The developing infant gut microbiota: mathematical predictions of the effects of oligosaccharides

Versluis, D.M.

Chapter 1

Introduction

The human gut starts entirely or nearly entirely free of bacteria [1, 2], but is quickly colonised by many bacterial species [3, 4]. In a matter of days the bacterial population grows to around 10^{10} bacterial cells per gram feces [5, 6]. Together, these bacteria form a dynamic community in the gut. This community is known as the infant gut microbiota. Influenced by factors such as nutrition and the environment, each infant develops a unique microbiota. These microbiotas are very different from those typically found in adults, and deserve to be studied on their own. The infant gut microbiota has been recognised as important for infant health since the 19th century [7], but until recently it was difficult to identify and quantify the major bacterial species involved [8]. With metagenomics it is now possible to identify hundreds of different species in infant feces [3, 4, 9], and to analyse their metabolic capacity [10, 11], but it is still unclear why the infant gut microbiota reaches the different compositions that it does, and how this process is influenced by nutrition [12]. To develop new hypotheses, we develop mathematical models in this thesis. For these models, we make use of genome-scale metabolic models (GEMs) [10, 11]. GEMs are reconstructions of bacterial metabolic networks, based on bacterial genetics and enzyme functions [13]. Each GEM consist of a list of reactions that a particular organism can perform [13]. Each reaction has a set of input and output molecules. Together, these reactions and molecules form a network model. Constraint-based modelling can generate predictions for the flow of molecules through this network [13]. A GEM can be used to create a stoichiometric matrix that includes the stoichiometry of each reaction. Constraint-based modelling then applies additional constraints to these

stoichiometric matrices [13]. These constraints can be based on our knowledge of the bacterial environment, such as by constraining some reactions based on the availability of specific nutrients. By using these constraints, constraint-based modelling can create predictions for what flows of molecules through the network are plausible under different environmental conditions [13]. Constraint-based modelling predicts metabolism as sets of flux rates for input fluxes, internal fluxes, and output fluxes. These represent the flow of nutrients and metabolites into, through, and out of the cell. One internal flux rate is usually taken as a proxy for the biomass production rate. We use constraint-based modelling to make predictions on individual cells, and integrate many of these predictions to create predictions for the mechanisms behind the infant gut microbiota. These predictions may be used to guide future *in vitro* and *in vivo* research into the infant gut microbiota, and inform potential nutritional interventions. In this chapter we will first explore the typical ecology of the infant gut, followed by how constraint-based modelling works and how it can be used to provide hypotheses on the mechanisms of the infant gut microbiota.

1.1 The infant gut microbiota

1.1.1 Infant health

The infant gut microbiota plays an important role in infant health [14]. Bacteria of the genus *Bifidobacterium* are very common in the infant gut, and are associated with health benefits for infants, such as a lower chance to experience colic [15] or be underweight [16], and reduced inflammation [17]. These positive effects of *Bifidobacterium* are partially explained through its direct interaction with the immune system [17], and partially by the metabolites it produces [18, 19]. Other bacterial species are also associated with improved infant health. For example, butyrate production by species such as *Anaerobutyricum hallii* is associated with a reduced chance to develop allergies or colic [20, 21, 22, 23, 24]. Butyrate is the preferred source of nutrition for the infant's gut colonocytes [25], which may explain its beneficial effects. More generally, many species in the infant gut, including *Bifidobacterium* and butyrate producers, also release other acids into the gut [18, 26]. Acids produced by the microbiota can prevent the growth of pathogens such as some pathogenic *Escherichia coli* strains [19, 18]. Finally, some species in the infant gut can consume the mucin that is secreted by the gut wall [27], which may also affect infant health [28]. This

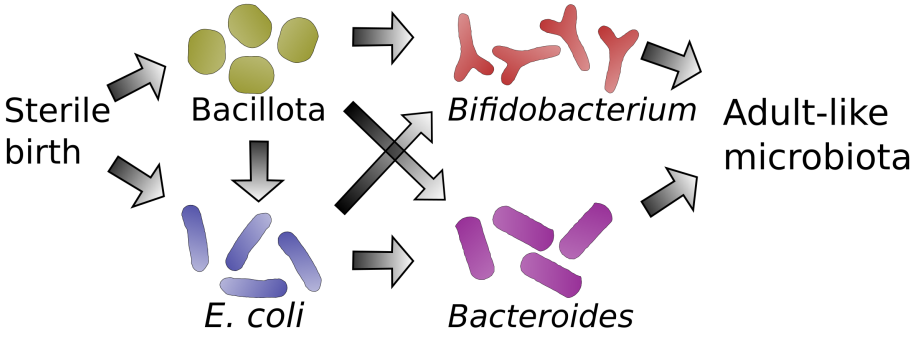

Figure 1.1: Typical temporal progression of dominant bacterial groups in the infant gut microbiota
Arrows indicate common transitions, but other transitions can also occur, and more community states exist. The presence and prevalence of trajectories other than *E. coli* to *Bifidobacterium* to adult-like vary greatly across studies. The trajectories presented here are based on [31, 32, 3, 33, 34]

consumption is associated with increased pathogen susceptibility in mice [28], possibly because mucin consumption exposes the gut wall directly to the bacteria in the gut [29]. A high abundance of *Bifidobacterium* may protect against mucin consumption in breastfed infants [30]. We will address the mechanics behind this in chapter 4. Because of these health effects, we are particularly interested in examining how and why the microbiota reaches the composition that it does, and how this composition may be influenced.

1.1.2 Microbiota composition

Studies of the healthy infant gut microbiota are necessarily limited to fecal samples [35]. These function as a proxy for the actual infant gut microbiota in the infant gut, but are not an unbiased representation of the whole gut microbiota [36, 35, 37]. Species that attach more strongly to the mucin may be under-represented in feces, for example [36, 35]. Initial studies into the composition of the infant gut microbiota were also limited to easily culturable organisms from feces. Though *E. coli* was first isolated and cultured from infant feces [7], other species were harder to culture. Even with the introduction of qPCR, incorrect primers caused a systematic under-reporting of *Bifidobacterium* species in feces until recently [8]. Metagenomic techniques became available to study the gut microbiota in the mid-2000's, and allow

for the analysis of all genomic material in a fecal sample [38]. This provided a culture and primer-independent view of the infant gut microbiota for the first time. However, bias is still present with this method, for example due to species-specific differences in DNA extraction efficiency [39], and some species that are detected with culture-based methods may not be detected by metagenomics [37]. Some general temporal patterns have been observed from the metagenomic analysis of infant fecal samples. These patterns vary greatly due to many factors, only some of which are known. Nonetheless, analysis of fecal samples currently provides the best available view on the infant gut microbiota. These analyses predict that for the first days after birth the infant gut microbiota in fecal samples is typically dominated by Enterobacteriaceae, particularly *E. coli* [32, 16, 31, 40]. This is typically followed by a succession to a microbiota dominated by *Bifidobacterium* species, particularly *Bifidobacterium longum* [32, 16, 31, 40]. However, this succession does not always occur. Various possible paths are visualised in Fig. 1.1. The gut microbiota of some infants (e.g. up to 17.6% in [33]) are instead dominated by Bacillota species such as *Lactobacillus* [33] or *Staphylococcus* [32], and some other infant gut microbiotas are dominated by *Bacteroides* species [34, 40, 3], such as *Bacteroides vulgatus* [3]. Regardless of community type, the microbiota typically shifts to an adult-like composition after several months. This shift often coincides with weaning [41, 40]. Such an adult-like microbiota is usually dominated by Bacillota or *Bacteroides* species [3, 40].

In addition to these very abundant species, many other species are present in the infant gut microbiota at all timepoints [3, 31]. These may include populations of the health-promoting butyrate producing bacteria [3, 20], but may also include pathogens [42] and species such as *Ruminococcus gnavus* and *Cutibacterium avidum*, whose health effects are unclear [43, 44].

Next to the temporal variability of the microbiota there is likely also spatial variation in the abundance of particular species within the microbiota. Such a spatial distribution has been observed in piglets [45] and adult humans [46, 47, 48, 49]. These studies revealed a complex distribution of species and bacterial interactions across the gut. For example, in piglets the relative abundance of *Lactobacillus* is highest proximally [46, 47], while *Bacteroides* is more abundant distally [47]. In adult humans *Bacteroides* species are also more abundant in the distal colon, and *Enterococcus* species are more abundant in the proximal colon [48]. Similarly, Proteobacteria are more abundant close to the colon wall in adult humans, while Bacillota are more abundant in the lumen of the colon [49]. Similar patterns may exist in the infant

gut, but there is no specific knowledge on how the microbiota is spatially organized in the infant gut.

1.1.3 Causes of variation in microbiota composition

The above described patterns are generalisations for the infant gut microbiota, and it is important to note that the presence and timing of these patterns vary greatly across infants [32, 33, 31]. Some of this variation may be explained by differences in the techniques that are used to analyse the microbiota [8], but even within the same studies a great variability is usually found between individuals of the same age [32, 33, 31]. Many studies have attempted to explain some of this variation, and it seems that many different factors interact [50, 51]. We will highlight some of these factors: cesarean or vaginal birth [3], antibiotic usage [52], infant nutrition [53], and region of birth [54]. Together, these can explain some, but not all variation in the observed infant gut microbiota.

Analysis of fecal samples has shown that vaginally born infants typically acquire some bacterial species from the mother's gut microbiota and vaginal microbiota [3, 55]. Although the influence of these sources fades over time, they can greatly influence the species present in the infant gut microbiota in the first days and weeks [56, 55]. Due to the the selection pressures in the infant gut, the effect of birth type on relative abundances of species is typically much smaller. The composition of the infant gut microbiota represents neither the mother's gut microbiota nor the vaginal microbiota [3, 56]. In infants born by cesarean section the relation between the mother's gut and vaginal microbiota is disturbed, and the infant's gut microbiota is richer in species acquired from the skin microbiota and other sources in the environment [3]. This typically leads to a microbiota that is dominated by *E. coli* for longer, and a later transition to *Bifidobacterium* [16]. Cesarean section is also associated with a lower abundance of *Bacteroides* species in the infant gut microbiota [3] and an overall lowered bacterial diversity [55]. However, these are only general patterns, and some infants born by cesarean section have a microbiota indistinguishable from the typical microbiota of vaginally born infants [51]. Altogether, infants born vaginally are expected to have a microbiota that is, on average, more beneficial for the infant's health than those that are born by cesarean section [3]. Breastfeeding may partially compensate for the negative effects on the microbiota of a cesarean section birth [50].

Infants are sometimes treated with antibiotics, which can greatly affect the gut

microbiota even when not administered orally [57]. Antibiotics cause a decrease in *Bifidobacterium* [57, 58], an increase in *E. coli* [59, 60, 58], and a decrease in total diversity [57, 58]. This effect is strongest in vaginally born infants [52]. The microbiota usually recovers, but often only partially [57]. Changes in nutrition may aid recovery [61], but it is unclear how consistently effective these are.

A very important difference in nutrition between infants is the type and quantity of prebiotic oligosaccharides [62]. Prebiotic oligosaccharides are short chains of sugars that cannot be digested by the infant, but can be digested by certain species in the infant gut microbiota, primarily *Bifidobacterium* species [63, 64, 65]. When prebiotic oligosaccharides are present, *Bifidobacterium* species can become very abundant [53]. All human milk contains prebiotic oligosaccharides, but the type and quantity varies based on age and genetics [66, 67]. Milk oligosaccharides also occur in nearly all mammal species [68], even those that do not produce lactose, such as the Virginia opossum (*Didelphis virginiana*)[69, 70]. Milk oligosaccharides are hypothesised to have evolved specifically to stimulate a beneficial infant gut microbiota [65, 70]. Many, but not all, human infant formulas (artificial infant milk) also contain prebiotic oligosaccharides, which may be identical to a type present in human milk, or unique to infant formula [71]. When present, these oligosaccharides also stimulate *Bifidobacterium* species [53]. This creates variability in the microbiota between infants fed formula with oligosaccharides, formula without oligosaccharides, and breastfed infants [53]. Additional variation is introduced by the variable presence of *Bifidobacterium* species. Of the *Bifidobacterium* species, *Bifidobacterium longum* ssp. *infantis* is considered the best adapted to consuming human milk oligosaccharides [65]. However, infants are only likely to acquire *B. longum* ssp. *infantis* in regions with historically high rates of breastfeeding [54].

The birth conditions, nutrition and environment of the infant are known to be important for the composition of the gut microbiota, but it remains unclear how they are mediated. We must turn to bacterial metabolism and ecology to answer why and how these factors ultimately contribute to the observed compositions of the microbiota.

1.1.4 Bacterial metabolism and ecology

Because bacteria are excreted from the gut regularly, it is assumed that most bacterial populations in the gut, or at least the bacterial populations found in feces,

need to grow rapidly to maintain a constant population in the gut [72]. Bacteria need nutrients for growth, which can be derived from three sources in the infant gut: nutrition undigested by the host, such as lactose from milk, nutrients released by the host, such as intestinal mucin, and metabolites produced by other bacteria, such as lactate. The type and amount of each of these sources greatly influence the abundance of bacterial species [53, 73, 74, 75]. Because bacteria can derive nutrition from multiple sources, this also leads to complex interactions between bacteria. Bacteria may, for example, simultaneously compete with each other for host-derived nutrition and benefit from consuming each other's metabolites. We will discuss the sources of nutrition and the possible interactions these allow.

Human milk and infant formula contain carbohydrates, proteins, and fats that can be taken up by the small intestine of the infants [66]. A small quantity of these compounds will inevitably evade uptake, and be available to the microbiota[76, 77]. It is unclear exactly how much lactose is available to the microbiota, but it is estimated to be around 2% [78]. As infant feces does not typically contain lactose [79], the microbiota likely consumes all lactose that reaches the colon [80]. The presence of lactose in infant formula greatly increases the abundance of *Bifidobacterium* [75, 81]. This indicates that at least some of the sugars in infant formula become available to the microbiota. The presence of lactose in the gut may lead to competitive relationships between bacteria.

Human milk and many infant formulas also contain prebiotic oligosaccharides. Prebiotic oligosaccharides serve a unique role because they are resistant to uptake by the infant, and so are nearly entirely available to the microbiota. Specialized enzymes are required to break down these oligosaccharides, so they are not consumed by as many different species as lactose [63, 64]. As mentioned before, prebiotic oligosaccharides are consumed primarily by species of the genus *Bifidobacterium* [63, 64]. Both intracellular and extracellular enzymes for degradation of prebiotic oligosaccharides exist in this genus [82, 83], and the enzymes used depend on both the *Bifidobacterium* species and the oligosaccharide that is digested [82, 83, 84, 85, 86]. Competition occurs between *Bifidobacterium* species for uptake of these oligosaccharides [83].

A second way in which oligosaccharides become available to the microbiota in the colon is through mucin. Mucin is continuously released from the gut wall by dedicated goblet cells, and consists largely of oligosaccharides [29]. The composition of these oligosaccharides depends on age and genetics [29, 73]. Several species, including *Bifidobacterium bifidum* and *Bacteroides* species, can consume these oligosaccharides

[87, 27]. This typically occurs extracellularly, allowing other species to profit from the breakdown products too [27]. Some infants have a microbiota that consumes very little mucin [30]. This indicates that it is possible for this source of nutrition to remain unused.

Extracellular digestion of oligosaccharides, both from mucin and from prebiotics, often causes some breakdown products to be lost to the environment. This makes the products available to other species as 'public goods' [88, 27]. This is also known as substrate cross-feeding [89], and may negatively affect the species that performs the extracellular digestion [89, 88]. Many bacterial species in the infant gut also excrete metabolites, such as lactate, that can serve as nutrition for other species [90, 74], which is known as metabolite cross-feeding [89]. This type of cross-feeding is typically commensalist, as it does not negatively affect the producer of metabolites [89]. Cross-feeding is an important factor in shaping the composition of the infant gut microbiota [27, 91]. Many species, including the main producers of the beneficial butyrate, rely on cross-feeding to acquire nutrients [92]. These species can use both substrate cross-feeding and metabolite cross-feeding [27]. Conversely, cross-feeding is also used by some pathogens to establish themselves in the infant gut microbiota [93, 94]. Some species in the infant gut, such as *Veillonella dispar*, even lack the ability to consume sugars altogether [95]. This species relies entirely on alternative carbon sources such as lactate [74]. The production of cross-feeding substrates is variable - for example *Bifidobacterium* only produces the cross-feeding substrate lactate when sugars are abundant [90]. Other cross-feeding substrates can also be produced by *Bifidobacterium*, such as propane-1,2-diol and fucose, but these depend on the composition of the oligosaccharides consumed [96, 82].

Because our knowledge of the infant gut is limited to *in vitro* experiments and analysis of fecal samples it remains unclear how much production and consumption of cross-feeding substrates occurs *in vivo*, and by what species. This makes it difficult to design, for example, nutritional interventions that stimulate species that cross-feed, such as the butyrate producers associated with improved infant health [20]. *In vitro*, *Bifidobacterium* may consume a specific oligosaccharide and produce metabolites that are consumed by butyrate-producing bacteria. These butyrate-producing bacteria are associated with improved colonocyte health *in vivo*. However, *in vivo* the metabolites produced by *Bifidobacterium* may instead be taken up by other species, that are not beneficial (Fig. 1.2). Even when specific combinations of species can be studied *in vitro*, it is not feasibly to study many combinations due to constraints on time and

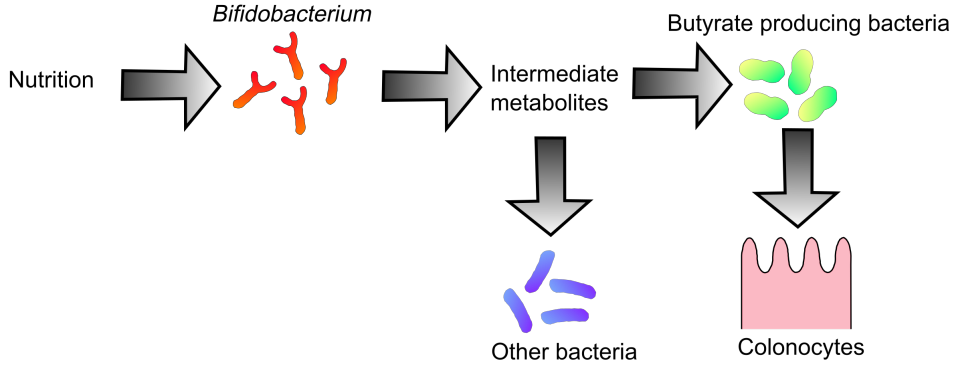

Figure 1.2: Schematic of potential cross-feeding in the infant gut
Nutrition, particularly prebiotic oligosaccharides, is consumed by *Bifidobacterium*, which produces intermediate metabolites such as lactate. These may be taken up by butyrate-producing bacteria, which produce butyrate that can be consumed by colonocytes. However, they may also be taken up by other bacteria that do not produce butyrate.

materials. More generally, it remains unclear how and why the ecology of the infant gut is formed, and what factors are most important in shaping them. Why are these particular species dominant, and others present at a low abundance? Can we shape the microbiota to have a certain composition? We turn to modelling of bacterial metabolism and ecology to try to provide more predictions and hypothesis to guide future experiments.

1.2 Modelling the infant gut microbiota

In this thesis we will use constraint-based modelling techniques to make predictions and create hypotheses for the mechanics of the infant gut. We will first present an overview of the methods used to model bacterial metabolism, followed by modelling of metabolic interactions over time and space.

1.2.1 Modelling bacterial metabolism

Many techniques exist for modelling bacterial metabolism and bacterial interactions, including systems of partial differential equations [97], inferred-network based methodology [98], and constraint-based modelling [99]. We will focus on constraint-based modelling, because of its ability to model very complex metabolisms and metabolic interactions without precise kinetic parameters.

The constraint based modelling techniques we discuss here start with a metabolic network reconstruction [13, 100]. A metabolic network reconstruction consists of a list of reactions, performed by enzymes, each of which has a set of input and output molecules. For example, the beta-galactosidase reaction takes the disaccharide lactose as input, and produces the monosaccharides glucose and galactose. Another reaction may convert galactose to glucose. Together, these reactions and molecules form a reconstructed metabolic network [13, 100] (Fig. 1.3A). From this network model, a stoichiometric matrix can be created that contains the stoichiometric coefficients of each reaction [13]. Many possible configurations of fluxes could flow through this system - we call these configurations solutions. Together, these solution form a solution space. Constraint-based modelling places constraints on the stoichiometric matrix to represent, for example, physical or chemical constraints [13]. This limits the solution space. The shape of the resulting solution space informs what set or sets of fluxes are plausible [13]. Concretely, constraint-based modelling can, as we will discuss in detail later, make predictions for the metabolic inputs and outputs of a bacterial population, what internal reactions are used in the process (Fig. 1.3 B), and how all this depends on the available inputs.

The constraint-based study of bacterial metabolic networks emerged from the study of capacitated flow networks [100]. Capacitated flow networks are graphs where each edge has a certain capacity. Capacitated flow networks were first used in an American study of Soviet railway networks [104]. These networks, where cities are nodes and railway lines are edges, could be analysed to show which lines were most vital to the functioning of the transportation network [104]. Similar analysis can be performed on graph representations of metabolic network models. Here, the nodes of a graph represent enzymes, and the edges linking them represent the molecules that the enzymes can convert into other molecules. Capacity limits can be imposed to represent the limited availability of certain molecules. Early analysis of bacterial metabolic networks models focused on analyzing very simplified networks that con-

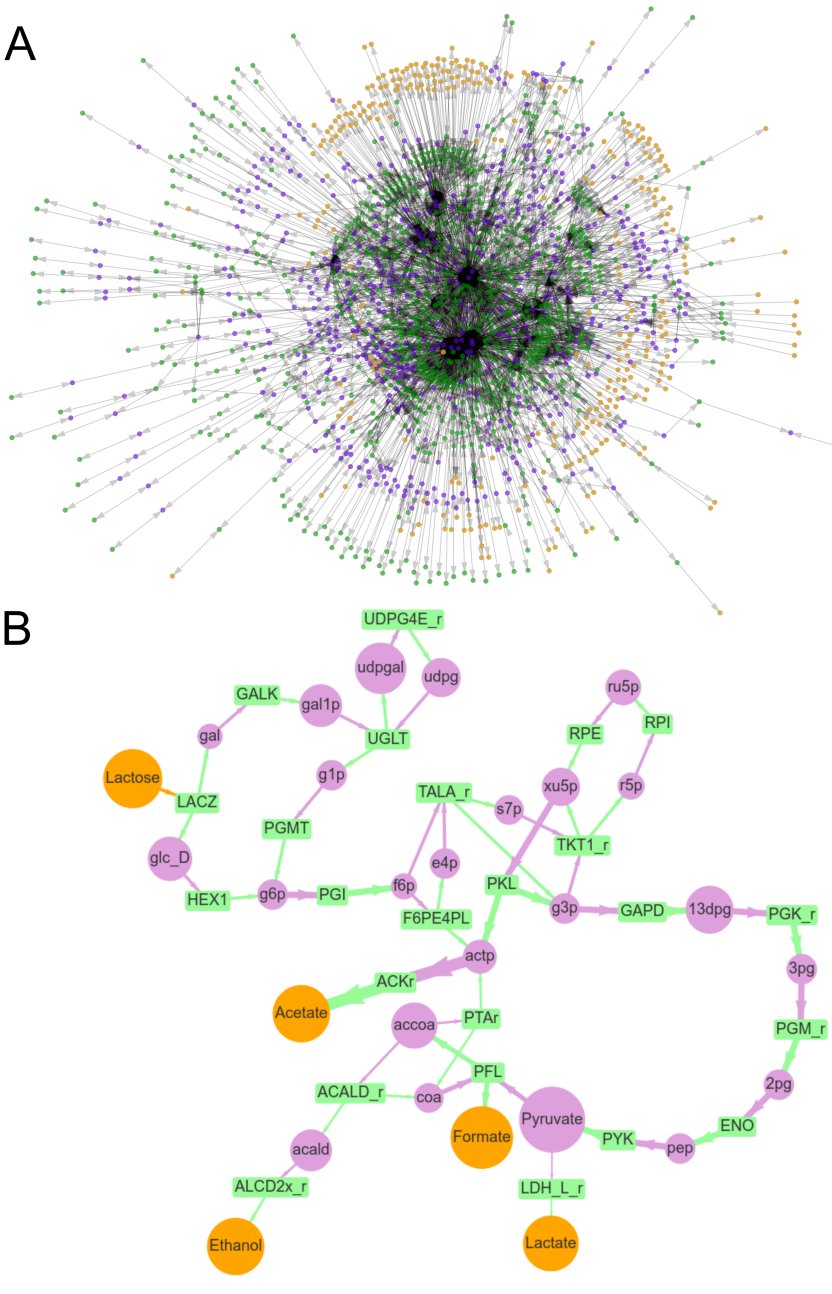

Figure 1.3: Visualisations of a genome-scale metabolic network model of *Bifidobacterium longum* Purple represents intermediate metabolites, orange represents metabolites that can be exchanged with the environment, and green represents reactions. Arrows indicate whether metabolites are consumed or produced by the reactions. GEM and names used from [101]. (A) Visualisation of the whole metabolic network model, created with ModelExplorer [102]. (B) Visualisation of a flux balance analysis solution of the network of A, when optimising ATP production with lactose and water as only inputs, and an enzymatic constraint [103]. Transport reactions and co-factors are not depicted.

tained only a few reactions [100]. More recent techniques allow for the reconstruction of genome-scale metabolic models (GEMs) that may contain thousands of reactions, a large part of the metabolism of a species [10]. These GEMs are generated by analysing the complete genome of a bacterial species, identifying the enzymes and metabolic functions present, and assigning appropriate reactions [10]. Further gap-filling and curation allows for the creation of metabolic models that match much, but not all, of the available *in vitro* data on, for example, which carbohydrates could be fermented by which bacterial species[10]. We use the AGORA database in this work, which contains over 800 GEMs of species within the human gut microbiota, including all major infant gut microbiota species [10]. These models encompass the consumption and creation of sugars, some polysaccharides, and amino acids, but not the prebiotic oligosaccharides that are common in the infant gut, or factors such as protein regulation. The lack of protein regulation means that the constitutive expression of enzymes, which occurs in many bacteria [105, 106], is not represented. These models thus provide a robust basis for modelling, but not a complete representation of metabolism. The models need to be extended to address some more specific questions, such as questions about prebiotic oligosaccharides.

From a GEM, bacterial metabolism can then be analysed using several constraint-based techniques, such as flux variability analysis [107], elementary flux mode analysis [108], and flux balance analysis (FBA) [13]. We will focus on FBA. FBA can create predictions for what fluxes flow through the reactions of a metabolic network under certain conditions. Concretely, FBA predicts a solution for the network. This solution consists of a set of fluxes $\vec{f}$, which includes input fluxes, internal fluxes, output fluxes, and flux through a reaction that is assumed to be a proxy for biomass production. To do this, FBA uses a stoichiometric matrix S generated from the GEM that contains the stoichiometry of all metabolic reactions [13]. Together, all these reactions are assumed to be in steady state:

$$S \cdot \vec{f} = 0,$$

where $\vec{f}$ is a vector of the metabolic fluxes through each reaction in the network, in quantity per time unit per population unit. One reaction is marked as the 'biomass reaction'. FBA assumes that the bacterium regulates its metabolism in such a way that flux through the biomass reaction is optimised. Given these assumptions, FBA uses linear programming to calculate what set or sets of fluxes $\vec{f}$ optimise this biomass

reaction. This set of fluxes is the prediction FBA makes of the bacteria's metabolism. It contains the input and output of metabolites, as well as a growth rate. To place the model in a more specific context, many additional constraints can be used in FBA. For example, reactions that can take up metabolites from the the environment, F_{in}, can be constrained by an upper bound F_{ub} that represents the availability of metabolites from the environment:

$$\vec{F_{in}} \leq \vec{F_{ub}}. \tag{1.1}$$

Additional limits can be set on the network to represent limited enzymatic capacity, such as by limiting the total amount of flux [100, 109]:

$$\sum \vec{f} \leq a \tag{1.2}$$

where the enzymatic constraint a is in quantity per time unit per population unit. Many additional techniques exist to constrain metabolism. For example, thermodynamic constraints can be introduced [110], or membrane occupancy constraints [111]. These constraints allow for more realistic predictions of bacterial metabolism by allowing for, for example, substrate concentration-dependent shifts in metabolism [112]. Specific reactions can also be added or disabled to represent mutated bacterial populations [13].

1.2.2 Modelling bacterial metabolic interactions

Initial models of bacterial metabolism represented a single population of bacteria exchanging nutrients and metabolites with the environment (Fig 1.4A). FBA was used to calculate a steady state, or a set of steady states, which could then be analysed [113, 114]. However, bacterial species are rarely alone in their ecosystem. Later models incorporated multiple species, which could each have their own network of reactions [115]. These populations were coupled to each other by their shared use of the same metabolite pool (Fig. 1.4B). Again, FBA was used to calculate steady states which could be analysed. This allowed for the modelling of competition and cooperation between species. However, many ecological processes are dynamic, and do not maintain a single steady state. To represent these conditions dynamic FBA (dFBA) was developed [116]. In dFBA the model is run in timesteps, and a steady state is calculated for each timestep. Within each timestep the populations are

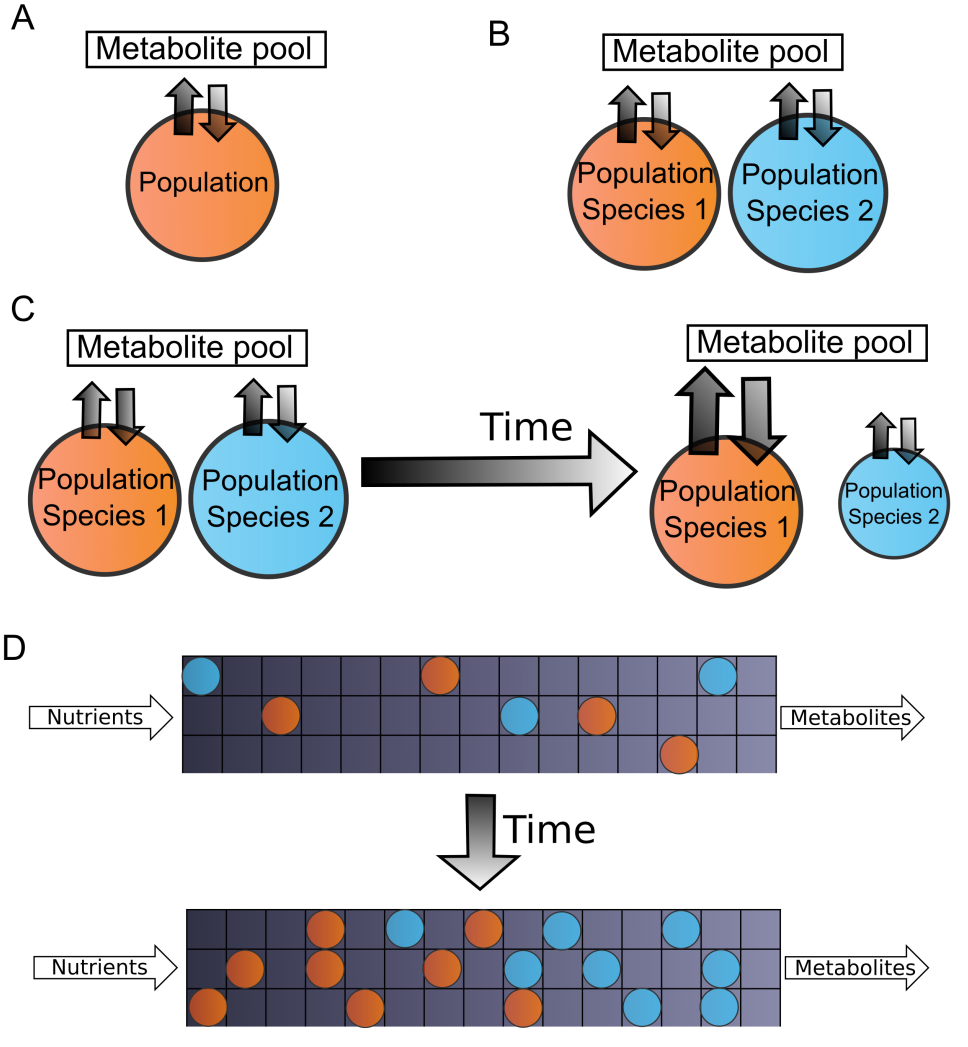

Figure 1.4: Schematics of several ways of modelling bacterial metabolism
(A) A single population exchanging metabolites with a metabolite pool in steady state (B) Two populations exchanging metabolites with the same metabolite pool in steady state (C) Two populations exchanging metabolites with the same metabolite pool in non-steady state, i.e. with dynamic FBA (D) Two species, each with many small populations, on a spatial grid that regulates the exchange of metabolites in non-steady state, i.e. with dynamic FBA. Populations may also move across the grid.

assumed to be in a quasi-steady state [117]. This means that there is no accumulation of nutrients or metabolites within the cells, so that all nutrients taken up are converted into either growth or metabolites. By updating the metabolite pool and population sizes each timestep based on the quasi-steady state solution, and then recalculating the FBA, changes over time can be modelled for one ore more populations [118] (Fig. 1.4C). When multiple species are used this is known as dynamic multi-species metabolic modeling (DMMM)[118]. DMMMs are capable of capturing ecological processes that vary over time, but the spatial dimension is also important in many ecological contexts [119, 46, 120]. To represent this, spatial DMMMs have been developed [121, 122] (Fig. 1.4D). These models also include a spatial dimension that influences how bacterial populations can interact over time [121, 122]. Bacterial populations have specific locations, and exchange of metabolites between populations depends on diffusion and flow [121]. Bacterial populations can also move across locations, so that spatial separation can occur. The addition of spatial mechanics allows the system to model complex competitive relationships between bacterial colonies that would not exist in a well-mixed condition [121]. This approach also allows for the modelling of the spatial dimension of the gut, which we will discuss next.

1.2.3 Modelling the gut microbiota

From the spatial DMMM approach, several models have been developed that provide explanations or predictions for observed effects in the gut microbiota. Van Hoek and Merks [122] based all species in their DMMM on a *Lactobacillus plantarum* genome-scale metabolic model enhanced with other common gut bacteria pathways. They let this model evolve to fill several niches in a spatial simulation of the adult gut [122]. These evolved species also displayed spatial variation, as cross-feeding took place more distally in the gut than primary consumption. The BacArena framework was the first to model the gut microbiota with a variety of species-specific GEMs [123]. This allows the model to make predictions on the metabolic and spatial roles of specific species in the microbiota. Specifically, BacArena predicted that the presence of intestinal mucins drives bacterial diversity, and that a gradient of mucin consumption exists in the gut [123]. Later work with the SteadyCom framework reproduced the spatial variation between facultative and strict anaerobes in the adult human gut, but not the Bacillota/*Bacteroides* ratio typically observed [124]. In this thesis, we will apply the DMMM modelling approach to the infant gut microbiota to create

hypothesis and make predictions specific to the infant gut. We base our model on the spatial model of Van Hoek and Merks [122], which we combine with the GEMs of the AGORA database [10].

1.3 Thesis outline

The thesis is structured as follows. In **chapter 2** we introduce the multiscale mathematical model of the infant gut microbiota. We use the model to examine the role of factors such as enzyme limitations and oxygen availability in shaping succession in the infant gut. In **chapter 3** we expand the model to include the prebiotic oligosaccharides galacto-oligosaccharides (GOS) and 2'-fucosyllactose (2'-FL), and examine how the addition of these prebiotics influence the model predictions. We focus in particular on the possible effects on the health-promoting *Bifidobacterium* species and butyrate producers. In **chapter 4** we further expand the model by including mucins, and the extracellular digestion of mucins, GOS, and 2'-FL. This extracellular digestion allows for cross-feeding to occur in new ways. We examine why breastfeeding may reduce mucin consumption in the infant gut, and explain the different abundance of species through differences in their metabolism, in particular whether or not they produced public goods. In **chapter 5** we examine the effects of disturbances by factors such as antibiotics on the infant gut microbiota, how the microbiota recovers from these disturbances, and how this may be influenced by prebiotic oligosaccharides. Finally, in **chapter 6** we discuss the methods and findings of the thesis in a broader context. We discuss the assumptions and weaknesses of our modelling approach, and give recommendations on future improvements and explorations.