque concentration, yielding 1.85 mg/g (1.58 mg/g) for the high LDL profile.

XXVII) Figure 3.4.24 shows collagen plaque concentration, yielding 4.87 mg/g (6.26 mg/g) for the high LDL profile.

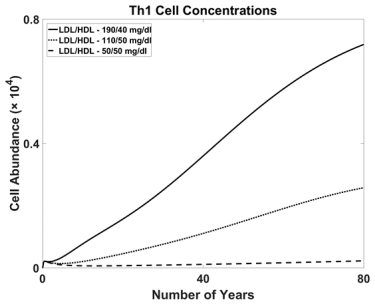
XXVIII) Figure 3.4.25 shows PDGF plaque concentration, yielding 1048 pg/g (1381 pg/g) and 2 pg/g (279 pg/g) for high and low LDL profiles, respectively.

XXIX) Figure 3.4.26 shows foam cell aggregation after the parameter determining rate of differentiation to T_h1 cells has been increased by 10% and the parameter determining the rate of differentiation to T_h2 cells has been decreased by 10%. This has led to a modest increase in foam cell concentrations for a high LDL profile.

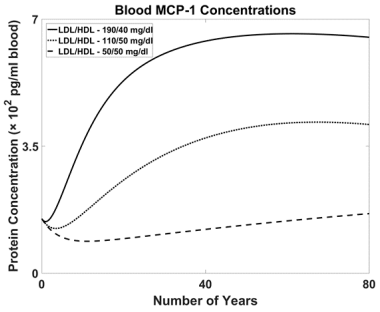
XXX) Figure 3.4.27, Figure 3.4.28 and Figure 3.4.29 shows oxidized LDL concentrations and smooth muscle cell and foam cell abundance when the LDL and HDL are switched from 190 mg/dl and 40 mg/dl, respectively, to 50mg/dl and 50mg/dl, respectively after 40 years, demonstrating plaque reduction.

XXXI) Figure Figure 3.4.30 shows oxidised LDL plaque concentration depending on LDL profile. At 80 years, the high LDL profile yields 36.8 μg/g (38.6 μg/g) and the low LDL profile yields 2.6 μg/g (0.25 μg/g).

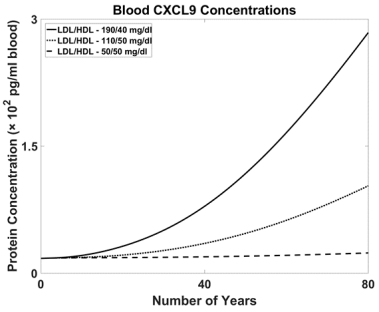
1)



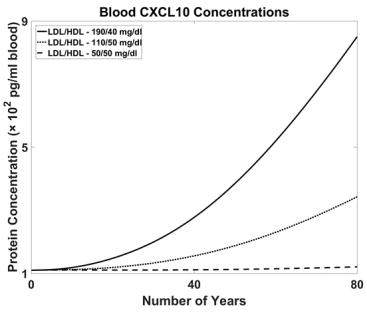
2)



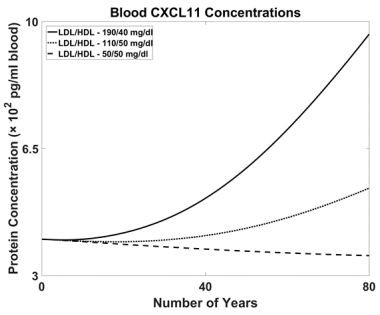
3)



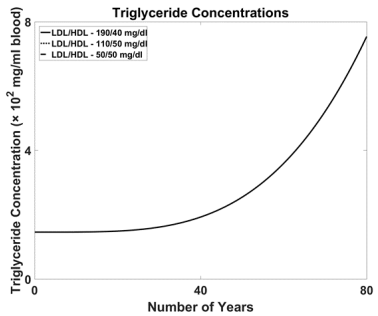
4)



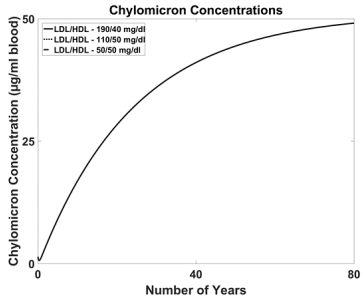
5)



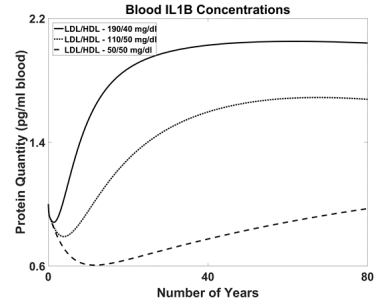
6)



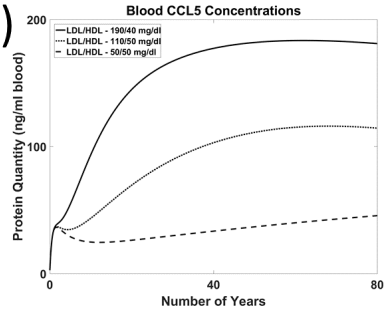
7)



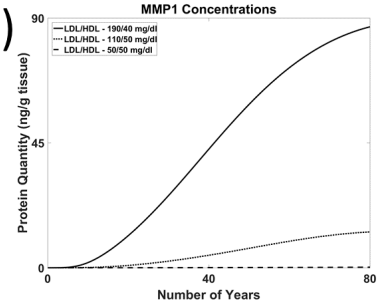
8)



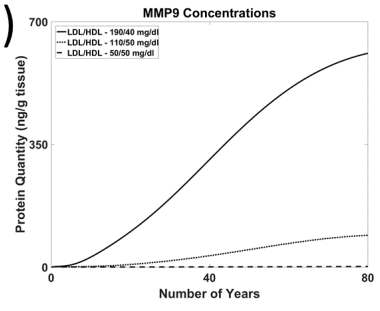
9)



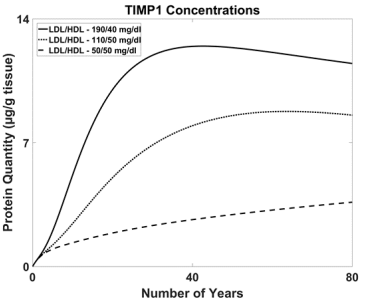
10)



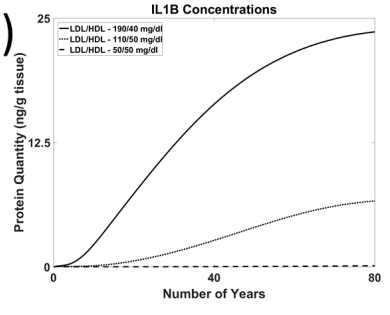
11)



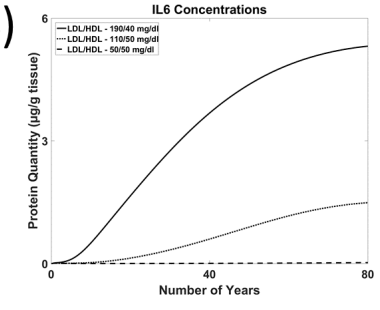
12)



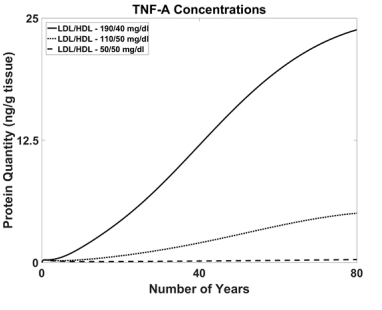
13)



14)



15)



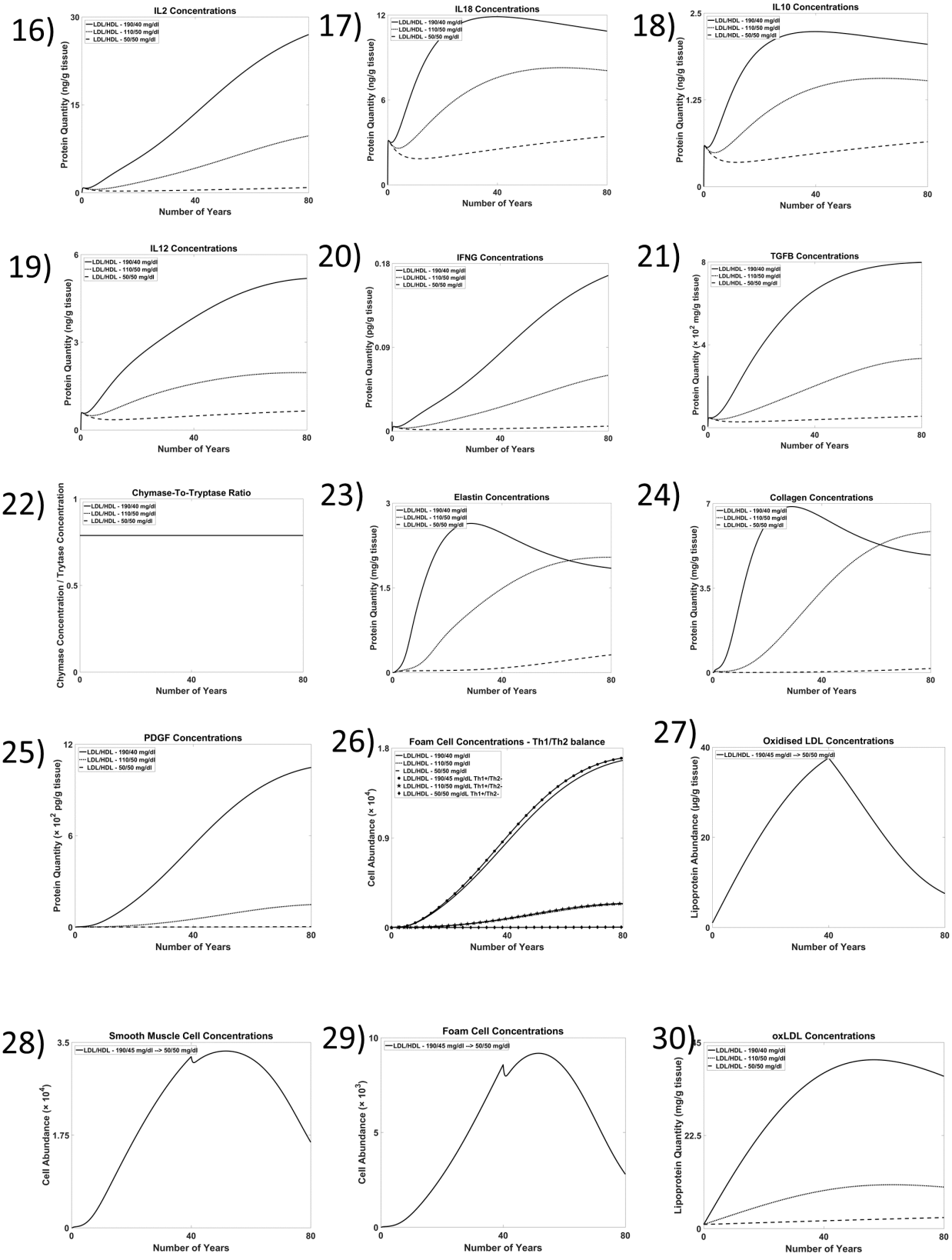


Figure 3.4: The performance of the model for clinical requirements determined from the literature

3.3.2 Clinical Results

In addition to addressing the constraints outlined above, the model also agrees with the following clinical results.

XXXII) Blockade of endogenous IL-12 has been shown to reduce atherogenesis (Hauer et al., 2005). Figure 3.5.1 shows that with a 75% reduction to the rate parameters describing IL-12 production, foam cell abundance is significantly reduced for high and mid LDL profiles.

XXXIII) Deficiency of ABCA1 function impairs reverse cholesterol transport and increases atheroma size (Westerterp et al., 2013). Figure 3.5.2 shows that with a reduction in the initial ABCA1 concentration by 90%, foam cell concentration is increased across the lifetime of the simulation.

XXXIV) Deficiency of MCSF reduces monocyte/macrophage circulation and plaque formation (Qiao et al., 1997). Figure 3.5.3 shows that with a reduction in the initial MCSF concentrations from 100 μg/mg of tissue to 0, macrophage abundance drops significantly within the plaque.

XXXV) T-cells abundance is reduced as a result of IFNGR knockout (Gupta et al., 1997). Figure 3.5.4 shows that decreasing the k_{cat} rate parameter describing IFNG production by 50% reduces cell abundance within the plaque.

XXXVI) IL-18 has been shown to be atherogenic (Whitman, 2002). Figure 3.5.5 shows that increasing the rate parameter describing IL-18 production by 50%, increases smooth muscle cell recruitment within the plaque.

XXXVII) Reduction in proteoglycan concentration reduces intimal oxLDL concentrations (Delgado-Roche et al., 2015). Figure 3.5.7 shows that decreasing the initial concentration of proteoglycan concentration from 500 μg/mg of tissue to 100 pg/mg of tissue reduces the concentration of oxidized LDL within the plaque.

XXXVIII) Increasing activity of matrix metalloproteinases leads to degraded extracellular matrix (Adiguzel et al., 2009). Figure 3.5.8 shows that doubling the binding rate parameter between extra cellular matrix and matrix metalloproteinases significantly reduces collagen concentrations.

XXXIX) PLA2 concentration has been shown to correlate with atherogenesis (Vickers et al., 2009). Figure 3.5.9 shows that a reduction in the initial PLA2 concentrations by 90% reduces the foam cell concentration within the plaque.

XL) Increasing PDGF activity increases smooth muscle cell abundance (Ferns et al., 1991). Figure 3.5.9 shows that increasing the rate parameter describing PGDF production by 200% increases smooth muscle cell recruitment in the plaque.

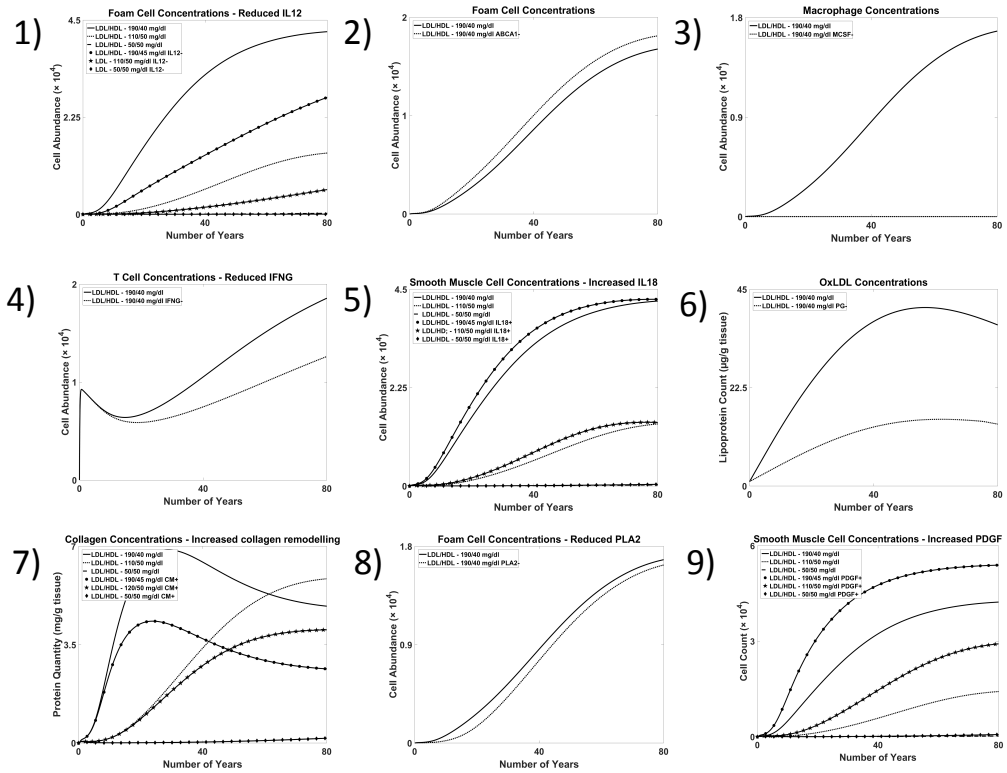


Figure 3.5**:** The performance of the model for further clinical observations

3.3.3 Reusability of the model

The visual map is available as a file encoded using the SBGN-ML format and the mathematical model is available as a file encoded using the SBML format, both from the supplementary material. The mathematical model has been deposited into the European Bioinformatics Institute's BioModels repository (MODEL1710020000).

The files can be opened, edited and analysed in software supporting the SBGNML and SBML standards. SBML compliant software includes Copasi (Bergmann et al., 2017), Cytoscape with the cy3SBML plugin (König et al., 2012) and Dizzy (Ramsey et al., 2005). Figure 3.6

shows the graphical map opened in three representative SBGN compliant editors: Newt (http://web.newteditor.org/), PathVisio (Kutmon et al., 2015) and VANTED with SBGN-ED extension (Czauderna et al., 2011) along with a subsection of the plain text XML file.

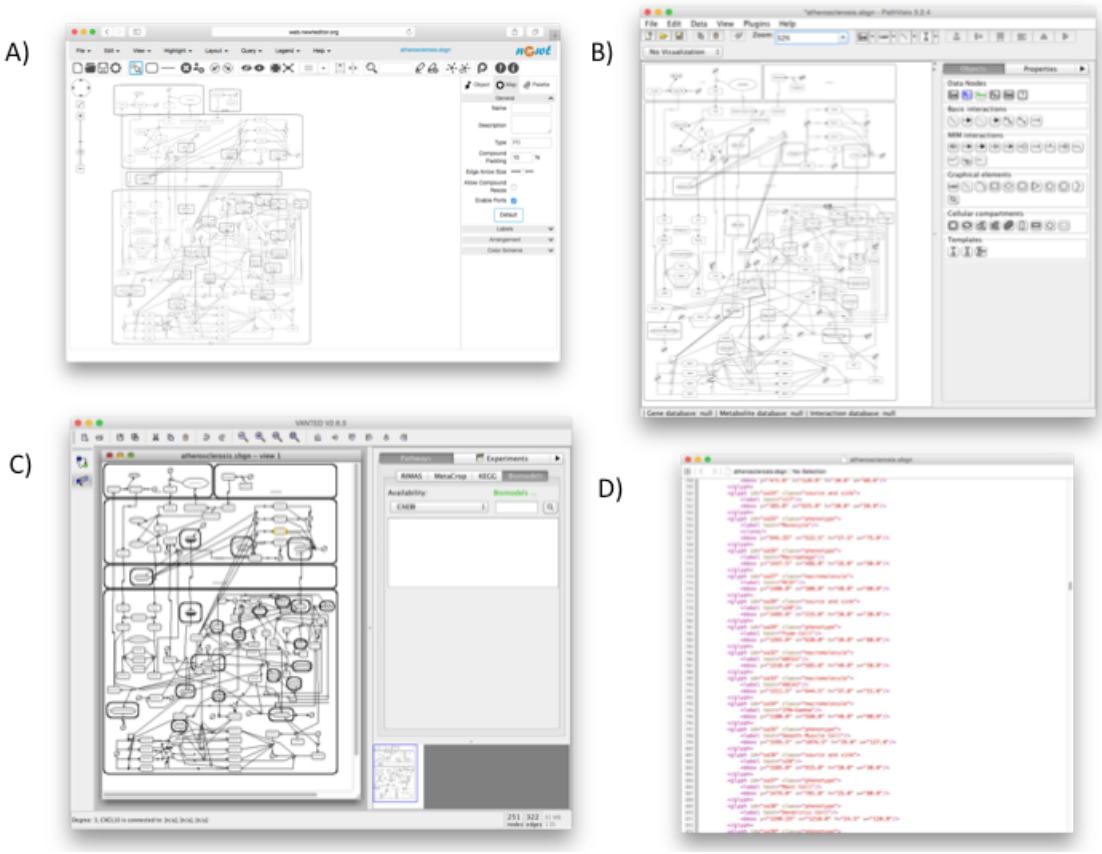


Figure 3.6: The model viewed in using the A) Newt B) PathVisio and C) VANTED platforms and D) viewed as plain text XML.

3.3.4 Therapeutic hypothesis generation

Running the analysis described in section 3.2.2, we were able to determine the following drug combination as an optimal intervention that drove plaque regression: Ustekinumab (IL12R) – 7.6279, GSK1070806 (IL18R) - 7.5937, Bindarit (CCL2) – 36.9922, where concentrations are described as multiples of the corresponding inhibition constants, K_i. This combination was identified relatively quickly by the model, in 1080 minutes, with approximately optimal results being identified much more quickly, within 300 minutes. Figure 3.7B, Figure 3.7C and Figure 3.7D show the dynamics of the model after this intervention is applied at forty years. Here we can see that smooth muscle cells,

macrophages and foam cells and T_h1 cell counts are all driven down by the intervention across the second period of forty years.

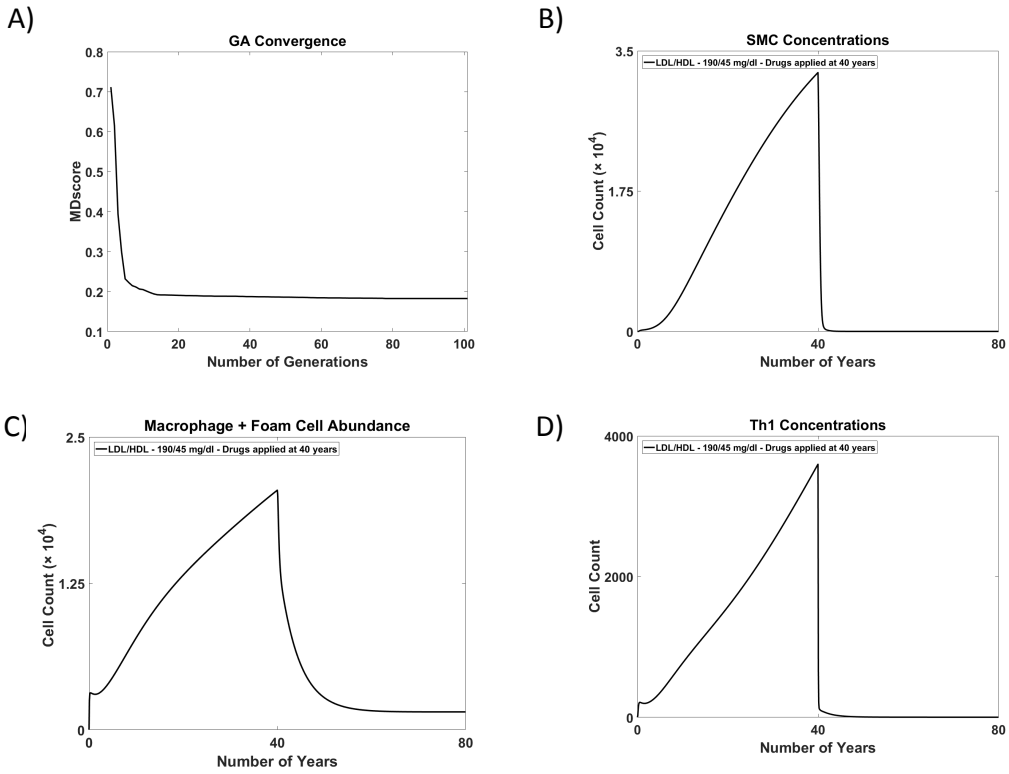


Figure 3.7: Convergence on an atheroprotective multi-drug intervention hypothesis

3.4 Discussion

Atherosclerotic plaques are highly challenging to study due to their location. Human *in vivo* studies present logistical and ethical challenges and there are few *in vitro* resources that can contribute increasing our understanding of plaque development. Whilst they are not a complete replacement for *in vitro* studies, computational studies have the potential to contribute to research in this area and to yield non-*in vivo* resources that can improve our understanding of the disease.

Here we have produced a predictive model of the dynamics of atherosclerosis, which we hope will serve as a resource for the cardiovascular research community that can be reused, refined and expanded in future. The model we have produced has the potential to contribute to therapy development through multiple avenues. Primarily, the model can be used to predict the consequences for the dynamics of atherosclerosis of interventions that target components of the pathways involved in the disease. This can be applied to single

drug development, by identifying the components of the model that have the greatest impact on foam cell accumulation and smooth muscle cell recruitment or to multi-drug interventions that achieve similar goals through the compound effect of suppressing multiple pathway components, but to smaller individual degrees (Benson et al., 2017). It is known that atherosclerosis is a comorbidity of diseases such as rheumatoid arthritis and depression (Gibson et al., 2017). By using proteomic data from studies of other diseases, this model can also be used to explore the role of atherosclerosis as a comorbidity of other conditions. Similarly it can be used to explore the possible off-target impact of therapies targeting a separate conditions, where the therapeutics are known to also target components of the pathways associated with atherosclerosis.

Although we often consider disease pathologies in isolation, atherosclerosis is part of a much larger network of interactions and we can use the model to explore the impact of interventions on the components of pathways and networks that regulate atherosclerosis. For example, it would be possible to extend the model to include PCSK9 metabolism in order to explore the impact of PCSK9 inhibitors on plaque development or to include JAK-STAT signalling to explore the role of immune signalling on atherosclerosis progression.

A predictive model of this type has the potential to move the discussion around disease from an understanding of behaviour of individual disease components (such as foam cell accumulation or smooth muscle cell recruitment) to an understanding of the dynamics of the network and of how, with the transition from health to disease, the network dynamics as a whole transition from healthy dynamics to disease dynamics.

The predictions of the model show broad agreement with observed clinical results. Because the model describes spatial effects and cellular function at extremely simple levels, it is unlikely to be able to recreate all clinical results exactly. Doing so would require a model of greater complexity across length scales. However, the model demonstrates order of magnitude agreement in almost all cases and shows the correct qualitative dose responses. In many cases, we found it particularly challenging to constrain parameters so as to ensure sufficiently large responses to changes in lipoprotein profile. As a result, the model does systematically over-estimate components of the model for lower LDL profiles and the

difference between high and low LDL profiles, although large, is not as great as that observed clinically.

As we have demonstrated, a model of this form can be used to study disease dynamics and to develop therapeutic hypothesis. Such a model can be adapted to individuals or to patient subgroups by tuning the parameters of the interactions. Creating parameterisations that are tailored to individuals could be achieved by optimizing the model to patient time course data or through computational inference from patient genome data. However, adapting the model in this way to represent the disease dynamics of individuals or patient subgroups would yield models of atherosclerosis that could support programmes of personalized or stratified medicine, facilitating the development of therapeutic hypotheses that are tailored to the patient or the patient subgroup.

CVD is a large burden on healthcare worldwide. Front line therapies for the primary and secondary prevention of atherosclerotic vascular disease include smoking cessation, lipid management, blood pressure control, optimal control of diabetes and the use of antiplatelet agents. By far the most commonly used class of lipid lowering drugs is statins, which inhibit HMG-CoA reductase. Ezetimibe, a cholesterol absorption inhibitor, may be used in patients who are statin intolerant or who are not achieving lipid targets on the highest maximally targeted dose of statin. A new, recently licenced class of drugs, proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors suppress LDLR degradation by PCSK9 are associated with significant reduction in serum LDL cholesterol concentration and cardiovascular events. Emerging drugs include Apolipoprotein B antisense drugs that suppress translation of ApoB, a key component of LDL and microsomal triglyceride transfer protein inhibitors that show significant LDL-C reduction.

Amongst the drugs identified as part of the multi-drug intervention, numerous studies of CCR2 inhibition in atherosclerosis have shown a reduction in intima media thickness and plaque area, lowered monocyte infiltration and increased lesion stability (Zhao, 2010). Propagermanium, an anti-inflammatory drug, has been shown to affect CCR2 function (Yokochi et al., 2001). Targeting CCR5 with the drug maraviroc has been shown to reduce atherogenesis in mice (Cipriani et al., 2013).

The scale of global CVD burden means that there is a pressing need to develop new pharmaceutical therapeutics in this area that both address clinical need and can sustain the pharmaceutical industry as intellectual property protection expires around current therapeutics. Multi-drug interventions of the type identified here have a vast untapped potential to contribute to future therapeutics in this way.

3.4.1 Acknowledgements

We are indebted to Patricia Navarro for assistance with figure production.

Chapter 4:

Variations in Protein Structure

4.1 Introduction

The study of protein-protein interactions (PPIs) is vital to improving our understanding of disease dynamics. PPIs are heavily involved in atherosclerosis, driving the subprocesses of lipoprotein oxidation (Yoshida and Kisugi, 2010), immune cell recruitment (Bobryshev, 2006), monocyte migration (Moore et al., 2013a), macrophage differentiation (Qiao et al., 1997), T cell differentiation (Tse et al., 2013a), extra cellular matrix (ECM) remodelling (Newby, 2005), and fibrous cap formation (Rudijanto, 2007). Using the model described in Chapter 3 of this thesis, we intend to show how variation in protein structure affects atherosclerosis dynamics *in silico*. Protein structure determines function, and subsequently molecular processes and biochemical events driven by PPIs are influenced by structural changes. Proteins obtain their 3D structure through protein folding, where polypeptides spontaneously form their functional structure based on their amino acid sequence. Protein folds have been greatly studied by the scientific community in recent years, with great progress seen in protein structure prediction strategies. Variants caused by single nucleotide polymorphisms (SNPs), insertions or deletions (indels) and other mutations which lead to a change in the amino acid sequence of a protein can cause significant functional differences. Studying the variation in proteins involved in atherogenesis can be used to hypothesise whether a mutation would disturb the dynamics of atherosclerosis and perform part of a basis for an *in silico* learning platform.

4.1.1 The 1000 Genome Project

The 1000 Genome Project was created in 2008, designed to curate a collection of variations discovered within human genomes (Abecasis et al., 2012; Durbin et al., 2010). Reductions in cost of next generation sequencing techniques, such as whole genome sequencing (WGS) and whole exome sequencing (WES), led to the creation of the project devised to provide a publicly-available database of human genetic variation. In phase 3 of the 1000 Genome Project, whole genome sequence data, generated from the Illumina sequencing platform, derived from 2504 samples was integrated into the repository for use by the wider scientific community (Sudmant et al., 2015). These 2504 samples were taken from unrelated individuals with genetic lineage from 26 different populations. 84.4 million variants were found across phase 3 of the project. The majority of these variants (81.3 million) were found on the autosomes. Single Nucleotide Polymorphisms (SNPs) consist of 96.1% of

variant sites on the autosomes, and indels are responsible for 3.8%. The average person is expected to have between 4.1 million and 5 million variations that differ from the 1000 Genome Project reference genome (Auton et al., 2015). Most variations found within humans happen at low frequency; more than 65% of allele variants occur in less than 0.2% of the population. These rarer variants are more likely to be specific to particular population subgroups, whereas nearly all variants found in phase 3 of the 1000 genome project with a variant allele frequency (VAF) greater than 2% were found in all continents (Sudmant et al., 2015).

4.1.2 Protein Folding

Simulating polypeptides folding into unique structures has been a problem tackled by computational chemists for years (Dorn et al., 2014). Complex native protein states form rapidly (in the order of nanoseconds), making it challenging for experimental techniques to obtain an accurate picture of the folding process at the atomic level. Even in cases where proteins do not share an evolutionary origin, proteins often share structural qualities with other proteins. Structural Classification of Proteins (SCOP), a manually-curated repository describing the structural relationships between proteins, contains 1195 different 'folds' or 'motifs' found within multiple protein structures with no common homologues (Murzin et al., 1995). SCOP has been extended (SCOPe) in recent years, using automation alongside manual curation to provide classifications of protein superfamilies (Chandonia et al., 2017). Identifying the amino acid sequences that lead to these motifs has led to significant advances in the prediction of three-dimensional structure from amino acid sequences.

4.1.3 Protein structure/function

As of July 2017, approximately 88 million protein sequences exist in the UniProt knowledgebase (Bateman et al., 2015), and 122,500 experimentally validated structures exist in the Protein Data Bank (PDB) (Rose et al., 2015), meaning that there are validated structures for approximately 0.14% of proteins with a known amino acid sequence.

Structures can be divided into four categories, or levels, depending on their complexity. Primary structure, the linear amino acid sequence held together by covalent bonds, is defined by the gene associated with the protein. The N-terminus, the initial section of the

protein created by the ribosome during protein synthesis, often contains a signal peptide sequence that ensures that the protein is delivered to the intended organelle. During and after protein synthesis, polypeptide chains fold into their secondary and tertiary structures. Post-translational modification (PTM), which can occur before or after the folding process, can lead to the removal of this signal peptide as part of assembly of the mature form of the protein. Amino acid sequences are conventionally read from N-terminus to C-terminus.

Secondary structure involves the formation of motifs, or individual segments of proteins. These motifs, such as alpha helices or beta sheets, typically form before the formation of the full tertiary structure. Local segments of the tertiary structure form as part of a transitional state, where a pattern of hydrogen bonds form in the backbone of the polypeptide chain at the motif location, before folding of the three-dimensional structure is complete. Predicting the secondary structure of a protein from its amino acid sequence is an important step in tertiary structure prediction.

The three-dimensional shape of a protein is referred to as its tertiary structure. Tertiary structures differ from quaternary structures as they are composed of a single polypeptide chain. This polypeptide 'backbone' is a result of the linked collection of residues defined as the primary structure. Side chains, which attach to the main polypeptide chain at the alpha-carbon atoms, fundamentally define the tertiary structure, chemical properties and interaction mechanisms of a protein. This three-dimensional shape gives the protein the ability to achieve its biological function.

Quaternary structure refers to the three-dimensional complex formed by arrangement of multiple tertiary structure subunits, including oligomers and complexes formed through PPIs. Certain proteins are only active in their oligomer form. Most protein structures perform their biological function when part of a quaternary structure, rather than in their monomer form. Complex formation can lead to activation or inhibition of one of the subunits of the complex.

4.1.4 Protein Data Bank

The Protein Data Bank (PDB) is a publically available repository of protein structures curated through experimental methods (Rose et al., 2015). Structures contained within the PDB are stored as a series of atomic coordinates in a .pdb file. The PDB contains tertiary and quaternary structures for more than 122,000 proteins (12/7/17), derived using X-ray crystallography, nuclear magnetic resonance (NMR), electron microscopy, microelectron diffraction and combinations of these methods. The PDB file format, .pdb, is the standard format for collating three-dimensional molecular structures, and allows for the storage of information pertaining to atomic coordinates, secondary structure, residue types, temperature factors, chain identifiers, atom names and occupancy. The PDB is the primary source of structural biology data online, and it is the usual source for homologous data used by bioinformatics tools.

4.1.5 Predicting Tertiary Structure

The tertiary structure of a protein is determined by its amino acid sequence. The almost instantaneous nature of protein folding leads to a series of questions. What forces cause the protein to quickly conform to its native structure? How is the protein capable of avoiding unwanted conformations? These questions are beyond the scope of this thesis; however they are questions that require some elucidation to provide high quality predictions of tertiary protein structure from amino acid sequences.

Protein structural prediction methods can be advantageous in filling in the gaps in our knowledge due to the significant quantity of known amino acid sequences with no known structure. As these methods develop, understanding of how protein folding mechanisms operate will improve.

Currently, protein structure prediction tools can be inserted into one of four broad categories:

4.1.5.1 Homology Modelling

Homology modelling relies on the concept of conservation of structure between similar proteins (Martí-Renom et al., 2000). Evolutionarily linked proteins generally have similar sequences and structure. Through the identification of proteins that share a common

ancestry, or homologues, then we can identify structural motifs that are likely to remain in our target protein. Sequence alignment and structural templates with expected shared motifs are then combined to predict a protein structure.

4.1.5.2 Threading

Threading is a similar process to homology modelling; however proteins of the same fold are identified and used to infer a protein structure, rather than predicting structure through comparison to evolutionary homologues (Xu et al., 2008). The number of folds that have been identified is relatively small compared to the number of proteins in the PDB, which allows for the statistical inference of protein structures from amino acid sequences (Chandonia et al., 2017; Rose et al., 2015).

4.1.5.3 Ab Initio

Ab Initio, or *de novo,* protein structure prediction relies on calculating protein structures with no available homologue or protein folding data. (Hardin et al., 2002). The most efficient protein structure prediction methods utilise structural data taken from the PDB, including from non-homologous proteins. This is a computationally intensive task, and as such these methods are more likely to be successful on relatively small proteins (fewer than 120 amino acids).

4.1.5.4 Secondary Structure

Secondary structural elements such as alpha helices or beta sheets can be identified from amino acid sequences. While these predictions will not provide a tertiary protein structure, insight into protein structure and function can be derived from a secondary structure prediction.

4.1.6 Critical Assessment of Protein Structure Prediction (CASP)

There are multiple different bioinformatics tools that predict a protein structure from an amino acid sequence, using a wide range of different methods and algorithms to reach their predicted result. The Protein Structure Prediction Centre have developed a series of assessments, titled 'The Critical Assessment of Protein Structure Prediction' (CASP), designed to objectively test available methods for protein structure prediction (Moult et al., 2016). CASP has been organised every two years since 1994, and has been described as "the

world championship of protein structure prediction" due to its competitive nature (Moult et al., 2014, 2016). This provides the research community and potential users with a non-partisan review and assessment of the predictive power of all of the algorithms involved in the study.

At its core, CASP participants are provided with an amino acid sequence and asked to provide a tertiary structure to be compared with a recently-elucidated experimentally-derived structure. In CASP11, performed in 2014, 207 modelling groups entered from approximately 100 research laboratories (Moult et al., 2016).

Categories where modelling performance is judged include homology modelling methods, *ab initio* methods, model refinement, model accuracy prediction and contact point and binding site prediction. Model refinement and model accuracy prediction have seen significant improvements recently (Heo et al., 2013; Moult et al., 2016). Refinement methods can additionally help with minimising template bias.

Recently solved protein structures through protein crystallography or NMR spectroscopy are held back by the PDB. CASP12 entrants were subsequently provided with 90 models for which to predict tertiary structure, with independent assessors ranking each entrant within each category.

4.1.7 I-TASSER

I-TASSER (Iterative Threading ASSEmbly Refinement) is a tool developed by the Yang Zhang Lab, designed to heuristically predict protein structure and function from amino acid sequences (Roy et al., 2010). I-TASSER has been consistently successful in recent community-wide CASP experiments, ranked as the top protein structure prediction method four times out of the last five experiments. In summary, I-TASSER will initially establish a secondary structure for a given amino acid sequence by matching the query sequence to an existing database (to select homologues) and performing multiple alignments with these evolutionary relatives (McGuffin et al., 2000). Multiple threading programs are then used to predict templates, and their quality is ranked based on the significance of the threading alignment. Successfully aligned continuous fragments from these threading-generated structures are used to assemble conformations, with *ab initio* modelling used to build other

unaligned regions. A modified Monte Carlo simulation technique is used to assemble these fragments. An additional structural assembly simulation is completed, with additional constraints implemented from the initial threading alignments. A collection of highly detailed reports on the I-TASSER methodology can be found in (Roy et al., 2010; Wu et al., 2007; Yang et al., 2014; Zhang, 2008).

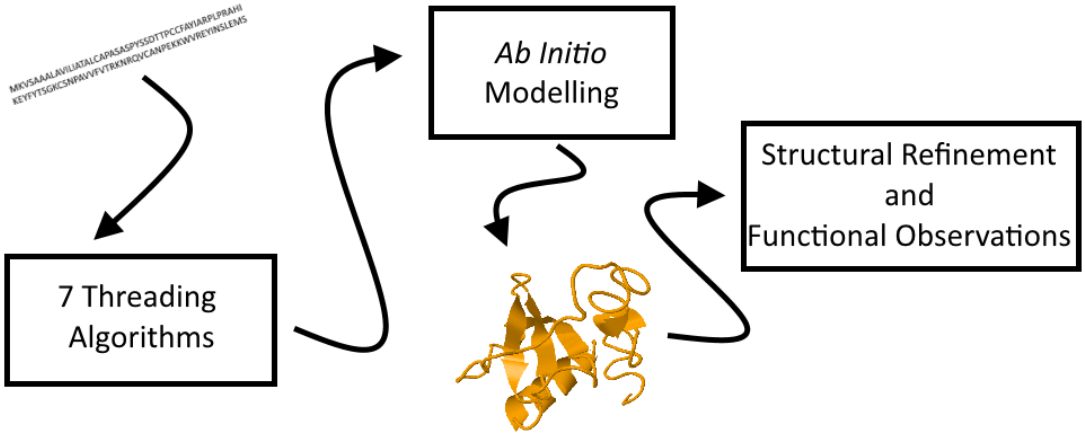


Figure 4.1: Summary of I-TASSER methodology

4.1.7.1 Evaluating Quaternary Structure

An oligomeric protein structure is a complex that consists of two or more monomers (or tertiary structures), held together by non-covalent bonds. Homo-oligomers are complexes that consist of identical parts, whereas hetero-oligomers are composed of different monomers. For example, interferon gamma (a pro-inflammatory cytokine secreted by T_h1 cells that is a potent driver of atherosclerosis (McLaren and Ramji, 2009)) is a homodimer in its functional form, whereas interferon gamma receptor (the associated receptor required for cell signalling) is a heterodimer, built from two separate proteins coded by the genes IFNGR1 and IFNGR2 (Tau and Rothman, 1999). Oligomeric structures can be composed of different numbers of subunits.

Certain proteins can only perform their biological function when they become part of a complex. Environmental differences can alter function for the same complex (Nooren, 2003).

Methods for predicting oligomeric status and quaternary structure from primary structure have improved significantly within the last few years (Biasini et al., 2014; Mukherjee and

Zhang, 2011). Docking methods have been used to predict dimeric structure from tertiary structure (Davis and Baker, 2009), and methods have been developed to predict homo-oligomeric status and structure from tertiary structure (Baek et al., 2017). Continuous Automated Model Evaluation (CAMEO) is a community effort to evaluate tertiary and quaternary structure prediction methods (Haas et al., 2013). Docking of rigid-body tertiary structures is the primary method used to model complex structures; however homology modelling methods and methods to model complexes from primary structure have seen success recently.

4.1.8 Comparing Protein Structures

Algorithms to compare protein structures and quantify differences are continually progressing. There are two primary categories of structural comparison methods, depending on the existence of an alignment between matching residues within the structure. The root-mean-square derivation (RMSD) uses an optimal alignment and compares the pairwise distance between residues. However, when this optimal alignment is not given then alignments require to be identified, or alignment-independent algorithms must be used (Kufareva and Abagyan, 2012).

Predicted protein structures often see an increased amount of variation and noise at the terminals due to fewer constraints from other residues in the system. RMSD evaluates the distances between each residue pairing equally, potentially misrepresenting the overall accuracy if there is significant variance at either terminal. The TM-Score is a method that weights the terminal distances less than the mid-protein distances, removing instances where significant terminal variation gives misleading results (Zhang and Skolnick, 2005).

4.2 Methods

4.2.1 Sequence Isolation and Structure Prediction

The model of atherosclerosis derived in Chapter 3 is used as a basis for studying the structures of proteins involved in atherosclerosis. These proteins are shown in Table 4.1 as they are described in the mathematical model.

Table 4.1: Proteins described in Chapter 3's Model of Atherosclerosis

ABCA1	ABCG1	CCL2 in Intima	CCL2 in Lumen	CXCL9
CXCL10	CXCL11	Chymase	Collagen	EGF
Elastin	Interferon Gamma	IL-1B	IL-10	IL-12
IL-17	IL-18	IL-2	IL-21	IL-33
IL-4	IL-5	IL-6	MCSF	MMP1
MMP2	MMP3	MMP9	MMP13	PDGF
PLA2	Proteoglycans	SMAse	TGF-Beta	TIMP1
TIMP2	TIMP3	TIMP4	TNF-Alpha	Tryptase
TACE				

A literature mining process was undertaken to establish the genes that encode each of these proteins, and whether the protein is an oligomer in its active form. For each protein, a Google Scholar search was undertaken with the query 'PROTEINNAME gene', and the top five papers were searched for a gene name to correspond with each protein. Four of these proteins describe a class of proteins (collagen, elastin, proteoglycans and tryptase) rather than a single macromolecule, and are not included within this dataset. Experimentally validated versions of these proteins within the protein data bank were found and their PDB codes, encoding genes and the proteins they interact with in the model is described in Table 4.2. References for each of these interactions can be found in Appendix 3, Table 1. The oligomeric status of the non-monomer proteins are described in Table 4.3.

Table 4.2: Gene names, PDB codes and interaction data for proteins contained within Chapter 3's model of atherosclerosis

No.	Protein 1	Protein 1 Gene Name	Protein 2	Protein 2 Gene Name	PDB Code	Notes
1	TACE	ADAM17	TIMP3	TIMP3	3CKI	
2	MCP1	CCL2	CCR2	CCR2	3IFD (1) 5T1A (2)	
3	RANTES	CCL5	CCR5	CCR5	1HRJ (1) 4MBS(2)	
4	Chymase	CMA1	MMP1	MMP1	1NN6	
5	MCSF	CSF1	CSF1R	CSF1R	4WRL	
6	CXCL10	CXCL10	CXCR3	CXCR3	1o80	
7	CXCL11	CXCL11	CXCR3	CXCR3	1RJT	
8	CXCL9	CXCL9	CXCR3	CXCR3		
9	EGF	EGF	EGFR	EGFR	1IVO	
10	IFNG	IFNG	IFNGR	IFNGR1	1FG9	IFNG is Homodimeric

No.	Protein 1	Protein 1 Gene Name	Protein 2	Protein 2 Gene Name	PDB Code	Notes
				IFNGR2		IFNGR is heterodimeric
11	IL1B	IL1B	IL1R	IL1R1		
12	IL10	IL10	IL10R	IL10RA	1J7V	
13	IL12	IL12A IL12B	IL12R	IL12R1, IL12R2	1F45	IL12 and IL12R are Heterodimeric
14	IL17A	IL17A	IL17RA	IL17RA	4SHA	
15	IL18	IL18	IL18R	IL18RA IL18RAP	4R6U	IL18R is Heterodimeric
16	IL2	IL2	IL2R	IL2RA	2B5I	
17	IL21	IL21	IL21R	IL21R	3TGX	
18	IL33	IL33	IL1R	IL1R1	4KC3	
19	IL4	IL4	IL4R	IL4R	3BPL	
20	IL5	IL5	IL5R	IL5RA	3QT2	IL5 is Homodimeric
21	IL6	IL6	IL6R	IL6R	4J4L	
22	PDGFA	PDGFA	PDGFRA	PDGFRA		
23	PDGFB	PDGFB	PDGFRB	PDGFRB	3MJG	PDGFB is Homodimeric
24	PDGFA	PDGFA	PDGFRB	PDGFRB		
25	PDGFB	PDGFB	PDGFRA	PDGFRA		
26	TGFB	TGFB1	TGFBR	TGFBR1	3KFD	TGFB is Homodimeric
				TGFBR2		TGFBR is Heterodimeric
27	TIMP1	TIMP1	MMP1	MMP1	2J0T	
28	TIMP1	TIMP1	MMP2	MMP2		
29	TIMP1	TIMP1	MMP3	MMP3	1OO9	
30	TIMP1	TIMP1	MMP9	MMP9	1L6J	
31	TIMP1	TIMP1	MMP13	MMP13	2E2D	
32	TIMP2	TIMP2	MMP1	MMP1		
33	TIMP2	TIMP2	MMP2	MMP2	1GXD	
34	TIMP2	TIMP2	MMP3	MMP3		
35	TIMP2	TIMP2	MMP9	MMP9		
36	TIMP2	TIMP2	MMP13	MMP13		
37	TIMP3	TIMP3	MMP1	MMP1		
38	TIMP3	TIMP3	MMP2	MMP2		
39	TIMP3	TIMP3	MMP3	MMP3		
40	TIMP3	TIMP3	MMP9	MMP9		
41	TIMP3	TIMP3	MMP13	MMP13		
42	TIMP4	TIMP4	MMP1	MMP1		
43	TIMP4	TIMP4	MMP2	MMP2		

No.	Protein 1	Protein 1 Gene Name	Protein 2	Protein 2 Gene Name	PDB Code	Notes
44	TIMP4	TIMP4	MMP3	MMP3		
45	TIMP4	TIMP4	MMP9	MMP9		
46	TNFA	TNF	TNFR	TNFSFR1A TNFSFR1B	3ALQ	TNFR is Heterodimeric

In Table 4.2, PDB Codes are provided for bound complexes containing the two protein structures shown. Where experimental data can only be found for either Protein 1 or Protein 2 within the interaction, the PDB code is appended with (1) or (2).

Table 4.3: Oligomers involved in atherosclerosis model

Protein Name	Oligomeric Status	Gene Names
Interferon-Gamma Receptor	Heterodimer	IFNGR1 IFNGR2
Interferon-Gamma	Homodimer	IFNG
Interleukin 12	Heterodimer	IL12A IL12B
Interleukin 12 Receptor	Heterodimer	IL12RA IL12RB
Interleukin 17	Homodimer	IL17A
Interleukin 18 Receptor	Heterodimer	IL18R1 IL18RAP
Transforming Growth Factor Beta	Homodimer	TGFB1
Transforming Growth Factor Beta Receptor	Heterodimer	TGFBR1 TGFBR2
Tumor Necrosis Factor Alpha Receptor	Heterodimer	TNFRSF1A TNFRSF1B

Amino acid sequences were identified for all relevant genes within the phase 3 study of the 1000 Genomes Project. Due to a lack of information defining which splice variants are involved in atherogenesis, all protein forming splice variants were included in the data set. Data from the 1000 Genome Project can be accessed through Ensembl, a genome database and browser that contains, displays, annotates and analyses data such as gene sequence, variation and homology (Yates et al., 2016). The Ensembl API can be used to obtain direct access to an underlying MySQL database, allowing for the development of a software pipeline to generate protein structures for a collection of mutations from a gene name (Yates et al., 2014).

Due to concerns about computational intensity for work done within chapters 4 and 5 within this thesis, we restricted the dataset to variants that exist in more than the following percentage of the population:

$$p = \frac{L + 200}{1200} * 0.02$$

Where: p = probability of mutation existing within population; L = Amino Acid sequence length

This equation was generated to allow for a sufficient spread of mutations across each of the relevant proteins while limiting our total computational time to under 300,000 core hours. A list of all sequences isolated can be found in Appendix 4.

Tertiary protein structures were predicted using I-TASSER 4.4 on the Irish Centre for High-End Computing (ICHEC) Fionn cluster – across 240 2.4GHz Intel Cores with 64GB of RAM. I-TASSER was run with the following parameters set:

Table 4.4: I-TASSER Parameters

-runstyle	'parallel'	Allows for multithreading
-nmodel	5	Returns 5 different models for assessment
-hours	30	Limits each internal modelling simulation to 30 hours. N.B. This does not limit total simulation time to 30 hours.

I-TASSER will provide multiple structures (in the form of .pdb files) and other predictions at the end of a simulation, e.g. secondary structure, solvent accessibility, gene ontology terms. For each I-TASSER run, multiple structural confirmations are created and subsequently clustered together to then create a final prediction. Multiple clusters are then created, and the most accurate prediction is then isolated through the use of a scoring function. For more information on how the global quality assessment (known as the C-Score) is calculated, see (Zhang, 2008).

The prediction with the highest C-Score was isolated, and treated as our predicted tertiary structure. Proteins isolated in this manner were named in the format: {Gene Name} – {Splice Variant} – {Mutation Location}. For example, CCL2-001-64 is a protein structure representing the splice variant CCL2-001 found within the 1000 Genome Project database with a mutation causing a change in the 64^{th} amino acid.

4.2.2 Alignment

TM-Align is used to compare monomeric protein structures (Zhang and Skolnick, 2005). Experimentally validated structures as detailed in Table 4.2 have been isolated from the protein data bank, and I-TASSER structures as predicted from the wildtype have been aligned to these experimentally validated structures. An observed accuracy value is calculated by utilising root-mean-square-deviations (RMSDs) and TM-Scores, obtained while performing this alignment using TM-Align. I-TASSER will predicted values for both RMSD and TM-Score to assess model quality, and these values are compared to our observed accuracy. An initial alignment is chosen by aligning secondary structure elements (helices, sheets or coils) to maximise the TM-Score, an evaluation of similarity between two tertiary structures that prioritises alignment in the centre of the protein due to the increased likelihood of error and variance in the protein terminals. A TM-Score of greater than 0.5 suggests correct topology when comparing structure, while a TM-Score lower than 0.17 is almost ensured to be a random correlation (Xu and Zhang, 2010).

RMSDs and TM-Scores calculated in this way are used to evaluate the error in the predicted protein structures, and to give confidence in the quality assessments of I-TASSER where native structures are unavailable.

4.2.3 Accuracy

I-TASSER provides a local per residue ‘L-score’ as well as a global C-score to quantify prediction accuracy. An L-score has been calculated for each protein model using secondary structure information and predictions of solvent accessibility to give a per-residue maximal RMSD (in Angstrom) from the native structure. For more information on the ResQ algorithms used to calculate the residue specific accuracy, see (Yang et al., 2016). To study the accuracy of these local scoring mechanisms, each model with an experimentally

validated structure within the PDB as described in Table 4.2 has been aligned to this PDB structure, and a per-residue RMSD has been calculated to study the quality of these L-scores.

A collection of graphs have been generated to compare this L-Score value to an observed accuracy, shown in Figure 4.6. For each wildtype structure with a known experimentally-derived structure in the PDB defined in Table 4.2, TM-Align was used to align the predicted and the experimental structure and a per-residue pairwise RMSD was calculated to evaluate the predicted L-score. For structures where the alignment was non-continuous, the largest continuous sequence was graphed.

Once the level of accuracy of the L-scores has been ascertained, a heatmap has been created to display prediction accuracy. Selected heatmaps can be found in Figure 4.5, and heatmaps for each protein structure calculated can be found in Appendix 4. Heatmaps were generated in Jmol (The Jmol Team, 2007) using the commands:

background white; ribbon; cartoons only; color property temperature "low" range 0 25; write image filename.png

These commands provide a white background, remove individual atoms, show the protein secondary structure in cartoon form, colour the proteins from red to yellow over a range of 0 Angstrom to 25 Angstrom, and write the image to a file respectively.

4.2.4 Variance Heatmap

Once a collection of PDB files had been generated for all included mutations, we wanted to compare our structures. TM-Align was used for each mutation to provide an optimal superposition with its corresponding wildtype. TM-Align will align two proteins by using a variation of the Levitt-Gerstein weight factor, weighting closer residue pairs more than further distant pairs. This prevents large variation at terminals, a common occurrence in protein prediction, from suggesting that a protein is a less accurate prediction that it actually is (Zhang and Skolnick, 2005).

A per-residue root-mean-square-deviation (RMSD) is calculated for each alignment with the wildtype. For each residue, the standard error of the mean (SEM) is calculated for all of the

mutations associated with a particular splice variant of a protein. These SEM values are then heatmapped onto the wildtype of the protein, giving a coloured description of the protein areas most likely to be highly varied.

The primary aim of this chapter is to study how protein structures involved in atherosclerosis vary between population subgroups. A heatmap image has been generated for each splice variant defined in Table 4.6. For each of the protein structures isolated as described in section 4.2.1, an alignment was performed with the wildtype using TM-Align. Using the optimal alignment as calculated with this method, an RMSD value was calculated for each pair of residues. The SEM for each of these per-residue RMSDs was taken.

A heatmap was generated by manually replacing column 10 (b-factor or temperature factor) in the pdb file for the wildtype with this SEM value, and utilising Jmol with the commands:

background white; ribbon; cartoons only; color property temperature "bgyor" range 0 50; write image filename.png

These commands provide a white background, remove individual atoms, show the protein secondary structure in cartoon form, colour the proteins from blue to green to red over a range of 0 Angstrom to 50 Angstrom, and write the image to a file respectively. The precise colour scheme used can be found in Figure 4.6.

This has allowed us to generate a heatmap showing the residues and locations within the protein with the highest variance.

4.2.5 Computational Intensity

Using I-TASSER to predict a tertiary structure from an amino acid sequence is a computationally intensive task. Due to resource constraints, total computational time was estimated by relating amino acid sequence length to computational time. Before the entire data set of amino acids was processed to produce tertiary structures, thirty-three structures were predicted using I-TASSER to allow us to estimate the total time required. Sequences were included in this alpha test if their sequence length was:

AALength = k * 25 where k = 1,2,3,4… and AALength between 50 – 500

AALength = k * 50 + 500 where k = 1,2,3,4… and AALength between 500 – 1200

Tertiary protein structures were predicted using I-TASSER 4.4 on the Irish Centre for High-End Computing (ICHEC) Fionn cluster – across 240 2.4GHz Intel Cores with 64GB of RAM.

Sequences of the exact length required were not available with our preliminary dataset, so in those cases the closest available sequence length was taken. The sequences predicted during this benchmarking test are detailed in Table 4.5:

Table 4.5: Protein Structures used in I-TASSER benchmarking test

Sequence Name	Sequence Length	Computational Time (Core Hours)
PDGFC-002-14	50	12.9
OLR1-005-45	77	17.8
CCL2-001-69	99	16.8
CXCL9-001-1	125	23.5
CD40-201-124	151	31.6
IFNG-001-162	166	42.1
CD40-002-8	203	59.1
TIMP4-001-19	224	63.5
CMA1-001-1	247	78.6
TPSB2-001-18	275	96.3
GLG1-007-139	305	226.2
TPSG1-001-24	321	166.9
CXCR1-001-1	350	188.2
CCR2-201-233	374	216.2
EGFR-005-13	405	110.1
TGFBR1-003-232	426	180.1
LPL-001-20	475	150.5
TGFBR1-001-17	503	345.9
ICAM1-001-39	532	281.4
TGFBR2-002-522	592	424.4
LCP1-006-319	627	466.6
VCAM1-002-11	647	492.7
ABCG1-005-28	677	706.6
EGFR-002-703	705	380.7
VCAM1-001-11	739	590.1
PLA2G6-002-258	752	727.0
SELP-006-19	768	757.4
PLA2G6-001-53	806	764.2
IL4R-001-82	825	746.6

Sequence Name	Sequence Length	Computational Time (Core Hours)
PDGFRA-001-1000	1089	817.1
EGF-003-892	1166	771.8
GLG1-003-80	1192	771.2
EGFR-001-458	1210	861.8

N.B. Some of the sequences involved in this alpha test were not included in the final version of the model, due to the preliminary status of the model at the time this work was completed.

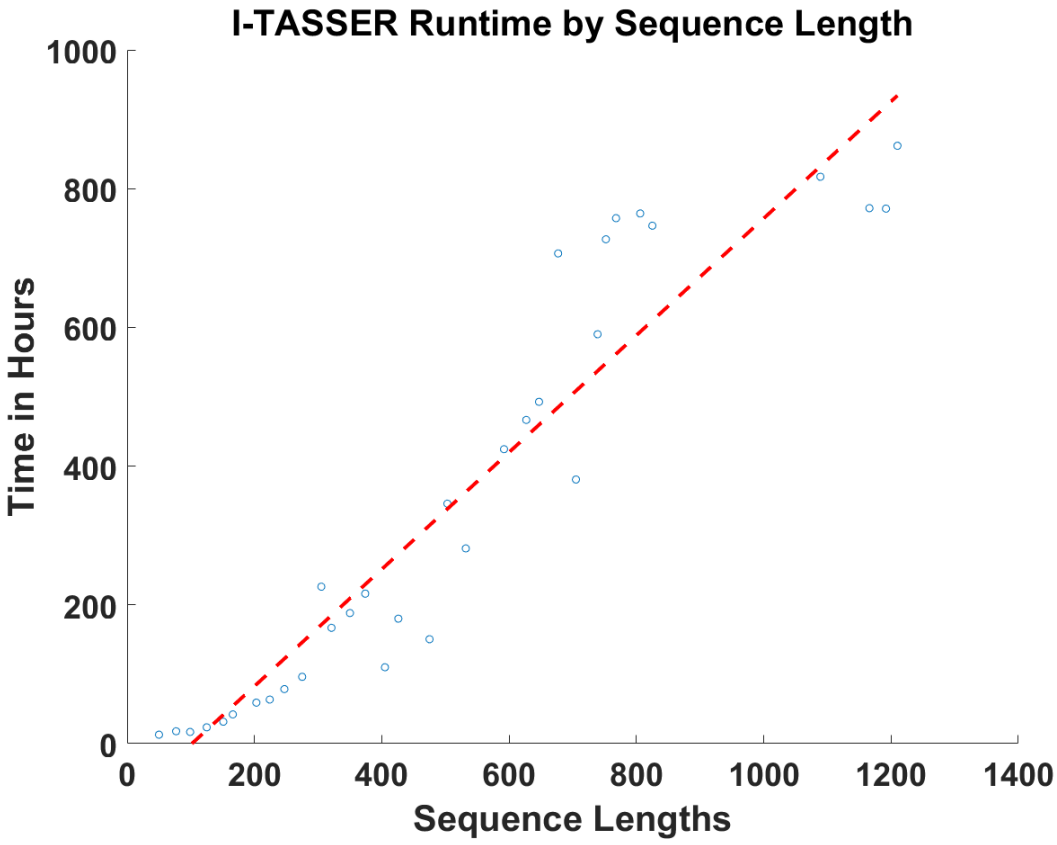


Figure 4.2: Comparing I-TASSER runtime to sequence length

4.3 Results

830 amino acid sequences were isolated from the 1000 Genome Project Phase 3 MySQL database, over a total of 65 genes and 146 splice variants, based on the gene names described in Table 4.2 and formula described in section 4.2.1. I-TASSER was used to predict a tertiary protein structure for each of these protein structures in the form of a .pdb file.

Sequence lengths range from 37 to 1210 amino acids long with a mean of 388.29 (median 247).

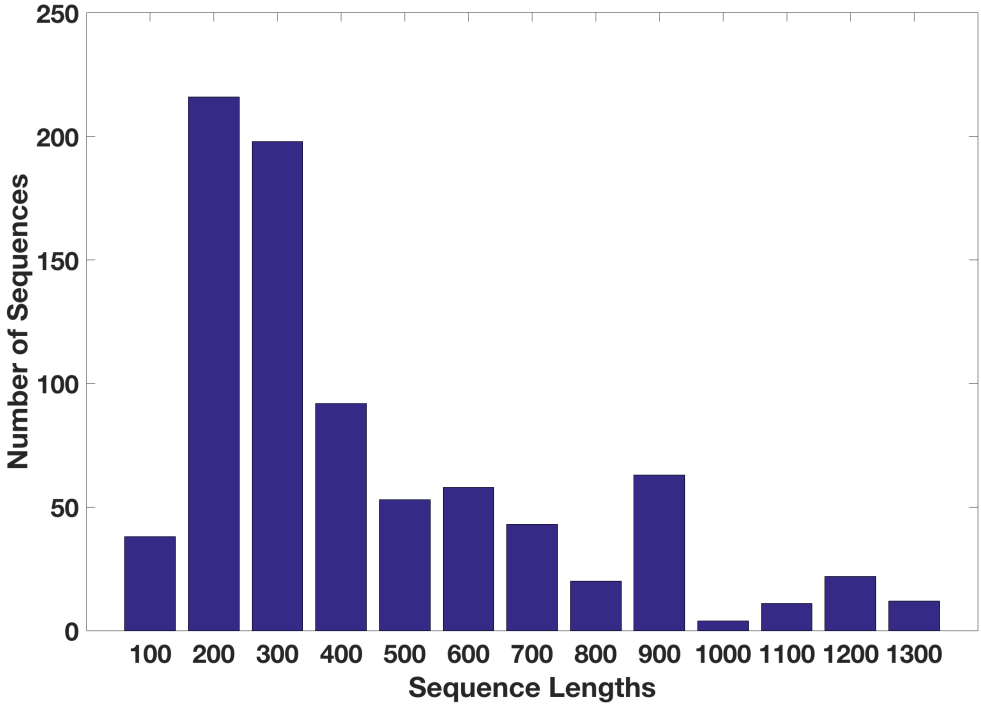


Figure 4.3: Distribution of Sequence Lengths isolated from Phase 3 of the 1000 Genome Project

Table 4.6: All isolated splice variants and corresponding number of structures, including mutations

Splice Variant Name	Number of Structures	Splice Variant Name	Number of Structures	Splice Variant Name	Number of Structures
ABCA1-001	5	IL12A-001	5	IL5RA-003	4
ABCA1-004	3	IL12A-005	14	IL5RA-004	4
ABCG1-001	1	IL12A-006	13	IL5RA-005	4
ABCG1-002	1	IL12B-001	9	IL5RA-006	4
ABCG1-004	1	IL12RB1-001	9	IL5RA-007	4
ABCG1-006	1	IL12RB1-002	9	IL5RA-201	5
ADAM17-001	5	IL12RB1-003	8	IL6-001	12
CCL2-001	3	IL12RB2-001	6	IL6-003	8

Splice Variant Name	Number of Structures	Splice Variant Name	Number of Structures	Splice Variant Name	Number of Structures
CCL2-004	1	IL12RB2-004	8	IL6-004	6
CCL5-001	6	IL12RB2-201	7	IL6-005	11
CCL5-002	6	IL12RB2-202	8	IL6-006	14
CCR2-001	3	IL17A-001	8	IL6-201	12
CCR2-002	2	IL17RA-001	5	IL6R-001	4
CCR2-201	3	IL17RA-201	4	IL6R-003	2
CCR5-001	4	IL18-001	6	IL6R-201	2
CD36-001	4	IL18-003	6	MMP1-001	3
CD36-006	4	IL18-006	6	MMP13-001	2
CD36-019	1	IL18R1-001	4	MMP13-002	2
CMA1-001	14	IL18R1-201	4	MMP13-201	2
CMA1-002	16	IL18R1-202	7	MMP2-001	2
CSF1-001	7	IL18RAP-001	1	MMP2-008	4
CSF1-002	7	IL18RAP-002	2	MMP3-001	3
CSF1-003	4	IL1B-001	1	MMP9-001	7
CSF1-004	5	IL1R1-001	4	PDGFA-001	12
CSF1R-001	4	IL1R1-006	4	PDGFA-002	14
CSF1R-201	10	IL1R1-013	3	PDGFB-001	11
CXCL10-001	9	IL2-001	5	PDGFB-002	17
CXCL11-001	4	IL21-001	3	PDGFRA-001	8
CXCL11-003	4	IL21-201	3	PDGFRA-002	1
CXCL9-001	6	IL21R-001	3	PDGFRB-001	5
CXCR3-001	1	IL21R-003	3	PLA2G6-001	3
CXCR3-002	1	IL21R-006	3	PLA2G6-002	1
EGF-001	9	IL2RA-001	2	TGFB1-001	4
EGF-002	9	IL2RA-002	13	TGFBR1-001	3
EGF-003	7	IL2RA-004	5	TGFBR1-003	1
EGFR-001	3	IL33-001	6	TGFBR2-001	2
EGFR-002	4	IL33-201	6	TGFBR2-002	2
EGFR-003	2	IL33-202	12	TIMP1-001	14
EGFR-004	3	IL33-203	12	TIMP1-002	11
EGFR-005	1	IL4-001	11	TIMP1-003	10
EGFR-201	2	IL4-002	10	TIMP2-001	4
EGFR-202	1	IL4-201	8	TIMP2-003	3
IFNG-001	4	IL4R-001	13	TIMP2-008	6
IFNGR1-001	5	IL4R-004	13	TIMP3-001	8
IFNGR1-201	5	IL4R-201	14	TIMP4-001	11
IFNGR2-001	4	IL4R-202	1	TNF-001	2
IL10-001	8	IL5-001	8	TNFSFR1A-001	3
IL10RA-001	7	IL5RA-001	4	TNFSFR1B-001	5

Splice Variant Name	Number of Structures	Splice Variant Name	Number of Structures	Splice Variant Name	Number of Structures
IL10RB-001	5	IL5RA-002	4		

4.3.1 Global Accuracy

Each I-TASSER run provides up to five models of each structure based on the sequence given. The C-Score is calculated for each of these models as described in section 4.2.1.

Of these 830 sequences, the mean C-score is -1.4134, the maximum is 1.55 and the minimum is -5. These structures have a mean TM-score and RMSD of 0.552 +- 0.127 and 9.71 +-3.75 Angstrom respectively. The model returned with the highest C-score is chosen as the final prediction.

Figure 4.4: Images of CCL2-001 (left) and CCR2-001 (right) generated by I-TASSER, displayed in ribbon format using JMol

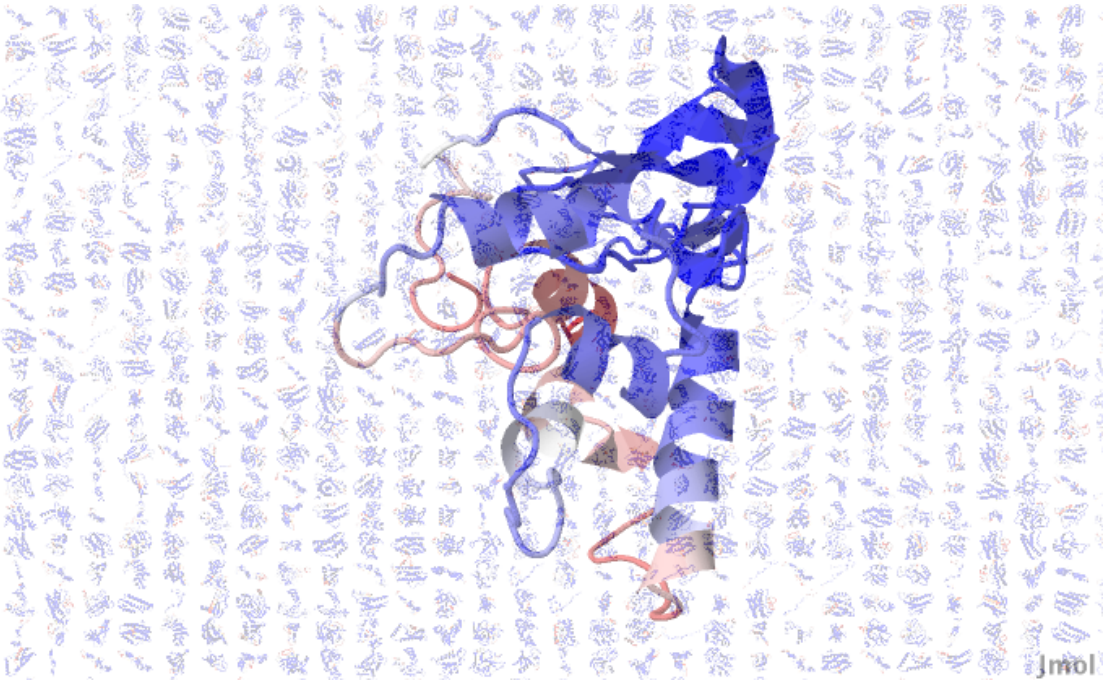


Figure 4.5: Accuracy heatmap of PDGFA-001 on a mosaic of accuracy heatmaps from all structures within our dataset

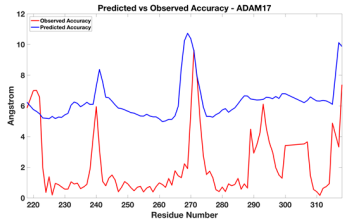
Heatmaps to demonstrate the residue specific quality (or L-score) of the predictions were generated as described in section 4.2.3 and are shown in collectively in Figure 4.5 and individually in Appendix 4.

4.3.2 Error

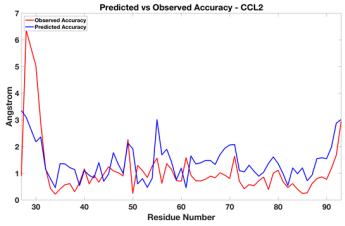
A subset of the generated wildtype models can be compared to experimentally observed structures. For each of the proteins with a PDB code included in Table 4.2, TM-Align was used to calculate an optimal alignment between the model and the experimental structure, and the distance was calculated for each pairwise residue to provide an observed accuracy value. Of the 65 different genes that we have predicted protein structures for, 53 contain experimentally derived complexes within the PDB. 53 pair-wise alignments were performed with the experimentally validated structures and the predicted ones, with a mean TM-Score of 0.558. Twenty-six structures have a TM-score higher than 0.5 (suggesting correct topology), and none have a TM-score lower than 0.17 (suggesting a random correlation) (Xu and Zhang, 2010).

The graphs in Figure 4.6 show the pairwise accuracy between residues as defined in section 4.2.3.

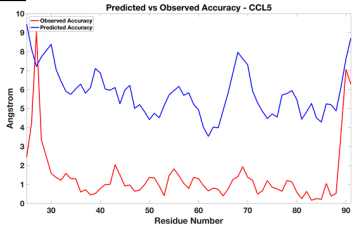
ADAM17



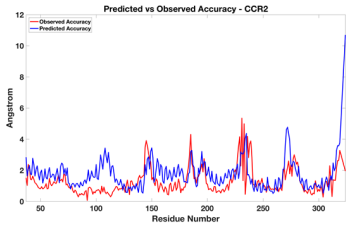
CCL2



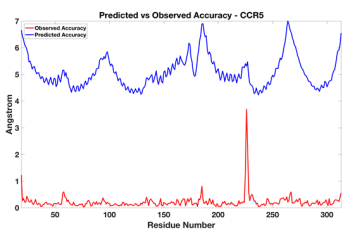
CCL5



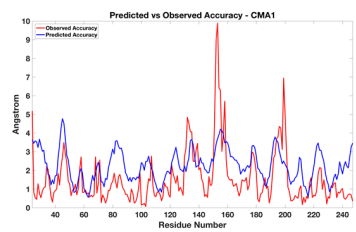
CCR2



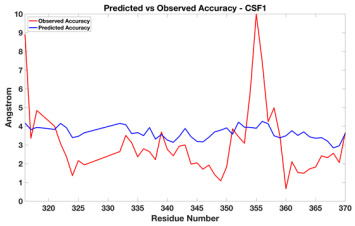
CCR5



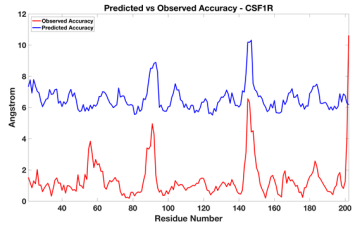
CMA1



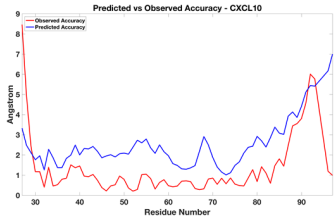
CSF1



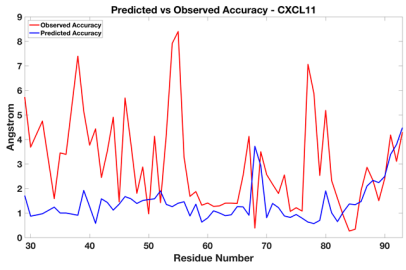
CSF1R



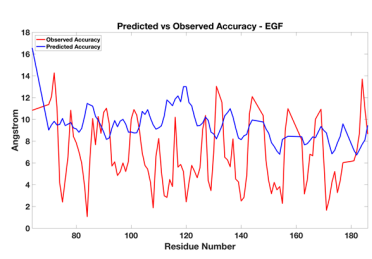
CXCL10



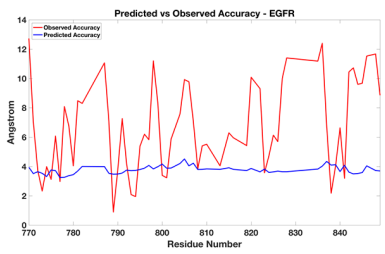
CXCL11



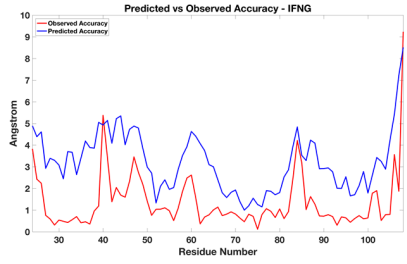
EGF



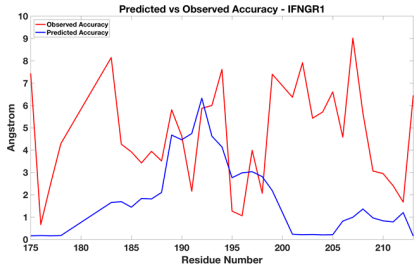
EGFR



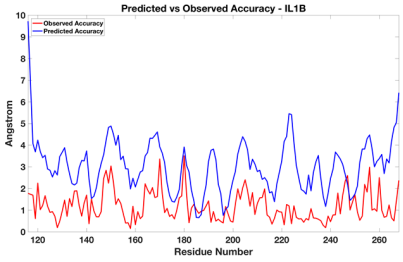
IFNG



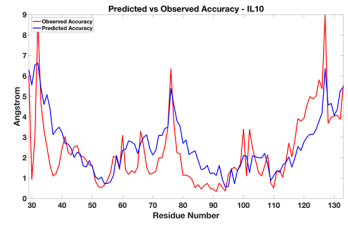
IFNGR1



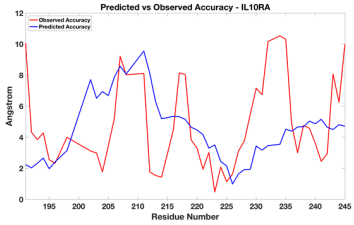
IL1B



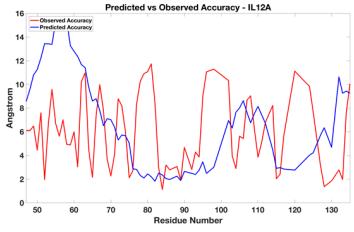
IL10



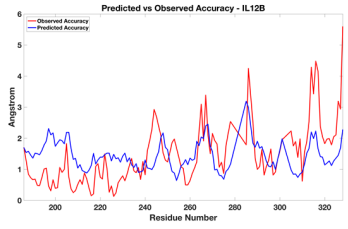
IL10RA



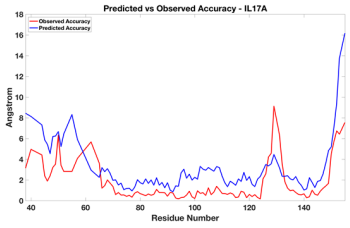
IL12A



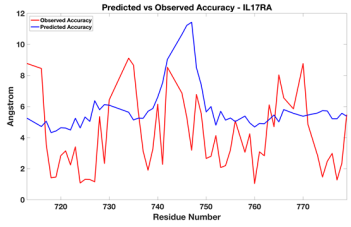
IL12B



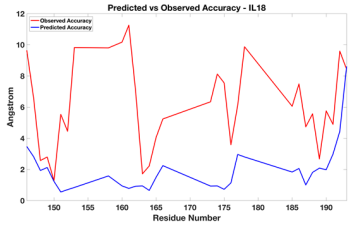
IL17A



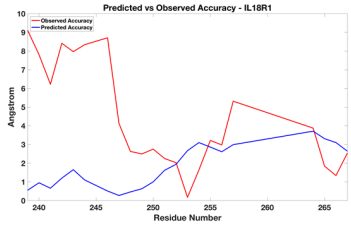
IL17RA



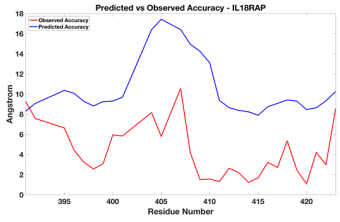
IL18



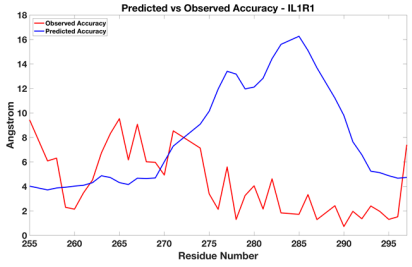
IL18R1



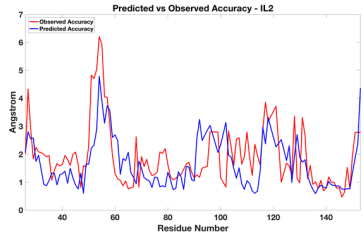
IL18RAP



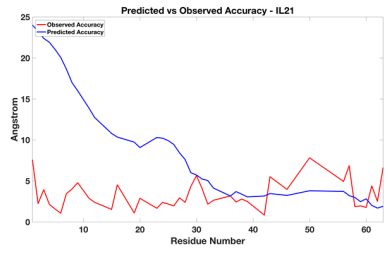
IL1R1



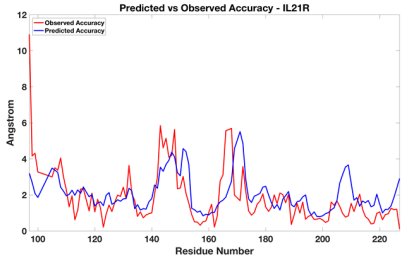
IL2



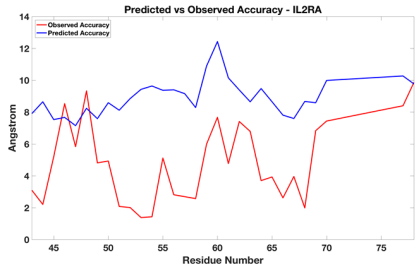
IL21



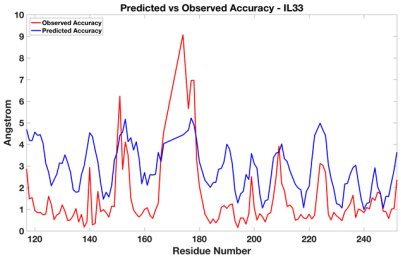
IL21R



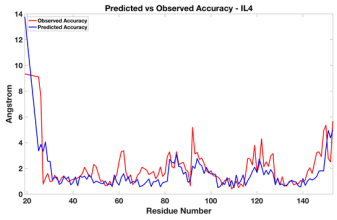
IL2RA



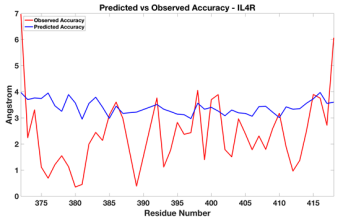
IL33



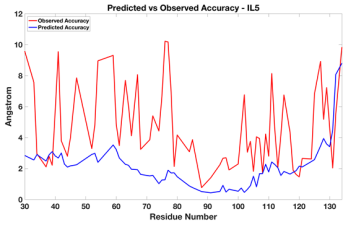
IL4



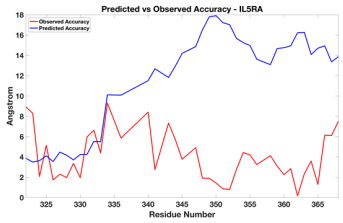
IL4R



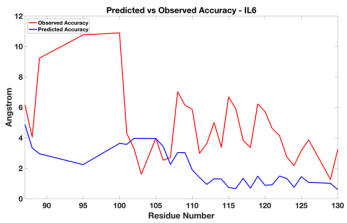
IL5



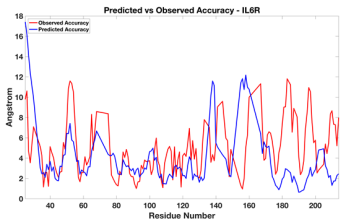
IL5RA



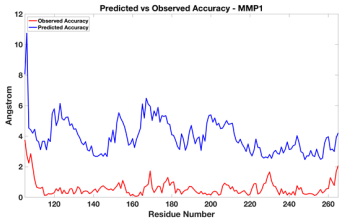
IL6



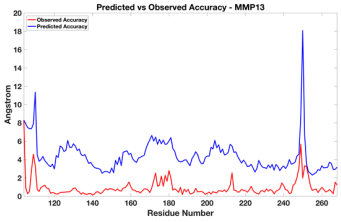
IL6R



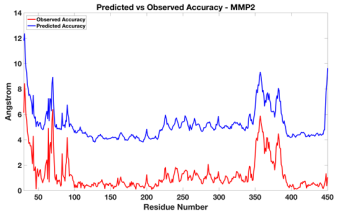
MMP1



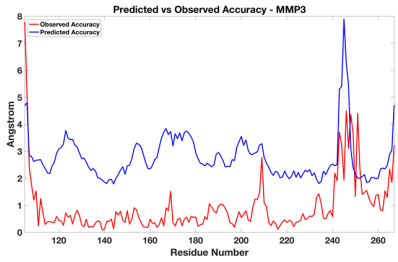
MMP13



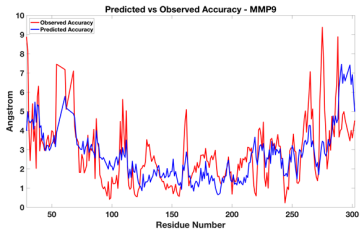
MMP2



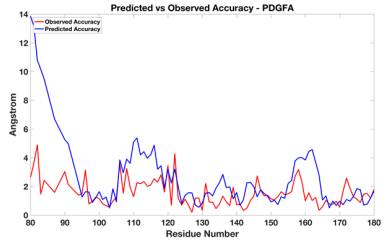
MMP3



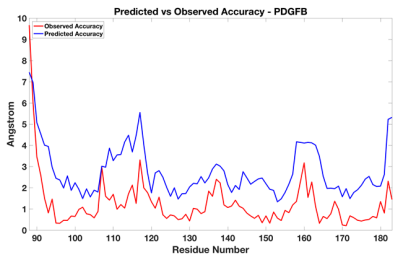
MMP9



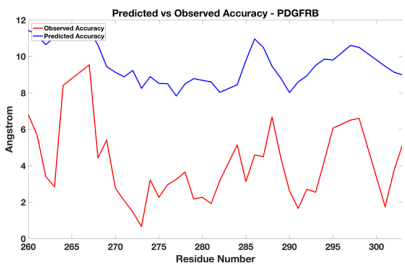
PDGFA



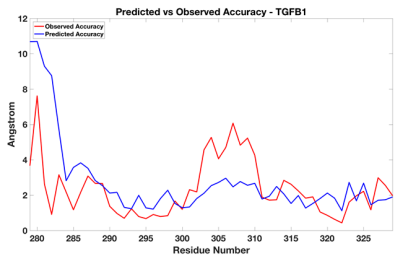
PDGFB



PDGFRB



TGFB1



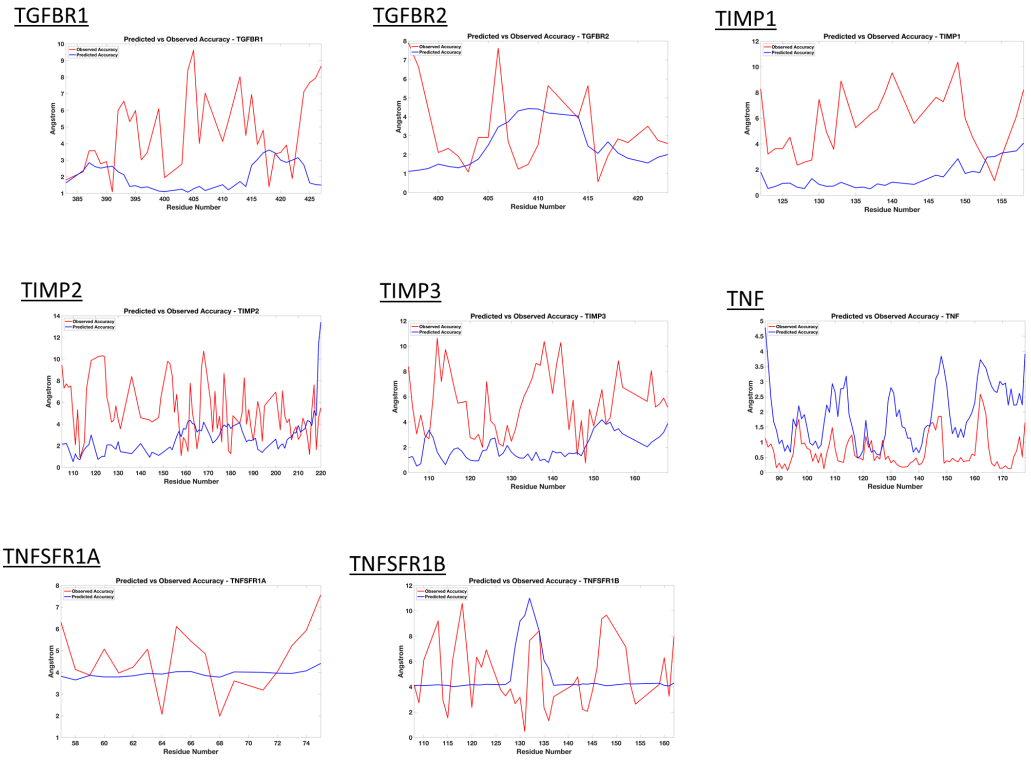


Figure 4.6: 53 graphs comparing observed accuracy (red) to predicted accuracy (blue)

TM-Scores for each of these alignments are shown in Table 4.7.

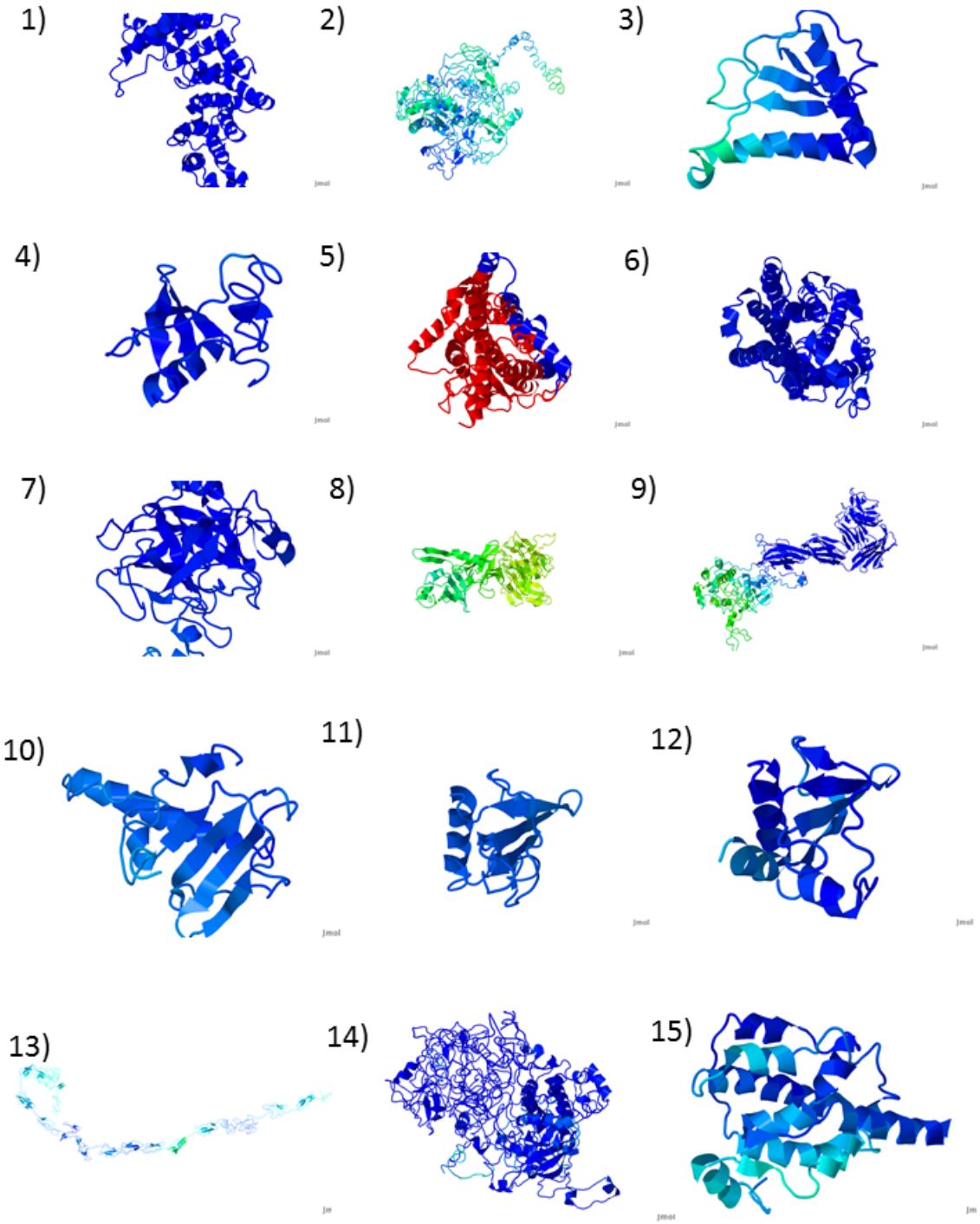
Table 4.7: TM-Scores for alignments between predicted and experimentally derived wildtypes

Protein Name	TM-Score	Protein Name	TM-Score
ADAM17-001	0.74	IL21R-001	0.54
CCL2-001	0.87	IL2RA-001	0.37
CCL5-001	0.80	IL33-001	0.89
CCR2-001	0.73	IL4-001	0.83
CCR5-001	0.85	IL4R-001	0.33
CMA1-001	0.91	IL5-001	0.30
CSF1-001	0.40	IL5RA-001	0.34
CSF1R-001	0.63	IL6-001	0.23
CXCL10-001	0.79	IL6R-001	0.45
CXCL11-001	0.49	MMP1-001	0.98
EGF-001	0.22	MMP13-001	0.95

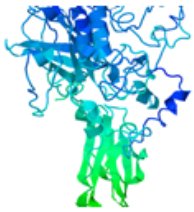
Protein Name	TM-Score	Protein Name	TM-Score
EGFR-001	0.36	MMP2-001	0.96
IFNG-001	0.60	MMP3-001	0.92
IFNGR1-001	0.33	MMP9-001	0.79
IL10-001	0.55	PDGFA-001	0.86
IL10RA-001	0.33	PDGFB-001	0.87
IL12A-001	0.22	PDGFRB-001	0.35
IL12B-001	0.94	TGFB1-001	0.68
IL17A-001	0.82	TGFBR1-001	0.34
IL17RA-001	0.33	TGFBR2-001	0.31
IL18-001	0.21	TIMP1-001	0.30
IL18R1-001	0.31	TIMP2-001	0.34
IL18RAP-001	0.26	TIMP3-001	0.28
IL1B-001	0.93	TNF-001	0.97
IL1R1-001	0.23	TNFSFR1A-001	0.27
IL2-001	0.82	TNFSFR1B-001	0.25
IL21-001	0.26		

4.3.3 Heatmaps

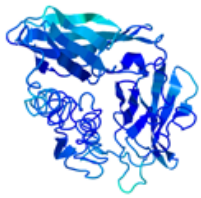
Of the sequences isolated as described in section 4.2.1, mutations are available for 61 proteins. A heatmap was generated for each of these proteins as described in section 4.2.4.


1)
2)
3)
4)
5)
6)
7)
8)
9)
10)
11)
12)
13)
14)
15)
0 Å
25 Å

16)



17)



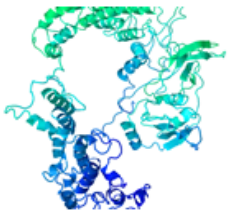
18)

19)

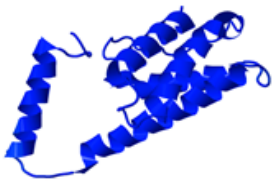
20)

21)

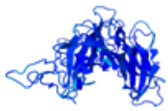
22)



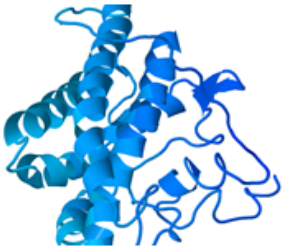
23)



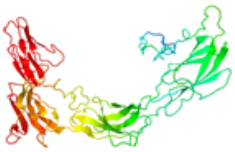
24)



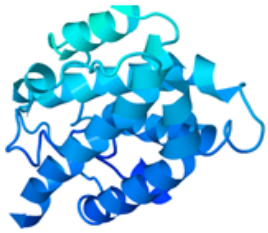
25)



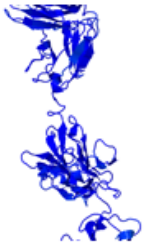
26)



27)

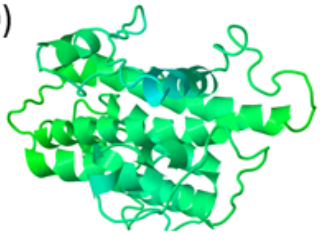


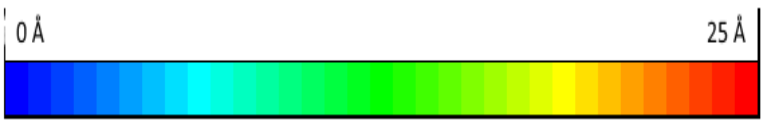
28)



29)

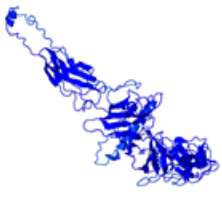
30)



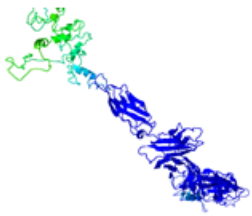


31)

32)

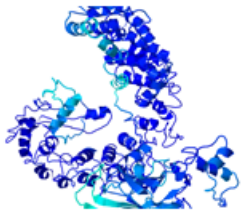


33)

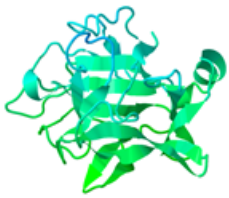


34)

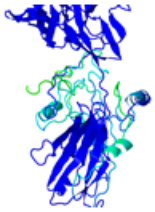
35)



36)



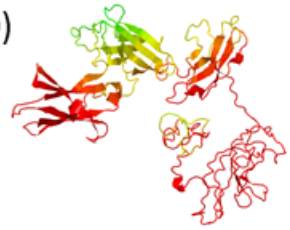
37)



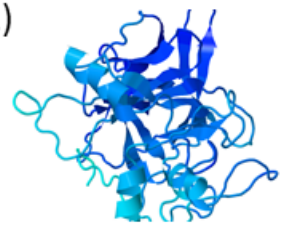
38)

39)

40)

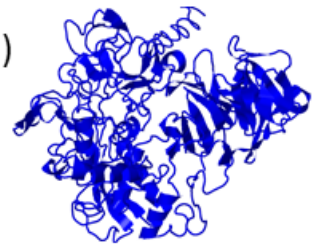


41)

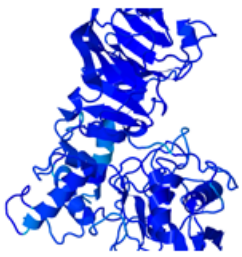


42)

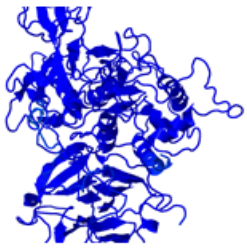
43)

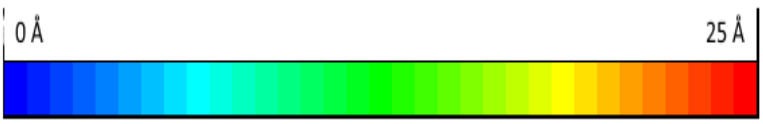


44)



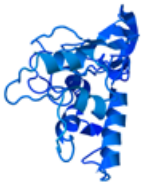
45)



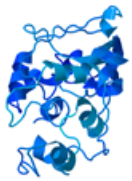


46)

47)



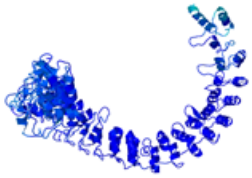
48)



49)

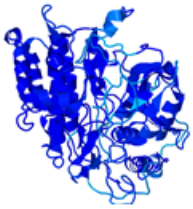
50)

51)



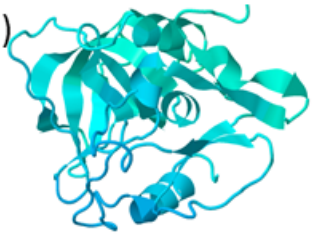
52)

53)

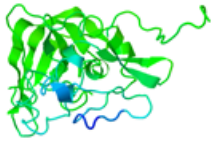


54)

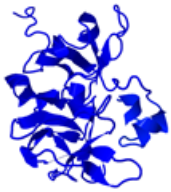
55)



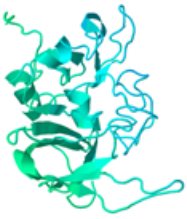
56)



57)

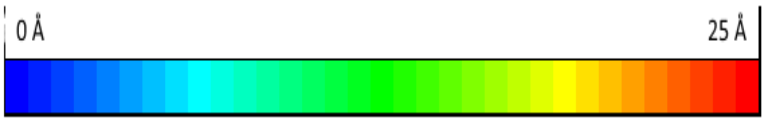


58)



59)

60)



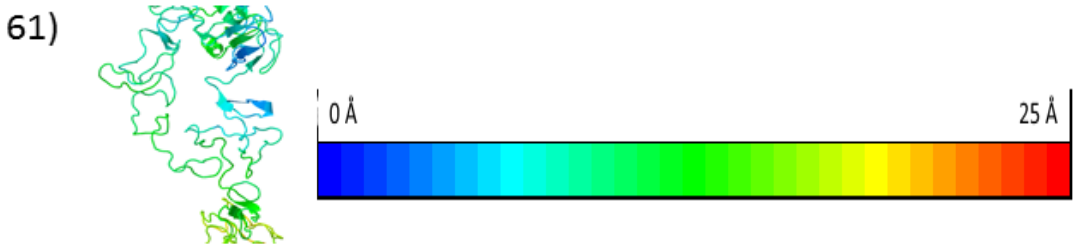


Figure 4.7: 61 heatmaps displaying the variance within protein structures across a population.

Table 4.8: Corresponding gene names for the heatmaps shown in Figure 4.7

Heatmap No	Gene Name	Heatmap No	Gene Name
1	ABCA1	32	IL12RB1
2	ADAM17	33	IL12RB2
3	CCL2	34	IL17A
4	CCL5	35	IL17RA
5	CCR2	36	IL18
6	CCR5	37	IL18R1
7	CMA1	38	IL18RAP
8	CSF1	39	IL21
9	CSF1R	40	IL21R
10	CXCL9	41	IL33
11	CXCL10	42	MMP1
12	CXCL11	43	MMP2
13	EGF	44	MMP3
14	EGFR	45	MMP9
15	IFNG	46	MMP13
16	IFNGR1	47	PDGFA
17	IFNGR2	48	PDGFB
18	IL1R1	49	PDGFRA
19	IL2	50	PDGFRB
20	IL2RA	51	PLA2G6
21	IL4	52	TGFB1
22	IL4R	53	TGFBR1
23	IL5	54	TGFBR2
24	IL5RA	55	TIMP1
25	IL6	56	TIMP2
26	IL6R	57	TIMP3
27	IL10	58	TIMP4
28	IL10RA	59	TNF
29	IL10RB	60	TNFSFR1A

30	IL12A	61	TNFSFR1B
31	IL12B		

With this gradient colour scheme, blue areas of the protein show the least variance, while green/red sections correspond to increased variance across the population within these areas. For each protein heatmap, the range and SEM of temperature values were calculated to study which proteins have particular areas where structural variation has an increased effect. Of the 61 proteins available, 12 have a mean SEM greater than 15 Angstrom, due to certain areas of the protein with higher structural variance than others. Five proteins have a maximum variance greater than 50 Angstrom, leading to the general red colour. The protein showing the most variance with this methodology is CCR2 (heatmap number 5, mean SEM 89.66), showing significant structural variation across isolated mutations.

4.4 Discussion

4.4.1 Mutations linked to disease

DbSNP, a database connecting variations to clinical data and other sequence resources, and ClinVar, a repository linking variations to phenotype, were searched for each mutation isolated from the 1000 Genome Project to isolate variants with known clinical significance to disease (Landrum et al., 2014; Sherry et al., 2001).

IL4R-001-75 (rs1805010) has multiple known clinical outcomes. This polymorphism can lead to a slow progression to acquired immunodeficiency syndrome (AIDS) (Soriano et al., 2005), and additionally cause resistance to atopy (Franjkovic et al., 2005). Interleukin 4 receptor has been shown to promote differentiation of T_h0 cells into T_h2 cells through IL-4 signalling and to regulate production of the IgE antibody (Katona et al., 1991; Lee and Hirani, 2006).

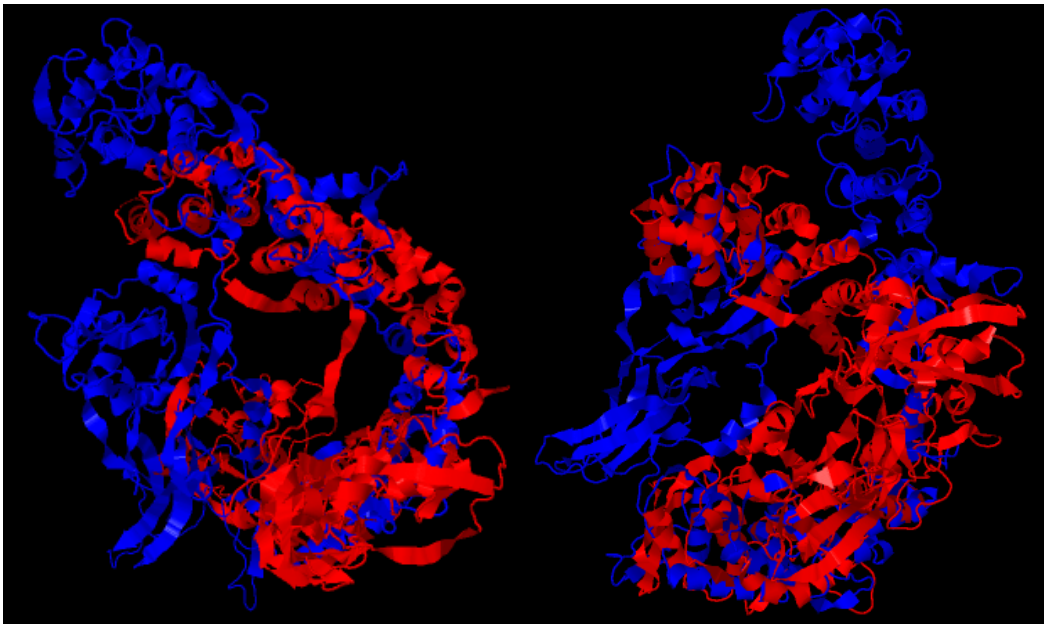


Figure 4.8: (Left) - IL4R-001 (Blue) aligned with IL4R-001-75 (Red)

(Right) - IL4R-001 (Blue) aligned with IL4R-001-503 (Red)

The I-TASSER generated model of IL4R-001 has a C-score of -1.63 and an estimated RMSD of 12.4 +- 4.3 A, so while the model may have altered backbone structure and imperfect side chain conformations when compared to the native structure, we can be confident that the model has the correct topology. For the variant IL4R-001-75, we see a C-Score of -2.14 with an estimated RMSD of 13.8 +- 3.9. As shown in Figure 4.6 (Graph 32), the predicted accuracy of our wildtype is reliable.

The mutation has caused a significant structural difference between models. A TM-Score of 0.32569 has been calculated between the two models, which is a score suggesting that similarity between the two structures is only slightly better than random (Zhang and Skolnick, 2005).

IL4R-001-503 (rs1805015) is another mutation in interleukin 4 receptor which is known to be protective against atopy (Franjkovic et al., 2005). For the variant IL4R-001-503, we see a C-Score of -2.12 with an estimated RMSD of 13.7 +- 4.0, which is remarkably similar to the accuracy scores of IL4R-001-75. We see similar structural significance with the wildtype as IL4R-001-75, with a TM-Score of 0.33986.

However, our models of IL4R-001-75 and IL4R-001-503 show strong structural similarity. With a TM-Score of 0.84648 and an RMSD of 4.16A, we can see that the tertiary structure of IL4R with these mutations is remarkably similar.

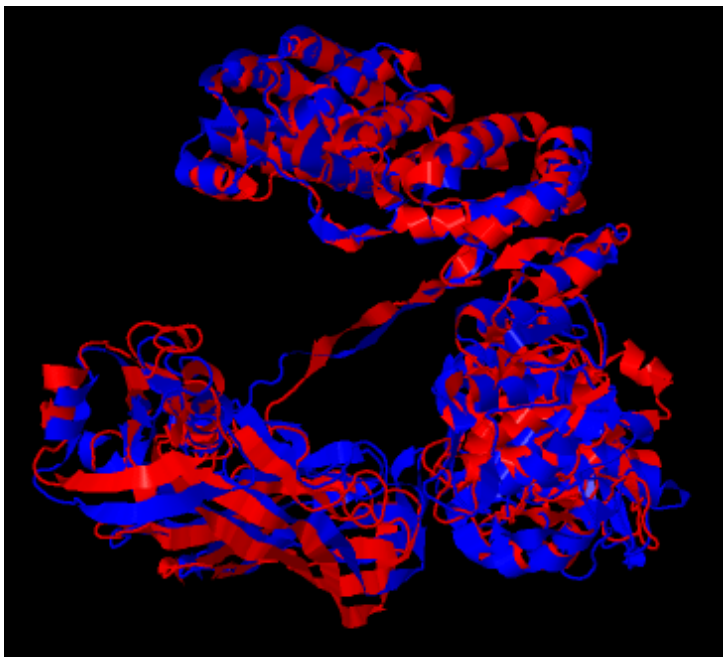


Figure 4.9: Alignment between IL4R-001-75 and IL4R-001-503

We know that existence of either of these mutations within an individual gives a protection against atopy, and that IL4R-001-75 gives a susceptibility to AIDS. With the known link between protein structure and protein function, the structural similarity between these proteins warrants investigation into the function of IL4R-001-503.

ABCA1-001-219 (rs2230806) is a mutation in ATP-binding cassette transporter A1 that has been shown to reduce CHD risk in individuals with familial hypercholesterolemia (Cenarro et al., 2003). ABCA1 is a transporter of various biological entities across cellular membranes which is involved in cholesterol efflux from atheromae. We see similar alignment between our models of ABCA1-001-219 and its wildtype, with a TM-score of 0.83924 and an RMSD of 3.12, indicating that the structural effect of the mutation is minor. However, our C-Scores for the wildtype and mutation are -3.31 and -3.40 respectively, so the quality of the structural prediction may be a factor when comparing these proteins.

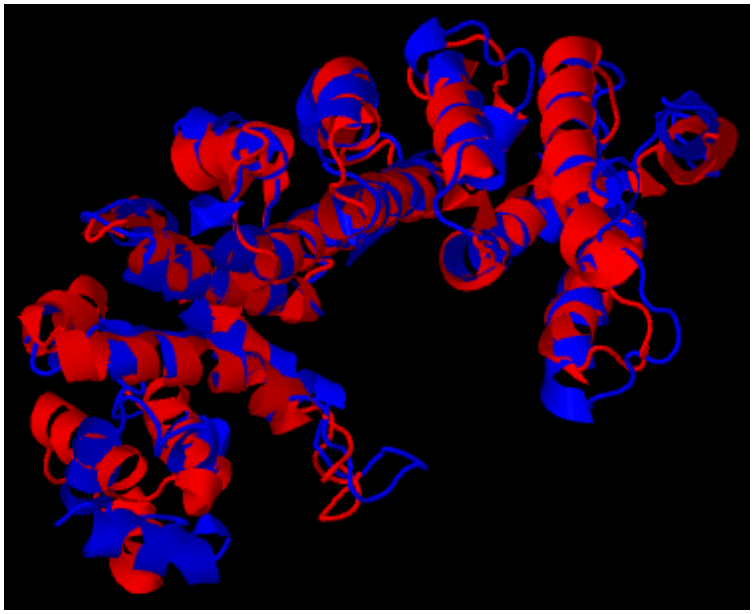


Figure 4.10: ABCA1-001 aligned with the protective mutation ABCA1-001-219

4.4.2 General Discussion

I-TASSER is the current industry standard in protein-structure prediction from amino acid sequences, having been the leader in the community wide CASP experiments 4 times in the last 5 contests, and as such we can have confidence in our results and the ability of I-TASSER to assess the quality of its work. Our alignments to the wildtype show that a strong C-score gives an accurate assessment of model quality and that L-scores are a strong assessment of the per-residue accuracy of a structure. Although there are structures where the I-TASSER assessment of quality is reduced, overall confidence in structural accuracy is justified.

The graphs in Figure 4.6 which relate predicted structural accuracy to observed accuracy show that the L-Score is a strong indicator of quality; however it is not perfect. Of all of the graphs, 10 show a global disparity between observed and predicted accuracy: CXCL11, EGFR, IFNGR1, IL18, IL5, IL6, TGFBR1, TIMP1, TIMP2 and TIMP3. However, each of these structures show a weak alignment to the predicted structure, as only 1 structure (CXCL11 with 0.49) with a TM-Score when aligned to its wildtype of greater than 0.37 shows this disparity. In summary, when we are confident in our global result, we can be confident in our local results.

Similar structure between proteins implies to similar function. By assessing which individual mutations cause the greatest variations in protein structures, we can isolate a collection of structures that are more likely to have different function than the wildtypes. If we can identify which mutations are more likely to cause variation in active sites of proteins, this procedure could be refined further.

Relating sequence length to model quality can be done in multiple ways. By relating the sequence length to the C-Score of the structure, we see no correlation:

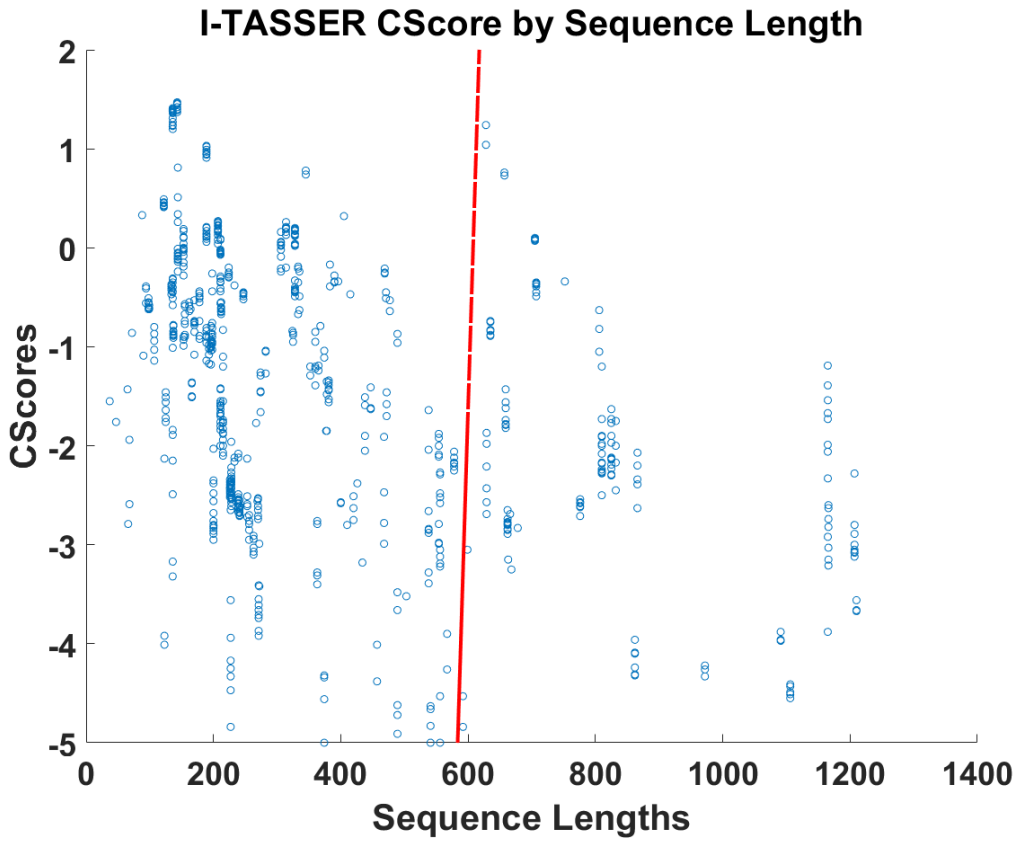


Figure 4.11: Comparison of C-Score and Sequence Length

However, by comparing the predicted RMSD to the sequence length, we do see a correlation:

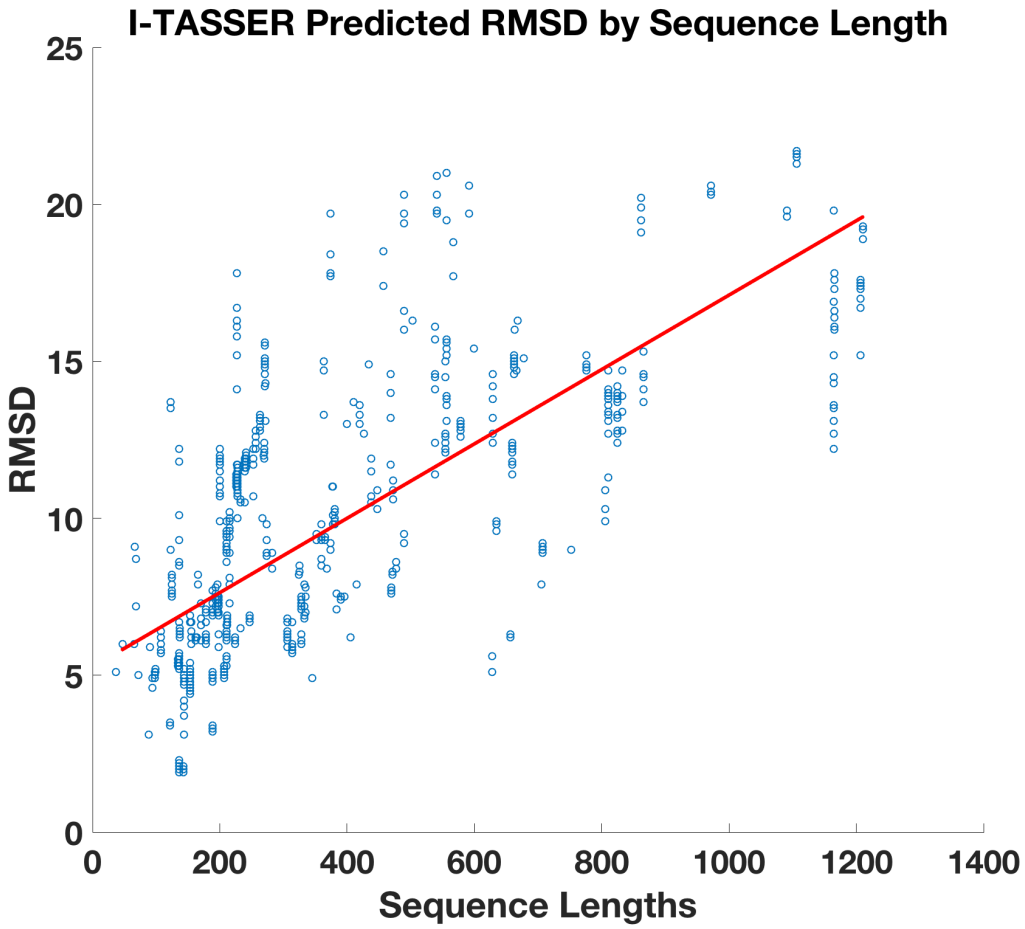


Figure 4.12: Comparison of RMSD and Sequence Length

The line of best fit drawn in Figure 4.11 and Figure 4.12 was drawn using Matlab's polyfit function. This is due to the inclusion of scoring functions from I-TASSER's subprocess threading algorithms which incorporate sequence length. However, I-TASSER studies have shown a correlation between C-Score and RMSD of 0.75 (Yang et al., 2014). In summary, structural quality can be assessed in multiple different ways and care should be taken to ensure awareness of the limitations of solely using one metric to study quality.

What qualities in a protein lead to a stronger more accurate prediction? With I-TASSER's reliance on homologue discovery to improve the results from threading algorithms, it is assumed that the number of homologues identified is related to prediction quality. However, only approximately 50% of known sequences can be shown to be related to a known experimentally derived protein structure (Fiser, 2017). *Ab initio* algorithms have

been shown to be more reliable with smaller proteins in cases where homologues are unavailable. Identification of known folds within the primary structure also provides a high quality prediction through threading methods. However, threading of individual segments has been shown to introduce significant noise into resulting models.

Variance heatmaps were generated using the standard error of the mean (SEM) rather than standard deviation, as this allowed us to account for small numbers of mutations for certain structures. Five heatmap models show a red section (an area where the pairwise SEM between residues is greater than 50A), showing a significant amount of variance. These severe mutations exist on regions of CCR2, IL21, IL21R, IL6R and TNFSFR1A; a subset of proteins with a wide range of C-Score values (-1.11, -0.64, -3.39, -2.47 and -1.39), suggesting a low likelihood of this variance being due to weaknesses in the prediction algorithms.

Identification of binding site location within a tertiary structure is an important part of elucidating function and drug design. However, experimentally-derived binding site information is limited. With accurate binding site data, the local variance could be used to assess the likelihood that a variation will be less severe, and subsequently more likely to be benign. Utilising a relative scale for the heatmap by colouring the protein based on the maximum variance seen within the structure itself rather than a fixed global value across all proteins would allow for the isolation of areas showing most variance and potentially be useful in identifying severe mutations. Particular heatmaps, such as IL12BR2, show increased variance in localised regions, which could be good or bad depending on the binding site location.

Future work in this area can focus on many different topics. Improving the quality of the predictions utilised through computational methods would benefit every computational pipeline where modelling is required. These methods are developing quickly and as each new structure is included in the Protein Data Bank, the average quality of each prediction increases for structures calculated through homology modelling methods. Increases in quality of binding site identification, intramolecular contact elucidation and unstructured region identification are qualities that the protein structure prediction community are striving for.

Identification of binding site areas is a crucial step that would be key in the study of protein function and drug development. Multiple bioinformatics tools have been developed to tackle this problem, focusing on sequence data or structural data (Xie and Hwang, 2015). CAMEO experiments have been set up, as a continuous evaluation alongside the bi-annual CASP tests, which evaluates the quality of ligand-binding sites as part of its analysis. Current results suggest that a combination of structural and sequence based methods are the current industry standards.

Gene Ontology (GO) is an attempt to create a standardised vocabulary to discuss genes, their attributes and their products. Through the identification of homologues during the model formation process, predictions can be made into the biological processes in which a protein is involved, their explicit function and the cellular component(s) involved in construction.

Due to computational intensity requirements, limitations were added to the number of structures that were predicted and analysed in this chapter. The heatmaps generated in section Figure 4.7 would be a more useful descriptor of the variation across the protein if all known mutations were included in the analysis. Where population data is available, weighting the mutations by rate of occurrence could give a clearer picture into which parts of a protein structure are more likely to be affected by mutation.

Even without doing all of these things, further analysis can be completed with the PDB file output. PDBs can be used for protein-protein docking, electrostatic potential analysis, drug discovery, motif identification and discovery, complex formation, function prediction, homologue identification, fold recognition and many more. PDB files can be used to estimate binding kinetics of protein-protein interactions, allowing for this dataset to be used as part of a programme to study how the dynamics of atherosclerosis differ between population subgroups.

Chapter 5:

Calculating Binding Kinetics

5.1 Introduction

Determining the rate of association and dissociation for a PPI is an important step in the study of binding kinetics, theoretical drug design and the general study of the biochemical interaction. These molecular interactions drive almost all cellular processes (Phizicky and Fields, 1995). Modification of proteins requires such an interaction to take place, and the rate of protein-protein association and dissociation influences the consequences of the interaction and subsequent downstream effects. Additionally, variation in binding kinetics can have significant effect on drug efficacy (Pan et al., 2013). Quantifying binding kinetics of PPIs can be completed experimentally or predicted computationally. Our intention is to use the mathematical model described in Chapter 3, and the protein structures derived in Chapter 4, to calculate binding kinetics for each interaction included within the model, ultimately leading to population subgroup specific model reparameterisation.

5.1.1 Kinetic rates - k_{on}, k_{off} and k_d

A simple interaction between an enzyme E and a substrate S to form a complex P can be described as follows:

$$E + S \rightleftarrows ES \rightleftarrows EP \rightleftarrows E + P$$

This outlines the processes of enzyme-substrate binding, substrate modification, product formation and enzyme unbinding. This interaction can be simplified if we remove the intermediate step, leading to the Michaelis-Menten Model (Johnson and Goody, 2012; Michaelis and Menten, 1913):

$$E + S \rightleftarrows ES \rightarrow E + P$$

The rate of association, k_{on}, is defined as the rate of binding between E and S to form the complex ES. The rate of dissociation, k_{off}, represents the rate of transient complex unbinding to return to their original separate parts. This simplified interaction represents the binding and unbinding of a ligand without an intermediate stage (a common phase of the interaction process); however it is sufficient to illustrate important concepts of PPIs. Binding

can change the physical properties of both E and S, including tertiary structure (Ahmad et al., 2013).

The binding affinity, or dissociation constant (k_d), can be described using a ratio of concentrations at equilibrium:

$$k_d = \frac{[E][S]}{[ES]}$$

The dissociation constant k_d is a measure of the tendency of a protein complex to separate into its original parts. This is not to be confused with the dissociation rate constant k_{off}, the speed of complex separation into its constituent parts. The binding constant (k_a) is defined as the inverse of the dissociation constant. At the point of equilibrium, the enzyme concentration at which half of the substrate is bound to enzyme is equal to the dissociation constant k_d. In other words, equilibrium is reached when the following equation is true:

$$[E][S]k_{on} = [ES]k_{off}$$

leading to:

$$k_d = \frac{k_{off}}{k_{on}}$$

5.1.2 Determination of binding kinetics

Enzyme kinetic parameters are determined by the structure and molecular properties of the molecules involved within a PPI. The 'lock and key' theory of protein docking, considering the dilemma of protein-protein docking as a problem where proteins are treated as rigid-body structures, is a simplification that reduces the quality of any knowledge gleaned from an interaction (Jorgensen, 1991). The solvent accessible surface area (SASA), alongside the accessibility of the binding site areas provide a hard-limit to the speed of the biochemical reaction due to the greater impositions added when binding to an obstructed site over an open reaction site. Molecular flexibility similarly affects the rate of binding, particularly if

the binding site is occluded. The upper limit of the association rate k_{on} is approximately 10^{10} M^{-1} s^{-1}; however this requires optimal electrostatic navigation (Schreiber et al., 2009). Amino acid hydrophobicity, vital in the protein folding process, is also a key player in the protein-protein binding process (Young et al., 1994). Ascertaining binding kinetics to allow for future quantitative study can be completed using Surface Plasmon Resonance (SPR), ultraviolet-visible spectroscopy or NMR through titration experiments (Hahnefeld et al., 2004; Katoh et al., 2003).

5.1.3 Electrostatics

Electrostatic properties within protein-protein interactions are a vital component in the determination of binding kinetics and complex stability. The electrostatic strength involved in an interaction between proteins varies depending on the pH and the ionic strength of the surrounding environment. An increase in ionic strength will decrease k_{on} while having an insignificant effect on k_{off} (Pan et al., 2013; Wendt et al., 1997). The electrostatic component of the total binding free energy involved in a PPI can be studied using continuum solvation methods, a modelling approach where the solvent is represented as a single continuous element rather than a collection of individual molecules. Electrostatic potentials can be calculated using the Poisson-Boltzmann equation when temperature, ionic strength, molecular structure, atomic radii and atomic charges are known (Baker et al., 2001). Alterations to electrostatic potentials of proteins to improve steering within PPIs have been shown to lead to increases in the association rate (Phillip et al., 2009). The arrangement of charged and polar atoms within the three dimensional structure of the polypeptides, and subsequent electrostatic interaction formation within the complex, are involved in the determination of specificity and binding affinity of a PPI (Erijman et al., 2014).

5.1.4 Estimating binding rates and computational methods

The study and development of computational methods leading to calculation of PPI association rates has seen significant progress recently. Reducing a collection of microscopic interactions into solely a pair of proteins to allow for the study of a particular PPI is a cornerstone of calculating and estimating binding kinetics. This simplification does not take into account changes in rate parameters due to conformational changes, transitional states

and the effect of alternate molecules in the system, although by definition a rate constant ignores these also. Molecular Dynamics (MD) and Brownian Dynamics (BD) based algorithms have allowed for characterization of binding kinetics by estimating the effects on binding caused by electrostatics, hydrodynamic forces and structural flexibility (Pang and Zhou, 2017; Qin et al., 2011; Yu et al., 2015b). Methods predicting binding kinetics solely from protein primary structure have been developed which look at the attributes of amino acids (Yugandhar and Gromiha, 2014). The dissociation rate (the ratio of k_{on} to k_{off}) is significantly easier to predict than the individual kinetic parameters k_{on} and k_{off} (Xue et al., 2016).

Providing a qualitative measurement of k_{on} and k_{off} using molecular and Brownian dynamics solutions has many potential benefits within an *in silico* platform of PPI analysis. However, MD simulations that include all atoms within the structures, backbone and side-chain flexibility and a sufficiently fine time-step to study the interaction will quickly become too computationally intensive to be feasible. BD, a simplified system where protein components of lower priority are removed from the simulation yet are included as a collective random force, has been used by multiple groups as part of a package to computationally assess association rates (Qin et al., 2011; Yu et al., 2015b). Treating proteins to be docked as rigid bodies gives a significant reduction of degrees of freedom within the system, and permits a longer time-step to be used.

5.1.5 Bioinformatics tools for estimation of binding kinetics

Bioinformatics approaches to protein ligand binding have been developed with increasing complexity over the recent years. Molecular dynamics methods have been used to solve docking and association rate problems; however this can be a particularly computationally intensive task. Calculating binding kinetics for interactions is a highly complex problem with multiple degrees of freedom, requiring reduction of complexity before a practical method of solution is available.

TransComp is a web server designed to predict the association rate between two proteins through the “transient complex theory”, which can be found at

http://pipe.sc.fsu.edu/transcomp/ (as of 31/09/17). Interaction between two proteins R and L, forming a native complex C, is defined as:

$$R + L \rightleftarrows R^*L \rightarrow C$$

where R*L is a transient complex, a structure where distance between subunits is similar to the native complex although short range interactions have not commenced leading to complex formation. If the rate of the transient complex forming the native complex through conformational rearrangement, or R*L → C, is significantly faster than the rate of unbinding, or R*L → R + L, then the transient complex theory used in TransComp states that the association rate is approximately equal to the rate of transient complex formation, or R + L → R*L (Qin et al., 2011). In short, the theory holds when $k_{on} >> k_{cat}$. This situation describes a diffusion controlled reaction, a rapid reaction type where TransComp has improved accuracy in predicting k_{on} values.

The "transient complex theory" estimates k_{on} using the equation:

$$k_{on} = k_{a0} e^{(-\frac{\Delta G^*_{el}}{k_B T})}$$

where k_{a0} is the rate of random diffusion led transient complex formation. The Boltzmann factor $e^{-E/kT}$ represents the electrostatic potential impact on the binding mechanisms (where $E = \Delta G^*_{el}$, the electrostatic interaction energy for formation of R*L) (Alsallaq and Zhou, 2008; Qin et al., 2011).

SDA7 is a tool developed to estimate protein-protein association rate constants using a Brownian Dynamics approach to bind rigid solutes. SDA has been used in studies of association of soluble proteins to membrane bound proteins (Spaar et al., 2009) and drug-receptor complexes (Elsawy et al., 2012). SDA7 uses the Northrup-Allison-McCammon method to calculate the association constant k_{on} (Northrup et al., 1984), using ready-made interaction grids to calculate forces acting on individual atoms, reducing computational time. SDA7 will produce a range of results, requiring the definition of reaction criteria to isolate a specific result.

Predicting how tightly a protein-ligand complex is bound is a vital step in estimating the dissociation rate of a ligand-protein complex. PRODIGY is a bioinformatics tool that estimates the dissociation constant k_d by looking at the number of connections between each protein and individual surface features, utilizing the identified correlation between the number of contact points between two proteins in a complex and its binding affinity while considering structural properties away from the interface (Kastritis et al., 2011, 2014). PRODIGY has been shown to be a more accurate predictor of kinetic parameters than similar models. By using the binding affinity benchmarking algorithm established in Kastritis *et al.* (Kastritis et al., 2014), PRODIGY has become one of the best performing protein-protein binding affinity predictors that are publically available (Xue et al., 2016). The chemical and physical properties that are significant to dissociation are not as well understood as association.

5.1.6 Other binding kinetics prediction methods

The problem of modelling and combining the variety of factors that determine protein-protein kinetic rates, including structural diffusion, electrostatics, short-range interaction properties and solvent properties, has seen significant effort applied to it by the scientific community in recent years. Methods have been developed to consider the involvement of macromolecular crowding (Martinez et al., 2015; Mereghetti et al., 2010). Two-dimensional models considering the role of hydrodynamic forces are available, however standalone results from these algorithms are not sufficiently accurate (Cheung and Konstantopoulos, 2011). Machine learning strategies have been developed to significantly reduce computational time while predicting association rates (Xie et al., 2017).

5.1.7 The quest for native quaternary structure

Most proteins within a cell become functional when part of a complex or a quaternary structure, including simple dimers or more advanced oligomers and complexes. PPIs are vital in biological processes such as cell signaling, proteolysis and protein transport (Aldridge et al., 2006; Saunders et al., 2005; Yvan-Charvet et al., 2010). Proteins which remain in their monomer form and do not form higher order complexes are in the minority (Poupon and

Janin, 2010). Treating a structure as a monomer during proteomic study rather than in its complex form can lead to missing physiological elements, including binding site alterations and changes in long-range electrostatic forces.

Quaternary structures include oligomers and complexes with different quantities of subunits. Complex formation often serves to activate or inhibit one of the protein subunits (e.g. in atherogenesis MMP-1 is inhibited when binding with TIMP-1 to form a complex). Ascertaining the quaternary structure, either experimentally or computationally, encounters similar problems to identifying protein tertiary structure (see section 4.2.1). Experimentally derived complexes are stored within the PDB and are available for use by the wider research community (Rose et al., 2015). Transitory complexes, often involved in enzyme catalysis as described in section 5.1.5, are short lived, while oligomers remain in their quaternary structure long-term. Lack of knowledge of the forces in use to determine complex formation can lead to difficulty in identifying the oligomeric status of a protein, the type of protein that it would bind to and the kinetics involved in the binding process.

CASP started a study of quaternary structure accuracy analysis in 2010; however complexes predicted using the CASP9 targets were described as 'very poor' (Mariani et al., 2011). CAPRI (Critical Assessment of PRediction of Interactions), a community study into modelling quaternary structure from tertiary structure and evaluating docking procedures, have recently collaborated with CASP in order to assess oligomeric model quality and to evaluate docking procedures. Recent results have shown that the oligomeric state of a protein (i.e. the number of monomer subunits) can be predicted from its tertiary structure, that tetrameric structure is particularly challenging to predict and that analysis of PPIs can still provide useful insight even if input models are not highly accurate (Lensink et al., 2016).

Experimental elucidation of quaternary structure can be performed with multiple experimental techniques. Electron microscopy can be used to ascertain large macromolecular complex structures due to the high resolution provided (Nayeem et al., 1994). X-ray crystallography can also be used; however crystallization can be particularly challenging for quaternary structures (Ilari and Savino, 2008). NMR has also been used,

however, with a lower resolution which can lead to some ambiguity in the smaller details of the structures and the interactions (Lukin et al., 2003).

Homology modelling methods are generally considered to provide the most accurate structural models of tertiary structures from amino acid sequences. However, this can often lead to template bias, where structures are more likely to take physiological properties from the homologue templates used in model formation. Refinement methods to improve predicted structure quality and alleviate template bias have seen significant progress in recent years in both global whole-structure accuracy and in local regions. Potential Energy Minimization (PEM) through alterations to the backbone, side chain rebuilding to remove steric clashes and structural relaxation can lead to refined structural targets (Heo et al., 2013, 2016). Other approaches including molecular dynamics based refinements, elastic network models, knowledge and fragment based approaches, and optimization of hydrogen bond networks have been successfully utilized in CASP experiments (Nugent et al., 2014).

5.1.8 MM-Align

Comparing oligomeric and complex structures can be performed with MM-Align, which provides an alignment for two proteins that is independent of primary sequence (Mukherjee and Zhang, 2009). TM-Align will perform a similar function for comparing monomers, however multimer functionality is not provided (Zhang and Skolnick, 2005). Chains for input proteins could be merged into a single contrived monomer and used with TM-Align, although MM-Align provides additional functionality that prevents cross-chain alignments. MM-Align will provide an RMSD and a TM-Score as a similarity score between proteins, although unlike TM-Align the TM-Score also assesses alignment coverage.

5.1.9 Docking

Protein docking is the process of converting protein subparts into a single protein complex. Evaluating how proteins interact with each other through computational docking methods is an important step in computational drug design due to the ability of certain therapeutics to modulate or block PPIs. The most common types of docking algorithms, including geometric hashing, Monte-Carlo simulations or fast generalized Fourier transformation, involve

searching spatial arrangements of the proteins involved and using an energy-minimizing scoring function to isolate a particular configuration (Ritchie, 2008). These methods will produce a list of potential structures, including the correct near-native solution; however isolating this structure from many is a prevalent challenge to bioinformaticians interested in this field. Knowledge of solvent accessibility, binding site details and local docking approaches can all be used to limit the number of potential models available and obtain near-native structure. Methods for predicting homo-oligomeric structure have shown success in recent CASP/CAPRI experiments (Baek et al., 2017).

5.1.10 Hex and Rosetta

There are currently more than 60 different docking tools available to predict quaternary structure from protein subparts (Pagadala et al., 2017). Each of these tools are designed to combine protein subunits together to form a complex or oligomer, however multiple different methods are used for docking, providing varying accuracy depending on the methods and inputs used. HexDock is a fast Fourier transform (FFT) based protein docking algorithm (Ritchie and Venkatraman, 2010). Hex utilizes a 'spherical polar Fourier technique' (SPF), where rotational correlations have been used as a replacement for a standard Cartesian grid. FRODOCK, a fast-rotational-docking approach has performed benchmarking tests showing a significant performance improvement using this method (Garzon et al., 2009). Multiple potential configurations are produced by Hex, which are ordered by an energy based scoring function to provide an optimal docking solution.

RosettaDock is a tool which can be used to perform local docking. Local docking requires knowledge of the binding locations of each protein. Local docking searches a subspace of the total possible spatial configurations, requiring the user to provide an approximation of the output structure by ensuring that binding sites are facing towards each other. Local docking provides a structure closer to the native, as computational time spent searching for the correct spatial conformation is significantly reduced, allowing for a focus on short-range binding-site interactions and the removal of ambiguity between possible near-native combinations.

5.1.11 Free energy difference

The binding of two molecules in a solvent can be described by the Gibbs Free Energy, a description of the total energy that can be used in an interaction plus entropic factors, where:

$$\Delta G = -RT * \ln(1/K_d)),$$

$$\Delta G = \Delta H - T\,\Delta S$$

where ΔG is the change in free energy, ΔH and ΔS are the sum of the relevant enthalpy and entropy respectively, T is the temperature in Kelvin and R is the ideal gas constant. A k_d value of 1×10^{-12} M has a Gibbs free energy of approximately -7×10^4 J/mol, and an reduction of 3×10^4 J/mol in energy corresponds to a k_d reduction in the region of five orders of magnitude (György M. Keserü and David C. Swinney, 2015).

Studying the difference in Gibbs free energy between a mutation and its wildtype can provide insights into binding kinetics, structural changes and protein stability. Changes in binding affinity caused by mutations can be combined with population genetics to predict disease-causing mutations and become part of a stratified medicine program.

5.1.12 PDB2PQR

The study of electrostatics and its role in PPI has been made simpler through the use of continuum solvation models. To perform electrostatic potential calculations, commonly performed with MD and BD simulations, full all-atom structural data and atomic charges and radii are required. PDB2PQR is a tool designed to automate preprocessing tasks for molecular electrostatics calculations, including adding missing heavy atoms, assigning force field values including atomic charges and radii, and generating a PQR file (similar to a PDB file where the B-factor and occupancy values have been replaced with atomic charge and radius values).

Brownian dynamics and molecular dynamics simulations of PPIs require an empirical force field to allow for energy-based calculations. MD and BD simulations relate local and global

structural energy to forces at the atomic-level leading to structural coordinates over time, representing the docking process. Force fields, such as AMBER (Assisted Model Building with Energy Refinement) or CHARMM (Chemistry at Harvard Macromolecular Mechanics) contain a set of potential energy functions, which are used to calculate the forces undertaken by each individual atom (Brooks et al., 1983; Weiner et al., 1986).

5.2 Methods

5.2.1 Establish Atherosclerosis Interactions

Using the model of atherosclerosis described in Chapter 3 and the protein structure models generated in Chapter 4, association rates were predicted for a series of interactions involved in atherogenesis as detailed below. Using the proteins detailed in Table 4.2 , the following interactions will be studied as part of this procedure:

Table 5.1: Interactions related to Chapter 3's Model of Atherosclerosis

Interaction Number	Ligand	Receptor	Notes
1	ADAM17	TIMP3	
2	CCL2	CCR2	
3	CCL5	CCR5	
4	CMA1	MMP1	
5	CSF1	CSF1R	
6	CXCL10	CXCR3	
7	CXCL11	CXCR3	
8	CXCL9	CXCR3	
9	EGF	EGFR	
10	IFNG	IFNGR	IFNG is Homodimeric IFNGR is Heterodimeric
11	IL1B	IL1R1	
12	IL10	IL10RA	
13	IL12	IL12R	IL12 and IL12R are Heterodimeric
14	IL17A	IL17RA	
15	IL18	IL18R	IL18R is Heterodimeric
16	IL2	IL2R	

Interaction Number	Ligand	Receptor	Notes
17	IL21	IL21R	
18	IL33	IL1RL1	
19	IL4	IL4R	
20	IL5	IL5R	IL5 is Homodimeric
21	IL6	IL6R	
22	PDGFA	PDGFRA	
23	PDGFB	PDGFRB	PDGFB is Homodimeric
24	PDGFA	PDGFRB	
25	PDGFB	PDGFRA	
26	TGFB	TGFBR	TGFB is Homodimeric TGFBR is Heterodimeric
27	TIMP1	MMP1	
28	TIMP1	MMP2	
29	TIMP1	MMP3	
30	TIMP1	MMP9	
31	TIMP1	MMP13	
32	TIMP2	MMP1	
33	TIMP2	MMP2	
34	TIMP2	MMP3	
35	TIMP2	MMP9	
36	TIMP2	MMP13	
37	TIMP3	MMP1	
38	TIMP3	MMP2	
39	TIMP3	MMP3	
40	TIMP3	MMP9	
41	TIMP3	MMP13	
42	TIMP4	MMP1	
43	TIMP4	MMP2	
44	TIMP4	MMP3	
45	TIMP4	MMP9	
46	TNFA	TNFR	TNFR is Heterodimeric

5.2.2 Establishing oligomeric status

A literature mining process was undertaken to establish the oligomeric status of each protein structure predicted. These oligomers are detailed in Table 5.2.

Table 5.2: Oligomers involved in atherosclerosis model

Protein Name	Oligomeric Status	Gene Names
Interferon-Gamma	Homodimer (Farrar and Schreiber, 1993)	IFNG
Interferon-Gamma Receptor	Heterodimer (Farrar and Schreiber, 1993)	IFNGR1 IFNGR2
Interleukin 2 Receptor	Heteroligomer (Malek and Castro, 2010)	IL2RA IL2RB IL2RG
Interleukin 5	Homodimer (Takatsu, 2011)	IL5
Interleukin 10	Homodimer (Syto et al., 1998)	IL10
Interleukin 10 Receptor	Tetrameric (Verma et al., 2016)	IL10RA IL10RB
Interleukin 12	Heterodimer (Mazzeo et al., 2002)	IL12A IL12B
Interleukin 12 Receptor	Heterodimer (Gately et al., 1998)	IL12RA IL12RB
Interleukin 17	Homodimer (Liu et al., 2013)	IL17A
Interleukin 18 Receptor	Heterodimer (Kato et al., 2003)	IL18R1 IL18RAP
Platelet Derived Growth Factor Beta	Homodimer (Bhandari et al., 1994)	PDGFA PDGFB
Transforming Growth Factor Beta	Homodimer (Esebanmen and Langridge, 2017)	TGFB1
Transforming Growth Factor Beta Receptor	Heterodimer (Cheifetz et al., 1988)	TGFBR1 TGFBR2
Tumor Necrosis Factor Alpha Receptor	Heterodimer (Wang and Al-lamki, 2013)	TNFRSF1A TNFRSF1B

5.2.3 Obtaining structures from PDB

Sequences isolated from the 1000 Genome Project for the 65 wildtype proteins were aligned with each sequence in the PDB. If a sequence showed >95% sequence identity with an experimentally derived structure within the PDB for both proteins within an interaction as described in Table 5.1, it was considered to be a representative model for use in docking simulations. In the absence of suitable knowledge of the native complex, Hex was used to predict quaternary structure of our protein structures. Hex procedures used are detailed in section 5.2.8.

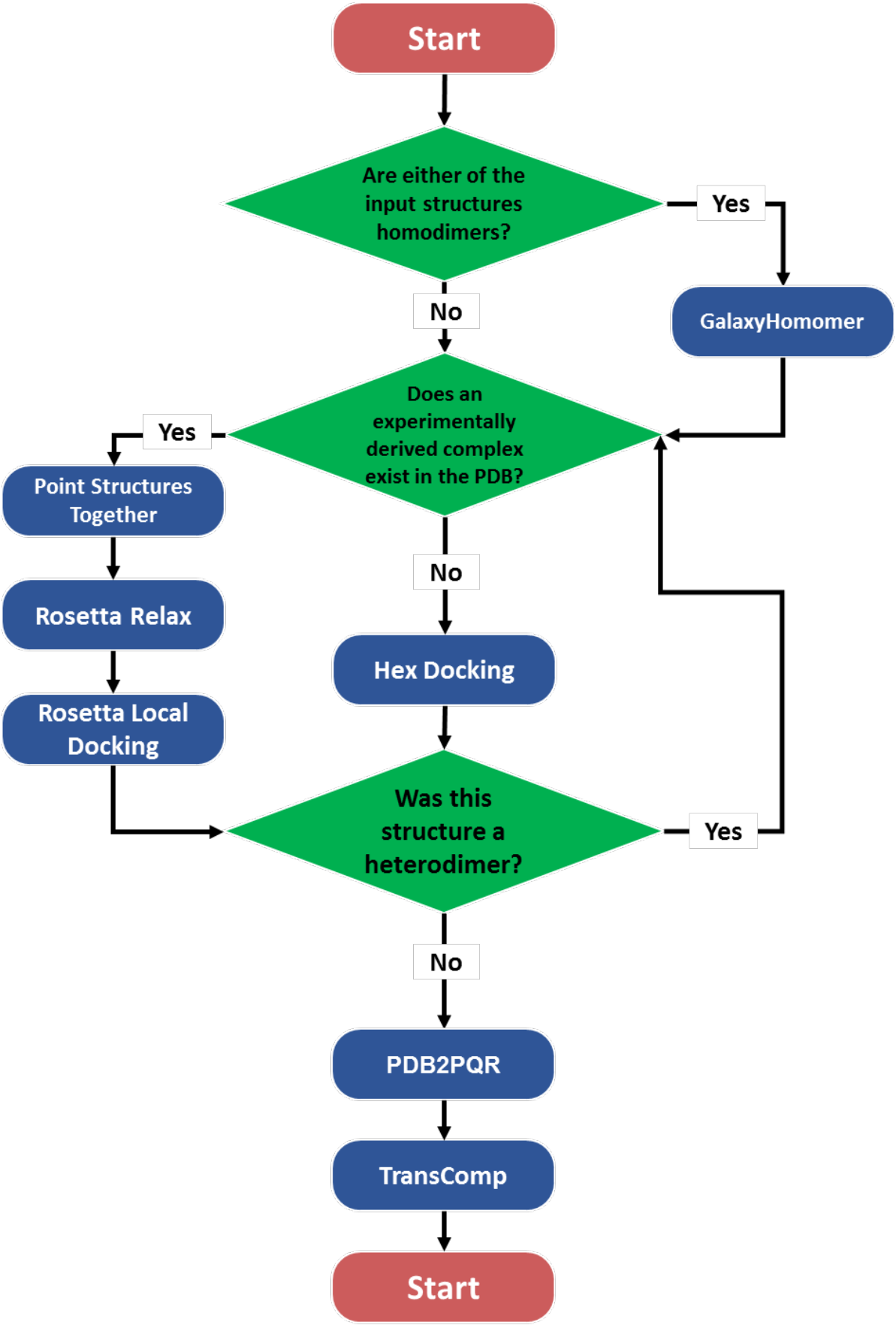


Figure 5.1: Flowchart describing procedure undertaken for each predicted structure

5.2.4 Local Docking Preparation

To provide a complex as close to the native as possible, a local docking procedure was undertaken. A local docking procedure requires knowledge of the native binding sites, and binding simulations are performed on a subset of the total number of possible configurations rather than performing a computationally intensive global search. Limiting the conformational space to areas near to the provided input where possible allows for the removal of wildly incorrect arrangements and significantly reduces computational time.

For each of our interactions as detailed in Table 5.1, predicted protein structures derived as detailed in Chapter 4 were used. For these interactions, if a native structure exists within the PDB, residues involved in the binding site were isolated from the experimental structure. Complexes were opened in JMol (The Jmol Team, 2007) and all atoms within 3 angstrom of the opposing structure were found with the command:

'select within(3.0, :' AChain ') and within(3.0, :' BChain ')'

All residues containing atoms isolated in this way were considered to be near-binding-site. Sequences isolated from the 1000 Genome Project were aligned to sequences for the experimentally derived protein to find the associated near-binding-site residues within the structural models.

To allow for the local docking procedure to take place effectively, near-binding-sites for each model isolated as previously described are positioned to face the other before local docking begins. For each protein structure within the dataset generated in Chapter 4, a near-binding-site configuration is created for each structure with the wildtype of the opposing structure within the interaction. A six-dimensional rotational space search for each protein is undertaken using Matlab, in order to minimize the mean distance between each atom in the near-binding-site for each protein, generating rotation and translation matrices to ensure that pairwise near-binding-site residue distance is rotationally minimized. Proteins are rotated by pi/6 radians on each axis around their center, and the distance to the near-binding-site residues on the opposing protein is calculated, and the minimum distance is

selected as the correct rotation to be applied. Proteins are moved so that their centers are 100Å apart before the rotational matrices are calculated, and returned to 10Å apart after rotation has been calculated. Rotational and translational matrices are applied to our PDB files using Chimera, due to its conservation of secondary structure under rotation and its automation capability (Pettersen et al., 2004).

5.2.5 Rosetta Relax

Where suitable knowledge of the native complex was available, RosettaCommons docking was used to predict complex structure for each interaction defined in Table 5.1, using the mutation data isolated in Chapter 4.

Once a structure was obtained with rotational near-binding-site-distance minimization, these structures are relaxed using the Rosetta 'Relax' protocol. Relax performs a simple all-atom refinement that acts as a pre-processing step for future docking, through continued side-chain repackaging and energy minimization (Conway et al., 2014; Khatib et al., 2011; Tyka et al., 2011).

Rosetta Relax Protocol was run with the following parameters:

Table 5.3: Rosetta Relax Parameters

Parameter Name	Value	Notes
-nstruct	50	Number of runs to perform
-ex1		Adds additional side-chain rotamers, highly recommended by the developers
-ex2		Adds additional side-chain rotamers, highly recommended by the developers
-use_input_sc		Uses current rotamers within current Monte-Carlo simulations – recommended by developers
-no_optH	False	Performs hydrogen atom position optimization
-flip_HNQ		Permits alterations of HIS, ASN, and GLN during optimization of hydrogen placements – recommended by developers

-overwrite		Allows results to be overwritten when performing multiple runs, added for algorithmic simplicity

The Chimera derived structures are detailed in the previous step were provided as input, with an output location for the PDB file and the score file specified using the commands –in:file:s, -out:path:pdb and –out:path:scorefile respectively. RosettaRelax returns multiple relaxed protein structures, ordered by the energy based scoring function talaris2014 (O'Meara et al., 2015).

Relax analysis was performed on Google Cloud, on a virtual machine of type n1-standard-1 (1 vCPU, 3.75 GB memory) with an Intel Sandy Bridge CPU platform, running Ubuntu 16.04.

5.2.6 Local Docking

RosettaDock was used to form quaternary protein structures due to its highly customizable local docking procedures, providing complexes closer to the native than global docking procedures at a reduced computational cost. The relaxed structures, derived using the near-binding-site information as described in section 5.2.4 and the Rosetta Relax protocol are used as an input for a Rosetta local dock. Structures with binding pockets facing towards each other are provided as an input with the following parameters:

Table 5.4: Rosetta Docking Parameters

Parameter Name	Value	Notes
-nstruct	500	Number of runs to perform
-unboundrot	~Input File~	
-partners	~ReceptorChain_LigandChain~	e.g. A_B to dock proteins A and B
-dock_pert	3 8	
-ex1		
-ex2aro		
-out:suffix	_local_dock	

Docking analysis was performed on Google Cloud, on a virtual machine of type n1-standard-1 (1 vCPU, 3.75 GB memory) with an Intel Sandy Bridge CPU platform, running Ubuntu 16.04.

RosettaDock returns a collection of potential structures, each characterized by two energy-based scoring functions, the total score and interface score, as defined in (O'Meara et al., 2015). The interface score is calculated as the total complex score minus the score for each monomer subunit. As recommended by the developers, 500 complex structural models were generated through this method, and the model with the lowest total score was chosen to be our model of quaternary structure, with interface score being used when total scores were identical.

5.2.7 Homodimers

Homodimeric structures were calculated using GalaxyHomomer, a methodology designed to predict homo-oligomer structure from tertiary structure or amino acid sequences. Tertiary structures for the homodimeric proteins described in Table 5.2 were used as an input for GalaxyHomomer. Input parameters for oligomeric state and regions to be refined were not given as an input. Up to five homo-oligomer results are returned at the end of the predictive process, where priority was given to the sequence-based modelling result with the highest sequence identity. In the absence of available templates in the PDB with similar sequence to our input protein, structure-based templates were utilized if they have a TM-Score higher than 0.5 when aligned to the input protein. Failing this, if no appropriate templates were found then *ab initio* methods were used. All outputs isolated from GalaxyHomomer were given additional refinement using GalaxyRefineComplex, an algorithm designed to improve model structures through side chain rebuilding (Heo et al., 2016). Due to the small amount of homodimers involved in our dataset (48), analyses were performed on the corresponding web servers – http://galaxy.seoklab.org/refine and http://galaxy.seoklab.org/homomer. Both web servers are running on Linux clusters of 2.33GHz Intel Xeon 8-core processors (4 servers for GalaxyRefine, 12 for GalaxyHomomer). The source code can be downloaded from the aforementioned websites (August 2017).

5.2.8 Hex Docking

In the absence of a representative template in the PDB to provide near-binding-site data, Hex (a fast Fourier transform (FFT) based protein docking algorithm) was used to predict a complex structure (Ritchie and Venkatraman, 2010).

Hex simulations are run using the parameters:

Table 5.5: Hex Docking Parameters

Parameter Name	Value	Notes
Correlation Type	Shape + Electro	
Compute Device	CPU	Hex has functionality to perform quick docking simulations on a GPU unit, however this functionality was unavailable on our chosen hardware
Sampling Method	Range Angles	
Solutions	2000	Up to a maximum of 2000 solutions were provided – however Hex will provide fewer if conformations are not found
Grid Dimension	0.5	

Hex docking was performed on a Windows 7 machine, with an Intel Core i7-4770 @ 3.4GHz processor and an Nvidia Quadro K600 graphics card.

Heteroligomeric structures were predicted using RosettaCommons Docking, a collection of steps designed to predict binding dynamics from unbound protein subunits.

5.2.9 Selecting a tool to calculate k_{on}

SDA7 and TransComp were both considered to calculate association rate using our predicted protein structures. To compare SDA and TransComp results, a collection of 13 structures with known k_{on} were found and used as inputs for both methods. Outputs were then compared to the experimental values to ascertain which method to be used for the larger dataset. A list of the structures with association rates predicted and their results can be found in section 5.3.4 (Table 5.9).

5.2.10 PDB2PQR

Hex and Rosetta outputs were converted into PQR files using PDB2PQR, using the AMBER forcefield. This PQR file of a protein-protein complex was then split into its two individual subunits, subA.pqr and subB.pqr as part of preprocessing for upcoming TransComp use.

5.2.11 TransComp alterations

While TransComp is an excellent tool providing strong predictions of association kinetics from tertiary structure, some minor alterations were required to be made from the downloadable version of TransComp. Due to a compatibility issue with version 16.04 of Ubuntu, line 13 of Sample.sh required to be changed to 'nconfOne=100000' to prevent excessive rotational and translational space sampling. In addition, due to the size of the macromolecules returned by RosettaCommons Docking and Hex, coordinates in PQR files required 8 digits to be scored, removing necessary whitespace from the input file. To correct this, additional whitespace was added into all PQR files where necessary. Both of these changes were made in consultation with the developers.

5.2.12 TransComp

TransComp was used to generate association rates for each interaction subunit. Prerequisites include APBS (Adaptive Poisson-Boltzmann Solver) (Baker et al., 2001) for performing electrostatics calculations, and Gnuplot (a command line graphing tool) (Williams et al., 2013). APBS 1.5 and GnuPlot 4.6.0 were installed on a virtual machine on Google Cloud of type n1-standard-1 (1 vCPU, 3.75 GB memory) with an Intel Sandy Bridge CPU platform, running Ubuntu 16.04.

The solvent ionic strength selected was 0.15M for all interactions as a standardized, realistic value.

Occasionally, TransComp would be unable to generate a transient complex from the hex output due to difficulty in forming a single transient complex, believing that multiple transitory structures were involved in the interaction. In these situations, the next strongest

prediction from Hex or Rosetta was chosen, step 5.2.10 was repeated on the new complex and another TransComp run was launched.

However, when performing docking and subsequent association rate calculation then not all of these proteins are considered in these oligomeric states.

5.2.13 PRODIGY

PRODIGY was used to predict a dissociation constant k_d as described in section 5.1.5. Using the relation:

$$k_d = \frac{k_{off}}{k_{on}}$$

PRODIGY results were combined with TransComp results to predict a k_{off} value. PRODIGY analysis was performed on a virtual machine on Google Cloud of type n1-standard-1 (1 vCPU, 3.75 GB memory) with an Intel Sandy Bridge CPU platform, running Ubuntu 16.04.

5.3 Results

5.3.1 Homodimers

Within the collection of protein structures calculated as described in Chapter 4, 5 different proteins are considered homodimers, with 27 mutations giving a total of 32 different homodimeric structures to study. The breakdown of these proteins can be seen in Table 5.6.

Table 5.6: Homodimeric structures involved in atherosclerosis model

Protein Name	Number of Structures	PDB Code	Gene Names
Interferon-Gamma	4	1EKU	*IFNG*
Interleukin 5	8	3QT2	*IL5*
Interleukin 17	8	4HR9	*IL17A*
Platelet Derived Growth Factor Beta	28	1PDG	*PDGFB*

For each of the homodimeric structures detailed in Table 5.6, structural alignment between the predicted structure and the experimentally derived structure (PDB codes can be found in Table 4.2) using MM-Align (as described in section 5.1.8) has given an RMSD of 3.78Å, 3.31Å, 2.66Å and 2.31Å for IFNG, IL5, IL17A and PDGFB respectively. TM-Scores given were more unreliable at 0.18443, 0.10971, 0.80273 and 0.57694 respectively, however this is due to additional protein chains contained within the experimentally derived PDB file.

The following figures show a boxplot representing the collection of mutations associated with each interaction detailed in Table 5.1. Outliers are shown if the RMSD is less than q1 – 1.5 × (q3 – q1) or greater than q3 + 1.5 × (q3 – q1), where q1 and q3 are the 25th percentile and 75th percentiles of the sample data for each complex.

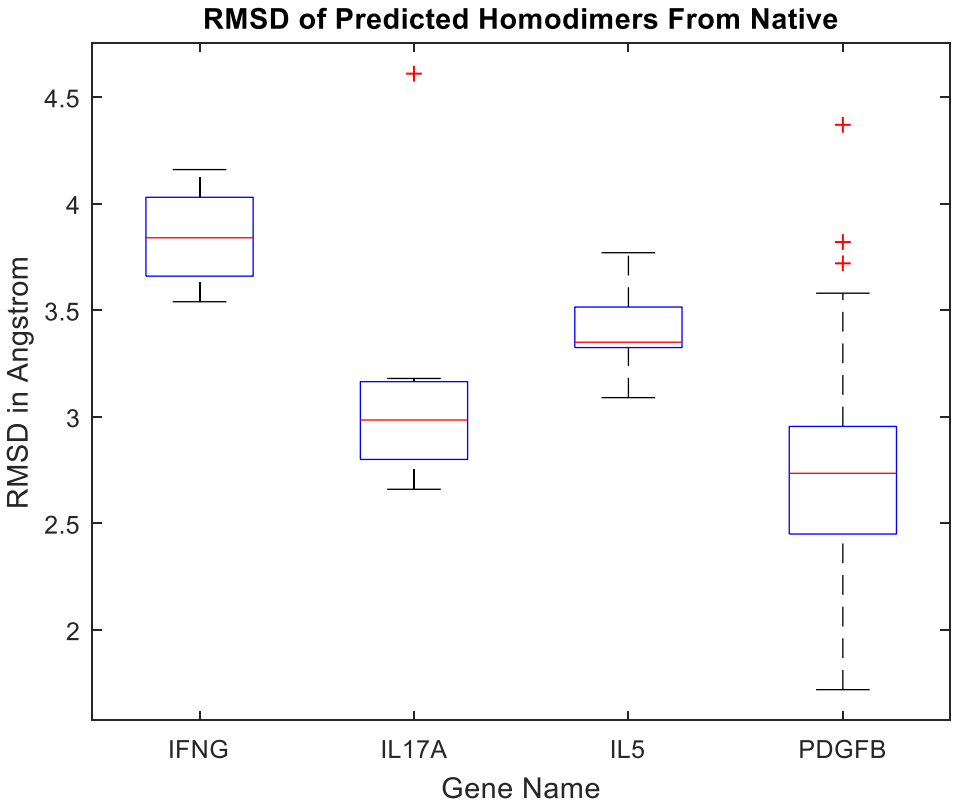


Figure 5.2: Box plots showing RMSD for each homodimeric structure related to the native structure

5.3.2 Heterodimers

Table 5.7: Heterodimeric structures involved in atherosclerosis model

Protein Name	Oligomeric Status	PDB Code	Gene Names
Interferon-Gamma Receptor	Heterodimer	N/A	IFNGR1 IFNGR2
Interleukin 12	Heterodimer	1F45	IL12A IL12B
Interleukin 12 Receptor	Heterodimer	N/A	IL12RA IL12RB
Interleukin 18 Receptor	Heterodimer	3WO4	IL18R1 IL18RAP
Transforming Growth Factor Beta Receptor	Heterodimer	3KFD	TGFBR1 TGFBR2
Tumor Necrosis Factor Alpha Receptor	Heterodimer	N/A	TNFRSF1A TNFRSF1B

Of the 6 heterodimeric structures included in our dataset as detailed in Table 5.7, structural alignment was performed between the predicted structure and the experimentally derived structure for the 3 proteins where appropriate PDB codes were found. RMSD scores for the wildtypes of IL12, IL18R and TGFBR were 2.91Å, 1.32Å and 8.83Å respectively.

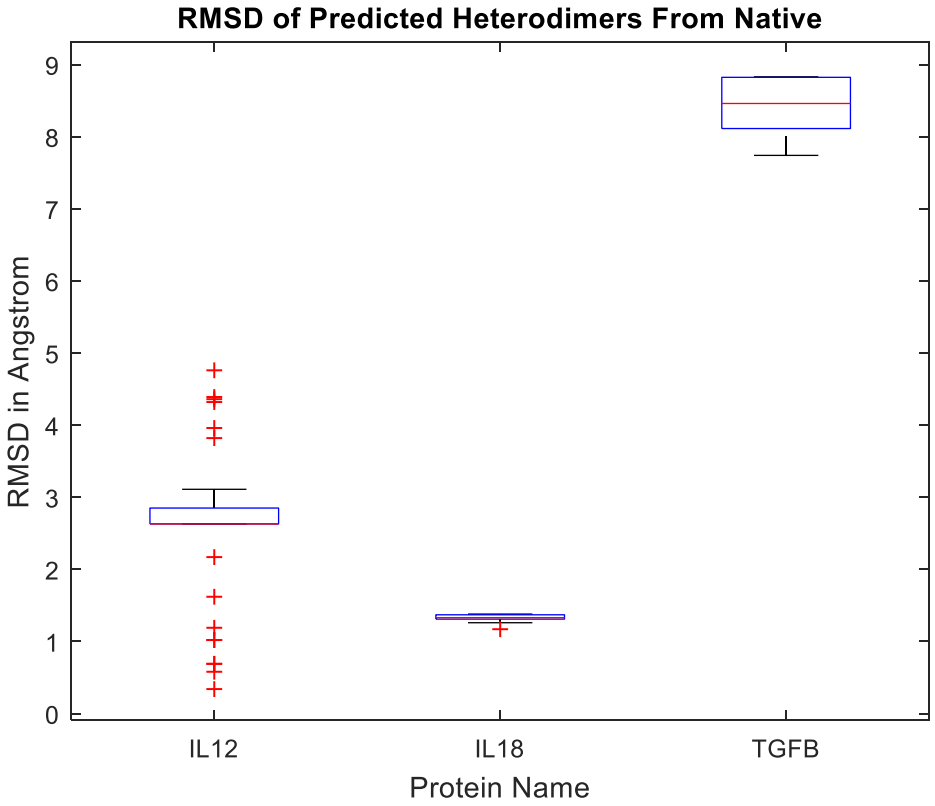


Figure 5.3: Box plots showing RMSD for each heterodimeric structure related to the native structure

5.3.3 Rosetta Docked Structures

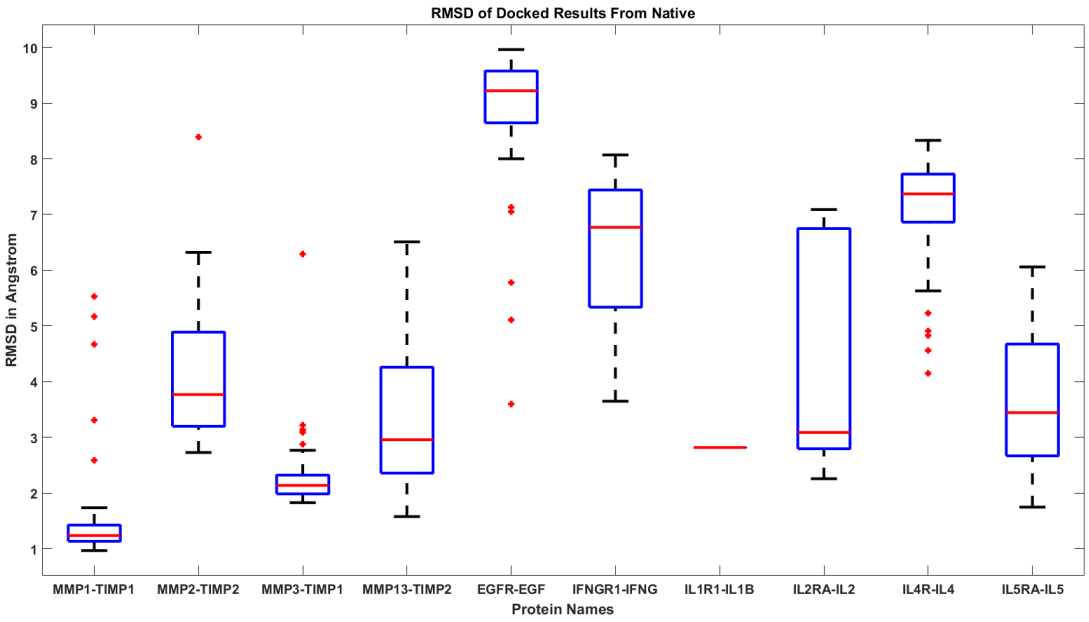
The following structures have been docked together to make a complex using Rosetta, using the procedures detailed in section 5.2.6:

Table 5.8: Proteins using Rosetta Local Docking

Protein 1	**Protein 2**	**PDB Code**
EGF	EGFR	1IVO
IL10	IL10R1	1J7V
TIMP1	MMP3	1OO9
IL2	IL2RA	2B5I
TIMP2	MMP13	2E2D
TIMP1	MMP1	2J0T
IL4	IL4R	3BPL

ADAM17	TIMP3	3CKI
PDGFB	PDGFRB	3MJG
IL1B	IL1R	1ITB
IL5	IL5R	3QT2
IL17A	IL17RA	4HSA
IL6	IL6R	4J4L
IL33	IL1R	4KC3
IL18	IL18R	4R6U
CSF1	CSF1R	4WRL

Each mutation was aligned using MM-Align to the experimentally derived wildtype associated with the interaction as detailed in Table 5.8.



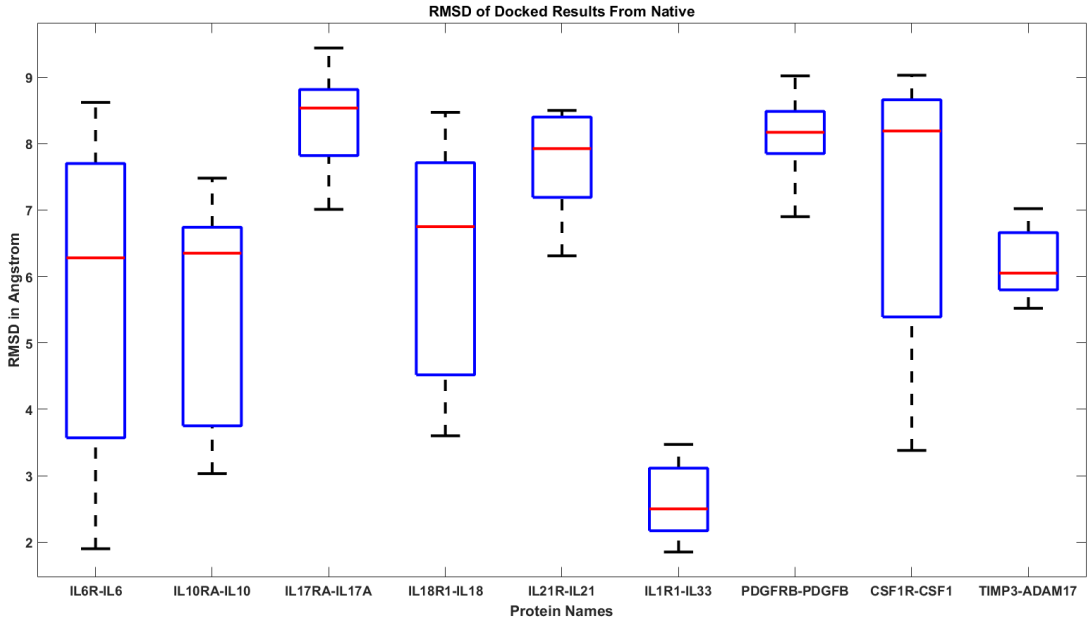


Figure 5.4: Box Plots for docked proteins

5.3.4 Testing TransComp and SDA

To evaluate the accuracy of TransComp and SDA, a collection of experimentally derived structures with known association rates were used as inputs as detailed in section 5.2.9. Results are shown in Table 5.9.

Table 5.9: Benchmarking for TransComp and webSDA for known k_{on} values

PDB	C	Description	IS	Ex K_{on}	TC	T Acc	SDA	Range	SDARC	SDA Acc
2J0T	A:D	MMP1-TIMP1 (Troeberg et al., 2002)	230	5.2E+04	8.9E+04	1.7E+00	7.10E+6	2.30E+8	3/8.5	1.37E+2
1UEA	A:B	MMP3-TIMP1 (Troeberg et al., 2002)	230	2.0E+05	5.6E+04	2.8E-01	6.60E+6	4.00E+8	3/7.0	3.30E+1
1GXD	A:C	MMP2/TIMP2 (Olson et al., 1997)	166	1.4E+04	1.9E+05	1.4E+01	5.20E+6	2.00E+9	4/5.5	3.71E+2
1KTZ	A:B	TGFB/TGFBR (Baardsnes et al., 2009)	160	7.4E+05	5.6E+06	7.6E+00	4.21E+7	1.90E+9	3/6.0	5.69E+1
1P9M	A:B	IL6/IL6R	150	2.7E+07	6.5e+05	4.1E+01	6.14E+6	6.26E+8	3/6.0	1.36E+1

		(Hammacher, 1996)								
1IAR	A:B	IL4/IL4R (Shen et al., 1996)	150	1.9E+07	5.3E+07	2.8E+00	5.40E+7	1.40E+9	4/5.0	2.84E+0
2ERJ	A:C	IL2A/IL2R (Myszka et al., 1996)	150	7.8E+06	3.6E+07	4.6E+00	3.60E+7	1.70E+9	4/4.0	4.62E+0
2ERJ	B:D	IL2B/IL2R (Myszka et al., 1996)	150	1.3E+06	9.8E+05	7.5E-01	2.60E+6	5.00E+8	4/6.0	2.00E+0
1FYH	A:B	IFNG/IFNGR (Sadir et al., 1998)	150	7.3E+06	8.5E+04	1.2E-02	6.10E+6	9.10E+9	4/6.0	8.36E-1
1EQY	A:S	Actin/Gelsolin (SELVE and WEGNER, 1987)	110	2.5E+04	2.4E+05	9.6E+00	6.20E+6	3.50E+8	2/7.5	2.48E+2
2PCC	A:B	cytochrome c peroxidase / cytochrome (Mei et al., 1996)	150	1.3E+09	4.3E+09	3.3E+00	5.70E+9	9.80E+8	4/6.0	4.38E+0
1UDI	E:L	uracil-DNA glycosylase / inhibitor (Bennett et al., 1993)	100	1.5E+08	5.8E+08	3.9E+00	9.00E+8	7.20E+6	4/6.0	6.00E+0

PDB – PDB Code for experimentally derived structure; C – Chain IDs in PDB; Desc – Description of the interaction; IS – Ionic Strength (in mM); Ex k_{on} – Experimentally derived k_{on} value; TC – TransComp prediction; T Acc – TC/(Ex k_{on}); SDA – webSDA prediction; Range – Upper limit of SDA prediction; SDARC – Reaction criteria used in SDA to isolate best result (No of Contact Points/Distance Apart); SDA Acc – (SDA/(Ex k_{on})

Within this proof-of-concept test, 10 of our TransComp results were within an order of magnitude from the reported results, while only 6 SDA7 results were within the same window. Additionally, SDA7 returns a range of results, requiring the definition of reaction criteria which would return a range of results that would need to be arbitrarily reduced into a singular value. As such, TransComp was selected as the tool of choice for calculating association rates within this project.

5.3.5 TransComp and PRODIGY Results

TransComp was used to calculate association rates were calculated for 1119 complexes involved in atherosclerosis. Of these complexes, 835 returned an association rate. Failure in calculating an association rate was due to the interaction requiring an intermediary step, or computational intensity restrictions. Association rates were not calculated when an individual interaction would take longer than 3 months to process on a single core virtual machine. An exhaustive list of association rate values can be found in Appendix 5. The median of the association rates is 1.5×10^{5}, with an interquartile range of 2.142×10^{6}.

Some results obtained were evidently erroneous. For example, the binding of CSF1R-001-245 to CSF1-001 returned a k_{on} value of 3.81×10^{-188}. Such issues however do not seem to be an issue for the fast-binding proteins; TransComp estimation of the association rate between MMP3-001-45 and TIMP3-001 is 5.61×10^{10}, which is a value within the range of expected protein association rates (Schreiber et al., 2009). To filter out these results, only results with a k_{on} value between 10^{0} and 10^{10} were used.

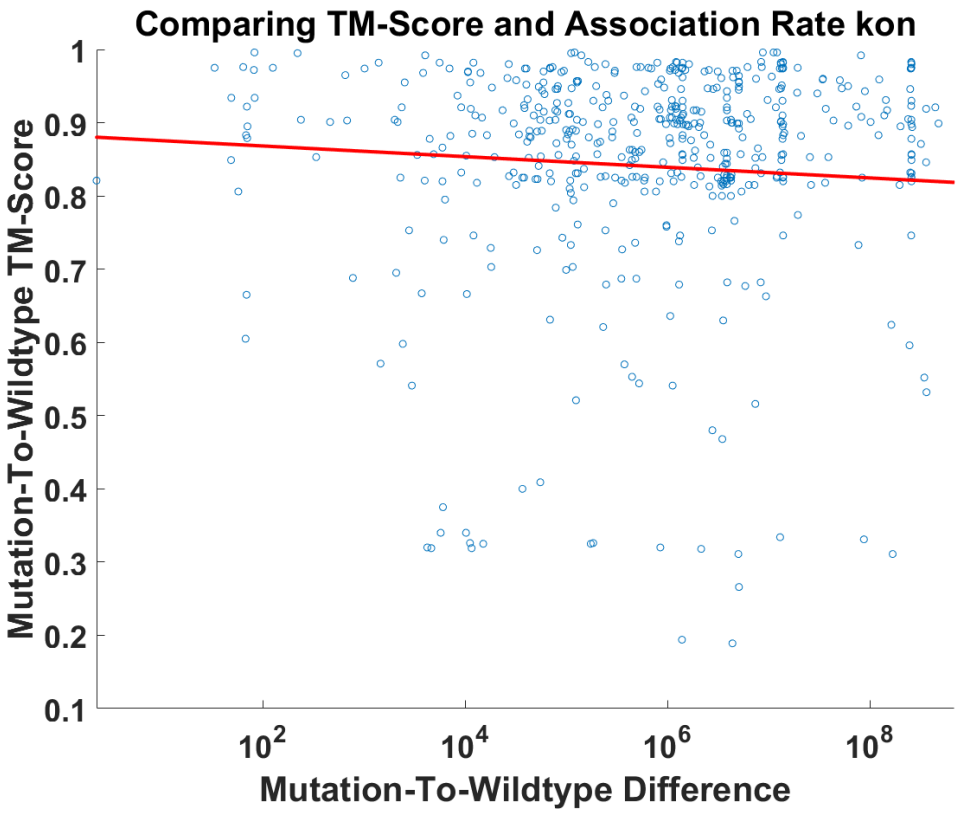


Figure 5.5: Comparing TM-Score for each mutation to its wildtype to the difference between their association rates

MATLAB's polyfit function was used to generate the line of best fit. Figure 5.5 shows a comparison between the variance in k_{on} between a mutation and its wildtype alongside the TM-Score for an alignment between the mutated protein and its wildtype. The line of best fit was generated in Matlab using the polyfit method. Figure 5.5 suggests a correlation, albeit minor, between increased variation within a protein structure leading to increased variation in association rate. This is expected behavior in direction and magnitude, with it being intuitive that a decrease in protein similarity would lead to an increase in variation of association within binding.

PRODIGY was used to estimate a dissociation rate k_d to allow for the generation of binding kinetic parameter k_{off} as defined in section 5.2.13. The collection of k_d results compared to TM-Score is shown in Figure 5.6. The mean of the k_d values is 1.517 $x10^{-5}$, with an interquartile range of 7.3677 $x10^{-6}$. Figure 5.6 shows a similar correlation to Figure 5.5.

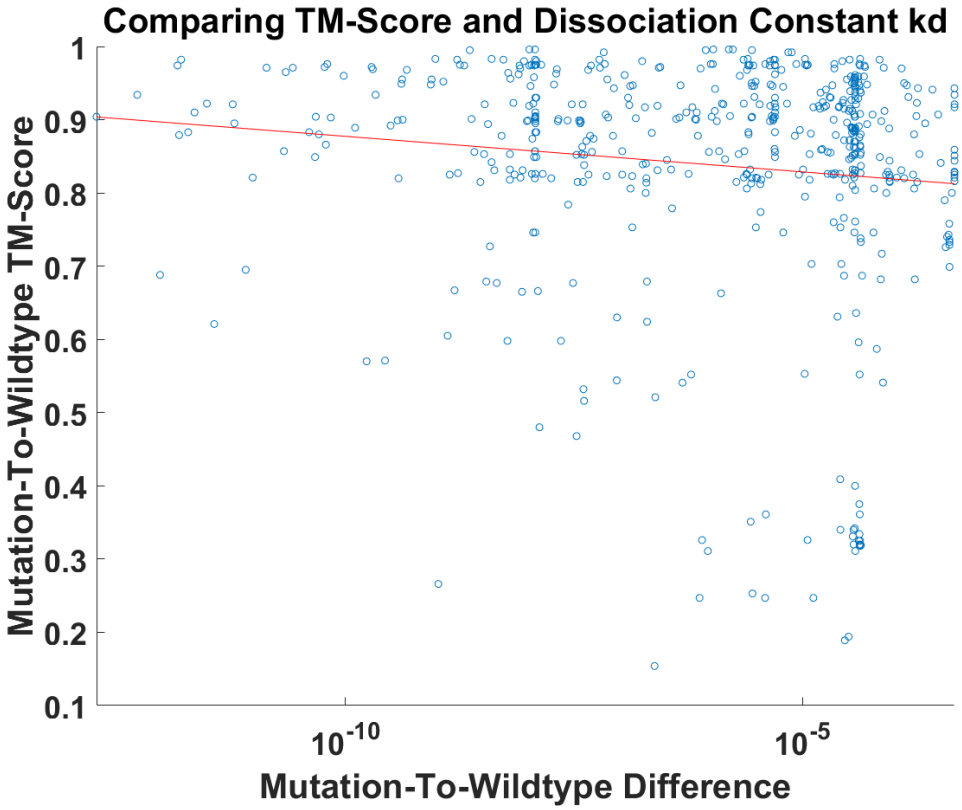


Figure 5.6: Comparing TM-Score for each mutation to its wildtype to the difference between their dissociation constants

5.4 Discussion

5.4.1 How accurate are our results?

A collection of association rates have been estimated for each of the interactions detailed in Table 5.1. The accuracy of our oligomeric and complex structures has been assessed in sections 5.3.1, 5.3.2 and 5.3.3. RMSD is a useful tool for evaluating the similarity between our computationally established structures and the experimentally derived native. While we used the TM-Score as an assessment method of protein similarity for our tertiary structures in Chapter 4, the RMSD is a more useful algorithm with our generated complexes due to the nature of the TM-score to assess superposition coverage as well as structural accuracy. Due to the sparse nature of experimentally derived proteins relevant to this study found in the PDB, structures with sequence identity >95% were chosen, while disregarding coverage. For example, the extracellular domain of epidermal growth factor receptor (EGFR) was used

while assessing interaction 9 (Appendix 3, Table 1) in the absence of whole protein data (Ogiso et al., 2002).

Homodimeric wildtype structures derived from amino acid sequences during this chapter have a mean RMSD of 3.015A, while our heterodimeric structures have a mean RMSD of 4.35A. The sample size is too small to make sweeping conclusions, however strong near-native structures have been established for our wildtype data, giving confidence in the structures generated using mutation data.

Six of the seven oligomeric structures predicted returned with an RMSD of less than 3.8A from the native. TGFBR shows a distinct difference from the native with an RMSD of 8.83A. This could be due to the low quality prediction of its constituent parts, with the C-score for TGFBR1-001 and TGFBR2-001 being -3.52 and -3.9 respectively.

5.4.2 Compound error

When pipelining quantitative prediction methods together, compound error can quickly become a problem. To get to the stage where we can predict association rates from amino acid sequences in this thesis, we have converted primary structure into tertiary structure, into quaternary structure (potentially involving oligomers), into an association and dissociation rate prediction. Each of these stages introduces a degree of uncertainty, decreasing confidence in our results. In addition, it is distinctly possible that a minor discrepancy early within the algorithmic pipeline can propagate into a major divergence from the expected result by the time all aspects of the methodology have been applied.

While we can assess accuracy for individual components of our methodology such as aligning predicted structures or association rates to experimentally validated ones, calculating how a variation in an uncertain area of a protein structure prediction will affect our k_{on} prediction is a significantly more challenging task. However, as confidence grows in the predictive power of each of the subsections of our pipeline, confidence grows in the complete methodology. Protein structure prediction methods and docking methods are improving year-on-year (Lensink et al., 2016), new structures are experimentally derived

and force-fields and continuum solvation methods used in association rate prediction are updated regularly (Baker et al., 2001; Dolinsky et al., 2007), leading to a continued refinement of the algorithm as a whole.

5.4.3 Why choose TransComp?

Comparison tests between SDA7 and TransComp showed comparable results between the two, with a slight quality improvement with TransComp results. In addition, SDA7 had larger computational times and a range of possible results were returned, only providing a scalar result with the addition of reaction criteria (i.e. the distance between proteins where the short-range interactions begin to occur and the number of contact points involved in the interaction). These reasons pushed us towards using TransComp to calculate association rates for our collection of interactions.

5.4.4 What can we do with this information?

Association and dissociation rate parameters are vital in drug development. Kinetics can be used to calculate equilibrium states and to estimate how quickly the system responds to external stimuli, such as a drug or another competitor. Within the context of systems biology, binding kinetics can be used to parameterize a mathematical model to provide an increased level of biological relevance. Currently, having ascertained a collection of association rates for a variety of mutations, existing pathways can be reparametrized referring to each mutation.

5.4.5 How can we make this better?

While sections of the described pipeline are state-of-the-art and utilize the strongest publically available algorithms developed by the computational biology community, improvements could be made to the procedure if computational intensity constraints were removed. Increasing the amount of time I-TASSER spends on refinement on structural areas with insufficient homologous data could lead to an increase in the quality of the input structures, and subsequently the output binding kinetics. Not all of the computationally-calculated docked structures were used as TransComp inputs. Due to computational intensity issues, IL12 binding to IL12R (interaction 13 in Table 5.1) was unable to be

completed in its full form. As both IL12 and IL12R are heterodimeric, the total number of atoms within the system became so large that each mutation would have taken more than a year to process on a single core system. PDGF suffered a similar fate, with both PDGFRA and PDGFRB being large receptors with 1090 and 1107 amino acids long respectively. Removal of computational constraints would have allowed for the calculation of these values. The quality of our quaternary structure is highly dependent on the availability of structural information leading to an informed docking simulation. Increasing the number of docks produced by both Rosetta and Hex to be slimmed down using their respective scoring functions would increase the possibility of a near-native structure being isolated. Rosetta global docking algorithms could also have been used instead of Hex if not for the great computational expense required in doing so. However, one of the biggest improvements that could be made would be through using flexible docking strategies instead of the rigid-body hybrid used in Rosetta local docking.

5.4.6 Diffusion limited interactions

While the TransComp-derived k_{on} values that we have ascertained range between 10^{-188} and 10^{11}, the TransComp methodology is more robust with association rates greater than $1x\ 10^{5}$. As defined in section 5.1.5, regimes where the transient complex is converted into the final complex significantly faster and more often than it unbinds into its constituent parts (or 'diffusion-limited') can be estimated more accurately with the TransComp methodology. This conformation happens when $10^{5} < k_{on} < 10^{10}$ (Qin et al., 2011). For slower proteins, where the short-range interactions required to finalize complex formation are less reliable, estimations used in TransComp become less biologically accurate. While we can have more confidence in our results as k_{on} is greater, values in the lower half of the spectrum are still useful for us later in this thesis. The estimation utilized in TransComp methodology of $k_{on} \approx k_D$ (the rate of transient complex formation) is no longer accurate, with $k_{on} \approx k_r * k_D$, where k_r is equal to the speed of complex formation from the transient complex divided by the rate of unbinding from the transient complex into constituent parts.

5.4.7 Future Work

The development of a pipeline to calculate association and dissociation rates from two input amino acid sequence has the potential to benefit mathematical modelling and computational drug design communities. Kinetic parameter availability is sparse for most known PPIs, and the ability to initially parameterize a pathway in the absence of experimentally derived values will be beneficial in initial development of mathematical models. Additionally, calculating binding kinetics for mutated structures can be used as part of a stratified medicine program when used to reparametrize disease pathways.

Chapter 6:

Population Subgroup Specific Atherosclerosis Dynamics

6.1 Introduction

Computational models are a means of solving a problem, not the end result. Without relating model results back to pathway dynamics to study the underlying biology and biochemistry then computational models, while interesting, serve little practical use. Computational models can be described as a tool to provide insight into disease dynamics. Ordinary differential equation (ODE) models of disease, such as the model of atherosclerosis defined in Chapter 3, relate variation in the quantity of one biological entity to variation in all others included within the model. Networks of biological processes which utilize experimentally derived data to maximize biological relevance have led to remarkable insights into disease pathogenesis (Akman et al., 2012; Fujita et al., 2014; Mazein et al., 2013; Mizuno et al., 2012; Morgan et al., 2016).

Development of technologies relating to '-omics' data has led to vast quantities of genomic, proteomic, metabolomic and transcriptomic data being generated within the last two decades. Placing this data in the public domain or facilitating methods for data sharing allows for the wider scientific community to benefit and utilize these data as part of a wide range of experiments (Chervitz et al., 2011). Similar to mathematical modelling, the challenge for computational biologists is to convert this massive dataset into information that can provide a beneficial research output. Combining mathematical models with '-omics' data, through adjustments such as reparameterisation or network optimization, can lead to new insights into biological processes. Adjusting quantitative mathematical models to align with variations noted in '-omics' data has the potential to enable systems pharmacology approaches based on stratified patient groups.

6.1.1 Using mathematical models to stratify patient groups

Reparameterising mathematical models as part of a program of personalized or stratified medicine is a fledgling field, however important work has been undertaken to demonstrate the possibilities. The core concept relies upon calculating a set of quantitative measurements of biological activity to stratify patient groups that can be applied to a stable and robust mathematical model allowing for the assessment of how pathway dynamics differ due to this stratification. Using a scoring system based on CD34 levels, mathematical

models have been developed which suggest a stratification method for patients with chronic myeloid leukemia (Brehme et al., 2016). Models for dopamine metabolism and folate-mediated one carbon metabolism have been developed and reparametrized based on experimentally-derived enzyme activity data, which has been connected to population genetics to provide a potential stratification methodology (Nijhout et al., 2015). Patient stratification has previously been achieved through reparameterisation of mathematical models, however gene expression levels were used rather than binding kinetics (Björnson et al., 2015). Combining enzyme kinetics and population genetics as part of a mathematical model reparameterisation program could lead to the discovery of different disease dynamics within subgroups, alongside the identification of biomarkers or optimal therapeutic responses.

6.1.2 Potential future uses of mathematical models

Computational models can provide a series of benefits regarding the elucidation of biological processes. Two-state discrete Boolean models have been used to represent biological processes such as apoptosis (Schlatter et al., 2009) and circadian rhythms (Akman et al., 2012). Fully quantitative models, such as ODE models, can generate therapeutic hypotheses such as the identification of predictive biomarkers and the generation of potential multi-drug therapies. Stratifying patient groups based on biomarker data and subsequent therapeutic selection have seen an increase in the success of breast cancer treatment (Yeo et al., 2014). Through the development of mathematical models of small sections of biological systems, and combining them once the underlying dynamics are well understood, larger models are being created by the systems biology community to represent whole biological systems. A whole cell model of *Mycoplasma genitalium* has been produced (Karr et al., 2012a). Ideally, if computational models can be developed to accurately represent the underlying physiology while representing whole biological systems, an *in silico* clinical trial could be developed before patient trials, giving insight into drug efficacy, potential side effects and allowing for study into the particular subtype or manifestation of the disorder.

6.1.3 Systems Pharmacology

Mathematical modelling has led to optimization of multi-drug therapeutics in multiple diseases. Models have been developed of breast cancer signaling networks where an optimal therapeutic blockade was hypothesized to bypass compensatory circuits and limit tumor growth, a hypothesis which was subsequently validated on mice (Kirouac et al., 2013). Systems pharmacology methods have been used to suggest drug targets for the treatment of Alzheimer's disease (Nicholas et al., 2013). Using systems pharmacology methods to study the effects of common drugs on atherosclerosis progression has been done before using a simplified model considering foam cells, macrophages, oxidized LDL, endothelial cell disruption and haemodynamics (Pichardo-Almarza et al., 2015).

6.1.4 Binding Kinetics

The atherosclerosis model as described in Chapter 3 contains a series of equations following the law of mass action and Michaelis-Menten kinetics. The law of mass action considers a simple interaction between two reactants A and R (Voit et al., 2015):

$$A + R \rightleftarrows AR$$

where the reaction velocity is modelled as proportional to the reactant quantities:

Equation 6-1

$$\frac{dAR}{dt} = k_{cat} * A * R$$

Whereas Michaelis-Menten equations take this a little further and include an intermediary step and a final product P (Johnson and Goody, 2012; Michaelis and Menten, 1913):

$$A + R \rightleftarrows AR \rightarrow A + P$$

Connecting the relationship between the quantities of reactants and products in this manner allows for the development of models of whole biological systems; however a collection of assumptions are required in the development of a model of this type. This system considers an interaction with only one transitional state AR, and binding kinetics

would differ in a system with an increased number of transitory structures. It is assumed that protein A is not altered by binding and is freely available to bind to other entities once this interaction is complete, and that the transitory state binding can be reversed. Finally, it is presumed that all proteins in this interaction are uniformly accessible to each other. Within a system where these assumptions are true, the reaction velocity is modelled as:

$$\frac{dAR}{dt} = k_e * A_0$$

Where A_0 is the total concentration of protein A, and k_e is defined as:

$$k_e = \frac{k_{cat} * R}{k_M + R}$$

where the Michaelis Constant k_M is defined as:

Equation 6-2

$$k_M = \frac{k_{off} + k_{cat}}{k_{on}}$$

A derivation for these equations can be found in (Briggs and Haldane, 1925; Michaelis and Menten, 1913).

6.1.5 Computational Inhibition

For an interaction in a system described in section 6.1.4, an inhibitor can be introduced that can lead to a reduction in reaction velocity. These substances can be a natural part of a homeostatic biological system, such as epidermal growth factor (EGF) inhibiting SMC production of elastin (Ichiro et al., 1990), or as an external source like a targeted drug. Modelling the introduction of a drug into a biological system designed to compete with other reactants for a particular receptor can be completed with an extension to the Michaelis-Menten methodology:

$$\begin{array}{ccccc} A + R & \rightleftarrows & \mathrm{AR} & \rightleftarrows & \mathrm{A} + \mathrm{P} \\ \uparrow\downarrow & & \uparrow\downarrow & & \\ RI & & ARI & & \end{array}$$

Where A and R are two reactants and I is an inhibitor.

Competitive inhibition such as this is caused by A and I competing to bind to the same protein R. Reaction velocity in this scenario is modelled as:

$$\frac{dP}{dt} = \frac{k_{cat}[A][R]}{k_M\left(1 + \frac{I}{k_I}\right) + [R]}$$

Where I = Concentration of the inhibitor; k_I = inhibition constant. The inhibition constant is defined the concentration where half of the maximum inhibition is produced.

6.1.6 Scoring Function and Multi-Drug Therapeutic Hypotheses

Due to the complex nature of atherosclerosis and the quantity of biological entities included within the model of atherosclerosis as defined in Chapter 3, a scoring function is required to represent atheroma severity. Creating such a function and value allows for the direct comparison of atheromata simulated using the model and allows for the development of a system to minimize atheroma size based on theoretical drugs added to the system.

A scoring function is a mathematical method of simplifying a multi-dimensional system to a scalar value, vital when assessing the impact of perturbations to the system. When searching a large parameter space, such as finding an optimal multi-drug therapy, finding an accurate global minimum can be a challenging process. The solution space within our multi-drug system is ‘bumpy’, that is to say that many local minima exist, and as such finding the true global minimum of the problem requires some nuance. Simulated annealing is an optimization technique relying on probabilistic methods to find a global minimum where multiple local minima exist. Genetic algorithms are a technique which imitates the biological process of natural selection, where a parameter space is searched to find an optimized solution to a similar problem. A 'population' of individual solutions are generated according to a scoring function. At each generation, individual solutions are chosen at random to be

produce a solution for the next generation. As the number of generations increases, solutions evolve towards an optimal solution. Genetic algorithms have been shown to return a more accurate evaluation at an increased computational cost for circuit partitioning problems (Manikas and Cain, 1996) and a similar result with a reduced computational cost for regression problems concerning neural network ensembles (Soares et al., 2013).

6.1.7 Population genetics

Ascertaining genetic differences between population subgroups can be beneficial during computational analyses relating to stratified medicine. Quantification of changes in molecular interactions due to genetic variance can be used to reparameterize mathematical models, potentially leading to population specific insights into disease pathogenesis. Changes in binding kinetics in one specific interaction can cause unforeseen downstream effects, seemingly unrelated to the initial interaction. Through systems biology methods and mathematical modelling, downstream effects can be predicted and therapeutic hypotheses developed for particular interactions. Population genetics can provide insight into which interactions are likely to differ in separate population subgroups.

Within phase 3 of the 1000 Genomes Project, sequence and variant data exists for 5 super-populations consisting of a total of 26 different population subgroups (Auton et al., 2015). These populations are shown in Table 6.1.

Table 6.1: Five super-populations and twenty-six population subgroups considered within Phase 3 of the 1000 Genome Project

African	Yoruba – Nigeria
	Luhya – Kenya
	Gambian
	Mende – Sierra Leone
	Esan – Nigeria
	African-Americans in the south-west of the USA
	African-Caribbeans - Barbados

American – Mixed	Mexican-Americans in the south-west of the USA
	Puerto Rican
	Colombian
	Peruvian
East Asian	Han Chinese – China
	Japanese
	Southern Han Chinese – China
	Chinese Dai – Xishuangbanna, China
	Kinh - Vietnam
European	European Americans – Utah, USA
	Toscani - Italy
	Finnish
	British
	Iberian - Spain
South Asian	Gujarati Indian – Texas
	Punjabi – Pakistan
	Bengali – Bangladesh
	Sri Lankan Tamil – UK
	Indian Telugu - UK

6.1.8 Relating plaque morphology to severity

While the scoring function introduced in section 6.2.2 is an appropriate method of giving a snapshot of atheroma severity due to each of our mutations, this simplification is likely to avoid reporting on subtleties hidden within the model. Atheroma size could be related to the lesion content of macrophages, foam cells and smooth muscle cells. Collagen, elastin and other ECM elements are vital when considering plaque volume and lesion stability. Concentrations of MMPs and their inhibitors are key when considering chance of plaque rupture. Plasma concentrations of biological entities have the potential to be useful biomarkers. When searching the vast quantities of data isolated from our collection of

mathematical models developed within this chapter, mutations leading to variations in these quantities are more likely to be beneficial and clinically relevant.

6.2 Methods

The model of atherosclerosis as defined in Chapter 3 has been used as a basis for this study. A detailed list of interactions studied can be seen in Chapter 5 (Table 5.1). Parameters altered to simulate the effect of mutations can be found in Appendix 6.

6.2.1 Calculating k_M and k_{cat}

Kinetic parameters k_M and k_{cat} will be altered by the changes in k_{on} and k_{off} calculated in Chapter 5. Equation 6-2 (section 6.1.4) has been used to calculate how this mutation data will alter k_{cat} and k_M, to allow for per-mutation reparameterisation of the mathematical model. However, due to compound error in the derivation of binding kinetics from predicted structure and the estimation of certain kinetic parameters used in the model, the equality requires a correction factor in order to be valid. This correction factor, k_{corr}, is added as follows:

Equation 6-3

$$k_M = \frac{{k_{cat}}/{k_{corr}} + k_{off}}{k_{on}}$$

The correction factor k_{corr} was applied specifically to k_{cat} to prevent cases of a negative k_{cat} value being generated within our dataset.

For each interaction, k_{corr} has been calculated for the interaction between wildtypes, where k_{on} and k_{off} have been calculated as described in Chapter 5, and k_{cat} and k_M have been isolated from the mathematical model. Kinetic parameters k_{cat} and k_M are calculated for each mutation by combining the wildtype-derived k_{corr} with mutation-specific k_{off} and k_{on}.

Within the model of atherosclerosis defined in Chapter 3, interactions involving protein-protein interactions are formulated in one of three ways:

6.2.1.1 Law of Mass Action

Interactions involving mass action kinetics are modelled in the form:

$$v = k * A * R$$

Where k = k_{cat} / k_M; v = reaction velocity; A and R are reactants within the interaction. Through rearrangement of Equation 6-3, k_{cat} is calculated by:

Equation 6-4

$$k_{cat} = (k_{on} * k_M - k_{off}) * k_{corr}$$

Where k_M is set to be 1.

6.2.1.2 Michaelis-Menten Kinetics

Interactions involving Michaelis-Menten kinetics are modelled in the form:

$$v = \frac{V_{max} * R}{k_M + R}$$

Where V_{max} = k_{cat} * E_0; E_0 = Initial enzyme concentration (Johnson and Goody, 2012; Michaelis and Menten, 1913).

For each interaction that is modelled using Michaelis-Menten kinetics, k_{cat} is calculated by:

Equation 6-5

$$k_{cat} = (k_{on} * k_M - k_{off}) * k_{corr}$$

6.2.1.3 Michaelis-Menten Kinetics Involving Multiple Entities

Interactions involving Michaelis-Menten kinetics with multiple biological entities are modelled in the form:

$$v = \frac{k * E * S_1 * S_2 \ldots}{(k_{M1} + S_1)(k_{M2} + S_2) \ldots}$$

When focusing on a mutation within the interaction involving S_1, k_{cat} is calculated using:

Equation 6-6

$$k_{cat} = (k_{on} * k_{M1} - k_{off}) * k_{corr}$$

6.2.2 Scoring function

To provide a quick insight into the disease altering dynamics of a mutation, a scoring function has been developed to provide a scalar value representing a combination of atheroma size, cell density, protein content and collagen content. To assess the quality of a simulation, a cell count and the collagen content of total intimal protein (as a percentage) are calculated. The cell count includes total abundances of smooth muscle cells, macrophages, foam cells and T cells. The atheroma severity score, or A_{Score}, can be calculated as follows:

Equation 6-7

$$A_{Score} = \frac{\left(\frac{C}{C_{max}}\right) + \left(\frac{T_{Coll}}{M_{max}}\right)}{2}$$

Where C = Total Cell Count; C_{max} = 88479; T_{Coll} = Total Collagen Percentage; M_{max} = 0.7248;

C_{max} and M_{max} are the total cell abundance and percentage of total protein as collagen for a simulation over 80 years with an LDL concentration of 190mg/dl and an HDL concentration of 45mg/dl.

6.2.3 Generation of mutation-specific models

For each k_{on}-k_{off} pair representing a mutation as described in Chapter 5 and detailed in Appendix 5, the corresponding model equation has been altered and the model dynamics simulated (a full list of model equations can be found in Appendix 3). Of the 835 k_{on} values calculated, k_{cat} values were calculated where an appropriate k_{off} value exists. Each of these mutation k_{cat} values were paired with the interaction k_{corr} as defined in section 6.2.1. Wildtype values were successfully calculated for all interactions excluding MMP1→TIMP3, MMP1→TIMP4 and PDGF→ PDGFR, and as such mutations involving these interactions were not included in our dataset. Wildtype values were unobtainable for these three interactions after TransComp failure due to an expected intermediary stage when binding. Additionally, interactions were not considered where TransComp returned a results where $k_{on} < 1$. Observed association rate values range from in the order of 10^2 to 10^{10}, so removing all values lower than 10^0 was considered reasonable. After this trimming down of the dataset, 746 k_{on}-k_{off} pairs remained for reparameterisation. The A_{score} as defined in section 6.2.2 was calculated for each mutation as a quantification of atheroma severity.

6.2.4 Genetic Profiles

For each of the 31 populations described in Table 6.1, a genetic profile has been created to represent common mutations found within the collection of proteins considered within the model (as detailed in Table 4.2). Each population starts with a genetic profile consisting of wildtypes for each model protein. Population genetics data was isolated from the 1000 Genome Project using the EnsEMBL API as described in Section 4.2.1. For each population, a search was undertaken to isolate mutations within phase 3 of the 1000 Genome Project with a > 50% variant allele frequency (VAF). If such an allele exists for a population subgroup, their genetic profile is updated to include this mutation. Once complete, atherosclerosis dynamics were simulated for each genetic profile to allow for the study of population-specific pathogenesis differences.

6.2.5 Multi-Drug Therapies

For each of our genetic profiles generated as described in section 6.2.4, a multi-drug therapeutic optimization process is undertaken. A collection of 15 drugs shown in Table 6.2

are applied to each of our population specific models through inhibition equations as described in section 3.2.2.

Table 6.2: Collection of 15 drugs included in multi-drug therapy in silico minimization experiment

Interaction blocked	Drug	Citation
IL12 -> IL12R	Ustekinumab	(Sandborn et al., 2008)
IL18 -> IL18R		(Krumm et al., 2017)
IL4 -> IL4R		(Steinke, 2004)
IL10 -> IL10R	Rituximab	(Alas et al., 2001)
TGFB -> TGFBR	Galunisertib	(Herbertz et al., 2015)
IL6 -> IL6R	Tocilizumab	(Sebba, 2008)
IL21 -> IL21R		(Young et al., 2007)
MMP1 -> TIMP1 MMP1 -> TIMP2 MMP1 -> TIMP3 MMP1 -> TIMP4		(De Andrade Leite, 2009)
MMP2 -> TIMP1 MMP2 -> TIMP2 MMP2 -> TIMP3 MMP2 -> TIMP4	BiPS	(Lauzier et al., 2008)
MMP3 -> TIMP1 MMP3 -> TIMP2 MMP3 -> TIMP3 MMP3 -> TIMP4	NNGH	(Whitlock et al., 2007)
MMP9 -> TIMP1 MMP9 -> TIMP2 MMP9 -> TIMP3 MMP9 -> TIMP4	BiPS	(Lauzier et al., 2008)
MMP13 -> TIMP1 MMP13-> TIMP2 MMP13-> TIMP3 MMP13 -> TIMP4	CL 82198 hydrochloride	(Engel et al., 2005)
EGF -> EGFR	Tyrphostin 47	(Levitzki and Gazit, 1995)
oxLDL -> Foam Cells *Target is CD36*		(Geloen et al., 2012)
CXCL9 -> CXCR3 CXCL10 -> CXCR3 CXCL11 -> CXCR3		(Heise et al., 2005)

A genetic algorithm is applied to optimize the reduction of atheroma size and severity (represented using the A_{Score} as defined in section 6.2.2), while attempting to minimize the

number of drugs required to perform this task within this system. The drugs considered are the detailed in Table 6.2. The scoring function minimized by the genetic algorithm is:

$$MD_{score} = A_{Score} + 0.045D + 0.001 * \sum_{n=1}^{D} C(n)$$

Where the A_{Score} is as defined in section 6.2.2, D is the number of drugs applied to the system, and C(*n*) is the concentration of drug *n*. Drug concentrations were set to be between 0 and 1000 mM, and all k_i values were assumed to be 1 mM. Simulations were performed in Matlab using the ga() method, and the following parameters were included within the simulation.

Table 6.3: Genetic Algorithm Parameters

Generations	500
Population Size	2500
TimeLimit	18000 (in seconds)
InitialPopulation	0 for each drug

Simulations were performed on an Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz (Octo-core) CPU with 64GB of RAM running CentOS 7.

6.3 Results

6.3.1 Genetic Profiles for Population Subgroups

In order to develop a description of mutations that are representative of a population, genetic profiles were developed as described in section 6.2.4. In situations where multiple mutations exist for the same gene, and ultimately the same protein, the most common mutation was selected as part of the genetic profile. The 31 genetic profiles are shown below.

Table 6.4: Mutations included in genetic profiles created for model reparameterisation

Population Subgroup	Mutations
Han Chinese – China	EGFR-001-521 (rs2227983) MMP9-001-279 (rs17576)
Japanese – Japan	CCR2-001-64 (rs1799864) EGFR-001-521 (rs2227983)
Southern Han Chinese – China	CSF1R-001-362 (rs10079250) MMP9-001-279 (rs17576)
Chinese Dai – China	EGFR-001-521 (rs2227983) MMP9-001-279 (rs17576)
Kinh – Vietnam	EGFR-001-521 (rs2227983) MMP9-001-279 (rs17576)
European Americans – Utah, USA	IL10RA-001-351 (rs2229113)
Toscani – Italy	IL12RB1-001-214 (rs11575934)
Finnish – Finland	CSF1-001-408 (rs1058885) IL12RB1-001-214 (rs11575934)
British – UK	None
Iberian – Spain	None
Yorubi – Nigeria	IL4R-001-576 (rs1801275)
Luhya – Kenya	IL4R-001-576 (rs1801275)
Gambian – Gambia	IL4R-001-576 (rs1801275)
Mende – Sierra Leone	IL4R-001-576 (rs1801275)
Esan – Nigeria	IL4R-001-576 (rs1801275)
African Americans – USA	None
African Caribbeans - Barbados	IL4R-001-576 (rs1801275)
Mexican Ancestry – USA	None
Puerto Ricans – Puerto Rico	None
Colombian – Colombia	None
Peruvian – Peru	None
Gujarati Indian – Texas, USA	EGFR-001-521 (rs2227983)

Population Subgroup	Mutations
Punjabi - Pakistan	None
Bengali – Bangladesh	None
Sri Lankan Tamil – UK	None
Indian Telugu – UK	None
East Asian	EGFR-001-521 (rs2227983) MMP9-001-279 (rs17576)
European	None
African	IL4R-001-576 (rs1801275)
American	None
South Asian	None

6.3.2 New k_{cat} values

For each mutation that yields a change in model kinetics an updated k_{cat} value was created, calculated as described in section 6.2.1. An exhaustive list of these updated k_{cat} values can be found in Appendix 6. For the individual proteins contained within the genetic profiles derived in section 6.3.1, reparameterisation details can be found in Table 6.5.

Table 6.5: Reparameterisation details for mutations contained within genetic profiles

Name	Predicted k_{on}	Predicted k_{off}	Model k_{cat}	Model k_M	Updated k_{cat}	Equation No
EGFR-001-521	2.07E+5	2.0358	1938	100000	31341	58
MMP9-001-279	5.95E+5 3.98E+1 8.88E+7 8.55E+2	109.6585 0 7.722 4.22E-06	1.0E-7 1.0E-7 1.0E-7 1.0E-7	MA MA MA MA	8.26E-4 1.06E-7 8.07E-2 2.23E-11	50 51 52 53
CCR2-001-64	2.87E+5	3.34E-06	0.0011 9000	MA 500	0.000162 1388	16 65
CSF1R-001-362	7.32E+4	0.32838	0.0994	MA	5.31E-6	22
IL10RA-001-351	9.10E+04	3.10E-16	8.1882	100000	7.53	40

			0.644 0.0063	100000 100000	0.5926 0.0058	41 44
IL12RB1-001-214	4.15E+03	1.43E-18	8.0970	100000	21.4029	39
CSF1-001-408	5.13E+04	0.063048	0.0994	MA	0.1567	22
IL4R-001-576	4.12E+05	52.3652	8.1882	100000	288.3366	40

6.3.3 A collection of atherosclerosis models

The majority of mutations studied as defined in section 6.2.3 show a minor change in model dynamics. 86% of mutations (644 out of 746) studied are within an A_{score} range of 0.5-1.5.

The median A_{score} is 1, the mean score is 1.2223 and values range from 0.0147-90.0286. The top 10 most atheroprotective and atherogenic mutations are shown in Table 6.6.

Table 6.6: Top atheroprotective and atherogenic mutations within dataset

Top 10 Atheroprotective Variants	**A_{score}**	**Top 10 Atherogenic Variants**	**A_{score}**
CCR2-001-355	0.0147	CXCR3-002	90.0790
CCR2-002-64	0.0171	CXCL11-003-29	46.2952
CCL2-001-69	0.0207	CXCL11-001-73	18.5207
CSF1-002-461	0.0228	CXCL11-003	12.4330
CXCL10-001-58	0.0592	CXCL9-001-101	6.1296
CCL5-001-68	0.0625	CCL5-002-5	4.3096
CCL5-002-68	0.0625	CCL5-001-5	4.3096
IL4-201-134	0.0675	MMP9-001-279	1.7972
CCL5-001-40	0.0679	TIMP2-008-92	1.7972
CCL5-002-40	0.0679	TIMP1-002-10	1.7971

6.3.4 Population subgroup specific results

Utilising the subgroup specific genetic profiles, the model of atherogenesis was simulated over a period of 80 years, with genetic profile changes made as described in section 6.3.1.

Of the 31 populations included within phase 3 of the 1000 Genome Project, we have created 8 unique genetic profiles:

6.3.4.1 Profile 1 - Han Chinese from China, Chinese Dai from China and Kinh from Vietnam

Mutations in EGFR and MMP9 were found to be prevalent in more than 50% of individuals from these populations within the 1000 Genome Project dataset. MMP9 is a collagenase involved in the degradation of collagen and elastin. EGF is a key player in the regulation of elastin synthesis. Variation within rates of extra-cellular matrix synthesis and degradation within the *tunica intima* can have a distinct effect on plaque stability. These mutations demonstrate a distinct increase in concentrations of extra-cellular matrix within the plaque while having no effect on cellular contents.

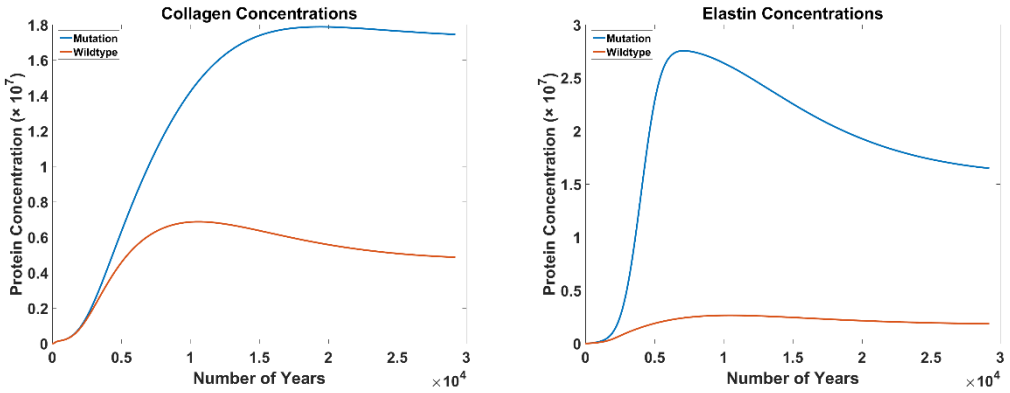


Figure 6.1: Plots showing the evolution of collagen and elastin concentrations over time for genetic profile 1 compared to the wildtype profile

6.3.4.2 Profile 2 - Japanese from Japan

Mutations in CCR2 and EGFR were identified as suitable for this genetic profile. Of all of the genetic profiles studied, this is the sole profile which breaks the stability of our model. The reduction in the rate of monocyte recruitment as seen due to the effect of mutation CCR2-001-64 has caused a severe reduction in the rate of monocyte recruitment, and subsequently atheroma size. Quantitatively, it is clear that this result is not representative of atherogenesis within a Japanese population. This result suggests that the mechanisms

included within the model representing monocyte recruitment are incomplete and should be expanded to prevent a single point of failure within the model.

However, qualitative results can still be gleaned from this methodology. A reduction in the rate of monocyte recruitment within our mathematical model would lead to a reduction in atheroma size, suggesting that a mutation prevalent within Japanese populations may be fundamentally atheroprotective.

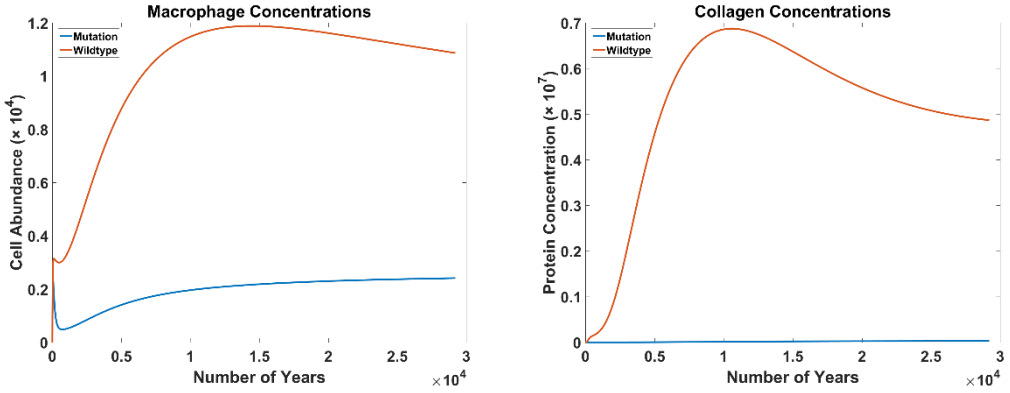


Figure 6.2: Plots showing the evolution of macrophage and elastin concentrations over time for genetic profile 2 compared to the wildtype profile

6.3.4.3 Profile 3 - Southern Han Chinese from China

MMP9 and CSF1R mutations were found to be prevalent in Southern Han Chinese populations. Alongside the lesion stability consequences caused by MMP9, a mutation in CSF1R alters the rate by which monocytes will differentiate into macrophages. This mutation has caused a slowing in the rate of differentiation, reducing the relevant k_{cat} value from 9.94 $x10^{-2}$ to 5.31 $x10^{-6}$. This mutation seems to have caused a minor effect on the cellular composition of the atheroma within the model.

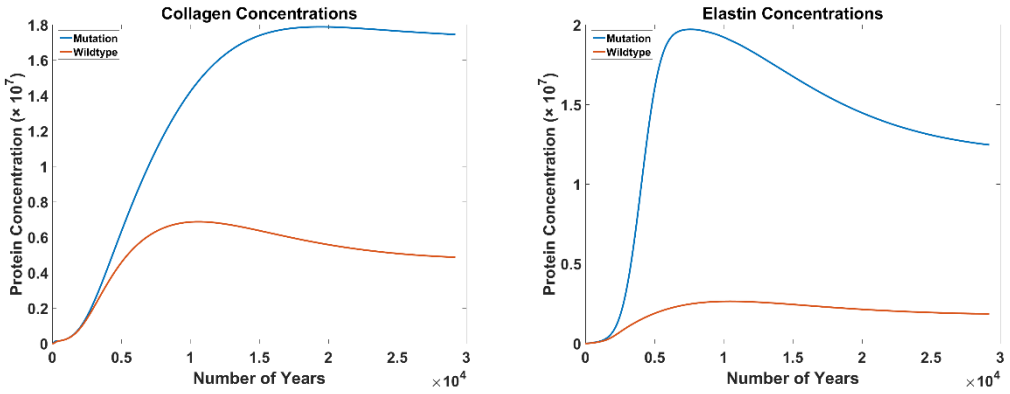


Figure 6.3: Plots showing the evolution of Collagen and Elastin concentrations over time for genetic profile 3 compared to the wildtype profile

6.3.4.4 Profile 4 - European Americans from Utah, USA

European Americans were found to have a high frequency of mutation in IL10RA within our dataset. Experimentally, the IL10 Receptor has been shown to be tetrameric (Verma et al., 2016), however to reduce the computational difficulty of calculating association rates for this interaction we only considered one alpha chain of the receptor. Within our model, this mutation has caused a minor atheroprotective effect, however this effect is almost negligible.

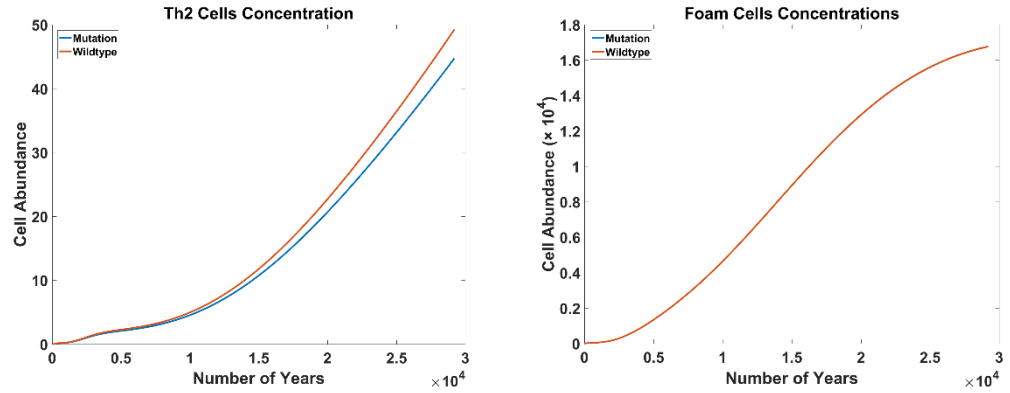


Figure 6.4: Plots showing the evolution of Collagen and Elastin concentrations over time for genetic profile 4 compared to the wildtype profile

6.3.4.5 Profile 5 - Toscani from Italy

Toscani individuals were shown to have a known mutation in IL12RB1. Similarly to IL10R as discussed in the European American dataset described in section 6.3.4.4, this receptor is

known to be experimentally heterodimeric, however to reduce computational costs in estimating association rates then the chain IL12RB1 was exclusively used. Interleukin 12 is a promotor of T_h1 differentiation in humans, leading to increased IFNG production driving atherogenesis.

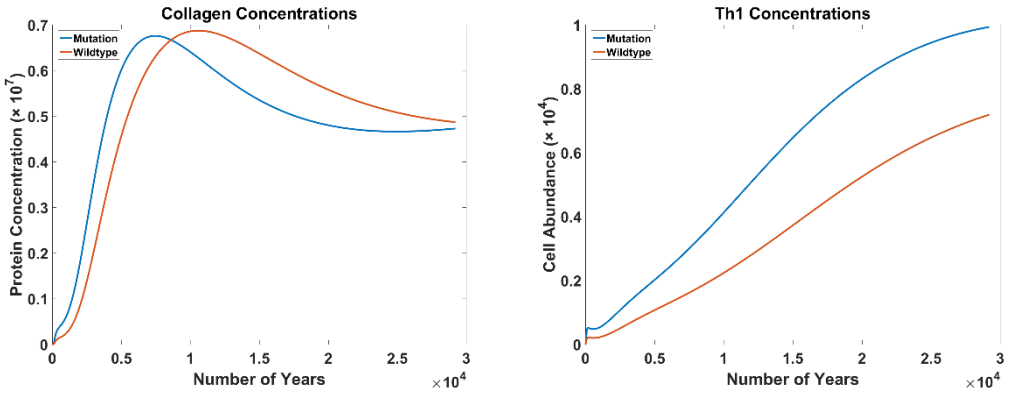


Figure 6.5: Plots showing the evolution of collagen and T_h1 cells over time for genetic profile 5 compared to the wildtype profile

6.3.4.6 Profile 6 - Finnish from Finland

Our Finnish genetic profile contains the same mutation in IL12RB1 as our Toscani population, alongside a mutation in macrophage colony stimulating factor. This CSF1 mutation causes a very minor pro-atherogenic effect, however the effect is so small that it is almost insignificant. This effect is engulfed by the mutative effect of IL12RB1 within this genetic profile.

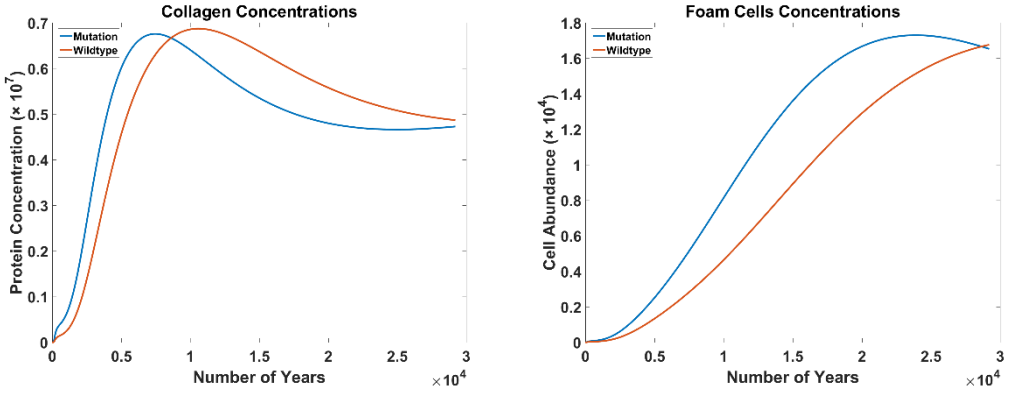


Figure 6.6: Plots showing the evolution of collagen and foam cells over time for genetic profile 6 compared to the wildtype profile

6.3.4.7 Profile 7 - Yorubi from Nigeria, Luhya from Kenya, Gambian from Gambia, Mende from Sierra Leone, Esan from Nigeria and African Caribbeans from Barbados

IL4R-001-576 is a mutation that was shown to be widespread within six of the population subgroups identified as part of the 1000 Genome Project. IL4 is a cytokine that encourages T_h2 differentiation. The T_h1-T_h2 balance is important within atherosclerosis, with increased T_h2 activity reducing the severity of the atheroma. This mutation has shown an anti-atherogenic effect through the reduction of cell abundance throughout the atheroma.

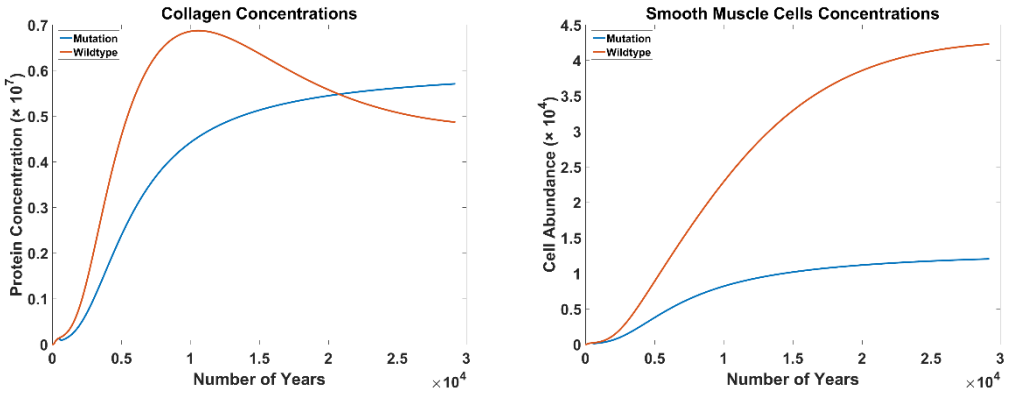


Figure 6.7: Plots showing the evolution of collagen and smooth muscle cells over time for genetic profile 7 compared to the wildtype profile

6.3.4.8 Profile 8 - Gujarati Indian from Texas, USA

EGFR is an important component of elastin synthesis. EGFR-001-521 is the mutation (also common within Han Chinese, Chinese Dai and Kinh populations) that is the sole constituent of our Gujarati Indian from Texas, USA genetic profile. This mutation causes a minor increase in elastin concentration within the *tunica intima*, increasing plaque size while also increasing plaque stability.

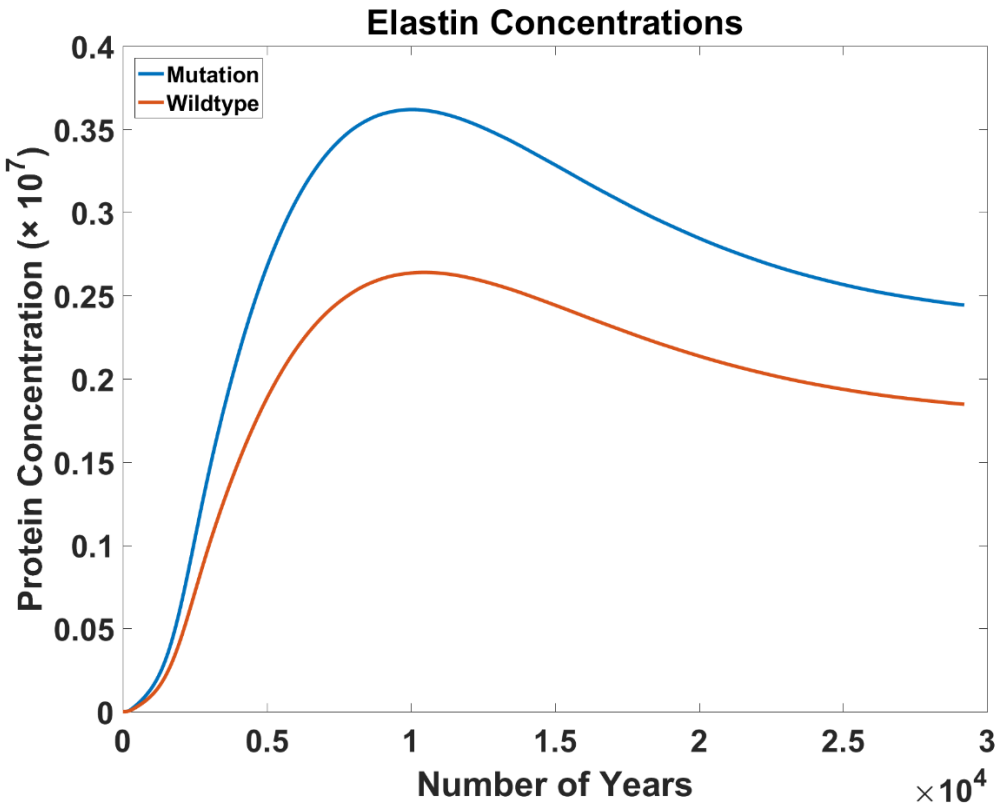


Figure 6.8: Plot showing the evolution of elastin concentrations over time for genetic profile 8 compared to the wildtype profile

6.3.4.9 All Others

Each of the other population subgroups included in this study showed no mutations that occurred in more than 50% of the subgroup members.

6.3.5 Population subgroup therapy optimization.

For each of our population subgroup models, atherosclerosis dynamics were simulated as described in section 6.2.4. Using each of our population subgroup models, a genetic algorithm was applied to calculate the multi-drug combination designed to minimize plaque size for each of our population subgroups as detailed in section 6.3.1. Results are shown below.

6.3.5.1 Wildtype Profile

The optimal multi-drug combination ascertained using a genetic algorithm. was:

Drug Target	Concentration
oxLDL → CD36	20.0704 mM
IL12 → IL12R	33.0863 mM

Returning a final MD_{score} of 0.3067. Applying no drugs to the system would provide an MD_{score} of 1, showing a significant reduction as a result of this multi-drug combination.

6.3.5.2 Profile 1 - Han Chinese from China, Chinese Dai from China and Kinh from Vietnam

The optimal multi-drug combination ascertained using a genetic algorithm was:

Drug Target	Concentration
IL18 → IL18R	12.1153 mM
IL12 → IL12R	12.2492 mM

Returning a final MD_{score} of 0.2585. Applying the optimal multi-drug combination for the wildtype profile returns an MD_{score} of 0.2860, suggesting an improved therapeutic response by switching to drugs with target the T_h1/T_h2 balance instead of limiting foam cell formation for this genetic profile.

6.3.5.3 Profile 2 - Japanese from Japan

The optimal multi-drug combination ascertained using a genetic algorithm was:

Drug Target	Concentration
IL12 → IL12R	14.78 mM

Returning a final MD_{score} of 0.1636. . Applying the optimal multi-drug combination for the wildtype profile returns an MD_{score} of 0.1903, suggesting that targeting oxLDL phagocytosis by immune cells is less efficacious with this genetic profile due to immune cell concentration being reduced due to the mutation CCR2-001-64.

6.3.5.4 Profile 3 - Southern Han Chinese from China

The optimal multi-drug combination ascertained using a genetic algorithm was:

Drug Target	**Concentration**
IL18 → IL18R	13.1614 mM
IL12 → IL12R	13.0325 mM

Returning a final MD_{score} of 0.2694. Wildtype therapeutics give an MD_{score} of 0.2904, so our optimal therapies for genetic profile 3 are similar to genetic profile 1.

6.3.5.5 Profile 4 - European Americans from Utah, USA

The optimal multi-drug combination ascertained using a genetic algorithm was:

Drug Target	**Concentration**
oxLDL → CD36	20.0704 mM
IL12 → IL12R	33.0863 mM

Returning a final MD_{score} of 0.3066. Alternative multi-drug therapeutics were not found for this population subgroup.

6.3.5.6 Profile 5 - Toscani from Italy

The optimal multi-drug combination ascertained using a genetic algorithm was:

Drug Target	**Concentration**
oxLDL → CD36	11.3650 mM
IL12 → IL12R	15.2277 mM
IL18 → IL18R	15.0994 mM

Returning a final MD_{score} of 0.3268. Wildtype therapeutics give an MD_{score} of 0.3713, suggesting that adding an additional drug could reduce atheroma risk for individuals in genetic profile 5, although increasing the potential for side effects.

6.3.5.7 Profile 6 - Finnish from Finland

The optimal multi-drug combination ascertained using a genetic algorithm was:

Drug Target	**Concentration**
oxLDL → CD36	11.3638 mM
IL12 → IL12R	15.2290 mM
IL18 → IL18R	15.1005 mM

Returning a final MD_{score} of 0.3268. Wildtype therapeutics give an MD_{score} of 0.3713, giving identical scores to genetic profile 5. The mutation CSF1-001-408, being the sole difference between our Finnish and Toscani genetic profiles, is proving to be negligible in multi-drug therapeutic hypotheses development.

6.3.5.8 Profile 7 - Yorubi from Nigeria, Luhya from Kenya, Gambian from Gambia, Mende from Sierra Leone, Esan from Nigeria and African Caribbeans from Barbados

The optimal multi-drug combination ascertained using a genetic algorithm was:

Drug Target	**Concentration**
oxLDL → CD36	12.1494 mM
IL12 → IL12R	13.2339 mM

Returning a final MD_{score} of 0.2452. Wildtype therapeutics give an MD_{score} of 0.2633, suggesting that a reduction in drug concentration may be beneficial to individuals in genetic profile 7. It is likely that this is simply a quirk of the MD_{score} function however.

6.3.5.9 Profile 8 - Gujarati Indian from Texas, USA

The optimal multi-drug combination ascertained using a genetic algorithm was:

Drug Target	**Concentration**

IL18 → IL18R	21.6088 mM
IL12 → IL12R	21.4817 mM

Returning a final MD_{score} of 0.2954. Wildtype therapeutics provide an MD_{score} of 0.2974, suggesting a very minor difference in benefit through patient stratification with this genetic profile.

Our multi-drug therapeutic experiments show that drug combinations exist to provide a reduced MD_{score} for seven out of our eight genetic profiles. Figure 6.9 compares the effect of wildtype drug combinations to patient subgroups combinations.

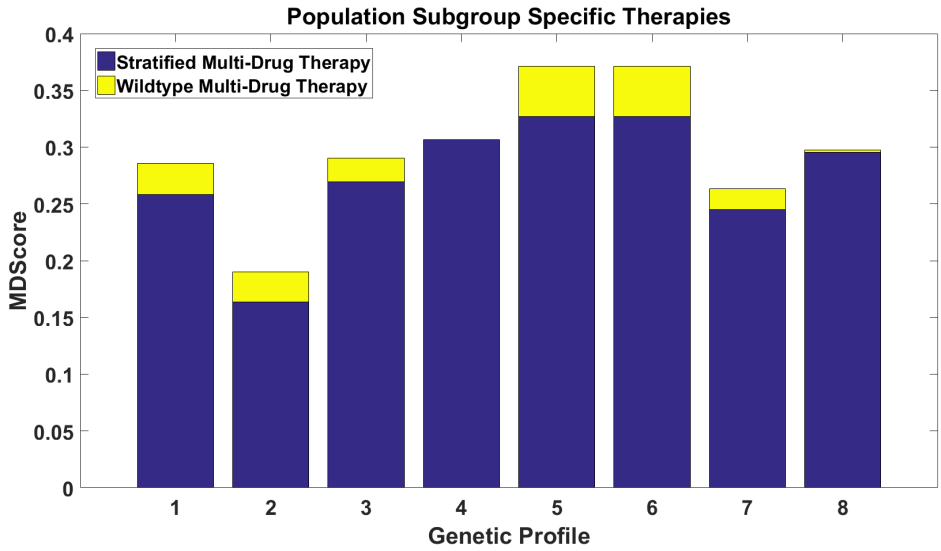


Figure 6.9: MD_{score} after wildtype and subgroup-specific therapies. Population subgroup specific therapies show a reduction in atheroma size

6.4 Discussion

6.4.1.1 Proatherogenic Mutations

Eight of the ten most atheroprotective variants involve reducing the rate of which immune cells are brought to the atheroma site. Proteins CCR2-001-355, CCR2-002-64, CCL2-001-69, CXCL10-001-58, CCL5-001-68, CCL5-002-68, CCL5-001-40 and CCL5-002-40 all cause a significant reduction in atheroma size due to the reduction of immune cells recruited to the lesion. Another mutation in the top 10 is IL4-201-134, causing a severe change in the T_h1/T_h2 balance within the atheroma, giving an A_{score} of 0.0675. Limiting immune cell

recruitment and disrupting the T_h1/T_h2 balance are two potential areas of investigation in the treatment of atherosclerosis. Finally, CSF1-002-461 severely inhibits the rate of monocyte differentiation into macrophages, preventing the frequent phagocytosis of oxLDL which drives atherogenesis.

6.4.1.2 Genetic Profiles

For each of the 8 population subgroups studied in this chapter, we have seen a change in pathway dynamics when reparameterisation has taken place to consider the effect of mutations commonly found within these populations. Within genetic profile 1 (Han Chinese from China, Chinese Dai from China and Kinh from Vietnam), a decrease in extra-cellular matrix remodelling has led to a total increase in concentrations of elastin and collagen. Increased collagen within an atheroma can lead to an increase in lesion size, however collagen can also be beneficial to plaque stability, reducing the likelihood of a cardiovascular event. While studies exist providing statistics related to CVD for each of these populations, none of them provide sufficient histological detail to use as validation for the conclusion from our genetic profile study (Chen et al., 2017; Lu et al., 2012; Minh, 2006). Similarly, genetic profile 3 (Southern Han Chinese) shows elastin degradation and a minor alteration in plaque cellular composition due to a mutation in CSF1R, however experimental data to corroborate this result is unavailable.

Japanese individuals within Japan present fewer symptoms of subclinical atherosclerosis than other populations. Indeed, Japan has a significantly fewer CVD related deaths than reported in other countries (Nojiri and Daida, 2017). Despite the stability issues with our Japanese genetic profile, the atheroprotective nature of the variation common within this genetic profile mirrors the reduction in subclinical atherosclerosis manifestation and subsequent CVD risk. There are multiple potential reasons for the reduced CVD risk in Japan, including a low-fat diet and significantly reduced obesity within the country. However, Japanese individuals who have followed a western lifestyle from childhood are shown to have a lower mortality rate from CVD than Caucasian individuals within the US, suggesting that genetics may have a part to play (Sekikawa et al., 2007, 2003).

Genetic profile 4 (European Americans from Utah, USA) has shown a minor change in the T_h1-T_h2 balance due to a mutation in IL10RA. Similarly, genetic profile 5 (Toscani from Italy) shows a minor change in this balance due to an IL12RB1 mutation. The effect of these mutations is very minor and would likely have an almost insignificant effect on cardiovascular risk. Studies into European American populations are limited (Johnson et al., 2011), however it has been shown that cardiovascular disease related morbidity and mortality in Italy is reduced when compared to other European countries, although this is likely to be primarily due to dietary factors (Palmieri et al., 2010).

Additionally, the Finnish genetic profile (genetic profile 6) shows a similar level of atherosclerotic risk to genetic profile 5, with the addition of a mutation in CSF1, causing a very minor change to atheroma cellular content. Despite Finland having the highest rate of CVD related deaths in the late 1960s, this number was significantly reduced in 2012 due to reduction in numbers of smokers and healthier diet leading to average lower serum cholesterol and blood pressure (Jousilahti et al., 2016; Vartiainen et al., 1999).

Genetic profile 7 (Yorubi from Nigeria, Luhya from Kenya, Gambian from Gambia, Mende from Sierra Leone, Esan from Nigeria and African Caribbeans from Barbados) show an alteration to the T_h1- T_h2 balance due to a mutation in IL4R. This mutation is expected to be atheroprotective. Europeans have a higher risk of cardiovascular disease than an African Caribbean population (Chaturvedi, 2003). Studies have been performed to assess CVD risk within Yoruba Nigerians and individuals in western Kenya, however these studies solely look at common risk factors (Chege, 2016; Deeg et al., 2008). Hypertension across Gambian, Sierra Leonean and Southern Nigerian populations has also been studied (Awad et al., 2014; Isara and Okundia, 2015).

The profile representing Gujarati Indians from Texas, USA (Profile 8) exclusively contains a mutation in EGFR, leading to an increase in elastin concentration within the artery wall. While studies exist looking at the burden of CVD in Gujarati Indians in Asia, a Texas-based population has not been studied (Sharma et al., 2015).

Multiple studies have shown an increased cardiovascular disease risk for African Americans and South Asian individuals (Gupta and Brister, 2006; Lloyd-Jones et al., 2010). Unfortunately, no mutations showed a greater than 50% prevalence within our African American or South Asian datasets.

6.4.2 Variations in Interleukin-18 affect plaque stability

IL18R1-001-232 (rs148457935) is a mutation which leads to an increase in k_{on} to $1.26x10^6$. Within the mathematical model, this has given a k_{cat} value of $2.4525\ x10^5$ to represent this mutation. Studies have shown that an increase in IL-18 concentration and activity leads to a reduction in collagen within the intima in apolipoprotein E deficient mice, creating plaque instability (De Nooijer et al., 2004). Increasing the activity of IL18 through reparameterisation will lead to similar behavior as incorporating overexpression into the model, suggesting that this behavior will be seen in an individual with the mutation IL18R1-001-232. Collagen concentrations for a collection of IL18 related mutations are shown below:

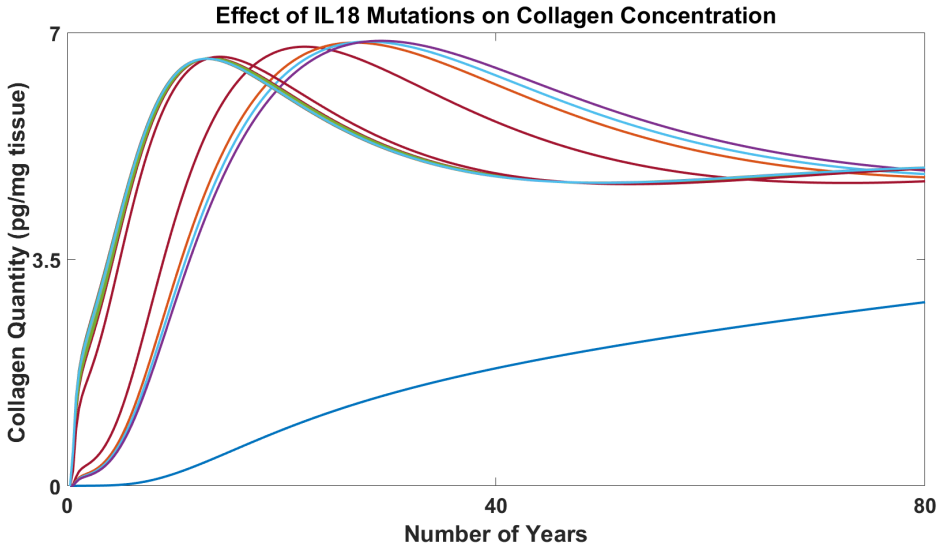


Figure 6.10: The effect of all mutations affecting the interaction between IL18 → IL18R1 on collagen concentrations

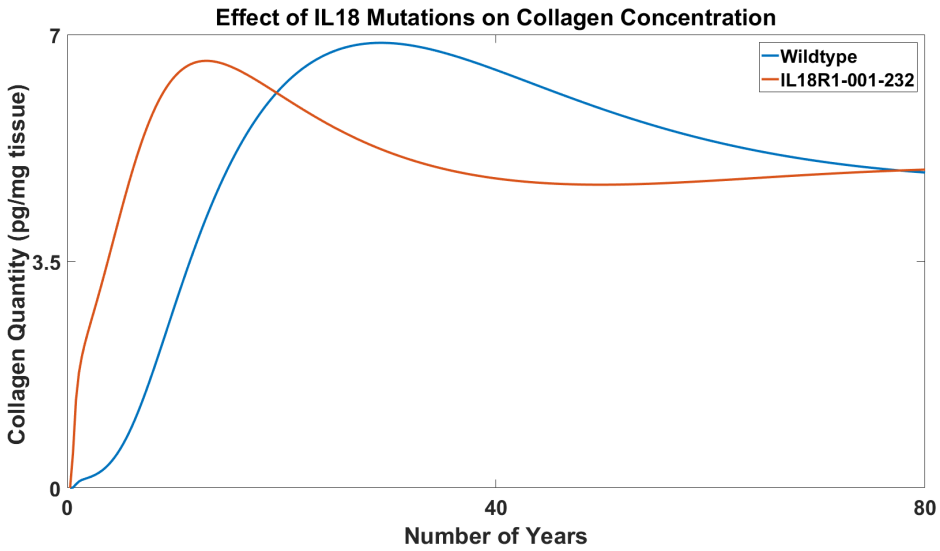


Figure 6.11: The effect of IL18R1-001-232 (rs148457935) on collagen concentrations

As seen in Figure 6.10, an increase in IL18 activity is expected to lead to an increase in the rate of collagen production in the beginning, however as time passes this concentration is expected to stop increasing and stabilize to approximately $5x10^6$ pg/ml, even with small variations in protein activity. With a large increase in IL18 activity represented by the mutation IL18R1-001-232, our initial spike in collagen concentration quickly turns into a reduction in total extra-cellular matrix within the plaque as collagen remodelling occurs.

6.4.3 Statins and PCSK9 Inhibitors

Two commonly prescribed drugs for the treatment of atherosclerosis-driven cardiovascular disease that are not included in our multi-drug experiment are statins and PCSK9 inhibitors. The mechanism of action of both statins and PCSK9 inhibitors involve the reduction of plasma LDL levels through the modification of the cholesterol metabolism pathway and the lipoprotein metabolism pathway respectively. Their therapeutic properties could be indirectly added to the model through a forced reduction of plasma LDL levels, however it was decided to solely focus on interactions included in the model to prevent altering of multi-drug therapeutic methodology for a specific case. Therapeutic options focused on the disease pathophysiology rather than the cause.

6.4.4 Multi-drug therapeutics

The introduction of fifteen different drugs was simulated for each genetic profile. Each of the multi-drug therapeutics simulations described in section 6.3.5 suggested a combination of three different drugs, while five genetic profiles suggest the introduction of a new drug targeting the interaction IL18 → IL18R to disrupt the T_h1/T_h2 balance. The nine genetic profiles outlined in section 6.3.1 suggest a collection of different drug combinations and concentrations to treat stratified patient groups.

6.4.5 Can we identify biomarkers in plasma from high concentrations within the atheroma?

Each of our genetic profiles can be studied for protein concentrations with significant variation from the wildtype profile for potential use as biomarkers. Genetic Profile 7 shows a significant increase in IL-4 concentration within the plaque, with a large increase occurring quickly after atherogenesis initiation.

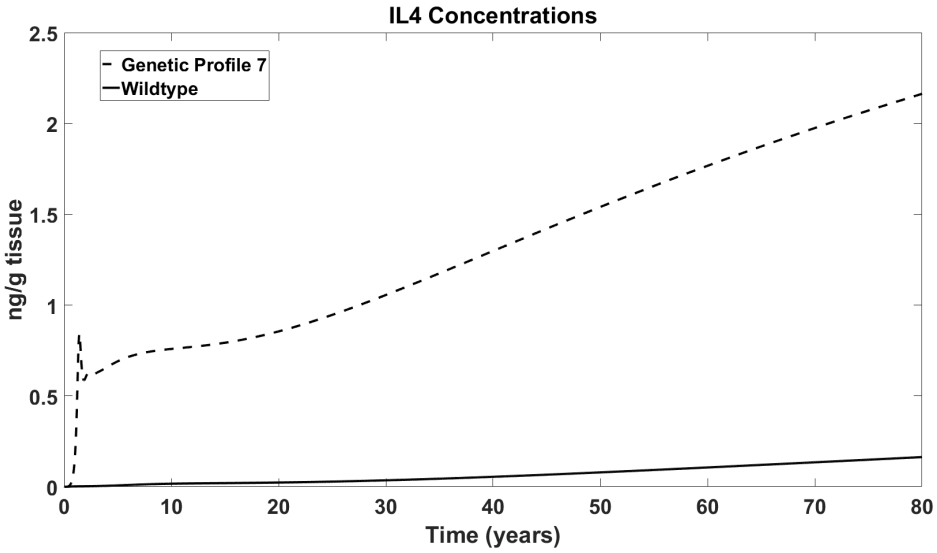


Figure 6.12: Interleukin 4 concentrations for genetic profile 7

Identification of biomarkers within population subgroups is a cornerstone of stratified medicine approaches. Within our mathematical model, concentrations are modelled for proteins found within the blood, the plaque, or both. The benefit of a biomarker within atherosclerosis is in the ability to utilize measurements of an entity as an indicator to

disease progression. Any potential biomarkers identified would be significantly less beneficial in practice if providing measurements proved challenging – i.e. if the biomarker was found within the plaque rather than within blood. In the case of IL-4, is it possible that we will see similar results within the plasma? The answer depends on the dynamics of the biomarker, tissue properties and the cytokine half-life; however further investigation is warranted.

6.4.6 Outliers

Particular results are obvious outliers. For example, TransComp estimated the association rate of CSF1R-001-245 with CSF1-001 to be 3.81×10^{-188} M^{-1} s^{-1}. Additionally, IL12RB1-001-214 was given a k_d value of 3.45×10^{-22} M by PRODIGY, leading to a very slow off-rate of 1.4305×10^{-18} s^{-1}. While these values give us a quantification that we can use, such extreme variations in binding kinetics have led to significant changes in k_{cat} values which heavily alter model dynamics, breaking model stability. As such, any obvious outliers with a k_{on} value less than 1 were removed from the dataset and model dynamics were not simulated. Providing a limit at the higher end of our collection of association rates was not necessary; the highest association rate calculated that altered model dynamics was the association between MMP3-001-45 and TIMP3-001 at 5.61×10^{10} M^{-1} s^{-1}, a reasonable value. High variations in k_{off} were less likely to provide a severe change to k_{cat}, leading to unrealistic disruptions to model dynamics, so k_{off} values were used whenever PRODIGY returned a value.

6.4.7 Other Diseases

Mutations shown to be atherogenic or atheroprotective could have opposite effects with other diseases. Asthma is an inflammation driven disease affecting the lungs, which shares inflammatory mechanisms with atherosclerosis (Lambrecht and Hammad, 2015). Asthma is a T_h2 driven disorder where T cell activation leads to mucus production, oedema and inflammation of the airways. Tipping the T_h1/T_h2 balance towards T_h2 cells is likely to increase the severity of the disorder (Bosnjak et al., 2011). Genetic profile 7 contains IL4R-001-576, a mutation which promotes T_h2 differentiation, altering the T_h1/T_h2 balance within atherosclerosis. African Caribbeans, a population subgroup included as part of genetic

profile 7, have an increased prevalence of asthma when compared to white UK individuals (Whitrow and Harding, 2010).

6.4.8 Conclusion

In this chapter we have combined the results detailed in chapters 3, 4 and 5 to develop an *in silico* platform to study how atherosclerosis dynamics differ between population subgroups. The creation of genetic profiles to represent individual populations allows for the reparameterisation of the model for these genetic profiles and the study of theoretical multi-drug interventions to optimize plaque reduction. Systems biology methods have the potential to be a strong predictor of how a perturbation will affect disease dynamics when the mathematical model of the disease and the disruption are both well-defined. The quantity and quality of publically available biological data relating to PPIs and of mathematical models relating to disease processes has significantly increased in recent years. Combining powerful models with vast datasets has the potential to lead to advances in stratified medicine; including population subgroup specific therapeutic hypotheses, biomarker identification and virtual clinical trial functionality.

Chapter 7:

Conclusion and Future Work

7.1 Overview

Computational and mathematical modelling of disease processes has a key part to play in the future of stratified and personalized medicine. Sequencing the genome of an individual to provide genetic insight which can be used to provide optimal therapeutics is an idea that has been considered, discussed and debated amongst healthcare professionals for years (Gonzalez-garay, 2015; Kalow, 2002; Nijhout et al., 2015). However, many significant advances in computational modelling are required before this becomes a reality. Improving methods to analyse how genetic variation affects the dynamics of particular diseases is an important step on the journey to idealized personalized medicine. Within this thesis, we have developed a computational model of atherosclerosis containing significantly more biological entities than any other available model of plaque formation. Validating model results through comparison with experimentally derived results has been performed extensively to ensure biological relevance. Multi-drug systems pharmacology approaches have been utilized to suggest drug combinations to minimize plaque growth, and an overall *in silico* platform for the study of atherosclerosis dynamics has been created.

Utilizing sequence data taken from phase 3 of the 1000 Genome Project, a collection of mutations relating to atherosclerosis were isolated. A protein-forming amino acid sequence was constructed for each of these sequences, and state-of-the-art protein folding methods were used to predict tertiary structures for each of these mutations. Comparative steps were taken to demonstrate the validity and accuracy of each of these predicted protein structures through alignment to experimental proteins found in the Protein Data Bank, and these mutations were heatmapped to show regions of proteins with increased structural variability. Docking methods were utilized to derive a quaternary structure, and to study the complexes formed during atherogenesis. Binding kinetics were calculated in the form of k_{on} and k_{off} values for each of the complexes formed with these docking procedures, and the effects of mutations on these kinetics were studied.

As part of a stratified medicine program designed to study the dynamics of atherosclerosis across population subgroups, variant allele frequency data isolated from the 1000 Genome Project were used to create a genetic profile for 31 different population subgroups. Binding kinetics for mutations involved in these genetic profiles were used to reparametrize the

mathematical model, to predict population specific changes in plaque development. Utilizing a genetic algorithm to find a global minimum to a scoring function related to atheroma size and plaque stability, multi-drug combinations to maximize plaque reduction were calculated to suggest population-specific therapeutics and progression biomarkers.

7.2 Have we satisfied our aims?

The aims of this thesis as outlined in Section 1.3 have been addressed. These statements of intent have all been satisfied within our previous chapters.

7.2.1 Aim 1 – Develop a computational model of atherosclerosis

After careful study of currently existing computational and mathematical models of atherosclerosis and plaque development, a literature review of this field was scripted and published, included in Chapter 2. These models were considered during the development of the mathematical model detailed in Chapter 3. The mathematical model developed as part of this thesis includes aspects of atherosclerosis that have not been modelled by other groups and models the concentrations of more biological entities than other models within this field (Parton et al., 2015). Extensive validation has been performed to ensure the biological relevance of the model.

7.2.2 Aim 2 – Study the variation in structure for proteins related to atherosclerosis

A collection of protein structures detailed in Chapter 4 were obtained using protein structure prediction methods that CASP experiments have proven to be highly reliable. Alignments to experimentally derived wildtypes were performed to ensure accuracy of our models. Quaternary structures including heterodimers, homodimers and complexes were derived using docking methods, and binding kinetics were calculated using these structures, as explained in Chapter 5.

7.2.3 Aim 3 – Predict how structural variance will change atherosclerosis dynamics

Rate parameters k_{on} and k_{off} were calculated and described in Chapter 5. The mathematical model detailed in Chapter 3 was reparametrized according to this binding kinetics data. Genetic profiles for population subgroups have been developed utilizing 1000 Genomes

Project population genetics data to allow for the prediction of how population subgroup-specific structural variance will alter atherosclerosis progression, as shown in Chapter 6.

Each of these aims have been achieved and comprehensively detailed within this thesis, assisting in the overall objective of studying the dynamics of atherosclerosis across population subgroups.

7.3 The need for collaboration

While the mathematical model developed shows extensive depth and complexity in the study of atherosclerosis, numerous biological processes and pathways significant to plaque formation are not included as part of this study. LDL concentration within the blood, one of the primary risk factors for cardiovascular disease, is considered as a constant within the mathematical model. This prevents the consideration of short-term fluctuations in concentrations, dietary changes, the process of cholesterol metabolism, the process of lipoprotein metabolism and lipoprotein subtypes. Mathematical models have been developed to study the role of statins in cardiovascular disease (Eussen et al., 2011), with statins being the most commonly used drug in the treatment of CVD. Considering other risk factors to atherosclerosis (such as diabetes mellitus, smoking and old age) expands the list of biological processes to consider while studying the process of atherogenesis as a whole. The list of risk factors that have been shown to have a connection to cardiovascular disease is so expansive that incorporating each of them into a model would require something akin to a computational model of an entire human (Grundy et al., 1999; M. Mooney and T. Mc Auley, 2015; Muntner et al., 2005; Powell, 1998; Shi et al., 2005). Recently, the first whole cell model was developed, representing *Mycoplasma genitalium* (Karr et al., 2012a), in what is still one of the largest computational models currently available to represent biological processes, combining data from more than 900 different publications. Finding accurate biological parameters for all interactions within this model of *M. genitalium* is an incredibly challenging task, requiring manual searching of a vast collection of published data to find potential parameter values. The volume of work required to develop a mathematical model covering a whole cell is enormous, even with *M. genitalium*'s small genome size (Karr et al., 2012b).

In the pursuit of true personalized medicine, interconnected models representing large biological systems are more likely to be able to predict side effects of model perturbations. Of course, the study of subsections and pathways can bring key insight into disease dynamics; however, larger and more comprehensive models of biological systems have the potential to bring new, stronger results. Gaps in our knowledge can be bridged through the introduction of systems biology approaches to biological systems, yet the systems concerned are so complex that it requires vast resources spanning multiple disciplines to bring computational biology results into clinical healthcare benefits. Developing a strong computational model of a disease process requires skills in areas such as mathematics, informatics and biochemistry. To provide expertise in individual disease areas and guidance on clinically beneficial results, input from domain experts and clinicians can be a significant benefit to the project. Development of a multi-scale model requires the summation of proteomics, genetics, metabolomics and transcriptomics, while looking at an assortment of cell types and tissues where interactions can take place. Collaborative efforts are the future of systems medicine to ensure that each of these areas are suitably covered.

7.4 The Mathematical Model of Atherosclerosis

The mathematical model developed in Chapter 3 is a keystone of the work outlined within this thesis. The model has been developed to SBGN (Le Novere et al., 2009) and SBML (Hucka et al., 2003) open standards and the SBML version has been deposited in BioModels, a public domain computational model database, to encourage collaboration and reuse (Novère, 2006). Results have shown the biological relevance of the mathematical model through reproduction of results from other biological studies, as detailed in section 3.3. Considering the concentration of each protein and cell type within the model using an ODE allows for the pharmacological modelling of drug interventions using hypothetical drugs acting on specific targets as well as known drugs.

However, despite the considerable detail included in the multi-scale model, it can in no way be considered a complete representation of atherosclerosis. Some cell subtypes, such as anti-inflammatory macrophages, have not been included within the mathematical model (Bobryshev et al., 2016). Subtypes of dendritic cells and mast cells have not been considered. Additions to the cytokine milieu could also be introduced, alongside heat shock

proteins and microRNAs to represent regulation of gene expression (Nazari-Jahantigh et al., 2014; Xu et al., 2012). Due to the restrictions imposed by adhering to SBML open standards, spatial factors are not considered within this model. This removes factors such as arterial remodelling and wall shear stress, which has been shown to be linked to atheroma development (Cunningham and Gotlieb, 2005). The removal of spatial considerations also limits study of atheroma location, and any considerations of atheroma size have to be linked to the morphological content of the artery wall.

Model outputs have been compared to experimental results to express the validity of the model. Experimental data have been found describing concentrations for a collection of proteins and cell types included within the model, and model simulations are compared to these data to ensure biological relevance. Perturbations in concentrations and activity of a selection of cytokines and chemokines within the model were introduced to predict how these changes influence atherogenesis. Variations in concentrations of IL-12 and IL-18 in mice have been shown to disrupt the T_h1/T_h2 balance and alter the cellular composition of the plaque (Elhage et al., 2003; Hauer et al., 2005). Reductions in proteoglycan concentrations and PLA2 activity lower plaque levels of oxidized LDL, leading to a reduction in atheroma size due to fewer foam cells being created (Delgado-Roche et al., 2015; Vickers et al., 2009). An increase in PDGF concentration boosts smooth muscle cell abundance, affecting plaque stability (Loppnow and Libby, 1990). Collagen concentrations are more challenging to predict when considering perturbations to the system. As plaque size grows and time progresses, collagen synthesis grows due to increased abundance of smooth muscle cells, while collagen remodelling increases at the same time due to increased production of MMPs (Newby, 2005; Rekhter, 1999). More severe lesions left long enough will have decreasing levels of collagen due to higher levels of MMPs within the system, increasing the likelihood of plaque rupture and a subsequent cardiac event (Adiguzel et al., 2009). Each of these studies have been reproduced *in silico* within the model. Multi-drug systems pharmacology experiments have suggested a benefit to limiting phagocytosis of oxLDL, or disrupting the T_h1/T_h2 balance by targeting IL12R and IL18R.

7.4.1 Future Improvements

While the development of this model of atherosclerosis and the subsequent pharmacological study of potential multi-drug interventions provides insight into atherosclerosis pathogenesis, some common therapeutics used to treat CVD and atherosclerosis are not represented within this model. Statins inhibit the synthesis of cholesterol by inhibition of HMG-CoA reductase, leading to a reduction of LDL within the blood (Stancu and Sima, 2001), and as such the effects of statin therapy can be considered within the model, albeit indirectly. Similarly, PCSK9 inhibitors prevent the degradation of LDL receptors in the liver, increasing the rate of LDL-C removal (Page and Watts, 2016). PCSK9 inhibitors lead to significant reduction of LDL concentrations within the liver, which could be indirectly included within the model similarly to statins. Creation of a 'supermodel' that includes atherosclerosis, cholesterol metabolism and the removal of plasma LDL by hepatocytes would allow for the study of multi-drug interventions that includes the subtleties involved with statin therapy and PCSK9 inhibition. Cholesterol metabolism has been successfully modelled previously using the SBML format (Mazein et al., 2013; Mc Auley et al., 2012) and more recently the role of aging in this process has been considered (Morgan et al., 2016). The connection between PCSK9 and LDL cholesterol levels have been modelled within specific mammalian systems (Hansen et al., 2017), and the pharmacological results of PCSK9 inhibition on plasma LDL has been modelled (Gadkar et al., 2014). Each of these models represent a biological process which is key to the progression of cardiovascular disease, and the creation of a combined model could lead to new therapeutic hypotheses in the treatment of this disease.

7.5 The Protein Structure Dataset

Predicting a tertiary structure for a protein from its amino acid sequence is a field that has seen significant attention from biologists and bioinformaticians in recent decades. Current state-of-the-art tools allow for the generation of a 3D structure in a timeframe ranging from hours to weeks depending on the length of the primary structure and the number of known homologues and folds for use in structure formation. Due to computational intensity restrictions, only a subset of all mutations contained within phase 3 of the 1000 Genomes Project were chosen for subsequent analysis in this study. The subset chosen was weighted to produce more mutations for shorter sequences, to prevent the majority of mutations

chosen affecting larger structures like receptors. This gave a broad spread of mutations across all proteins selected to represent atherosclerosis, allowing for further study of protein structure, binding kinetics and function. Comparison of predicted structures to known experimentally-derived structures was performed to evaluate accuracy of overall structure and I-TASSER accuracy estimation methods. Generation of heatmaps showing how a collection of mutations leads to variation in protein structure allows for the visualization of highly varied areas across populations, and when combined with binding site data could be useful in the recognition of the severity of a group of mutations. When clinical data exists linking a mutation to a particular disease, our dataset can be searched for variants with a strong alignment to either the wildtype or the disease-causing mutation to predict the likelihood of a mutation being disease-causing. The dataset generated was large and varied, benefiting later stratification exercises.

Improvements could still be made to our predicting protein structure procedures, however. Expanding our dataset beyond the 1000 Genome Project to similar sequencing projects such as UK10K (Consortium, 2016) would have increased the number of population subgroups with available data and increased the likelihood of the population genetics data used being more representative of these subgroups as a whole. The process of predicting tertiary structure could have been optimized by solely using homology modelling methods on structures where such relatives were available, significantly reducing computational time and freeing up such time for additional mutations, albeit with a minor drop in accuracy. Heatmaps could have been more informative if weighted with population genetics data. In general, our intentions to produce the most accurate predictions possible within our computational limitations to minimize compound error were successfully achieved in Chapter 4.

Experimental results obtained within this thesis show that C-scores are a good indicator of global accuracy of protein structure. Local accuracy predictions, values describing the per-residue RMSD as calculated by ResQ, are generally valuable yet are not 100% accurate (Yang et al., 2016). The collection of graphs shown in Figure 4.6 to study local accuracy of structural predictions demonstrate that while limits provided to variation from wildtype are beneficial, 100% accuracy is not guaranteed. A heatmap has been developed for each

protein to represent this local accuracy to allow for quick visualization of protein areas where we are more confident of accurate structure. An additional heatmap was created to show the variance in protein structure across the collection of mutations included in our dataset. Mutations linked to disease in IL4R and ABCA1 were identified and structural similarity with wildtype and other mutations was considered, suggesting protein structures that are structurally similar to disease causing mutations that may benefit from further investigation.

7.5.1 Future Improvements

As protein structure prediction methods become more accurate and less computationally intensive, large structural datasets can be developed which could be used as part of a multitude of studies. With the existence of repositories such as ClinVar (Landrum et al., 2014) detailing links between mutations and phenotypes, including mutations linked to disease, structural datasets could be compared to known disease-linked mutations to predict mutation severity (in the context of disease). CASP and CAMEO experiments have excellently constructed a competitive and rewarding atmosphere to drive progress with protein structure prediction methods, with the number of entrants growing year-on-year and increasing accuracy and speed being demonstrated by the winning entries consistently between competitions (Haas et al., 2013; Moult et al., 2016). The Protein Data Bank is the world-leading repository for experimentally derived protein structures, with 9,990 X-Ray structures and 455 NMR structures uploaded in 2016 (PDB, 2016). With the improvement of structural prediction methods, a separate repository containing predicted protein structures has the potential to be beneficial to the bioinformatics community at large. Caution would need to be taken to ensure a minimum level of predicted-structure quality; however, it has been demonstrated that even I-TASSER results returning a relatively weak prediction will still represent the correct model topology. As such, a predicted-structure repository could benefit the research community, particularly for larger proteins where performing structural prediction methods is a computationally intensive task.

7.6 Calculation of Binding Kinetics

Alterations in the velocity of binding and unbinding within a protein-protein interaction alter the thermodynamic properties of the interaction (Prakash, 2011). Through the

calculation of association and dissociation rates for a wildtype and a collection of mutations, predictions into interaction activity can be made. Electrostatic potentials between proteins have a significant impact on association rates (Ritchie, 2008). The power and the polarity of the electrostatic interaction between two reactants can lead to fast or slow binding rates, while electrostatics can also influence steering to ensure that reactants interact at their respective binding sites (Wade et al., 1998). As such, focusing on electrostatics alongside protein shape is a reasonable method of computationally studying and prediction association rates for a PPI. TransComp has many benefits for use as part of a study of PPI association rates, including access to a publically available web server for testing and to return computationally intensive results in a short period of time, and internal determination of reaction criteria leading to a single scalar result being returned; removing an additional assumption or computational step which would have been required if SDA7 results were used (Martinez et al., 2015; Qin et al., 2011). Using TransComp to calculate association rates for 1119 interactions required significant computational resources, in contrast to PRODIGY which was able to estimate k_d vales. Approximately 500,000 core hours were required to predict association rates for 835 interactions involved within our model of atherosclerosis, and multiple big data challenges needed to be overcome in terms of computational time and storing data output.

Validating TransComp results demonstrated that results were generally within an order of magnitude of experimentally derived association rate values. While this is still a large margin of error to work with, it does provide sufficient accuracy for us to progress into reparameterisation experiments. TransComp has shown increased accuracy for proteins with association rates in the range of 10^5 - 10^{10} $M^{-1}s^{-1}$, due to the transient complex approximation utilized as part of its methodology (Qin et al., 2011). As such, association rates in the lower half of the results spectrum with a value of 10^0 - 10^4 $M^{-1}s^{-1}$ are less likely to be accurate.

A collection of TransComp results with low k_{on} values were obvious outliers. For example CSF1R-001-245 binding to CSF-001 returns an association rate of 3.81 $x10^{-188}$. This breakdown in kinetics prediction is caused by a very large electrostatic potential between two reactants being calculated, giving a powerful repelling force. Additionally, multiple

interactions within our dataset did not return a result from either TransComp or PRODIGY. TransComp was unable to predict association rate values for complexes with more than one transitory state, which was common for some proteins within our dataset. PRODIGY was generally more reliable; however, issues were had with some complexes (primarily IL5 → IL5RA and its mutations) where the two complex subunits were considered to be too far apart for k_D estimation. These issues led to a significant reduction in the size of our binding kinetics dataset, from an initial 1119 to 835. Fortunately, binding kinetics were successfully obtained for each mutation and wildtype included in the genetic profiles defined in section 6.3.1, allowing for successful reparameterisation experiments despite our reduction in dataset size.

Derivation of binding kinetics are the product of a collection of bioinformatics tools. GalaxyHomomer (Baek et al., 2017), RosettaDock (Davis and Baker, 2009), Hex (Macindoe et al., 2010), TransComp (Qin et al., 2011), PRODIGY (Xue et al., 2016) and PDB2PQR (Dolinsky et al., 2007) were all used to provide a collection of association rates. Homodimeric structures, predicted using GalaxyHomomer, gave dimeric structures similarly close to the native as local docking results provided by local docking despite GalaxyHomomer not requiring contact points as an input. Benchmarking results showed that TransComp showed increased accuracy over SDA7 when estimating k_{on}, even for association rates as low as 10^4 M^{-1} s^{-1}. Variation in association rate and dissociation constant showed a minor correlation with association rate.

7.7 Future Improvements

Improving dataset coverage and improving association and dissociation rate accuracy are two obvious improvements that could be made to our methodology. Predicting whether a complex is likely to fail TransComp analysis before spending considerable computational resources on the run itself could allow for the reduction of computational time, allowing resources to be put to better use. Utilising similar methods to calculate k_{on} and k_{off}, rather than combining the Brownian Dynamics simulation style of TransComp and the contact-based methods of PRODIGY could reduce bias introduced by both methods. New methods of predicting binding kinetics utilizing machine learning techniques have been developed recently, potentially reducing computational time and providing more accurate predictions

in the lower half of the rate spectrum (Xie et al., 2017). As new methodologies and algorithms become available then increases in speed and accuracy could provide a platform for scaling up this methodology as part of a disease pathway reparameterisation platform.

7.8 Atherosclerosis Dynamics Across Population Subgroups

Combining Chapter 3's model of atherosclerosis with Chapter 4's protein structure data and the binding kinetics data predicted in Chapter 5 allows for the development of an *in silico* learning platform for patient stratification and the study of atherogenesis dynamics. Utilizing binding kinetics data for a collection of mutations to investigate disease dynamics and predict pathogenesis differences for each mutation has provided a platform for evaluation of its downstream effects. Experimental data has shown encouraging corroborative results when compared to disease dynamics of our genetic profiles. Multi-drug therapeutics designed for the minimization of plaque size when inhibitors are added to the system suggests different therapeutics were optimized for different population subgroups. The generation of 746 models of atherosclerosis representing individual mutations with effects on disease dynamics provides a demonstration of model robustness and a proof of concept that reparameterisation can provide prediction on therapeutic effects and '-omics' data driven disease dynamics as part of an *in silico* testing system.

Increased accuracy in binding kinetics data would have improved the quality of our reparameterisation experiments. Experimental derivation of binding kinetics or protein structure would have led to a reparameterisation that was closer to the native, underlying biology. The genetic profiles developed in chapter 6 are designed to represent individual population subgroups; however, due to the limited number of mutations included within the dataset isolated in chapter 4, these genetic profiles contained a relatively small number of variations. The A_{score}, designed to provide a scalar value to represent atheroma size and severity, could have been calculated in a multitude of ways. The MD_{score}, studying the efficacy of the multi-drug interventions, could similarly have been altered by varying the weights attached to introducing a new drug into the system or increasing drug concentrations.

The genetic profiles created in section 6.3.1 were designed to relate our mutation data to particular population subgroups to investigate the potential of altering atherosclerosis dynamics and therapeutic optimization for each of these subgroups. Unfortunately, due to limitations placed on our dataset, only 8 unique genetic profiles were created. However, interesting results were obtained for each of these genetic profiles. The mutation CCR2-001-64 (rs1799864), found in 50.96% of Japanese individuals within phase 3 of the 1000 Genome Project, reduces the rate of monocyte recruitment to the site of endothelial damage, a potential influencing factor for the reduced rate of cardiovascular disease in a Japanese population. Similarly, IL12RB1-001-214 (rs11575934) is commonly found in Toscani and Finnish populations (58.87% and 55.55% respectively), a mutation which alters the T cell constituency of the lesion while only making a minor change to atheroma severity, giving an opportunity for alternative therapeutics for these population subgroups. Alternative genetic profiles were found and potential multi-drug therapeutics were found for each of our genetic profiles through the use of a genetic algorithm, where different drug combinations providing a greater reduction in atheroma size and severity were found for 7 of our 8 genetic profiles.

7.8.1 Future Improvements

Expanding the size of the genetic profiles to include all mutations within the 1000 Genome Project (Auton et al., 2015), or similar projects like UK10K (Consortium, 2016) could have provided a reparameterisation more representative of the population. Starting from these improved genetic profiles and utilising the entire computational pathway outlined in this thesis for mutations, by predicting a tertiary structure, binding kinetics and reparameterisation, would have given credence to the biological relevance of each of the genetic profiles. Creation of a ‘supermodel’ as postulated in section 7.4.1 would have allowed for the introduction of other common CVD related drugs into the system, such as PCSK9 inhibitors and statins. Additionally, a ‘supermodel’ could provide insights into multi-drug therapeutic dynamics and provide a collection of other mechanisms which could be responsible for population subgroup specific differences in atherosclerosis dynamics. Using commonly known drugs with known inhibitory constant (k_i) values could have allowed for the study of concentrations to be more accurate and allowed for the study of absolute values rather than relative concentrations.

7.9 Compound error

All of the work undertaken within this thesis can be considered as part of the same methodology, an algorithmic road with the final goal of atherosclerosis patient stratification. From the development of the mathematical model, to tertiary structure prediction, to binding kinetics derivation, and finally reparameterisation, each chapter is a key part of the methodology with which following chapters require to correctly function. Chaining together algorithmic sections with predictive elements requires awareness and handling of compound error. Minor errors in results in the earlier part of the overall methodology combine with minor errors in other subsections leading to an overall larger error in the end. Reduction of error in individual subsections where possible can lead to larger improvements in the quality of our final results, leading to general improvements as technologies develop and knowledge grows. Improvements and additions to the publically available model of atherosclerosis through parameterization updates or network optimization can reduce the error and improve the quality of our final multi-drug therapeutics. Improvements in the prediction algorithms for obtaining tertiary structures from amino acid sequence would lead to docking improvements and subsequently more accurate binding kinetic information. Due to the close-knit and intertwined nature of each chapter within this thesis, improvements to the output data from any chapter could lead to a more accurate end result.

7.10 Systems Medicine

The question remains of how our systems biology and systems pharmacology methods can be beneficial and provide a clinical benefit in the future. With the advent of big data within biological sciences, access to transcriptomic, genomic, proteomic and metabolomic data is easier than ever. These datasets have the potential to be exceptionally beneficial; however focusing on only one of these '-omics' areas has the potential to be isolating to a scientist. Our mathematical model of atherosclerosis can be used as a platform to combine datasets from each of these technologies into an overlying vision of disease dynamics. Potential molecular targets for drugs have been described for an overall population and additionally for subgroups of these populations as part of a program of stratified medicine, and overall therapeutic hypotheses have been postulated. Significant developments are required to

convert therapeutic hypotheses generated by mathematical models into clinical benefits, however bridging knowledge gaps through network development and mathematical modelling can provide illuminating clinically significant analyses.

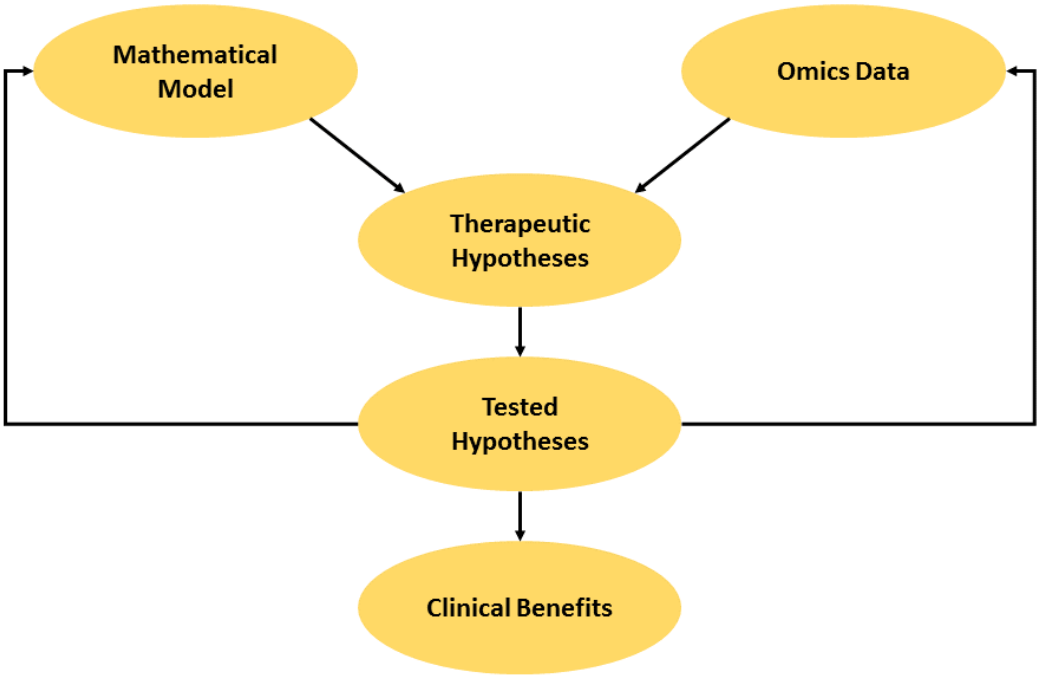


Figure 7.1: Systems Medicine Methodology

7.11 Future work

The work included in this thesis utilizes state-of-the-art technologies in protein structure prediction, docking and prediction of binding kinetics to develop therapeutic hypotheses of atherosclerosis. The natural progression of this work is to expand the size and scope of the model of atherosclerosis to include lipoprotein metabolism and cholesterol metabolism. Once a successful model has been developed to include these biological processes, other models could be combined on the way to the ultimate goal of developing a 'virtual human.' In the more immediate future, it would be beneficial to use the same methodology on a different network of disease or biological function. Many of the protein structures and binding kinetics data generated as part of this project are important parts of inflammatory disease, allowing for cross-use in a similar study looking at a disease with a core inflammatory element, such as rheumatoid arthritis or asthma. An interesting 'wet-lab' study would be to create protein structures with the mutations included in our genetic

profiles using technologies such as CRISPR and experimentally validate our binding kinetic results. These more accurate results could then be used as an optimization of this work. If endarterectomy samples could be obtained, quantification of concentrations of proteins within plaques would be beneficial for the development of an increasingly accurate mathematical model. Any and all of these developments would lead to improvements in the treatment of atherosclerosis utilizing stratified medicine.

Appendices

Appendix 1

In addition to the two manuscripts detailed during Chapter 2 and Chapter 3 of this thesis, the following manuscript was published during this PhD programme:

Benson, H., Watterson, S., Sharman, J., Mpamhanga, C., Parton, A., Southan, C., Harmar, A., and Ghazal, P.(2017) Is systems pharmacology ready to impact upon therapy development? A study on the cholesterol biosynthesis pathway. British Journal of Pharmacology, doi: 10.1111/bph.14037.

The following publication was submitted during this PhD programme and is currently awaiting peer-review response:

Alexander Mazein, Marek Ostaszewski, Inna Kuperstein, Steven Watterson, Nicolas Le Novère, Diane Lefaudeux, Bertrand De Meulder, Johann Pellet, Irina Balaur, Mansoor Saqi, Maria Manuela Nogueira, Feng He, Andrew Parton, Nathanaël Lemonnier, Piotr Gawron, Stephan Gebel, Pierre Hainaut, Markus Ollert, Ugur Dogrusoz, Emmanuel Barillot, Andrei Zinovyev, Reinhard Schneider, Rudi Balling and Charles Auffray. Systems medicine disease maps: community-driven comprehensive representation of disease mechanisms; Nature Biotechnology

Appendix 3

Model Construction

A list of cell types involved in atherosclerosis was assembled from KEGG maps (Kanehisa et al., 2014; Wixon and Kell, 2000) and a collection of review articles of atherosclerosis (see Supplementary table 4). For each of these cell types, PubMed and Google Scholar searches were undertaken with the query *'CELLNAME atherosclerosis review'*. The results were sorted by *'Best Match'/'relevance'*, and the top 25 results were selected. Each article was searched for references to extracellular proteins and other biological entities connected to these cells. Each entity found in this way was listed as a provisional candidate for our model. Another combined PubMed and Google Scholar search was undertaken with the query '*ENTITYNAME atherosclerosis'*, to identify experimental data to show its source or presence within an atheroma (or another compartment contained within the model) and its influence (however minor) on atherogenesis. If it the entity met this criteria, it was included. Interactions were added where it was demonstrated that a biological entity directly influences another within the model. A list of the proteins included is shown in Appendix 3 - Supplementary Table 5 and a list of the lipoproteins is shown in Appendix 3 - Supplementary Table 6.

The resulting model contains five compartments, representing areas of the body where disease subprocesses take place: the tunica intima, endothelium, lumen, liver and intestine.

A search for 'atherosclerosis' on PubMed returns approximately 125,000 results (2016 figure https://www.ncbi.nlm.nih.gov/pubmed/?term=atherosclerosis) making a complete manual curation unrealistic. Once our initial network had been developed it was split into the collection of sub-processes shown in Appendix 3 - Supplementary Table 7. A PubMed search was performed for a collection of keywords related to these sub-processes, improving the completeness of the model.

The model was assembled using CellDesigner (Funahashi et al., 2003, 2008) to facilitate exporting files into Systems Biology Markup Language (SBML) (Hucka et al., 2003) and Systems Biology Graphical Notation (SBGN) Process Description formats (Le Novere et al.,

2009). All proteins were given their unique name from the HGNC (HUGO Gene Nomenclature Committee) (Gray et al., 2015) to prevent ambiguity.

For mathematical expediency, it is assumed that all entities within each compartment occupy the same space, negating the need for diffusion coefficients. Interactions were described mathematically with either the Law of Mass Action (LOMA) (Voit et al., 2015) or Michaelis-Menten kinetics (Johnson and Goody, 2012; Michaelis and Menten, 1913). Degradation and cell death are modelled using LOMA. Lipoprotein delipidation, transfer and formation is modelled using LOMA. Cytokine production, excluding TNF-Alpha, within the *tunica intima* is modeled using LOMA. Cytokine production within the artery lumen is modeled using Michaelis-Menten kinetics. Oxidation of lipoproteins, foam cell formation and SMC migration are modelled using a modified form of Michaelis-Menten kinetics. All other interactions are modelled using LOMA. A Hill coefficient has been added to equations regarding lipoprotein transfer to ensure accuracy in downstream oxidised LDL concentrations for different LDL concentrations. A list of the specific equations for each species can be found in Appendix 3 - Supplementary Table 1.

In order to find biologically relevant parameter values, a literature mining process was undertaken. Enzyme rate parameters were included in the model were either taken from studies which calculated them directly, determined through fitting species concentrations to experimental data or made an estimation where appropriate experimental data was not obtainable. Due to the lack of parameterisation data available, kinetic parameters were taken from studies of any mammalian system. A list of parameter values used can be found in Appendix 3 - Supplementary Table 2. A list of concentrations used to estimate unknown parameters can be found in Appendix 3 – Supplementary Table 3.

Model Simulation

The model was replicated in MATLAB and was simulated using the non-stiff differential equation solver function ‘ode23t’. To ensure that the model had been replicated accurately,

the SBML version of the model was also simulated using the SBML ODE Solver built into CellDesigner (Funahashi et al., 2003; Machné et al., 2006). Initial conditions for each entity were estimated using control group data from cardiovascular disease studies. Concentrations of LDL and HDL in the blood were kept constant, to reflect a stable patient context.

The resulting model contains 89 ordinary differential equations (ODEs), which are detailed in Appendix 3 - Supplementary Table 1, which are used to describe the development over time of 89 different biological entities.

Supplementary Table 1. The equations describing the model of atherosclerosis.

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
1 - LDL_L	LDL (in Lumen)	190000	(Lusis, 2000)	µg/dL	$\frac{dLDL_L}{dt} = 0$
2 – LDL_I	LDL (in Intima)	0	(Tabas et al., 2007)	µg/mg tissue	$\frac{dLDL_I}{dt} = j_{LL} * {LDL_L}^3 - k_{lpb} * LDL_I * PG$
3 – PBLDL	Proteoglycan bound LDL	0	(Young and McEneny, 2001) (Zalewski and Macphee, 2005) (Lusis, 2000)	µg/mg tissue	$\frac{dPBLDL}{dt} = k_{lpb} * LDL_I - \frac{k_{oxl} * PBLDL * Oxy * PLA * SMase * Lipoxy}{k_{moxl} + PBLDL}$
4 – PGs	Proteoglycans	10	(Tabas et al., 2007)	Ng/mg tissue	$\frac{dPG}{dt} = 0$
5 – Oxy	Free Oxygen Radicals	500	(Young and McEneny, 2001) (Steinberg and Witztum, 2010)	units	$\frac{dOxy}{dt} = 0$
6 – PLA	PLA2	10	(Tellis and Tselepis, 2009)	Ng/mg tissue	$\frac{dPLA}{dt} = 0$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
7 – SMAse	SMAse	10	(Kinnunen and Holopainen, 2002)	Ng/mg tissue	$\frac{dSMA}{dt} = 0$
8 – Lipoxy	Lipooxygenase	10	(Funk, 2006)	Ng/mg tissue	$\frac{dLI}{dt} = 0$
9 – OxLDL	Oxidized Low Density Lipoprotein	1000000	(Yoshida and Kisugi, 2010)	fg/mg tissue	$\frac{dOxy}{dt} = \frac{k_{oxl} * PBLDL * Oxy * PLA * SMase * Lipoxy}{k_{moxl} + PBLDL} - \frac{k_{fcr} * Mac * Oxy * IFNG}{((k_{mfc} + Mac) * (k_{mfco} + oxLDL))}$
10 – HDL_L	High Density Lipoprotein in Lumen	40000	(Lusis, 2000)	µg/dl	$\frac{dHDLL}{dt} = 0$
11 – HDL_I	High Density Lipoprotein in Intima	0	(von Eckardstein et al., 2001)	µg/mg tissue	$\frac{dHDLI}{dt} = j_{hl} * HDLL^2 - k_{hpb} * HDLI * PGs$
12 – PBHDL	Proteoglycan Bound High Density Lipoprotein	0	(Tabas et al., 2007) (Cobbold et al., 2002b)	µg/mg tissue	$\frac{dPBHDL}{dt} = k_{hpb} * HDL_I - \frac{k_{oxh} * PBHDL * Oxy * PLA * SMase * Lipoxy}{k_{moxh} + PBHDL} - k_{clhdli*FC*ABCA1*ABCG1*PBHDL}$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
13 – OxHDL	Oxidized HDL	0	(Heinecke, 2006)	fg/mg tissue	$\frac{doxLDL}{dt} = \frac{k_{oxh} * PBLDL * Oxy * PLA * SMase * Lipoxy}{k_{moxh} + PBHDL}$
14 – MCP1 in Lumen	Monocyte Chemoattractant Protein 1	150000	(Valente et al., 2014) (Pirillo et al., 2013) (Harrington, 2000)	Fg/ml	$\frac{dMCP1L}{dt} = \frac{k_{mcpecMM} * EC * oxLDL * IL1BL}{((km_{mcpil1b} + IL1bL)(km_{mcpec} + oxLDL)} - d_{mcpl} * MCP1L$
15 – Endothelial Cells	Endothelial Cells	4000	(Davignon and Ganz, 2004)	Cells/ mm^2	$\frac{dEC}{dt} = 0$
16 – Monocytes	Monocytes in Lumen	20000	(Bobryshev, 2006)	Cells/ ml	$\frac{dMonoL}{dt} = k_{mono} * MCP1L - k_{monoL} * MonoL$
17 – Monocytes	Monocytes in Intima	0	(Bobryshev, 2006) (Qiao et al., 1997)	Total Cells	$\frac{dMonoI}{dt} = k_{monoL} * MonoL - k_{mdiff} * MonoI * MCSF$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
18 – Foam Cell	Foam Cells	0	(Yu et al., 2013) (McLaren and Ramji, 2009) (Ohashi et al., 2005) (Oram and Lawn, 2001) (Yvan-Charvet et al., 2010)	Total Cells	$\frac{dFC}{dt} = \frac{k_{fc} * Mac * Oxy * IFNG}{((k_{mfc} + Mac) * (k_{mfco} + oxLDL))} - d_{fc} * FC - k_{rct} * PBHDL * ABCA1 * ABCG1 * FC$
19 – IFNG	Interferon Gamma	10	(McLaren and Ramji, 2009)	Fg/mg tissue	$\frac{dIFNG}{dt} = \frac{k_{IFNG1} * Th1}{k_{mifng1} + Th1} + \frac{k_{IFNG17} + Th17}{k_{mifng17} + Th17} - d_{IFNG} * IFNG$
20 – MCSF	Monocyte Colony Stimulating Factor 1	100	(Qiao et al., 1997)	Fg/mg tissue	$\frac{dMCSF}{dt} = k_{MCSF} * Mac - d_{MCSF} * MCSF$
21 – MCP-1	MCP-1 in Intima	1	(Harrington, 2000)	Fg/mg tissue	$\frac{dMCPI}{dt} = k_{mcpi} * FC - d_{mcpi} * MCPI - \frac{k_{mcpi} * MCPI}{k_{mmcpi} + MCPI}$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
22 – Macrophages	Macrophages	0	(McLaren and Ramji, 2009) (Qiao et al., 1997) (Schrijvers et al., 2007)	Total Cells	$\frac{dMac}{dt} = k_{mdiff} * MonoI * MCSF - \frac{k_{fc} * Mac * Oxy * IFNG}{\left(\left(k_{mfc} + Mac\right) * \left(k_{mfco} + oxLDL\right)\right)} - d_{mc} * Mac$
23 – CXCL9	Chemokine (C-X-C motif) ligand 9	17400	(Zernecke et al., 2008)	Pg/ml	$\frac{dC9}{dt} = \frac{k_{cx9} * EC * IFNG}{k_{mcx9} + IFNG} - d_{ifng} * C9$
24 – CXCL10	Chemokine (C-X-C motif) ligand 10	110000	(Zernecke et al., 2008)	Pg/ml	$\frac{dC10}{dt} = \frac{k_{cx10} * EC * IFNG}{k_{mcx10} + IFNG} - d_{ifng} * C10$
25 – CXCL11	Chemokine (C-X-C motif) ligand 11	400000	(Zernecke et al., 2008)	Pg/ml	$\frac{dC11}{dt} = \frac{k_{cx11} * EC * IFNG}{k_{mcx11} + IFNG} - d_{ifng} * C11$
26 – T Cell	T Cells in Lumen	500000	(Tse et al., 2013b)	Cells/ ml	$\frac{dTL}{dt} = \frac{k_{mm2tl} * C9 * C10 * c11 * C}{((kt_{mcx9} + C9) * (kt_{mcx10} + C10) * (kt_{mcx11} + } - k_{tt} * TL$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
27 – T Cell	T Cells in Intima	0	(Tse et al., 2013a) (Lee et al., 1999) (Badimon, 2012) (Lee and Hirani, 2006) (Mallat et al., 1999) (Merhi-Soussi et al., 2005) (Grainger, 2007) (Tse et al., 2013a)	Total Cells	$\frac{dTI}{dt} = k_{tt} * TL - \frac{k_{th1} * TI * IL12 * IL18}{((km_{il12} + IL12) * (km_{IL18} + IL18))} - \frac{k_{th2} * TI * IL4 * IL10}{((km_{IL4} + IL4) * (km_{IL10} + IL10))} - \frac{k_{th17} * TI * IL10 * TGF}{((km_{IL10} + IL10) * (km_{TGF} + TGF))} - \frac{k_{fh} * TI * IL6 * IL21}{((km_{IL6} + IL6) * (km_{IL21} + IL21))} - \frac{k_{reg} * TI * IL10 * TGF}{((km_{IL10} + IL10) * (km_{TGF} + TGF))} - k_{tnk} * TI$
28 – IL1B	Interleukin 1 Beta	1	(Moore et al., 2013b) (Bennett et al., 2016)	Ng/mg tissue	$\frac{dIL1B}{dt} = k_{il1m} * FC + k_{il1s} * SMC - d_{il1} * IL1$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
29 – IL2	Interleukin 2	1	(Upadhya et al., 2004)	Ng/mg tissue	$\frac{dIL2}{dt} = k_{il2t1} * Th1 + k_{il2c} * DC - d_{il2} * IL2$
30 – IL4	Interleukin 4	1	(Mallat et al., 2009)	Ng/mg tissue	$\frac{dIL4}{dt} = k_{il4t2} * Th2 + k_{il4tfh} * Tfh - d_{il4} * IL4$
31 – IL5	Interleukin 5	1	(Mallat et al., 2009)	Ng/mg tissue	$\frac{dIL5}{dt} = k_{il5} * Th2 - d_{il5} * IL5$
32 – IL6	Interleukin 6	1	(Moore et al., 2013b) (Bennett et al., 2016)	Ng/mg tissue	$\frac{dIL6}{dt} = k_{il6M} * Mac + k_{il6S} * SMC + k_{il6C} * MC - d_{il6} * IL6$
33 – IL10	Interleukin 10	1	(Pastrana et al., 2012) (Tsiantoulas et al., 2014)	Ng/mg tissue	$\frac{dIL10}{dt} = k_{il10r} * Treg + k_{il10b1} * B1 + k_{il10b2} * B2 + k_{il10t2} * Th2 + k_{il10m} * Mac - d_{il10} * IL10$
34 – IL12	Interleukin 12	1	(Lee et al., 1999) (Sartori et al., 1997)	Ng/mg tissue	$\frac{dIL12}{dt} = k_{il12b} * B2 + k_{il12d} * DC + k_{il12f} * FC + k_{il12m} * Mac - d_{il12} * IL12$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
35 – IL18	Interleukin 18	1	(Tse et al., 2013a) (Xu et al., 2000)	Ng/mg tissue	$\frac{dIL18}{dt} = k_{il18} * Mac - d_{il18} * IL18$
36 – IL21	Interleukin 21	1	(Tse et al., 2013a)	Ng/mg tissue	$\frac{dIL21}{dt} = k_{il21} * Tfh - d_{il21} * IL21$
37 – IL33	Interleukin 33	100	(Demyanets et al., 2011)	Ng/mg tissue	$\frac{dIL33}{dt} = 0$
38 – TGFB	Transforming Growth Factor Beta	250	(Pastrana et al., 2012)	Ng/mg tissue	$\frac{dTGF}{dt} = k_{TGFr} * Treg + k_{TGFs} * SMC - d_{tgf} * TGF$
39 – Th1	T helper 1 Cells	0	(Tse et al., 2013a) (Mallat et al., 2009) (Gerdes et al., 2002)	Total Cells	$\frac{dTh1}{dt} = \frac{k_{th1} * TI * IL12 * IL18}{((km_{IL12} + IL12) * (km_{IL18} + IL18))} - d_{th1} * Th1$
40 – Th2 Cells	T helper 2 Cells	0	(Tse et al., 2013a)	Total Cells	$\frac{dTh2}{dt} = \frac{k_{th2} * TI * IL4 * IL10}{((km_{IL4} + IL4) * (km_{IL10} + IL10))} - d_{th2} * Th2$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
41 – Th17 Cells	T helper 17 Cells	0	(Tse et al., 2013a)	Total Cells	$\frac{dTh17}{dt} = k_{th17} * \frac{TI * IL10 * TGF}{((km_{IL10} + IL10) * (km_{TGF} + TGF))} - d_{th17} * Th17$
42 – TFH Cells	T familial helper cells	0	(Tse et al., 2013a)	Total Cells	$\frac{dTfh}{dt} = k_{fh} * \frac{TI * IL6 * IL21}{((km_{IL6} + IL6) * (km_{IL21} + IL21))} - d_{tfh} * Tfh$
43 – TNK Cells	Natural Killer T Cells	0	(Tse et al., 2013a)	Total Cells	$\frac{dTnk}{dt} = k_{tnk} * TI - d_{tnk} * Tnk$
44 – Treg Cells	Regulatory T Cells	0	(Tse et al., 2013a) (Pastrana et al., 2012)	Total Cells	$\frac{dTreg}{dt} = k_{reg} * \frac{TI * IL10 * TGF}{((km_{IL10} + IL10) * (km_{TGF} + TGF))} - d_{treg} * Treg$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
45 – TNFA	Tumor Necrosis Factor Alpha	0	(Moore et al., 2013b) (Theoharides et al., 2012) (Damsker et al., 2010) (Amour et al., 1998)	Ng/mg tissue	$\frac{dTNF}{dt} = k_{TNFFC} * FC * {TACE}/{(1 + \frac{TIMP3}{k_{tt3i}})} + k_{TNFMC} * MC * {TACE}/{(1 + \frac{TIMP3}{k_{tt3i}})} + k_{TNFT1} * Th1 * {TACE}/{(1 + \frac{TIMP3}{k_{tt3i}})} + k_{TNFTh17} * Th17 * {TACE}/{(1 + \frac{TIMP3}{k_{tt3i}})} - d_{TNF} * TNF$
46 – MMP1	Matrix Metalloproteinase 1	0	(Newby, 2005) (Saunders et al., 2005)	Ng/mg tissue	$\frac{dMMP1}{dt} = k_{MMP1} * FC * Chy * Try - d_{MMP1} * MMP1 - c_{mmp1} * MMP1 * (TIMP1 + TIMP2 + TIMP3 + TIMP4)$
47 – MMP3	Matrix Metalloproteinase 3	0	(Newby, 2005)	Ng/mg tissue	$\frac{dMMP3}{dt} = k_{MMP3} * FC * Chy * Try - d_{MMP3} * MMP3 - c_{mmp3} * MMP3 * (TIMP1 + TIMP2 + TIMP3 + TIMP4)$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
48 – MMP9	Matrix Metalloproteinase 9	0	(Newby, 2005) (Ardi et al., 2007)	Ng/mg tissue	$\frac{dMMP9}{dt} = k_{MMP9} * FC * Neut - d_{MMP9} * MMP9 - c_{mmp9} * MMP9 * (TIMP1 + TIMP2 + TIMP3 + TIMP4)$
49 – MMP13	Matrix Metalloproteinase 13	0	(Newby, 2005)	Ng/mg tissue	$\frac{dMMP13}{dt} = k_{MMP13} * FC - d_{MMP13} * MMP13 - c_{mmp13} * MMP13 * (TIMP1 + TIMP2 + TIMP3 + TIMP4)$
50 – TIMP1	Tissue Inhibitor of MMPs 1	0	(Newby, 2008)	Ng/mg tissue	$\frac{dTIMP1}{dt} = k_{TIMP1} * Mac * Chy - d_{TIMP1} * TIMP1 - c_{TIMP1} * TIMP1 * (MMP1 + MMP2 + MMP3 + MMP9 + MMP13)$
51 – TIMP2	Tissue Inhibitor of MMPs 2	0	(Newby, 2008)	Ng/mg tissue	$\frac{dTIMP2}{dt} = k_{TIMP2} * Mac - d_{TIMP2} * TIMP2 - c_{TIMP2} * TIMP2 * (MMP1 + MMP2 + MMP3 + MMP9 + MMP13)$
52 – TIMP3	Tissue Inhibitor of MMPs 3	0	(Newby, 2008)	Ng/mg tissue	$\frac{dTIMP3}{dt} = k_{TIMP3} * Mac - d_{TIMP3} * TIMP3 - c_{TIMP3} * TIMP3 * (MMP1 + MMP2 + MMP3 + MMP9 + MMP13)$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
53 – TIMP4	Tissue Inhibitor of MMPs 4	0	(Koskivirta et al., 2006)	Ng/mg tissue	$\frac{dTIMP4}{dt} = k_{TIMP4} * Mac - d_{TIMP4} * TIMP4 - c_{TIMP4} * TIMP4 * (MMP1 + MMP2 + MMP3 + MMP9 + MMP13)$
54 – Chymase	Chymase	0	(Bot et al., 2011)	Ng/mg tissue	$\frac{dChy}{dt} = k_{Chy} * MC - d_{Chy} * Chy$
55 – Tryptase	Tryptase	0	(Ramalho et al., 2013)	Ng/mg tissue	$\frac{dTry}{dt} = k_{Try} * MC - d_{Try} * Try$
56 – PDGF	Platelet Derived Growth Factor	0	(Schwartz, 1997)	Ng/mg tissue	$\frac{dPDGF}{dt} = k_{PDGF} * SMC - d_{PDGF} * PDGF$
57 – EGF	Epidermal growth factor	1	(Ichiro et al., 1990)	Ng/mg tissue	$\frac{dEGF}{dt} = 0$
58 – Elastin	Elastin	0	(Rudijanto, 2007) (Yamamoto et al., 1997) Ichiro (Mecham et al., 1997)	Pg/mg tissue	$\frac{dEla}{dt} = k_{ela} * \frac{SMC * TGFB}{(km_{ela} * \left(1 + \frac{EGF}{ki_{egf}}\right) + SMC)} - d_{ela} * Ela * (MMP1 + MMP2 + MMP3 + MMP9 + MMP13)$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
59 – Collagen	Collagen	0	(Rudijanto, 2007) (Butticè et al., 2006) (Lauer-Fields et al., 2002)	Pg/mg tissue	$\frac{dColl}{dt} = k_{Coll} * SMC - d_{Coll} * Coll - re_{coll}(MMP1 + MMP2 + MMP3 + MMP9 + MMP13)$
60 – B1 Cells	B1 Cells	10	(Perry et al., 2012)	Total Cells	$\frac{dB1}{dt} = 0$
61 – B2 Cells	B2 Cells	10	(Perry et al., 2012)	Total Cells	$\frac{dB2}{dt} = 0$
62 – Dendritic Cells	Dendritic Cells	10	(Subramanian and Tabas, 2014)	Total Cells	$\frac{dDC}{dt} = 0$
63 – Neutrophil	Neutrophil	10	(Hartwig et al., 2014)	Total Cells	$\frac{dNeut}{dt} = 0$
64 – Mast Cells	Mast Cells	10	(Theoharides et al., 2012)	Total Cells	$\frac{dMC}{dt} = 0$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
65 – SMCs	Smooth Muscle Cells	0	(Viedt et al., 2002) (Schwartz, 1997)	Total Cells	$\frac{dSMC}{dt} = k_{SMC} * \frac{MCP1 * PDGF}{(km_{SMCP} + PDGF) * (km_{SMCC} + MCP1)} - d_{SMC} * SMC$
66 – Platelets	Platelets	10	(Lievens and von Hundelshausen, 2011)	Total Cells	$\frac{dPl}{dt} = 0$
67 – IL1B	IL1B Lumen	1000	(Dinarello, 2009)	Fg/ml	$\frac{dIL1BL}{dt} = \frac{k_{IL1BL} * Pl * MonoL}{km_{IL1b} + MonoL} - d_{IL1Bl} * IL1BL$
68 – CCL5	CCL5	2700	(Zernecke and Weber, 2010)	Pg/ml	$\frac{dCCL5}{dt} = \frac{k_{CCL5} * Pl * MonoL}{km_{CCL5} + MonoL} - d_{CCL5} * CCL5$
69 – ABCA1	ATP-binding cassette transporter member 1	10	(Wang et al., 2007)	Ng/mg tissue	$\frac{dABCA1}{dt} = 0$
70 – ABCG1	ATP-binding cassette sub-family G member 1	10	(Wang et al., 2007)	Ng/mg tissue	$\frac{dABCG1}{dt} = 0$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
71 – CLHDL	Cholesterol Laden HDL in Intima	0	(Wang et al., 2007) (Oram and Lawn, 2001) (Zhang et al., 2013)	fg/mg tissue	$\frac{dCLHDLI}{dt} = k_{CLHDLI} * FC * ABCA1 * ABCG1 * PBHDL - k_{CHLDLL} * CLHDLI$
72 – CLHDL	Cholesterol Laden HDL in Liver	20	(Zhang et al., 2013)	fg/mg tissue	$\frac{dCLHDLL}{dt} = k_{CHLDLL} * CLHDLI - k_{EHDL} * CLHDLL$
73 – Empty HDL Liver	Empty HDL Liver	20	(Staels and Fonseca, 2009)	fg/mg tissue	$\frac{dEHDL}{dt} = k_{EHDL} * CLHDLL - k_{HDLT} * EHDL$
74 – Non Activated Macrophages	Non Activated Macrophages	0	(Rosenson et al., 2012)	Total Cells	$\frac{dNAM}{dt} = k_{NAM} * FC * ABCA1 * ABCG1 * PBHDL - d_{NAM} * NAM$
75 – IL17	Interleukin 17	0	(Madhur et al., 2011)	ng/mg tissue	$\frac{dIL17}{dt} = k_{il17} * Th17 - d_{il17} * IL17$
76 – Bile Acids	Bile Acids	0	(Staels and Fonseca, 2009)	Nmol/ g	$\frac{dBile}{dt} = k_{bile} * CLHDLL$
77 – Cholesterol	Cholesterol in Intestine	20	(Redgrave, 1970)	Units	$\frac{dCholI}{dt} = 0$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
78 – Chylomicron Intestine	Chylomicrons	0	(Redgrave, 1970)	units	$\frac{dChyI}{dt} = k_{ChyI} * CholI - k_{chyl} * ChyI$
79 – Chylomicron	Chylomicron Lumen	1.4	(Tomkin and Owens, 2012) (Goldberg, 1996)	Ng/ml	$\frac{dChyL}{dt} = k_{chyl} * ChyI - k_{chyrl} * ChyL * LPL$
80 – Chylomicron Renmants	Chylomicron Renmants in Lumen	0	(Sakurai et al., 2005)	Ng/ml	$\frac{dChyRL}{dt} = k_{chyrl} * ChyL * LPL - k_{iccr} * ChyRL$
81 – LPL	Lipoprotein Lipase	20	(Goldberg, 1996)	Ng/ml	$\frac{dLPL}{dt} = 0$
82 – Cholesterol in Liver	Cholesterol in Liver	0	(Cooper, 1997) (Hu et al., 2008) (Zambon et al., 2003)	ng/mg tissue	$\frac{dCholL}{dt} = k_{iccr} * ChyRL$
83 – VLDL Liver	Very Low Density Lipoprotein Liver	0	(Mason, 1998)	ng/mg tissue	$\frac{dVLDLLi}{dt} = k_{VLDLL} * CholL - k_{VLDLt} * VLDLLi$

Node ID	Entity Name	Initial Concentration	Citations	Units	Equation
84 – VLDL Lumen	Very Low Density Lipoprotein Lumen	2000	(Mason, 1998) (Alaupovic et al., 1986)	µg/ml	$\frac{dVLDLL}{dt} = k_{VLDLt} * VLDLLi - k_{IDL} * VLDLL * LPL$
85 – IDL Lumen	Intermediate Density Lipoprotein Lumen	0	(Alaupovic et al., 1986)	µg/ml	$\frac{dIDL}{dt} = k_{IDL} * VLDLL * LPL - k_{LDL} * IDL * LPL$
86 – TGs Lumen	Triglycerides Lumen	146000	(Talayero and Sacks, 2011)	µg/ml	$\frac{dTGs}{dt} = k_{TGv} * VLDLL * LPL + k_{TGi} * IDL * LPL$
87 – FFAs	Free Fatty Acids	10	(Alaupovic et al., 1986)	µg/ml	$\frac{dFFA}{dt} = k_{FFAv} * VLDLL * LPL + k_{FFAi} * IDL * LPL$
88 – TACE	TACE	1	(Canault et al., 2006)	fg/mg tissue	$\frac{dTACE}{dt} = 0$
89 – MMP2	Matrix Metalloproteinase 2	1	(Dick et al., 2011)	Ng/mg tissue	$\frac{dMMP2}{dt} = k_{MMP2} * FC - d_{MMP2} * MMP2 - c_{mmp2} * MMP2 * (TIMP1 + TIMP2 + TIMP3 + TIMP4)$

Supplementary Table 2. The parameters used in the model of atherosclerosis.

Parameter Name	Description	Value	Notes
j_{ll}	Rate of LDL Transfer into intima	5 x10^{-13}	Found: 1 x10^{-3} – (Cilla et al., 2014) 3.84 x10^{-5} µM/s – (Cobbold et al., 2002a) 1 x10^{-7}mg/s - (McKay et al., 2005) Not used due to Hill coefficient
k_{lpb}	Rate of proteoglycans binding to LDL	266	(Camejo et al., 1993)
k_{oxl}	Vmax for rate of oxidization of PB-LDL	10625	(McKay et al., 2005)
k_{moxl}	Km for oxLDL production rate	123	(McKay et al., 2005)
k_{moxh}	Km for oxHDL production rate	180625	(McKay et al., 2005)
k_{oxh}	Vmax for rate of oxidization of PB-HDL	0.10625	(McKay et al., 2005)
j_{hl}	Rate of HDL Transfer into intima	4.625 x10^{-7}	Found: 1.2x10^{-3} µM/s – (Cobbold et al., 2002a) 1e-8mg/s - (McKay et al., 2005) Not used due to Hill coefficient
k_{hpb}	Rate of proteoglycans binding to HDL	266	Estimated – Compared to k_{lpb}
k_{fcr}	Rate of phagocytosis of oxLDL by macrophages	16200	Estimated – See Table 5

Parameter Name	Description	Value	Notes
$k_{mcpecMM}$	Vmax MCP1 Production	0.1350	Estimated – See Table 5
km_{mcpec}	MM Constant of MCP1 Production from ECs	1.05×10^{7}	Estimated – See Table 5
$Km_{mcpil1b}$	MM Constant of MCP1 Production due to IL1B	2	Estimated – See Table 5
K_{mcpl}	VMax of MCP1 transfer from intima	1.7886×10^{-6}	Estimated
Km_{mcpl}	MM for MCP1 transfer from intima	62.5	Estimated
K_{mcpi}	Rate of MCP1 Production from Foam Cells	100	Estimated – See Table 5
D_{mcpl}	Degradation rate of MCP1 in lumen	0.00065	Found: 1.73 (Friedman and Hao, 2014) Not used to maintain protein levels in plasma
D_{mcpi}	Degradation rate of MCP1 in intima	1.73	(Owen and Sherratt, 1997)
D_{mcsf}	Degradation rate of MCSF	4.1472	(Chen et al., 2012)
K_{mono}	Recruitment rate of monocytes in lumen	0.00105125	Estimated – See Table 5
K_{monoL}	Transfer rate of monocytes to intima	0.0080375	Estimated – See Table 5
K_{mdiff}	Rate of monocyte differentiation into macrophages	0.0994	(Bulelzai and Dubbeldam, 2012; Mao et al., 1986)

Parameter Name	Description	Value	Notes
k_{fc}	Vmax of oxLDL phagocytosis by macropages	540	Found: $1.2x10^{-18}$ m^3/cell/s – (Cilla et al., 2014) 1e-6/s – (Díaz-zuccarini et al., 2014) Not used due to equation format
km_{fc}	MM Constant for Foam Cell consumption of oxLDL for macrophages	6000000	Estimated – See Table 5
km_{fco}	MM Constant for Foam Cell comsumption of oxLDL for oxLDL	6000000	Estimated – See Table 5
D_{fc}	Death rate of foam cells	0.0075	Found: 0.3 (Friedman and Hao, 2014) Not used due to disruption to model dynamics
D_{mc}	Death rate of macrophages	0.05	Estimated – See Table 5
K_{mcsf}	VMax rate of MCSF production	50	(Gruber and Gerrard, 1992)
K_{ifng1}	VMax rate of IFNG from Th1 Cells	1.2E6	(Hao and Friedman, 2014; Tsukaguchi et al., 2011)
Km_{ifng1}	MM Coefficient for IFNG Production from Th1 Cells	$7.5 x10^7$	Estimated
Km_{ifng17}	MM Coefficient for IFNG Production from Th17 Cells	$1.2 x10^6$	Estimated
K_{ifng17}	Production rate of IFNG from Th17	$7.5 x10^7$	Estimated

Parameter Name	Description	Value	Notes
	Cells		
D_{ifng}	Degradation rate of IFNG	0.69	(Hao and Friedman, 2014)
K_{rct}	Rate of reverse cholesterol transport	1.5×10^{-8}	Estimated
K_{cx9}	VMax rate of CXCL9	27360	Estimated – See Table 5
K_{cx10}	VMax rate of CXCL10	10000	Estimated – See Table 5
K_{cx11}	VMax rate of CXCL11	4400	Estimated – See Table 5
K_{mcx9}	Michaelis Menten coefficient for CXCL9	950000000	Estimated – See Table 5
K_{mcx10}	Michaelis Menten coefficient for CXCL10	125000000	Estimated – See Table 5
K_{mcx11}	Michaelis Menten coefficient for CXCL11	65000000	Estimated – See Table 5
K_{ccl5}	Production rate of CCL5	11000	Estimated – See Table 5
K_{mccl5}	MM coefficient for CCL5	10000000	Estimated – See Table 5
D_{cx9}	Degradation rate of CXCL9	0.000006	Found: 2.3148×10^{-5} /s –Cilla Not used to maintain protein levels in plasma
D_{cx10}	Degradation rate of CXCL10	0.000005	Found: 2.3148×10^{-5} /s –Cilla Not used to maintain protein levels

Parameter Name	Description	Value	Notes
			in plasma
D_{cx11}	Degradation rate of CXCL11	0.00000625	Found: 2.3148 x10^{-5} /s –Cilla Not used to maintain protein levels in plasma
D_{ccl5}	Degradation rate of CCL5	0.005125	Found: 2.3148 x10^{-5} /s –Cilla Not used to maintain protein levels in plasma
Kt_{mcx9}	MM Constant for T Cell Recruitment due to CXCL9	25000	Estimated – See Table 5
Kt_{mcx10}	MM Constant for T Cell Recruitment due to CXCL10	120000	Estimated – See Table 5
Kt_{mcx11}	MM Constant for T Cell Recruitment due to CXCL11	120000	Estimated – See Table 5
kt_{mc}	MM Constant for T Cell Recruitment due to CCL5	2000	Estimated – See Table 5
K_{mm2tl}	Rate of T Cell Recruitment due to CXCL9/10/11	1000	Estimated – See Table 5
K_{tt}	Rate of T Cell transfer	0.0004	Estimated
K_{th1}	Rate of T Cell differentiation into Th1 Cells	8.097	Estimated
K_{th2}	Rate of T Cell differentiation into Th2 Cells	8.1882	Estimated
K_{th17}	Rate of T Cell	0.644	Estimated

Parameter Name	Description	Value	Notes
	differentiation into Th17 Cells		
K_{tfh}	Rate of T Cell differentiation into Tfh Cells	66.435	Estimated
K_{tnk}	Rate of T Cell differentiation into TNK Cells	0.000014	Estimated
K_{treg}	Rate of T Cell differentiation into Treg Cells	0.00625	Estimated
K_{il1m}	Rate of IL1 production from macrophages	0.8	Estimated – See Table 5
K_{il1s}	Rate of IL1 production from SMCs	0.8	Estimated – See Table 5
D_{il1}	Degradation rate for IL1	2	(Cilla et al., 2014; Zhao et al., 2005)
K_{il2c}	Rate of IL2 Production from Dendritic Cells	7.5	Estimated - (Tsukaguchi et al., 2011) – See Table 5
K_{il2t1}	Rate of IL2 Production from Th1 Cells	7.5	Estimated - (Tsukaguchi et al., 2011) – See Table 5
D_{il2}	Degradation rate for IL2	2	(Cilla et al., 2014; Zhao et al., 2005)
K_{il4t2}	Rate of IL4 Production by T_h2 Cells	1	(Tsukaguchi et al., 2011) - estimated
K_{il4tfh}	Rate of IL4 Production by T_{fh} Cells	1	(Tsukaguchi et al., 2011) - estimated
D_{il4}	Degradation rate	2	(Cilla et al., 2014; Zhao et al., 2005)

Parameter Name	Description	Value	Notes
	for IL4		
K_{il5}	Rate of IL5 Production	0.9	Estimated
D_{il5}	Degradation rate for IL5	2	(Cilla et al., 2014; Zhao et al., 2005)
K_{il6m}	Rate of IL6 Production by macrophages	180	Estimated – See Table 5
K_{il6s}	Rate of IL6 Production by SMCs	180	Estimated – See Table 5
K_{il6c}	Rate of IL6 Production by Mast Cells	180	Estimated – See Table 5
D_{il6}	Degradation rate for IL6	2	(Cilla et al., 2014; Zhao et al., 2005)
K_{il10r}	Rate of IL10 Production by T_{reg} Cells	0.375	Estimated
K_{il10b1}	Rate of IL10 Production by B1 Cells	0.375	Estimated
K_{il10b2}	Rate of IL10 Production by B2 Cells	0.375	Estimated
K_{il10t2}	Rate of IL10 Production by T_h2 Cells	0.375	Estimated
K_{il10m}	Rate of IL10 Production by macrophages	0.375	Estimated
D_{il10}	Degradation rate for IL10	2	(Cilla et al., 2014; Zhao et al., 2005)
K_{il12b}	Rate of IL12	0.375	Estimated based on (Engele et al.,

Parameter Name	Description	Value	Notes
	Production by B2 Cells		2002)
K_{il12d}	Rate of IL12 Production by Dendritic Cells	0.375	Estimated based on (Engele et al., 2002)
K_{il12m}	Rate of IL12 Production by Macrophages	0.375	Estimated based on (Engele et al., 2002)
K_{il12f}	Rate of IL12 Production by Foam Cells	0.375	Estimated based on (Engele et al., 2002)
D_{il12}	Degradation rate for IL12	2	(Cilla et al., 2014; Zhao et al., 2005)
K_{il18}	Rate of IL18 Production	2	Estimated
D_{il18}	Degradation rate for IL18	2	(Cilla et al., 2014; Zhao et al., 2005)
K_{il21}	Rate of IL21 Production	21	Estimated – See Table 5
D_{il21}	Degradation rate for IL21	2	(Cilla et al., 2014; Zhao et al., 2005)
K_{tgfr}	Rate of TGFB Production by T_{reg} Cells	0.03	Estimated
K_{tgfs}	Rate of TGFB Production by SMCs	0.03	Estimated
K_{tgfm}	Rate of TGFB Production by Macrophages	0.03	Estimated
K_{mtgf}	Michaelis Menten constant for TGFB	100000	Estimated
D_{tgf}	Degradation rate	2	(Cilla et al., 2014; Zhao et al., 2005)

Parameter Name	Description	Value	Notes
	for TGF		
D_{th1}	Death rate of Th1 Cells	0.06	(Asquith et al., 2006)
D_{th2}	Death rate of Th2 Cells	0.06	(Asquith et al., 2006)
D_{th17}	Death rate of Th17 Cells	0.06	(Asquith et al., 2006)
D_{tfh}	Death rate of TFH Cells	0.06	(Asquith et al., 2006)
D_{tnk}	Death rate of TNK Cells	0.06	(Asquith et al., 2006)
D_{treg}	Death rate of Treg Cells	0.06	(Asquith et al., 2006)
D_{t0}	Death rate of naïve T Cells	0.02	Estimated
K_{tnfFC}	Production rate of TNFA from Foam Cells	2.2	Estimated – See Table 5
K_{tnfMC}	Production rate of TNFA from Mast Cells	2.2	Estimated – See Table 5
K_{tnfT1}	Production rate of TNFA from Th1 Cells	2.2	Estimated – See Table 5
K_{tnfT17}	Production rate of TNFA from Th17 Cells	2.2	Estimated – See Table 5
D_{tnf}	Degradation rate f TNFA	2	(Cilla et al., 2014; Zhao et al., 2005)
K_{mmp1}	Production rate of MMP1	0.002	Estimated – See Table 5
D_{mmp1}	Degradation rate of MMP1	4.32	(Kim and Friedman, 2010)

Parameter Name	Description	Value	Notes
K_{mmp2}	Production rate of MMP2	0.4	(Hao and Friedman, 2014)
D_{mmp2}	Degradation rate of MMP2	4.32	(Kim and Friedman, 2010)
k_{mmp3}	Production rate of MMP3	0.0004	Estimated
D_{mmp3}	Degradation rate of MMP3	4.32	(Kim and Friedman, 2010)
K_{mmp9}	Production rate of MMP9	20	Estimated – See Table 5
D_{mmp9}	Degradation rate of MMP9	4.32	(Kim and Friedman, 2010)
K_{mmp13}	Production rate of MMP13	0.25	Estimated – See Table 5
D_{mmp13}	Degradation rate of MMP13	4.32	(Kim and Friedman, 2010)
K_{timp1}	Production rate of TIMP1	200	Estimated – See Table 5
D_{timp1}	Degradation rate of TIMP1	20	Estimated based on (Nunes et al., 2011)
K_{timp2}	Production rate of TIMP2	200	(Hao and Friedman, 2014)
D_{timp2}	Degradation rate of TIMP2	20	Estimated based on (Nunes et al., 2011)
K_{timp3}	Production rate of TIMP3	200	(Hao and Friedman, 2014)
D_{timp3}	Degradation rate of TIMP3	20	Estimated based on (Nunes et al., 2011)
K_{timp4}	Production rate of TIMP4	200	(Hao and Friedman, 2014)
D_{timp4}	Degradation rate of TIMP4	20	Estimated based on (Nunes et al., 2011)

Parameter Name	Description	Value	Notes
C_{mmp1}	Rate of MMP1 binding to TIMPs	$1x10^{-7}$	Estimated
C_{mmp2}	Rate of MMP2 binding to TIMPs	$1x10^{-7}$	Estimated
C_{mmp3}	Rate of MMP3 binding to TIMPs	$1x10^{-7}$	Estimated
C_{mmp9}	Rate of MMP9 binding to TIMPs	$1x10^{-7}$	Estimated
C_{mmp13}	Rate of MMP13 binding to TIMPs	$1x10^{-7}$	Estimated
C_{timp1}	Rate of TIMP1 binding to MMPs	$1x10^{-7}$	Estimated
C_{timp2}	Rate of TIMP2 binding to MMPs	$1x10^{-7}$	Estimated
C_{timp3}	Rate of TIMP3 binding to MMPs	$1x10^{-7}$	Estimated
C_{timp4}	Rate of TIMP4 binding to MMPs	$1x10^{-7}$	Estimated
K_{chy}	Production rate of Chymase	0.0033125	Estimated – See Table 5
D_{chy}	Degradation rate of Chymase	0.0003125	Estimated – See Table 5
K_{try}	Production rate of Tryptase	0.0042	Estimated – See Table 5
D_{try}	Degradation rate of Tryptase	0.0003125	Estimated – See Table 5
K_{pdgf}	Production rate of PDGF	7500	Estimated – See Table 5
D_{pdgf}	Degradation rate for PDGF	2.4	(Fok, 2012)
K_{ela}	Production rate for Elastin	0.01	Estimated

Parameter Name	Description	Value	Notes
D_{ela}	Degradation rate for Elastin	2.57 x10^{-7}	(Friedman and Hao, 2014)
K_{mela}	MM Constant for Elastin Production	100000	Estimated
K_{tt3i}	Inhibition constant for TIMP3-TACE binding	1000000	Estimated
K_{coll}	Production rate for Collagen	25	Estimated – See Table 5
D_{coll}	Degradation rate for Collagen	0.0333	Estimated
K_{smc}	Proliferation rate for SMCs	9000	Estimated – See Table 5
K_{msmcp}	MM Constant for proliferation due to MCP1	500	Estimated – See Table 5
K_{msmcc}	MM Constant for proliferation due to PDGF	500	Estimated
D_{smc}	Death rate of SMCs	0.144	(Bennett et al., 1995)
K_{il1bl}	Production rate of IL1B	10.125	Estimated – See Table 5
K_{mil1b}	MM Constant for IL1B Production	50000	Estimated – See Table 5
D_{il1bl}	Degradation rate of IL1B	0.0003125	Estimated – See Table 5
K_{clhdli}	Rate of HDL consumption of cholesterol in RCT Process	0.0000006	Estimated
K_{clhdll}	Rate of transfer of CL-HDL from intima	0.0003125	Estimated

Parameter Name	Description	Value	Notes
K_{ehdl}	Rate of cholesterol removal from HDL	0.0003125	Estimated
K_{hdlt}	Rate of HDL removal from liver	0.0003125	Estimated
K_{bile}	Rate of bile acid production	0.000856	(Mc Auley et al., 2012)
K_{chyi}	Rate of chylomicron formation in intestine	0.3125	Estimated
K_{chyl}	Rate of chylomicron transfer to lumen	0.0001125	Estimated
K_{chyrl}	Rate of chylomicron remnants production	0.0006125	Estimated
K_{iccr}	Rate of cholesterol formation in liver	0.01	Estimated
K_{vldll}	Rate of VLDL production in liver	1	(Eussen et al., 2011)
$K\__{ldlt}$	Rate of VLDL transfer to intima	0.016	(Mc Auley et al., 2012)
K_{idl}	Rate of VLDL delipidation	3.125×10^{-8}	Estimated
K_{ldl}	Rate of IDL delipidation	3.125×10^{-10}	Estimated
K_{tgv}	Rate of TG production in liver from VLDL	2.5125×10^{-8}	Estimated
K_{tgi}	Rate of TG production from IDL	2.5125×10^{-8}	Estimated
K_{ffav}	Rate of Free Fatty	3.125×10^{-10}	Estimated

Parameter Name	Description	Value	Notes
	Acid production from VLDL		
K_{ffai}	Rate of Free Fatty Acid production from IDL	3.125×10^{-10}	Estimated
K_{nam}	Non-activated macrophages production	0. 000000015	Estimated
D_{nam}	Death rate of non-activated macrophages	0.33	Estimated
K_{il17}	Production rate of IL17	30	Estimated
D_{il17}	Degradation rate of IL17	2	(Cilla et al., 2014; Zhao et al., 2005)
K_{mil12}	MM Constant for IL12 Production	100000	Estimated
K_{mil18}	MM Constant for IL18 Production	100000	Estimated
K_{mil4}	MM Constant for IL4 Production	100000	Estimated
K_{mil10}	MM Constant for IL10 Production	100000	Estimated
K_{mtgf}	MM Constant for TGFB Production	100000	Estimated
K_{mil6}	MM Constant for IL6 Production	1000000	Estimated
K_{mil21}	MM Constant for IL21 Production	1000000	Estimated

Supplementary Table 3. Boundary conditions for the model of atherosclerosis.

LDL	40-190 mg/dL	(Lemieux et al., 2015)
oxLDL	1-40 ng/mg tissue	(Nishi et al., 2002)(Hoff et al., 1978)
HDL	40-70 mg/dL	(Lemieux et al., 2015)
MCP1 – Lumen	113 - 754 pg/ml	(Arakelyan et al., 2005)
Monocytes – Lumen	20000-80000 cells/ml	(Johnsen et al., 2005)
CXCL9 – Lumen	17.4-271.2 pg/ml	(Yu et al., 2015a)
CXCL10 – Lumen	127-956.5 pg/ml	(Tavakolian Ferdousie et al., 2017)
CXCL11 – Lumen	420-1062 pg/ml	(Kao et al., 2003)
T Cells – Lumen	500000-1500000 cells/ml	(Backteman et al., 2012)
IL1B – Lumen	0.28-2.12pg/ml	(Di Iorio et al., 2003)
CCL5 - Lumen	2.7ng/ml -176ng/ml	(Kathiresan et al., 2011)
TGs – Lumen	58-1005mg/dl	(Gotto, 1998)
MMP1	18-104 ng/g tissue	(Molloy et al., 2004)
MMP9	121-722 ng/g tissue	(Molloy et al., 2004)
TIMP1	5340-12380 ng/g tissue	(Molloy et al., 2004)
IL1B Intima	12-24 ng/g tissue	(Molloy et al., 2004)
IL6	1475-5146 ng/g tissue	(Molloy et al., 2004)
TNFA	15-27 ng/g tissue	(Molloy et al., 2004)
IL2	15.3-24 pg/mg protein	(Ragino et al., 2012)
MCP1	150-650 pg/mg protein	(Ragino et al., 2012)
IL18	2-10.7 pg/mg protein	(Ragino et al., 2012)
IL10	1.5 – 2.3 pg/mg tissue	(Stein et al., 2008)

IL12	3.6 – 4.6 pg/mg tissue	(Stein et al., 2008)
IFNG	20-182 fg/mg tissue	(Grufman et al., 2014)
TGF-Beta	0.33-0.76 ng/ µg tissue	(Herder et al., 2012)
Chymase	107.775 (No units)	(Ramalho et al., 2013)
Tryptase	135.113 (No units)	(Ramalho et al., 2013)
Elastin	1.580 ng/g tissue	(Gonçalves et al., 2003)
Collagen	6.260 ng/g tissue	(Gonçalves et al., 2003)
Chylomicrons	1.4-52.6 µg/ml	(Sakai et al., 2003)

Supplementary Table 4: Cell types involved in our model of atherosclerosis

Monocytes in Intima	Monocytes in Lumen	Platelets	Naïve CD4+ T Cells in Lumen	Endothelial Cells
Naïve CD4+ T Cells in Intima	Th1 Cells	Th2 Cells	Th17 Cells	TNK Cells
TReg Cells	TFH Cells	Macrophages in Intima	Foam Cells	B1 Cells
B2 Cells	Dendritic Cells	Mast Cells	Neutrophils	Smooth Muscle Cells

Supplementary Table 5: Proteins involved in our model of atherosclerosis

ABCA1	ABCG1	CCL2 in Intima	CCL2 in Lumen	CXCL9
CXCL10	CXCL11	Chymase	Collagen	EGF
Elastin	Interferon Gamma	IL-1B in Intima	IL-10	IL-12
IL-17	IL-18	IL-1B in Lumen	IL-2	IL-21
IL-33	IL-4	IL-5	IL-6	MCSF
MMP1	MMP2	MMP3	MMP9	MMP13
PDGF	PLA2	Proteoglycans	SMAse	TGF-Beta
TIMP1	TIMP2	TIMP3	TIMP4	TNF-Alpha
Tryptase				

Supplementary Table 6: All Lipoprotein Types contained within the SBML model of atherosclerosis

Low-density Lipoprotein in Lumen	High Density Lipoprotein in Lumen	Very High Density Lipoprotein in Lumen	Intermediate-density Lipoprotein in Lumen	Chylomicrons in Lumen
Chylomicron Remnants in Lumen	Chylomicron in Intestine	Empty HDL in Liver	Cholesterol-Laden HDL in Liver	Very Low Density Lipoprotein in Liver
Low Density Lipoprotein in	High Density Lipoprotein in	Oxidized Low Density	Oxidized High Density	Proteoglycan Bound High

Intima	Intima	Lipoprotein in Intima	Lipoprotein in Intima	Density Lipoprotein in Intima
Proteoglycan bound High Density Lipoprotein in Intima	Cholesterol Laden High Density Lipoprotein in Intima			

Supplementary Table 7: Subprocesses of atherosclerosis

Lipoprotein mass transfer into the intima
Lipoprotein retention within the intima
Lipoprotein oxidation
Reverse Cholesterol Transport
Monocyte recruitment and adhesion
Monocyte migration
Monocyte to macrophage differentiation
Phagocytosis of oxidized LDL
T Cell Differentiation
Matrix metalloproteinases and their inhibitors
Smooth muscle cell proliferation
VLDL delipidation
Extra Cellular Matrix

Supplementary Table 8 – Additional boundary conditions

Total Cell Count per plaque area	25-4000 cells/mm^2	(Brandl et al., 1997)
Average Plaque Area	10.9-15.2mm^2	(von Birgelen et al., 1998)
Total Cell Count	217.5 – 62000 cells	(von Birgelen et al., 1998)(Brandl et al., 1997)
Dry Weight of Aeortic Wall	1.08mg/mm^2	(Manley and Mullinger,

(Atherosclerotic)		1967)
Dry Weight of Aeortic Wall (Normal)	0.44mg/mm^2	(Manley and Mullinger, 1967)
Estimated SMC Concentration	23738 cells in advanced plaque	(Bonanno et al., 2000) (von Birgelen et al., 1998)(Brandl et al., 1997)
Estimated Macrophage Concentration	21593 cells in advanced plaque	(Bonanno et al., 2000) (von Birgelen et al., 1998)(Brandl et al., 1997)
Estimated T Cell Concentration	17168 cells in advanced plaque	(Bonanno et al., 2000) (von Birgelen et al., 1998)(Brandl et al., 1997)

The mathematical model detailed above has been implemented in MATLAB, using the method 'ode23t'. Ode23t is a MATLAB function designed to solve systems of moderately stiff ODEs. In order to validate the biological relevance of the model, we have obtained a collection of experimentally derived data from the publically available literature.

Appendix 4

Due to the size of the sequences and heatmaps included in this appendix, this data can be found at the following link:

https://www.dropbox.com/s/9klm1xkr0cv6w9r/Appendix4Large.docx?dl=0

Or

https://goo.gl/58u8io

Appendix 5

Binding kinetics calculated as described in Chapter 5.

Receptor Name	Ligand Name	k_{on}	k_d	k_{off}
CCR2-001	CCL2-001	1.86E+06	3.77E-08	7.01E-02
CCR2-001	CCL2-001-69	5.52E+04	3.92E-09	2.16E-04
CCR2-001	CCL2-001-71	2.54E+05	4.89E-10	1.24E-04
CCR2-001	CCL2-004	2.02E+05	1.77E-06	3.57E-01
CCR2-001-355	CCL2-001	2.73E+04	4.03E-10	1.10E-05
CCR2-001-64	CCL2-001	2.87E+05	1.17E-11	3.34E-06
CCR2-002	CCL2-001	1.87E+05	4.52E-08	8.45E-03
CCR2-002-64	CCL2-001	3.92E+04	2.97E-09	1.16E-04
CCR2-201	CCL2-001	8.37E+05	3.41E-08	2.85E-02
CCR2-201-355	CCL2-001	7.05E+05	1.66E-09	1.17E-03
CCR2-201-64	CCL2-001	3.47E+05	7.10E-11	2.46E-05
CCR5-001	CCL5-001	5.94E+03	9.67E-11	5.74E-07
CCR5-001	CCL5-001-40	1.07E+01	3.55E-11	3.80E-10
CCR5-001	CCL5-001-5	4.25E+05	4.08E-09	1.73E-03
CCR5-001	CCL5-001-56	1.98E+03	8.70E-11	1.72E-07
CCR5-001	CCL5-001-67	1.08E+04	1.18E-10	1.27E-06
CCR5-001	CCL5-001-68	5.64E+00	4.79E-10	2.70E-09
CCR5-001	CCL5-002	5.94E+03	9.67E-11	5.74E-07
CCR5-001	CCL5-002-40	1.07E+01	3.55E-11	3.80E-10
CCR5-001	CCL5-002-5	4.25E+05	4.08E-09	1.73E-03
CCR5-001	CCL5-002-56	1.98E+03	8.70E-11	1.72E-07
CCR5-001	CCL5-002-67	1.08E+04	1.18E-10	1.27E-06
CCR5-001	CCL5-002-68	5.64E+00	4.79E-10	2.70E-09
CCR5-001-223	CCL5-001	1.71E+04	9.76E-10	1.67E-05
CCR5-001-335	CCL5-001	3.43E+03	7.37E-08	2.53E-04
CCR5-001-55	CCL5-001	5.29E+03	1.19E-10	6.29E-07
CSF1R-001	CSF1-001	1.37E+09	4.08E-06	5.59E+03
CSF1R-001	CSF1-001-292	1.67E+10	1.73E-05	2.88E+05
CSF1R-001	CSF1-001-408	2.16E+09	1.65E-06	3.55E+03
CSF1R-001	CSF1-001-438	1.34E+04	1.36E-06	1.82E-02
CSF1R-001	CSF1-001-489	8.55E+05	4.66E-05	3.99E+01
CSF1R-001	CSF1-002-292	3.18E+09	3.33E-06	1.06E+04
CSF1R-001	CSF1-002-408	5.13E+04	1.23E-06	6.30E-02
CSF1R-001	CSF1-002-461	2.01E+00	7.81E-05	1.57E-04
CSF1R-001	CSF1-002-489	7.68E+09	9.52E-08	7.31E+02
CSF1R-001	CSF1-002-523	1.35E+02	1.63E-07	2.20E-05
CSF1R-001	CSF1-003	1.68E+04	1.08E-06	1.82E-02
CSF1R-001	CSF1-003-292	1.63E+09	5.52E-05	8.99E+04
CSF1R-001	CSF1-003-373	2.81E-28	2.48E-28	6.96E-56

CSF1R-001	CSF1-004	1.65E+04	2.81E-07	4.64E-03
CSF1R-001-245	CSF1-001	3.81E-188	3.82E-06	1.45E-193
CSF1R-001-362	CSF1-001	7.32E+04	4.49E-06	3.28E-01
CSF1R-201	CSF1-001	1.86E+05	7.56E-06	1.41E+00
CSF1R-201-153	CSF1-001	9.29E+09	2.44E-05	2.26E+05
CSF1R-201-241	CSF1-001	3.19E+05	4.41E-06	1.41E+00
CSF1R-201-245	CSF1-001	4.95E+09	2.54E-05	1.26E+05
CSF1R-201-268	CSF1-001	1.85E+07	2.69E-04	4.98E+03
CSF1R-201-279	CSF1-001	4.99E-140	5.20E-06	2.59E-145
CSF1R-201-298	CSF1-001	1.91E+05	1.00E-06	1.92E-01
CSF1R-201-32	CSF1-001	2.57E+04	5.19E-06	1.33E-01
CSF1R-201-60	CSF1-001	7.69E+03	5.51E-05	4.24E-01
CXCR3-001	CXCL10-001	3.50E+06	1.97E-07	6.91E-01
CXCR3-001	CXCL10-001-18	6.91E+06	6.57E-10	4.54E-03
CXCR3-001	CXCL10-001-29	1.52E+06	1.75E-08	2.66E-02
CXCR3-001	CXCL10-001-33	6.93E+04	3.63E-09	2.52E-04
CXCR3-001	CXCL10-001-58	1.29E+02	2.28E-09	2.94E-07
CXCR3-001	CXCL10-001-66	6.43E-02	3.03E-07	1.95E-08
CXCR3-001	CXCL10-001-68	7.13E+03	2.95E-05	2.10E-01
CXCR3-001	CXCL10-001-80	4.92E+06	1.30E-06	6.37E+00
CXCR3-001	CXCL10-001-96	1.39E+05	1.65E-10	2.29E-05
CXCR3-001	CXCL11-001	3.38E+03	1.71E-10	5.77E-07
CXCR3-001	CXCL11-001-29	6.04E+02	1.38E-07	8.35E-05
CXCR3-001	CXCL11-001-72	5.50E+04	7.25E-08	3.99E-03
CXCR3-001	CXCL11-001-73	1.12E+06	7.63E-05	8.54E+01
CXCR3-001	CXCL11-003	7.55E+05	6.76E-08	5.10E-02
CXCR3-001	CXCL11-003-29	2.73E+06	2.61E-05	7.12E+01
CXCR3-001	CXCL11-003-72	5.27E+04	1.38E-07	7.25E-03
CXCR3-001	CXCL11-003-73	4.22E+02	4.88E-07	2.06E-04
CXCR3-001	CXCL9-001	1.82E+05	3.56E-06	6.48E-01
CXCR3-001	CXCL9-001-1	3.31E+05	9.08E-08	3.00E-02
CXCR3-001	CXCL9-001-101	1.95E+07	6.51E-08	1.27E+00
CXCR3-001	CXCL9-001-125	4.24E+05	4.56E-07	1.93E-01
CXCR3-001	CXCL9-001-40	2.76E+06	7.99E-07	2.21E+00
CXCR3-001	CXCL9-001-71	1.69E+05	3.11E-09	5.26E-04
CXCR3-002	CXCL10-001	1.02E+06	6.45E-06	6.58E+00
CXCR3-002	CXCL11-001	1.81E+04	3.88E-05	7.02E-01
CXCR3-002	CXCL9-001	2.81E+08	6.31E-08	1.77E+01
EGFR-001	EGF-001	1.28E+04	1.29E-06	1.66E-02
EGFR-001	EGF-001-151	2.31E+09	3.05E-07	7.03E+02
EGFR-001	EGF-002-389	8.14E+04	2.43E-05	1.98E+00
EGFR-001-521	EGF-001	2.07E+05	9.84E-06	2.04E+00
EGFR-002	EGF-001	1.36E+05	1.91E-06	2.59E-01
EGFR-002-521	EGF-001	2.13E-74	2.22E-07	4.72E-81

EGFR-002-640	EGF-001	3.11E+04	4.39E-08	1.37E-03
EGFR-002-703	EGF-001	5.12E+04	1.91E-07	9.76E-03
EGFR-003	EGF-001	7.60E+04	3.36E-07	2.56E-02
EGFR-003-521	EGF-001	1.28E+04	6.60E-08	8.45E-04
EGFR-004-476	EGF-001	3.43E+08	6.70E-07	2.30E+02
EGFR-005	EGF-001	7.45E-144	4.60E-08	3.43E-151
EGFR-201	EGF-001	9.34E+04	5.30E-05	4.95E+00
EGFR-201-521	EGF-001	2.21E+03	6.29E-08	1.39E-04
EGFR-202	EGF-001	5.16E+04	1.03E-05	5.33E-01
IFNGR1-001	IFNG-001	7.34E+06	4.10E-08	3.01E-01
IFNGR1-001	IFNG-001-160	3.31E+09	6.39E-08	2.12E+02
IFNGR1-001	IFNG-001-162	1.09E+09	1.00E-08	1.09E+01
IFNGR1-001	IFNG-001-72	4.60E+07	2.44E-09	1.12E-01
IFNGR1-001-14	IFNG-001	1.08E+07	7.12E-09	7.69E-02
IFNGR1-001-180	IFNG-001	7.55E+03	2.36E-11	1.78E-07
IFNGR1-001-335	IFNG-001	3.68E+08	7.16E-10	2.64E-01
IFNGR1-001-467	IFNG-001	1.01E+07	2.77E-08	2.80E-01
IFNGR1-201	IFNG-001	3.60E+08	1.05E-08	3.77E+00
IFNGR1-201-14	IFNG-001	3.65E+08	1.47E-06	5.37E+02
IFNGR1-201-180	IFNG-001	3.34E+07	1.17E-09	3.91E-02
IFNGR1-201-46	IFNG-001	8.43E+02	3.78E-11	3.18E-08
IFNGR1-201-61	IFNG-001	1.47E+02	3.70E-12	5.44E-10
IFNGR2-001	IFNG-001	3.29E+01	6.33E-09	2.08E-07
IFNGR2-001	IFNG-001-160	2.44E+03	3.89E-10	9.48E-07
IFNGR2-001	IFNG-001-162	5.83E+06	1.80E-09	1.05E-02
IFNGR2-001	IFNG-001-72	3.68E+02	8.74E-12	3.21E-09
IFNGR2-001-64	IFNG-001	4.69E+03	9.63E-12	4.52E-08
IL10RA-001	IL10-001	9.89E+04	4.07E-04	4.02E+01
IL10RA-001	IL10-001-15	1.07E+06	2.16E-06	2.31E+00
IL10RA-001	IL10-001-169	8.11E+04	4.72E-09	3.82E-04
IL10RA-001	IL10-001-19	2.09E+05	1.43E-10	2.99E-05
IL10RA-001	IL10-001-20	5.75E+05	6.08E-08	3.49E-02
IL10RA-001	IL10-001-45	8.12E+03	8.67E-06	7.04E-02
IL10RA-001	IL10-001-71	1.05E+05	2.38E-05	2.50E+00
IL10RA-001	IL10-001-72	1.50E+05	3.71E-05	5.56E+00
IL10RA-001-159	IL10-001	3.88E+05	4.56E-05	1.77E+01
IL10RA-001-224	IL10-001	4.60E+04	8.04E-05	3.70E+00
IL10RA-001-233	IL10-001	1.75E-01	1.21E-13	2.12E-14
IL10RA-001-351	IL10-001	9.10E+04	3.41E-21	3.10E-16
IL10RA-001-420	IL10-001	1.45E-16	6.79E-26	9.84E-42
IL12RB1-001	IL12A-001	1.57E+03	1.18E-29	1.85E-26
IL12RB1-001	IL12A-001-211	5.24E+05	7.36E-07	3.86E-01
IL12RB1-001	IL12A-001-82	2.36E+04	4.92E-10	1.16E-05
IL12RB1-001	IL12A-005	8.10E+04	6.69E-08	5.42E-03

IL12RB1-001	IL12A-005-115	3.74E+04	4.69E-11	1.75E-06
IL12RB1-001	IL12A-005-153	1.56E+06	8.55E-11	1.33E-04
IL12RB1-001	IL12A-005-164	4.57E+04	1.48E-09	6.77E-05
IL12RB1-001	IL12A-005-200	4.49E+07	1.14E-09	5.14E-02
IL12RB1-001	IL12A-005-3	6.81E+05	1.45E-18	9.87E-13
IL12RB1-001	IL12A-005-81	5.39E+07	1.06E-05	5.71E+02
IL12RB1-001	IL12A-006	9.19E+02	9.06E-11	8.33E-08
IL12RB1-001	IL12A-006-139	8.94E+06	3.62E-12	3.23E-05
IL12RB1-001	IL12A-006-157	1.50E+06	2.32E-06	3.48E+00
IL12RB1-001	IL12A-006-177	1.42E+05	8.60E-07	1.22E-01
IL12RB1-001	IL12A-006-224	4.11E+04	1.17E-09	4.82E-05
IL12RB1-001	IL12A-006-28	3.49E-05	2.17E-14	7.58E-19
IL12RB1-001	IL12A-006-3	7.19E+04	6.78E-07	4.88E-02
IL12RB1-001	IL12A-006-81	1.44E+05	2.88E-07	4.14E-02
IL12RB1-001-156	IL12A-001	5.08E+01	3.63E-07	1.85E-05
IL12RB1-001-214	IL12A-001	4.15E+03	3.45E-22	1.43E-18
IL12RB1-001-352	IL12A-001	8.56E-35	1.08E-21	9.24E-56
IL12RB1-001-365	IL12A-001	7.64E-09	1.09E-18	8.33E-27
IL12RB1-001-378	IL12A-001	3.08E+03	1.45E-06	4.47E-03
IL12RB1-001-91	IL12A-001	8.55E-15	8.17E-27	6.98E-41
IL12RB1-002-156	IL12A-001	5.08E+01	3.63E-07	1.85E-05
IL12RB1-002-214	IL12A-001	1.14E-15	3.45E-22	3.93E-37
IL12RB1-002-378	IL12A-001	3.08E+03	1.45E-06	4.47E-03
IL12RB1-002-47	IL12A-001	3.53E-22	2.62E-19	9.24E-41
IL12RB1-002-91	IL12A-001	8.55E-15	8.17E-27	6.98E-41
IL12RB1-003-201	IL12A-001	9.05E+06	7.78E-06	7.04E+01
IL17RA-001	IL17A-001	4.45E-02	4.61E-13	2.05E-14
IL17RA-001	IL17A-001-134	7.48E+00	2.60E-05	1.94E-04
IL17RA-001	IL17A-001-43	1.47E-13	4.60E-06	6.77E-19
IL17RA-001	IL17A-001-69	2.43E+03	9.95E-05	2.42E-01
IL17RA-001	IL17A-001-71	5.42E+08	3.29E-06	1.78E+03
IL17RA-001	IL17A-001-74	3.18E-03	5.07E-06	1.61E-08
IL17RA-001	IL17A-001-95	1.79E-84	5.83E-07	1.04E-90
IL17RA-001-367	IL17A-001	1.80E-03	1.73E-11	3.11E-14
IL17RA-001-607	IL17A-001	4.46E+03	5.29E-07	2.36E-03
IL17RA-001-691	IL17A-001	6.96E+00	4.40E-07	3.06E-06
IL18R1-001	IL18-001	4.16E+01	6.93E-11	2.88E-09
IL18R1-001	IL18-001-127	1.37E-04	7.21E-11	9.88E-15
IL18R1-001	IL18-001-164	1.49E-01	9.41E-13	1.40E-13
IL18R1-001	IL18-001-22	3.26E+03	1.96E-10	6.40E-07
IL18R1-001	IL18-001-47	4.57E+01	1.16E-09	5.29E-08
IL18R1-001	IL18-001-63	2.89E+03	1.03E-07	2.98E-04
IL18R1-001	IL18-003	5.86E+03	7.81E-10	4.58E-06
IL18R1-001	IL18-003-123	1.06E+04	6.05E-11	6.41E-07

IL18R1-001	IL18-003-160	8.74E+01	2.26E-10	1.97E-08
IL18R1-001	IL18-003-22	2.06E+00	2.05E-11	4.21E-11
IL18R1-001	IL18-003-43	5.12E+01	3.60E-11	1.84E-09
IL18R1-001	IL18-003-59	1.50E+04	2.79E-10	4.19E-06
IL18R1-001	IL18-006	4.16E+01	6.93E-11	2.88E-09
IL18R1-001	IL18-006-127	1.37E-04	7.21E-11	9.88E-15
IL18R1-001	IL18-006-164	1.49E-01	9.41E-13	1.40E-13
IL18R1-001	IL18-006-22	3.26E+03	1.96E-10	6.40E-07
IL18R1-001	IL18-006-47	4.57E+01	1.16E-09	5.29E-08
IL18R1-001	IL18-006-63	2.89E+03	1.03E-07	2.98E-04
IL18R1-001-170	IL18-001	3.81E+03	7.25E-09	2.76E-05
IL18R1-001-232	IL18-001	1.26E+06	2.27E-08	2.86E-02
IL18R1-001-423	IL18-001	1.05E+03	9.98E-11	1.05E-07
IL18R1-201	IL18-001	4.16E+01	6.93E-11	2.88E-09
IL18R1-201-170	IL18-001	3.81E+03	7.25E-09	2.76E-05
IL18R1-201-232	IL18-001	1.26E+06	2.27E-08	2.86E-02
IL18R1-201-423	IL18-001	1.05E+03	9.98E-11	1.05E-07
IL18R1-202	IL18-001	5.78E+04	2.11E-11	1.22E-06
IL18R1-202-100	IL18-001	3.05E-02	2.52E-10	7.67E-12
IL18R1-202-117	IL18-001	5.63E+05	4.04E-09	2.27E-03
IL18R1-202-139	IL18-001	3.42E+04	2.47E-09	8.46E-05
IL18R1-202-162	IL18-001	1.45E+05	3.95E-10	5.72E-05
IL18R1-202-53	IL18-001	4.32E+04	1.76E-10	7.59E-06
IL18R1-202-84	IL18-001	2.00E+04	3.11E-10	6.21E-06
IL1R1-001	IL1B-001	4.41E-29	2.41E-20	1.06E-48
IL1R1-001	IL33-001-124	1.56E+07	4.11E-05	6.40E+02
IL1R1-001	IL33-001-153	5.44E-01	4.23E-06	2.30E-06
IL1R1-001	IL33-001-201	1.48E+04	1.08E-05	1.60E-01
IL1R1-001	IL33-001-263	1.05E+07	1.84E-06	1.93E+01
IL1R1-001	IL33-001-83	5.05E+06	2.96E-07	1.50E+00
IL1R1-001	IL33-201-31	2.11E+05	2.08E-06	4.40E-01
IL1R1-001	IL33-201-4	2.43E+10	6.68E-06	1.62E+05
IL1R1-001	IL33-201-75	4.26E+04	7.48E-07	3.19E-02
IL1R1-001	IL33-201-96	3.19E+07	1.18E-04	3.76E+03
IL1R1-001	IL33-202-108	2.46E+00	1.24E-05	3.05E-05
IL1R1-001	IL33-202-180	6.98E+10	4.26E-05	2.97E+06
IL1R1-001	IL33-202-39	2.06E+11	1.08E-04	2.22E+07
IL1R1-001	IL33-202-55	4.28E+10	6.31E-05	2.70E+06
IL1R1-001	IL33-202-57	1.57E-22	9.83E-21	1.54E-42
IL1R1-001	IL33-203	8.31E+06	2.48E-05	2.06E+02
IL1R1-001	IL33-203-31	1.13E+09	5.75E-07	6.50E+02
IL1R1-001	IL33-203-39	2.56E+09	1.86E-04	4.77E+05
IL1R1-001	IL33-203-4	3.80E+04	1.05E-05	3.99E-01
IL1R1-001	IL33-203-95	2.84E+08	7.89E-05	2.24E+04

IL1R1-001-510	IL33-001	2.74E-01	7.30E-07	2.00E-07
IL1R1-006-344	IL33-001	6.87E+08	2.97E-06	2.04E+03
IL1R1-006-425	IL33-001	2.00E+09	4.42E-08	8.85E+01
IL1R1-013	IL33-001	2.18E+05	1.91E-11	4.16E-06
IL1R1-013-124	IL33-001	4.09E+07	2.99E-11	1.22E-03
IL1R1-013-344	IL33-001	1.91E+09	2.74E-11	5.23E-02
IL1RL1-001	IL33-201-96	9.25E+06	3.41E-06	3.16E+01
IL1RL1-001	IL33-202-108	2.32E-37	9.75E-29	2.26E-65
IL1RL1-001-176	IL33-001	4.22E+04	9.91E-12	4.18E-07
IL1RL1-002-216	IL33-001	4.10E+02	1.23E-14	5.04E-12
IL1RL1-008-180	IL33-001	7.79E+00	1.30E-13	1.01E-12
IL1RL1-201-127	IL33-001	2.62E+09	2.01E-05	5.26E+04
IL21R-001	IL21-001	2.46E+05	3.53E-09	8.69E-04
IL21R-001	IL21-001-135	3.18E+08	4.69E-08	1.49E+01
IL21R-001	IL21-001-40	1.25E+10	4.26E-08	5.33E+02
IL21R-001	IL21-201	1.22E+05	3.55E-10	4.33E-05
IL21R-001	IL21-201-135	1.34E+09	3.79E-07	5.08E+02
IL21R-001	IL21-201-40	2.19E+05	1.82E-09	3.97E-04
IL21R-001-318	IL21-001	1.31E+05	2.67E-05	3.50E+00
IL21R-001-484	IL21-001	1.77E+00	4.50E-11	7.96E-11
IL21R-003	IL21-001	2.46E+05	3.53E-09	8.69E-04
IL21R-003-318	IL21-001	1.31E+05	2.67E-05	3.50E+00
IL21R-003-484	IL21-001	1.54E+06	2.03E-07	3.13E-01
IL21R-006	IL21-001	2.46E+05	3.53E-09	8.69E-04
IL21R-006-318	IL21-001	2.28E+05	1.26E-05	2.87E+00
IL21R-006-484	IL21-001	1.54E+06	2.03E-07	3.13E-01
IL2RA-001	IL2-001	7.02E+01	3.27E-12	2.29E-10
IL2RA-001	IL2-001-140	3.07E+02	3.46E-12	1.06E-09
IL2RA-001	IL2-001-21	4.84E-01	1.75E-12	8.48E-13
IL2RA-001	IL2-001-43	1.64E+04	4.35E-11	7.14E-07
IL2RA-001	IL2-001-54	4.15E-04	1.26E-12	5.21E-16
IL2RA-001-153	IL2-001	1.52E+03	2.75E-10	4.18E-07
IL2RA-002	IL2-001	1.10E+02	3.64E-12	4.01E-10
IL2RA-002-111	IL2-001	2.30E+05	6.95E-12	1.60E-06
IL2RA-002-148	IL2-001	3.77E+03	1.57E-09	5.91E-06
IL2RA-002-159	IL2-001	2.14E+03	1.14E-11	2.44E-08
IL2RA-002-161	IL2-001	3.73E+05	1.74E-10	6.50E-05
IL2RA-002-181	IL2-001	3.02E+00	1.32E-09	4.00E-09
IL2RA-002-191	IL2-001	3.54E+05	3.82E-09	1.35E-03
IL2RA-002-200	IL2-001	8.43E+02	2.33E-12	1.96E-09
IL2RA-002-25	IL2-001	9.34E+06	1.29E-06	1.21E+01
IL2RA-002-39	IL2-001	1.62E+08	2.00E-07	3.25E+01
IL2RA-002-42	IL2-001	3.53E+06	9.38E-08	3.31E-01
IL2RA-002-61	IL2-001	1.04E+04	1.28E-08	1.33E-04

IL2RA-002-91	IL2-001	1.53E+00	8.57E-09	1.31E-08
IL2RA-004	IL2-001	8.24E+04	4.31E-12	3.55E-07
IL2RA-004-111	IL2-001	1.24E+05	2.46E-07	3.05E-02
IL2RA-004-153	IL2-001	4.44E+05	1.06E-05	4.69E+00
IL2RA-004-244	IL2-001	5.21E+05	9.32E-08	4.85E-02
IL2RA-004-91	IL2-001	1.05E-04	4.94E-12	5.19E-16
IL4R-001	IL4-001	1.17E+04	4.37E-05	5.12E-01
IL4R-001	IL4-001-100	3.49E-37	6.22E-26	2.17E-62
IL4R-001	IL4-001-105	2.33E+09	1.56E-05	3.64E+04
IL4R-001	IL4-001-109	3.35E+03	1.21E-09	4.05E-06
IL4R-001	IL4-001-152	1.30E+05	9.35E-06	1.21E+00
IL4R-001	IL4-001-22	4.25E-21	3.16E-35	1.34E-55
IL4R-001	IL4-001-26	2.49E+06	1.74E-06	4.34E+00
IL4R-001	IL4-001-30	1.84E+09	2.76E-06	5.08E+03
IL4R-001	IL4-001-53	1.34E+07	2.73E-06	3.66E+01
IL4R-001	IL4-001-98	3.01E+07	9.04E-07	2.72E+01
IL4R-001	IL4-002	5.40E+03	1.03E-06	5.54E-03
IL4R-001	IL4-002-102	1.18E-16	9.57E-14	1.13E-29
IL4R-001	IL4-002-136	2.66E+06	6.21E-07	1.65E+00
IL4R-001	IL4-002-22	2.10E+06	6.57E-05	1.38E+02
IL4R-001	IL4-002-26	1.20E+07	3.55E-07	4.26E+00
IL4R-001	IL4-002-30	1.49E+03	1.14E-11	1.70E-08
IL4R-001	IL4-002-82	1.71E+06	1.54E-06	2.62E+00
IL4R-001	IL4-002-84	5.39E+04	6.01E-06	3.24E-01
IL4R-001	IL4-002-89	1.60E+05	3.59E-09	5.75E-04
IL4R-001	IL4-002-93	7.67E+07	5.26E-08	4.04E+00
IL4R-001	IL4-201	6.87E-08	7.57E-06	5.20E-13
IL4R-001	IL4-201-134	2.46E+10	1.09E-04	2.68E+06
IL4R-001	IL4-201-22	2.45E+08	2.43E-06	5.96E+02
IL4R-001	IL4-201-26	1.77E+04	1.72E-06	3.04E-02
IL4R-001	IL4-201-30	6.18E-03	6.09E-06	3.76E-08
IL4R-001	IL4-201-67	7.50E+08	1.41E-06	1.06E+03
IL4R-001	IL4-201-82	1.07E+06	4.96E-06	5.30E+00
IL4R-001-185	IL4-001	1.67E+08	5.75E-06	9.61E+02
IL4R-001-400	IL4-001	5.41E+03	3.30E-05	1.78E-01
IL4R-001-436	IL4-001	1.29E+07	1.54E-06	1.99E+01
IL4R-001-503	IL4-001	1.56E+03	1.76E-05	2.75E-02
IL4R-001-576	IL4-001	4.12E+05	1.27E-04	5.24E+01
IL4R-001-579	IL4-001	8.56E+05	1.42E-06	1.22E+00
IL4R-001-75	IL4-001	5.93E+02	3.23E-05	1.92E-02
IL4R-001-752	IL4-001	1.85E+05	2.13E-06	3.94E-01
IL4R-001-82	IL4-001	3.08E+09	6.24E-06	1.92E+04
IL4R-001-97	IL4-001	2.15E-09	1.29E-05	2.77E-14
IL4R-004	IL4-001	5.82E+01	7.99E-08	4.65E-06

IL4R-004-185	IL4-001	2.15E+06	5.01E-07	1.08E+00
IL4R-004-400	IL4-001	5.58E+04	2.14E-07	1.19E-02
IL4R-004-431	IL4-001	1.30E+06	3.68E-07	4.78E-01
IL4R-004-436	IL4-001	2.86E-33	5.20E-06	1.49E-38
IL4R-004-503	IL4-001	1.74E+04	7.41E-06	1.29E-01
IL4R-004-576	IL4-001	1.82E+02	8.95E-11	1.63E-08
IL4R-004-579	IL4-001	1.59E+04	7.38E-06	1.17E-01
IL4R-004-75	IL4-001	1.95E+05	4.45E-05	8.67E+00
IL4R-004-752	IL4-001	2.67E+04	8.03E-07	2.14E-02
IL4R-004-786	IL4-001	1.88E+04	8.69E-06	1.63E-01
IL4R-004-82	IL4-001	8.68E+07	8.04E-06	6.98E+02
IL4R-004-97	IL4-001	4.89E+05	8.34E-06	4.08E+00
IL4R-201	IL4-001	2.97E-08	1.90E-14	5.64E-22
IL4R-201-170	IL4-001	4.99E+00	6.00E-06	2.99E-05
IL4R-201-385	IL4-001	1.05E+06	2.05E-05	2.15E+01
IL4R-201-416	IL4-001	6.07E+07	1.28E-05	7.76E+02
IL4R-201-421	IL4-001	2.95E+05	1.33E-04	3.91E+01
IL4R-201-488	IL4-001	8.17E+07	7.70E-07	6.29E+01
IL4R-201-561	IL4-001	5.62E+05	5.79E-07	3.25E-01
IL4R-201-564	IL4-001	3.34E-06	7.13E-12	2.38E-17
IL4R-201-60	IL4-001	1.26E+05	2.15E-06	2.71E-01
IL4R-201-660	IL4-001	2.16E+06	1.51E-05	3.26E+01
IL4R-201-67	IL4-001	1.63E+04	1.37E-06	2.24E-02
IL4R-201-737	IL4-001	9.11E+05	1.11E-05	1.01E+01
IL4R-201-771	IL4-001	4.36E-178	3.84E-06	1.67E-183
IL4R-201-82	IL4-001	7.78E+06	8.21E-07	6.38E+00
IL5RA-001	IL5-001	8.21E+04	2.88E-05	2.36E+00
IL5RA-001	IL5-001-104	3.30E+05	0.00E+00	0.00E+00
IL5RA-001	IL5-001-17	5.88E+04	1.15E-05	6.76E-01
IL5RA-001	IL5-001-40	1.91E+07	1.07E-05	2.03E+02
IL5RA-001	IL5-001-45	2.26E+04	1.32E-05	2.98E-01
IL5RA-001	IL5-001-49	8.75E+05	3.56E-05	3.12E+01
IL5RA-001	IL5-001-6	4.51E+05	1.30E-05	5.85E+00
IL5RA-001	IL5-001-93	4.41E+04	1.21E-05	5.35E-01
IL5RA-001-129	IL5-001	3.95E+06	1.01E-04	3.99E+02
IL5RA-001-359	IL5-001	5.71E+05	7.39E-05	4.22E+01
IL5RA-001-59	IL5-001	6.35E+05	4.87E-05	3.09E+01
IL5RA-002	IL5-001	4.01E+05	2.89E-05	1.16E+01
IL5RA-002-129	IL5-001	3.00E+05	0.00E+00	0.00E+00
IL5RA-002-262	IL5-001	6.36E+04	0.00E+00	0.00E+00
IL5RA-002-59	IL5-001	1.88E+06	0.00E+00	0.00E+00
IL5RA-003	IL5-001	8.18E+05	0.00E+00	0.00E+00
IL5RA-003-129	IL5-001	1.94E+05	0.00E+00	0.00E+00
IL5RA-003-262	IL5-001	5.15E+05	0.00E+00	0.00E+00

IL5RA-003-59	IL5-001	1.95E+05	5.06E-05	9.86E+00
IL5RA-004	IL5-001	3.32E+04	3.06E-05	1.02E+00
IL5RA-004-129	IL5-001	8.37E+06	1.99E-04	1.66E+03
IL5RA-004-359	IL5-001	4.30E+05	1.61E-07	6.91E-02
IL5RA-004-59	IL5-001	1.75E+05	5.98E-05	1.05E+01
IL5RA-005	IL5-001	1.61E+05	8.21E-05	1.32E+01
IL5RA-005-129	IL5-001	4.63E+03	3.75E-06	1.74E-02
IL5RA-005-262	IL5-001	1.18E+06	3.35E-04	3.96E+02
IL5RA-005-59	IL5-001	8.29E+05	7.75E-05	6.42E+01
IL5RA-006	IL5-001	2.47E+05	2.15E-05	5.30E+00
IL5RA-006-129	IL5-001	1.90E+05	1.08E-04	2.05E+01
IL5RA-006-262	IL5-001	8.65E+05	0.00E+00	0.00E+00
IL5RA-006-59	IL5-001	2.64E+05	3.57E-06	9.42E-01
IL5RA-007	IL5-001	6.87E+04	0.00E+00	0.00E+00
IL5RA-007-129	IL5-001	5.85E+05	0.00E+00	0.00E+00
IL5RA-007-359	IL5-001	1.25E+05	1.64E-05	2.05E+00
IL5RA-007-59	IL5-001	2.17E+05	0.00E+00	0.00E+00
IL5RA-201	IL5-001	7.17E+04	1.53E-06	1.10E-01
IL5RA-201-129	IL5-001	6.12E+05	0.00E+00	0.00E+00
IL5RA-201-264	IL5-001	2.75E+03	1.52E-07	4.17E-04
IL5RA-201-287	IL5-001	4.95E+05	0.00E+00	0.00E+00
IL5RA-201-59	IL5-001	7.49E+04	1.14E-07	8.51E-03
IL6R-001	IL6-001	1.30E+05	3.79E-05	4.93E+00
IL6R-001	IL6-001-104	2.76E+06	1.03E-06	2.85E+00
IL6R-001	IL6-001-116	2.31E+04	4.53E-08	1.05E-03
IL6R-001	IL6-001-152	7.44E+04	1.40E-07	1.04E-02
IL6R-001	IL6-001-162	2.68E+04	1.59E-06	4.26E-02
IL6R-001	IL6-001-2	5.50E+05	1.41E-06	7.73E-01
IL6R-001	IL6-001-31	6.57E+03	4.11E-07	2.70E-03
IL6R-001	IL6-001-32	3.32E+05	4.87E-06	1.62E+00
IL6R-001	IL6-001-55	2.35E+05	2.71E-06	6.37E-01
IL6R-001	IL6-001-6	6.27E+05	2.34E-07	1.46E-01
IL6R-001	IL6-001-7	1.73E+05	2.19E-06	3.78E-01
IL6R-001	IL6-001-79	1.48E+06	9.33E-07	1.38E+00
IL6R-001	IL6-003	7.14E+05	9.62E-06	6.87E+00
IL6R-001	IL6-003-110	3.78E+05	3.26E-08	1.23E-02
IL6R-001	IL6-003-28	4.97E+04	2.58E-08	1.28E-03
IL6R-001	IL6-003-3	2.11E+05	4.87E-06	1.03E+00
IL6R-001	IL6-003-40	6.21E+06	5.72E-06	3.55E+01
IL6R-001	IL6-003-76	2.60E+04	3.53E-06	9.17E-02
IL6R-001	IL6-003-89	1.60E+05	3.31E-05	5.30E+00
IL6R-001	IL6-003-92	9.60E+05	1.26E-05	1.20E+01
IL6R-001	IL6-004	1.96E+05	3.59E-06	7.03E-01
IL6R-001	IL6-004-28	9.33E+05	2.53E-07	2.36E-01

IL6R-001	IL6-004-3	9.90E+05	3.92E-06	3.88E+00
IL6R-001	IL6-004-40	2.42E+05	1.77E-04	4.28E+01
IL6R-001	IL6-004-76	4.13E+03	7.67E-09	3.17E-05
IL6R-001	IL6-004-86	2.78E+03	3.06E-09	8.51E-06
IL6R-001	IL6-005	1.63E+04	1.18E-06	1.93E-02
IL6R-001	IL6-005-129	7.10E+04	7.14E-06	5.07E-01
IL6R-001	IL6-005-139	7.26E+04	1.49E-07	1.08E-02
IL6R-001	IL6-005-2	7.70E+04	8.62E-07	6.64E-02
IL6R-001	IL6-005-32	9.82E+04	6.00E-07	5.89E-02
IL6R-001	IL6-005-56	2.12E+05	2.92E-06	6.19E-01
IL6R-001	IL6-005-6	4.91E+04	1.82E-08	8.92E-04
IL6R-001	IL6-005-8	3.26E+05	8.33E-07	2.72E-01
IL6R-001	IL6-005-81	6.95E+04	4.49E-06	3.12E-01
IL6R-001	IL6-005-9	1.56E+05	6.48E-08	1.01E-02
IL6R-001	IL6-005-93	2.61E+02	2.31E-07	6.02E-05
IL6R-001	IL6-006	1.13E+04	4.68E-06	5.29E-02
IL6R-001	IL6-006-104	1.68E+05	2.36E-06	3.96E-01
IL6R-001	IL6-006-116	2.29E+05	3.57E-06	8.17E-01
IL6R-001	IL6-006-152	2.12E+04	8.65E-07	1.83E-02
IL6R-001	IL6-006-165	1.38E+03	7.30E-08	1.01E-04
IL6R-001	IL6-006-168	1.34E+04	1.23E-05	1.65E-01
IL6R-001	IL6-006-186	2.28E+04	2.47E-06	5.63E-02
IL6R-001	IL6-006-2	3.55E+03	1.93E-07	6.84E-04
IL6R-001	IL6-006-31	4.69E+06	9.42E-06	4.42E+01
IL6R-001	IL6-006-32	1.10E+06	1.59E-05	1.74E+01
IL6R-001	IL6-006-55	2.24E+06	1.50E-05	3.36E+01
IL6R-001	IL6-006-6	6.54E+05	2.14E-05	1.40E+01
IL6R-001	IL6-006-7	1.99E+06	6.36E-08	1.27E-01
IL6R-001	IL6-006-79	3.72E+05	3.68E-07	1.37E-01
IL6R-001	IL6-201	1.12E+05	6.15E-07	6.89E-02
IL6R-001	IL6-201-104	1.81E+04	6.09E-09	1.10E-04
IL6R-001	IL6-201-116	7.90E+04	2.99E-06	2.36E-01
IL6R-001	IL6-201-152	6.06E+03	4.42E-09	2.68E-05
IL6R-001	IL6-201-162	9.62E+04	2.27E-08	2.18E-03
IL6R-001	IL6-201-2	5.76E+05	4.87E-06	2.80E+00
IL6R-001	IL6-201-31	1.46E+03	8.87E-08	1.29E-04
IL6R-001	IL6-201-32	1.19E+06	5.09E-06	6.06E+00
IL6R-001	IL6-201-55	9.18E-03	6.02E-10	5.52E-12
IL6R-001	IL6-201-6	1.43E+07	2.85E-05	4.08E+02
IL6R-001	IL6-201-7	6.29E+03	1.68E-08	1.06E-04
IL6R-001	IL6-201-79	5.67E+05	6.72E-07	3.81E-01
IL6R-001-358	IL6-001	1.99E+04	6.38E-08	1.27E-03
IL6R-001-385	IL6-001	1.85E+05	1.19E-05	2.20E+00
IL6R-001-65	IL6-001	9.34E+04	1.81E-07	1.69E-02

IL6R-003	IL6-001	2.10E+05	4.98E-05	1.05E+01
IL6R-003-65	IL6-001	1.85E+05	5.16E-07	9.55E-02
IL6R-201	IL6-001	1.38E+05	2.87E-05	3.97E+00
IL6R-201-65	IL6-001	8.88E+04	4.89E-08	4.34E-03
MMP1-001	CMA1-001	5.05E+06	4.19E-10	2.12E-03
MMP1-001	CMA1-001-151	3.02E+06	4.06E-09	1.23E-02
MMP1-001	CMA1-001-157	2.58E+07	2.26E-07	5.84E+00
MMP1-001	CMA1-001-183	1.00E+04	7.32E-10	7.32E-06
MMP1-001	CMA1-001-221	1.82E+07	1.51E-07	2.74E+00
MMP1-001	CMA1-001-226	1.96E+05	1.04E-08	2.05E-03
MMP1-001	CMA1-001-33	7.91E+05	8.43E-10	6.67E-04
MMP1-001	CMA1-001-38	1.19E-01	1.46E-09	1.74E-10
MMP1-001	CMA1-001-46	7.19E+05	6.88E-07	4.95E-01
MMP1-001	CMA1-001-48	6.72E+04	9.26E-07	6.23E-02
MMP1-001	CMA1-001-63	1.21E+06	7.55E-08	9.14E-02
MMP1-001	CMA1-001-66	3.64E+06	2.16E-08	7.85E-02
MMP1-001	CMA1-001-69	4.79E+08	2.74E-08	1.31E+01
MMP1-001	CMA1-001-98	3.26E+02	4.52E-11	1.47E-08
MMP1-001	CMA1-002	2.47E+05	8.35E-09	2.06E-03
MMP1-001	CMA1-002-1	3.93E+01	8.47E-12	3.33E-10
MMP1-001	CMA1-002-106	4.74E+03	6.82E-09	3.23E-05
MMP1-001	CMA1-002-110	7.73E-01	4.97E-14	3.84E-14
MMP1-001	CMA1-002-115	1.28E+08	4.10E-05	5.25E+03
MMP1-001	CMA1-002-124	5.75E+07	4.65E-07	2.68E+01
MMP1-001	CMA1-002-126	3.90E+07	4.72E-05	1.84E+03
MMP1-001	CMA1-002-40	1.82E+07	1.40E-07	2.55E+00
MMP1-001	CMA1-002-46	1.67E+06	1.14E-07	1.90E-01
MMP1-001	CMA1-002-48	3.22E+03	6.47E-09	2.08E-05
MMP1-001	CMA1-002-59	2.21E+04	6.16E-06	1.36E-01
MMP1-001	CMA1-002-72	1.57E+07	2.07E-08	3.25E-01
MMP1-001	CMA1-002-75	5.09E+06	1.60E-09	8.12E-03
MMP1-001	CMA1-002-77	3.92E+07	2.07E-06	8.12E+01
MMP1-001	CMA1-002-93	1.25E+07	5.15E-10	6.43E-03
MMP1-001	CMA1-002-95	5.53E+07	1.20E-08	6.61E-01
MMP1-001	TIMP1-001	1.40E+06	5.02E-06	7.02E+00
MMP1-001	TIMP1-001-10	8.70E+04	2.00E-07	1.74E-02
MMP1-001	TIMP1-001-102	5.75E+05	4.01E-07	2.31E-01
MMP1-001	TIMP1-001-105	4.84E+04	2.04E-06	9.89E-02
MMP1-001	TIMP1-001-136	4.38E+05	5.98E-07	2.62E-01
MMP1-001	TIMP1-001-171	1.53E+00	1.93E-11	2.95E-11
MMP1-001	TIMP1-001-185	1.69E+06	9.80E-06	1.66E+01
MMP1-001	TIMP1-001-192	1.97E+05	5.62E-07	1.11E-01
MMP1-001	TIMP1-001-203	3.85E-02	3.73E-11	1.43E-12
MMP1-001	TIMP1-001-33	1.88E+05	3.46E-08	6.50E-03

MMP1-001	TIMP1-001-50	1.74E+03	1.27E-11	2.20E-08
MMP1-001	TIMP1-001-51	1.92E+05	4.95E-09	9.49E-04
MMP1-001	TIMP1-001-76	1.40E+01	2.34E-10	3.27E-09
MMP1-001	TIMP1-001-77	2.71E+02	1.61E-06	4.37E-04
MMP1-001	TIMP1-002	2.19E+05	1.57E-05	3.44E+00
MMP1-001	TIMP1-002-10	9.59E+03	2.70E-06	2.59E-02
MMP1-001	TIMP1-002-102	1.59E+05	1.52E-05	2.41E+00
MMP1-001	TIMP1-002-105	3.73E+04	4.45E-06	1.66E-01
MMP1-001	TIMP1-002-115	7.68E+04	1.12E-05	8.62E-01
MMP1-001	TIMP1-002-116	5.66E+05	1.80E-06	1.02E+00
MMP1-001	TIMP1-002-33	4.09E+00	2.22E-12	9.08E-12
MMP1-001	TIMP1-002-50	6.70E-07	1.16E-12	7.76E-19
MMP1-001	TIMP1-002-51	2.59E+02	5.46E-12	1.42E-09
MMP1-001	TIMP1-002-76	4.52E+06	1.42E-05	6.40E+01
MMP1-001	TIMP1-002-77	5.65E+03	5.74E-06	3.24E-02
MMP1-001	TIMP1-003	6.87E+01	4.02E-12	2.76E-10
MMP1-001	TIMP1-003-107	1.68E+05	7.50E-05	1.26E+01
MMP1-001	TIMP1-003-12	3.36E+05	2.46E-06	8.27E-01
MMP1-001	TIMP1-003-121	2.02E+04	6.49E-09	1.31E-04
MMP1-001	TIMP1-003-128	2.48E+04	6.30E-07	1.56E-02
MMP1-001	TIMP1-003-13	7.56E+05	1.58E-05	1.19E+01
MMP1-001	TIMP1-003-139	5.11E+02	1.87E-11	9.56E-09
MMP1-001	TIMP1-003-38	1.92E+05	1.57E-05	3.00E+00
MMP1-001	TIMP1-003-41	3.02E+06	5.96E-07	1.80E+00
MMP1-001	TIMP1-003-72	4.09E+01	1.78E-12	7.29E-11
MMP1-001	TIMP2-001	8.87E+01	1.55E-13	1.37E-11
MMP1-001	TIMP2-003	1.75E+00	3.72E-15	6.51E-15
MMP1-001	TIMP2-003-130	2.50E+01	6.91E-08	1.73E-06
MMP1-001	TIMP2-003-139	4.36E+06	2.68E-11	1.17E-04
MMP1-001	TIMP2-008	1.75E+00	3.72E-15	6.51E-15
MMP1-001	TIMP2-008-130	8.49E+06	6.37E-11	5.41E-04
MMP1-001	TIMP2-008-139	3.00E+05	1.39E-11	4.16E-06
MMP1-001	TIMP2-008-42	1.07E+04	7.70E-07	8.24E-03
MMP1-001	TIMP2-008-54	7.49E+00	5.96E-11	4.46E-10
MMP1-001	TIMP2-008-69	3.91E+03	4.71E-10	1.84E-06
MMP1-001	TIMP3-001-142	7.42E+07	5.36E-05	3.98E+03
MMP1-001	TIMP3-001-186	4.17E+06	8.76E-05	3.65E+02
MMP1-001	TIMP3-001-22	1.21E+08	5.46E-05	6.60E+03
MMP1-001	TIMP4-001-109	1.65E+05	1.41E-06	2.33E-01
MMP1-001	TIMP4-001-141	4.89E+03	8.14E-07	3.98E-03
MMP1-001	TIMP4-001-19	1.68E+07	4.20E-05	7.05E+02
MMP1-001	TIMP4-001-206	3.57E+07	1.43E-04	5.09E+03
MMP1-001	TIMP4-001-214	2.28E+05	5.58E-12	1.27E-06
MMP1-001-374	CMA1-001	2.00E+06	1.28E-09	2.55E-03

MMP1-001-374	TIMP1-001	4.78E+05	5.32E-07	2.54E-01
MMP1-001-374	TIMP4-001	8.72E+04	1.02E-04	8.93E+00
MMP1-001-406	CMA1-001	9.47E+06	8.30E-09	7.86E-02
MMP1-001-406	TIMP1-001	8.68E+01	6.05E-15	5.25E-13
MMP1-001-406	TIMP3-001	1.73E+07	4.87E-06	8.42E+01
MMP13-001	TIMP1-001	2.56E+08	1.23E-08	3.14E+00
MMP13-001	TIMP1-001-10	3.47E+09	4.64E-07	1.61E+03
MMP13-001	TIMP1-001-102	1.87E+07	1.63E-06	3.04E+01
MMP13-001	TIMP1-001-105	5.05E+08	1.43E-07	7.21E+01
MMP13-001	TIMP1-001-136	4.42E+05	2.20E-08	9.74E-03
MMP13-001	TIMP1-001-171	1.35E+08	5.54E-12	7.48E-04
MMP13-001	TIMP1-001-185	1.55E+08	4.29E-07	6.65E+01
MMP13-001	TIMP1-001-192	6.14E+08	1.51E-05	9.30E+03
MMP13-001	TIMP1-001-203	5.64E+07	2.21E-04	1.25E+04
MMP13-001	TIMP1-001-33	3.29E+08	1.33E-06	4.36E+02
MMP13-001	TIMP1-001-50	1.18E+08	5.50E-07	6.49E+01
MMP13-001	TIMP1-001-51	7.48E+05	1.03E-06	7.73E-01
MMP13-001	TIMP1-001-76	6.02E+02	4.69E-13	2.83E-10
MMP13-001	TIMP1-001-77	6.97E+08	5.18E-08	3.61E+01
MMP13-001	TIMP1-002	7.65E+04	6.59E-08	5.04E-03
MMP13-001	TIMP1-002-10	1.23E+04	4.32E-10	5.31E-06
MMP13-001	TIMP1-002-102	6.71E+06	6.00E-09	4.03E-02
MMP13-001	TIMP1-002-105	1.08E+05	1.32E-09	1.42E-04
MMP13-001	TIMP1-002-115	1.50E+06	9.39E-10	1.41E-03
MMP13-001	TIMP1-002-116	6.17E+04	3.02E-08	1.86E-03
MMP13-001	TIMP1-002-33	4.10E+04	6.23E-08	2.55E-03
MMP13-001	TIMP1-002-50	4.14E+01	2.10E-08	8.71E-07
MMP13-001	TIMP1-002-51	3.22E+01	1.10E-11	3.55E-10
MMP13-001	TIMP1-002-76	2.17E+05	5.81E-08	1.26E-02
MMP13-001	TIMP1-002-77	6.25E+05	6.00E-08	3.75E-02
MMP13-001	TIMP1-003	1.41E+06	5.46E-09	7.70E-03
MMP13-001	TIMP1-003-107	5.40E+05	7.24E-08	3.91E-02
MMP13-001	TIMP1-003-12	4.92E+06	4.78E-05	2.35E+02
MMP13-001	TIMP1-003-121	2.13E+06	2.92E-06	6.22E+00
MMP13-001	TIMP1-003-128	1.45E+07	1.23E-08	1.79E-01
MMP13-001	TIMP1-003-13	2.95E+06	2.62E-08	7.73E-02
MMP13-001	TIMP1-003-139	3.26E+05	3.29E-05	1.07E+01
MMP13-001	TIMP1-003-38	1.52E+06	4.29E-05	6.52E+01
MMP13-001	TIMP1-003-41	3.06E+00	5.95E-13	1.82E-12
MMP13-001	TIMP1-003-72	1.84E+03	1.69E-11	3.11E-08
MMP13-001	TIMP2-001	1.29E+06	7.49E-09	9.66E-03
MMP13-001	TIMP2-001-119	3.72E+07	1.17E-09	4.35E-02
MMP13-001	TIMP2-003	2.53E+05	3.27E-07	8.27E-02
MMP13-001	TIMP2-003-130	2.41E+06	4.20E-05	1.01E+02

MMP13-001	TIMP2-003-139	4.33E-06	2.71E-11	1.17E-16
MMP13-001	TIMP2-008	1.19E+05	4.37E-06	5.20E-01
MMP13-001	TIMP2-008-130	7.51E+05	2.24E-05	1.68E+01
MMP13-001	TIMP2-008-139	4.33E-06	2.71E-11	1.17E-16
MMP13-001	TIMP2-008-42	4.41E+06	2.24E-05	9.86E+01
MMP13-001	TIMP2-008-54	1.13E+05	2.87E-07	3.24E-02
MMP13-001	TIMP2-008-69	9.28E+03	2.49E-06	2.31E-02
MMP13-001	TIMP3-001	5.51E+06	1.79E-04	9.86E+02
MMP13-001	TIMP3-001-142	1.63E+05	1.16E-05	1.89E+00
MMP13-001	TIMP3-001-201	3.02E+04	1.21E-05	3.66E-01
MMP13-001	TIMP3-001-22	1.27E+07	8.85E-05	1.12E+03
MMP13-001	TIMP3-001-37	3.09E+07	1.67E-04	5.17E+03
MMP13-001-158	TIMP1-001	3.99E+08	3.00E-06	1.20E+03
MMP13-001-158	TIMP2-001	1.93E+05	1.99E-05	3.85E+00
MMP13-001-158	TIMP3-001	1.52E+07	3.41E-05	5.18E+02
MMP13-002	TIMP1-001	1.43E+05	1.96E-09	2.80E-04
MMP13-002	TIMP2-001	2.85E+06	1.36E-05	3.88E+01
MMP13-002	TIMP3-001	1.06E+06	4.53E-05	4.80E+01
MMP13-002-158	TIMP1-001	1.09E+08	2.02E-07	2.20E+01
MMP13-002-158	TIMP2-001	2.16E+05	5.08E-06	1.10E+00
MMP13-201	TIMP1-001	1.30E+08	1.18E-11	1.54E-03
MMP13-201	TIMP2-001	8.82E+04	5.50E-06	4.85E-01
MMP13-201	TIMP3-001	1.32E+09	1.04E-07	1.38E+02
MMP13-201-158	TIMP1-001	1.42E+06	1.92E-08	2.73E-02
MMP13-201-158	TIMP2-001	1.12E+05	2.76E-05	3.09E+00
MMP13-201-158	TIMP3-001	4.42E+09	3.31E-09	1.46E+01
MMP2-001	TIMP1-001	1.38E+07	1.20E-08	1.65E-01
MMP2-001	TIMP1-001-10	5.72E+07	9.86E-11	5.64E-03
MMP2-001	TIMP1-001-102	7.20E+06	1.17E-12	8.39E-06
MMP2-001	TIMP1-001-105	1.17E+06	1.36E-09	1.59E-03
MMP2-001	TIMP1-001-136	4.18E+03	2.95E-13	1.23E-09
MMP2-001	TIMP1-001-171	1.92E+00	3.11E-17	5.98E-17
MMP2-001	TIMP1-001-185	6.82E+05	4.86E-14	3.32E-08
MMP2-001	TIMP1-001-192	5.02E+07	1.74E-09	8.74E-02
MMP2-001	TIMP1-001-203	1.64E+05	5.21E-14	8.55E-09
MMP2-001	TIMP1-001-33	1.49E+07	6.19E-10	9.22E-03
MMP2-001	TIMP1-001-50	1.92E+05	1.38E-14	2.65E-09
MMP2-001	TIMP1-001-51	9.59E+06	4.38E-12	4.20E-05
MMP2-001	TIMP1-001-76	1.99E+05	3.01E-12	5.99E-07
MMP2-001	TIMP1-001-77	2.55E+07	4.20E-08	1.07E+00
MMP2-001	TIMP1-002	2.39E+05	5.93E-06	1.42E+00
MMP2-001	TIMP1-002-10	3.00E+03	7.34E-11	2.20E-07
MMP2-001	TIMP1-002-102	9.45E+05	3.79E-07	3.58E-01
MMP2-001	TIMP1-002-105	2.39E+00	3.86E-13	9.22E-13

MMP2-001	TIMP1-002-115	1.94E+00	2.33E-11	4.51E-11
MMP2-001	TIMP1-002-116	1.93E+03	1.58E-10	3.04E-07
MMP2-001	TIMP1-002-33	4.96E+05	9.64E-07	4.78E-01
MMP2-001	TIMP1-002-50	5.45E+06	7.73E-09	4.21E-02
MMP2-001	TIMP1-002-51	5.20E+05	3.68E-10	1.91E-04
MMP2-001	TIMP1-002-76	2.48E+04	4.88E-09	1.21E-04
MMP2-001	TIMP1-002-77	1.96E+05	9.69E-08	1.90E-02
MMP2-001	TIMP1-003	1.47E+00	5.12E-11	7.53E-11
MMP2-001	TIMP1-003-107	7.17E+05	6.64E-09	4.76E-03
MMP2-001	TIMP1-003-12	3.84E+05	1.40E-08	5.36E-03
MMP2-001	TIMP1-003-121	9.91E+05	4.07E-10	4.03E-04
MMP2-001	TIMP1-003-128	2.41E+04	1.67E-09	4.02E-05
MMP2-001	TIMP1-003-13	2.91E-01	1.86E-11	5.42E-12
MMP2-001	TIMP1-003-139	5.82E+04	8.31E-11	4.84E-06
MMP2-001	TIMP1-003-38	3.81E+04	8.01E-09	3.05E-04
MMP2-001	TIMP1-003-41	6.71E+05	2.99E-09	2.01E-03
MMP2-001	TIMP1-003-72	6.10E+00	1.34E-12	8.19E-12
MMP2-001	TIMP2-001	5.45E+00	8.20E-15	4.47E-14
MMP2-001	TIMP2-001-119	3.77E+05	1.20E-05	4.53E+00
MMP2-001	TIMP2-001-146	4.75E+05	4.96E-06	2.36E+00
MMP2-001	TIMP2-003	2.42E+06	3.94E-05	9.54E+01
MMP2-001	TIMP2-003-130	6.14E+06	3.93E-06	2.41E+01
MMP2-001	TIMP2-003-139	9.41E+04	7.78E-05	7.32E+00
MMP2-001	TIMP2-008	2.09E+04	6.51E-05	1.36E+00
MMP2-001	TIMP2-008-130	7.27E+04	8.58E-06	6.24E-01
MMP2-001	TIMP2-008-42	5.92E+04	1.22E-04	7.24E+00
MMP2-001	TIMP2-008-54	2.55E+06	4.85E-05	1.24E+02
MMP2-001	TIMP2-008-69	1.41E+04	5.90E-06	8.32E-02
MMP2-001	TIMP3-001	8.47E-01	3.31E-07	2.81E-07
MMP2-001	TIMP3-001-121	8.33E+07	4.61E-06	3.84E+02
MMP2-001	TIMP3-001-186	3.11E+00	3.46E-06	1.07E-05
MMP2-001	TIMP4-001	4.28E+06	9.00E-05	3.85E+02
MMP2-001	TIMP4-001-103	1.78E+04	4.20E-06	7.47E-02
MMP2-001	TIMP4-001-109	4.55E+05	2.52E-07	1.15E-01
MMP2-001	TIMP4-001-141	5.81E+05	2.44E-09	1.42E-03
MMP2-001	TIMP4-001-143	1.23E+05	4.95E-07	6.08E-02
MMP2-001	TIMP4-001-19	7.12E+04	7.06E-06	5.03E-01
MMP2-001	TIMP4-001-206	5.15E+04	7.77E-06	4.00E-01
MMP2-001	TIMP4-001-214	7.34E+05	9.06E-09	6.65E-03
MMP2-001	TIMP4-001-24	6.64E+05	1.06E-04	7.01E+01
MMP2-001-447	TIMP1-001	9.50E+07	7.92E-08	7.52E+00
MMP2-001-447	TIMP2-001	4.01E+03	7.67E-06	3.08E-02
MMP2-001-447	TIMP3-001	1.40E+05	7.34E-06	1.03E+00
MMP2-008	TIMP1-001	1.06E+07	7.18E-11	7.61E-04

MMP2-008	TIMP2-001	1.96E+06	7.11E-06	1.39E+01
MMP2-008-333	TIMP1-001	7.59E+08	1.43E-08	1.08E+01
MMP2-008-333	TIMP2-001	1.11E+05	8.76E-07	9.72E-02
MMP2-008-333	TIMP3-001	2.20E+02	3.14E-06	6.91E-04
MMP2-008-371	TIMP1-001	2.25E+07	1.67E-09	3.76E-02
MMP2-008-371	TIMP2-001	1.20E+05	1.06E-06	1.28E-01
MMP2-008-371	TIMP3-001	8.31E+01	1.90E-06	1.58E-04
MMP2-008-545	TIMP1-001	1.79E+06	1.15E-10	2.06E-04
MMP2-008-545	TIMP2-001	1.11E+07	1.74E-06	1.93E+01
MMP3-001	TIMP1-001	7.79E+04	2.82E-06	2.20E-01
MMP3-001	TIMP1-001-10	9.99E+03	4.62E-08	4.62E-04
MMP3-001	TIMP1-001-102	7.00E-04	1.00E-09	7.02E-13
MMP3-001	TIMP1-001-105	5.61E+01	2.60E-12	1.46E-10
MMP3-001	TIMP1-001-136	2.82E+05	7.60E-05	2.14E+01
MMP3-001	TIMP1-001-171	2.37E+06	5.39E-06	1.28E+01
MMP3-001	TIMP1-001-185	5.41E+05	3.49E-06	1.89E+00
MMP3-001	TIMP1-001-192	6.66E+04	3.08E-06	2.05E-01
MMP3-001	TIMP1-001-203	4.33E-03	6.37E-14	2.76E-16
MMP3-001	TIMP1-001-33	5.51E+05	5.11E-06	2.82E+00
MMP3-001	TIMP1-001-50	3.06E+04	2.12E-06	6.50E-02
MMP3-001	TIMP1-001-51	3.67E+02	1.96E-14	7.19E-12
MMP3-001	TIMP1-001-76	3.36E+05	2.32E-05	7.80E+00
MMP3-001	TIMP1-001-77	2.51E+04	4.19E-06	1.05E-01
MMP3-001	TIMP1-002	3.58E+06	3.37E-05	1.20E+02
MMP3-001	TIMP1-002-10	2.79E+05	4.05E-06	1.13E+00
MMP3-001	TIMP1-002-102	8.70E+04	1.35E-07	1.18E-02
MMP3-001	TIMP1-002-105	2.95E+05	2.17E-05	6.40E+00
MMP3-001	TIMP1-002-115	1.40E+06	4.00E-05	5.60E+01
MMP3-001	TIMP1-002-116	7.70E+03	8.32E-12	6.41E-08
MMP3-001	TIMP1-002-33	1.61E+05	2.17E-05	3.49E+00
MMP3-001	TIMP1-002-50	1.46E+05	2.56E-06	3.74E-01
MMP3-001	TIMP1-002-51	5.20E+05	7.33E-05	3.81E+01
MMP3-001	TIMP1-002-76	4.13E+04	2.49E-05	1.03E+00
MMP3-001	TIMP1-002-77	1.78E+05	1.11E-05	1.97E+00
MMP3-001	TIMP1-003	2.43E+07	1.95E-06	4.74E+01
MMP3-001	TIMP1-003-107	8.66E+05	7.66E-06	6.64E+00
MMP3-001	TIMP1-003-12	1.18E+05	1.40E-05	1.65E+00
MMP3-001	TIMP1-003-121	1.47E+05	4.55E-05	6.69E+00
MMP3-001	TIMP1-003-128	7.65E+05	2.27E-07	1.74E-01
MMP3-001	TIMP1-003-13	1.93E+07	1.43E-05	2.75E+02
MMP3-001	TIMP1-003-139	8.99E+04	5.47E-07	4.92E-02
MMP3-001	TIMP1-003-38	8.45E+06	2.94E-07	2.49E+00
MMP3-001	TIMP1-003-41	3.20E+05	3.02E-05	9.65E+00
MMP3-001	TIMP1-003-72	9.46E+03	1.62E-07	1.53E-03

MMP3-001	TIMP2-001	2.00E+04	1.36E-10	2.72E-06
MMP3-001	TIMP2-003	1.49E+04	3.52E-06	5.25E-02
MMP3-001	TIMP2-003-130	5.11E+05	9.96E-07	5.09E-01
MMP3-001	TIMP2-003-139	4.82E+06	2.69E-07	1.29E+00
MMP3-001	TIMP2-008	7.70E+06	1.30E-05	1.00E+02
MMP3-001	TIMP2-008-130	2.10E+05	1.05E-06	2.20E-01
MMP3-001	TIMP2-008-139	1.07E+06	6.13E-08	6.56E-02
MMP3-001	TIMP2-008-42	2.33E+06	8.33E-09	1.94E-02
MMP3-001	TIMP2-008-54	1.51E+06	3.30E-10	4.99E-04
MMP3-001	TIMP2-008-69	3.18E+05	4.56E-08	1.45E-02
MMP3-001	TIMP3-001	8.89E+05	2.11E-04	1.87E+02
MMP3-001	TIMP3-001-121	1.28E+07	4.82E-05	6.17E+02
MMP3-001	TIMP3-001-142	1.24E+05	1.50E-04	1.86E+01
MMP3-001	TIMP3-001-186	3.05E+05	1.03E-04	3.14E+01
MMP3-001	TIMP3-001-196	8.54E+09	1.90E-04	1.62E+06
MMP3-001	TIMP3-001-22	7.21E+06	3.03E-05	2.18E+02
MMP3-001	TIMP3-001-37	9.50E+06	2.02E-05	1.91E+02
MMP3-001	TIMP4-001	4.97E+04	2.02E-05	1.00E+00
MMP3-001	TIMP4-001-141	1.19E+03	1.43E-11	1.70E-08
MMP3-001	TIMP4-001-214	2.15E-08	1.99E-15	4.27E-23
MMP3-001	TIMP4-001-24	1.07E+07	3.01E-04	3.22E+03
MMP3-001	TIMP4-001-28	7.71E+08	1.21E-04	9.33E+04
MMP3-001-45	TIMP1-001	6.05E+07	3.58E-08	2.16E+00
MMP3-001-45	TIMP3-001	5.61E+10	3.41E-04	1.91E+07
MMP3-001-45	TIMP4-001	2.20E+07	2.19E-04	4.82E+03
MMP3-001-96	TIMP1-001	2.41E+05	1.91E-06	4.61E-01
MMP9-001	TIMP1-001	7.20E+01	6.14E-12	4.42E-10
MMP9-001	TIMP1-001-10	7.49E+02	7.56E-11	5.66E-08
MMP9-001	TIMP1-001-102	8.52E+05	1.77E-04	1.50E+02
MMP9-001	TIMP1-001-105	2.81E+06	4.72E-06	1.33E+01
MMP9-001	TIMP1-001-136	1.25E+06	1.16E-06	1.45E+00
MMP9-001	TIMP1-001-171	4.17E+00	8.05E-12	3.36E-11
MMP9-001	TIMP1-001-185	1.45E+05	2.41E-09	3.49E-04
MMP9-001	TIMP1-001-192	5.34E+05	1.82E-05	9.74E+00
MMP9-001	TIMP1-001-203	1.89E+00	1.92E-14	3.62E-14
MMP9-001	TIMP1-001-33	2.72E+00	9.20E-12	2.50E-11
MMP9-001	TIMP1-001-50	1.26E+04	8.38E-12	1.06E-07
MMP9-001	TIMP1-001-51	2.07E+03	5.37E-11	1.11E-07
MMP9-001	TIMP1-001-76	1.84E+05	1.24E-07	2.28E-02
MMP9-001	TIMP1-001-77	2.27E+05	3.87E-05	8.78E+00
MMP9-001	TIMP1-002	6.56E+05	7.85E-08	5.15E-02
MMP9-001	TIMP1-002-10	5.25E+06	1.07E-07	5.63E-01
MMP9-001	TIMP1-002-102	3.00E+04	3.54E-05	1.06E+00
MMP9-001	TIMP1-002-105	4.12E+06	2.59E-05	1.07E+02

MMP9-001	TIMP1-002-115	1.21E+04	6.00E-05	7.26E-01
MMP9-001	TIMP1-002-116	1.14E+06	1.30E-04	1.48E+02
MMP9-001	TIMP1-002-33	3.41E+03	2.59E-09	8.83E-06
MMP9-001	TIMP1-002-50	1.08E+06	1.21E-04	1.30E+02
MMP9-001	TIMP1-002-51	2.37E+01	5.28E-11	1.25E-09
MMP9-001	TIMP1-002-76	2.35E+03	1.07E-07	2.50E-04
MMP9-001	TIMP1-002-77	4.51E+03	5.13E-09	2.31E-05
MMP9-001	TIMP1-003	1.16E+06	2.18E-05	2.52E+01
MMP9-001	TIMP1-003-107	5.63E+03	4.54E-12	2.55E-08
MMP9-001	TIMP1-003-12	3.88E+04	4.87E-08	1.89E-03
MMP9-001	TIMP1-003-121	1.08E+03	4.68E-12	5.06E-09
MMP9-001	TIMP1-003-128	7.04E+03	8.05E-09	5.66E-05
MMP9-001	TIMP1-003-13	3.88E+01	1.80E-09	6.97E-08
MMP9-001	TIMP1-003-139	1.46E+03	1.68E-09	2.45E-06
MMP9-001	TIMP1-003-38	1.64E+05	9.73E-10	1.60E-04
MMP9-001	TIMP1-003-41	2.69E+04	1.77E-07	4.77E-03
MMP9-001	TIMP1-003-72	1.97E+02	1.09E-08	2.15E-06
MMP9-001	TIMP2-001	3.74E+01	0.00E+00	0.00E+00
MMP9-001	TIMP2-003	4.45E+03	3.35E-10	1.49E-06
MMP9-001	TIMP2-003-130	6.09E+04	1.40E-08	8.55E-04
MMP9-001	TIMP2-003-139	8.54E+04	2.97E-10	2.54E-05
MMP9-001	TIMP2-008	4.45E+03	3.35E-10	1.49E-06
MMP9-001	TIMP2-008-130	1.78E-04	2.97E-15	5.28E-19
MMP9-001	TIMP2-008-139	8.54E+04	2.97E-10	2.54E-05
MMP9-001	TIMP2-008-42	1.59E+07	1.32E-05	2.10E+02
MMP9-001	TIMP2-008-54	1.81E+05	1.27E-06	2.30E-01
MMP9-001	TIMP2-008-69	4.20E+04	1.30E-09	5.45E-05
MMP9-001	TIMP3-001	1.10E+02	5.35E-11	5.89E-09
MMP9-001	TIMP3-001-121	3.02E+05	1.44E-09	4.33E-04
MMP9-001	TIMP3-001-142	5.30E+01	1.32E-07	7.01E-06
MMP9-001	TIMP3-001-186	3.49E+05	8.31E-09	2.90E-03
MMP9-001	TIMP3-001-196	4.91E+05	1.45E-08	7.11E-03
MMP9-001	TIMP3-001-201	1.66E+05	2.19E-12	3.63E-07
MMP9-001	TIMP3-001-22	3.18E+04	3.06E-09	9.72E-05
MMP9-001	TIMP3-001-37	1.95E+04	3.35E-09	6.52E-05
MMP9-001	TIMP4-001	3.84E+06	4.62E-04	1.77E+03
MMP9-001	TIMP4-001-103	2.20E+05	8.19E-12	1.80E-06
MMP9-001	TIMP4-001-109	2.30E+05	2.46E-07	5.67E-02
MMP9-001	TIMP4-001-112	2.32E+03	1.35E-08	3.13E-05
MMP9-001	TIMP4-001-141	2.37E+03	2.71E-13	6.43E-10
MMP9-001	TIMP4-001-143	1.54E+04	1.32E-09	2.03E-05
MMP9-001	TIMP4-001-19	1.07E+06	2.92E-05	3.12E+01
MMP9-001	TIMP4-001-206	3.74E+03	1.11E-07	4.14E-04
MMP9-001	TIMP4-001-214	6.31E+02	9.38E-09	5.92E-06

MMP9-001	TIMP4-001-24	3.02E+03	3.98E-08	1.20E-04
MMP9-001	TIMP4-001-28	8.71E+03	1.91E-08	1.67E-04
MMP9-001-20	TIMP1-001	2.34E+01	6.67E-12	1.56E-10
MMP9-001-20	TIMP3-001	2.74E+01	2.68E-10	7.35E-09
MMP9-001-20	TIMP4-001	7.86E+02	2.08E-09	1.64E-06
MMP9-001-239	TIMP1-001	2.42E+03	2.71E-13	6.55E-10
MMP9-001-239	TIMP3-001	9.20E+03	3.40E-09	3.13E-05
MMP9-001-239	TIMP4-001	8.42E+01	6.09E-10	5.12E-08
MMP9-001-279	TIMP1-001	5.95E+05	1.84E-04	1.10E+02
MMP9-001-279	TIMP2-001	3.98E+01	0.00E+00	0.00E+00
MMP9-001-279	TIMP3-001	8.88E+07	8.70E-08	7.72E+00
MMP9-001-279	TIMP4-001	8.55E+02	4.94E-09	4.22E-06
MMP9-001-38	TIMP1-001	1.92E+06	1.32E-05	2.53E+01
MMP9-001-38	TIMP3-001	1.80E+06	6.26E-07	1.13E+00
MMP9-001-38	TIMP4-001	3.62E+02	3.70E-09	1.34E-06
MMP9-001-574	TIMP1-001	5.32E+02	1.19E-08	6.30E-06
MMP9-001-574	TIMP3-001	2.25E+03	2.15E-08	4.83E-05
MMP9-001-574	TIMP4-001	2.25E-03	1.12E-10	2.52E-13
MMP9-001-668	TIMP1-001	1.05E+06	5.86E-05	6.15E+01
MMP9-001-668	TIMP3-001	5.16E-05	3.34E-10	1.72E-14
MMP9-001-668	TIMP4-001	2.68E-04	2.95E-12	7.91E-16
PDGFRA-001-761	PDGFA-001	3.99E+04	1.23E-14	4.91E-10
PDGFRA-002	PDGFA-001	1.83E+07	7.28E-07	1.33E+01
PDGFRB-001	PDGFB-002-204	7.23E+08	6.75E-07	4.88E+02
TGFBR1-001	TGFB1-001	5.89E+04	1.62E-10	9.51E-06
TGFBR1-001	TGFB1-001-10	1.81E+04	1.13E-08	2.05E-04
TGFBR1-001	TGFB1-001-25	8.20E-02	8.51E-12	6.98E-13
TGFBR1-001	TGFB1-001-263	2.14E+03	2.90E-10	6.20E-07
TGFBR1-001-17	TGFB1-001	5.01E+04	1.62E-10	8.09E-06
TGFBR1-001-19	TGFB1-001	9.31E+04	1.62E-10	1.50E-05
TGFBR1-003	TGFB1-001	1.81E+05	8.27E-10	1.50E-04
TGFBR2-001	TGFB1-001	2.21E+03	2.20E-08	4.87E-05
TGFBR2-001	TGFB1-001-10	2.17E+05	7.39E-08	1.60E-02
TIMP3-001	ADAM17-001	5.97E+05	2.59E-06	1.54E+00
TIMP3-001	ADAM17-001-202	5.85E-105	3.47E-06	2.03E-110
TIMP3-001-121	ADAM17-001	2.53E+03	5.07E-06	1.28E-02
TIMP3-001-142	ADAM17-001	8.63E+03	1.17E-07	1.01E-03
TIMP3-001-186	ADAM17-001	4.82E+06	3.02E-07	1.46E+00
TIMP3-001-196	ADAM17-001	1.51E+05	1.47E-07	2.21E-02
TIMP3-001-201	ADAM17-001	3.15E+05	5.18E-09	1.63E-03
TIMP3-001-22	ADAM17-001	1.96E+08	6.92E-07	1.36E+02
TIMP3-001-37	ADAM17-001	2.47E+05	5.78E-07	1.43E-01
TNFSFR1B-001	TNF-001	8.94E+04	2.76E-08	2.46E-03
TNFSFR1B-001	TNF-001-84	1.13E+04	8.96E-11	1.01E-06

TNFSFR1B-001-187	TNF-001	5.19E+02	6.97E-09	3.62E-06
TNFSFR1B-001-203	TNF-001	4.44E+06	2.92E-05	1.30E+02
TNFSFR1B-001-232	TNF-001	1.47E+06	3.21E-05	4.72E+01
TNFSFR1B-001-264	TNF-001	1.04E+05	9.82E-08	1.02E-02

Appendix 6

Model reparameterisations as described in Chapter 6

ReceptorName	LigandName	kcatWT	kcatMut	AScore
CCR2-001	CCL2-001	0.0010513 9000	0.0010513 9000	
CCR2-001	CCL2-001-69	0.0010513 9000	3.1198e-05 267.0968	0.03
CCR2-001	CCL2-001-71	0.0010513 9000	0.00014356 1229.0323	0.33
CCR2-001	CCL2-004	0.0010513 9000	0.00011417 977.4177	0.26
CCR2-001-355	CCL2-001	0.0010513 9000	1.543e-05 132.0968	0.02
CCR2-001-64	CCL2-001	0.0010513 9000	0.00016221 1388.7097	0.38
CCR2-002	CCL2-001	0.0010513 9000	0.00010569 904.8387	0.24
CCR2-002-64	CCL2-001	0.0010513 9000	2.2155e-05 189.6774	0.03
CCR2-201	CCL2-001	0.0010513 9000	0.00047306 4050	0.73
CCR2-201-355	CCL2-001	0.0010513 9000	0.00039846 3411.2904	0.69
CCR2-201-64	CCL2-001	0.0010513 9000	0.00019612 1679.0323	0.45
CCR5-001	CCL5-001	1000	1000	1.00
CCR5-001	CCL5-001-40	1000	1.8013	0.10
CCR5-001	CCL5-001-5	1000	71548.8215	5.93
CCR5-001	CCL5-001-56	1000	333.3333	0.72
CCR5-001	CCL5-001-67	1000	1818.1818	1.11
CCR5-001	CCL5-001-68	1000	0.94949	0.09
CCR5-001	CCL5-002	1000	1000	1.00
CCR5-001	CCL5-002-40	1000	1.8013	0.10
CCR5-001	CCL5-002-5	1000	71548.8215	5.93
CCR5-001	CCL5-002-56	1000	333.3333	0.72
CCR5-001	CCL5-002-67	1000	1818.1818	1.11
CCR5-001	CCL5-002-68	1000	0.94949	0.09
CCR5-001-223	CCL5-001	1000	2878.7879	1.20
CCR5-001-335	CCL5-001	1000	577.4411	0.86
CCR5-001-55	CCL5-001	1000	890.5724	0.97
CSF1R-001	CSF1-001	9.94E-02	9.94E-02	
CSF1R-001	CSF1-001-292	9.94E-02	1.21E+00	1.00
CSF1R-001	CSF1-001-408	9.94E-02	1.57E-01	1.00

CSF1R-001	CSF1-001-438	9.94E-02	9.72E-07	1.00
CSF1R-001	CSF1-001-489	9.94E-02	6.20E-05	1.00
CSF1R-001	CSF1-002-292	9.94E-02	2.31E-01	1.00
CSF1R-001	CSF1-002-408	9.94E-02	3.72E-06	1.00
CSF1R-001	CSF1-002-461	9.94E-02	1.46E-10	0.03
CSF1R-001	CSF1-002-489	9.94E-02	5.57E-01	1.00
CSF1R-001	CSF1-002-523	9.94E-02	9.79E-09	1.00
CSF1R-001	CSF1-003	9.94E-02	1.22E-06	1.00
CSF1R-001	CSF1-003-292	9.94E-02	1.18E-01	1.00
CSF1R-001	CSF1-003-373	9.94E-02	2.04E-38	0.03
CSF1R-001	CSF1-004	9.94E-02	1.20E-06	1.00
CSF1R-001-245	CSF1-001	9.94E-02	2.76E-198	0.03
CSF1R-001-362	CSF1-001	9.94E-02	5.31E-06	1.00
CSF1R-201	CSF1-001	9.94E-02	1.35E-05	1.00
CSF1R-201-153	CSF1-001	9.94E-02	6.74E-01	1.00
CSF1R-201-241	CSF1-001	9.94E-02	2.31E-05	1.00
CSF1R-201-245	CSF1-001	9.94E-02	3.59E-01	1.00
CSF1R-201-268	CSF1-001	9.94E-02	1.34E-03	1.00
CSF1R-201-279	CSF1-001	9.94E-02	3.62E-150	0.03
CSF1R-201-298	CSF1-001	9.94E-02	1.39E-05	1.00
CSF1R-201-32	CSF1-001	9.94E-02	1.86E-06	1.00
CSF1R-201-60	CSF1-001	9.94E-02	5.58E-07	1.00
CXCR3-001	CXCL10-001	1000	1000	
CXCR3-001	CXCL10-001-18	1000	1974.2857	1.13
CXCR3-001	CXCL10-001-29	1000	434.2857	0.78
CXCR3-001	CXCL10-001-33	1000	19.8	0.21
CXCR3-001	CXCL10-001-58	1000	0.036857	0.09
CXCR3-001	CXCL10-001-66	1000	1.84E-05	0.09
CXCR3-001	CXCL10-001-68	1000	2.0371	0.10
CXCR3-001	CXCL10-001-80	1000	1405.7143	1.06
CXCR3-001	CXCL10-001-96	1000	39.7143	0.30
CXCR3-002	CXCL10-001	1000	291.4286	0.68
CXCR3-001	CXCL11-001	1000	1000	
CXCR3-001	CXCL11-001-29	1000	178.6982	0.58
CXCR3-001	CXCL11-001-72	1000	16272.1893	2.03
CXCR3-001	CXCL11-001-73	1000	331360.9465	27.27
CXCR3-001	CXCL11-003	1000	223372.7811	18.13
CXCR3-001	CXCL11-003-29	1000	807692.3075	68.95
CXCR3-001	CXCL11-003-72	1000	15591.716	1.99
CXCR3-001	CXCL11-003-73	1000	124.8521	0.51
CXCR3-002	CXCL11-001	1000	5355.0296	1.32
CXCR3-001	CXCL9-001	1000	1000	1.00
CXCR3-001	CXCL9-001-1	1000	1818.6813	1.11
CXCR3-001	CXCL9-001-101	1000	107142.8572	8.67

CXCR3-001	CXCL9-001-125	1000	2329.6703	1.18
CXCR3-001	CXCL9-001-40	1000	15164.8352	1.96
CXCR3-001	CXCL9-001-71	1000	928.5714	0.98
CXCR3-002	CXCL9-001	1000	1543956.044	134.63
EGFR-001	EGF-001	1938	1938	
EGFR-001	EGF-001-151	1938	349748437.5	0.50
EGFR-001	EGF-002-389	1938	12324.4687	0.70
EGFR-001-521	EGF-001	1938	31341.0937	0.60
EGFR-002	EGF-001	1938	20591.25	0.64
EGFR-002-521	EGF-001	1938	3.23E-75	1.19
EGFR-002-640	EGF-001	1938	4708.7344	0.86
EGFR-002-703	EGF-001	1938	7752	0.77
EGFR-003	EGF-001	1938	11506.875	0.71
EGFR-003-521	EGF-001	1938	1938	1.00
EGFR-004-476	EGF-001	1938	51932343.75	0.50
EGFR-005	EGF-001	1938	1.13E-144	1.19
EGFR-201	EGF-001	1938	14141.3437	0.68
EGFR-201-521	EGF-001	1938	334.6078	1.15
EGFR-202	EGF-001	1938	7812.5625	0.77
IFNGR1-001	IFNG-001	16200 27360 10000 4400	16200 27360 10000 4400	1.00
IFNGR1-001	IFNG-001-160	16200 27360 10000 4400	7305449.423 12338092.6431 4509536.7847 1984196.1853	1.06
IFNGR1-001	IFNG-001-162	16200 27360 10000 4400	2405722.1454 4062997.2752 1485013.624 653405.9946	1.06
IFNGR1-001	IFNG-001-72	16200 27360 10000 4400	101525.8895 171465.9401 62670.2997 27574.9319	1.03
IFNGR1-001-14	IFNG-001	16200 27360 10000 4400	23836.5131 40257.2207 14713.8965 6474.1144	1.01
IFNGR1-001-180	IFNG-001	16200 27360 10000 4400	16.6635 28.1428 10.2861 4.5259	0.12
IFNGR1-001-335	IFNG-001	16200 27360 10000 4400	812207.1172 1371727.5204 501362.3978 220599.455	1.05

IFNGR1-001-467	IFNG-001	16200 27360 10000 4400	22291.5534 37647.9564 13760.218 6054.4959	1.01
IFNGR1-201	IFNG-001	16200 27360 10000 4400	794550.433 1341907.3569 490463.2153 215803.8147	1.05
IFNGR1-201-14	IFNG-001	16200 27360 10000 4400	805584.6789 1360544.9591 497275.2044 218801.0899	1.05
IFNGR1-201-180	IFNG-001	16200 27360 10000 4400	73716.6242 124499.1826 45504.0872 20021.7984	1.02
IFNGR1-201-46	IFNG-001	16200 27360 10000 4400	1.8606 3.1423 1.1485 0.50534	0.12
IFNGR1-201-61	IFNG-001	16200 27360 10000 4400	0.32444 0.54795 0.20027 0.08812	0.11
IFNGR2-001	IFNG-001	16200 27360 10000 4400	16200 27360 10000 4400	
IFNGR2-001	IFNG-001-160	16200 27360 10000 4400	1201458.9737 2029130.6991 741641.3374 326322.1884	1.05
IFNGR2-001	IFNG-001-162	16200 27360 10000 4400	2870699101.1248 4848291793.3131 1772036474.1641 779696048.6322	1.06
IFNGR2-001	IFNG-001-72	16200 27360 10000 4400	181203.6486 306032.8267 111854.1033 49215.8055	1.04
IFNGR2-001-64	IFNG-001	16200 27360 10000 4400	2309361.7167 3900255.3191 1425531.9149 627234.0426	1.06
IL10RA-001	IL10-001	8.1882 0.644 0.00625	8.1882 0.644 0.00625	
IL10RA-001	IL10-001-15	8.1882 0.644 0.00625	88.5882 6.9674 0.067619	0.84

IL10RA-001	IL10-001-169	8.1882 0.644 0.00625	6.7145 0.52809 0.0051251	1.00
IL10RA-001	IL10-001-19	8.1882 0.644 0.00625	17.3037 1.3609 0.013208	1.00
IL10RA-001	IL10-001-20	8.1882 0.644 0.00625	47.6058 3.7442 0.036337	0.98
IL10RA-001	IL10-001-45	8.1882 0.644 0.00625	0.67228 0.052874 0.00051314	1.00
IL10RA-001	IL10-001-71	8.1882 0.644 0.00625	8.6932 0.68372 0.0066355	1.00
IL10RA-001	IL10-001-72	8.1882 0.644 0.00625	12.4189 0.97674 0.0094793	1.00
IL10RA-001-159	IL10-001	8.1882 0.644 0.00625	32.1236 2.5265 0.02452	1.00
IL10RA-001-224	IL10-001	8.1882 0.644 0.00625	3.8085 0.29953 0.002907	1.00
IL10RA-001-233	IL10-001	8.1882 0.644 0.00625	1.4489e-05 1.1395e-06 1.1059e-08	1.00
IL10RA-001-351	IL10-001	8.1882 0.644 0.00625	7.5341 0.59256 0.0057508	1.00
IL10RA-001-420	IL10-001	8.1882 0.644 0.00625	1.2005e-20 9.4419e-22 9.1633e-24	1.00
IL12RB1-001	IL12A-001	8.10E+00	8.10E+00	
IL12RB1-001	IL12A-001-211	8.10E+00	2.70E+03	0.94
IL12RB1-001	IL12A-001-82	8.10E+00	1.22E+02	0.95
IL12RB1-001	IL12A-005	8.10E+00	4.18E+02	0.94
IL12RB1-001	IL12A-005-115	8.10E+00	1.93E+02	0.94
IL12RB1-001	IL12A-005-153	8.10E+00	8.05E+03	0.93
IL12RB1-001	IL12A-005-164	8.10E+00	2.36E+02	0.94
IL12RB1-001	IL12A-005-200	8.10E+00	2.32E+05	0.93
IL12RB1-001	IL12A-005-3	8.10E+00	3.51E+03	0.94
IL12RB1-001	IL12A-005-81	8.10E+00	2.78E+05	0.93
IL12RB1-001	IL12A-006	8.10E+00	4.74E+00	0.97
IL12RB1-001	IL12A-006-139	8.10E+00	4.61E+04	0.93
IL12RB1-001	IL12A-006-157	8.10E+00	7.74E+03	0.93
IL12RB1-001	IL12A-006-177	8.10E+00	7.32E+02	0.94
IL12RB1-001	IL12A-006-224	8.10E+00	2.12E+02	0.94
IL12RB1-001	IL12A-006-28	8.10E+00	1.80E-07	0.15

IL12RB1-001	IL12A-006-3	8.10E+00	3.71E+02	0.94
IL12RB1-001	IL12A-006-81	8.10E+00	7.43E+02	0.94
IL12RB1-001-156	IL12A-001	8.10E+00	2.62E-01	0.44
IL12RB1-001-214	IL12A-001	8.10E+00	2.14E+01	0.99
IL12RB1-001-352	IL12A-001	8.10E+00	4.41E-37	0.15
IL12RB1-001-365	IL12A-001	8.10E+00	3.94E-11	0.15
IL12RB1-001-378	IL12A-001	8.10E+00	1.59E+01	1.00
IL12RB1-001-91	IL12A-001	8.10E+00	4.41E-17	0.15
IL12RB1-002-156	IL12A-001	8.10E+00	2.62E-01	0.44
IL12RB1-002-214	IL12A-001	8.10E+00	5.88E-18	0.15
IL12RB1-002-378	IL12A-001	8.10E+00	1.59E+01	1.00
IL12RB1-002-47	IL12A-001	8.10E+00	1.82E-24	0.15
IL12RB1-002-91	IL12A-001	8.10E+00	4.41E-17	0.15
IL12RB1-003-201	IL12A-001	8.10E+00	4.67E+04	0.93
IL18R1-001	IL18-001	8.10E+00	8.10E+00	
IL18R1-001	IL18-001-127	8.10E+00	2.67E-05	0.15
IL18R1-001	IL18-001-164	8.10E+00	2.90E-02	0.19
IL18R1-001	IL18-001-22	8.10E+00	6.35E+02	0.94
IL18R1-001	IL18-001-47	8.10E+00	8.90E+00	1.00
IL18R1-001	IL18-001-63	8.10E+00	5.63E+02	0.94
IL18R1-001	IL18-003	8.10E+00	1.14E+03	0.94
IL18R1-001	IL18-003-123	8.10E+00	2.06E+03	0.94
IL18R1-001	IL18-003-160	8.10E+00	1.70E+01	0.99
IL18R1-001	IL18-003-22	8.10E+00	4.01E-01	0.52
IL18R1-001	IL18-003-43	8.10E+00	9.97E+00	1.00
IL18R1-001	IL18-003-59	8.10E+00	2.92E+03	0.94
IL18R1-001	IL18-006	8.10E+00	8.10E+00	
IL18R1-001	IL18-006-127	8.10E+00	2.67E-05	0.15
IL18R1-001	IL18-006-164	8.10E+00	2.90E-02	0.19
IL18R1-001	IL18-006-22	8.10E+00	6.35E+02	0.94
IL18R1-001	IL18-006-47	8.10E+00	8.90E+00	1.00
IL18R1-001	IL18-006-63	8.10E+00	5.63E+02	0.94
IL18R1-001-170	IL18-001	8.10E+00	7.42E+02	0.94
IL18R1-001-232	IL18-001	8.10E+00	2.45E+05	0.93
IL18R1-001-423	IL18-001	8.10E+00	2.04E+02	0.94
IL18R1-201	IL18-001	8.10E+00	8.10E+00	
IL18R1-201-170	IL18-001	8.10E+00	7.42E+02	0.94
IL18R1-201-232	IL18-001	8.10E+00	2.45E+05	0.93
IL18R1-201-423	IL18-001	8.10E+00	2.04E+02	0.94
IL18R1-202	IL18-001	8.10E+00	1.13E+04	0.93
IL18R1-202-100	IL18-001	8.10E+00	5.94E-03	0.16
IL18R1-202-117	IL18-001	8.10E+00	1.10E+05	0.93
IL18R1-202-139	IL18-001	8.10E+00	6.66E+03	0.93
IL18R1-202-162	IL18-001	8.10E+00	2.82E+04	0.93

IL18R1-202-53	IL18-001	8.10E+00	8.41E+03	0.93
IL18R1-202-84	IL18-001	8.10E+00	3.89E+03	0.94
IL1R1-001	IL1B-001	1.35E-01	1.35E-01	
IL21R-001	IL21-001	6.64E+01	6.64E+01	
IL21R-001	IL21-001-135	6.64E+01	8.59E+04	0.67
IL21R-001	IL21-001-40	6.64E+01	3.38E+06	0.43
IL21R-001	IL21-201	6.64E+01	3.29E+01	1.00
IL21R-001	IL21-201-135	6.64E+01	3.62E+05	0.57
IL21R-001	IL21-201-40	6.64E+01	5.91E+01	1.00
IL21R-001-318	IL21-001	6.64E+01	3.54E+01	1.00
IL21R-001-484	IL21-001	6.64E+01	4.78E-04	1.01
IL21R-003	IL21-001	6.64E+01	6.64E+01	
IL21R-003-318	IL21-001	6.64E+01	3.54E+01	1.00
IL21R-003-484	IL21-001	6.64E+01	4.16E+02	0.98
IL21R-006	IL21-001	6.64E+01	6.64E+01	
IL21R-006-318	IL21-001	6.64E+01	6.16E+01	1.00
IL21R-006-484	IL21-001	6.64E+01	4.16E+02	0.98
IL4R-001	IL4-001	8.19E+00	8.19E+00	
IL4R-001	IL4-001-100	8.19E+00	2.44E-40	1.00
IL4R-001	IL4-001-105	8.19E+00	1.63E+06	0.11
IL4R-001	IL4-001-109	8.19E+00	2.34E+00	1.00
IL4R-001	IL4-001-152	8.19E+00	9.10E+01	0.83
IL4R-001	IL4-001-22	8.19E+00	2.97E-24	1.00
IL4R-001	IL4-001-26	8.19E+00	1.74E+03	0.38
IL4R-001	IL4-001-30	8.19E+00	1.29E+06	0.11
IL4R-001	IL4-001-53	8.19E+00	9.38E+03	0.19
IL4R-001	IL4-001-98	8.19E+00	2.11E+04	0.14
IL4R-001	IL4-002	8.19E+00	3.78E+00	1.00
IL4R-001	IL4-002-102	8.19E+00	8.26E-20	1.00
IL4R-001	IL4-002-136	8.19E+00	1.86E+03	0.37
IL4R-001	IL4-002-22	8.19E+00	1.47E+03	0.41
IL4R-001	IL4-002-26	8.19E+00	8.40E+03	0.20
IL4R-001	IL4-002-30	8.19E+00	1.04E+00	1.00
IL4R-001	IL4-002-82	8.19E+00	1.20E+03	0.44
IL4R-001	IL4-002-84	8.19E+00	3.77E+01	0.99
IL4R-001	IL4-002-89	8.19E+00	1.12E+02	0.79
IL4R-001	IL4-002-93	8.19E+00	5.37E+04	0.12
IL4R-001	IL4-201	8.19E+00	4.81E-11	1.00
IL4R-001	IL4-201-134	8.19E+00	1.72E+07	0.10
IL4R-001	IL4-201-22	8.19E+00	1.71E+05	0.11
IL4R-001	IL4-201-26	8.19E+00	1.24E+01	1.00
IL4R-001	IL4-201-30	8.19E+00	4.33E-06	1.00
IL4R-001	IL4-201-67	8.19E+00	5.25E+05	0.11
IL4R-001	IL4-201-82	8.19E+00	7.49E+02	0.51

IL4R-001-185	IL4-001	8.19E+00	1.17E+05	0.12
IL4R-001-400	IL4-001	8.19E+00	3.79E+00	1.00
IL4R-001-436	IL4-001	8.19E+00	9.03E+03	0.19
IL4R-001-503	IL4-001	8.19E+00	1.09E+00	1.00
IL4R-001-576	IL4-001	8.19E+00	2.88E+02	0.65
IL4R-001-579	IL4-001	8.19E+00	5.99E+02	0.55
IL4R-001-75	IL4-001	8.19E+00	4.15E-01	1.00
IL4R-001-752	IL4-001	8.19E+00	1.29E+02	0.77
IL4R-001-82	IL4-001	8.19E+00	2.16E+06	0.11
IL4R-001-97	IL4-001	8.19E+00	1.50E-12	1.00
IL4R-004	IL4-001	8.19E+00	4.07E-02	1.00
IL4R-004-185	IL4-001	8.19E+00	1.50E+03	0.41
IL4R-004-400	IL4-001	8.19E+00	3.91E+01	0.99
IL4R-004-431	IL4-001	8.19E+00	9.10E+02	0.48
IL4R-004-436	IL4-001	8.19E+00	2.00E-36	1.00
IL4R-004-503	IL4-001	8.19E+00	1.22E+01	1.00
IL4R-004-576	IL4-001	8.19E+00	1.27E-01	1.00
IL4R-004-579	IL4-001	8.19E+00	1.11E+01	1.00
IL4R-004-75	IL4-001	8.19E+00	1.36E+02	0.76
IL4R-004-752	IL4-001	8.19E+00	1.87E+01	1.00
IL4R-004-786	IL4-001	8.19E+00	1.32E+01	1.00
IL4R-004-82	IL4-001	8.19E+00	6.07E+04	0.12
IL4R-004-97	IL4-001	8.19E+00	3.42E+02	0.63
IL4R-201	IL4-001	8.19E+00	2.08E-11	1.00
IL4R-201-170	IL4-001	8.19E+00	3.49E-03	1.00
IL4R-201-385	IL4-001	8.19E+00	7.35E+02	0.52
IL4R-201-416	IL4-001	8.19E+00	4.25E+04	0.13
IL4R-201-421	IL4-001	8.19E+00	2.06E+02	0.70
IL4R-201-488	IL4-001	8.19E+00	5.72E+04	0.12
IL4R-201-561	IL4-001	8.19E+00	3.93E+02	0.61
IL4R-201-564	IL4-001	8.19E+00	2.34E-09	1.00
IL4R-201-60	IL4-001	8.19E+00	8.82E+01	0.84
IL4R-201-660	IL4-001	8.19E+00	1.51E+03	0.41
IL4R-201-67	IL4-001	8.19E+00	1.14E+01	1.00
IL4R-201-737	IL4-001	8.19E+00	6.38E+02	0.54
IL4R-201-771	IL4-001	8.19E+00	3.05E-181	1.00
IL4R-201-82	IL4-001	8.19E+00	5.44E+03	0.24
IL6R-001	IL6-001	6.64E+01	6.64E+01	
IL6R-001	IL6-001-104	6.64E+01	1.41E+03	0.94
IL6R-001	IL6-001-116	6.64E+01	1.18E+01	1.00
IL6R-001	IL6-001-152	6.64E+01	3.80E+01	1.00
IL6R-001	IL6-001-162	6.64E+01	1.37E+01	1.00
IL6R-001	IL6-001-2	6.64E+01	2.81E+02	0.99
IL6R-001	IL6-001-31	6.64E+01	3.36E+00	1.00

IL6R-001	IL6-001-32	6.64E+01	1.70E+02	0.99
IL6R-001	IL6-001-55	6.64E+01	1.20E+02	1.00
IL6R-001	IL6-001-6	6.64E+01	3.20E+02	0.98
IL6R-001	IL6-001-7	6.64E+01	8.84E+01	1.00
IL6R-001	IL6-001-79	6.64E+01	7.56E+02	0.96
IL6R-001	IL6-003	6.64E+01	3.65E+02	0.98
IL6R-001	IL6-003-110	6.64E+01	1.93E+02	0.99
IL6R-001	IL6-003-28	6.64E+01	2.54E+01	1.00
IL6R-001	IL6-003-3	6.64E+01	1.08E+02	1.00
IL6R-001	IL6-003-40	6.64E+01	3.17E+03	0.90
IL6R-001	IL6-003-76	6.64E+01	1.33E+01	1.00
IL6R-001	IL6-003-89	6.64E+01	8.18E+01	1.00
IL6R-001	IL6-003-92	6.64E+01	4.91E+02	0.98
IL6R-001	IL6-004	6.64E+01	1.00E+02	1.00
IL6R-001	IL6-004-28	6.64E+01	4.77E+02	0.98
IL6R-001	IL6-004-3	6.64E+01	5.06E+02	0.98
IL6R-001	IL6-004-40	6.64E+01	1.24E+02	1.00
IL6R-001	IL6-004-76	6.64E+01	2.11E+00	1.00
IL6R-001	IL6-004-86	6.64E+01	1.42E+00	1.00
IL6R-001	IL6-005	6.64E+01	8.33E+00	1.00
IL6R-001	IL6-005-129	6.64E+01	3.63E+01	1.00
IL6R-001	IL6-005-139	6.64E+01	3.71E+01	1.00
IL6R-001	IL6-005-2	6.64E+01	3.93E+01	1.00
IL6R-001	IL6-005-32	6.64E+01	5.02E+01	1.00
IL6R-001	IL6-005-56	6.64E+01	1.08E+02	1.00
IL6R-001	IL6-005-6	6.64E+01	2.51E+01	1.00
IL6R-001	IL6-005-8	6.64E+01	1.67E+02	0.99
IL6R-001	IL6-005-81	6.64E+01	3.55E+01	1.00
IL6R-001	IL6-005-9	6.64E+01	7.97E+01	1.00
IL6R-001	IL6-005-93	6.64E+01	1.33E-01	1.01
IL6R-001	IL6-006	6.64E+01	5.77E+00	1.00
IL6R-001	IL6-006-104	6.64E+01	8.59E+01	1.00
IL6R-001	IL6-006-116	6.64E+01	1.17E+02	1.00
IL6R-001	IL6-006-152	6.64E+01	1.08E+01	1.00
IL6R-001	IL6-006-165	6.64E+01	7.05E-01	1.00
IL6R-001	IL6-006-168	6.64E+01	6.85E+00	1.00
IL6R-001	IL6-006-186	6.64E+01	1.17E+01	1.00
IL6R-001	IL6-006-2	6.64E+01	1.81E+00	1.00
IL6R-001	IL6-006-31	6.64E+01	2.40E+03	0.92
IL6R-001	IL6-006-32	6.64E+01	5.62E+02	0.97
IL6R-001	IL6-006-55	6.64E+01	1.14E+03	0.95
IL6R-001	IL6-006-6	6.64E+01	3.34E+02	0.98
IL6R-001	IL6-006-7	6.64E+01	1.02E+03	0.95
IL6R-001	IL6-006-79	6.64E+01	1.90E+02	0.99

IL6R-001	IL6-201	6.64E+01	5.72E+01	1.00
IL6R-001	IL6-201-104	6.64E+01	9.25E+00	1.00
IL6R-001	IL6-201-116	6.64E+01	4.04E+01	1.00
IL6R-001	IL6-201-152	6.64E+01	3.10E+00	1.00
IL6R-001	IL6-201-162	6.64E+01	4.92E+01	1.00
IL6R-001	IL6-201-2	6.64E+01	2.94E+02	0.99
IL6R-001	IL6-201-31	6.64E+01	7.46E-01	1.00
IL6R-001	IL6-201-32	6.64E+01	6.08E+02	0.97
IL6R-001	IL6-201-55	6.64E+01	4.69E-06	1.01
IL6R-001	IL6-201-6	6.64E+01	7.31E+03	0.84
IL6R-001	IL6-201-7	6.64E+01	3.21E+00	1.00
IL6R-001	IL6-201-79	6.64E+01	2.90E+02	0.99
IL6R-001-358	IL6-001	6.64E+01	1.02E+01	1.00
IL6R-001-385	IL6-001	6.64E+01	9.45E+01	1.00
IL6R-001-65	IL6-001	6.64E+01	4.77E+01	1.00
IL6R-003	IL6-001	6.64E+01	1.07E+02	1.00
IL6R-003-65	IL6-001	6.64E+01	9.45E+01	1.00
IL6R-201	IL6-001	6.64E+01	7.05E+01	1.00
IL6R-201-65	IL6-001	6.64E+01	4.54E+01	1.00
MMP1-001	CMA1-001	2.00E-03	2.00E-03	
MMP1-001	CMA1-001-151	2.00E-03	1.20E-03	1.00
MMP1-001	CMA1-001-157	2.00E-03	1.02E-02	1.01
MMP1-001	CMA1-001-183	2.00E-03	3.96E-06	1.00
MMP1-001	CMA1-001-221	2.00E-03	7.21E-03	1.01
MMP1-001	CMA1-001-226	2.00E-03	7.76E-05	1.00
MMP1-001	CMA1-001-33	2.00E-03	3.13E-04	1.00
MMP1-001	CMA1-001-38	2.00E-03	4.71E-11	1.00
MMP1-001	CMA1-001-46	2.00E-03	2.85E-04	1.00
MMP1-001	CMA1-001-48	2.00E-03	2.66E-05	1.00
MMP1-001	CMA1-001-63	2.00E-03	4.79E-04	1.00
MMP1-001	CMA1-001-66	2.00E-03	1.44E-03	1.00
MMP1-001	CMA1-001-69	2.00E-03	1.90E-01	1.02
MMP1-001	CMA1-001-98	2.00E-03	1.29E-07	1.00
MMP1-001	CMA1-002	2.00E-03	9.78E-05	1.00
MMP1-001	CMA1-002-1	2.00E-03	1.56E-08	1.00
MMP1-001	CMA1-002-106	2.00E-03	1.88E-06	1.00
MMP1-001	CMA1-002-110	2.00E-03	3.06E-10	1.00
MMP1-001	CMA1-002-115	2.00E-03	5.07E-02	1.02
MMP1-001	CMA1-002-124	2.00E-03	2.28E-02	1.01
MMP1-001	CMA1-002-126	2.00E-03	1.54E-02	1.01
MMP1-001	CMA1-002-40	2.00E-03	7.21E-03	1.01
MMP1-001	CMA1-002-46	2.00E-03	6.61E-04	1.00
MMP1-001	CMA1-002-48	2.00E-03	1.28E-06	1.00
MMP1-001	CMA1-002-59	2.00E-03	8.75E-06	1.00

MMP1-001	CMA1-002-72	2.00E-03	6.22E-03	1.00
MMP1-001	CMA1-002-75	2.00E-03	2.02E-03	1.00
MMP1-001	CMA1-002-77	2.00E-03	1.55E-02	1.01
MMP1-001	CMA1-002-93	2.00E-03	4.95E-03	1.00
MMP1-001	CMA1-002-95	2.00E-03	2.19E-02	1.01
MMP1-001-374	CMA1-001	2.00E-03	7.92E-04	1.00
MMP1-001-406	CMA1-001	2.00E-03	3.75E-03	1.00
MMP1-001	TIMP1-001	1.00E-07	1.00E-07	
MMP1-001	TIMP1-001-10	1.00E-07	6.21E-09	1.00
MMP1-001	TIMP1-001-102	1.00E-07	4.11E-08	1.00
MMP1-001	TIMP1-001-105	1.00E-07	3.46E-09	1.00
MMP1-001	TIMP1-001-136	1.00E-07	3.13E-08	1.00
MMP1-001	TIMP1-001-171	1.00E-07	1.09E-13	1.00
MMP1-001	TIMP1-001-185	1.00E-07	1.21E-07	1.00
MMP1-001	TIMP1-001-192	1.00E-07	1.41E-08	1.00
MMP1-001	TIMP1-001-203	1.00E-07	2.75E-15	1.00
MMP1-001	TIMP1-001-33	1.00E-07	1.34E-08	1.00
MMP1-001	TIMP1-001-50	1.00E-07	1.24E-10	1.00
MMP1-001	TIMP1-001-51	1.00E-07	1.37E-08	1.00
MMP1-001	TIMP1-001-76	1.00E-07	1.00E-12	1.00
MMP1-001	TIMP1-001-77	1.00E-07	1.94E-11	1.00
MMP1-001	TIMP1-002	1.00E-07	1.56E-08	1.00
MMP1-001	TIMP1-002-10	1.00E-07	6.85E-10	1.00
MMP1-001	TIMP1-002-102	1.00E-07	1.14E-08	1.00
MMP1-001	TIMP1-002-105	1.00E-07	2.66E-09	1.00
MMP1-001	TIMP1-002-115	1.00E-07	5.49E-09	1.00
MMP1-001	TIMP1-002-116	1.00E-07	4.04E-08	1.00
MMP1-001	TIMP1-002-33	1.00E-07	2.92E-13	1.00
MMP1-001	TIMP1-002-50	1.00E-07	4.79E-20	1.00
MMP1-001	TIMP1-002-51	1.00E-07	1.85E-11	1.00
MMP1-001	TIMP1-002-76	1.00E-07	3.23E-07	1.00
MMP1-001	TIMP1-002-77	1.00E-07	4.04E-10	1.00
MMP1-001	TIMP1-003	1.00E-07	4.91E-12	1.00
MMP1-001	TIMP1-003-107	1.00E-07	1.20E-08	1.00
MMP1-001	TIMP1-003-12	1.00E-07	2.40E-08	1.00
MMP1-001	TIMP1-003-121	1.00E-07	1.44E-09	1.00
MMP1-001	TIMP1-003-128	1.00E-07	1.77E-09	1.00
MMP1-001	TIMP1-003-13	1.00E-07	5.40E-08	1.00
MMP1-001	TIMP1-003-139	1.00E-07	3.65E-11	1.00
MMP1-001	TIMP1-003-38	1.00E-07	1.37E-08	1.00
MMP1-001	TIMP1-003-41	1.00E-07	2.16E-07	1.00
MMP1-001	TIMP1-003-72	1.00E-07	2.92E-12	1.00
MMP1-001-374	TIMP1-001	1.00E-07	3.41E-08	1.00
MMP1-001-406	TIMP1-001	1.00E-07	6.20E-12	1.00

MMP1-001	TIMP2-001	1.00E-07	1.00E-07	
MMP1-001	TIMP2-003	1.00E-07	1.97E-09	1.00
MMP1-001	TIMP2-003-130	1.00E-07	2.82E-08	1.00
MMP1-001	TIMP2-003-139	1.00E-07	4.92E-03	1.00
MMP1-001	TIMP2-008	1.00E-07	1.97E-09	1.00
MMP1-001	TIMP2-008-130	1.00E-07	9.57E-03	1.00
MMP1-001	TIMP2-008-139	1.00E-07	3.38E-04	1.00
MMP1-001	TIMP2-008-42	1.00E-07	1.21E-05	1.00
MMP1-001	TIMP2-008-54	1.00E-07	8.44E-09	1.00
MMP1-001	TIMP2-008-69	1.00E-07	4.41E-06	1.00
MMP13-001	TIMP1-001	1.00E-07	1.00E-07	
MMP13-001	TIMP1-001-10	1.00E-07	1.36E-06	1.00
MMP13-001	TIMP1-001-102	1.00E-07	7.30E-09	1.00
MMP13-001	TIMP1-001-105	1.00E-07	1.97E-07	1.00
MMP13-001	TIMP1-001-136	1.00E-07	1.73E-10	1.00
MMP13-001	TIMP1-001-171	1.00E-07	5.27E-08	1.00
MMP13-001	TIMP1-001-185	1.00E-07	6.05E-08	1.00
MMP13-001	TIMP1-001-192	1.00E-07	2.40E-07	1.00
MMP13-001	TIMP1-001-203	1.00E-07	2.20E-08	1.00
MMP13-001	TIMP1-001-33	1.00E-07	1.29E-07	1.00
MMP13-001	TIMP1-001-50	1.00E-07	4.61E-08	1.00
MMP13-001	TIMP1-001-51	1.00E-07	2.92E-10	1.00
MMP13-001	TIMP1-001-76	1.00E-07	2.35E-13	1.00
MMP13-001	TIMP1-001-77	1.00E-07	2.72E-07	1.00
MMP13-001	TIMP1-002	1.00E-07	2.99E-11	1.00
MMP13-001	TIMP1-002-10	1.00E-07	4.80E-12	1.00
MMP13-001	TIMP1-002-102	1.00E-07	2.62E-09	1.00
MMP13-001	TIMP1-002-105	1.00E-07	4.22E-11	1.00
MMP13-001	TIMP1-002-115	1.00E-07	5.86E-10	1.00
MMP13-001	TIMP1-002-116	1.00E-07	2.41E-11	1.00
MMP13-001	TIMP1-002-33	1.00E-07	1.60E-11	1.00
MMP13-001	TIMP1-002-50	1.00E-07	1.62E-14	1.00
MMP13-001	TIMP1-002-51	1.00E-07	1.26E-14	1.00
MMP13-001	TIMP1-002-76	1.00E-07	8.48E-11	1.00
MMP13-001	TIMP1-002-77	1.00E-07	2.44E-10	1.00
MMP13-001	TIMP1-003	1.00E-07	5.51E-10	1.00
MMP13-001	TIMP1-003-107	1.00E-07	2.11E-10	1.00
MMP13-001	TIMP1-003-12	1.00E-07	1.92E-09	1.00
MMP13-001	TIMP1-003-121	1.00E-07	8.32E-10	1.00
MMP13-001	TIMP1-003-128	1.00E-07	5.66E-09	1.00
MMP13-001	TIMP1-003-13	1.00E-07	1.15E-09	1.00
MMP13-001	TIMP1-003-139	1.00E-07	1.27E-10	1.00
MMP13-001	TIMP1-003-38	1.00E-07	5.94E-10	1.00
MMP13-001	TIMP1-003-41	1.00E-07	1.20E-15	1.00

MMP13-001	TIMP1-003-72	1.00E-07	7.19E-13	1.00
MMP13-001-158	TIMP1-001	1.00E-07	1.56E-07	1.00
MMP13-002	TIMP1-001	1.00E-07	5.59E-11	1.00
MMP13-002-158	TIMP1-001	1.00E-07	4.26E-08	1.00
MMP13-201	TIMP1-001	1.00E-07	5.08E-08	1.00
MMP13-201-158	TIMP1-001	1.00E-07	5.55E-10	1.00
MMP13-001	TIMP2-001	1.00E-07	1.00E-07	
MMP13-001	TIMP2-001-119	1.00E-07	2.88E-06	1.00
MMP13-001	TIMP2-003	1.00E-07	1.96E-08	1.00
MMP13-001	TIMP2-003-130	1.00E-07	1.87E-07	1.00
MMP13-001	TIMP2-003-139	1.00E-07	3.36E-19	1.00
MMP13-001	TIMP2-008	1.00E-07	9.22E-09	1.00
MMP13-001	TIMP2-008-130	1.00E-07	5.82E-08	1.00
MMP13-001	TIMP2-008-139	1.00E-07	3.36E-19	1.00
MMP13-001	TIMP2-008-42	1.00E-07	3.42E-07	1.00
MMP13-001	TIMP2-008-54	1.00E-07	8.76E-09	1.00
MMP13-001	TIMP2-008-69	1.00E-07	7.19E-10	1.00
MMP13-001-158	TIMP2-001	1.00E-07	1.50E-08	1.00
MMP13-002	TIMP2-001	1.00E-07	2.21E-07	1.00
MMP13-002-158	TIMP2-001	1.00E-07	1.67E-08	1.00
MMP13-201	TIMP2-001	1.00E-07	6.84E-09	1.00
MMP13-201-158	TIMP2-001	1.00E-07	8.68E-09	1.00
MMP13-001	TIMP3-001	1.00E-07	1.00E-07	
MMP13-001	TIMP3-001-142	1.00E-07	2.96E-09	1.00
MMP13-001	TIMP3-001-201	1.00E-07	5.48E-10	1.00
MMP13-001	TIMP3-001-22	1.00E-07	2.31E-07	1.00
MMP13-001	TIMP3-001-37	1.00E-07	5.61E-07	1.00
MMP13-001-158	TIMP3-001	1.00E-07	2.76E-07	1.00
MMP13-002	TIMP3-001	1.00E-07	1.92E-08	1.00
MMP13-201	TIMP3-001	1.00E-07	2.40E-05	1.00
MMP13-201-158	TIMP3-001	1.00E-07	8.02E-05	1.00
MMP2-001	TIMP1-001	1.00E-07	1.00E-07	
MMP2-001	TIMP1-001-10	1.00E-07	4.14E-07	1.00
MMP2-001	TIMP1-001-102	1.00E-07	5.22E-08	1.00
MMP2-001	TIMP1-001-105	1.00E-07	8.48E-09	1.00
MMP2-001	TIMP1-001-136	1.00E-07	3.03E-11	1.00
MMP2-001	TIMP1-001-171	1.00E-07	1.39E-14	1.00
MMP2-001	TIMP1-001-185	1.00E-07	4.94E-09	1.00
MMP2-001	TIMP1-001-192	1.00E-07	3.64E-07	1.00
MMP2-001	TIMP1-001-203	1.00E-07	1.19E-09	1.00
MMP2-001	TIMP1-001-33	1.00E-07	1.08E-07	1.00
MMP2-001	TIMP1-001-50	1.00E-07	1.39E-09	1.00
MMP2-001	TIMP1-001-51	1.00E-07	6.95E-08	1.00
MMP2-001	TIMP1-001-76	1.00E-07	1.44E-09	1.00

MMP2-001	TIMP1-001-77	1.00E-07	1.85E-07	1.00
MMP2-001	TIMP1-002	1.00E-07	1.73E-09	1.00
MMP2-001	TIMP1-002-10	1.00E-07	2.17E-11	1.00
MMP2-001	TIMP1-002-102	1.00E-07	6.85E-09	1.00
MMP2-001	TIMP1-002-105	1.00E-07	1.73E-14	1.00
MMP2-001	TIMP1-002-115	1.00E-07	1.41E-14	1.00
MMP2-001	TIMP1-002-116	1.00E-07	1.40E-11	1.00
MMP2-001	TIMP1-002-33	1.00E-07	3.59E-09	1.00
MMP2-001	TIMP1-002-50	1.00E-07	3.95E-08	1.00
MMP2-001	TIMP1-002-51	1.00E-07	3.77E-09	1.00
MMP2-001	TIMP1-002-76	1.00E-07	1.80E-10	1.00
MMP2-001	TIMP1-002-77	1.00E-07	1.42E-09	1.00
MMP2-001	TIMP1-003	1.00E-07	1.07E-14	1.00
MMP2-001	TIMP1-003-107	1.00E-07	5.20E-09	1.00
MMP2-001	TIMP1-003-12	1.00E-07	2.78E-09	1.00
MMP2-001	TIMP1-003-121	1.00E-07	7.18E-09	1.00
MMP2-001	TIMP1-003-128	1.00E-07	1.75E-10	1.00
MMP2-001	TIMP1-003-13	1.00E-07	2.11E-15	1.00
MMP2-001	TIMP1-003-139	1.00E-07	4.22E-10	1.00
MMP2-001	TIMP1-003-38	1.00E-07	2.76E-10	1.00
MMP2-001	TIMP1-003-41	1.00E-07	4.86E-09	1.00
MMP2-001	TIMP1-003-72	1.00E-07	4.42E-14	1.00
MMP2-001-447	TIMP1-001	1.00E-07	6.88E-07	1.00
MMP2-008	TIMP1-001	1.00E-07	7.68E-08	1.00
MMP2-008-333	TIMP1-001	1.00E-07	5.50E-06	1.00
MMP2-008-371	TIMP1-001	1.00E-07	1.63E-07	1.00
MMP2-008-545	TIMP1-001	1.00E-07	1.30E-08	1.00
MMP2-001	TIMP2-001	1.00E-07	1.00E-07	1.00
MMP2-001	TIMP2-001-119	1.00E-07	6.92E-03	1.00
MMP2-001	TIMP2-001-146	1.00E-07	8.72E-03	1.00
MMP2-001	TIMP2-003	1.00E-07	4.44E-02	1.00
MMP2-001	TIMP2-003-130	1.00E-07	1.13E-01	1.00
MMP2-001	TIMP2-003-139	1.00E-07	1.73E-03	1.00
MMP2-001	TIMP2-008	1.00E-07	3.83E-04	1.00
MMP2-001	TIMP2-008-130	1.00E-07	1.33E-03	1.00
MMP2-001	TIMP2-008-42	1.00E-07	1.09E-03	1.00
MMP2-001	TIMP2-008-54	1.00E-07	4.68E-02	1.00
MMP2-001	TIMP2-008-69	1.00E-07	2.59E-04	1.00
MMP2-001-447	TIMP2-001	1.00E-07	7.36E-05	1.00
MMP2-008	TIMP2-001	1.00E-07	3.60E-02	1.00
MMP2-008-333	TIMP2-001	1.00E-07	2.04E-03	1.00
MMP2-008-371	TIMP2-001	1.00E-07	2.20E-03	1.00
MMP2-008-545	TIMP2-001	1.00E-07	2.04E-01	1.00
MMP2-001	TIMP3-001	1.00E-07	1.00E-07	

MMP2-001	TIMP3-001-121	1.00E-07	9.83E+00	1.00
MMP2-001	TIMP3-001-186	1.00E-07	3.67E-07	1.00
MMP2-001-447	TIMP3-001	1.00E-07	1.65E-02	1.00
MMP2-008-333	TIMP3-001	1.00E-07	2.60E-05	1.00
MMP2-008-371	TIMP3-001	1.00E-07	9.81E-06	1.00
MMP2-001	TIMP4-001	1.00E-07	1.00E-07	
MMP2-001	TIMP4-001-103	1.00E-07	4.16E-10	1.00
MMP2-001	TIMP4-001-109	1.00E-07	1.06E-08	1.00
MMP2-001	TIMP4-001-141	1.00E-07	1.36E-08	1.00
MMP2-001	TIMP4-001-143	1.00E-07	2.87E-09	1.00
MMP2-001	TIMP4-001-19	1.00E-07	1.66E-09	1.00
MMP2-001	TIMP4-001-206	1.00E-07	1.20E-09	1.00
MMP2-001	TIMP4-001-214	1.00E-07	1.72E-08	1.00
MMP2-001	TIMP4-001-24	1.00E-07	1.55E-08	1.00
MMP3-001	TIMP1-001	1.00E-07	1.82E-09	1.00
MMP3-001	TIMP1-001-10	1.00E-07	2.33E-10	1.00
MMP3-001	TIMP1-001-102	1.00E-07	1.64E-17	1.00
MMP3-001	TIMP1-001-105	1.00E-07	1.31E-12	1.00
MMP3-001	TIMP1-001-136	1.00E-07	6.59E-09	1.00
MMP3-001	TIMP1-001-171	1.00E-07	5.54E-08	1.00
MMP3-001	TIMP1-001-185	1.00E-07	1.26E-08	1.00
MMP3-001	TIMP1-001-192	1.00E-07	1.56E-09	1.00
MMP3-001	TIMP1-001-203	1.00E-07	1.01E-16	1.00
MMP3-001	TIMP1-001-33	1.00E-07	1.29E-08	1.00
MMP3-001	TIMP1-001-50	1.00E-07	7.15E-10	1.00
MMP3-001	TIMP1-001-51	1.00E-07	8.58E-12	1.00
MMP3-001	TIMP1-001-76	1.00E-07	7.85E-09	1.00
MMP3-001	TIMP1-001-77	1.00E-07	5.86E-10	1.00
MMP3-001	TIMP1-002	1.00E-07	8.36E-08	1.00
MMP3-001	TIMP1-002-10	1.00E-07	6.52E-09	1.00
MMP3-001	TIMP1-002-102	1.00E-07	2.03E-09	1.00
MMP3-001	TIMP1-002-105	1.00E-07	6.89E-09	1.00
MMP3-001	TIMP1-002-115	1.00E-07	3.27E-08	1.00
MMP3-001	TIMP1-002-116	1.00E-07	1.80E-10	1.00
MMP3-001	TIMP1-002-33	1.00E-07	3.76E-09	1.00
MMP3-001	TIMP1-002-50	1.00E-07	3.41E-09	1.00
MMP3-001	TIMP1-002-51	1.00E-07	1.21E-08	1.00
MMP3-001	TIMP1-002-76	1.00E-07	9.65E-10	1.00
MMP3-001	TIMP1-002-77	1.00E-07	4.16E-09	1.00
MMP3-001	TIMP1-003	1.00E-07	5.68E-07	1.00
MMP3-001	TIMP1-003-107	1.00E-07	2.02E-08	1.00
MMP3-001	TIMP1-003-12	1.00E-07	2.76E-09	1.00
MMP3-001	TIMP1-003-121	1.00E-07	3.43E-09	1.00
MMP3-001	TIMP1-003-128	1.00E-07	1.79E-08	1.00

MMP3-001	TIMP1-003-13	1.00E-07	4.51E-07	1.00
MMP3-001	TIMP1-003-139	1.00E-07	2.10E-09	1.00
MMP3-001	TIMP1-003-38	1.00E-07	1.97E-07	1.00
MMP3-001	TIMP1-003-41	1.00E-07	7.48E-09	1.00
MMP3-001	TIMP1-003-72	1.00E-07	2.21E-10	1.00
MMP3-001-45	TIMP1-001	1.00E-07	1.41E-06	1.00
MMP3-001-96	TIMP1-001	1.00E-07	5.63E-09	1.00
MMP3-001	TIMP2-001	1.00E-07	4.67E-10	1.00
MMP3-001	TIMP2-003	1.00E-07	3.48E-10	1.00
MMP3-001	TIMP2-003-130	1.00E-07	1.19E-08	1.00
MMP3-001	TIMP2-003-139	1.00E-07	1.13E-07	1.00
MMP3-001	TIMP2-008	1.00E-07	1.80E-07	1.00
MMP3-001	TIMP2-008-130	1.00E-07	4.91E-09	1.00
MMP3-001	TIMP2-008-139	1.00E-07	2.50E-08	1.00
MMP3-001	TIMP2-008-42	1.00E-07	5.44E-08	1.00
MMP3-001	TIMP2-008-54	1.00E-07	3.53E-08	1.00
MMP3-001	TIMP2-008-69	1.00E-07	7.43E-09	1.00
MMP3-001	TIMP3-001	1.00E-07	2.08E-08	1.00
MMP3-001	TIMP3-001-121	1.00E-07	2.99E-07	1.00
MMP3-001	TIMP3-001-142	1.00E-07	2.90E-09	1.00
MMP3-001	TIMP3-001-186	1.00E-07	7.13E-09	1.00
MMP3-001	TIMP3-001-196	1.00E-07	2.00E-04	1.00
MMP3-001	TIMP3-001-22	1.00E-07	1.68E-07	1.00
MMP3-001	TIMP3-001-37	1.00E-07	2.22E-07	1.00
MMP3-001-45	TIMP3-001	1.00E-07	1.31E-03	1.00
MMP3-001	TIMP4-001	1.00E-07	1.16E-09	1.00
MMP3-001	TIMP4-001-141	1.00E-07	2.78E-11	1.00
MMP3-001	TIMP4-001-214	1.00E-07	5.02E-22	1.00
MMP3-001	TIMP4-001-24	1.00E-07	2.50E-07	1.00
MMP3-001	TIMP4-001-28	1.00E-07	1.80E-05	1.00
MMP3-001-45	TIMP4-001	1.00E-07	5.14E-07	1.00
MMP9-001	TIMP1-001	1.00E-07	1.00E-07	
MMP9-001	TIMP1-001-10	1.00E-07	1.04E-06	0.97
MMP9-001	TIMP1-001-102	1.00E-07	1.18E-03	0.90
MMP9-001	TIMP1-001-105	1.00E-07	3.90E-03	0.90
MMP9-001	TIMP1-001-136	1.00E-07	1.74E-03	0.90
MMP9-001	TIMP1-001-171	1.00E-07	5.79E-09	1.00
MMP9-001	TIMP1-001-185	1.00E-07	2.01E-04	0.90
MMP9-001	TIMP1-001-192	1.00E-07	7.42E-04	0.90
MMP9-001	TIMP1-001-203	1.00E-07	2.63E-09	1.00
MMP9-001	TIMP1-001-33	1.00E-07	3.78E-09	1.00
MMP9-001	TIMP1-001-50	1.00E-07	1.75E-05	0.91
MMP9-001	TIMP1-001-51	1.00E-07	2.88E-06	0.95
MMP9-001	TIMP1-001-76	1.00E-07	2.56E-04	0.90

MMP9-001	TIMP1-001-77	1.00E-07	3.15E-04	0.90
MMP9-001	TIMP1-002	1.00E-07	9.11E-04	0.90
MMP9-001	TIMP1-002-10	1.00E-07	7.29E-03	0.90
MMP9-001	TIMP1-002-102	1.00E-07	4.17E-05	0.91
MMP9-001	TIMP1-002-105	1.00E-07	5.72E-03	0.90
MMP9-001	TIMP1-002-115	1.00E-07	1.68E-05	0.92
MMP9-001	TIMP1-002-116	1.00E-07	1.58E-03	0.90
MMP9-001	TIMP1-002-33	1.00E-07	4.74E-06	0.94
MMP9-001	TIMP1-002-50	1.00E-07	1.50E-03	0.90
MMP9-001	TIMP1-002-51	1.00E-07	3.29E-08	1.00
MMP9-001	TIMP1-002-76	1.00E-07	3.26E-06	0.95
MMP9-001	TIMP1-002-77	1.00E-07	6.26E-06	0.93
MMP9-001	TIMP1-003	1.00E-07	1.61E-03	0.90
MMP9-001	TIMP1-003-107	1.00E-07	7.82E-06	0.93
MMP9-001	TIMP1-003-12	1.00E-07	5.39E-05	0.91
MMP9-001	TIMP1-003-121	1.00E-07	1.50E-06	0.97
MMP9-001	TIMP1-003-128	1.00E-07	9.78E-06	0.92
MMP9-001	TIMP1-003-13	1.00E-07	5.39E-08	1.00
MMP9-001	TIMP1-003-139	1.00E-07	2.03E-06	0.96
MMP9-001	TIMP1-003-38	1.00E-07	2.28E-04	0.90
MMP9-001	TIMP1-003-41	1.00E-07	3.74E-05	0.91
MMP9-001	TIMP1-003-72	1.00E-07	2.74E-07	0.99
MMP9-001-20	TIMP1-001	1.00E-07	3.25E-08	1.00
MMP9-001-239	TIMP1-001	1.00E-07	3.36E-06	0.95
MMP9-001-279	TIMP1-001	1.00E-07	8.26E-04	0.90
MMP9-001-38	TIMP1-001	1.00E-07	2.67E-03	0.90
MMP9-001-574	TIMP1-001	1.00E-07	7.39E-07	0.98
MMP9-001-668	TIMP1-001	1.00E-07	1.46E-03	0.90
MMP9-001	TIMP2-001	1.00E-07	1.00E-07	
MMP9-001	TIMP2-003	1.00E-07	1.19E-05	0.92
MMP9-001	TIMP2-003-130	1.00E-07	1.63E-04	0.90
MMP9-001	TIMP2-003-139	1.00E-07	2.28E-04	0.90
MMP9-001	TIMP2-008	1.00E-07	1.19E-05	0.92
MMP9-001	TIMP2-008-130	1.00E-07	4.76E-13	1.00
MMP9-001	TIMP2-008-139	1.00E-07	2.28E-04	0.90
MMP9-001	TIMP2-008-42	1.00E-07	4.25E-02	0.90
MMP9-001	TIMP2-008-54	1.00E-07	4.84E-04	0.90
MMP9-001	TIMP2-008-69	1.00E-07	1.12E-04	0.90
MMP9-001-279	TIMP2-001	1.00E-07	1.06E-07	1.00
MMP9-001	TIMP3-001	1.00E-07	1.00E-07	
MMP9-001	TIMP3-001-121	1.00E-07	2.75E-04	0.90
MMP9-001	TIMP3-001-142	1.00E-07	4.82E-08	1.00
MMP9-001	TIMP3-001-186	1.00E-07	3.17E-04	0.90
MMP9-001	TIMP3-001-196	1.00E-07	4.46E-04	0.90

MMP9-001	TIMP3-001-201	1.00E-07	1.51E-04	0.90
MMP9-001	TIMP3-001-22	1.00E-07	2.89E-05	0.91
MMP9-001	TIMP3-001-37	1.00E-07	1.77E-05	0.91
MMP9-001-20	TIMP3-001	1.00E-07	2.49E-08	1.00
MMP9-001-239	TIMP3-001	1.00E-07	8.36E-06	0.93
MMP9-001-279	TIMP3-001	1.00E-07	8.07E-02	0.90
MMP9-001-38	TIMP3-001	1.00E-07	1.64E-03	0.90
MMP9-001-574	TIMP3-001	1.00E-07	2.05E-06	0.96
MMP9-001-668	TIMP3-001	1.00E-07	4.69E-14	1.00
MMP9-001	TIMP4-001	1.00E-07	1.00E-07	
MMP9-001	TIMP4-001-103	1.00E-07	5.73E-09	1.00
MMP9-001	TIMP4-001-109	1.00E-07	5.99E-09	1.00
MMP9-001	TIMP4-001-112	1.00E-07	6.04E-11	1.00
MMP9-001	TIMP4-001-141	1.00E-07	6.17E-11	1.00
MMP9-001	TIMP4-001-143	1.00E-07	4.01E-10	1.00
MMP9-001	TIMP4-001-19	1.00E-07	2.79E-08	1.00
MMP9-001	TIMP4-001-206	1.00E-07	9.74E-11	1.00
MMP9-001	TIMP4-001-214	1.00E-07	1.64E-11	1.00
MMP9-001	TIMP4-001-24	1.00E-07	7.87E-11	1.00
MMP9-001	TIMP4-001-28	1.00E-07	2.27E-10	1.00
MMP9-001-20	TIMP4-001	1.00E-07	2.05E-11	1.00
MMP9-001-239	TIMP4-001	1.00E-07	2.19E-12	1.00
MMP9-001-279	TIMP4-001	1.00E-07	2.23E-11	1.00
MMP9-001-38	TIMP4-001	1.00E-07	9.43E-12	1.00
MMP9-001-574	TIMP4-001	1.00E-07	5.86E-17	1.00
MMP9-001-668	TIMP4-001	1.00E-07	6.98E-18	1.00
TGFBR1-001	TGFB1-001	0.644 0.00625 8.1882	0.644 0.00625 8.1882	
TGFBR1-001	TGFB1-001-10	0.644 0.00625 8.1882	0.1979 0.0019206 2.5162	1.00
TGFBR1-001	TGFB1-001-25	0.644 0.00625 8.1882	8.9657e-07 8.7012e-09 1.14e-05	1.00
TGFBR1-001	TGFB1-001-263	0.644 0.00625 8.1882	0.023398 0.00022708 0.2975	1.00
TGFBR1-001-17	TGFB1-001	0.644 0.00625 8.1882	0.54778 0.0053162 6.9648	1.00
TGFBR1-001-19	TGFB1-001	0.644 0.00625 8.1882	1.0179 0.009879 12.9426	1.00
TGFBR1-003	TGFB1-001	0.644 0.00625 8.1882	1.979 0.019206 25.1624	1.00

TGFBR2-001	TGFB1-001	0.644 0.00625 8.1882	0.644 0.00625 8.1882	
TGFBR2-001	TGFB1-001-10	0.644 0.00625 8.1882	63.2344 0.61369 803.9997	0.53
TIMP3-001	ADAM17-001	2.2 2.2 2.2 2.2	2.2 2.2 2.2 2.2	
TIMP3-001	ADAM17-001-202	2.2 2.2 2.2 2.2	2.1558e-110 2.1558e-110 2.1558e-110 2.1558e-110	1.00
TIMP3-001-121	ADAM17-001	2.2 2.2 2.2 2.2	0.0093233 0.0093233 0.0093233 0.0093233	1.00
TIMP3-001-142	ADAM17-001	2.2 2.2 2.2 2.2	0.031802 0.031802 0.031802 0.031802	1.00
TIMP3-001-186	ADAM17-001	2.2 2.2 2.2 2.2	17.7622 17.7622 17.7622 17.7622	1.00
TIMP3-001-196	ADAM17-001	2.2 2.2 2.2 2.2	0.55645 0.55645 0.55645 0.55645	1.00
TIMP3-001-201	ADAM17-001	2.2 2.2 2.2 2.2	1.1608 1.1608 1.1608 1.1608	1.00
TIMP3-001-22	ADAM17-001	2.2 2.2 2.2 2.2	722.2794 722.2794 722.2794 722.2794	1.00
TIMP3-001-37	ADAM17-001	2.2 2.2 2.2 2.2	0.91022 0.91022 0.91022 0.91022	1.00

References

Abdulhannan, P., Russell, D.A., and Homer-Vanniasinkam, S. (2012). Peripheral arterial disease: A literature review. Br. Med. Bull. *104*, 21–39.

Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.

Adiguzel, E., Ahmad, P.J., Franco, C., and Bendeck, M.P. (2009). Collagens in the progression and complications of atherosclerosis. Vasc. Med. *14*, 73–89.

Ahmad, E., Rabbani, G., Zaidi, N., Khan, M.A., Qadeer, A., Ishtikhar, M., Singh, S., and Khan, R.H. (2013). Revisiting ligand-induced conformational changes in proteins: essence, advancements, implications and future challenges. J. Biomol. Struct. Dyn. *31*, 630–648.

Ahmed, S.A., and Giddens, D.P. (1983). Velocity measurements in steady flow through axisymmetric stenoses at moderate Reynolds numbers. J. Biomech. *16*, 505–516.

Ai, L., and Vafai, K. (2006). A coupling model for macromolecule transport in a stenosed arterial wall. Int. J. Heat Mass Transf. *49*, 1568–1591.

Ajmera, I., Swat, M., Laibe, C., Novere, N.L., and Chelliah, V. (2013). The impact of mathematical modeling on the understanding of diabetes and related complications. CPT Pharmacometrics Syst Pharmacol *2*, e54.

Akboga, M.K., Canpolat, U., Yayla, C., Ozcan, F., Ozeke, O., Topaloglu, S., and Aras, D. (2015). Association of Platelet to Lymphocyte Ratio With Inflammation and Severity of Coronary Atherosclerosis in Patients With Stable Coronary Artery Disease. Angiology *67*, 89–95.

Akman, O.E., Watterson, S., Parton, A., Binns, N., Millar, A.J., and Ghazal, P. (2012). Digital clocks: simple Boolean models can quantitatively describe circadian systems. J. R. Soc. Interface *9*, 2365–2382.

Alas, S., Emmanouilides, C., and Bonavida, B. (2001). Inhibition of interleukin 10 by Rituximab results in down-regulation of Bcl-2 and sensitization of B-cell non-Hodgkin's lymphoma to apoptosis. Clin. Cancer Res. *7*, 709–723.

Alaupovic, P., Wang, C., Mcconathy, W.J., Weiser, D., and Downs, D. (1986). Lipolytic Degradation of Human Very Low Density Lipoproteins by Human Milk Lipoprotein Lipase : The Identification of Lipoprotein B as the Main Lipoprotein Degradation Product ' to be the main reason for their compositional difference from native LDLz . T. *244*.

Aldridge, B.B., Burke, J.M., Lauffenburger, D.A., and Sorger, P.K. (2006). Physicochemical modelling of cell signalling pathways. Nat. Cell Biol. *8*, 1195–1203.

Alexander, R.W. (1995). Hypertension and the Pathogenesis of Atherosclerosis : Oxidative Stress and the Mediation of Arterial Inflammatory Response: A New Perspective. Hypertension *25*, 155–161.

Alsallaq, R., and Zhou, H.-X. (2008). Electrostatic Rate Enhancement and Transient Complex of Protein-Protein

Association. Proteins *71*, 320–335.

Amour, A., Slocombe, P.M., Webster, A., Butler, M., Knight, C.G., Smith, B.J., Stephens, P.E., Shelley, C., Hutton, M., Knäuper, V., et al. (1998). TNF-α converting enzyme (TACE) is inhibited by TIMP-3. FEBS Lett. *435*, 39–44.

De Andrade Leite, S.R. (2009). Inhibitors of human collagenase, MMP1. Eclet. Quim. *34*, 87–102.

Arakelyan, A., Petrkova, J., Hermanova, Z., Boyajyan, A., Lukl, J., and Petrek, M. (2005). Serum levels of the MCP-1 chemokine in patients with ischemic stroke and myocardial infarction. Mediators Inflamm. *2005*, 175–179.

Ardi, V.C., Kupriyanova, T.A., Deryugina, E.I., and Quigley, J.P. (2007). Human neutrophils uniquely release TIMP-free MMP-9 to provide a potent catalytic stimulator of angiogenesis. Pnas *104*, 20262–20267.

Asquith, B., Edwards, C.T.T., Lipsitch, M., and McLean, A.R. (2006). Inefficient cytotoxic T lymphocyte-mediated killing of HIV-1-infected cells in vivo. PLoS Biol. *4*, 583–592.

Assmann, G., Cullen, P., and Schulte, H. (2002). Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the Prospective Cardiovascular Münster (PROCAM) study. Circulation *105*, 310–315.

Auffray, C., and Hood, L. (2012). Editorial: Systems biology and personalized medicine - the future is now. Biotechnol J *7*, 938–939.

Auton, A., Abecasis, G.R., Altshuler, D.M., Durbin, R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Donnelly, P., Eichler, E.E., et al. (2015). A global reference for human genetic variation. Nature *526*, 68–74.

Awad, M., Setareh-Shenas, S., Robert Pixton, J., Soliman, C., Czer, L.S.C., Ruzza, A., and Mirocha, J. (2014). Prevalence of hypertension in the Gambia and Sierra Leone, western Africa: a cross-sectional study. Cardiovasc. J. Afr. *25*, 269–278.

Baardsnes, J., Hinck, C.S., Hinck, A.P., and O'Connor-McCourt, M.D. (2009). T??R-II discriminates the high- and low-affinity TGF-?? isoforms via two hydrogen-bonded ion Pairs. Biochemistry *48*, 2146–2155.

Backteman, K., Andersson, C., Dahlin, L.-G., Ernerudh, J., and Jonasson, L. (2012). Lymphocyte subpopulations in lymph nodes and peripheral blood: a comparison between patients with stable angina and acute coronary syndrome. PLoS One *7*, e32691.

Badimon, L. (2012). Interleukin-18: A potent pro-inflammatory cytokine in atherosclerosis. Cardiovasc. Res. *96*, 172–175.

Badimon, L., and Vilahur, G. (2014). Thrombosis formation on atherosclerotic lesions and plaque rupture. J. Intern. Med. *276*, 618–632.

Baek, M., Park, T., Heo, L., Park, C., and Seok, C. (2017). GalaxyHomomer: a web server for protein homo-

oligomer structure prediction from a monomer sequence or structure. Nucleic Acids Res. *40*, W294–W297.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. U. S. A. *98*, 10037–10041.

Bateman, A., Martin, M.J., O'Donovan, C., Magrane, M., Apweiler, R., Alpi, E., Antunes, R., Arganiska, J., Bely, B., Bingley, M., et al. (2015). UniProt: A hub for protein information. Nucleic Acids Res. *43*, D204–D212.

Bennett, M.R., Evan, G., and Schwartz, S.M. (1995). Apoptosis of Human Vascular Smooth Muscle Cells Derived from Normal Vessels and Coronary Atherosclerotic Plaques.

Bennett, M.R., Sinha, S., and Owens, G.K. (2016). Vascular Smooth Muscle Cells in Atherosclerosis. Circ. Res. *118*, 692–702.

Bennett, S.E., Schimerlik, M.I., and Mosbaugh, D.W. (1993). Kinetics of the uracil-DNA glycosylase/inhibitor protein association. Ung interaction with Ugi, nucleic acids, and uracil compounds. J. Biol. Chem. *268*, 26879–26885.

Benson, H., Watterson, S., Sharman, J., Mpamhanga, C., Parton, A., Southan, C., Harmar, A., and Ghazal, P. Is systems pharmacology ready to impact upon therapy development? A study on the cholesterol biosynthesis pathway. Br. J. Pharmacol. n/a-n/a.

Bentzon, J.F., Otsuka, F., Virmani, R., and Falk, E. (2014). Mechanisms of plaque formation and rupture. Circ. Res. *114*, 1852–1866.

Berg, E.L. (2014). Systems biology in drug discovery and development. Drug Discov. Today *19*, 113–125.

Bergmann, F.T., Hoops, S., Klahn, B., Kummer, U., Mendes, P., Pahle, J., and Sahle, S. (2017). COPASI and its applications in biotechnology. J. Biotechnol. *261*, 215–220.

Bhandari, B., Grandaliano, G., and Abboud, H.E. (1994). Platelet-derived growth factor (PDGF) BB homodimer regulates PDGF A- and PDGF B-chain gene transcription in human mesangial cells. Biochem. J. *297*, 385–388.

Bhattacharya, B.S., Sweby, P.K., Minihane, A.M., Jackson, K.G., and Tindall, M.J. (2014). A mathematical model of the sterol regulatory element binding protein 2 cholesterol biosynthesis pathway. J Theor Biol *349*, 150–162.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., et al. (2014). SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. *42*, 252–258.

von Birgelen, C., Mintz, G.S., de Vrey, E.A., Kimura, T., Popma, J.J., Airiian, S.G., Leon, M.B., Nobuyoshi, M., Serruys, P.W., and de Feyter, P.J. (1998). Atherosclerotic coronary lesions with inadequate compensatory enlargement have smaller plaque and vessel volumes: observations with three dimensional intravascular ultrasound in vivo. Heart *79*, 137–142.

Björnson, E., Mukhopadhyay, B., Asplund, A., Pristovsek, N., Cinar, R., Romeo, S., Uhlen, M., Kunos, G., Nielsen,

J., and Mardinoglu, A. (2015). Stratification of Hepatocellular Carcinoma Patients Based on Acetate Utilization. Cell Rep. *13*, 2014–2026.

Blum, A., and Miller, H.I. (1996). The role of inflammation in atherosclerosis. Isr J Med Sci *32*, 1059–1065.

Bobryshev, Y. V. (2006). Monocyte recruitment and foam cell formation in atherosclerosis. Micron *37*, 208–222.

Bobryshev, Y. V., Ivanova, E.A., Chistiakov, D.A., Nikiforov, N.G., and Orekhov, A.N. (2016). Macrophages and Their Role in Atherosclerosis: Pathophysiology and Transcriptome Analysis. Biomed Res. Int. *2016*.

Boden, W.E. (2000). High-density lipoprotein cholesterol as an independent risk factor in cardiovascular disease: assessing the data from framingham to the veterans affairs high-density lipoprotein intervention trial. Am. J. Cardiol. *86*, 19–22.

Boisvert, W.A. (2004). Modulation of atherogenesis by chemokines. Trends Cardiovasc. Med. *14*, 161–165.

Bonanno, E., Mauriello, A., Partenzi, A., Anemona, L., and Spagnoli, L.G. (2000). Flow cytometry analysis of atherosclerotic plaque cells from human carotids: A validation study. Cytometry *39*, 158–165.

Bosnić, Z., Vračar, P., Radovi, M.D., Devedžić, G., Filipović, N.D., and Kononenko, I. (2012). Mining data from hemodynamic simulations for generating prediction and explanation models. IEEE Trans. Inf. Technol. Biomed. *16*, 248–254.

Bosnjak, B., Stelzmueller, B., Erb, K.J., and Epstein, M.M. (2011). Treatment of allergic asthma: modulation of Th2 cells and their responses. Respir. Res. *12*, 114.

Bot, I., Bot, M., Van Heiningen, S.H., Van Santbrink, P.J., Lankhuizen, I.M., Hartman, P., Gruener, S., Hilpert, H., Van Berkel, T.J.C., Fingerle, J., et al. (2011). Mast cell chymase inhibition reduces atherosclerotic plaque progression and improves plaque stability in ApoE-/- mice. Cardiovasc. Res. *89*, 244–252.

Boussel, L., Arora, S., Rapp, J., Rutt, B., Huston, J., Parker, D., Yuan, C., Bassiouny, H., and Saloner, D. (2009). Atherosclerotic plaque progression in carotid arteries: monitoring with high-spatial-resolution MR imaging--multicenter trial. Radiology *252*, 789–796.

Brandl, R., Richter, T., Haug, K., Wilhelm, M.G., Maurer, P.C., and Nathrath, W. (1997). Topographic Analysis of Proliferative Activity in Carotid Endarterectomy Specimens by Immunocytochemical Detection of the Cell Cycle Related Antigen Ki-67. Circulation *96*, 3360–3368.

Brehme, M., Koschmieder, S., Montazeri, M., Copland, M., Oehler, V.G., Radich, J.P., Brümmendorf, T.H., and Schuppert, A. (2016). Combined Population Dynamics and Entropy Modelling Supports Patient Stratification in Chronic Myeloid Leukemia. Sci. Rep. *6*, 24057.

Briggs, G.E., and Haldane, J.B.S. (1925). A Further Note on the Kinetics of Enzyme Action. Biochem. J. *19*, 1037–1038.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. (1983). CHARMM: A

program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem. *4*, 187–217.

Bucher, J., Riedmaier, S., Schnabel, A., Marcus, K., Vacun, G., Weiss, T.S., Thasler, W.E., Nüssler, A.K., Zanger, U.M., and Reuss, M. (2011). A systems biology approach to dynamic modeling and inter-subject variability of statin pharmacokinetics in human hepatocytes. BMC Syst. Biol. *5*, 66.

Budu-Grajdeanu, P., Schugart, R.C., Friedman, A., Valentine, C., Agarwal, A.K., and Rovin, B.H. (2008). A mathematical model of venous neointimal hyperplasia formation. Theor. Biol. Med. Model. *5*, 2.

Bulelzai, M.A., and Dubbeldam, J.L. (2012). Long time evolution of atherosclerotic plaques. J Theor Biol *297*, 1–10.

Butticè, G., Miller, J., Wang, L., and Smith, B.D. (2006). Interferon-γ induces major histocompatibility class II transactivator (CIITA), which mediates collagen repression and major histocompatibility class II activation by human aortic smooth muscle cells. Circ. Res. *98*, 472–479.

Cai, J.M., Hatsukami, T.S., Ferguson, M.S., Small, R., Polissar, N.L., and Yuan, C. (2002). Classification of human carotid atherosclerotic lesions with in vivo multicontrast magnetic resonance imaging. Circulation *106*, 1368–1373.

Calvez, V., and Ebde, A. (2010). Mathematical modelling of the atherosclerotic plaque formation. 1–16.

Camejo, G., Fager, G., Rosengren, B., Hurt-Camejo, E., and Bondjers, G. (1993). Binding of low density lipoproteins by proteoglycans synthesized by proliferating and quiescent human arterial smooth muscle cells. J. Biol. Chem. *268*, 14131–14137.

Canault, M., Peiretti, F., Kopp, F., Bonardo, B., Bonzi, M.F., Coudeyre, J.C., Alessi, M.C., Juhan-Vague, I., and Nalbone, G. (2006). The TNF alpha converting enzyme (TACE/ADAM17) is expressed in the atherosclerotic lesions of apolipoprotein E-deficient mice: Possible contribution to elevated plasma levels of soluble TNF alpha receptors. Atherosclerosis *187*, 82–91.

Casa, L.D.C., and Ku, D.N. (2014). High Shear Thrombus Formation under Pulsatile and Steady Flow. Cardiovasc. Eng. Technol. *5*, 154–163.

Cenarro, A., Artieda, M., Castillo, S., Mozas, P., Reyes, G., Tejedor, D., Alonso, R., Mata, P., Pocoví, M., and Civiera, F. (2003). A common variant in the ABCA1 gene is associated with a lower risk for premature coronary heart disease in familial hypercholesterolaemia. J. Med. Genet. *40*, 163–168.

Chandonia, J.M., Fox, N.K., and Brenner, S.E. (2017). SCOPe: Manual Curation and Artifact Removal in the Structural Classification of Proteins ??? extended Database. J. Mol. Biol. *429*, 348–355.

Chaturvedi, N. (2003). ETHNIC DIFFERENCES IN CARDIOVASCULAR DISEASE. Heart *89*, 681–686.

Chege, P.M. (2016). Multiple cardiovascular disease risk factors in rural Kenya: evidence from a health and demographic surveillance system using the WHO STEP-wise approach to chronic disease risk factor surveillance. South African Fam. Pract. *58*, 54–61.

Cheifetz, S., Bassols, A., Stanley, K., Ohta, M., Greenberger, J., and Massagué, J. (1988). Heterodimeric transforming growth factor beta. Biological properties and interaction with three types of cell surface receptors. J. Biol. Chem. *263*, 10783–10789.

Chelliah, V., Juty, N., Ajmera, I., Ali, R., Dumousseau, M., Glont, M., Hucka, M., Jalowicki, G., Keating, S., Knight-Schrijver, V., et al. (2015). BioModels: ten-year anniversary. Nucleic Acids Res *43*, D542-8.

Chen, J., and Lu, X.Y. (2004). Numerical investigation of the non-Newtonian blood flow in a bifurcation model with a non-planar branch. J. Biomech. *37*, 1899–1911.

Chen, D., Roda, J.M., Marsh, C.B., Eubank, T.D., and Friedman, A. (2012). Hypoxia Inducible Factors-Mediated Inhibition of Cancer by GM-CSF: A Mathematical Model. Bull. Math. Biol. *74*, 2752–2777.

Chen, W.-W., Gao, R.-L., Liu, L.-S., Zhu, M.-L., Wang, W., Wang, Y.-J., Wu, Z.-S., LI, H.-J., GU, D.-F., YANG, Y.-J., et al. (2017). China cardiovascular diseases report 2015: a summary. J. Geriatr. Cardiol. *14*, 1–10.

Cheng, C., Van Haperen, R., De Waard, M., Van Damme, L.C.A., Tempel, D., Hanemaaijer, L., Van Cappellen, G.W.A., Bos, J., Slager, C.J., Duncker, D.J., et al. (2005). Shear stress affects the intracellular distribution of eNOS: Direct demonstration by a novel in vivo technique. Blood *106*, 3691–3698.

Cheng, C., Tempel, D., Van Haperen, R., Van Der Baan, A., Grosveld, F., Daemen, M.J.A.P., Krams, R., and De Crom, R. (2006). Atherosclerotic lesion size and vulnerability are determined by patterns of fluid shear stress. Circulation *113*, 2744–2753.

Chervitz, S.A., Deutsch, E.W., Field, D., Parkinson, H., Quackenbush, J., Rocca-serra, P., Sansone, S., Jr, C.J.S., Chris, F., Taylor, R., et al. (2011). Data Standards for Omics Data: The Basis of Data Sharing and Reuse.

Cheung, L.S.L., and Konstantopoulos, K. (2011). An analytical model for determining two-dimensional receptor-ligand kinetics. Biophys. J. *100*, 2338–2346.

Chiurlia, E., D'Amico, R., Ratti, C., Granata, A.R., Romagnoli, R., and Modena, M.G. (2005). Subclinical coronary artery atherosclerosis in patients with erectile dysfunction. J. Am. Coll. Cardiol. *46*, 1503–1506.

Chrysochou, C., and Kalra, P.A. (2010). Atherosclerotic renovascular disease and the heart. J. Ren. Care *36*, 146–153.

Chua, H.N., and Roth, F.P. (2011). Discovering the targets of drugs via computational systems biology. J Biol Chem *286*, 23653–23658.

Cilla, M., Peña, E., and Martínez, M. a (2014). Mathematical modelling of atheroma plaque formation and development in coronary arteries. J. R. Soc. Interface *11*, 20130866.

Cipriani, S., Francisci, D., Mencarelli, A., Renga, B., Schiaroli, E., D'Amore, C., Baldelli, F., and Fiorucci, S. (2013). Efficacy of the CCR5 Antagonist Maraviroc in Reducing Early, Ritonavir-Induced Atherogenesis and Advanced Plaque Progression in Mice<span hwp:id="article-t. Circulation *127*, 2114 LP-2124.

Cobbold, C.A., Sherratt, J.A., and Maxwell, S.R.J. (2002a). Lipoprotein oxidation and its significance for atherosclerosis: a mathematical approach. Bull Math Biol *64*, 65–95.

Cobbold, C.A., Sherratt, J.A., and Maxwell, S.R. (2002b). Lipoprotein oxidation and its significance for atherosclerosis: a mathematical approach. Bull Math Biol *64*, 65–95.

Cohen, A., Myerscough, M.R., and Thompson, R.S. (2014). Athero-protective effects of High Density Lipoproteins (HDL): An ODE model of the early stages of atherosclerosis. Bull Math Biol *76*, 1117–1142.

Colby, S.L., and Ortman, J.M. (2015). Projections of the size and composition of the US population: 2014 to 2060. Curr. Popul. Reports P25-1143.

Consortium, T.U. (2016). The UK10K project identifies rare variants in health and disease. *526*, 82–90.

Conway, P., Tyka, M.D., DiMaio, F., Konerding, D.E., and Baker, D. (2014). Relaxation of backbone bond geometry improves protein energy landscape modeling. Protein Sci. *23*, 47–55.

Cooper, A.D. (1997). Hepatic uptake of chylomicron remnants. J. Lipid Res. *38*, 2173–2192.

Cooper, C.J., Murphy, T.P., and Cutlip, D.E. (2014). Stenting and medical therapy for atherosclerotic renal-artery stenosis. J. Vasc. Surg. *59*, 873.

Cuchel, M., and Rader, D.J. (2006). Macrophage reverse cholesterol transport: Key to the regression of atherosclerosis? Circulation *113*, 2548–2555.

Cuellar, A.A., Lloyd, C.M., Nielsen, P.F., Bullivant, D.P., Nickerson, D.P., and Hunter, P.J. (2003). An overview of CellML 1.1, a biological model description language. Simulation-Transactions Soc. Model. Simul. Int. *79*, 740–747.

Cunningham, K.S., and Gotlieb, A.I. (2005). The role of shear stress in the pathogenesis of atherosclerosis. Lab. Invest. *85*, 9–23.

Cuomo, F., Roccabianca, S., Dillon-Murphy, D., Xiao, N., Humphrey, J.D., and Figueroa, C.A. (2017). Effects of age-associated regional changes in human central artery mechanics on systemic hemodynamics revealed by computational modeling. Am. J. Hypertens. 1–21.

Cushing, S.D., Berliner, J.A., Valente, A.J., Territo, M.C., Navab, M., Parhami, F., Gerrity, R., Schwartz, C.J., and Fogelman, A.M. (1990). Minimally modified low density lipoprotein induces monocyte chemotactic protein 1 in human endothelial cells and smooth muscle cells. Proc. Natl. Acad. Sci. U. S. A. *87*, 5134–5138.

Czauderna, T., Klukas, C., Schreiber, F., and Ideker, T. (2011). Editing, validating and translating of SBGN maps. Bioinformatics *27*, 2340–2341.

D'Agostino, R.B., Vasan, R.S., Pencina, M.J., Wolf, P.A., Cobain, M., Massaro, J.M., and Kannel, W.B. (2008). General cardiovascular risk profile for use in primary care: The Framingham heart study. Circulation *117*, 743–753.

Damsker, J.M., Hansen, A.M., and Caspi, R.R. (2010). Th1 and Th17 cells. *1183*, 211–221.

Davies, M.J. (2000). The pathophysiology of acute coronary syndromes. Heart *83*, 361 LP-366.

Davignon, J., and Ganz, P. (2004). Role of endothelial dysfunction in atherosclerosis. Circulation *109*, III27-32.

Davis, I.W., and Baker, D. (2009). RosettaLigand Docking with Full Ligand and Receptor Flexibility. J. Mol. Biol. *385*, 381–392.

Dedner, A., Klöfkorn, R., Nolte, M., and Ohlberger, M. (2010). A generic interface for parallel and adaptive discretization schemes: Abstraction principles and the Dune-Fem module. Comput. (Vienna/New York) *90*, 165–196.

Deeg, M., Baiyewu, O., Gao, S., Ogunniyi, A., Shen, J., Gureje, O., Taylor, S., Murrell, J., Unverzagt, F., Smith-Gamble, V., et al. (2008). A comparison of cardiovascular disease risk factor biomarkers in African Americans and Yoruba Nigerians. Ethn. Dis. *18*, 427–433.

Deepa, T.K., Binu, L.S., and Sukesh, a. K. (2009). Modelling Blood Flow and Analysis of Atherosclerotic Plaque Rupture under G-Force. 2009 3rd Int. Conf. Bioinforma. Biomed. Eng. 1–4.

Delgado-Roche, L., Brito, V., Acosta, E., Pérez, A., Fernández, J.R., Hernández-Matos, Y., Griñán, T., Soto, Y., León, O.S., Marleau, S., et al. (2015). Arresting progressive atherosclerosis by immunization with an anti-glycosaminoglycan monoclonal antibody in apolipoprotein E-deficient mice. Free Radic. Biol. Med. *89*, 557–566.

Demyanets, S., Konya, V., Kastl, S.P., Kaun, C., Rauscher, S., Niessner, A., Pentz, R., Pfaffenberger, S., Rychli, K., Lemberger, C.E., et al. (2011). Interleukin-33 induces expression of adhesion molecules and inflammatory activation in human endothelial cells and in human atherosclerotic plaques. Arterioscler. Thromb. Vasc. Biol. *31*, 2080–2089.

Dhooge, A., Govaerts, W., and Kuznetsov, Y.A. (2003). MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs. ACM Trans. Math. Softw. *29*, 141–164.

Díaz-zuccarini, V., Tomaso, G. Di, Agu, O., and Pichardo-almarza, C. (2014). Towards personalised management of atherosclerosis via computational models in vascular clinics : technology based on patient-speci fi c simulation approach. 1–6.

Dick, W., Zhu, C., Björkegren, J., Skogsberg, J., and Eriksson, P. (2011). MMP-2 and MMP-9 are prominent matrix metalloproteinases during atherosclerosis development in the Ldlr -/-Apob 100/100 mouse. Int. J. Mol. Med. *28*, 247–253.

van Dijk, R.A., Duinisveld, A.J.F., Schaapherder, A.F., Mulder-Stapel, A., Hamming, J.F., Kuiper, J., de Boer, O.J., van der Wal, A.C., Kolodgie, F.D., Virmani, R., et al. (2015). A change in inflammatory footprint precedes plaque instability: a systematic evaluation of cellular aspects of the adaptive immune response in human atherosclerosis. J. Am. Heart Assoc. *4*, e001403-.

Dinarello, C.A. (2009). Immunological and inflammatory functions of the interleukin-1 family. Annu. Rev. Immunol. *27*, 519–550.

Dolinsky, T.J., Czodrowski, P., Li, H., Nielsen, J.E., Jensen, J.H., Klebe, G., and Baker, N.A. (2007). PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res. *35*, 522–525.

Dorn, M., E Silva, M.B., Buriol, L.S., and Lamb, L.C. (2014). Three-dimensional protein structure prediction: Methods and computational strategies. Comput. Biol. Chem. *53*, 251–276.

Douglas Channon, KM., G. (2014). The pathogenesis of atherosclerosis. Medicine (Baltimore). *42*, 480–484.

Durbin, R.M., Altshuler, D.L., Durbin, R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Collins, F.S., De La Vega, F.M., Donnelly, P., et al. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061–1073.

von Eckardstein, a, Nofer, J.R., and Assmann, G. (2001). High density lipoproteins and arteriosclerosis. Role of cholesterol efflux and reverse cholesterol transport. Arterioscler. Thromb. Vasc. Biol. *21*, 13–27.

Eftaxiopoulos, D. a., and Atkinson, C. (2005). A nonlinear, anisotropic and axisymmetric model for balloon angioplasty. Proc. R. Soc. A Math. Phys. Eng. Sci. *461*, 1097–1128.

Elhage, R., Jawien, J., Rudling, M., Ljunggren, H., Takeda, K., Akira, S., and Bayard, F. (2003). R educed atherosclerosis in interleukin-18 deficient apolipoprotein E- knockout mice. Cardiovasc. Res. *59*, 234–240.

Elsawy, K.M., Twarock, R., Lane, D.P., Verma, C.S., and Caves, L.S.D. (2012). Characterization of the ligand receptor encounter complex and its potential for in silico kinetics-based drug development. J. Chem. Theory Comput. *8*, 314–321.

Engel, C.K., Pirard, B., Schimanski, S., Kirsch, R., Habermann, J., Klingler, O., Schlotte, V., Weithmann, K.U., and Wendt, K.U. (2005). Structural Basis for the Highly Selective Inhibition of MMP-13. Chem. Biol. *12*, 181–189.

Engele, M., Castiglione, K., Schwerdtner, N., Wagner, M., Bölcskei, P., Röllinghoff, M., and Stenger, S. (2002). Induction of TNF in Human Alveolar Macrophages As a Potential Evasion Mechanism of Virulent Mycobacterium tuberculosis. J. Immunol. *168*, 1328–1337.

Erijman, A., Rosenthal, E., and Shifman, J.M. (2014). How structure defines affinity in protein-protein interactions. PLoS One *9*.

Esebanmen, G.E., and Langridge, W.H.R. (2017). The role of TGF-beta signaling in dendritic cell tolerance. Immunol. Res. *65*, 987–994.

Ethier, C.R. (2002). Computational modeling of mass transfer and links to atherosclerosis. Ann. Biomed. Eng. *30*, 461–471.

Eussen, S.R., Rompelberg, C.J., Klungel, O.H., and van Eijkeren, J.C. (2011). Modelling approach to simulate reductions in LDL cholesterol levels after combined intake of statins and phytosterols/-stanols in humans.

Lipids Health Dis. *10*, 187.

Fabregat, A., Sidiropoulos, K., Garapati, P., Gillespie, M., Hausmann, K., Haw, R., Jassal, B., Jupe, S., Korninger, F., McKay, S., et al. (2016). The reactome pathway knowledgebase. Nucleic Acids Res. *44*, D481–D487.

Faratian, D., Goltsov, A., Lebedeva, G., Sorokin, A., Moodie, S., Mullen, P., Kay, C., Um, I.H., Langdon, S., Goryanin, I., et al. (2009). Systems biology reveals new strategies for personalizing cancer medicine and confirms the role of PTEN in resistance to trastuzumab. Cancer Res *69*, 6713–6720.

Farrar, M.A., and Schreiber, R.D. (1993). The Molecular Cell Biology of Interferon-gamma and its Receptor. Annu. Rev. Immunol. *11*, 571–611.

Federation, W.H. (2012). Urbanization and Cardiovascular Disease Fact Sheet.

Feig, J.E., Rong, J.X., Shamir, R., Sanson, M., Vengrenyuk, Y., Liu, J., Rayner, K., Moore, K., Garabedian, M., and Fisher, E.A. (2011). HDL promotes rapid atherosclerosis regression in mice and alters inflammatory properties of plaque monocyte-derived cells. Proc. Natl. Acad. Sci. U. S. A. *108*, 7166–7171.

Feret, J., Danos, V., Krivine, J., Harmer, R., and Fontana, W. (2009). Internal coarse-graining of molecular systems. Proc. Natl. Acad. Sci. U. S. A. *106*, 6453–6458.

Fernández-Alvira, J.M., Fuster, V., Dorado, B., Soberón, N., Flores, I., Gallardo, M., Pocock, S., Blasco, M.A., and Andrés, V. (2016). Short Telomere Load, Telomere Length, and Subclinical Atherosclerosis: The PESA Study. J. Am. Coll. Cardiol. *67*, 2467–2476.

Ferns, G.A., Raines, E.W., Sprugel, K.H., Motani, A.S., Reidy, M.A., and Ross, R. (1991). Inhibition of neointimal smooth muscle accumulation after angioplasty by an antibody to PDGF. Science (80-.). *253*, 1129 LP-1132.

Filipovic, N., and Kojic, M. (2004). Computer simulations of blood flow with mass transport through the carotid artery bifurcation. Theor. Appl. Mech. *31*, 1–33.

Filipovic, N., Meunier, N., Fotiadis, D., and Parodi, O. (2011a). Three-dimensional Numerical Simulation of Plaque Formation in the Arteries • Transport of LDL from the arterial lumen to the.

Filipovic, N., Fotiadis, D., Pelosi, W., and Parodi, O. (2011b). Experimental and computer model of plaque formation in the artery. 10th Int. Work. Biomed. Eng. BioEng 2011 1–4.

Filipovic, N., Teng, Z., Radovic, M., Saveljic, I., Fotiadis, D., and Parodi, O. (2013). Computer simulation of three-dimensional plaque formation and progression in the carotid artery. Med Biol Eng Comput *51*, 607–616.

Filipovic, N., Isailovic, V., Milosevic, Z., Nikolic, D., Saveljic, I., Radovic, M., Nikolic, M., Cirkovic-Andjelkovic, B., Themis, E., Fotiadis, D., et al. (2017). Computational modeling of plaque development in the coronary arteries. In CMBEBIH 2017: Proceedings of the International Conference on Medical and Biological Engineering 2017, A. Badnjevic, ed. (Singapore: Springer Singapore), pp. 269–274.

Finney, A., and Hucka, M. (2003). Systems biology markup language: Level 2 and beyond. Biochem Soc Trans *31*, 1472–1473.

Fiser, A. (2017). Comparative Protein Structure Modelling BT - From Protein Structure to Function with Bioinformatics. D. J. Rigden, ed. (Dordrecht: Springer Netherlands), pp. 91–134.

Fok, P.-W. (2012). Mathematical model of intimal thickening in atherosclerosis: vessel stenosis as a free boundary problem. J Theor Biol *314*, 23–33.

Franjkovic, I., Gessner, A., König, I., Kissel, K., Bohnert, A., Hartung, A., Ohly, A., Ziegler, A., Hackstein, H., and Bein, G. (2005). Effects of common atopy-associated amino acid substitutions in the IL-4 receptor alpha chain on IL-4 induced phenotypes. Immunogenetics *56*, 808–817.

Frank T. Bergmann, B.E.S. and M.H. (2015). SBML Software Summary.

Friedman, A., and Hao, W. (2014). A Mathematical Model of Atherosclerosis with Reverse Cholesterol Transport and Associated Risk Factors. Bull Math Biol.

Friedman, M.H., and Ehrlich, L.W. (1975). Effect of spatial variations in shear on diffusion at the wall of an arterial branch. Circ. Res. *37*, 446–454.

Frolkis, A., Knox, C., Lim, E., Jewison, T., Law, V., Hau, D.D., Liu, P., Gautam, B., Ly, S., Guo, A.C., et al. (2009). SMPDB: The small molecule pathway database. Nucleic Acids Res. *38*, 480–487.

Fujita, K.A., Ostaszewski, M., Matsuoka, Y., Ghosh, S., Glaab, E., Trefois, C., Crespo, I., Perumal, T.M., Jurkowski, W., Antony, P.M.A., et al. (2014). Integrating pathways of parkinson's disease in a molecular interaction map. Mol. Neurobiol. *49*, 88–102.

Funahashi, A., Morohashi, M., Kitano, H., and Tanimura, N. (2003). CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. BIOSILICO *1*, 159–162.

Funahashi, A., Matsuoka, Y., Jouraku, A., Morohashi, M., Kikuchi, N., and Kitano, H. (2008). CellDesigner 3.5: A versatile modeling tool for biochemical networks. Proc. IEEE *96*, 1254–1265.

Funk, C.D. (2006). Lipoxygenase pathways as mediators of early inflammatory events in atherosclerosis. Arterioscler. Thromb. Vasc. Biol. *26*, 1204–1206.

Fuster, V. (1995). Elucidation off the role of plaque instability and rupture in acute coronary events. Am. J. Cardiol. *76*.

Gabriel, S.A., Ding, Y., Feng, Y., and Gear, J.A. (2014). Deposition-driven Growth in Atherosclerosis Modelling. 2–5.

Gadkar, K., Budha, N., Baruch, a, Davis, J.D., Fielder, P., and Ramanujan, S. (2014). A Mechanistic Systems Pharmacology Model for Prediction of LDL Cholesterol Lowering by PCSK9 Antagonism in Human Dyslipidemic Populations. CPT Pharmacometrics Syst. Pharmacol. *3*, e149.

Garzon, J.I., Lopéz-Blanco, J.R., Pons, C., Kovacs, J., Abagyan, R., Fernandez-Recio, J., and Chacon, P. (2009). FRODOCK: A new approach for fast rotational protein-protein docking. Bioinformatics *25*, 2544–2551.

Gately, M.K., Renzetti, L.M., Magram, J., Stern, A.S., Adorini, L., Gubler, U., and Presky, D.H. (1998). THE INTERLEUKIN-12/INTERLEUKIN-12-RECEPTOR SYSTEM: Role in Normal and Pathologic Immune Responses. Annu. Rev. Immunol. *16*, 495–521.

Geloen, A., Helin, L., Geeraert, B., Malaud, E., Holvoet, P., and Marguerie, G. (2012). CD36 inhibitors reduce postprandial hypertriglyceridemia and protect against diabetic dyslipidemia and atherosclerosis. PLoS One *7*, 1–12.

Gerdes, N., Sukhova, G.K., Libby, P., Reynolds, R.S., Young, J.L., and Schönbeck, U. (2002). Expression of Interleukin (IL)-18 and Functional IL-18 Receptor on Human Vascular Endothelial Cells, Smooth Muscle Cells, and Macrophages: Implications for Atherogenesis. J. Exp. Med. *195*, 245–257.

Gessaghi, V.C., Raschi, M. a., Tanoni, D.Y., Perazzo, C. a., and Larreteguy, A.E. (2011). Growth model for cholesterol accumulation in the wall of a simplified 3D geometry of the carotid bifurcation. Comput. Methods Appl. Mech. Eng. *200*, 2117–2125.

Getz, G.S., and Reardon, C.A. (2012). Animal models of Atherosclerosis. Arterioscler. Thromb. Vasc. Biol. *32*, 1104–1115.

Gibson, D.S., Drain, S., Kelly, C., McGilligan, V., McClean, P., Atkinson, S.D., Murray, E., McDowell, A., Conway, C., Watterson, S., et al. (2017). Coincidence versus consequence: opportunities in multi-morbidity research and inflammation as a pervasive feature. Expert Rev. Precis. Med. Drug Dev. *2*, 147–156.

Girke, S., Klofkorn, R., and Ohlberger, M. (2014). Efficient Parallel Simulation of Atherosclerotic Plaque Formation Using Higher Order Discontinuous Galerkin Schemes.

Glagov, S., Zarins, C., Giddens, D.P., and Ku, D.N. (1988). Hemodynamics and atherosclerosis. Insights and perspectives gained from studies of human arteries. Arch. Pathol. Lab. Med. *112*, 1018–1031.

Go, A.S., Mozaffarian, D., Roger, V.L., Benjamin, E.J., Berry, J.D., Blaha, M.J., Dai, S., Ford, E.S., Fox, C.S., Franco, S., et al. (2014). Heart disease and stroke statistics--2014 update: a report from the American Heart Association.

Goldberg, I.J. (1996). Lipoprotein lipase and lipolysis: central roles in lipoprotein metabolism and atherogenesis. J. Lipid Res. *37*, 693–707.

Gomez-Cabrero, D., Compte, A., and Tegner, J. (2011). Workflow for generating competing hypothesis from models with parameter uncertainty. Interface Focus *1*, 438–449.

Gonçalves, I., Moses, J., Dias, N., Pedro, L.M., Fernandes e Fernandes, J., Nilsson, J., and Ares, M.P.S. (2003). Changes related to age and cerebrovascular symptoms in the extracellular matrix of human carotid plaques. Stroke. *34*, 616–622.

Gonzalez-garay, M.L. (2015). The road from next-generation sequencing to personalized medicine. *11*, 523–544.

Goriely, A., and Vandiver, R. (2010). On the mechanical stability of growing arteries. IMA J. Appl. Math. (Institute Math. Its Appl. *75*, 549–570.

Gotto, A.M. (1998). Triglyceride as a risk factor for coronary artery disease. Am J Cardiol *82*, 25.

Grainger, D.J. (2007). TGF-?? and atherosclerosis in man. Cardiovasc. Res. *74*, 213–222.

Gray, K.A., Yates, B., Seal, R.L., Wright, M.W., and Bruford, E.A. (2015). Genenames.org: The HGNC resources in 2015. Nucleic Acids Res. *43*, D1079–D1085.

Green, J., Waters, S., Cummings, L., van den Berg, J., Siggers, J., and Grief, a (2002). Atherosclerotic plaque rupture.

Gruber, M.F., and Gerrard, T.L. (1992). Production of macrophage colony-stimulating factor (M-CSF) by human monocytes is differentially regulated by GM-CSF, TNF??, and IFN-?? Cell. Immunol. *142*, 361–369.

Grufman, H., Schiopu, A., Edsfeldt, A., Bj??rkbacka, H., Nitulescu, M., Nilsson, M., Persson, A., Nilsson, J., and Gon??alves, I. (2014). Evidence for altered inflammatory and repair responses in symptomatic carotid plaques from elderly patients. Atherosclerosis *237*, 177–182.

Grundy, S.M., Benjamin, I.J., Burke, G.L., Chait, A., Eckel, R.H., Howard, B. V, Mitch, W., Smith, S.C., and Sowers, J.R. (1999). Diabetes and cardiovascular disease: a statement for healthcare professionals from the American Heart Association. Circulation *100*, 1134–1146.

Grzegorczyk, M., Husmeier, D., Edwards, K.D., Ghazal, P., and Millar, A.J. (2008). Modelling non-stationary gene regulatory processes with a non-homogeneous Bayesian network and the allocation sampler. Bioinformatics *24*, 2071–2078.

Guerriero, M.L., Prandi, D., Priami, C., and Quaglia, P. (2009). Process Calculi Abstractions for Biology. In Algorithmic Bioprocesses, pp. 463–486.

Gupta, M., and Brister, S. (2006). Is South Asian ethnicity an independent cardiovascular risk factor? Can. J. Cardiol. *22*, 193–197.

Gupta, S., Pablo, A.M., Jiang, X. c, Wang, N., Tall, A.R., and Schindler, C. (1997). IFN-gamma potentiates atherosclerosis in ApoE knock-out mice. J. Clin. Invest. *99*, 2752–2761.

Guy, R.D., Fogelson, A.L., and Keener, J.P. (2007). Fibrin gel formation in a shear flow. Math. Med. Biol. *24*, 111–130.

Guyton, J.R., and Klemp, K.F. (1996). Development of the lipid-rich core in human atherosclerosis. Arterioscler. Thromb. Vasc. Biol. *16*, 4–11.

György M. Keserü, and David C. Swinney (2015). Chapter 1: Thermodynamics and Kinetics of Drug Binding.

Haas, J., Roth, S., Arnold, K., Kiefer, F., Schmidt, T., Bordoli, L., and Schwede, T. (2013). The protein model portal - A comprehensive resource for protein structure and model information. Database *2013*, 1–9.

Hahnefeld, C., Drewianka, S., and Herberg, F.W. (2004). Determination of kinetic data using surface plasmon resonance biosensors. Methods Mol. Med. *94*, 299–320.

Hammacher, a. (1996). Influence of Interleukin-6 (IL-6) Dimerization on Formation of the High Affinity Hexameric IL-6 Receptor Complex. J. Biol. Chem. *271*, 20138–20144.

Hansen, R.J., Berna, M.J., Sperry, A.E., Beyer, T.P., Wroblewski, V.J., Schroeder, K.M., and Eacho, P.I. (2017). Quantitative characterization of the mechanism of action and impact of a "proteolysis-permitting" anti-PCSK9 antibody. MAbs *9*, 285–296.

Hansson, G. (2005). Inflammation, atherosclerosis, and coronary artery disease. N. Engl. J. Med. *353*, 1685–1695.

Hao, W., and Friedman, A. (2014). The LDL-HDL profile determines the risk of atherosclerosis: a mathematical model. PLoS One *9*, e90497.

Hao, W., Gong, S., Wu, S., Xu, J., Go, M.R., Friedman, A., and Zhu, D. (2017). A mathematical model of aortic aneurysm formation. PLoS One *12*.

Hardin, C., Pogorelov, T. V., and Luthey-Schulten, Z. (2002). Ab initio protein structure prediction. Curr. Opin. Struct. Biol. *12*, 176–181.

Harrington, J.R. (2000). The role of MCP-1 in atherosclerosis. Stem Cells *18*, 65–66.

Harrison, D., Griendling, K.K., Landmesser, U., Hornig, B., and Drexler, H. (2003). Role of oxidative stress in atherosclerosis. Am. J. Cardiol. *91*, 7A–11A.

Hartwig, H., Silvestre Roig, C., Daemen, M., Lutgens, E., and Soehnlein, O. (2014). Neutrophils in atherosclerosis. A brief overview. Hamostaseologie *35*.

Hauer, A.D., Uyttenhove, C., De Vos, P., Stroobant, V., Renauld, J.C., Van Berkel, U.C., Van Snick, J., and Kuiper, J. (2005). Blockade of interleukin-12 function by protein vaccination attenuates atherosclerosis. Circulation *112*, 1054–1062.

Hecht, F. (2012). New development in freefem+. J. Numer. Math. *20*, 251–265.

Heidenreich, P.A., Trogdon, J.G., Khavjou, O.A., Butler, J., Dracup, K., Ezekowitz, M.D., Finkelstein, E.A., Hong, Y., Johnston, S.C., Khera, A., et al. (2011). Forecasting the future of cardiovascular disease in the United States: A policy statement from the American Heart Association. Circulation *123*, 933–944.

Heinecke, J.W. (2006). Lipoprotein oxidation in cardiovascular disease: chief culprit or innocent bystander? J. Exp. Med. *203*, 813–816.

Heise, C.E., Pahuja, a, Hudson, S.C., Mistry, M.S., Putnam, a L., Gross, M.M., Gottlieb, P. a, Wade, W.S., Kiankarimi, M., Schwarz, D., et al. (2005). Pharmacological characterization of CXC chemokine receptor 3 ligands and a small molecule antagonist. J Pharmacol Exp Ther *313*, 1263–1271.

Heo, L., Park, H., and Seok, C. (2013). GalaxyRefine: Protein structure refinement driven by side-chain repacking. Nucleic Acids Res. *41*, 384–388.

Heo, L., Lee, H., and Seok, C. (2016). GalaxyRefineComplex: Refinement of protein-protein complex model structures driven by interface repacking. Sci. Rep. *6*, 32153.

Herbertz, S., Sawyer, J.S., Stauber, A.J., Gueorguieva, I., Driscoll, K.E., Estrem, S.T., Cleverly, A.L., Desaiah, D., Guba, S.C., Benhadji, K.A., et al. (2015). Clinical development of galunisertib (Ly2157299 monohydrate), a small molecule inhibitor of transforming growth factor-beta signaling pathway. Drug Des. Devel. Ther. *9*, 4479–4499.

Herder, C., Peeters, W., Zierer, A., de Kleijn, D.P. V, Moll, F.L., Karakas, M., Roden, M., Meisinger, C., Thorand, B., Pasterkamp, G., et al. (2012). TGF-??1 content in atherosclerotic plaques, TGF-??1 serum concentrations and incident coronary events. Eur. J. Clin. Invest. *42*, 329–337.

Hoff, H., Gaubatz, J.W., and Gotto Jr, A.M. (1978). Apo B concentration in the normal human aorta. Biochem. Biophys. Res. Commun. *85*, 1424–1430.

Holzapfel, G. a, Gasser, T.C., and Ogden, R.W. (2000). A new constitutive framework for arterial wall mechanics and a comperative study of material models. J. Elast. *61*, 1–48.

Hong, Y.M. (2010). Atherosclerotic cardiovascular disease beginning in childhood. Korean Circ. J. *40*, 1–9.

Hood, L. (2013). Systems Biology and P4 Medicine: Past, Present, and Future. Rambam Maimonides Med. J. *4*, e0012.

Hu, L., van der Hoogt, C.C., Espirito Santo, S.M.S., Out, R., Kypreos, K.E., van Vlijmen, B.J.M., Van Berkel, T.J.C., Romijn, J. a, Havekes, L.M., van Dijk, K.W., et al. (2008). The hepatic uptake of VLDL in lrp-ldlr-/-vldlr-/- mice is regulated by LPL activity and involves proteoglycans and SR-BI. J. Lipid Res. *49*, 1553–1561.

Huang, Y., Rumschitzki, D., Chien, S., and Weinbaum, S. (1997). A fiber matrix model for the filtration through fenestral pores in a compressible arterial intima. Am. J. Physiol. *272*, H2023–H2039.

Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., Kitano, H., Arkin, A.P., Bornstein, B.J., Bray, D., Cornish-Bowden, A., et al. (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics *19*, 524–531.

Hussein, A.A., Uno, K., Wolski, K., Kapadia, S., Schoenhagen, P., Tuzcu, E.M., Nissen, S.E., and Nicholls, S.J. (2011). Peripheral arterial disease and progression of coronary atherosclerosis. J. Am. Coll. Cardiol. *57*, 1220–1225.

Iasiello, M., Vafai, K., Andreozzi, A., and Bianco, N. (2016). Low-density lipoprotein transport through an arterial wall under hyperthermia and hypertension conditions - An analytical solution. J. Biomech. *49*, 193–204.

Ibragimov, A.I., McNeal, C.J., Ritter, L.R., and Walton, J.R. (2005). A mathematical model of atherogenesis as an

inflammatory response. Math Med Biol *22*, 305–333.

Ichiro, T., Tajima, S., and Nishikawa, T. (1990). Prederential Inhibition of Elastin Synthesis by Epidermal Growth Factor in Chick Aortic Smooth Muscle Cells. Biochem. Biophys. Res. Commun. *168*, 850–856.

Van Iersel, M.P., Villéger, A.C., Czauderna, T., Boyd, S.E., Bergmann, F.T., Luna, A., Demir, E., Sorokin, A., Dogrusoz, U., Matsuoka, Y., et al. (2012). Software support for SBGN maps: SBGN-ML and LibSBGN. Bioinformatics *28*, 2016–2021.

Ilari, A., and Savino, C. (2008). Protein Structure Determination by X-ray Crystallography.

International Diabetes Federation (2013). IDF Diabetes Atlas, 6th Edn.

Di Iorio, A., Ferrucci, L., Sparvieri, E., Cherubini, A., Volpato, S., Corsi, A., Bonafè, M., Franceschi, C., Abate, G., and Paganelli, R. (2003). Serum IL-1?? levels in health and disease: A population-based study. "The InCHIANTI study." Cytokine *22*, 198–205.

Isara, A.R., and Okundia, P.O. (2015). The burden of hypertension and diabetes mellitus in rural communities in southern Nigeria. Pan Afr. Med. J. *20*, 103.

Ishibashi, H., Sunamura, M., and Karino, T. (1995). Flow patterns and preferred sites of intimal thickening in end-to-end anastomosed vessels. Surgery *117*, 409–420.

Jansen, M., Pfaffelhuber, P., Hoffmann, M.M., Puetz, G., and Winkler, K. (2016). In silico modeling of the dynamics of low density lipoprotein composition via a single plasma sample. J. Lipid Res. *57*, 882–893.

Johnsen, S.H., Fosse, E., Joakimsen, O., Mathiesen, E.B., Stensland-Bugge, E., Njølstad, I., and Arnesen, E. (2005). Monocyte count is a predictor of novel plaque formation: A 7-year follow-up study of 2610 persons without carotid plaque at baseline the Troms?? study. Stroke *36*, 715–719.

Johnson, K., and Goody, R. (2012). The Original Michaelis Constant. Biochemistry *50*, 8264–8269.

Johnson, A.D., Yanek, L.R., Chen, M., Faraday, N., Martin, G., Tofler, G., Lin, S.J., Kraja, A.T., Province, M. a, Becker, D.M., et al. (2011). Genome-wide meta-analyses identifies 7 loci associated with platelet aggregation in response to agonists. *42*, 608–613.

Johnston, B.M., Johnston, P.R., Corney, S., and Kilpatrick, D. (2006). Non-Newtonian blood flow in human right coronary arteries: Transient simulations. J. Biomech. *39*, 1116–1128.

Jorgensen, W.L. (1991). Rusting of the lock and key model for protein-ligand binding. Science *254*, 954–955.

Joris, I., and Majno, G. (1978). Atherosclerosis and inflammation. Adv Exp Med Biol *104*, 227–243.

Joseph, A., Ackerman, D., Talley, J.D., Johnstone, J., and Kupersmith, J. (1993). Manifestations of coronary atherosclerosis in young trauma victims-An autopsy study. J. Am. Coll. Cardiol. *22*, 459–467.

Jousilahti, P., Laatikainen, T., Peltonen, M., Borodulin, K., Männistö, S., Jula, A., Salomaa, V., Harald, K., Puska, P., and Vartiainen, E. (2016). Primary prevention and risk factor reduction in coronary heart disease mortality

among working aged men and women in eastern Finland over 40 years: population based observational study. Bmj i721.

Kaazempur-Mofrad, M.R., and Ethier, C.R. (2001). Mass transport in an anatomically realistic human right coronary artery. Ann. Biomed. Eng. *29*, 121–127.

Kalow, W. (2002). Pharmacogenetics and personalised medicine. Fundam. Clin. Pharmacol. *16*, 337–342.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res *42*, 199–205.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. *45*, D353–D361.

Kao, J., Kobashigawa, J., Fishbein, M.C., Robb MacLellan, W., Burdick, M.D., Belperio, J.A., and Strieter, R.M. (2003). Elevated serum levels of the CXCR3 chemokine ITAC are associated with the development of transplant coronary artery disease. Circulation *107*, 1958–1961.

Karr, J.R., Sanghvi, J.C., Macklin, D.N., Gutschow, M. V, Jacobs, J.M., Bolival Jr., B., Assad-Garcia, N., Glass, J.I., and Covert, M.W. (2012a). A whole-cell computational model predicts phenotype from genotype. Cell *150*, 389–401.

Karr, J.R., Sanghvi, J.C., MacKlin, D.N., Gutschow, M. V., Jacobs, J.M., Bolival, B., Assad-Garcia, N., Glass, J.I., and Covert, M.W. (2012b). A whole-cell computational model predicts phenotype from genotype. Cell *150*, 389–401.

Kastritis, P.L., Moal, I.H., Hwang, H., Weng, Z., Bates, P.A., Bonvin, A.M.J.J., and Janin, J. (2011). A structure-based benchmark for protein-protein binding affinity. Protein Sci. *20*, 482–491.

Kastritis, P.L., Rodrigues, J.P.G.L.M., Folkers, G.E., Boelens, R., and Bonvin, A.M.J.J. (2014). Proteins feel more than they see: Fine-tuning of binding affinity by properties of the non-interacting surface. J. Mol. Biol. *426*, 2632–2652.

Kathiresan, S., Reilly, M.P., Samani, N.J., Schunkert, H., Erdmann, J., Assimes, T.L., Boerwinkle, E., Hall, A., Hengstenberg, C., K??nig, I.R., et al. (2011). RANTES/CCL5 and risk for coronary events: Results from the MONICA/KORA Augsburg case-cohort, Athero-express and CARDIoGRAM studies. PLoS One *6*.

Kato, Z., Jee, J., Shikano, H., Mishima, M., Ohki, I., Ohnishi, H., Li, A., Hashimoto, K., Matsukuma, E., Omoya, K., et al. (2003). The structure and binding mode of interleukin-18. Nat Struct Mol Biol *10*, 966–971.

Katoh, E., Louis, J.M., Yamazaki, T., Gronenborn, A.M., Torchia, D. a, and Ishima, R. (2003). A solution NMR study of the binding kinetics and the internal dynamics of an HIV-1 protease-substrate complex. Protein Sci. *12*, 1376–1385.

Katona, I.M., Urban Jr., J.F., Kang, S.S., Paul, W.E., and Finkelman, F.D. (1991). IL-4 requirements for the generation of secondary in vivo IgE responses. J. Immunol. *146*, 4215–4221.

Khatib, F., Cooper, S., Tyka, M.D., Xu, K., Makedon, I., Popovic, Z., Baker, D., and Players, F. (2011). Algorithm discovery by protein folding game players. Proc. Natl. Acad. Sci. U. S. A. *108*, 18949–18953.

El Khatib, N., Génieys, S., and Volpert, V. (2007). Atherosclerosis Initiation Modeled as an Inflammatory Process. Math. Model. Nat. Phenom. *2*, 126–141.

Kilic, A., and Mandal, K. (2012). Heat shock proteins: pathogenic role in atherosclerosis and potential therapeutic implications. Autoimmune Dis *2012*, 502813.

Kim, Y., and Friedman, A. (2010). Interaction of Tumor with Its Micro-environment: A Mathematical Model. Bull. Math. Biol. *72*, 1029–1068.

Kinnunen, P.K.J., and Holopainen, J.M. (2002). Sphingomyelinase Activity of LDL: A Link between Atherosclerosis, Ceramide, and Apoptosis? Trends Cardiovasc. Med. *12*, 37–42.

Kirouac, D.C., Du, J.Y., Lahdenranta, J., Overland, R., Yarar, D., Paragas, V., Pace, E., McDonagh, C.F., Nielsen, U.B., and Onsum, M.D. (2013). Computational Modeling of ERBB2-Amplified Breast Cancer Identifies Combined ErbB2/3 Blockade as Superior to the Combination of MEK and AKT Inhibitors. Sci. Signal. *6*, ra68-ra68.

Kirpalani, A., Park, H., Butany, J., Johnston, K.W., and Ojha, M. (1999). Velocity and wall shear stress patterns in the human right coronary artery. J. Biomech. Eng. *121*, 370–375.

Kolodgie, F.D., Gold, H.K., Burke, A.P., Fowler, D.R., Kruth, H.S., Weber, D.K., Farb, A., Guerrero, L.J., Hayase, M., Kutys, R., et al. (2003). Intraplaque hemorrhage and progression of coronary atheroma. N. Engl. J. Med. *349*, 2316–2325.

König, M., Dräger, A., and Holzhütter, H.-G. (2012). CySBML: a Cytoscape plugin for SBML. Bioinformatics *28*, 2402–2403.

Koskivirta, I., Rahkonen, O., Mäyränpää, M., Pakkanen, S., Husheem, M., Sainio, A., Hakovirta, H., Laine, J., Jokinen, E., Vuorio, E., et al. (2006). Tissue inhibitor of metalloproteinases 4 (TIMP4) is involved in inflammatory processes of human cardiovascular pathology. Histochem. Cell Biol. *126*, 335–342.

Kraemer, R. (2000). Regulation of cell migration in atherosclerosis. Curr. Atheroscler. Rep. *2*, 445–452.

Krumm, B., Meng, X., Xiang, Y., and Deng, J. (2017). Identification of small molecule inhibitors of Interleukin-18. Sci. Rep. *7*, 483.

Kufareva, I., and Abagyan, R. (2012). Homology Modeling. *857*, 231–257.

Kutmon, M., van Iersel, M.P., Bohler, A., Kelder, T., Nunes, N., Pico, A.R., and Evelo, C.T. (2015). PathVisio 3: An Extendable Pathway Analysis Toolbox. PLoS Comput. Biol. *11*, 1–13.

Kutmon, M., Riutta, A., Nunes, N., Hanspers, K., Willighagen, E.L., Bohler, A., Mélius, J., Waagmeester, A., Sinha, S.R., Miller, R., et al. (2016). WikiPathways: Capturing the full diversity of pathway knowledge. Nucleic Acids Res. *44*, D488–D494.

Kwiatkowska, M.Z., and Heath, J.K. (2009). Biological pathways as communicating computer systems. J. Cell Sci. *122*, 2793–2800.

Lafont, A. (2003). Basic aspects of plaque vulnerability. Heart *89*, 1262–1267.

Lambrecht, B.N., and Hammad, H. (2015). The immunology of asthma. Nat Immunol *16*, 45–56.

Landrum, M.J., Lee, J.M., Riley, G.R., Jang, W., Rubinstein, W.S., Church, D.M., and Maglott, D.R. (2014). ClinVar: Public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. *42*, 980–985.

Lauer-Fields, J.L., Juska, D., and Fields, G.B. (2002). Matrix metalloproteinases and collagen catabolism. Biopolym. - Pept. Sci. Sect. *66*, 19–32.

Lauzier, M.-C., Robitaille, G.A., Chan, D.A., Giaccia, A.J., and Richard, D.E. (2008). (2R)-[(4-Biphenylylsulfonyl)amino]-N-hydroxy-3-phenylpropionamide (BiPS), a matrix metalloprotease inhibitor, is a novel and potent activator of hypoxia-inducible factors. Mol. Pharmacol. *74*, 282–288.

Le, N.A., and Walter, M.F. (2007). The role of hypertriglyceridemia in atherosclerosis. Curr. Atheroscler. Rep. *9*, 110–115.

Lee, Y.W., and Hirani, A.A. (2006). Role of interleukin-4 in atherosclerosis. Arch. Pharm. Res. *29*, 1–15.

Lee, T.S., Yen, H.C., Pan, C.C., and Chau, L.Y. (1999). The role of interleukin 12 in the development of atherosclerosis in ApoE-deficient mice. Arter. Thromb Vasc Biol *19*, 734–742.

Lemieux, I., Lamarche, B., Couillard, C., Pascot, a, Cantin, B., Bergeron, J., Dagenais, G.R., and Després, J.P. (2015). Total cholesterol/HDL cholesterol ratio vs LDL cholesterol/HDL cholesterol ratio as indices of ischemic heart disease risk in men: the Quebec Cardiovascular Study. Arch. Intern. Med. *161*, 2685–2692.

Lendon, C., Born, G. V, Davies, M.J., and Richardson, P.D. (1992). Plaque fissure: the link between atherosclerosis and thrombosis. Nouv. Rev. Fr. Hematol. *34*, 27–29.

Lensink, M.F., Velankar, S., Kryshtafovych, A., Huang, S.Y., Schneidman-Duhovny, D., Sali, A., Segura, J., Fernandez-Fuentes, N., Viswanath, S., Elber, R., et al. (2016). Prediction of homoprotein and heteroprotein complexes by protein docking and template-based modeling: A CASP-CAPRI experiment. Proteins Struct. Funct. Bioinforma. 323–348.

Levitzki, A., and Gazit, A. (1995). Tyrosine kinase inhibition: an approach to drug development. Science *267*, 1782–1788.

Lewis, N.E., Schramm, G., Bordbar, A., Schellenberger, J., Andersen, M.P., Cheng, J.K., Patel, N., Yee, A., Lewis, R.A., Eils, R., et al. (2010). Large-scale in silico modeling of metabolic interactions between cell types in the human brain. Nat Biotechnol *28*, 1279–1285.

Li, C., Donizelli, M., Rodriguez, N., Dharuri, H., Endler, L., Chelliah, V., Li, L., He, E., Henry, A., Stefan, M.I., et al. (2010). BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic

models. BMC Syst Biol *4*, 92.

Li, Z.Y., Howarth, S.P.S., Tang, T., and Gillard, J.H. (2006). How critical is fibrous cap thickness to carotid plaque stability? A flow-plaque interaction model. Stroke *37*, 1195–1199.

Libby, P. (2002). Inflammation and Atherosclerosis. Circulation *105*, 1135–1143.

Libby, P., Lichtman, A., and Hansson, G. (2013). Immune Effector Mechanisms Implicated in Atherosclerosis: From Mice to Humans. Immunity *38*, 1092–1104.

Lichtman, A.H., Chin, J., Schmidt, J.A., and Abbas, A.K. (1988). Role of interleukin 1 in the activation of T lymphocytes. Proc. Natl. Acad. Sci. U. S. A. *85*, 9699–9703.

Lievens, D., and von Hundelshausen, P. (2011). Platelets in atherosclerosis. Thromb. Haemost. *106*, 827–838.

Lima, J.A.C., Desai, M.Y., Steen, H., Warren, W.P., Gautam, S., and Lai, S. (2004). Statin-Induced Cholesterol Lowering and Plaque Regression After 6 Months of Magnetic Resonance Imaging–Monitored Therapy. Circulation *110*, 2336 LP-2341.

Little, M.P., Gola, A., and Tzoulaki, I. (2009). A model of cardiovascular disease giving a plausible mechanism for the effect of fractionated low-dose ionizing radiation exposure. PLoS Comput. Biol. *5*.

Liu, B., and Tang, D. (2010). Computer Simulations of Atherosclerotic Plaque Growth in Coronary Arteries. Mol Cell Biomech *7*, 193–202.

Liu, S., Song, X., Chrunyk, B.A., Shanker, S., Hoth, L.R., Marr, E.S., and Griffor, M.C. (2013). Crystal structures of interleukin 17A and its complex with IL-17 receptor A. *4*, 1888.

Lloyd-Jones, D., Adams, R.J., Brown, T.M., Carnethon, M., Dai, S., De Simone, G., Ferguson, T.B., Ford, E., Furie, K., Gillespie, C., et al. (2010). Executive summary: Heart disease and stroke statistics-2010 update: A report from the american heart association. Circulation *121*.

Lonergan, M., Senn, S.J., McNamee, C., Daly, A.K., Sutton, R., Hattersley, A., Pearson, E., and Pirmohamed, M. (2017). Defining drug response for stratified medicine. Drug Discov. Today *22*, 173–179.

Loppnow, H., and Libby, P. (1990). Proliferating or interleukin 1-activated human vascular smooth muscle cells secrete copious interleukin 6. J. Clin. Invest. *85*, 731–738.

Lovren, F., Pan, Y., Quan, A., Singh, K.K., Shukla, P.C., Gupta, N., Steer, B.M., Ingram, A.J., Gupta, M., Al-Omran, M., et al. (2012). MicroRNA-145 targeted therapy reduces atherosclerosis. Circulation *126*.

Lu, H., Talbot, S., Robertson, K.A., Watterson, S., Forster, T., Roy, D., and Ghazal, P. (2015). Rapid proteasomal elimination of 3-hydroxy-3-methylglutaryl-CoA reductase by interferon-gamma in primary macrophages requires endogenous 25-hydroxycholesterol synthesis. Steroids.

Lu, X., Wang, L., Chen, S., He, L., Yang, X., Shi, Y., Cheng, J., Zhang, L., Charles Gu, C., Huang, J., et al. (2012). Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery

disease. Nat. Genet. *44*, 890–894.

Lukin, J.A., Kontaxis, G., Simplaceanu, V., Yuan, Y., Bax, A., and Ho, C. (2003). Quaternary structure of hemoglobin in solution. Proc. Natl. Acad. Sci. U. S. A. *100*, 517–520.

Lusis, A.J. (2000). Atherosclerosis. Nature *407*, 233–241.

M. Mooney, K., and T. Mc Auley, M. (2015). Cardiovascular disease and healthy ageing. J. Integr. Cardiol. *1*, 76–78.

Machné, R., Finney, A., Müller, S., Lu, J., Widder, S., and Flamm, C. (2006). The SBML ODE Solver Library: A native API for symbolic and fast numerical analysis of reaction networks. Bioinformatics *22*, 1406–1407.

Macindoe, G., Mavridis, L., Venkatraman, V., Devignes, M.D., and Ritchie, D.W. (2010). HexServer: An FFT-based protein docking server powered by graphics processors. Nucleic Acids Res. *38*, 445–449.

Madhur, M.S., Funt, S.A., Li, L., Vinh, A., Chen, W., Lob, H.E., Iwakura, Y., Blinder, Y., Rahman, A., Quyyumi, A.A., et al. (2011). Role of interleukin 17 in inflammation, atherosclerosis, and vascular function in apolipoprotein e-deficient mice. Arterioscler. Thromb. Vasc. Biol. *31*, 1565–1572.

Malek, T.R., and Castro, I. (2010). Interleukin-2 Receptor Signaling: At the Interface between Tolerance and Immunity. Immunity *33*, 153–165.

Mallat, Z., Besnard, S., Duriez, M., Deleuze, V., Emmanuel, F., Bureau, M.F., Soubrier, F., Esposito, B., Duez, H., Fievet, C., et al. (1999). Protective role of interleukin-10 in atherosclerosis. Circ. Res. *85*, e17–e24.

Mallat, Z., Taleb, S., Ait-Oufella, H., and Tedgui, A. (2009). The role of adaptive T cell immunity in atherosclerosis. J. Lipid Res. *50 Suppl*, S364–S369.

Manikas, T.W., and Cain, J.T. (1996). Genetic Algorithms vs . Simulated Annealing : A Comparison of Approaches for Solving the Circuit Partitioning Problem Genetic Algorithms vs . Simulated Annealing : A Comparison of Approaches for Solving the Circuit Partitioning Problem by.

Manley, G., and Mullinger, R.N. (1967). MUCOPOLYSACCHARIDES OF ATHEROSCLEROTIC PLAQUES AND PLATELETS.

Mao, H., Polliackq, A., Vivian, B., Bfziv, S., Bimn, S., Monocyte, K.W., and B, V. (1986). Parameters Affecting the In Vitro Maturation of Human Monocytes to Macrophages. Int. J. Cell Cloning *185*, 167–185.

Mariani, V., Kiefer, F., Schmidt, T., Haas, J., and Schwede, T. (2011). Assessment of template based protein structure predictions in CASP9. Proteins Struct. Funct. Bioinforma. *79*, 37–58.

Martí-Renom, M.A., Stuart, A.C., Fiser, A., Sánchez, R., Melo, F., and Šali, A. (2000). Comparative Protein Structure Modeling of Genes and Genomes. Annu. Rev. Biophys. Biomol. Struct. *29*, 291–325.

Martinez, M., Bruce, N.J., Romanowska, J., Kokh, D.B., Ozboyaci, M., Yu, X., Öztürk, M.A., Richter, S., and Wade, R.C. (2015). SDA 7: A modular and parallel implementation of the simulation of diffusional association

software. J. Comput. Chem. *36*, 1631–1645.

Mason, T.M. (1998). The role of factors that regulate the synthesis and secretion of very-low-density lipoprotein by hepatocytes. Crit. Rev. Clin. Lab. Sci. *35*, 461–487.

Mazein, A., Watterson, S., Hsieh, W.Y., Griffiths, W.J., and Ghazal, P. (2013). A comprehensive machine-readable view of the mammalian cholesterol biosynthesis pathway. Biochem Pharmacol *86*, 56–66.

Mazzeo, D., Sacco, S., Di Lucia, P., Penna, G., Adorini, L., Panina-Bordignon, P., and Ghezzi, P. (2002). Thiol Antioxidants Inhibit the Formation of the Interleukin-12 Heterodimer: a Novel Mechanism for the Inhibition of Il-12 Production. Cytokine *17*, 285–293.

Mc Auley, M.T., and Mooney, K.M. (2015). Computationally Modeling Lipid Metabolism and Aging: A Mini-review. Comput. Struct. Biotechnol. J. *13*, 38–46.

Mc Auley, M.T., Wilkinson, D.J., Jones, J.J., and Kirkwood, T.B. (2012). A whole-body mathematical model of cholesterol metabolism and its age-associated dysregulation. BMC Syst Biol *6*, 130.

McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server. Bioinformatics *16*, 404–405.

McKay, C., McDee, S., Mottram, N., Mulholland, T., Wilson, S., Kennedy, S., and Wadsworth, R. (2005). Towards a Model of Atherosclerosis. Univ. Strat. 1–29.

McLaren, J.E., and Ramji, D.P. (2009). Interferon gamma: A master regulator of atherosclerosis. Cytokine Growth Factor Rev. *20*, 125–135.

Mecham, R.P., Broekelmann, T.J., Fliszar, C.J., Shapiro, S.D., Welgus, H.G., and Senior, R.M. (1997). Elastin Degradation by Matrix Metalloproteinases. J. Biol. Chem. *272*, 18071–18076.

Mehrabadi, M., Casa, L.D.C., Aidun, C.K., and Ku, D.N. (2016). A Predictive Model of High Shear Thrombus Growth. Ann. Biomed. Eng. *44*, 2339–2350.

Mei, H., Wang, K., McKee, S., Wang, X., Waldner, J.L., Pielak, G.J., Durham, B., and Millett, F. (1996). Control of formation and dissociation of the high-affinity complex between cytochrome c and cytochrome c peroxidase by ionic strength and the low-affinity binding site. Biochemistry *35*, 15800–15806.

Mel'nyk, T.A. (2017). A mathematical model of the atherosclerosis development in thin blood vessels and its asymptotic approximation. 1–20.

Meng, T.C., Somani, S., and Dhar, P. (2004). Modeling and simulation of biological systems with stochasticity. In Silico Biol. *4*, 293–309.

Mereghetti, P., Gabdoulline, R.R., and Wade, R.C. (2010). Brownian dynamics simulation of protein solutions: Structural and dynamical properties. Biophys. J. *99*, 3782–3791.

Merhi-Soussi, F., Kwak, B.R., Magne, D., Chadjichristos, C., Berti, M., Pelli, G., James, R.W., MacH, F., and

Gabay, C. (2005). Interleukin-1 plays a major role in vascular inflammation and atherosclerosis in male apolipoprotein E-knockout mice. Cardiovasc. Res. *66*, 583–593.

Meyer, G., Merval, R., and Tedgui, A. (1996). Effects of pressure-induced stretch and convection on low-density lipoprotein and albumin uptake in the rabbit aortic wall. Circ. Res. *79*, 532–540.

Michaelis, L., and Menten, M.L. (1913). Biochem. Zeitschrift *49*, 333–369.

Minh, H. Van (2006). Epidemiology of cardiovascular disease in rural Vietnam.

Mitchell, S., and Mendes, P. (2013). A computational model of liver iron metabolism. PLoS Comput Biol *9*, e1003299.

Mizuno, S., Iijima, R., Ogishima, S., Kikuchi, M., Matsuoka, Y., Ghosh, S., Miyamoto, T., Miyashita, A., Kuwano, R., and Tanaka, H. (2012). AlzPathway: a comprehensive map of signaling pathways of Alzheimer's disease. BMC Syst. Biol. *6*, 1–10.

Mizzi, C., Peters, B., Mitropoulou, C., Mitropoulos, K., Katsila, T., Agarwal, M.R., van Schaik, R.H., Drmanac, R., Borg, J., and Patrinos, G.P. (2014). Personalized pharmacogenomics profiling using whole-genome sequencing. Pharmacogenomics *15*, 1223–1234.

Molloy, K.J., Thompson, M.M., Schwalbe, E.C., Bell, P.R.F., Naylor, A.R., and Loftus, I.M. (2004). Comparison of levels of matrix metalloproteinases, tissue inhibitor of metalloproteinases, interleukins, and tissue necrosis factor in carotid endarterectomy specimens from patients on versus not on statins preoperatively. Am. J. Cardiol. *94*, 144–146.

Moore, K.J., and Tabas, I. (2011). Macrophages in the pathogenesis of atherosclerosis. Cell *145*, 341–355.

Moore, K.J., Sheedy, F.J., and Fisher, E. a (2013a). Macrophages in atherosclerosis: a dynamic balance. Nat. Rev. Immunol. *13*, 709–721.

Moore, K.J., Sheedy, F.J., and Fisher, E. a (2013b). Macrophages in atherosclerosis: a dynamic balance. Nat. Rev. Immunol. *13*, 709–721.

Moreno, P.R. (2010). Vulnerable Plaque: Definition, Diagnosis, and Treatment. Cardiol. Clin. *28*, 1–30.

Morgan, A.E., Mooney, K.M., Wilkinson, S.J., Pickles, N.A., and Mc Auley, M.T. (2016). Mathematically modelling the dynamics of cholesterol metabolism and ageing. BioSystems *145*, 19–32.

Morrow, D.A. (2016). Myocardial Infarction: A Companion to Braunwald's Heart Disease.

Moult, J., Fidelis, K., Kryshtafovych, A., Schwede, T., and Tramontano, A. (2014). Critical assessment of methods of protein structure prediction (CASP) - round x. Proteins Struct. Funct. Bioinforma. *82*, 1–6.

Moult, J., Fidelis, K., Kryshtafovych, A., Schwede, T., and Tramontano, A. (2016). Critical assessment of methods of protein structure prediction: Progress and new directions in round XI. Proteins Struct. Funct. Bioinforma. 4–14.

Mozaffarian, D., Benjamin, E.J., Go, a. S., Arnett, D.K., Blaha, M.J., Cushman, M., de Ferranti, S., Despres, J.-P., Fullerton, H.J., Howard, V.J., et al. (2014). Heart Disease and Stroke Statistics--2015 Update: A Report From the American Heart Association.

Mozaffarian, D., Benjamin, E.J., Go, A.S., Arnett, D.K., Blaha, M.J., Cushman, M., Das, S.R., de Ferranti, S., Despr??s, J.P., Fullerton, H.J., et al. (2015). Heart Disease and Stroke Statistics???2016 Update: A Report From the American Heart Association.

Mpairaktaris, D.G., Soulis, J. V., and Giannoglou, G.D. (2017). Low density lipoprotein transport through patient-specific thoracic arterial wall. Comput. Biol. Med. *89*, 115–126.

Muddyman, D., Smee, C., Griffin, H., and Kaye, J. (2013). Implementing a successful data-management framework: the UK10K managed access model. Genome Med *5*, 100.

Mukherjee, S., and Zhang, Y. (2009). MM-align: A quick algorithm for aligning multiple-chain protein complex structures using iterative dynamic programming. Nucleic Acids Res. *37*.

Mukherjee, S., and Zhang, Y. (2011). Protein-protein complex structure predictions by multimeric threading and template recombination. Structure *19*, 955–966.

Muntner, P., He, J., Astor, B.C., Folsom, A.R., and Coresh, J. (2005). Traditional and Nontraditional Risk Factors Predict Coronary Heart Disease in Chronic Kidney Disease: Results from the Atherosclerosis Risk in Communities Study. J. Am. Soc. Nephrol. *16*, 529–538.

Murzin, A.G., Brenner, S.E., Hubbard, T., and Chothia, C. (1995). SCOP: A structural classification of proteins database for the investigation of sequences and structures. J. Mol. Biol. *247*, 536–540.

Myers, J.G., Moore, J.A., Ojha, M., Johnston, K.W., and Ethier, C.R. (2001). Factors influencing blood flow patterns in the human right coronary artery. Ann. Biomed. Eng. *29*, 109–120.

Myszka, D.G., Arulanantham, P.R., Sana, T., Wu, Z., Morton, T.A., Ciardelli, T.L., Myszka, D.G., Arulanantham, P.R., Sana, T., Wu, Z., et al. (1996). Kinetic analysis of ligand binding to interleukin-2 receptor complexes created on an optical biosensor surface Kinetic analysis of ligand binding to interleukin-2 receptor complexes created on an optical biosensor surface. 2468–2478.

Nabel, E.G., Ganz, P., Gordon, J.B., Alexander, R.W., and Selwyn, A.P. (1988). Dilation of normal and constriction of atherosclerotic coronary arteries caused by the cold pressor test. Circulation *77*, 43–52.

Nakashima, Y., Chen, Y.X., Kinukawa, N., and Sueishi, K. (2002). Distributions of diffuse intimal thickening in human arteries: Preferential expression in atherosclerosis-prone arteries from an early age. Virchows Arch. *441*, 279–288.

Nayeem, N., Green, T.P., Martin, I.L., and Barnard, E.A. (1994). Quaternary Structure of the Native GABAA Receptor Determined by Electron Microscopic Image Analysis. J. Neurochem. *62*, 815–818.

Nazari-Jahantigh, M., Egea, V., Schober, A., and Weber, C. (2014). MicroRNA-specific regulatory mechanisms in

atherosclerosis. J Mol Cell Cardiol.

Nelson, D.E., Ihekwaba, A.E., Elliott, M., Johnson, J.R., Gibney, C.A., Foreman, B.E., Nelson, G., See, V., Horton, C.A., Spiller, D.G., et al. (2004). Oscillations in NF-kappaB signaling control the dynamics of gene expression. Science (80-.). *306*, 704–708.

Nematollahi, A., Shirani, E., Mirzaee, I., and Sadeghi, M.R. (2012). Numerical simulation of LDL particles mass transport in human carotid artery under steady state conditions. Sci. Iran. *19*, 519–524.

Neužil, J., Thomas, S.R., and Stocker, R. (1996). Requirement for, promotion, or inhibition by α-tocopherol of radical- induced initiation of plasma lipoprotein lipid peroxidation. Free Radic. Biol. Med. *22*, 57–71.

Newby, A.C. (2005). Dual Role of Matrix Metalloproteinases (Matrixins) in Intimal Thickening and Atherosclerotic Plaque Rupture. Physiol. Rev. *85*, 1–31.

Newby, A.C. (2008). Metalloproteinase expression in monocytes and macrophages and its relationship to atherosclerotic plaque instability. Arterioscler. Thromb. Vasc. Biol. *28*, 2108–2114.

NHLBI (2016). NHLBI - What are the Signs and Symptoms of Atherosclerosis?

Nicholas, T., Duvvuri, S., Leurent, C., Raunig, D., Rapp, T., Iredale, P., Rowinski, C., Carr, R., Roberts, P., Spiros, A., et al. (2013). Systems pharmacology modeling in neuroscience: Prediction and outcome of PF-04995274, a 5-HT4 partial agonist, in a clinical scopolamine impairment trial. Adv. Alzheimer's Dis. *2*, 83–98.

Nichols M Luengo-Fernandez R, Leal J, Gray A, Scarborough P, Rayner M, T.N. (2012). European Cardiovascular Disease Statistics 2012 (European Heart Network, Brussels, European Society of Cardiology, Sophia Antipolis).

Nijhout, H.F., Best, J.A., and Reed, M.C. (2015). Using mathematical models to understand metabolism, genes, and disease. BMC Biol. *13*, 79.

Nishi, K., Itabe, H., Uno, M., Kitazato, K.T., Horiguchi, H., Shinno, K., and Nagahiro, S. (2002). Oxidized LDL in carotid plaques and plasma associates with plaque instability. Arterioscler. Thromb. Vasc. Biol. *22*, 1649–1654.

Nojiri, S., and Daida, H. (2017). Atherosclerotic Cardiovascular Risk in Japan. Japanese Clin. Med. *8*, 1179066017712713.

De Nooijer, R., Von Der Thüsen, J.H., Verkleij, C.J.N., Kuiper, J., Jukema, J.W., Van Der Wall, E.E., Van Berkel, T.J.C., and Biessen, E.A.L. (2004). Overexpression of IL-18 decreases intimal collagen content and promotes a vulnerable plaque phenotype in apolipoprotein-E-deficient mice. Arterioscler. Thromb. Vasc. Biol. *24*, 2313–2319.

Nooren, I.M. a. (2003). NEW EMBO MEMBER'S REVIEW: Diversity of protein-protein interactions. EMBO J. *22*, 3486–3492.

Northrup, S., Allison, S., and McCammon, A. (1984). Brownian dynamics simulation of diffusion-influenced bimolecular reactions. J. Chem. Phys. *80*, 1517–1524.

Novère, N. Le (2006). BioModels Database , a curated resource of annotated published models. Bioinformatics *691*, 2006–2006.

Le Novere, N., Bornstein, B., Broicher, A., Courtot, M., Donizelli, M., Dharuri, H., Li, L., Sauro, H., Schilstra, M., Shapiro, B., et al. (2006). BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. Nucleic Acids Res *34*, D689-91.

Le Novere, N., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A., Demir, E., Wegner, K., Aladjem, M.I., Wimalaratne, S.M., et al. (2009). The Systems Biology Graphical Notation. Nat Biotechnol *27*, 735–741.

Le Novère, N., Finney, A., Hucka, M., Bhalla, U.S., Campagne, F., Collado-Vides, J., Crampin, E.J., Halstead, M., Klipp, E., Mendes, P., et al. (2005). Minimum information requested in the annotation of biochemical models (MIRIAM). Nat. Biotechnol. *23*, 1509–1515.

Nugent, T., Cozzetto, D., and Jones, D.T. (2014). Evaluation of predictions in the CASP10 model refinement category. Proteins Struct. Funct. Bioinforma. *82*, 98–111.

Nunes, G.L.C., Simões, A., Dyszy, F.H., Shida, C.S., Juliano, M.A., Juliano, L., Gesteira, T.F., Nader, H.B., Murphy, G., Chaffotte, A.F., et al. (2011). Mechanism of heparin acceleration of tissue inhibitor of metalloproteases-1 (TIMP-1) degradation by the human neutrophil elastase. PLoS One *6*.

O'Meara, M.J., Leaver-Fay, A., Tyka, M.D., Stein, A., Houlihan, K., Dimaio, F., Bradley, P., Kortemme, T., Baker, D., Snoeyink, J., et al. (2015). Combined covalent-electrostatic model of hydrogen bonding improves structure prediction with Rosetta. J. Chem. Theory Comput. *11*, 609–622.

O'Keefe, J.H., Cordain, L., Harris, W.H., Moe, R.M., and Vogel, R. (2004). Optimal low-density lipoprotein is 50 to 70 mg/dl. J. Am. Coll. Cardiol. *43*, 2142 LP-2146.

Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J.H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., et al. (2002). Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell *110*, 775–787.

Ohashi, R., Mu, H., Wang, X., Yao, Q., and Chen, C. (2005). Reverse cholesterol transport and cholesterol efflux in atherosclerosis. QJM - Mon. J. Assoc. Physicians *98*, 845–856.

Olgac, U., Kurtcuoglu, V., and Poulikakos, D. (2008). Computational modeling of coupled blood-wall mass transport of LDL: effects of local wall shear stress. Am. J. Physiol. Heart Circ. Physiol. *294*, H909–H919.

Olson, M.W., Gervasi, D.C., Mobashery, S., and Fridman, R. (1997). Kinetic Analysis of the Binding of Human Matrix Metalloproteinase-2 and -9 to Tissue Inhibitor of Metalloproteinase (TIMP)-1 and TIMP-2. J. Biol. Chem. *272*, 29975–29983.

OpenStax CNX (2013). OpenStax, Structure and Function of Blood Vessels.

Oram, J.F., and Lawn, R.M. (2001). ABCA1. The gatekeeper for eliminating excess tissue cholesterol. J. Lipid Res. *42*, 1173–1179.

Organisation, W.H. (2014). Facts about aging.

Organisation, W.H. (2015). Cardiovascular Diseases Fact Sheet.

Otvos, J.D., Mora, S., Shalaurova, I., Greenland, P., Mackey, R.H., and Goff Jr., D.C. (2011). Clinical implications of discordance between low-density lipoprotein cholesterol and particle number. J Clin Lipidol *5*, 105–113.

Ougrinovskaia, A., Thompson, R.S., and Myerscough, M.R. (2010). An ODE model of early stages of atherosclerosis: mechanisms of the inflammatory response. Bull Math Biol *72*, 1534–1561.

Owen, M.R., and Sherratt, J. a (1997). Pattern formation and spatiotemporal irregularity in a model for macrophage-tumour interactions. J. Theor. Biol. *189*, 63–80.

Pagadala, N.S., Syed, K., and Tuszynski, J. (2017). Software for molecular docking: a review. Biophys. Rev. *9*, 91–102.

Page, M.M., and Watts, G.F. (2016). PCSK9 inhibitors - mechanisms of action. Aust. Prescr. *39*, 164–167.

Palinski, W., and Napoli, C. (2002). The fetal origins of atherosclerosis: maternal hypercholesterolemia, and cholesterol-lowering or antioxidant treatment during pregnancy influence in utero programming and postnatal susceptibility to atherogenesis. FASEB J. *16*, 1348–1360.

Palmieri, L., Bennett, K., Giampaoll, S., and Capewell, S. (2010). Explaining the decrease in coronary heart disease mortality in Italy between 1980 and 2000. Am. J. Public Health *100*, 684–692.

Pan, A.C., Borhani, D.W., Dror, R.O., and Shaw, D.E. (2013). Molecular determinants of drug-receptor binding kinetics. Drug Discov. Today *18*, 667–673.

Pang, X., and Zhou, H. (2017). Rate Constants and Mechanisms of Protein- Ligand Binding. Annu. Rev. Biophys. *46*, 105–130.

Parton, A., McGilligan, V., O'Kane, M., Baldrick, F.R., and Watterson, S. (2015). Computational modelling of atherosclerosis. Briefings Bioinforma. .

Pastrana, J.L., Sha, X., Virtue, A., Mai, J., Cueto, R., Lee, I.A., Wang, H., and Yang, X.-F. (2012). Regulatory T cells and Atherosclerosis. J. Clin. Exp. Cardiolog. *2012*, 2.

Pattnaik, P. (2005). Surface plasmon resonance: applications in understanding receptor-ligand interaction. Appl. Biochem. Biotechnol. *126*, 79–92.

PDB, T.R. (2016). PDB Statistics.

Pedersen, M., and Plotkin, G. (2008). A language for biochemical systems. In Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 63–82.

Perry, H.M., Bender, T.P., and McNamara, C.A. (2012). B cell subsets in atherosclerosis. Front Immunol *3*, 373.

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. J. Comput. Chem. *25*, 1605–1612.

Phillip, Y., Sherman, E., Haran, G., and Schreiber, G. (2009). Common crowding agents have only a small effect on protein-protein interactions. Biophys. J. *97*, 875–885.

Phizicky, E.M., and Fields, S. (1995). Protein-protein interactions: methods for detection and analysis. Microbiol. Rev. *59*, 94–123.

Pichardo-Almarza, C., Metcalf, L., Finkelstein, A., and Diaz-Zuccarini, V. (2015). Using a systems pharmacology approach to study the effect of statins on the early stage of atherosclerosis in humans. CPT Pharmacometrics Syst. Pharmacol. *4*, 41–50.

Pinto, F.J. (2014). Cardiologia. *33*.

Pirillo, A., Norata, G.D., and Catapano, A.L. (2013). LOX-1, OxLDL, and atherosclerosis. Mediators Inflamm. *2013*.

Placzek, S., Schomburg, I., Chang, A., Jeske, L., Ulbrich, M., Tillack, J., and Schomburg, D. (2017). BRENDA in 2017: New perspectives and new tools in BRENDA. Nucleic Acids Res. *45*, D380–D388.

Postiglione, A., and Napoli, C. (1995). Hyperlipidaemia and atherosclerotic cerebrovascular disease. Curr. Opin. Lipidol. *6*, 236–242.

Poupon, A., and Janin, J. (2010). Analysis and Prediction of Protein Quaternary Structure BT - Data Mining Techniques for the Life Sciences. O. Carugo, and F. Eisenhaber, eds. (Totowa, NJ: Humana Press), pp. 349–364.

Powell, J.T. (1998). Vascular damage from smoking: disease mechanisms at the arterial wall. Vasc. Med. *3*, 21–28.

Prakash, M.K. (2011). Insights on the role of (Dis)order from protein - Protein interaction linear free-energy relationships. J. Am. Chem. Soc. *133*, 9976–9979.

Prosi, M., Zunino, P., Perktold, K., and Quarteroni, A. (2005). Mathematical and numerical models for transfer of low-density lipoproteins through the arterial walls: a new methodology for the model set up with applications to the study of disturbed lumenal flow. J Biomech *38*, 903–917.

Qiao, J.H., Tripathi, J., Mishra, N.K., Cai, Y., Tripathi, S., Wang, X.P., Imes, S., Fishbein, M.C., Clinton, S.K., Libby, P., et al. (1997). Role of macrophage colony-stimulating factor in atherosclerosis: studies of osteopetrotic mice. Am. J. Pathol. *150*, 1687–1699.

Qin, S., Pang, X., and Zhou, H.X. (2011). Automated prediction of protein association rate constants. Structure *19*, 1744–1751.

Quarteroni, A., Veneziani, A., and Zunino, P. (2002). Mathematical and Numerical Modeling of Solute Dynamics in Blood Flow and Arterial Walls. SIAM J. Numer. Anal. *39*, 1488–1511.

Ragino, Y.I., Chernyavski, A.M., Polonskaya, Y. V, Volkov, A.M., and Kashtanova, E. V (2012). Activity of the inflammatory process in different types of unstable atherosclerotic plaques. Bull Exp Biol Med *153*, 186–189.

Ramalho, L.S., Oliveira, L.F., Cavellani, C.L., Ferraz, M.L.D.F., De Oliveira, F.A., Miranda Corrêa, R.R., De Paula Antunes Teixeira, V., and De Lima Pereira, S.A. (2013). Role of mast cell chymase and tryptase in the progression of atherosclerosis: Study in 44 autopsied cases. Ann. Diagn. Pathol. *17*, 28–31.

Ramsey, S., Orrell, D., and Bolouri, H. (2005). Dizzy: stochastic simulation of large-scale genetic regulatory networks. J. Bioinform. Comput. Biol. *3*, 415–436.

Rappitsch, G., Perktold, K., and Pernkopf, E. (1997). Numerical Modelling of Shear-Dependent Mass Transfer in Large Arteries. *857*, 847–857.

Raza, S., McDerment, N., Lacaze, P.A., Robertson, K., Watterson, S., Chen, Y., Chisholm, M., Eleftheriadis, G., Monk, S., O'Sullivan, M., et al. (2010). Construction of a large scale integrated map of macrophage pathogen recognition and effector systems. BMC Syst Biol *4*, 63.

Redgrave, T.G. (1970). Formation of cholesteryl ester-rich particulate lipid during metabolism of chylomicrons. J. Clin. Invest. *49*, 465–471.

Rekhter, M.D. (1999). Collagen synthesis in atherosclerosis: Too much and not enough. Cardiovasc. Res. *41*, 376–384.

Resnick, N., Yahav, H., and Shay-Salit, a (2003). Fluid shear stress and the vascular endothelium: for better and for worse. Prog. Biophys. … *81*, 177–199.

Ritchie, D.W. (2008). Recent progress and future directions in protein-protein docking. Curr. Protein Pept. Sci. *9*, 1–15.

Ritchie, D.W., and Venkatraman, V. (2010). Ultra-fast FFT protein docking on graphics processors. Bioinformatics *26*, 2398–2405.

Rollins, B.J., Walz, A., and Baggiolini, M. (1991). Recombinant human MCP-1/JE induces chemotaxis, calcium flux, and the respiratory burst in human monocytes. Blood *78*, 1112–1116.

Rose, P.W., Prlić, A., Bi, C., Bluhm, W.F., Christie, C.H., Dutta, S., Green, R.K., Goodsell, D.S., Westbrook, J.D., Woo, J., et al. (2015). The RCSB Protein Data Bank: Views of structural biology for basic and applied research and education. Nucleic Acids Res. *43*, D345–D356.

Rosenson, R.S., Brewer, H.B., Davidson, W.S., Fayad, Z. a., Fuster, V., Goldstein, J., Hellerstein, M., Jiang, X.C., Phillips, M.C., Rader, D.J., et al. (2012). Cholesterol efflux and atheroprotection: Advancing the concept of reverse cholesterol transport. Circulation *125*, 1905–1919.

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc *5*, 725–738.

Roy, S., Biswas, S., Khanna, S., Gordillo, G., Bergdall, V., Green, J., Marsh, C.B., Gould, L.J., and Sen, C.K. (2009).

Characterization of a preclinical model of chronic ischemic wound. Physiol. Genomics *37*, 211–224.

Rudijanto, A. (2007). The role of vascular smooth muscle cells on the pathogenesis of atherosclerosis. Acta Med Indones *39*, 86–93.

Sadir, R., Forest, E., and Lortat-Jacob, H. (1998). The heparan sulfate binding sequence of interferon-gamma increased the on rate of the interferon-gamma-interferon-gamma receptor complex formation. J. Biol. Chem. *273*, 10919–10925.

Sakai, N., Uchida, Y., Ohashi, K., Hibuse, T., Saika, Y., Tomari, Y., Kihara, S., Hiraoka, H., Nakamura, T., Ito, S., et al. (2003). Measurement of fasting serum apoB-48 levels in normolipidemic and hyperlipidemic subjects by ELISA. J. Lipid Res. *44*, 1256–1262.

Sakurai, A., Morita, S.-Y., Wakita, K., Deharu, Y., Nakano, M., and Handa, T. (2005). Effects of cholesterol in chylomicron remnant models of lipid emulsions on apoE-mediated uptake and cytotoxicity of macrophages. J. Lipid Res. *46*, 2214–2220.

Sandborn, W.J., Feagan, B.G., Fedorak, R.N., Scherl, E., Fleisher, M.R., Katz, S., Johanns, J., Blank, M., and Rutgeerts, P. (2008). A Randomized Trial of Ustekinumab, a Human Interleukin-12/23 Monoclonal Antibody, in Patients With Moderate-to-Severe Crohn's Disease. Gastroenterology *135*, 1130–1141.

Santoli, D., Yang, Y.C., Clark, S.C., Kreider, B.L., Caracciolo, D., and Rovera, G. (1987). Synergistic and antagonistic effects of recombinant human interleukin (IL) 3, IL-1 alpha, granulocyte and macrophage colony-stimulating factors (G-CSF and M-CSF) on the growth of GM-CSF-dependent leukemic cell lines. J. Immunol. *139*, 3348–3354.

Sartori, a, Ma, X., Gri, G., Showe, L., Benjamin, D., and Trinchieri, G. (1997). Interleukin-12: an immunoregulatory cytokine produced by B cells and antigen-presenting cells. Methods *11*, 116–127.

Sato, K., Niki, E., and Shimasaki, H. (1990). Free radical-mediated chain oxidation of low density lipoprotein and its synergistic inhibition by vitamin E and vitamin C. Arch. Biochem. Biophys. *279*, 402–405.

Saunders, W.B., Bayless, K.J., and Davis, G.E. (2005). MMP-1 activation by serine proteases and MMP-10 induces human capillary tubular network collapse and regression in 3D collagen matrices. J. Cell Sci. *118*, 2325–2340.

Schiopu, A., Frendéus, B., Jansson, B., Söderberg, I., Ljungcrantz, I., Araya, Z., Shah, P.K., Carlsson, R., Nilsson, J., and Fredrikson, G.N. (2007). Recombinant Antibodies to an Oxidized Low-Density Lipoprotein Epitope Induce Rapid Regression of Atherosclerosis in Apobec-1-/-/Low-Density Lipoprotein Receptor-/- Mice. J. Am. Coll. Cardiol. *50*, 2313–2318.

Schlatter, R., Schmich, K., Vizcarra, I.A., Scheurich, P., Sauter, T., Borner, C., Ederer, M., Merfort, I., and Sawodny, O. (2009). ON/OFF and beyond - A Boolean model of apoptosis. PLoS Comput. Biol. *5*.

Schreiber, G., Haran, G., and Zhou, H.-X. (2009). Fundamental aspects of protein− protein association kinetics.

Chem. Rev. *109*, 839–860.

Schrijvers, D.M., De Meyer, G.R.Y., Herman, A.G., and Martinet, W. (2007). Phagocytosis in atherosclerosis: Molecular mechanisms and implications for plaque progression and stability. Cardiovasc. Res. *73*, 470–480.

Schulze-Bauer, C.A.J., Mörth, C., and Holzapfel, G.A. (2003). Passive biaxial mechanical response of aged human iliac arteries. J. Biomech. Eng. *125*, 395–406.

Schunkert, H., König, I.R., Kathiresan, S., Reilly, M.P., Assimes, T.L., Holm, H., Preuss, M., Stewart, A.F.R., Barbalic, M., Gieger, C., et al. (2011). Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat. Genet. *43*, 333–338.

Schwartz, S.M. (1997). Perspectives series: cell adhesion in vascular biology. Smooth muscle migration in atherosclerosis and restenosis. J. Clin. Invest. *99*, 2814–2816.

Sebba, A. (2008). Tocilizumab: The first interleukin-6-receptor inhibitor. Am. J. Heal. Pharm. *65*, 1413–1418.

Sekikawa, A., Ueshima, H., Kadowaki, T., El-Saed, A., Okamura, T., Takamiya, T., Kashiwagi, A., Edmundowicz, D., Murata, K., Sutton-Tyrrell, K., et al. (2007). Less Subclinical Atherosclerosis in Japanese Men in Japan than in White Men in the United States in the Post–World War II Birth Cohort. Am. J. Epidemiol. *165*, 617–624.

Sekikawa, a, Horiuchi, B.Y., Edmundowicz, D., Ueshima, H., Curb, J.D., Sutton-Tyrrell, K., Okamura, T., Kadowaki, T., Kashiwagi, a, Mitsunami, K., et al. (2003). A “natural experiment” in cardiovascular epidemiology in the early 21st century. Heart *89*, 255–257.

SELVE, N., and WEGNER, A. (1987). pH???dependent rate of formation of the gelsolin???actin complex from gelsolin and monomeric actin. Eur. J. Biochem. *168*, 111–115.

Shahar, E., Heiss, G., Rosamond, W.D., and Szklo, M. (2008). Baldness and myocardial infarction in men: The atherosclerosis risk in communities study. Am. J. Epidemiol. *167*, 676–683.

Sharma, K.H., Sahoo, S., Shah, K.H., Patel, A.K., Jadhav, N.D., Parmar, M.M., and Patel, K.H. (2015). Are Gujarati Asian Indians “older” for their “vascular age” as compared to their “Chronological age”? QJM An Int. J. Med. *108*, 105–112.

Shaw, G.M., Chase, J.G., Starfinger, C., Smith, B.W., Hann, C.E., Desaive, T., and Ghuysen, A. (2007). Modelling the cardiovascular system. Crit. Care Resusc. *9*, 264–269.

Shen, B.J., Hage, T., and Sebald, W. (1996). Global and local determinants for the kinetics of interleukin-4/interleukin-4 receptor alpha chain interaction. A biosensor study employing recombinant interleukin-4-binding protein. Eur. J. Biochem. *240*, 252–261.

Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. *29*, 308–311.

Shi, Q., Vandeberg, J.F., Jett, C., Rice, K., Leland, M.M., Talley, L., Kushwaha, R.S., Rainwater, D.L., Vandeberg, J.L., and Wang, X.L. (2005). Arterial endothelial dysfunction in baboons fed a high-cholesterol, high-fat diet.

Am. J. Clin. Nutr. *82*, 751–759.

Silva, T., Sequeira, A., Santos, R.F., and Tiago, J. (2013). Mathematical Modeling of Atherosclerotic Plaque Formation Coupled with a Non-Newtonian Model of Blood Flow. Conf. Pap. Med. *2013*, 1–14.

Singh, R.B., Mengi, S.A., Xu, Y.J., Arneja, A.S., and Dhalla, N.S. (2002). Pathogenesis of atherosclerosis: A multifactorial process. Exp Clin Cardiol *7*, 40–53.

SJ, N., Puri, R., Anderson, T., and al, et (2016). Effect of evolocumab on progression of coronary disease in statin-treated patients: The glagov randomized clinical trial. JAMA *316*, 2373–2384.

Sliwa, K., and Zilla, P. (2012). Rheumatic heart disease: The tip of the iceberg. Circulation *125*, 3060–3062.

Soares, S., Antunes, C.H., and Araújo, R. (2013). Comparison of a genetic algorithm and simulated annealing for automatic neural network ensemble development. Neurocomputing *121*, 498–511.

Soriano, A., Lozano, F., Oliva, H., García, F., Nomdedéu, M., De Lazzari, E., Rodríguez, C., Barrasa, A., Lorenzo, J.I., Del Romero, J., et al. (2005). Polymorphisms in the interleukin-4 receptor α chain gene influence susceptibility to HIV-1 infection and its progression to AIDS. Immunogenetics *57*, 644–654.

Soulis, J. V, Fytanidis, D.K., Lampri, O.P., and Giannoglou, G.D. (2016). Low Density Lipoprotein and Non-Newtonian Oscillating Flow Biomechanical Parameters for Normal Human Aorta. Cardiol. Res. *7*, 66–79.

Spaar, A., Flöck, D., and Helms, V. (2009). Association of cytochrome c with membrane-bound cytochrome c oxidase proceeds parallel to the membrane rather than in bulk solution. Biophys. J. *96*, 1721–1732.

Stadius, M.L., Rowan, R., Fleischhauer, J.F., Kernoff, R., Billingham, M., and Gown, A.M. (1992). Time course and cellular characteristics of the iliac artery response to acute balloon injury. An angiographic, morphometric, and immunocytochemical analysis in the cholesterol-fed New Zealand white rabbit. Arterioscler. Thromb. *12*, 1267–1273.

Staels, B., and Fonseca, V.A. (2009). Bile acids and metabolic regulation: mechanisms and clinical responses to bile acid sequestration. Diabetes Care *32 Suppl 2*.

Stancu, C., and Sima, A. (2001). Statins: mechanism of action and effects. J. Cell. Mol. Med. *5*, 378–387.

Stary, H., Chandler, A., Glagov, S., Guyton, J., INsull, W., Rosenfel, M., Schaffer, S., Schwartz, C., Wagner, W., and Wissler, R. (1994). A definition of intial, fatty streak, and intermediate lesions of atherosclerosis. Arterioscler. Thromb. Vasc. Biol. *14*, 840–856.

Stary, H.C., Chandler, a B., Dinsmore, R.E., Fuster, V., Glagov, S., Insull, W., Rosenfeld, M.E., Schwartz, C.J., Wagner, W.D., and Wissler, R.W. (1995). A definition of advanced types of atherosclerotic lesions and a histological classification of atherosclerosis. Arterioscler. Thromb. Vasc. Biol. *15*, 1512–1531.

Stein, A., Montens, H.P., Steppich, B., Busch, G., Brandl, R., and Ott, I. (2008). Circulating endothelial progenitor cells decrease in patients after endarterectomy. J. Vasc. Surg. *48*, 1217–1222.

Steinberg, D., and Witztum, J.L. (2010). Oxidized low-density lipoprotein and atherosclerosis. Arterioscler. Thromb. Vasc. Biol. *30*, 2311–2316.

Steinke, J.W. (2004). Anti-interleukin-4 therapy. Immunol. Allergy Clin. North Am. *24*, 599–614.

Stenmark, K.R., Yeager, M.E., El Kasmi, K.C., Nozik-Grayck, E., Gerasimovskaya, E. V, Li, M., Riddle, S.R., and Frid, M.G. (2013). The adventitia: essential regulator of vascular wall structure and function. Annu Rev Physiol *75*, 23–47.

Stoneman, V.E.A., and Bennett, M.R. (2004). Role of apoptosis in atherosclerosis and its therapeutic implications. Clin. Sci. (Lond). *107*, 343–354.

Stroud, J.S., Berger, S. a, and Saloner, D. (2002). Numerical analysis of flow through a severely stenotic carotid artery bifurcation. J. Biomech. Eng. *124*, 9–20.

Subramanian, M., and Tabas, I. (2014). Dendritic cells in atherosclerosis. Semin Immunopathol *36*, 93–102.

Sudmant, P.H., Rausch, T., Gardner, E.J., Handsaker, R.E., Abyzov, A., Huddleston, J., Zhang, Y., Ye, K., Jun, G., Hsi-Yang Fritz, M., et al. (2015). An integrated map of structural variation in 2,504 human genomes. Nature *526*, 75–81.

Sun, N., Wood, N.B., Hughes, A.D., Thom, S. a M., and Xu, X.Y. (2006). Fluid-wall modelling of mass transfer in an axisymmetric Stenosis: Effects of shear-dependent transport properties. Ann. Biomed. Eng. *34*, 1119–1128.

Sun, X., Vilar, S., and Tatonetti, N.P. (2013). High-throughput methods for combinatorial drug discovery. Sci Transl Med *5*, 205rv1.

Syto, R., Murgolo, N.J., Braswell, E.H., Mui, P., Huang, E., and Windsor, W.T. (1998). Structural and biological stability of the human interleukin 10 homodimer. Biochemistry *37*, 16943–16951.

Szodoray, P., Timar, O., Veres, K., Der, H., Szomjak, E., Lakos, G., Aleksza, M., Nakken, B., Szegedi, G., and Soltesz, P. (2006). Th1/Th2 Imbalance, Measured by Circulating and Intracytoplasmic Inflammatory Cytokines – Immunological Alterations in Acute Coronary Syndrome and Stable Coronary Artery Disease. Scand. J. Immunol. *64*, 336–344.

Tabas, I., Williams, K.J., and Bor??n, J. (2007). Subendothelial lipoprotein retention as the initiating process in atherosclerosis: Update and therapeutic implications. Circulation *116*, 1832–1844.

Takatsu, K. (2011). Interleukin-5 and IL-5 receptor in health and diseases. Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci. *87*, 463–485.

Talayero, B.G., and Sacks, F.M. (2011). The role of triglycerides in atherosclerosis. Curr Cardiol Rep *13*, 544–552.

Tall, A.R., and Yvan-Charvet, L. (2015). Cholesterol, inflammation and innate immunity. Nat Rev Immunol *15*, 104–116.

Tatonetti, N.P., Liu, T., and Altman, R.B. (2009). Predicting drug side-effects by chemical systems biology. Genome Biol *10*, 238.

Tau, G., and Rothman, P. (1999). Biologic functions of the IFN-gamma receptors. Allergy *54*, 1233–1251.

Tavakolian Ferdousie, V., Mohammadi, M., Hassanshahi, G., Khorramdelazad, H., Falahati-Pour, S.K., Mirzaei, M., Allah Tavakoli, M., Kamiab, Z., Ahmadi, Z., Vazirinejad, R., et al. (2017). Serum CXCL10 and CXCL12 chemokine levels are associated with the severity of coronary artery disease and coronary artery occlusion. Int. J. Cardiol. *233*, 23–28.

Tellis, C.C., and Tselepis, A.D. (2009). ??he role of lipoprotein-associated phospholipase A2 in atherosclerosis may depend on its lipoprotein carrier in plasma. Biochim. Biophys. Acta - Mol. Cell Biol. Lipids *1791*, 327–338.

The Jmol Team (2007). Jmol: an open-source Java viewer for chemical structures in 3D. Jmolsourceforgenet.

Theoharides, T.C., Alysandratos, K.D., Angelidou, A., Delivanis, D.A., Sismanopoulos, N., Zhang, B., Asadi, S., Vasiadi, M., Weng, Z., Miniati, A., et al. (2012). Mast cells and inflammation. Biochim. Biophys. Acta - Mol. Basis Dis. *1822*, 21–33.

Thiele, I., Swainston, N., Fleming, R.M., Hoppe, A., Sahoo, S., Aurich, M.K., Haraldsdottir, H., Mo, M.L., Rolfsson, O., Stobbe, M.D., et al. (2013). A community-driven global reconstruction of human metabolism. Nat Biotechnol *31*, 419–425.

Tomaso, G. Di, Díaz-Zuccarini, V., and Pichardo-Almarza, C. (2011). A Multiscale Model of Atherosclerotic Plaque Formation at Its Early Stage. IEEE Trans. Biomed. Eng. *58*, 3460–3463.

Tomkin, G.H., and Owens, D. (2012). The chylomicron: Relationship to atherosclerosis. Int. J. Vasc. Med. *2012*.

Townsend, N., Bhatnagar, P., Wilkins, E., Wickramasinghe, K., and Rayner, M. (2015a). Cardiovascular Disease Statistics 2015.

Townsend, N., Bhatnagar, P., Wilkins, E., Wickramasinghe, K., and Rayner, M. (2015b). Cardiovascular Disease Statistics 2015.

Troeberg, L., Tanaka, M., Wait, R., Shi, Y.E., Brew, K., and Nagase, H. (2002). E. coli Expression of TIMP-4 and comparative kinetic studies with TIMP-1 and TIMP-2: Insights into the interactions of TIMPS and matrix metalloproteinase 2 (Gelatinase A). Biochemistry *41*, 15025–15035.

Trogan, E., Feig, J.E., Dogan, S., Rothblat, G.H., Angeli, V., Tacke, F., Randolph, G.J., and Fisher, E.A. (2006). Gene expression changes in foam cells and the role of chemokine receptor CCR7 during atherosclerosis regression in ApoE-deficient mice. Proc. Natl. Acad. Sci. *103*, 3781–3786.

Trusheim, M.R., Burgess, B., Hu, S.X., Long, T., Averbuch, S.D., Flynn, A.A., Lieftucht, A., Mazumder, A., Milloy, J., Shaw, P.M., et al. (2011). Quantifying factors for the success of stratified medicine. Nat. Rev. Drug Discov. *10*, 817–833.

Tse, K., Tse, H., Sidney, J., Sette, A., and Ley, K. (2013a). T cells in atherosclerosis. Int Immunol *25*, 615–622.

Tse, K., Tse, H., Sidney, J., Sette, A., and Ley, K. (2013b). T cells in atherosclerosis. Int. Immunol. *25*, 615–622.

Tsiantoulas, D., Diehl, C.J., Witztum, J.L., and Binder, C.J. (2014). B cells and humoral immunity in atherosclerosis. Circ. Res. *114*, 1743–1756.

Tsukaguchi, K., Balaji, K., and Boom, W. (2011). CD4+ T Cell and gamma-delta T Cell Responses to Mycobacterium tuberculosis.

Tyka, M.D., Keedy, D.A., Andr??, I., Dimaio, F., Song, Y., Richardson, D.C., Richardson, J.S., and Baker, D. (2011). Alternate states of proteins revealed by detailed energy landscape mapping. J. Mol. Biol. *405*, 607–618.

Umaerus, M., Rosengren, B., Fagerberg, B., Hurt-Camejo, E., and Camejo, G. (2012). HDL2 interferes with LDL association with arterial proteoglycans: a possible athero-protective effect. Atherosclerosis *225*, 115–120.

Umemoto, T., Han, C.Y., Mitra, P., Averill, M.M., Tang, C., Goodspeed, L., Omer, M., Subramanian, S., Wang, S., Den Hartigh, L.J., et al. (2013). Apolipoprotein AI and high-density lipoprotein have anti-inflammatory effects on adipocytes via cholesterol transporters: ATP-binding cassette A-1, ATP-binding cassette G-1, and scavenger receptor B-1. Circ Res *112*, 1345–1354.

Upadhya, S., Mooteri, S., Peckham, N., and Pai, R.G. (2004). Atherogenic effect of interleukin-2 and antiatherogenic effect of interleukin-2 antibody in apo-E-deficient mice. Angiology *55*, 289–294.

Valente, A.J., Irimpen, A.M., Siebenlist, U., and Chandrasekar, B. (2014). OxLDL induces endothelial dysfunction and death via TRAF3IP2: Inhibition by HDL3 and AMPK activators. Free Radic. Biol. Med. *70*, 117–128.

Vartiainen, E., Puska, P., Jousilahti, P., and Korhonen, H.J. (1999). Cardiovascular Diseases and Risk Factors in Finland. Prev. Med. (Baltim). *29*, S124–S129.

Velikova, M., Van Scheltinga, J.T., Lucas, P.J.F., and Spaanderman, M. (2014). Exploiting causal functional relationships in Bayesian network modelling for personalised healthcare. Int. J. Approx. Reason. *55*, 59–73.

Verhoeven, B.A.N., De Vries, J.P.P.M., Pasterkamp, G., Ackerstaff, R.G.A., Schoneveld, A.H., Velema, E., De Kleijn, D.P. V, and Moll, F.L. (2005). Carotid atherosclerotic plaque characteristics are associated with microembolization during carotid endarterectomy and procedural outcome. Stroke *36*, 1735–1740.

Verma, R., Balakrishnan, L., Sharma, K., Khan, A.A., Advani, J., Gowda, H., Tripathy, S.P., Suar, M., Pandey, A., Gandotra, S., et al. (2016). A network map of Interleukin-10 signaling pathway. J. Cell Commun. Signal. *10*, 61–67.

Vert, J.P. (2010). Reconstruction of Biological Networks by Supervised Machine Learning Approaches. In Elements of Computational Systems Biology, pp. 165–188.

Vickers, K.C., Maguire, C.T., Wolfert, R., Burns, A.R., Reardon, M., Geis, R., Holvoet, P., and Morrisett, J.D. (2009). Relationship of lipoprotein-associated phospholipase A(2) and oxidized low density lipoprotein in carotid atherosclerosis. J. Lipid Res. *50*, 1735–1743.

Viedt, C., Vogel, J., Athanasiou, T., Shen, W., Orth, S., Kubler, W., and Kreuzer, J. (2002). Monocyte

chemoattractant protein-1 induces proliferation and interleukin-6 production in human smooth muscle cells by differential activation of nuclear factor-kappaB and activator protein-1. Arterioscler. Thromb. Vasc. Biol. *22*, 914–920.

Viggers, R.F., Wechezak, A.R., and Sauvage, L.R. (1986). An apparatus to study the response of cultured endothelium to shear stress. J Biomech Eng *108*, 332–337.

Vincent Calvez Nicolas Meunier, Annie Raoult, M.A.E. (2010). Mathematical modelling of the atherosclerotic plaqueformation. ESAIM Proc. *30*, 1–14.

Vincent Calvez Nicolas Meunier, Annie Raoult and Gabriela Rusnakova, J.G.H. (2010). Mathematical and numerical modelling of early atherosclerotic lesions. ESAIM Proc. *30*, 1–14.

Vinkers DJ, Stek ML, van der Mast RC, de Craen AJM, Le Cessie S, Jolles J, Westendorp RGJ, G.J. (2005). Generalized atherosclerosis, cognitive function, and depressive symptoms in old age.

Virmani, R., Virmani, R., Kolodgie, F.D., Kolodgie, F.D., Burke, A.P., Burke, A.P., Farb, A., Farb, A., Schwartz, S.M., and Schwartz, S.M. (2000). Lessons From Sudden Coronary Death. Arterioscler. Thromb. 1262–1275.

Voit, E.O., Martens, H.A., and Omholt, S.W. (2015). 150 Years of the Mass Action Law. PLoS Comput. Biol. *11*, 1–7.

Wada, S., Koujiya, M., and Karino, T. (2002). Theoretical study of the effect of local flow disturbances on the concentration of low-density lipoproteins at the luminal surface of end-to-end anastomosed vessels. Med. Biol. Eng. Comput. *40*, 576–587.

Wade, R.C., Gabdoulline, R.R., Ludemann, S.K., and Lounnas, V. (1998). Electrostatic steering and ionic tethering in enzyme-ligand binding: insights from simulations. Proc. Natl. Acad. Sci. *95*, 5942–5949.

Wagenseil, J.E., and Mecham, R.P. (2009). Vascular extracellular matrix and arterial mechanics. Physiol Rev *89*, 957–989.

Waltemath, D., Adams, R., Beard, D.A., Bergmann, F.T., Bhalla, U.S., Britten, R., Chelliah, V., Cooling, M.T., Cooper, J., Crampin, E.J., et al. (2011). Minimum information about a simulation experiment (MIASE). PLoS Comput. Biol. *7*.

Wang, H.H. (2001). Analytical models of atherosclerosis. Atherosclerosis *159*, 1–7.

Wang, J., and Al-lamki, R.S. (2013). Tumor Necrosis Factor Receptor 2 : Its Contribution to Acute Cellular Rejection and Clear Cell Renal Carcinoma. *2013*.

Wang, K., Sun, J., Zhou, S., Wan, C., Qin, S., Li, C., He, L., and Yang, L. (2013). Prediction of Drug-Target Interactions for Drug Repositioning Only Based on Genomic Expression Similarity. PLoS Comput. Biol. *9*.

Wang, X., Collins, H.L., Ranalletta, M., Fuki, I. V, Billheimer, J.T., Rothblat, G.H., Tall, A.R., and Rader, D.J. (2007). Macrophage ABCA1 and ABCG1, but not SR-BI, promote macrophage reverse cholesterol transport in vivo. J. Clin. Invest. *117*.

Watterson, S., and Ghazal, P. (2010). Use of logic theory in understanding regulatory pathway signaling in response to infection. Future Microbiol. *5*, 163–176.

Watterson, S., Marshall, S., and Ghazal, P. (2008). Logic models of pathway biology. Drug Discov. Today *13*, 447–456.

Watterson, S., Guerriero, M.L., Blanc, M., Mazein, A., Loewe, L., Robertson, K.A., Gibbs, H., Shui, G., Wenk, M.R., Hillston, J., et al. (2013). A model of flux regulation in the cholesterol biosynthesis pathway: Immune mediated graduated flux reduction versus statin-like led stepped flux reduction. Biochimie *95*, 613–621.

Weber, C., and Noels, H. (2011). Atherosclerosis: current pathogenesis and therapeutic options. Nat Med *17*, 1410–1422.

Weiner, S.J., Kollman, P.A., Nguyen, D.T., and Case, D.A. (1986). An all atom force field for simulations of proteins and nucleic acids. J. Comput. Chem. *7*, 230–252.

Weller, H.G., and Tabor, G. (1998). A tensorial approach to computational continuum mechanics using object-oriented techniques. Comput. Phys. *12*, 620–631.

Wendt, H., Leder, L., Härmä, H., Jelesarov, I., Baici, A., and Bosshard, H.R. (1997). Very rapid, ionic strength-dependent association and folding of a heterodimeric leucine zipper. Biochemistry *36*, 204–213.

Westerterp, M., Murphy, A.J., Wang, M., Pagler, T.A., Vengrenyuk, Y., Kappus, M.S., Gorman, D.J., Nagareddy, P.R., Zhu, X., Abramowicz, S., et al. (2013). Deficiency of ABCA1 and ABCG1 in Macrophages Increases Inflammation and Accelerates Atherosclerosis in Mice. Circ. Res. *112*, 10.1161/CIRCRESAHA.113.301086.

Whiting, C.H., and Jansen, K.E. (2001). A stabilized finite element method for the incompressible Navier-Stokes equations using a hierarchical basis. Int. J. Numer. Methods Fluids *35*, 93–116.

Whitlock, G. a, Dack, K.N., Dickinson, R.P., and Lewis, M.L. (2007). A novel series of highly selective inhibitors of MMP-3. Bioorg. Med. Chem. Lett. *17*, 6750–6753.

Whitman, S.C. (2002). Interleukin-18 Enhances Atherosclerosis in Apolipoprotein E-/- Mice Through Release of Interferon-gamma. Circ. Res. *90*, 34e–38.

Whitrow, M.J., and Harding, S. (2010). Asthma in Black African, Black Caribbean and South Asian adolescents in the MRC DASH study: a cross sectional analysis. BMC Pediatr. *10*, 18.

Williams, T., Kelley, C., and many others (2013). Gnuplot 4.6: an interactive plotting program.

Wixon, J., and Kell, D. (2000). The Kyoto encyclopedia of genes and genomes--KEGG. Yeast *17*, 48–55.

World Health Organisation - Mendis S Norrving B editors, P.P. (2011). Global Atlas on Cardiovascular Disease Prevention and Control.

Wu, S., Skolnick, J., and Zhang, Y. (2007). Ab initio modeling of small proteins by iterative TASSER simulations. BMC Biol. *5*, 17.

Xiao, H., Lu, M., Yang, T., and Shyy, J.Y.-J. (2013). SREBP2 Activation of NLRP3 Inflammasome in Endothelium Mediates Hemodynamic-Induced Atherosclerosis Susceptibility. Circulation *128*, 632–642.

Xie, Z.-R., and Hwang, M.-J. (2015). Methods for Predicting Protein–Ligand Binding Sites BT - Molecular Modeling of Proteins. A. Kukol, ed. (New York, NY: Springer New York), pp. 383–398.

Xie, Z.-R., Chen, J., and Wu, Y. (2017). Predicting Protein–protein Association Rates using Coarse-grained Simulation and Machine Learning. Sci. Rep. *7*, 46622.

Xu, J., and Zhang, Y. (2010). How significant is a protein structure similarity with TM-score = 0.5? Bioinformatics *26*, 889–895.

Xu, D., Trajkovic, V., Hunter, D., Leung, B.P., Schulz, K., Alastair Gracie, J., McInnes, I.B., and Liew, F.Y. (2000). IL-18 induces the differentiation of Th1 or Th2 cells depending upon cytokine milieu and genetic background. Eur. J. Immunol. *30*, 3147–3156.

Xu, J., Jiao, F., and Yu, L. (2008). Protein structure prediction using threading. Methods Mol. Biol. *413*, 91–121.

Xu, Q., Metzler, B., Jahangiri, M., and Mandal, K. (2012). Molecular chaperones and heat shock proteins in atherosclerosis. Am J Physiol Hear. Circ Physiol *302*, 506–514.

Xue, C., Friedman, A., and Sen, C.K. (2009). A mathematical model of ischemic cutaneous wounds. Proc. Natl. Acad. Sci. U. S. A. *106*, 16782–16787.

Xue, L.C., Rodrigues, J.P., Kastritis, P.L., Bonvin, A.M., and Vangone, A. (2016). PRODIGY: A web server for predicting the binding affinity of protein-protein complexes. Bioinformatics *32*, 3676–3678.

Yahagi, K., Kolodgie, F.D., Otsuka, F., Finn, A. V., Davis, H.R., Joner, M., and Virmani, R. (2016). Pathophysiology of native coronary, vein graft, and in-stent atherosclerosis. Nat. Rev. Cardiol. *13*, 79–98.

Yamamoto, M., Aoyagi, M., Tajima, S., Wachi, H., Fukai, N., Matsushima, Y., and Yamamoto, K. (1997). Increase in Elastin Gene Expression and Protein Synthesis in Arterial Smooth Muscle Cells Derived From Patients with Moyamoya Disease. Stroke *28*, 1733–1738.

Yang, N., and Vafai, K. (2006). Modeling of low-density lipoprotein (LDL) transport in the artery-effects of hypertension. Int. J. Heat Mass Transf. *49*, 850–867.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2014). The I-TASSER Suite: protein structure and function prediction. Nat. Methods *12*, 7–8.

Yang, J., Wang, Y., and Zhang, Y. (2016). ResQ: An Approach to Unified Estimation of B-Factor and Residue-Specific Error in Protein Structure Prediction. J. Mol. Biol. *428*, 693–701.

Yang, L., Wang, K., Chen, J., Jegga, A.G., Luo, H., Shi, L., Wan, C., Guo, X., Qin, S., He, G., et al. (2011). Exploring off-targets and off-systems for adverse drug reactions via chemical-protein interactome - clozapine-induced agranulocytosis as a case study. PLoS Comput. Biol. *7*.

Yang, M., Chesterman, C.N., and Chong, B.H. (1995). Recombinant PDGF enhances megakaryocytopoiesis in vitro. Br. J. Haematol. *91*, 285–289.

Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L., et al. (2016). Ensembl 2016. Nucleic Acids Res. *44*, D710–D716.

Yates, a., Beal, K., Keenan, S., McLaren, W., Pignatelli, M., Ritchie, G.R.S., Ruffier, M., Taylor, K., Vullo, a., and Flicek, P. (2014). The Ensembl REST API: Ensembl Data for Any Language. Bioinformatics *31*, 143–145.

Yeo, B., Turner, N.C., and Jones, A. (2014). An update on the medical management of breast cancer. Bmj *348*, g3608–g3608.

Yokochi, S., Hashimoto, H., Ishiwata, Y., Shimokawa, H., Haino, M., Terashima, Y., and Matsushima, K. (2001). An anti-inflammatory drug, propagermanium, may target GPI-anchored proteins associated with an MCP-1 receptor, CCR2. J.Interferon Cytokine Res. *21*, 389–398.

Yoshida, H., and Kisugi, R. (2010). Mechanisms of LDL oxidation. Clin. Chim. Acta *411*, 1875–1882.

Young, I.S., and McEneny, J. (2001). Lipoprotein oxidation and atherosclerosis. Biochem. Soc. Trans. *29*, 358–362.

Young, D.A., Hegen, M., Ma, H.L.R., Whitters, M.J., Albert, L.M., Lowe, L., Senices, M., Wu, P.W., Sibley, B., Leathurby, Y., et al. (2007). Blockade of the interleukin-21/interleukin-21 receptor pathway ameliorates disease in animal models of rheumatoid arthritis. Arthritis Rheum. *56*, 1152–1163.

Young, L., Jernigan, R.L., and Covell, D.G. (1994). A role for surface hydrophobicity in protein-protein recognition. Protein Sci. *3*, 717–729.

Yu, H.T., Oh, J., Chang, H.-J., Lee, S.-H., Shin, E.-C., and Park, S. (2015a). Serum monokine induced by gamma interferon as a novel biomarker for coronary artery calcification in humans. Coron. Artery Dis. *26*, 317–321.

Yu, T., Lloyd, C.M., Nickerson, D.P., Cooling, M.T., Miller, A.K., Garny, A., Terkildsen, J.R., Lawson, J., Britten, R.D., Hunter, P.J., et al. (2011). The Physiome Model Repository 2. Bioinformatics *27*, 743–744.

Yu, X., Martinez, M., Gable, A.L., Fuller, J.C., Bruce, N.J., Richter, S., and Wade, R.C. (2015b). WebSDA: A web server to simulate macromolecular diffusional association. Nucleic Acids Res. *43*, W220–W224.

Yu, X.H., Fu, Y.C., Zhang, D.W., Yin, K., and Tang, C.K. (2013). Foam cells in atherosclerosis. Clin. Chim. Acta *424*, 245–252.

Yuan, F., Chien, S., and Weinbaum, S. (1991). A new view of convective-diffusive transport processes in the arterial intima. J. Biomech. Eng. *113*, 314–329.

Yugandhar, K., and Gromiha, M.M. (2014). Protein-protein binding affinity prediction from amino acid sequence. Bioinformatics *30*, 3583–3589.

Yvan-Charvet, L., Wang, N., and Tall, A.R. (2010). Role of HDL, ABCA1, and ABCG1 transporters in cholesterol

efflux and immune responses. Arterioscler. Thromb. Vasc. Biol. *30*, 139–143.

Zalewski, A., and Macphee, C. (2005). Role of lipoprotein-associated phospholipase A2 in atherosclerosis: Biology, epidemiology, and possible therapeutic target. Arterioscler. Thromb. Vasc. Biol. *25*, 923–931.

Zambon, A., Bertocco, S., Vitturi, N., Polentarutti, V., Vianello, D., and Crepaldi, G. (2003). Relevance of hepatic lipase to the metabolism of triacylglycerol-rich lipoproteins. Biochem. Soc. Trans. *31*, 1070–1074.

Zernecke, A., and Weber, C. (2010). Chemokines in the vascular inflammatory response of atherosclerosis. Cardiovasc. Res. *86*, 192–201.

Zernecke, A., Shagdarsuren, E., and Weber, C. (2008). Chemokines in atherosclerosis an update. Arterioscler. Thromb. Vasc. Biol. *28*, 1897–1908.

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. BMC Bioinformatics *9*, 40.

Zhang, Y., and Skolnick, J. (2005). TM-align: A protein structure alignment algorithm based on the TM-score. Nucleic Acids Res. *33*, 2302–2309.

Zhang, S., Ritter, L.R., and Ibragimov, a. I. (2013). Foam cell formation in atherosclerosis: HDL and macrophage reverse cholesterol transport. Discret. Contin. Dyn. Syst. - Ser. S 825–835.

Zhao, Q. (2010). Dual targeting of CCR2 and CCR5: therapeutic potential for immunologic and cardiovascular diseases. J. Leukoc. Biol. *88*, 41–55.

Zhao, W., Oskeritzian, C.A., Pozez, A.L., and Schwartz, L.B. (2005). Cytokine production by skin-derived mast cells: endogenous proteases are responsible for degradation of cytokines. J. Immunol. *175*, 2635–2642.

Zimmer, S., Grebe, A., Bakke, S.S., Bode, N., Halvorsen, B., Ulas, T., Skjelland, M., Nardo, D. De, Labzin, L.I., Kerksiek, A., et al. (2016). HHS Public Access. *8*.

Zohdi, T.I., Holzapfel, G. a., and Berger, S. a. (2004). A phenomenological model for atherosclerotic plaque growth and rupture. J. Theor. Biol. *227*, 437–443.