Title (150 char)

Is systems pharmacology ready to impact upon therapy development? A study on the cholesterol biosynthesis pathway.

Helen Benson[1¶*], Steven Watterson[2§*], Joanna Sharman[1], Chido Mpamhanga[3±], Andrew Parton[2], Christopher Southan[1], Anthony Harmar[3], Peter Ghazal[45]

[1]Centre for Integrative Physiology, University of Edinburgh, Hugh Robson Building, George Square, Edinburgh, Scotland, EH8 9XD

[2]Northern Ireland Centre for Stratified Medicine, University of Ulster, C-Tric, Altnagelvin Hospital Campus, Derry, Northern Ireland, BT47 6SB

[3]Centre for Cardiovascular Science, University of Edinburgh, The Queen's Medical Research Institute, 47 Little France Crescent, Edinburgh, Scotland, EH16 4TJ

[4]Division of Pathway Medicine, University of Edinburgh Medical School, 49 Little France Crescent, Edinburgh, Scotland, EH16 4SB

[5]Centre for Synthetic and Systems Biology, CH Waddington Building, King's Buildings, Mayfield Road, Edinburgh, Scotland, EH9 3JD

[§]Corresponding author s.watterson@ulster.ac.uk

[*]joint First Authors.

[±]Current address: LifeArc, Accelerator Building, SBC Open Innovation Campus, Stevenage, SG1 2FX, UK

[¶]Current address: TB Section, Respiratory Disease Department, National Infection Service, Public Health England, 61 Colindale Avenue, London, NW9 5EQ

Running title (60 char)

The feasibility of systems pharmacology

Author contribution

This work was conceived by HB, SW, PG and AH. The analysis and data compilation was undertaken by HB, SW, JS, and AP. The manuscript was written by HB, SW, JS, CM, AP, CS and PG.

Abstract

Background and Purpose An ever-growing wealth of information on current drugs and their pharmacological effects is available from online databases. As our understanding of systems biology increases, we have the opportunity to predict, model and quantify how drug combinations can be introduced that outperform conventional single-drug therapies. Here, we explore the feasibility of such systems pharmacology approaches with an analysis of the mevalonate branch of the cholesterol biosynthesis pathway.

Experimental Approach Using open online resources, we assembled a computational model of the mevalonate pathway and compiled a set of inhibitors directed against targets in this pathway. We used computational optimisation to identify combination and dose options that show not only maximal efficacy of inhibition on the cholesterol producing branch but also minimal impact on the geranylation branch, known to mediate the side effects of pharmaceutical treatment.

Key Results We describe serious impediments to systems pharmacology studies arising from limitations in the data, incomplete coverage and inconsistent reporting.   By curating a more complete dataset, we demonstrate the utility of computational optimization for identifying multi-drug treatments with high efficacy and minimal off-target effects.

Conclusion and Implications We suggest solutions that facilitate systems pharmacology studies, based on the introduction of standards for data capture that increase the power of experimental data. We propose a systems pharmacology work-flow for the refinement of data and the generation of future therapeutic hypotheses.

Non-approved Abbreviations

API - Application Program Interface

BPS - British Pharmacological Society

BRENDA - Braunschweig Enzyme Database

CID – Compound Identifier

FDA – US Food and Drug Administration

FDFT1 - Farnesyl-diphosphate farnesyl transferase 1

FDPS - Farnesyl diphosphate synthase

GtoPdb - Guide to Pharmacology Database

HMGCR – Hydroxymethylglutaryl-CoA reductase

HMGCS1 - Hydroxymethylglutaryl-CoA synthase

HPC – High Performance Computing

KEGG - Kyoto Encyclopedia of Genes and Genomes

IDI1 - Isopentenyl diphosphate delta isomerase 1

IDI2 - Isopentenyl diphosphate delta isomerase 2

InChlKey - International Chemical Identifier

INN - International Nonproprietary Name

IUBMB – International Union of Biochemistry and Molecular Biology

IUPAC – International Union of Pure and Applied Chemistry

IUPHAR - International Union of Basic and Clinical Pharmacology

MIABE – Minimum Information About a Bioactive Entity

MIAPE – Minimum Information About a Proteomic Experiment

MIBBI – Minimum Information for Biological and Biomedical Investigations

MIMIx – Minimum Information About a Molecular Interaction Experiment

MVD - Diphosphomevalonate decarboxylase

N2S – Name to Structure

NC-IUBMB – Nomenclature Committee of the International Union of Biochemistry and Molecular Biology

NIH – US National Institute for Health

OBO Foundry - Open Biomedical Ontologies (OBO) Foundry

ODE – Ordinary Differential Equation

Open PHACTS - Open Pharmacological Concept Triple Store

PKPD – Pharmacokinetic-Pharmacodynamic

PMVK - Phosphomevalonate kinase

QC – Quality Control

SBGN – Systems Biology Graphical Notation

SBGN-ML - Systems Biology Graphical Notation Markup Language

SBML – Systems Biology Markup Language

SMILES - Simplified Molecular-Input Line-Entry System

STRENDA - Standards for Reporting Enzymology Data

WHO – World Health Organisation

Supporting Information

Our supporting information includes the following files.

Supporting figure 1. A sensitivity analysis of the impact of pairs of drugs on the pathway.

Supporting figure 2. A representative reaction from the IUPHAR/BPS GuidetoPharmacology (GtoPdb).

Supporting table 1. A description of the databases and online resources used in this study.

Supporting table 2. A description of the inhibitors used in our optimisation calculations.

Supporting table 3. A list of the best performing drug combinations identified as intervention hypotheses.

A SBGN-ML file describing the mevalonate pathway shown in Figure 1. This is machine readable and interpretable as a biological pathway and is encoded using SBGN, an XML based mark up.

A SBML file describing the kinetics of the pathway shown in Figure 1. This is encoded using a XML based markup to describe the interactions and parameters of the pathway.

Introduction

The expansion of available genomic and proteomic data has enhanced our understanding of biomolecular interaction networks. Consequently, the development of systems biology approaches has enabled us to better understand how cellular behaviour emerges from these networks (Boran, 2010a). Systems-level approaches have been used to predict the on- and off-target impacts of an intervention (Boran, 2010b) and to identify the most sensitive components in pathways that suggest candidate drug targets (Benson et al. 2013). They also have the untapped potential to suggest therapies comprising combinations of drugs chosen to strategically reprogram biomolecular interaction networks in order to drive the system from a diseased to a healthy state (Zhao et al. 2013; van Hasselt, 2015; Watterson, 2010). This approach, known as *Systems Pharmacology* (Boran, 2010b; Westerhoff, 2015) is underpinned by the expansion in pathway, pharmacology and medicinal chemistry databases.

For example, WikiPathways held 804 human pathways[1] with 253 added in 2015 (Kutmon et al. 2016). KEGG PATHWAY holds 518 pathway maps (Kanehisa, 2017)[2]. Reactome currently holds 2148 human pathways involving 10684 proteins and isoforms[3] (Croft, 2014; Fabregat, 2016). ChEMBL version 23 (Gaulton et al. 2017) includes 14675320 bioactivities and the IUPHAR/BPS Guide to Pharmacology (GtoPdb) contains 15281 curated interactions in its 2017.5 release (Southan et al. 2016). In 2016, the Food and Drug Administration (FDA) new drug approvals fell to 22 , following 45 approvals in 2015 (US Food and Drug Administration 2016a ; US Food and Drug Administration 2016b). According to DrugBank release 5.0, their distinct molecular count of approved small-molecule drugs is 2037 (Law et al, 2014).

As this catalogue of pharmacological interactions grows and our understanding of pathway systems expands, it will be advantageous to integrate these resources in order to devise new potential therapies. Drug combination based interventions represent an opportunity for therapy development that can yield one-size-fits-all or personalized/stratified therapies and they can target pathways precisely rather than perturbing entire networks. Two NIH workshop white papers have made a strong case for Systems Pharmacology (Sorger et al. 2011) as a way to reduce attrition in therapy, to stimulate drug development, to bridge the gap

---

[1] http://www.wikipathways.org/index.php/WikiPathways:Statistics
[2] http://www.kegg.jp/kegg/docs/statistics.html
[3] http://reactome.org/stats.html

between network biology and translational medicine and to enhance industrial-academic collaborations. Systems Pharmacology is also likely to impact upon genomic medicine (Westerhoff, 2015), Systems Pathology, Systems Biology and Pharmacometrics (van der Greef, 2005; Vicini, 2013) and the tools that could contribute to Systems Pharmacology have been described (Kell, 2014; Berger, 2009; Lehar, 2007).

Prior work under the domain of Systems Pharmacology, has primarily focussed on pharmacokinetic-pharmacodynamic (PKPD) modelling (Darwich, 2017). Industry has evaluated Systems Pharmacology as a tool to inform trial design in areas of cardiovascular disease, endocrinology, neurogenerative disease, respiratory disease, oncology and infectious disease (Visser, 2014) and to inform regulatory development (Visser, 2014; Peterson et al., 2015). There have a been a number of specific studies of nerve growth factor (Benson, 2013), coagulation (Wajima, 2009), innate immunity (Madrasi, 2014), cancer (Abaan, 2013) and atherosclerosis (Pichardo-Almarza, 2015).

However, whilst there is much enthusiasm for Systems Pharmacology as a tool to improve the efficacy and safety of the drug development pipeline (Trame, 2016; van der Graaf, 2011; Rostami-Hodgegan, 2012), the practical challenges of systematically amalgamating pharmacology and pathway biology in a coherent framework have not been adequately addressed.

Here, we describe a Systems Pharmacology study of the cholesterol biosynthesis pathway, detailing the barriers to progress that we encountered and suggesting solutions to these impediments, before proposing a model of how Systems Pharmacology studies could be conducted in future. In particular, we build a dynamic ordinary differential equation (ODE) model of the pathway, which we parameterize as far as possible from the literature. We identify relevant pharmacological agents that act on this pathway and parameterize them as far as possible from the literature and online databases. We then use computational optimization techniques to identify the drug combinations that are most effective at suppressing the outputs of the pathway that lead to cholesterol production and that minimise off-target effects. In completing our analysis, we identify many of the problems that prevent this type of work being undertaken routinely and we suggest solutions that would enable systems pharmacology to make a regular contribution to therapy development.

As explored in previous studies (Caspi et al. 2016; Watterson et al. 2013; Bhattacharya et al. 2014; Mazein et al. 2013; Mazein et al. 2011), the cholesterol biosynthesis pathway is critical to both cardiovascular health (Lewington et al. 2007; Henderson et al. 2016; Parton et al. 2016) and innate immunity (Blanc et al. 2011; Lu et al. 2015; Robertson et al. 2016). As the target of the statin class of drug, we would expect this pathway to be amongst the most thoroughly characterized and for this reason we have chosen it for our feasibility study of Systems Pharmacology. For simplicity, we focus on the segment of the pathway that transforms acetyl-CoA to squalene and that forks to produce geranylgeranyl-diphosphate. As a precursor to cholesterol, we would expect squalene synthesis to track cholesterol synthesis and so we use this as a proxy. The branch of the pathway that produces geranylgeranyl-diphosphate has been shown to mediate both the innate immune response (Blanc et al. 2011) and the myopathy side-effects associated with statin treatment (Wagner et al. 2011). Any intervention that demonstrates a minimal impact on this branch will avoid one of the significant side-effects associated with standard cholesterol lowering treatments.

Methods

Pathway production

We started from the representations available in KEGG (Kanehisa et al. 2014), MetaCyc (Caspi et al. 2016) and the GtoPdb (Southan et al. 2016) taking these resources to be representative of the community of online pathway databases. We reviewed the primary literature to establish the structure of the mevalonate portion of the cholesterol biosynthesis pathway, in particular the enzymes involved in the pathway, the reactions they catalyse, their subcellular localisation, the species in which they were identified and any known isoforms.

Diagrams of the pathway were created using the SBGN standardised notation (Le Novere et al. 2009), the yEd diagram software (yWorks GmbH, http://www.yworks.com/products/yed) and the SBGN-ED add-on to VANTED (Czauderna et al. 2010). From these diagrams, we built kinetic models as systems of Ordinary Differential Equations (ODEs).

The ODE model of this pathway was built using Michaelis-Menten kinetics to describe each step except the interactions consuming isopentenyl-PP and producing geranylgeranyl-PP and pre-squalene-PP. These steps were described using mass action kinetics in order to simplify the process of calculating the steady state of the model and hence the steady state behavior of the pathway. Mass action kinetics were justified by the expectation that the pathway

interactions would operate far from substrate saturation making the dynamics robust against small fluctuations in enzyme concentration. Mass action rate constants were calculated from the K_{cat}, K_m and K_i parameters as described elsewhere (Watterson et al. 2013) and enzyme concentrations were taken from experimentally measured values (Watterson et al. 2013).

The pathway map and the associated mathematical model are available from the supporting material as SBGN-ML (van Iersel et al. 2012) and SBML files (Hucka et al. 2003), respectively.

Pathway parameterisation

We identified the kinetic parameters that quantify each reaction unambiguously (K_m, K_{cat}) using the BRENDA enzyme database (Chang et al. 2015) and verified the values described against the primary literature. In many instances enzymes were associated with multiple kinetic parameter sets. We selected kinetic parameters based upon the following criteria: i) specificity to the wild-type enzyme in one of the three main mammalian model species: human, mouse or rat, ii) sourced from a primary literature reference describing *in vivo* or *in vitro* experimental data as opposed to computationally-derived structural modelling data and iii) sourced from a reference that could be accessed and therefore verified. For many enzymes this yielded a range of values for each parameter and where this was the case, we used the mean of the values obtained.

Inhibitor list

Inhibitor compounds not already indexed in GtoPdb were identified for each reaction from ChEMBL and BRENDA, databases that we took to be representative of the community of target databases. We included a compound in our set if it met three criteria: i) the enzyme used in the assay was wild-type from one of the three main mammalian model species: human, mouse or rat, ii) an experimentally determined reaction-specific inhibition constant (K_i) was reported and iii) the assay conditions were reported. Crucially, all data were checked against the primary literature references. Where this yielded a range of inhibition constants for nominally identical compounds the most potent K_i values were used.

We verified the correct chemical structures of the inhibitors by cross-referencing the original references against the online chemical databases PubChem (Kim et al. 2016) and

ChemSpider (Pence & Williams 2010). The actual chemical structures of the marketed statin drugs were established by checking the FDA labels and the International Non-Proprietary Name (INN)-assigned structures on the WHO MedNet site (https://mednet-communities.net/inn). Comparison of unique structural identifiers allowed us to identify duplicates within the ChEMBL, BRENDA and literature-derived dataset, and to establish whether the chemical structure reported in a given reference matched the marketed drug or research compound structures.

Curated content describing the enzymes in this pathway, their substrates and small molecule inhibitors was used to consolidate and expand GtoPdb using the same approach and guidelines as described elsewhere (Pawson et al. 2014). The enzymes, list of inhibitors and kinetic parameters are now all updated in the July 2016.3 release of GtoPdb.

Hypothesis generation

We combined ODE kinetic models, the pathway parameters and the inhibitor parameters to create a model describing the dynamics of the mevalonate pathway. We sought to identify the drug combination that would best suppress the production of squalene as a precursor for cholesterol, but would also maintain production of geranylgeranyl-diphosphate at the same levels as in the absence of any inhibitors, thereby eliminating a significant side-effect of treatment. Firstly, we identified the steady-state activity of the pathway in the absence of any inhibitors. Then we used computational optimisation to identify the drug combination that, at steady state, minimised squalene production, but left geranylgeranyl diphosphate production the same as in the absence of inhibitors.

This was implemented using the Genetic Algorithm function available on Matlab (MathWorks, http://www.mathworks.com) in parallel with a population size of 200 and a function tolerance of 10^{-6}. Matlab was chosen as the modelling platform for its flexibility, stability and comprehensive libraries. The genetic algorithm started with one instance of a set of drug concentrations where each drug was assigned a concentration equal to its K_i. A 199 further instances of sets of drug concentrations were automatically generated from this instance by adding Gaussian noise to the concentration of each drug (with standard deviation 1, the default setting). These 200 instances comprised the first generation of candidate interventions. All instances of sets of concentrations were evaluated for their efficacy at suppressing squalene synthesis whilst maintaining geranylgeranyl-PP production. 200 new

instances were created as a second generation of candidate interventions from the two most effective instances of the first generation and with the addition of Gaussian noise. The 200 new instances were then themselves evaluated with the two most effective instances used to generate a further 200 new instances, the third generation. This process was iterated until we arrived at instances from which no improvement in efficacy could be found for 20 consecutive generations at which point we interpreted the best performing instance identified thus far as optimal.

Nomenclature of Targets and Ligands

Key protein targets and ligands in this article are hyperlinked to corresponding entries in http://www.guidetopharmacology.org, the common portal for data from the IUPHAR/BPS Guide to PHARMACOLOGY (Southan et al., 2016), and are permanently archived in the Concise Guide to PHARMACOLOGY 2015/16 (Alexander et al., 2015).

Results

Pathway Production

We produced the model of the mevalonate arm of the cholesterol biosynthesis pathway shown in Figure 1 in SBGN notation, describing the sequence of metabolic steps that lead from acetyl-CoA and acetoacetyl-CoA consumption to squalene and geranylgeranyl-diphosphate production. This pathway comprises 12 steps (see Table 1), involving 10 enzymes and 14 metabolites.

The parameters required for the resulting ODE model are shown in Table 1. After pooling results across mouse, human and rat models we were able to obtain experimental values for only 12 out of the 24 required parameters. Across the studies reported, pH values ranged from 7.0 to 8.0 and temperatures ranged from 25°C to 37°C, although in some studies neither pH nor temperature values were given. When verified against the primary references, we found that one parameter value obtained from BRENDA was missing from the literature reference provided, suggesting that it had been misattributed (K_{cat}=0.023/s for HMGCR). A second parameter had been transcribed (for MVD) where the literature source contradicted itself, specifying K_m=10μM in the abstract and K_m=10mM in the manuscript. Because computational hypothesis generation is highly sensitive to the values of the parameters, ambiguous or inaccurate reporting can have a significant impact on any predictions made.

Substrates were reported in varying levels of structural detail. Common names were used that could refer to multiple explicit forms of a chemical structure. However, variations in the chirality and chemical structure can significantly affect substrate affinity. The relative enzyme concentrations had been inferred previously (Watterson et al, 2013) and are listed in Table 2.

Supplementary Table 1 compares representations of the cholesterol biosynthesis pathway across the main publicly available pathway and chemical databases. It includes a summary of cross-referencing between databases with standard identifiers for unambiguous representation, which will be essential for future cross-platform interoperability.

Inhibitors

The inhibitors obtained from GtoPdb, BRENDA and the literature, together with their inhibition constants (K_i), are listed in Table 3. Six of the ten enzyme targets had quantified parameters in humans. It was necessary to include two inhibitors that had been only reported for rat enzymes (L-659,699 for HMGCS1 and 3-hydroxy-3-methyl-6-phosphohexanoic acid for PMVK) in order to maximise coverage of the pathway. Two enzyme paralogues (IDI1 and IDI2) had no reported inhibitors with available K_i values, representing a region of the pathway that cannot currently be modulated in our modelling process. This can be contrasted with the enzymes HMGCR and FDPS, each of which had an extensive list of inhibitors. Inhibition constants could be obtained for 8 of the 10 enzymes in the pathway. Where reported, these values came from studies conducted across a range of pH levels from 6.8 to 7.5 and temperatures from 25°C to 37°C.

Both explicit structure and name-to-structure (n2s) ambiguities existed around the reporting of inhibitor entities. In some cases, the common or trade name of a compound was used, without specification of the exact chemical structure and stereochemistry. In other cases we found a different n2s assignment across different database resources or indeed within the same resource. For example, under the HMGCS1 entry of BRENDA, the same inhibitor is listed twice as L-659,699 and (E,E)-11-[3-(hydroxymethyl)-4-oxo-2-oxytanyl]-3,5,7-trimethyl-2,4-undecadienenoic acid.

Several results recorded in ChEMBL were transcribed against the incorrect drug target. Three inhibitors listed against the enzyme HMGCS1 describe results obtained from experiments with HMGCR (Balasubramanian et al. 1989). There were also cases where the incorrect species had been recorded. For example, the compound with ChEMBL ID CHEMBL88601 cited in one study (Procopiou et al. 1994) (ChEMBL document ID CHEMBL1151052) is listed against the human FDFT1 enzyme, while in fact the paper describes results for the yeast *Candida albicans* and rats.

Hypothesis generation

In order to complete the gaps in the available parameter sets, we proceeded by assuming that where parameters were taken from separate studies, the same metabolite chemical structures were referenced. For all the unknown parameters, we substituted a single representative value, obtained by averaging across all known corresponding parameters.

Calculating the steady state behavior of the pathway in the absence of inhibitors yielded the profile of flux shown on the left of Fig 2A, which we take to be wild-type behavior. Using computational optimisation, we identified the following drug combination that produced the steady state profile of flux shown in the middle of Fig 2A and in Fig 2B: L-659,699=0.0294nM, Rosuvastatin=2.60nM, Farnesyl Thiodiphosphate=0.0340nM, Cinnamic acid=0.00104nM, 6-fluoromevalonate 5-diphosphate=0.0213nM, zoledronic acid=9.97nM, BPH-628=5.86nM; Zaragozic acid A=0.755nM (see Table 3 and supplementary Tables 2 and 3). Here, the production of squalene, a precursor of cholesterol is heavily suppressed and the production of geranylgeranyl-PP is maintained at wild-type levels. In Figure 2B, we see specifically the flux at endpoints of the two pathway branches. With this drug combination, the flux from Geranylgeranyl-PP → Protein Prenylation is the same between the wild- type (inhibitor free) case and the optimal multi-drug intervention case. However, the flux from Squalene → Cholesterol synthesis has been significantly suppressed.

In Figs 2A and 2B, we compare the flux profiles for wild-type and optimal multi-drug interventions to the case where Rosuvastatin, a type of statin, is applied alone. This inhibitor targets the interaction catalysed by HMGCR and we chose a concentration sufficient to suppress the rate of squalene formation and consumption to the same extent as the multi-drug intervention. As can be seen in Fig 2B Rosuvastatin intervention impacts upon both branches

of the pathway, suppressing Geranylgeranyl-PP formation and protein prenylation as an off-target effect of treatment.

Interestingly, a concentration of 362nM Rosuvastatin was required to achieve the same level of squalene suppression as the multi-drug intervention. The greatest individual drug concentration required in the optimal multi-drug intervention was 9.97nM and the total combined concentration was 19.3nM, a dramatically lower dosage.

The value of drug combinations can also be seen in Supplementary Fig 1 where we consider the impact of pairs of drugs (Lehar et al, 2007). Here we see that drug pairs with targets above the fork inhibit the flux though both pathway endpoints (Supplementary Fig. 1A & B). Drug pairs with targets above and below the fork together inhibit the flux through the cholesterol synthesizing branch (Supplementary Fig. 1C & D). However, drug pairs with targets above and below the fork at high doses can have a low impact on the flux through the protein prenylation branch (Supplementary Fig. 1E & F). Critically, Supplementary Fig. 1 B & E show that concentrations can be selected that significant suppress the cholesterol synthesising branch, but that do not suppress the protein prenylation branch. The results demonstrate comparable impact to the multi-drug intervention described above, but at higher individual and combined concentrations.

In order to identify the optimal multi-drug combination, it was necessary to use a high performance computing (HPC) platform. However, the HPC demands were modest. Using a 8 node desktop computer running MATLAB in parallel, we can see that the score (a dimensionless value, greater than or equal to zero, that quantifies how effectively the best performing multi-drug intervention identified achieves our objective, with zero indicating success) converges rapidly on an effective drug combination. It successfully identified an optimal combination in 46 minutes and achieved an approximately optimal solution in less than 10 minutes.

The results of our curation of the pathway and the inhibitors that target it are available in GtoPdb at
http://www.guidetopharmacology.org/GRAC/FamilyDisplayForward?familyId=104
an example of which is shown in supplementary Fig 2.

The model produced is available from Biomodels.org (Chelliah et al. 2013) (ID: MODEL1506220000).

Discussion

The importance of Systems Pharmacology

Multi-drug interventions

Multi-drug approaches are already employed in areas including HIV and oncology (Thakur, 2012; Petrelli, 2008). However, the existing interventions have typically been developed heuristically, rather than through systematic studies of the pathways involved, requiring significant domain expertise and subjective judgment. Systems Pharmacology introduces objective metrics that have the potential to transform therapy development, yielding therapeutic hypotheses more rapidly and cost-effectively.

Many diseases are multifactorial in nature, involving multiple pathways in their pathology. Effective future therapies will likely employ multi-drug approaches that target multiple points in the network of pathways responsible (i.e. polypharmacology). Promiscuous drugs can be incorporated advantageously into the generation of these hypothetical interventions, provided that their interactions are known and parameterized.

Multi-drug approaches can minimise the pleotropic effects of an intervention. As we demonstrated for statins, where a single drug intervention suppressed the output of a pathway to the same extent as multiple drugs targeting the same pathway, not only was the dose of each of the multiple drugs significantly lower than the dose of the single drug, but also the combined dose of all of the multiple drugs was significantly lower than the dose of the single drug. This intrinsically reduces the risk from off-target or pleotropic effects for each drug.

The systems pharmacology approach allows us to predict multi-drug strategies that may be optimal to treat a disease and can be used as a prioritization triage for future drug development. It can support personalized and stratified medicine, where we adapt the parameter sets of the underlying models of pathway activity to represent an individual (for personalized medicine) or a subpopulation (for stratified medicine) and we develop interventions that are customized to be optimal for the patient or patient group. A challenge lies in developing optimised therapies so that they preferentially target key tissues. Pathway models and pharmacological interactions can be made tissue specific by generating a new

parameter set for each tissue. Hypothesis generation would then use optimization to determine an intervention that impacted upon a key pathway in a key tissue, leaving other pathways unchanged across all tissues and with a minimal impact on the key pathway in non-targeted tissues.

Drug development

Few multi-drug treatments make it through the development process. The number of combinational therapies listed in the Therapeutic Target Database at the time of writing is 115 (Qin et al. 2014). A combination therapy, LCZ696, with the brand name *Entresto,* was approved in 2015 and is in Phase III of clinical trials for the treatment of cardiovascular disorders. Establishing drug combinations using a conventional drug development pipeline creates significant challenges as development essentially replicates the single drug development process multiple times. Systems Pharmacology is therefore critical to expanding the range of multi-drug interventions available in a cost effective manner. Although it may add extra steps to the pre-clinical stages of the drug development process, it could have a significant positive impact on the cost-efficiency associated with each success by reducing the attrition rate in the later stages of the pipeline (Bowes et al. 2012).

Integrating our understanding of pharmacology and systems biology will also enable us to make better predictions of the behavior of individual drugs. For example, squalene synthase (FDFT1) has been investigated as a potential drug target that lies downstream of HMGCR, the target for statins, in the cholesterol biosynthesis pathway (see Figure 1). FDFT1 catalyses an interaction after the fork to geranylgeranyl-diphosphate production and it has been speculated that squalene synthase inhibitors might suppress cholesterol production without impacting on the geranylgeranyl-diphosphate producing branch, in contrast to statin treatment. However, squalene synthase inhibitors typically have K_i values orders of magnitude greater than the typical K_i for statins (See Table 3b). As a result, squalene synthase inhibitor concentrations are required to be orders of magnitude greater than statin concentrations to suppress the corresponding enzyme activity comparably. Such high concentrations risk unforeseen off-target effects, making squalene synthase inhibitors a higher risk drug to develop.

Systems level Analysis

At the heart of Systems Pharmacology is the growing recognition that we will only be able to truly understand the best ways to therapeutically intervene in physiological function by considering biology at a systems level. The network of interactions that mediate physiological function is a dynamical system and just as health and disease are different dynamical states of cells, tissues and organs, they also describe different dynamical states of the networks (Ahn, 2006). In a network context, dynamical states can comprise a single stable configuration of the whole network or a sequence of configurations that repeat cyclically and stably. However, it is the configuration (species concentrations, distributions and structural conformations) of the network as a whole, or at least of critical subnetworks, that relate to phenotype, rather than any single component of the network (Lewis, 1991).

Small networks often yield dynamics that are intuitive and predictable. However, as networks become larger and richer in structure, novel and often counter-intuitive dynamics can emerge and it will only be once we are able to build high-confidence models at this scale that the full potential of systems level studies will be realized (Aderem, 2005). Building high confidence networks at this scale is inherently challenging as we see here. Coherently and unambiguously parameterizing all the interactions of a network is a significant logistical challenge. However, we have also seen that doing so enables us to identify and address the side-effects of treatment whilst the therapy is being designed, rather than retroactively. Hence systems levels approaches are well suited to pharmacological applications.

Current impediments to Systems Pharmacology

Problem 1: Lack of systematic recording

The absence of systematic and rigorous descriptions of metabolites and pharmaceutical compounds poses a significant challenge. Example 1, <u>fluvastatin</u> consists of two enantiomers, represented by PubChem Compound Identifiers (CIDs) 1548972 and 446155, with the 3R, 5S enantiomer (CID 446155) being significantly more pharmacologically active than the other (Boralli et al. 2009; Di Pietro et al. 2006). Commercial preparations used *in vitro* often vary in their stereochemical composition, with both enantiomers available individually, as well as in a racemic mixture. However, authors did not always specify the stereochemical composition used despite this necessarily impacting upon the inhibition constant, K_i, reported. Example 2, mevalonate is a metabolite that occurs naturally in mammals as the <u>(R)-isomer form</u>. Sigma-Aldrich currently refers to its marketed version as

'(RS)-mevalonic acid'. However, in one study (Potter and Miziorko 1997) the metabolite is obtained from the supplier Sigma-Aldrich and it is recorded on BRENDA under the general name 'mevalonic acid' without unambiguous chemical identifiers such as SMILES or InChIKeys. The isomer form affects the parameterisation of the metabolite. Hence, this ambiguity creates potential inaccuracy in any resulting modelling.

Problem 2: Reporting of the wrong data

We found cases of incorrect or incomplete kinetic data reported in the primary literature that undermined the ability to model interactions. V_{max} values were regularly reported instead of K_{cat} values where V_{max} is related to K_{cat} by $V_{max} = K_{cat} \times$ (Enzyme concentration). For a V_{max} value to be reusable in subsequent studies, the enzyme concentration must also be reported alongside it. However, we regularly found this not to be the case, making most reported V_{max} values unusable.

Similarly, inhibitors were frequently parameterised by IC_{50} values instead of K_i values where K_i and IC_{50} are related by $K_i = IC_{50} / (1 + S/K_m)$ and S is the substrate concentration. For IC_{50} values to be reusable in future studies, the substrate concentrations must also be reported. Here again we found regular omissions that rendered most reported IC_{50} values unusable.

Solution (1 & 2): Introduce data capture standards that facilitate unambiguous reconstruction of the results without optimization.

Reporting must include clear and thorough descriptions of experimental configurations and unambiguous identification of chemical structures through the use of comprehensive and standard nomenclature. Past experience has shown that effective standards can be developed through open community exercises (e.g. SBML, SBGN). The necessity for appropriate standards has been recognised previously by the chemical biology and pharmacometric communities (Oprea et al. 2011; Swat et al. 2015).

Standards are already employed widely across the life sciences, frequently building upon ontologies (controlled vocabularies of biological/chemical entities and concepts). The International Union of Pure and Applied Chemistry (IUPAC), the International Union of Biochemistry and Molecular Biology (IUBMB) Joint Commission on Biochemical

Nomenclature, and the Nomenclature Committee of IUBMB (NC-IUBMB) have provided guidelines on biochemical descriptions and enzyme classification. A library of ontologies for the life sciences has also been proposed by the OBO Foundry (Smith et al. 2007). Standards and guidelines also exist for reporting biomedical studies, including the minimum information (MI) standards overseen by the MIBBI (Minimum Information for Biological and Biomedical Investigations) Foundry who include the STRENDA Commission (Standards for Reporting Enzymology Data) (Gardossi et al. 2010). The MI standards of direct relevance include the 'Minimum information about a bioactive entity' (MIABE) (Orchard et al. 2011) the 'Minimum information about a proteomic experiment' (MIAPE) (Taylor et al. 2007) and the 'Minimum information about a molecular interaction experiment' (MIMIx) (Orchard et al. 2007).

Problem 3: Curation errors

Online databases can contain errors. We have identified cases where the incorrect structures, enzyme targets, species and parameter values had been recorded. Errors were at low frequency but some would undermine Systems Pharmacology approaches and these fell into two groups: errors that derived from mistakes in the literature itself, such as from mis-interpretation of data, and errors that derived from the incorrect transcribing from the literature to the database. The former derive from verbatim acceptance of results from manuscripts following author error. The latter errors can be introduced by databases themselves, either from semi-automated triage tools or inadvertent curator mistakes and this can be associated with a lack of clarity in the original document. In the present study and for the GtoPdb, we reviewed the primary literature when expanding our datasets and re-curated existing database coverage.

Solution 3: Quality control in curation of results

Using teams of curators to validate each other's work can reduce errors. This can be arranged systematically into error-identifying or error-correcting curation quality control programmes. In an error-identifying programme, each result is independently curated twice and where disagreements are identified, the data is reviewed. Such approaches have been discussed within the International Society for Biocuration (Bateman 2010). However, the funding limitations of most public databases preclude this degree of validation. In an error-correcting programme, each result would be independently curated three times and where a disagreement is found, the consensus is accepted automatically as correct.

Systems Pharmacology for the future

A workflow for future studies and hypothesis generation

With an adequate set of standards and a well-characterised experimental system, it should be possible to develop intervention hypotheses that can be tested to inform future therapy development and to contribute to iterative refinement of databases. To make this a consistent, high confidence process, it would be advantageous to work in one experimental system. Such an experimental system could be *in vivo* or *in vitro*. However, an *in vitro* model would offer more control and consistency. Such an *in vitro* system would serve as a first approximation to *in vivo* physiology and would contribute to determining how *in vitro* parameters are mapped to *in vivo* parameters in order to maximise their value. An advantage of using an *in vitro* system is that it would lend itself to automated hypothesis generation and testing and it could be used to systematically search for new protein-protein and drug-target interactions. It has been suggested that artificial intelligence methods would be suitable for this purpose in the laboratory (King et al. 2004). Automation would both minimise the time required for study and reduce the risk of misreporting or mis-curation of the results.

Our current systems-level understanding has grown to a scale where manual manipulation is no longer feasible. Standards such as SBML, SBGN and SBGN-ML and repositories such as BioModels have been developed partially to address this and automated model development allows the full value of databases to be realized (Swainston et al. 2011). Open PHACTS (Williams et al. 2012) is a consortium responsible for a number of pharmacological and life science databases whose aims include the improvement of data availability through the use of data standards, the incorporation of contextual data through semantic web standards and the cross-platform linkage of datasets through an identity mapping service. Developing multi-drug hypotheses is a challenge that grows exponentially with the number of drugs and interactions considered. HPC resources are likely to be essential for this development.

The following work-flow would enable the process to be automated (see Figure 3).

I) Pharmacological literature seeds databases of pharmacological interactions.
II) Pharmacological and chemical databases containing sufficient information for experimental results to be reproduced accurately. Database APIs facilitate extraction of results for hypothesis generation.

III) Interaction literature seeds databases of biological pathways.

IV) Pathway databases containing sufficient information for experimental results to be reproduced accurately. Database APIs facilitate extraction of results for hypothesis generation.

V) Hypothesis generation for single drug and multi-drug interventions using data obtained through APIs from the pharmacological and pathway databases.

VI) Hypothesis testing. Success yields a candidate therapy and provides validation of the database. Failure initiates further exploration of the underlying interactions that in turn refine the databases.

VII) Candidate Intervention. Following success, the group of compounds enters an optimization pipeline that reduces them to a minimal set of lead compounds for pre-clinical testing to establish their efficacy and safety (see Table 4).

5. Conclusion

The growth in our understanding of pharmacological interactions and the continuing development of our ability to computationally model pathway biology will increasingly enable us to explore drug combinations that target multiple points on multiple pathways to reprogram system level behaviour. In this way, Systems Pharmacology may lead to more effective therapies with fewer side effects. Here, we explored this approach for the mevalonate arm of the cholesterol biosynthesis pathway and in doing so, we identify many of the current barriers to progress.

We attempted to build a systems pharmacology model of the mevalonate arm of the cholesterol biosynthesis pathway but gaps and inconsistencies in the data prevented us from achieving this to a high level of confidence. In particular, we found the lack of comprehensive and systematic parameterisations, experimental variation, ambiguity in structural detail and inappropriate and inaccurate reporting from the primary literature to be obstacles. That this should be the case for a pathway of such high biomedical and commercial significance was unexpected. For this reason our best current parameterisation represents a patchwork of values taken from multiple species and experimental configurations. Nonetheless, by completing gaps in our knowledge with representative values, we were able to demonstrate subtle reprogramming of pathway dynamics that may contribute significantly to drug development. We propose that these obstacles can be removed through the adoption of standards and quality control.

Although we focused on the mevalonate arm of cholesterol biosynthesis, this approach could be applied to any pathway of interest for which targets and ligands are known. However, before this can happen at a general level both the computational biology and the pharmacology communities must collaborate to remove the current barriers to progress.

Conflict of Interest

HB, JS, CM, CS and AH have served as curators for the IUPHAR/BPS GtoPdb.

This paper is dedicated to the memory of Prof. Anthony (Tony) J. Harmar, *Emeritus Professor of Pharmacology,* University of Edinburgh.

Cholesterol
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=2718

Acetyl-CoA
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3038

Squalene
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3054

Geranylgeranyl diphosphate
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3052

Isopentenyl-diphosphate
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3048

Pre-squalene-PP
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3053

Mevalonate portion of the cholesterol biosynthesis pathway
http://www.guidetopharmacology.org/GRAC/FamilyDisplayForward?familyId=104

acetoacetyl-CoA
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3039

HMGCR
http://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=639

MVD
http://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=642

L-659,699
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=5886

HMGCS1
http://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=638

3-hydroxy-3-methyl-6-phosphohexanoic acid
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3202

PMVK
http://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=641

IDI1
http://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=646

IDI2
http://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=647

FDPS
http://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=644

FDFT1
http://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=645

Rosuvastatin
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=2954

Farnesyl Thiodiphosphate
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3216

Cinnamic acid
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3203

6-fluoromevalonate 5-diphosphate
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3205

zoledronic acid
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3177

BPH-628
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3188

Zaragozic acid A
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3057
Fluvastatin
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=2951

Mevalonate
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3042

3-hydroxy-3-methylglutaryl-CoA
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3040

Mevalonate-P
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3046

Mevalonate-PP
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3047

Isopentenyl-PP
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3048

Dimethylallyl-PP
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3049

Geranyl-PP
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3051

Farnesyl-PP
http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=3050

References

Abaan OD, Polley EC, Davis SR, Zhu YJ, Bilke S, Walker RL, Pineda M, Gindin Y, Jiang Y, Reinhold WC, Holbeck SL., 2013. The exomes of the NCI-60 panel: a genomic resource for cancer biology and systems pharmacology. Cancer research 73(14):4372-82.

Aderem A, 2005. Systems biology: its practice and challenges. Cell 121(4): 511-3.

Ahn AC, Tewari M, Poon CS, Phillips RS, 2006. The Clinical Applications of a Systems Approach. PLoS Med 3(7): e209.

Alexander SP, Fabbro D, Kelly E, Marrion N, Peters JA, Benson HE, Faccenda E, Pawson AJ, Sharman JL, Southan C, Davies JA., 2015. The concise guide to PHARMACOLOGY 2015/16: Enzymes. British Journal of Pharmacology 172(24):6024-109.

Balasubramanian N, Brown P, Catt J, Han W, Parker R, Sit S et al., 1989. A potent, tissue-selective, synthetic inhibitor of HMG-CoA reductase. *Journal of medicinal chemistry*, 32(9), pp.2038–41. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2769677 [Accessed September 8, 2016].

Bateman A, 2010. Curators of the world unite: the International Society of Biocuration. *Bioinformatics (Oxford, England)*, 26(8), p.991. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20305270 [Accessed September 8, 2016].

Benson N, Matsuura T, Smirnov S, Demin O, Jones H, Dua P, et al., 2013. Systems pharmacology of the nerve growth factor pathway: use of a systems biology model for the identification of key drug targets using sensitivity analysis and the integration of physiology and pharmacology. *Interface focus*, 3(2), p.20120071. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24427523 [Accessed September 8, 2016].

Bento A, Gaulton A, Hersey A, Bellis L, Chambers J, Davies M, et al., 2014. The ChEMBL bioactivity database: an update. *Nucleic acids research*, 42(Database issue), pp.D1083-90. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24214965 [Accessed September 8, 2016].

Berger S, Iyengar, R, 2009. Network analyses in systems pharmacology. *Bioinformatics (Oxford, England)*, 25(19), pp.2466–72. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19648136 [Accessed September 8, 2016].

Bhattacharya B, Sweby P, Minihane A, Jackson K, Tindall M, 2014. A mathematical model of the sterol regulatory element binding protein 2 cholesterol biosynthesis pathway.

Journal of theoretical biology, 349, pp.150–62. Available at:
http://www.ncbi.nlm.nih.gov/pubmed/24444765 [Accessed September 8, 2016].

Blanc M, Hsieh WY, Robertson KA, Watterson S, Shui G, Lacaze P, et al., 2011. Host
defense against viral infection involves interferon mediated down-regulation of sterol
biosynthesis. *PLoS Biology*, 9(3), p.e1000598.

Boralli V, Coelho E, Sampaio S, Marques M, Lanchote V, 2009. Enantioselectivity in the
pharmacokinetic interaction between fluvastatin and lercanidipine in healthy volunteers.
Journal of clinical pharmacology, 49(2), pp.205–11. Available at:
http://www.ncbi.nlm.nih.gov/pubmed/19033449 [Accessed September 8, 2016].

Boran ADW, Iyengar R, 2010a. Systems approaches to polypharmacology and drug
discovery. *Current opinion in drug discovery & development*, 13(3), pp.297–309.
Available at: http://www.ncbi.nlm.nih.gov/pubmed/20443163 [Accessed September 8,
2016].

Boran ADW, Iyengar R, 2010b. Systems pharmacology. *Mount Sinai Journal of Medicine*,
77(4), pp.333–344. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20687178
[Accessed September 8, 2016].

Bowes J, Brown AJ, Hamon J, Jarolimek W, Sridhar A, Waldron G, et al., 2012. Reducing
safety-related drug attrition: the use of in vitro pharmacological profiling. *Nature
reviews. Drug discovery*, 11(12), pp.909–22. Available at:
http://www.ncbi.nlm.nih.gov/pubmed/23197038 [Accessed September 8, 2016].

Caspi R, Billington R, Ferrer L, Foerster H, Fulcher CA, Keseler IM, et al., 2016. The
MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of
pathway/genome databases. *Nucleic acids research*, 44(D1), pp.D471-80. Available at:
http://www.ncbi.nlm.nih.gov/pubmed/26527732 [Accessed September 8, 2016].

Chang A, Schomburg I, Placzek S, Jeske L, Ulbrich M, Xiao M, et al, 2015. BRENDA in
2015: exciting developments in its 25th year of existence. *Nucleic acids research*,
43(Database issue), pp.D439-46. Available at:
http://www.ncbi.nlm.nih.gov/pubmed/25378310 [Accessed September 8, 2016].

Chelliah V, Laibe C, Le Novère N, 2013. BioModels Database: a repository of mathematical
models of biological processes. *Methods in molecular biology (Clifton, N.J.)*, 1021,
pp.189–99. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23715986 [Accessed
September 8, 2016].

Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al., 2014. The Reactome
pathway knowledgebase. *Nucleic acids research*, 42(Database issue), pp.D472-7.

Available at: http://www.ncbi.nlm.nih.gov/pubmed/24243840 [Accessed September 8, 2016].

Czauderna T, Klukas C, Schreiber F, 2010. Editing, validating and translating of SBGN maps. *Bioinformatics (Oxford, England)*, 26(18), pp.2340–1. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20628075 [Accessed September 8, 2016].

Darwich AS, Ogungbenro K, Vinks AA, Powell JR, Reny JL, Marsousi N, Daali Y, Fairman D, Cook J, Lesko LJ, McCune JS., 2017. Why has model-informed precision dosing not yet become common clinical reality? Lessons from the past and a roadmap for the future. Clin Pharma & Ther [in press].

Di Pietro G, Coelho EB, Geleilete TM, Marques MP, Lanchote VL, 2006. Chiral evaluation of fluvastatin in human plasma by high-performance liquid chromatography electrospray mass spectrometry. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences*, 832(2), pp.256–61. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16480934 [Accessed September 8, 2016].

Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, et al., 2016. The Reactome pathway Knowledgebase. Nucleic Acids Res. 44(D1):D481-7.

Gardossi L, Poulsen PB, Ballesteros A, Hult K, Svedas VK, Vasić-Racki D. et al., 2010. Guidelines for reporting of biocatalytic reactions. *Trends in biotechnology*, 28(4), pp.171–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20149467 [Accessed September 8, 2016].

Gaulton A, Hersey A, Nowotka M, Bento AP, Chambers J, Mendez D, Mutowo P, Atkinson F, Bellis LJ, Cibrián-Uhalte E, Davies M, 2016. The ChEMBL database in 2017. Nucleic acids research. 45(D1):D945-54.

Henderson R, O'Kane M, McGilligan V, Watterson S, 2016. The genetics and screening of familial hypercholesterolaemia. *Journal of biomedical science*, 23, p.39. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27084339 [Accessed September 8, 2016].

Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, et al., 2003. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics (Oxford, England)*, 19(4), pp.524–31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12611808 [Accessed September 8, 2016].

Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M, *Data, information, knowledge and principle: back to metabolism in KEGG.* Nucleic Acids Research, 2014.

42(D1): p. D199-D205.

Kell DB, Goodacre R, 2014. Metabolomics and systems pharmacology: why and how to model the human metabolic network for drug discovery. *Drug discovery today*, 19(2), pp.171–82. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23892182 [Accessed September 8, 2016].

Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al., 2016. PubChem Substance and Compound databases. *Nucleic acids research*, 44(D1), pp.D1202-13. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26400175 [Accessed September 8, 2016].

King R, Whelan K, Jones F, Reiser P, Bryant C, Muggleton S, et al., 2004. Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature*, 427(6971), pp.247–52. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14724639 [Accessed September 8, 2016].

Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K, 2017, KEGG: new perspectives on genomes, pathways, diseases and drugs, Nucleic Acids Res. 45(D1):D353-D361.

Kutmon M, Riutta A, Nunes N, Hanspers K, Willighagen EL, Bohler A, et al., 2016. WikiPathways: capturing the full diversity of pathway knowledge. *Nucleic acids research*, 44(D1), D488-494.

Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, Maciejewski A, Arndt D, Wilson M, Neveu V, Tang A, 2013. DrugBank 4.0: shedding new light on drug metabolism. Nucleic acids research. 42(D1):D1091-7.

Lehár J, Zimmermann GR, Krueger AS, Molnar RA, Ledell JT, Heilbut AM, Short GF, Giusti LC, Nolan GP, Magid OA, Lee MS., 2007. Chemical combination effects predict connectivity in biological systems. Molecular systems biology 3(1) 80.

Le Novère N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, et al., 2009. The Systems Biology Graphical Notation. *Nature biotechnology*, 27(8), p.735.

Lewington S, Whitlock G, Clarke R, Sherliker P, Emberson J, Halsey J, et al., 2007. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *Lancet*, 370(9602), pp.1829–39. Available at: http://dx.doi.org/10.1016/S0140-6736(07)61778-4 [Accessed April 7, 2012].

Lewis JE, Glass L, Steady states, limit cycles and chaos in models of complex biological networks, Int. J. Bifurcation Chaos 01, 477 (1991).

Lu H, Talbot S, Robertson K, Watterson S, Forster T, Roy D, et al., 2015. Rapid proteasomal

elimination of 3-hydroxy-3-methylglutaryl-CoA reductase by interferon-γ in primary macrophages requires endogenous 25-hydroxycholesterol synthesis. *Steroids*, 99(Pt B), pp.219–29. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25759117 [Accessed September 8, 2016].

Madrasi K, Burns RN, Hendrix CW, Fossler MJ, Chaturvedula A., 2014. Linking the population pharmacokinetics of tenofovir and its metabolites with its cellular uptake and metabolism. CPT: pharmacometrics & systems pharmacology. 3(11): 1-6

Mazein A,Watterson S, Hsieh W, Griffiths W, Ghazal P, 2013. A comprehensive machine-readable view of the mammalian cholesterol biosynthesis pathway. *Biochemical pharmacology*, 86(1), pp.56–66. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23583456 [Accessed September 8, 2016].

Mazein A, Watterson S, Gibbs HC, Hsieh W, Abdullah MB, Blanc M, Griffiths WJ, Freeman TC, Ghazal P, 2011. Regulation and feedback of cholesterol metabolism. Nature Precedings, 10.1038/npre.2011.6315.1

Oprea T, May E, Leitão A, Tropsha A, 2011. Computational systems chemical biology. *Methods in molecular biology (Clifton, N.J.)*, 672, pp.459–88. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20838980 [Accessed September 8, 2016].

Orchard S, Salwinski L, Kerrien S, Montecchi-Palazzi L, Oesterheld M, Stümpflen V, et al., 2007. The minimum information required for reporting a molecular interaction experiment (MIMIx). *Nature biotechnology*, 25(8), pp.894–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17687370 [Accessed September 8, 2016].

Orchard S, Al-Lazikani B, Bryant S, Clark D, Calder E, Dix I, et al., 2011. Minimum information about a bioactive entity (MIABE). *Nature reviews. Drug discovery*, 10(9), pp.661–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21878981 [Accessed September 8, 2016].

Parton A, McGilligan V, O'Kane M, Baldrick F, Watterson S, 2016. Computational modelling of atherosclerosis. *Briefings in bioinformatics*, 17(4), pp.562–75. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26438419 [Accessed September 8, 2016].

Pawson A, Sharman J, Benson H, Faccenda E, Alexander S, Buneman O, et al., 2014. The IUPHAR/BPS Guide to PHARMACOLOGY: an expert-driven knowledgebase of drug targets and their ligands. *Nucleic acids research*, 42(Database issue), pp.D1098-106. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24234439 [Accessed September 8,

2016].

Pence H, Williams,A, 2010. ChemSpider: An Online Chemical Information Resource. *Journal of Chemical Education*, 87(11), pp.1123–1124. Available at: http://pubs.acs.org/doi/abs/10.1021/ed100697w [Accessed September 8, 2016].

Peterson MC, Riggs MM., 2015. FDA advisory meeting clinical pharmacology review utilizes a quantitative systems pharmacology (QSP) model: a watershed moment? CPT: Pharmacometrics & Systems Pharmacology 4(3):189-92.

Petrelli A, Giordano S, 2008. From single- to multi-target drugs in cancer therapy: when aspecificity becomes an advantage. *Current medicinal chemistry*, 15(5), pp.422–32. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18288997 [Accessed September 8, 2016].

Pichardo–Almarza C, Metcalf L, Finkelstein A, 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 & Systems Pharmacology 4(1): 41-50

Potter D, Miziorko H, *Identification of catalytic residues in human mevalonate kinase.* Journal of Biological Chemistry, 1997. **272**(41): p. 25449-25454.

Procopiou P, Bailey E, Bamford M, Craven A, Dymock B, Houston J, et al., 1994. The squalestatins: novel inhibitors of squalene synthase. Enzyme inhibitory activities and in vivo evaluation of C1-modified analogues. *Journal of medicinal chemistry*, 37(20), pp.3274–81. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7932554 [Accessed September 8, 2016].

Qin C, Zhang C, Zhu F, Xu F, Chen S, Zhang P, et al., 2014. Therapeutic target database update 2014: a resource for targeted therapeutics. *Nucleic acids research*, 42(Database issue), pp.D1118-23. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24265219 [Accessed September 8, 2016].

Robertson K, Hsieh W, Forster T, Blanc M, Lu H, Crick P, et al., 2016. An Interferon Regulated MicroRNA Provides Broad Cell-Intrinsic Antiviral Immunity through Multihit Host-Directed Targeting of the Sterol Pathway. *PLoS biology*, 14(3), p.e1002364. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26938778 [Accessed September 8, 2016].

Rostami-Hodjegan A., 2012. Physiologically based pharmacokinetics joined with in vitro–in vivo extrapolation of ADME: a marriage under the arch of systems

pharmacology. Clinical Pharmacology & Therapeutics 92(1): 50-61

Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, et al., 2007. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature biotechnology*, 25(11), pp.1251–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17989687 [Accessed September 8, 2016].

Sorger P, Allerheiligen S, Altmann R, Shoichet B, Lauffenburger D, Iyengar R, et al., 2011. Quantitative and Systems Pharmacology in the Post-genomic Era: New Approaches to Discovering Drugs and Understanding Therapeutic Mechanisms. *in QSP Workshop Group (R Ward Ed)*.

Southan C, Sharman J, Benson H, Faccenda E, Pawson A, Alexander S, et al., 2016. The IUPHAR/BPS Guide to PHARMACOLOGY in 2016: towards curated quantitative interactions between 1300 protein targets and 6000 ligands. *Nucleic acids research*, 44(D1), pp.D1054-68. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26464438 [Accessed September 8, 2016].

Swainston N, Smallbone K, Mendes P, Kell D, Paton N, 2011. The SuBliMinaL Toolbox: automating steps in the reconstruction of metabolic networks. *Journal of integrative bioinformatics*, 8(2), p.186. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22095399 [Accessed September 8, 2016].

Swat M, Moodie S, Wimalaratne S, Kristensen N, Lavielle M, Mari A, et al., 2015. Pharmacometrics Markup Language (PharmML): Opening New Perspectives for Model Exchange in Drug Development. *CPT: pharmacometrics & systems pharmacology*, 4(6), pp.316–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26225259 [Accessed September 8, 2016].

Taylor C, Paton N, Lilley K, Binz P, Julian R, Jones A, et al., 2007. The minimum information about a proteomics experiment (MIAPE). *Nature biotechnology*, 25(8), pp.887–93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17687369 [Accessed September 8, 2016].

Thakur D, Marchand B, 2012. Hybrid optimal control for HIV multi-drug therapies: a finite set control transcription approach. *Mathematical biosciences and engineering : MBE*, 9(4), pp.899–914. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23311427 [Accessed September 8, 2016].

Trame MN, Biliouris K, Lesko LJ, Mettetal JT., 2016. Systems pharmacology to predict drug safety in drug development. European Journal of Pharmaceutical Sciences

94:93-5.

US Food and Drug Adminstration, 2016a. Novel Drug Approvals for 2016. Available at: http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/ucm483775.htm [Accessed September 11, 2016].

US Food and Drug Administration, 2016b. Novel New Drugs 2015 Summary. Available at: https://www.fda.gov/drugs/developmentapprovalprocess/druginnovation/ucm430302.htm

van der Graaf PH, Benson N., 2011. Systems pharmacology: bridging systems biology and pharmacokinetics-pharmacodynamics (PKPD) in drug discovery and development. Pharmaceutical research 28(7): 1460-4.

van der Greef J, McBurney R, 2005. Innovation: Rescuing drug discovery: in vivo systems pathology and systems pharmacology. *Nature reviews. Drug discovery*, 4(12), pp.961–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16341061 [Accessed September 8, 2016].

van Hasselt J, van der Graaf P, 2015. Towards integrative systems pharmacology models in oncology drug development. *Drug discovery today. Technologies*, 15, pp.1–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26464083 [Accessed September 8, 2016].

Van Iersel M, Villéger A, Czauderna T, Boyd S, Bergmann F, Luna A, et al., 2012. Software support for SBGN maps: SBGN-ML and LibSBGN. *Bioinformatics (Oxford, England)*, pp.1–6.

Vicini P, van der Graaf P, 2013. Systems pharmacology for drug discovery and development: paradigm shift or flash in the pan? *Clinical pharmacology and therapeutics*, 93(5), pp.379–81. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23598453 [Accessed September 8, 2016].

Visser SA, Alwis DP, Kerbusch T, Stone JA, Allerheiligen SR., 2014. Implementation of quantitative and systems pharmacology in large pharma. CPT: Pharmacometrics & Systems Pharmacology 3(10) e142

Wagner B, Gilbert T, Hanai J, Imamura S, Bodycombe N, Bon R, et al., 2011. A small-molecule screening strategy to identify suppressors of statin myopathy. *ACS chemical biology*, 6(9), pp.900–4.

Wajima T, Isbister GK, Duffull SB., 2009. A comprehensive model for the humoral coagulation network in humans. Clinical pharmacology and therapeutics. 6(3): 290.

Watterson S, Guerriero M, Blanc M, Mazein A, Loewe L, Robertson K, 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(3), pp.613–21.

Watterson S, Ghazal P, 2010. Use of logic theory in understanding regulatory pathway signaling in response to infection, Future Microbiol 5(2): 163–176

Westerhoff H, Nakayama S, Mondeel T, Barberis M, 2015. Systems Pharmacology: An opinion on how to turn the impossible into grand challenges. *Drug discovery today. Technologies*, 15, pp.23–31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26464087 [Accessed September 8, 2016].

Williams A, Harland L, Groth P, Pettifer S, Chichester C, Willighagen E, et al., 2012. Open PHACTS: semantic interoperability for drug discovery. *Drug discovery today*, 17(21–22), pp.1188–98. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22683805 [Accessed September 8, 2016].

Zhao S, Nishimura T, Chen Y, Azeloglu E, Gottesman O, Giannarelli C, et al., 2013. Systems pharmacology of adverse event mitigation by drug combinations. *Science translational medicine*, 5(206), p.206ra140. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24107779 [Accessed September 8, 2016].
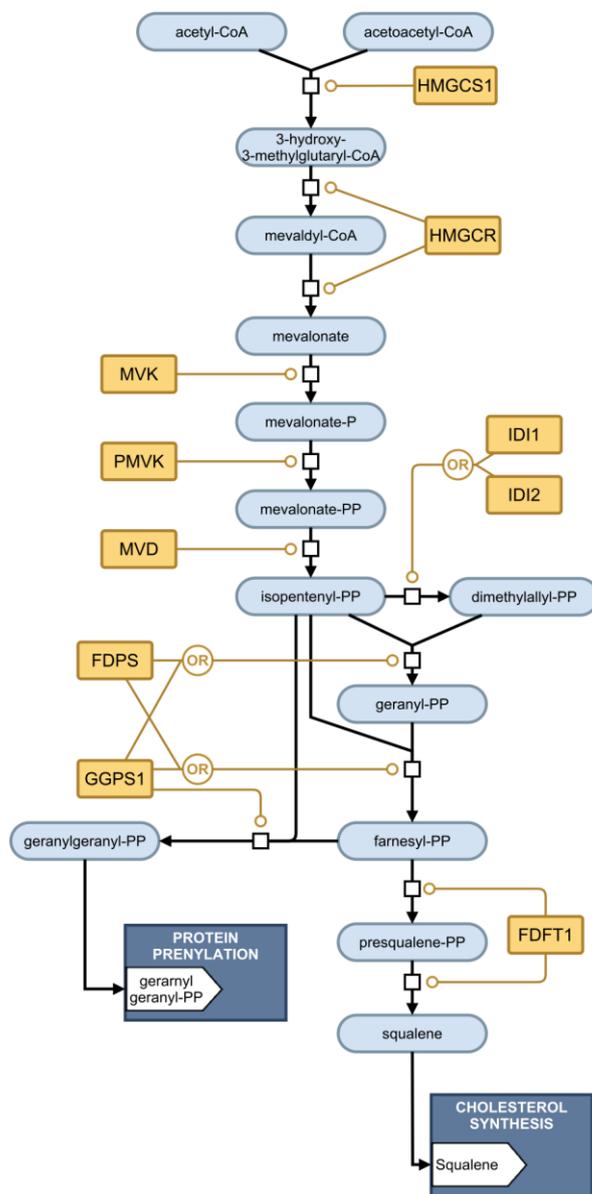
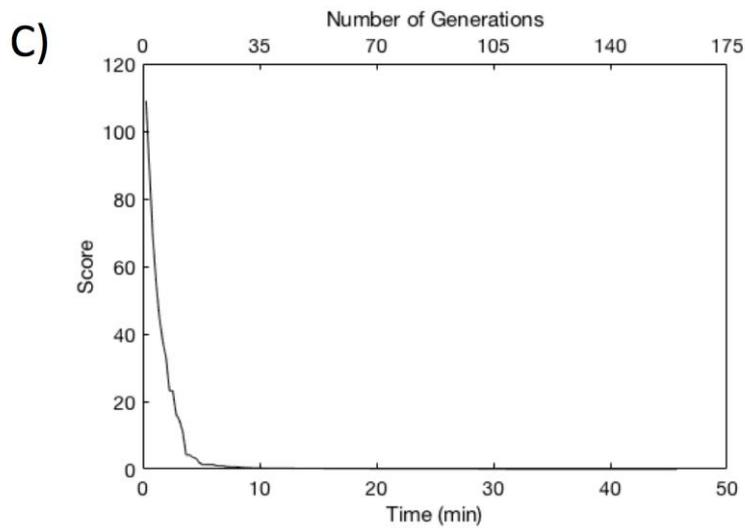Figure 1. The mevalonate arm of the cholesterol biosynthesis pathway.
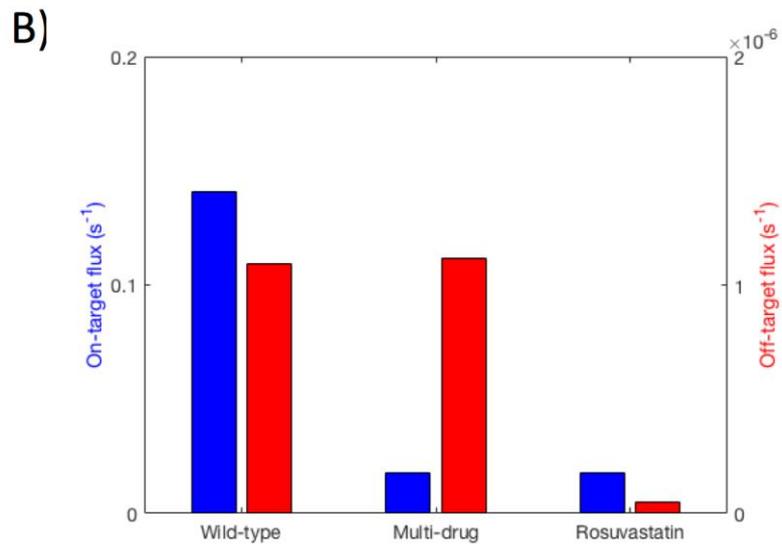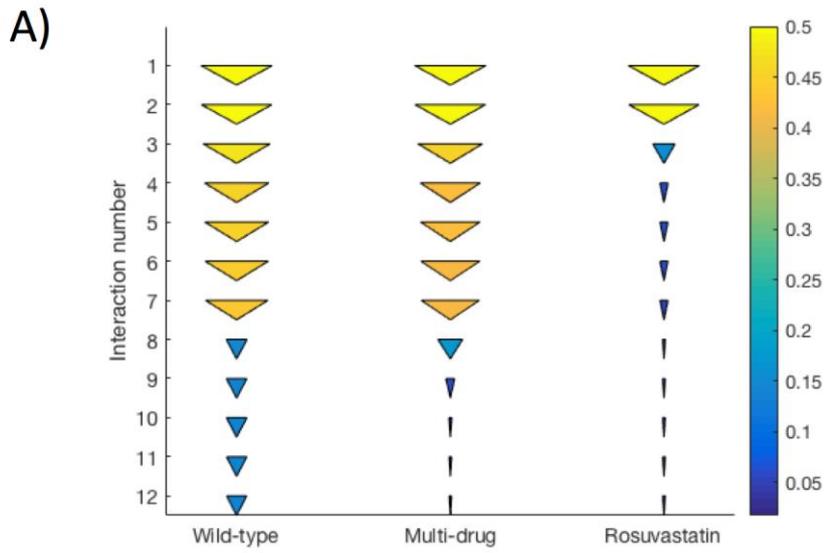
Figure 2. A) The profile of flux through the pathway shown in Figure 1 described as a cone-plot for the three scenarios: wild-type (treatment free), optimised multi-drug intervention and single-drug statin-like intervention. Cone size and colour both describe flux level. We show only the flux leading to cholesterol synthesis (the flux to protein prenylation is described in Fig 2B). Interactions are numbered by their product: (1: 3-hydroxy-3-methylglutaryl-CoA; 2: melvaldyl-CoA, 3: mevalonate, 4: mevalonate-P, 5: mevalonate-PP, 6: isopentenyl-PP, 7: dimethylallyl-PP, 8: geranyl-PP, 9: farnesyl-PP, 10: presqualene-PP, 11: Squalene, 12: Cholesterol synthesis). B) The flux through the endpoints of the two branches for the three scenarios: wild-type, optimised multi-drug intervention and single drug statin-like intervention. Flux through the squalene/cholesterol synthesis branch is shown in blue. Flux through the geranylgeranyl-pp/protein prenylation branch is shown in red. The statin concentration has been selected to ensure that the flux through the cholesterol synthesis branch is the same as in the multi-drug intervention. C) Convergence on the optimal multi-drug intervention that suppresses cholesterol synthesis whilst minimising off target effects, shown against time and against generations of the genetic algorithm.

Figure 3. The proposed Systems Pharmacology Workflow.

Table 1. A list of the enzymes of the lanosterol branch of the cholesterol synthesis pathway, with gene and protein identifiers and EC numbers.

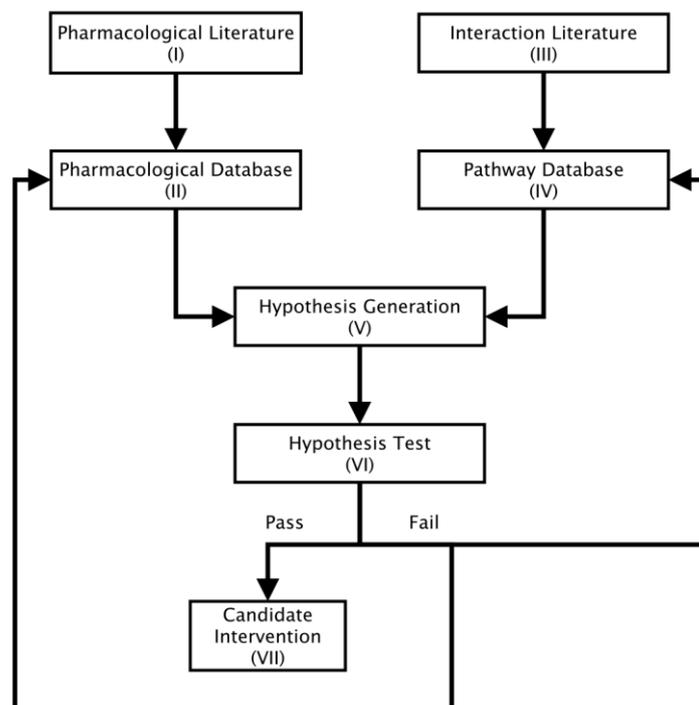| E.C number	Enzyme/GtoPdb target ID	UniProt ID	HGNC ID	IUBMB Enzyme approved name	Reaction Catalysed
2.3.3.10	HMGCS1/638	Q01581	5007	Hydroxymethylglutaryl-CoA synthase	acetyl CoA + H2O + acetoacetyl CoA = (S)-3-hydroxy-3-methylglutaryl-CoA + coenzyme A
1.1.1.34	HMGCR/639	P04035	5006	Hydroxymethylglutaryl-CoA reductase (NAPDH)	(S)-3-hydroxy-3-methylglutaryl-CoA + 2 NADPH = mevaldyl CoA + 2NADP+ mevaldyl CoA + 2NADP+ = (R)-mevalonate + coenzyme A + 2 NADP+
2.7.1.36	MVK/640	Q03426	7530	Mevalonate kinase	ATP + (R)-mevalonate =ADP + (R)-5-phosphomevalonate
2.7.4.2	PMVK/641	Q15126	9141	Phosphomevalonate kinase	ATP + (R)-5-phosphomevalonate =ADP + (R)-5-diphosphomevalonate
4.1.1.33	MVD/642	P53602	7529	Diphosphomevalonate decarboxylase	ATP + (R)-5-diphosphomevalonate = ADP + phosphate + isopentenyl diphosphate + CO_2
5.3.3.2	IDI1 and IDI2*/646 & 647	Q13907/Q9BXS1	5387/ 23487	Isopentenyl-diphosphate delta isomerase	Isopentenyl diphosphate =dimethylallyl diphosphate
2.5.1.1 2.5.1.10	FDPS/644	P14324	3631	Farnesyl diphosphate synthase	dimethylallyl diphosphate + isopentenyl diphosphate = diphosphate + geranyl diphosphate geranyl diphosphate + isopentenyl diphosphate = diphosphate + trans,trans-farnesyl diphosphate
2.5.1.1 2.5.1.10 2.5.1.29	GGPS1/643	O95749	4249	Farnesyltranstransferase	dimethylallyl diphosphate + isopentenyl diphosphate = diphosphate +geranyl diphosphate geranyl diphosphate + isopentenyl diphosphate = diphosphate + trans,trans-farnesyl diphosphate trans,trans-farnesyl diphosphate + isopentenyl diphosphate = diphosphate + geranylgeranyl diphosphate

2.5.1.21	FDFT1/645	P37268	3629	Farnesyl-diphosphate farnesyl transferase 1	2 trans,trans-farnesyl diphosphate = diphosphate + presqualene diphosphate
					presqualene diphosphate + NAD(P)H + H+ = trans-squalene + diphosphate + NAD(P)+

Footnote   Reported substrates, kinetic values and details of the experimental studies from which they were obtained, along with references are recorded.

Please note that ligands outlined in the table are listed using the nomenclature from the original literature. Where the reference did not specify the isomer used experimentally, it was assumed the racemate was used.

Table 1. A list of the enzymes of the lanosterol branch of the cholesterol synthesis pathway, with gene and protein identifiers and EC numbers.

E.C number	Km(mM)/pmid	Reported Substrate/GtoPdb Ligand ID	Kcat (s-1)/pmid	Organism	Reported Conditions	Mean Km (mM)	Substituted mean Km	Substituted mean Kcat
2.3.3.10	0.009/6118268	Acetyl-CoA/3038	--	Rattus norvegicus	absence of acetoacetyl-CoA, hydrolysis reaction	0.0785		6.651575
	0.2/6118268	Acetyl-CoA/3038	--	Rattus norvegicus	0.01M acetoacetyl-CoA			
	0.073/19706283	Acetyl-CoA/3038	--	Homo sapiens	--			
	0.076/19706283	Acetyl-CoA/3038	--	Homo sapiens	--			
	0.084/19706283	Acetyl-CoA/3038	--	Homo sapiens	--			
	0.029/7913309	Acetyl-CoA/3038	--	Homo sapiens	--			
1.1.1.34	0.006/4985697	3-hydroxy-3-methylglutaryl-CoA/3040	--	Rattus norvegicus	Only one enantiomer	0.0765		0.0023
	0.012/4985697	3-hydroxy-3-methylglutaryl-CoA/3040	--	Rattus norvegicus	--			
	0.01/10392455	3-hydroxy-3-methylglutaryl-CoA/3040	--	Mus musculus	Enzyme from tumour			
	0.014/10392455	3-hydroxy-3-methylglutaryl-CoA/3040	--	Mus musculus	Enzyme from liver and tumour			
	0.015/10392455	3-hydroxy-3-methylglutaryl-CoA/3040	--	Mus musculus	Enzyme from liver, implanted tumour			
	0.019/10392455	3-hydroxy-3-methylglutaryl-CoA/3040	--	Mus musculus	Enzyme from liver, implanted tumour			
	0.024/10392455	3-hydroxy-3-methylglutaryl-CoA/3040	--	Mus musculus	Enzyme from liver			
	0.07/16128575	3-hydroxy-3-methylglutaryl-CoA/3040	--	Homo sapiens	--			
	0.6/--	3-hydroxy-3-methylglutaryl-CoA/3040	--	Homo sapiens	pH 7.5/Temp not specified			
	0.068/18446881	hydroxymethylglutaryl-CoA	0.023/18446881	Rattus norvegicus	--			
	0.004/7077140	S-3-hydroxy-3-methylglutaryl-CoA/3040		Rattus norvegicus	--			
2.7.1.36	0.035/14680942	(RS)-mevalonate/3056	--	Rattus norvegicus	pH 7.5/25C	0.0337		
	0.035/17964869	(RS)-mevalonate/3056	21.9/18302342	Rattus norvegicus	pH 7.5/34C			
	0.0408/18302342	(RS)-mevalonate/3056	--	Homo sapiens	pH 7.5/30C			
	0.024/9325256	mevalonate/3056	--	Homo sapiens	pH 7.0/30C			
2.7.4.2	0.034/17902708	(R)-5-phosphomevalonate/3046	--	Homo sapiens	pH7.0/30C	0.034		6.651575
4.1.1.33	0.02/8744421	5-diphosphomevalonate/3055	--	Rattus norvegicus	--	0.0262		
	0.0289/18823933	5-diphosphomevalonate/3055	4.5/18823933	Homo sapiens	30C			
	0.036/16626865	5-diphosphomevalonate/3055	--	Rattus norvegicus	--			
	0.036/17888661	5-diphosphomevalonate/3055	--	Rattus norvegicus	--			
	0.01/11913522	mevalonate diphosphate/3055	--	Mus musculus	pH 7.2			
5.3.3.2	0.0228/17202134	isopentenyl diphosphate/3048	--	Homo sapiens	pH 8.0	0.0279		6.651575
	0.033/8806705	isopentenyl diphosphate/3048	--	Homo sapiens				
2.5.1.1	--						0.0351375	6.651575
2.5.1.10	--						0.0351375	6.651575
2.5.1.1	--	--	--	--			0.0351375	6.651575

								0.0351375	6.651575
2.5.1.10	--	--	--	--					
2.5.1.29	0.0029/17846065	isopentenyl diphosphate/3048	--	Rattus norvegicus	pH 7.0/37C		0.0027		
	0.003/16698791	isopentenyl diphosphate/3048	--	Homo sapiens	pH 7.7/37C				
	0.00071/17846065	trans,trans-farnesyl diphosphate/3050	--	Rattus norvegicus	pH 7.0/37C				
	0.0042/16698791	trans,trans-farnesyl diphosphate/3050	0.204/16698791	Homo sapiens	pH 7.7/37C				
2.5.1.21	0.0023/9473303	farnesyl diphosphate/2910	--	Homo sapiens			0.0016		6.651575
	0.001/1569107	trans-farnesyl diphosphate/3050	--	Rattus norvegicus					

Footnote   Reported substrates, kinetic values and details of the experimental studies from which they were obtained, along with references are recorded.

Please note that ligands outlined in the table are listed using the nomenclature from the original literature.

Where the reference did not specify the isomer used experimentally, it was assumed the racemate was used.

Table 2. Normalised enzyme levels

Enzyme	level
HMGCS1	1441
HMGCR	258
MVK	76
PMVK	874
MVD	111
IDI1	2707
IDI2	--
FDPS	7029
GGPS1	86
FDFT1	3425

Table 3. List of inhibitors for each of the enzymes in the lanosterol branch of the cholesterol synthesis pathway, with Ki values and references.

E.C number	Enzyme	Inhibitor name / GtoPdb Ligand ID	InChi Key	Approved drug?
2.3.3.10	HMGCS1	L-659,699/5886	ODCZJZWSXPVLAW-KXCGKLMDSA-N	No
1.1.1.34	HMGCR	Rosuvastatin/2954	BPRHUIZQVSMCRT-YXWZHEERSA-N	Yes
		Rosuvastatin/2954	BPRHUIZQVSMCRT-YXWZHEERSA-N	Yes
		Rosuvastatin/2954	BPRHUIZQVSMCRT-YXWZHEERSA-N	Yes
		Cerivastatin/2950	SEERZIQQUAZTOL-ANMDKAQQSA-N	Yes
		Cerivastatin/2950	SEERZIQQUAZTOL-ANMDKAQQSA-N	Yes
		Atorvastatin/2949	XUKUURHRXDUEBC-KAYWLYCHSA-N	Yes
		Atorvastatin/2949	XUKUURHRXDUEBC-KAYWLYCHSA-N	Yes
		Lovastatin/2739	PCZOHLXUXFIOCF-BXMDZJJMSA-N	Yes
		Lovastatin/2739	PCZOHLXUXFIOCF-BXMDZJJMSA-N	Yes
		Simvastatin/2955	RYMZZMVNJRMUDD-HGQWONQESA-N	Yes
		CHEMBL39312/7991	VWKZOIOUHUHQKZ-HZPDHXFCSA-N	No
		CHEMBL39102/7993	XKZCNQAYFRBCKR-HNNXBMFYSA-N	No
		Fluvastatin/2951	FJLGEFLZQAZZCD-MCBHFWOFSA-N	Yes
2.7.1.36	MVK	Farnesyl thiodiphosphate/3216	DRADWUUFBCYMDM-UHFFFAOYSA-L	No
2.7.4.2	PMVK	Cinnamic acid/3203	WBYWAXJHAXSJNI-VOTSOKGWSA-N	No
		Isoferulic acid**		No
		3-hydroxy-3-methyl-6-phosphohexanoic acid/3202	XRCIRZGXKWCWNQ-UHFFFAOYSA-N	No
		p-coumaric acid/5787	NGSWKAQJJWESNS-ZZXKWVIFSA-N	No
4.1.1.33	MVD	6-fluoromevalonate 5-diphosphate/3205	GLNCOGHKIHKSA-UHFFFAOYSA-N	No
		2-fluoromevalonate 5-diphosphate/3204	WPXHWHACORBSDS-UHFFFAOYSA-N	No
		diphosphoglycolyl proline/3206	CDFDGXYBANXCPC-UHFFFAOYSA-N	No
		CHEMBL1160330/7994	YERUUUBBRAPJND-UHFFFAOYSA-N	No
		CHEMBL1160328/7996	YIGLDWRZXXHIGZ-ZCFIWIBFSA-N	No
		P'-geranyl 2-fluoromevalonate 5-diphosphate/3207	ACYPMTKDKJZHBJ-MDWZMJQESA-N	No
		P'-geranyl 3,5,9-trihydroxy-3-methylnonanate 9-diphosphate/5621	PMUQIJKCGIYWGT-GZTJUZNOSA-N	No
5.3.3.2	IDI1	--		
5.3.3.2	IDI2	--		
2.5.1.1, 2.5.1.10	FDPS	Zoledronic acid/3177	XRASPMIURGNCCH-UHFFFAOYSA-N	Yes
		Zoledronic acid/3177	VWKZOIOUHUHQKZ-HZPDHXFCSA-N	Yes
		Risedronate/3176	IIDJRNMFWXDHID-UHFFFAOYSA-N	Yes
		Risedronate/3176	IIDJRNMFWXDHID-UHFFFAOYSA-N	Yes
		NE58062/3166	XUCBNFJYKWKAMN-UHFFFAOYSA-N	No
		NE97220/3171	NAIJOBGUXRHQJW-UHFFFAOYSA-N	No
		NE97220/3171	NAIJOBGUXRHQJW-UHFFFAOYSA-N	No
		NE58018/3168	XXNASZAYANFLID-UHFFFAOYSA-N	No
		NE58018/3168	XXNASZAYANFLID-UHFFFAOYSA-N	No
2.5.1.1, 2.5.1.10, 2.5.1.29	GGPS1	BPH-628/3188	MPBUFKZCEBTBSK-UHFFFAOYSA-N	No
		BPH-608/7977	YXQQNSYZOQHKHD-UHFFFAOYSA-N	No
		BPH-675/7975	MZVWVRVNMXTDAK-UHFFFAOYSA-N	No
		BPH-629/7976	BYVXAUZOTGITQZ-UHFFFAOYSA-N	No
		BPH-676/7978	NWIARQRYIRVYCM-UHFFFAOYSA-N	No
2.5.1.21	FDFT1	Zaragozic acid A/3057	DFKDOZMCHOGOBR-NCSQYGPNSA-N	No
		CHEMBL24362/3105	FBPJEWKDFUWVKV-UHFFFAOYSA-N	No
		CHEMBL1208103/3120	HGDWHTASNMRJMP-UHFFFAOYSA-N	No
		CHEMBL1207858/3127	AGJZDRXKAQZWEP-UHFFFAOYSA-N	No
		BPH-830/3121	GNETVUVZFYJATO-UHFFFAOYSA-N	No
		SQ-109/7997	JFIBVDBTCDTBRH-REZTVBANSA-N	No
		[1-(hydroxycarbamoyl)-4-(3-phenoxyphenyl)butyl]phosphonate/3120	HGDWHTASNMRJMP-UHFFFAOYSA-N	No
		compound 13 [PMID: 19456099]/3127	AGJZDRXKAQZWEP-UHFFFAOYSA-N	No
		(1-methyl-1-{[3-(3-phenoxyphenyl)propyl]carbamoyl}ethyl)phosphonate/3127	AGJZDRXKAQZWEP-UHFFFAOYSA-N	No

**Interaction not listed on GtoPdb. These reactions were selected from either BRENDA or ChEMBL to complete the dataset required for the modelling process.