Microbial Risk Analysis xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Microbial Risk Analysis

journal homepage: www.elsevier.com/locate/mran

Using thermodynamic parameters to calibrate a mechanistic dose-response for infection of a host by a virus

Paul Gale

59 Fairway Avenue, Tilehurst, Reading, Berkshire, UK

ARTICLE INFO

Keywords:
Dose-response
Virus
Entropy
Enthalpy
Mucin

ABSTRACT

Assessing the risk of infection from emerging viruses or of existing viruses jumping the species barrier into novel hosts is limited by the lack of dose response data. The initial stages of the infection of a host by a virus involve a series of specific contact interactions between molecules in the host and on the virus surface. The strength of the interaction is quantified in the literature by the dissociation constant (K_d) which is determined experimentally and is specific for a given virus molecule/host molecule combination. Here, two stages of the initial infection process of host intestinal cells are modelled, namely escape of the virus in the oral challenge dose from the innate host defenses (e.g. mucin proteins in mucus) and the subsequent binding of any surviving virus to receptor molecules on the surface of the host epithelial cells. The strength of virus binding to host cells and to mucins may be quantified by the association constants, K_a and K_{mucin}, respectively. Here, a mechanistic dose-response model for the probability of infection of a host by a given virus dose is constructed using K_a and K_{mucin} which may be derived from published K_d values taking into account the number of specific molecular interactions. It is shown that the effectiveness of the mucus barrier is determined not only by the amount of mucin but also by the magnitude of K_{mucin}. At very high K_{mucin} values, slight excesses of mucin over virus are sufficient to remove all the virus according to the model. At lower K_{mucin} values, high numbers of virus may escape even with large excesses of mucin. The output from the mechanistic model is the probability (p_1) of infection by a single virion which is the parameter used in conventional dose-response models to predict the risk of infection of the host from the ingested dose. It is shown here how differences in K_a (due to molecular differences in an emerging virus strain or new host) affect p_1, and how these differences in K_a may be quantified in terms of two thermodynamic parameters, namely enthalpy and entropy. This provides the theoretical link between sequencing data and risk of infection. Lack of data on entropy is a limitation at present and may also affect our interpretation of K_d in terms of infectivity. It is concluded that thermodynamic approaches have a major contribution to make in developing dose-response models for emerging viruses.

1. Introduction

Microbiological risk assessment (MRA) requires a dose-response relationship to translate the exposure (i.e. number of pathogen particles entering the host through a given route) into the probability of infection. Infection by an oral pathogen is defined as the multiplication of organisms within the host, followed by excretion (Haas et al., 1999) and, for the purpose of the work here does not include progression of disease or the host acquired immune response. Obtaining dose-response data for humans has generally relied on volunteer challenge experiments e.g. *Cryptosporidium parvum* in students (Okhuysen et al., 1998) or using outbreak data to back-calculate the relationship between measured exposures and infection rates (Teunis et al., 2004). There are limitations to both approaches particularly with emerging pathogens for which the exposure routes may not be fully elucidated, and for pathogens with serious clinical outcomes, e.g. *Zaire ebolavirus* (EBOV). Furthermore zoonotic viruses emerge through jumping the species barrier from an animal source to humans, e.g. Nipah virus (NiV) and

Abbreviations: Asp, aspartate; Cr, host cell receptor; CRD, carbohydrate-recognition domain; EBOV, Zaire ebolavirus; G, Gibbs free energy; ΔG_a, change in Gibbs free energy on association of virus and cell; GI, gastrointestinal; GP, glycoprotein; H, enthalpy; ΔH_a, change in enthalpy on association of virus and cell; $\Delta\Delta H_a$, change in ΔH_a; HA, haemagglutinin; HBGA, histoblood group antigen; HeV, Hendra virus; k, on/off rate constant; K_a, K_{mucin}, association constants; K_d, dissociation constant for two molecules bound to each other; L, Avogadro number; M, molar (moles dm^{-3}); MRA, microbiological risk assessment; MBP, mannose binding protein; MERS-CoV, MERS coronavirus; n, number of GP/Cr molecular contacts per virus/host cell binding; NiV, Nipah virus; NoV, norovirus; pfu, plaque-forming unit; NPC1, Niemann-Pick C1 protein; Phe, phenylalanine; PL, phospholipid; PRR, pathogen recognition receptor; R, ideal gas constant; S, entropy; ΔS_a, change in entropy on association of virus and cell; SPR, surface plasmon resonance; T, temperature; TIM-1, T-cell immunoglobulin and mucin domain protein 1; VSV, vesicular stomatitis virus

E-mail address: paul@galleryofbirds.com.

https://doi.org/10.1016/j.mran.2018.01.002
Received 9 November 2017; Received in revised form 29 December 2017; Accepted 3 January 2018

EBOV, and in this respect the dose response would be for a one-off event that may be inefficient and difficult to reproduce without large numbers of animals. An additional complication is that the pathogen may adapt to the new host, such that its infectivity increases. This is well established for filoviruses in laboratory animals where the infectivity per plaque-forming unit (pfu) may change by several orders of magnitude with passaging (Gale et al., 2016), and has recently been demonstrated for EBOV Makona adapting to humans through an amino acid substitution in its glycoprotein during the recent catastrophic outbreak in West Africa (Diehl et al., 2016; Urbanowicz et al., 2016). That outbreak also raised many questions regarding the unknown potential for companion animals (cats and dogs) to serve either as a reservoir or vector for the virus and so be involved in transmission of EBOV to humans and other animals. The absence of dose-response data for EBOV in humans limits development of MRAs for the risk of infection of citizens in the EU for example from EBOV in illegally imported bushmeat. Indeed, it has been proposed that the infectivity to humans of an EBOV pfu may differ not only from bushmeat samples from different wildlife species (e.g. fruits bats and nonhuman primates) but also from different individuals of the same species depending on the degree of host adaptation (Gale et al., 2016). In effect no two pieces of bushmeat from EBOV-infected wildlife may be the same in terms of infectivity to humans, although this remains to be proved. There is clearly a need for novel approaches to calibrate dose-response relationships for the purposes of MRA for emerging pathogens.

The infection process of a host cell can be broken down into the component steps and modelled mathematically (Handel et al., 2014) and the probability of infection can be expressed as a function of the combined probabilities of each step (Gale et al., 2014). These steps include overcoming the initial host defenses, binding of the virion to its host cell receptor, entry to the host cell (i.e. internalisation and uncoating of the virion), and replication, capsid assembly and budding (Gale et al., 2014). Previously it was demonstrated that a dose-response model could, in part, be parameterized using thermodynamic data for some of the key molecular interactions in the infection process (Gale, 2017). The beauty of thermodynamic data is that they can be measured experimentally by biochemists (in some cases just using molecular components e.g. cloned virus protein and host receptor protein (Wang et al., 2016)) and do not involve live animal or human volunteer studies, which is a major advantage for dangerous pathogens. Furthermore the effect of amino acid substitutions in the host receptors on binding affinity can be measured directly (Yuan et al., 2015). The possibility of applying thermodynamics is further developed here for two of the key steps in the infection process of a host by a virus. The first step modelled is the probability of the virus overcoming the innate host defenses posed by mucin protein molecules and the pathogen recognition receptors (PRRs) produced by the host. Mucins have sugar units on their surface which bind to components on the surface of the virus, for example the haemagglutinnin (HA) glycoprotein molecules of influenza virus (de Graaf and Fouchier, 2014) or the VP1 of norovirus (NoV) (de Rougemont et al., 2011). Mucus present in the respiratory tract hampers influenza virus infection and in the case of humans predominantly contains α2,3-sialic acid receptors. Indeed influenza viruses with α2,3 specificity were inhibited by human mucins (de Graaf and Fouchier, 2014). The PRRs include the mannose binding protein (MBP) which has carbohydrate-recognition domains (CRD) which bind to regularly repeating sugar units on pathogen surface (Taylor and Drickamer, 2006). The second step modelled here is the binding of the virus to its specific receptors on the host cell surface. The approach here is developed for a generic faecal/oral virus such as NoV and rotavirus which infects epithelial cells lining the intestine (Boshuizen et al., 2005; de Rougemont et al., 2011), but could be applied to influenza A viruses which are inhaled and infect cells of the trachea and lung (de Graaf and Fouchier, 2014).

This paper first gives an overview of a mechanistic dose response model to introduce two probability parameters, namely the fraction, F_v, of virus escaping the mucin defense barrier and the fraction, F_c, of host cells with bound virus. The Methods section sets out a difference equation method to model F_v and F_c as a function of the mucin: virus ratio and virus dose in the intestine, respectively. Central to determining F_v and F_c is the strength of binding of the virus to the mucin and host cell as defined by the equilibrium constants K_{mucin} and K_a respectively. In the Theory section, the application of published data on the binding of the virus surface envelop glycoprotein (GP) to host cell receptor (Cr) molecules or to mucin molecules is reviewed in terms of determining K_{mucin} and K_a in order to parameterize the dose-response. Particular reference is made to using the dissociation constant K_d which is routinely determined experimentally for virus GPs binding to Cr molecules (Gambaryan et al., 2005; Raman et al., 2014; Yuan et al., 2016). It is then shown how the strength of virus/host cell binding (i.e. the magnitude of K_a) may be predicted from changes in two thermodynamic parameters, namely enthalpy (H) and entropy (S). The effects of amino acid changes at the contact surfaces of the virus GP and Cr on the enthalpy are considered with a view to the future parameterization of dose-response models based on genetic sequencing data. Entropy changes are also considered both in terms of virus binding and also in the interpretation of K_d data.

2. Methods

2.1. Overview of the development of a mechanistic dose-response model for infection in the intestine

The model parameters and variables are summarised in Table 1. On ingestion of the initial virus challenge dose, $V_{initial}$, by the host there are a number of immediate host defences in the mouth and gastrointestinal (GI) tract including the mucus barrier, decoy receptors and the innate immune system that selectively bind and hence remove the virus (McGuckin et al., 2011). For example the histoblood group antigens (HBGAs) are genetically determined glycans to which NoV selectively binds and are present on both decoy receptors in the saliva and on mucin, the main protein component of mucus (Shanker et al., 2011). The total number of viruses surviving the mucin barrier, and getting through to the intestine is given by

$$\boldsymbol{V_{intestine} = F_V \times V_{initial}} \quad (1)$$

where F_v is the fraction of free virus, i.e. that not bound to mucin. As shown in Fig. 1, F_v can be modelled by two parameters, namely the total number, Muc_{total}, of mucin molecules in the mucus in the saliva and GI tract and an association constant, K_{mucin} (defined below) that quantifies the strength of binding of the virus to a mucin molecule. Thus by inserting F_v from Fig. 1 for a given mucin concentration into Eq. (1), the total number of free virus particles in the intestine and available to initiate infection of the epithelium may be modelled. On reaching the intestinal epithelium, a free virus particle binds to the surface of a host cell. The probability of infection of the host, p_{host}, equals the probability of successful infection of at least one cell and is related to the number of cells (C.V) with bound virus by:-

$$\boldsymbol{p_{host} = 1 - (1 - p_{cell})^{C.V}} \quad (2)$$

where p_{cell} is the probability of successful infection of a host cell given a virus has bound to its surface. Thus the more cells with bound virus then the greater the chance that infection will be successful in at least one of them. The probability p_{cell} depends on ability of the bound virus to enter the cell, replicate and bud (Gale et al., 2014; Gale, 2017) and is not discussed further here. Now

$$\boldsymbol{C.V = F_c C_{total}} \quad (3)$$

where F_c is the fraction of cells with bound virus, and C_{total} is the number of cells in the host intestinal epithelium. As shown in Fig. 2, F_c is directly proportional to the total number of virus particles in the intestine, $V_{intestine}$, and is also dependent on the strength of the binding