Bulletin of Mathematical Biology (2004) **66**, 523–537
doi:10.1016/j.bulm.2003.09.001

Use of a Quantitative Gene Expression Assay based on Micro-array Techniques and a Mathematical Model for the Investigation of Chlamydial Generation Time

D. P. WILSON*
School of Mathematical Sciences,
Queensland University of Technology,
GPO Box 2434,
Brisbane Qld 4001,
Australia
E-mail: d.wilson@fsc.qut.edu.au

S. MATHEWS AND C. WAN
School of Life Sciences,
Queensland University of Technology,
Brisbane Qld 4001,
Australia

A. N. PETTITT AND D. L. S. MCELWAIN
School of Mathematical Sciences,
Queensland University of Technology,
Brisbane Qld 4001,
Australia

Chlamydia is an important pathogen which possesses a unique developmental cycle. We used real-time PCR technology to measure gene transcript levels in *Chlamydia trachomatis* strain L2. By measuring 16S rRNA transcript levels, and developing a mathematical model of the chlamydial developmental cycle fitting the data, we predict an average generation time of approximately 2.6 h. Additionally, potentially this modelling also provides the foundation for the application of emerging micro-array technology in which identification of the gene signals that trigger a chlamydial body to start replicating or transform to its infectious form can be made possible.

1. INTRODUCTION

Infection with *Chlamydia* is common around the world where human chlamydial infections are predominantly caused by *Chlamydia trachomatis* and *Chlamydia pneumonia*. *C. trachomatis* infection leads to trachoma, the highest cause

*Author to whom correspondence should be addressed.

0092-8240/04/030523 + 15 $30.00/0

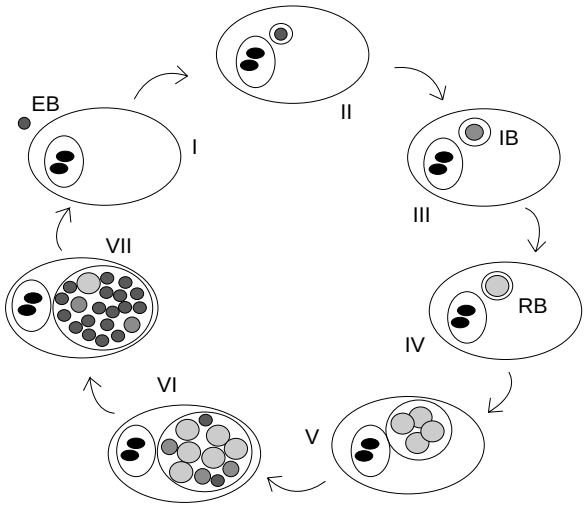

Figure 1. Schematic representation of the developmental cycle of *Chlamydia*. The different stages of chlamydial development are illustrated commencing when the EB attaches to the host cell (I) and is endocytosed within the first 2 h PI and remains within the inclusion (II). From 2–8 h PI the EB commences transformation via an IB (III) to become a RB (IV). The RBs multiply within the inclusion (V) before some RBs convert to EBs and some RBs continue to divide after 12 h PI (IV). RBs that have committed to become EBs are classified as IBs. Late development (VII) is identified when RB multiplication has slowed down and RB to EB conversion is at the highest rate (24–48 h PI).

of preventable blindness, and is also the most common human sexually transmitted disease. *C. pneumoniae* causes a broad spectrum of respiratory disease, including pneumonia and is associated as an etiological agent for atherosclerosis. Although chlamydial infections can be effectively treated with antibiotics, many infections are asymptomatic resulting in worsening disease when left undiagnosed and untreated.

Chlamydiae are obligate intracellular parasites with a unique developmental cycle characterised by the inter-conversion between morphologically different forms (see Fig. 1). Development commences when the infectious and metabolically inert chlamydial cell, termed the elementary body (EB), attaches and enters a susceptible host cell where it remains in the host-derived vacuole known as the inclusion. The EB reorganises its outer membrane and DNA structure during the first 6–8 h post-infection (PI) prior to developing into a replicative reticulate body (RB). The metamorphosis of EB to RB includes passage via an intermediate body (IB), and although this form of chlamydiae is not part of the classic developmental cycle it is relevant to our investigations. The RBs subsequently divide by binary fission until approximately 12–18 h PI when chlamydial development becomes asynchronous with some RBs converting to EBs and others still replicating. The re-organisation of RB to EB involves an IB that cannot be classified as either EB or RB. As development continues into the late phase of the cycle more RBs

are committed to becoming EBs and the proportion of EBs in the inclusion increases. The end of the developmental cycle is characterised by the release of infectious EBs from the host cell, the whole process taking approximately 40–72 h for *C. trachomatis*.

Current estimates for RB doubling times range from 2 h (McClarty, 1994) to 3 h (Mathews *et al*., 1999). Additionally, the rate of RB commitment to EBs is not known apart from the relatively simple modelling by Mathews *et al*. based on two general term sequences although this lacks a firm biological basis. Because chlamydial development is asynchronous it is difficult to determine the number of chlamydial bodies and progression of infection. Here, mathematical modelling can be utilised to track and predict the population of each type of chlamydial body throughout the developmental cycle. We use recently developed quantitative real-time polymerase chain reaction (PCR) technology to measure gene transcript levels (RNA) of the chlamydial 16S rRNA gene. PCR increases the number of complimentary DNA (cDNA) copies generated from the RNA. We obtain more real-time PCR data than has been obtained previously and we develop a mathematical model to fit the data in order to obtain a better approximation for the RB doubling time, and the rate of RB to EB commitment. We predict, based on our data, that for *C. trachomatis* L2 the doubling time is between 2.56 and 2.63 h (95% confidence) with a best fitting doubling time of 2.59 h.

2. PCR Experiment Materials and Methods

Despite the importance of chlamydial disease, the developmental expression of relatively few genes has been determined. The major difficulties in quantitative analysis of chlamydial gene expression are the inability to culture *Chlamydia* in a host-free environment, the asynchronous nature of chlamydial development, measuring the low level transcripts within the host cell background and difficulty in standardising for the number of chlamydiae within an inclusion. Although many studies have been significant in reporting quantitative gene expression, most have not been standardised for the number of chlamydial particles to provide a measurement of gene expression. Consequently, the presence of increased mRNA or protein levels may be associated with increased chlamydiae per host cell. Nevertheless, the relative developmental expression profiles of several genes have been determined with *euo* being described as an early gene (Zhang *et al*., 1998), omcB, hctA and hctB as late genes (Koehler *et al*., 1990; Hackstadt *et al*., 1991; Perara *et al*., 1992) and many genes expressed constitutively, including the 16S rRNA gene (Engel and Ganem, 1987) and the groESL operon (Lundemose *et al*., 1990). More recent investigations have increased the repertoire of known developmental-stage-specific expression using quantitative RT-PCR (Shaw *et al*., 2000) and microarrays (Nicholson *et al*., 2003). The first global stage-specific gene expression analysis of *C. trachomatis* using microarray analysis divided the chlamydial genome

into seven differential expression profiles (Nicholson *et al*., 2003). The presence of constitutively expressed genes allows them to be used to standardise for the number of chlamydial particles during normal development.

The emergence of real-time PCR technology has allowed the development of an assay to accurately determine the level of relative gene transcripts (RNA) during chlamydial development (Mathews *et al*., 1999). Since the number of chlamydial particles increases per host cell during development and the 16S rRNA gene is constitutive, the 16S rRNA was chosen as a reference to standardise for the number of chlamydial particles. The study involved generating cDNA of the 16S rRNA by random priming and confirmed that random priming of total RNA (isolated from *C. trachomatis* infected HEp-2 cells) generated cDNA representative of the RNA levels. In the same study, Mathews *et al.* determined the relative mRNA level of the late-stage-specific *omcB* gene and the three RNA polymerase σ factor genes. The details of the experimental setup, method of data collection, and data used in the model are given in the Appendix.

3. MATHEMATICAL MODEL

3.1. ***Observed data.*** cDNA is a DNA copy of RNA and its amount has been shown to be a good approximation to the amount of RNA in the culture. The ratio of RNA to DNA differs between EBs and RBs. Every chlamydial particle has one copy of DNA. RBs have a RNA to DNA ratio of approximately $4:1$ and EBs have a ratio of approximately $1:1$. We assume that on average, IBs have a RNA to DNA ratio of $5:2$. The *omcB* transcript levels reflect the number of RBs committed to become EBs because this gene product is only found in the outer membrane of EBs and thus indicates the number of IBs. The previous data of Mathews *et al*. shows the levels of *omcB* cDNA and DNA over the developmental cycle. Since the data does not directly give the number of EBs, IBs and RBs we assume that in terms of EBs, IBs and RBs, the total DNA at time t is given by

$$\mathrm{DNA}(t) = k_1(E_B(t) + I_B(t) + R_B(t)), \tag{1}$$

the total cDNA at time t is given by

$$\mathrm{cDNA}(t) = k_2(2E_B(t) + 5I_B(t) + 8R_B(t)), \tag{2}$$

and the total *omcB* at time t is given by

$$omcB(t) = k_3 I_B(t), \tag{3}$$

where k_1, k_2 and k_3 are unknown constants.

3.2. ***Theoretical simulation.*** Previous mathematical modelling of the biology of the chlamydial developmental cycle has involved equations for the different stages of the cycle (Wilson *et al.*, 2003). However, the triggers within the inclusion that cause the end of one stage and the commencement of another stage are unknown. For example, it is not known how RBs that are purely dividing are prompted to commit to transform back to EBs. We postulate that the transition between different stages are continuous and are nutrient dependent since chlamydiae require access to the nutrient-rich cytoplasm of a host cell in order to construct the macromolecules [protein, DNA, RNA, lipids, lipopolysaccaride (LPS) and murein] that make up a chlamydial particle. Although RBs are purely dividing initially and the rate of commitment to IBs is zero, we assume that there is an increasing, yet small rate of transformation. Thus, we model the dynamics of all chlamydial bodies over the entire developmental cycle. We assume that nutrients are consumed by chlamydial particles. We let $N(t)$ be the concentration of nutrients at time t. The parameter t is a time parameter over one developmental cycle. We let $I_1(t)$, $R(t)$, $I_2(t)$, $E(t)$ represent the concentration at time t of chlamydial bodies at the different stages, namely, IBs transforming from EBs to RBs, RBs, IBs transforming from RBs to EBs, and EBs respectively. Assuming that nutrient consumption is governed by a mass-action law, we obtain

$$\frac{dN}{dt} = -[c_1 I_1(t) + c_2 R(t) + c_3 I_2(t) + c_4 E(t)]N(t), \tag{4}$$

where c_1, c_2, c_3 and c_4 are rate parameters representing the level of nutrient consumption by each type of chlamydial body. Here, since we are modelling an *in vitro* experiment we do not include a source of nutrients. A model of the *in vivo* developmental cycle would include a nutrient source. We assume that the chlamydial concentrations are changing according to the following model system equations:

$$\frac{dI_1}{dt} = -k_T I_1(t), \tag{5}$$

$$\frac{dR}{dt} = k_T I_1(t) + \frac{\ln 2}{d_t} R(t) - k\left(\frac{N_0 - N(t)}{N_0 + N(t)}\right)^n R(t), \tag{6}$$

$$\frac{dI_2}{dt} = k\left(\frac{N_0 - N(t)}{N_0 + N(t)}\right)^n R(t) - k_T I_2(t), \tag{7}$$

$$\frac{dE}{dt} = k_T I_2(t). \tag{8}$$

Here, k_T is the average rate of transformation of an IB as it progresses from an EB to a RB, d_t is the average doubling time of RBs, k is a rate parameter for RBs committing to IBs to become EBs, and N_0 is the initial nutrient level. We assume that the average time for an EB to transform to a RB is the same as the average time for a RB to transform to an EB, represented by the parameter k_T in equations (5)

and (8). The loss term in equation (6) represents the rate of RBs committing to EBs and is governed by nutrient deprivation (Timms and Mathews, 2002). Initially, the rate is very small but as the supply of nutrients decreases the rate of commitment increases to a maximal commitment rate. The parameter, n, is a shape factor. There is no data available to estimate the nutrient consumption of each chlamydial type and is an area for future investigation. However, nutrient uptake is most certainly higher for RBs than EBs because RBs are metabolically active while EBs are not. Additionally, RBs are seen in electron microscopy to be close to the inclusion membrane where they can gain access to host cell nutrients but EBs are mostly seen in the lumen of the inclusion where host cell nutrients are not as accessible (Rockey and Matsumoto, 1999). Then, to investigate nutrient depletion we look at three cases: (i) only RBs deplete host cell nutrients ($c_1 = c_3 = c_4 = 0$); (ii) RBs deplete host cell nutrients twice as readily as IBs do ($c_1 = c_3 = 1/2c_2$, $c_4 = 0$); (iii) all chlamydial particles contribute equally to the depletion of host cell nutrients ($c_1 = c_2 = c_3 = c_4$). We suspect that case (ii) is most realistic, although the exact contribution of nutrient depletion by each chlamydial body is unknown.

We nondimensionalise the system by introducing the variables

$$\hat{N}(t) = \frac{N(t)}{N_0}, \qquad \hat{T}(t) = \frac{I_1(t)}{T_0}, \qquad \hat{R}(t) = \frac{R(t)}{T_0},$$
$$\hat{I}(t) = \frac{I_2(t)}{T_0}, \qquad \hat{E}(t) = \frac{E(t)}{T_0}, \qquad \hat{t} = c_2 t,$$
$$\alpha_1 = \frac{k_T}{c_2}, \qquad \alpha_2 = \frac{\ln 2}{c_2 d_t}, \qquad \alpha_3 = \frac{k}{c_2},$$

where T_0 is the concentration of IBs at time $t = 0$. We obtain

$$\frac{dT}{dt} = -\alpha_1 T(t), \tag{9}$$
$$\frac{dR}{dt} = \alpha_1 T(t) + \alpha_2 R(t) - \alpha_3 \left(\frac{1 - N(t)}{1 + N(t)}\right)^n R(t), \tag{10}$$
$$\frac{dI}{dt} = \alpha_3 \left(\frac{1 - N(t)}{1 + N(t)}\right)^n R(t) - \alpha_1 I(t), \tag{11}$$
$$\frac{dE}{dt} = \alpha_1 I(t), \tag{12}$$
$$\frac{dN}{dt} = \begin{cases} -R(t)N(t), & \text{case (i)} \\ -[1/2T(t) + R(t) + 1/2I(t)]N(t), & \text{case (ii)} \\ -[T(t) + R(t) + I(t) + E(t)]N(t), & \text{case (iii)} \end{cases} \tag{13}$$

where we have dropped the *hat* notation for convenience. The initial conditions are $N(0) = 1$, $T(0) = 1$, $R(0) = 0$, $I(0) = 0$, $E(0) = 0$.

4. A SIMPLIFIED MODEL

Preliminary study with the full model led to the conclusion that n needed to be large to fit experimental data irrespective of the assumptions used for nutrient depletion. We know that biologically there is an intracellular trigger causing the commencement of RB commitment to become EBs, thus it is plausible that n is large in equation (10), effectively causing the rate of commitment to be expressed mathematically as a Heaviside function. Furthermore, the rate of RB commitment, $\alpha_3\left(\frac{1-N(r)}{1+N(r)}\right)^n$, can be expressed as

$$\alpha_3\left(\frac{1-N(r)}{1+N(r)}\right)^n = \alpha_3 \exp[n\ln(1-N) - n\ln(1+N)], \tag{14}$$

but since $\ln(1-N) = -N + \mathrm{O}(N^2)$ and $\ln(1+N) = N + \mathrm{O}(N^2)$,

$$\alpha_3\left(\frac{1-N(r)}{1+N(r)}\right)^n \approx \alpha_3 \exp(-2nN), \tag{15}$$

and thus the rate of commitment becomes significant only if $N \approx \frac{1}{n}$. Consequently, we reduce our model equations to the study of early dynamics and late dynamics of the developmental cycle, before and after the switching on of the RB commitment term. In addition, if n is large, the choice of parameters c_1, c_2, c_3 and c_4, cases (i), (ii) and (iii), and indeed any other choice also leads to similar predictions because, prior to the rate of commitment term becoming significant, the only variable with significant dynamics is $R(t)$. This follows since in the first part of the developmental cycle RBs are replicating by binary fission and this is the only significant event occurring.

4.1. ***Case 1: Early dynamics.*** The intracellular dynamics prior to the commitment of RBs to become IBs can be expressed by the following approximate system:

$$\frac{dT}{dt} = -\alpha_1 T(t), \tag{16}$$

$$\frac{dR}{dt} = \alpha_1 T(t) + \alpha_2 R(t), \tag{17}$$

with solution $T(t) = e^{-\alpha_1 t}$ and $R(t) = \frac{\alpha_1}{\alpha_1+\alpha_2}(e^{\alpha_2 t} - e^{-\alpha_1 t})$. In nondimensionalising the simple model, we have put $\bar{t} = t/t^*$ where the bar has been dropped for convenience and t^* is the time for one developmental cycle.

4.2. ***Case 2: Late dynamics.*** The intracellular dynamics following the trigger for RBs to become IBs can be approximated by the following system:

$$\frac{dR}{dt} = -(\alpha_3 - \alpha_2)R(t) \tag{18}$$

$$\frac{dI}{dt} = \alpha_3 R(t) - \alpha_1 I(t) \tag{19}$$

$$\frac{dE}{dt} = \alpha_1 I(t). \tag{20}$$

Here, we assume that the population of IBs converting to RBs is negligible because this process has already occurred. Assuming that the transition from early to late dynamics occurs quickly and that the trigger is at time $t = t_0$, the solution to the late dynamics model equations is given by

$$R(t) = \frac{\alpha_1}{\alpha_1 + \alpha_2}(e^{\alpha_2 t_0} - e^{-\alpha_1 t_0})e^{-(\alpha_3 - \alpha_2)(t - t_0)} \tag{21}$$

$$I(t) = \frac{\alpha_1 \alpha_3 (e^{\alpha_2 t_0} - e^{-\alpha_1 t_0})(e^{-(\alpha_3 - \alpha_2)(t - t_0)} - e^{-\alpha_1 (t - t_0)})}{(\alpha_1 + \alpha_2 - \alpha_3)(\alpha_1 + \alpha_2)} \tag{22}$$

$$E(t) = \frac{\alpha_1 \alpha_3 (e^{\alpha_2 t_0} - e^{-\alpha_1 t_0})}{(\alpha_1 + \alpha_2)(\alpha_3 - \alpha_2)} \times \left(1 - \frac{(\alpha_3 - \alpha_2)\exp(-\alpha_1 (t - t_0)) - \alpha_1 \exp(-(\alpha_3 - \alpha_2)(t - t_0))}{\alpha_3 - \alpha_2 - \alpha_1}\right). \tag{23}$$

Our simplified model also has four parameters to be determined but has the advantage that simple closed form solutions are available and the dependence of RB loss on nutrient availability is removed. Thus, the assumptions of cases (i)–(iii) in the full model of the contribution each chlamydial body makes to nutrient consumption is eliminated. It could be suggested that the simplified model supports the hypothesis of chlamydial development due to a developmental time clock (Ward, 2003). That is, a specific point in time during the cycle triggers the activation of the next stage in development.

5. Solution and Results

We fit our mathematical model to the cDNA data and from equation (2) take the parameter, k_2, to be

$$k_2 = \frac{\text{cDNA}(0)}{2E_B(0) + 5I_B(0) + 8R_B(0)} = 0.2\ \text{cDNA}(0) = 2.759 \times 10^6. \tag{24}$$

There are various methods that may be employed to fit parameters to the model equations. Additionally, an appropriate choice for the error function to be minimised must be made of which various choices can be made. Once estimates are made for the parameters, the uncertainty of these estimates is also important, and it

would be ideal if a probabilistic density function for each quantity that is being estimated is computed. When carrying out parameter estimation of non-linear models, a least squares or total least squares algorithm is very likely to give biased results in that the approach assumes independent, normally distributed (IND) forecast errors and nonlinear models will not yield IND errors even if the noise is IND (McSharry and Smith, 1999). A maximum likelihood approach would be adequate. However, since experiments have been performed in triplicate, we have some degree of confidence in the accuracy of our data and we use the sample variances as weights in non-linear weighted least squares and then employ a quasi-likelihood estimation technique to determine a profile likelihood for each of our estimated parameters. The best fitting set of parameters to the weighted least squares algorithm is consistent with the solution determined by a maximum likelihood algorithm if cDNA(t) is of the exponential family of distributions. Finally, we establish a 95% confidence interval for each estimated parameter.

A numerical algorithm minimized the weighted sum of squared residuals generated by

$$g(\underline{\alpha}) = \sum_{t=\text{experiment time pts.}} \left[\frac{\text{cDNA}(t) - \widehat{\text{cDNA}}(\underline{\alpha}, t)}{\text{var}(\text{cDNA}(t))} \right]^2, \tag{25}$$

where $\underline{\alpha}$ is a vector containing the set of parameters to be determined ($\alpha_1, \alpha_2, \alpha_3$ and n for the full system and $\alpha_1, \alpha_2, \alpha_3$ and t_0 for the simplified system). In the error function, cDNA(t) is the experimental level of cDNA data, var(cDNA(t)) is the sample variance, and $\widehat{\text{cDNA}}$ is the predicted cDNA level from our mathematical model. A fourth-order Runge–Kutta method was used to solve the full model system of coupled first-order initial value equations, and the closed form solutions was used in fitting for the simplified model. A steepest decent algorithm was used to obtain a search direction in parameter space for a better sequence of parameters until appropriate convergence of parameters was achieved. Various high, low and intermediate values were chosen as initial values for each parameter in order to obtain confidence that the final best fitting set of parameters gave rise to not just a local, but the global minimum of the error function g.

Our data fitting algorithm for the full model system was performed using a range of values of n whilst ensuring physically realistic results arise. Throughout each data fitting procedure n is kept fixed and optimal values for other parameters were determined. It was determined numerically that the integer parameter n, must be large ($O(10^3)$) for a reasonable fit to the experimental data. The parameter, n, is large reflecting the triggering that occurs biologically as RBs commit to transform. Thus, it seems reasonable to use the simplified model.

As stated earlier, there are various ways of fitting the model to the data. Other methods, such as a relative least squares technique, provides a fit that appears to be a better fit. However, we do not have as much confidence in this fit because the mean experimental data is not reliable enough. We note that the sample variance tends to increase with time over the developmental cycle. Then, we do not have

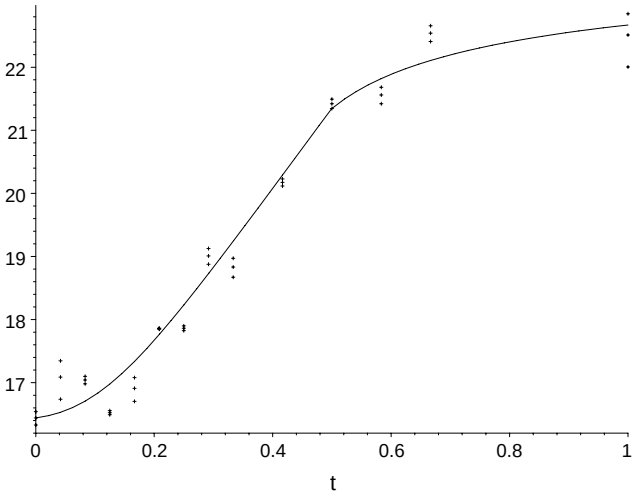

Figure 2. The cDNA data from our real-time PCR experiments, with standard error bars, and the best fitting curve for the simplified model. The plot shows log(cDNA) versus time.

as much confidence in predictions of late dynamics. However, the variance is relatively small over early dynamics and it is over this period we may have increased assurance of our predictions. The cDNA data and curve fitting using the best fitting parameter set for the simplified model are illustrated in Fig. 2. The simplified model yields profiles for each chlamydial population and cDNA levels that are in good agreement with simulations of the full model (full model simulation not shown). Apart from the obvious advantage of obtaining a closed form solution to the model system of equation, the simplified equations also provides considerable computational advantage when finding the best fitting parameter set.

5.1. ***Confidence intervals for parameter estimates.*** The confidence intervals are based on the ideas of profile likelihood [see, for example, Cox and Hinkley (1974)]. Here, we assume for a given choice of a single (scalar) parameter, ϕ, with the remainder given by (vector) $\underline{\psi}$, that the likelihood is given by $L(\phi, \underline{\psi})$. For fixed ϕ we maximise $L(\phi, \underline{\psi})$ giving $\hat{\underline{\psi}}(\phi)$ as the solution for $\underline{\psi}$. The profile likelihood method gives a 95% confidence interval as those values of ϕ satisfying

$$-2l(\phi, \hat{\underline{\psi}}(\phi)) \leq -2l(\hat{\phi}, \hat{\underline{\psi}}(\phi)) + K, \tag{26}$$

with $l(\phi, \underline{\psi}) = \log L(\phi, \underline{\psi})$ and K the upper 5% of χ_1^2. To provide an estimated 95% confidence interval for each parameter value we plot $g(\alpha_1, \hat{\alpha_2}, \hat{\alpha_3}, \hat{t_0})$ versus α_1 where the hat denotes the converged parameter value. Then, the two solutions of $g(\alpha_1, \hat{\alpha_2}, \hat{\alpha_3}, \hat{t_0}) = g_{\min} + K$, give the bounds on the 95% confidence interval. We apply a similar technique for each parameter. The results obtained using this technique are given below for an extended model.

However, the statistical model fitted is not adequate to explain all the lack of fit. From Fig. 2 it is obvious there is additional error to that from the replication of the experimental points for each value of t. As an aside, if the experiments were repeated in triplicate and the model were 'true' then one would expect, for each of the 3 sets of points, for given t, each point randomly lying above or below the mean. Instead, for each t, the three values all lie below or above the mean (fitted curve) except for $t = 1.0$. We need to extend the model to take account of this extra level of error and the implied extra uncertainty. For each t we assume the experimental cDNA(t) values (average of 3) measure a value of the mean cDNA level at time t, $\mu(t)$, without bias. However, these $\mu(t)$ are distributed about the true (model) values $m(t)$ with error with standard deviation σ. Writing $s(t)$ for the (estimated) standard deviation of experimental values (average of 3) for each t and $m(t)$ is given by cDNA($\underline{\alpha}, t$), we extend the criterion (25) to include the parameter σ^2. The quantity is given by -2 loglikelihood for this extended model assuming normal errors at each level and we take

$$g(\underline{\alpha}, \sigma) = \sum_t \left[\frac{\text{cDNA}(t) - \widehat{\text{cDNA}}(\underline{\alpha}, t)}{s(t)^2 + \sigma^2} \right]^2 + \sum_t \log(s(t)^2 + \sigma^2). \tag{27}$$

The previous profile likelihood procedure is followed by extending the set of parameters to include σ. It will make the confidence intervals wider. Also, as it stands, the $t = 0.2$ point is fitted too precisely because var(cDNA($t = 0.2$)) ~ 0. This new statistical model will allow for more error at $t = 0.2$ and might give a better overall fit.

When we incorporate the additional parameter, σ, to account for the error in the model, and use the error function given in equation (27) the best fitting set of parameters are $\alpha_1 = 2.06$, $\alpha_2 = 12.80$, $\alpha_3 = 13.45$, $t_0 = 0.49$ and $\sigma = 2.68 \times 10^7$. The mean RB doubling time, d_t, is evaluated as

$$d_t = \frac{\ln(2)t^*}{\alpha_2} = 2.59, \text{ approximately.} \tag{28}$$

We apply the technique described above, and illustrated in Fig. 3, to obtain 95% confidence intervals of $\alpha_1 \in (1.89, 2.23)$, $\alpha_2 \in (12.63, 12.97)$, $\alpha_3 \in (12.04, 15.49)$, $t_0 \in (0.48, 0.51)$, and $d_t \in (2.56, 2.63)$.

6. Summary and Conclusions

Chlamydia is a very difficult organism to propagate and determining the gene expression per chlamydial particle has been elusive until recently. We have used real-time PCR methodology to quantitatively measure gene transcript levels in *C. trachomatis* L2, obtaining 16S rRNA cDNA levels (a measure of the RNA)

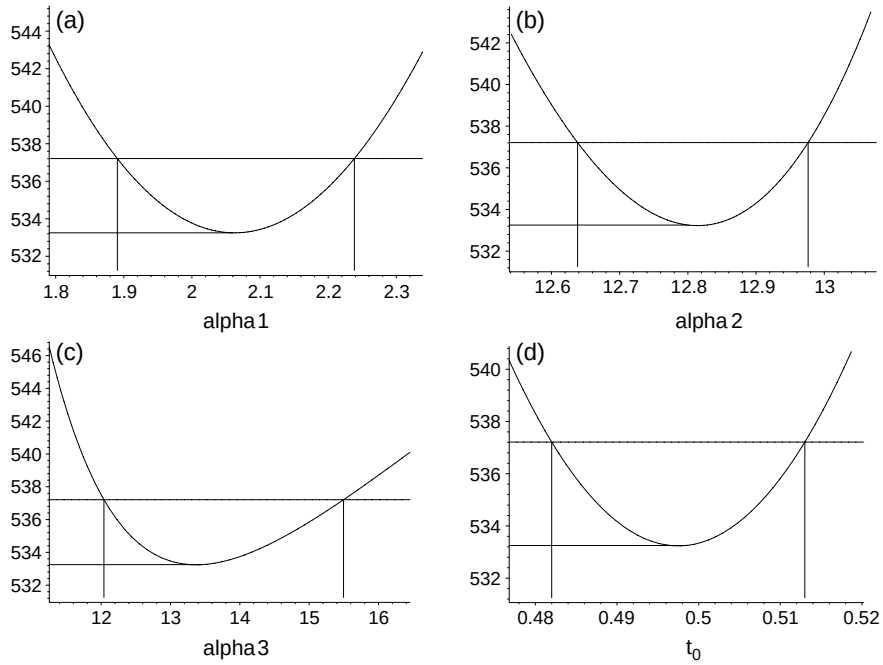

Figure 3. (a) $g(\alpha_1, \hat{\alpha_2}, \hat{\alpha_3}, \hat{t_0}, \hat{\sigma})$ versus α_1; (b) $g(\hat{\alpha_1}, \alpha_2, \hat{\alpha_3}, \hat{t_0}, \hat{\sigma})$ versus α_2; (c) $g(\hat{\alpha_1}, \hat{\alpha_2}, \alpha_3, \hat{t_0}, \hat{\sigma})$ versus α_3; (d) $g(\hat{\alpha_1}, \hat{\alpha_2}, \hat{\alpha_3}, t_0, \hat{\sigma})$ versus t_0. Here, the hat denotes the converged parameter value. The line segments denote the bounds on 95% confidence intervals.

for more developmental time-points than previously reported. 16S rRNA is an essential component of the protein synthesis in bacterial cells and the levels are a good measure of the number of chlamydial particles. We have developed a mathematical model of chlamydial intracellular replication and fitted it to the cDNA data. The model more naturally represents the processes within a chlamydial host cell inclusion than previous models of chlamydial development. We employed a numerical algorithm to obtain parameters of best fit. As a result, the cDNA data we obtained gives rise to our prediction of a mean RB doubling time of 2.59 with 95% confidence interval (2.56, 2.63). Our mathematical model allows the definition of specific stages of chlamydial development, monitoring progression of disease in a more precise manner than classical developmental cycle descriptions. Consequently, our modelling of the number of each type of chlamydial particle in the inclusion during development will also assist the understanding of immune response since chlamydial peptides presented to the cellular arm of the immune system depends on the number of chlamydial particles within each inclusion. In addition, an understanding of the population of chlamydial particles within the inclusion may allow development of more specific chemotherapeutic agents.

Potentially, our modelling can be extended to identify triggers causing EBs to commit to IBs converting to RBs as well as the reverse process. This modelling

would be an application to the emerging micro-array technology which has the potential to provide developmental expression for every chlamydial gene. The modelling of development will allow us to determine what genes are expressed at each specific stage of development, including the intermediate phases. We could potentially answer questions like 'expression of what gene signals the committment of an RB to become an EB?' which would reveal clearly defined targets for microbiologists in drug and vaccine development.

APPENDIX: EXPERIMENTAL SETUP AND METHOD OF DATA COLLECTION

The real-time PCR data used for this modelling was obtained using cDNA generated from *C. trachomatis* L2/434/Bu grown in HEp-2 cell monolayer cultures (as described in Mathews *et al*., 1999) consisting of 2.5×10^7 cells. Monolayer cultures were infected with *C. trachomatis* and triplicate samples (5×10^6 cells) harvested at 0, 2, 4, 6, 8, 10, 12, 14, 16, 20, 24, 28, 32 and 48 h PI by replacement of growth medium with 5 ml of Tri Reagent (Sigma) to allow for RNA isolation (according to the manufacturer's instructions). Genomic DNA was removed from the RNA by treatment with 20 U RNase-free DNase (Roche) and precipitation using 0.1 volume of 7.5 M ammonium-acetate and 2.5 volumes of 100% ethanol precipitation before three 75% ethanol washes. cDNA was generated using the 'Expand Reverse Transcriptase' kit (Roche) and 2 μg of random hexamers with 50 μg of total RNA (denatured at 75 °C for 10 min and quenced on ice), 1 U RNase inhibitor (Roche) and 1 U reverse transcriptase in a 12 μL reaction incubated at 42 °C for 1 h. RNA was removed from the cDNA with the addition of 1 U DNase-free RNase I (Roche) in 0.1 volume 0.2 M EDTA before precipitation with 0.1 volume 3.5 M sodium acetate and precipitated with three volumes of 100% ethanol prior to suspension in 100 μL of Tris-Buffer (pH 7).

The 16S rRNA PCR primers (ct16s-F and ct16sR) and DNA standards used for the real-time PCR are previously described (Mathews *et al*., 1999). The real-time PCR assays were done in triplicate for each developmental time-point using 50 ng of cDNA (or 10^8, 10^6, 10^4 and 10^2 copies of the 16S DNA standard), 1×Perkin-Elmer SYBR Green Mastermix (PE Biosystems) in a 15 μL reaction volume containing 1 μM of each primer. Amplification was done with the Rotorgene 2000 (Corbett Research, Sydney, Australia) using cycling parameters of: 1 cycle 95 °C for 3 min 45 cycles of 94 °C for 10 s/52 °C for 10 s/72 °C for 10 s with fluorescent acquisition at 74 °C. Melt curve parameters and data acquisition was set to range from 55 °C to 90 °C with a ramp rate of 10 °C/5 s. Quantitation and melt curve data analysis was performed on Rotorgene 2000 (Revision 4.4) software using melt curve analysis to verify correct target amplification. The copy number for each gene was compared to the copy number of 16S rRNA for the equivalent sample

and a relative level of transcripts to chlamydial particles calculated. The real-time PCR cDNA data is shown in Table A1.

Table A1. Experimental real-time PCR cDNA data for *Chlamydia trachomatis* L2.

Time (h PI)	cDNA level	Sample var	Time (h PI)	cDNA level	Sample var
0	1.38×10^7	1.35×10^{12}	14	1.80×10^8	3.18×10^{14}
2	2.64×10^7	3.88×10^{13}	16	1.51×10^8	3.24×10^{14}
4	2.52×10^7	1.44×10^{12}	20	5.78×10^8	6.55×10^{15}
6	1.50×10^7	1.54×10^{11}	24	2.01×10^9	1.43×10^{16}
8	2.21×10^7	1.08×10^{13}	28	2.31×10^9	5.76×10^{16}
10	5.69×10^7	1.60×10^{11}	32	6.15×10^9	3.73×10^{17}
12	5.72×10^7	3.10×10^{12}	48	5.98×10^9	3.63×10^{18}

REFERENCES

Cox, D. R. and D. V. Hinkley (1974). *Theoretical Statistics*, New York: Chapman and Hall.

Engel, J. N. and D. Ganem (1987). Chlamydial rRNA operons: gene organization and identification of putative tandem promoters. *J. Bacteriol.* **169**, 5678–5685.

Hackstadt, T., W. Baehr and Y. Ying (1991). *Chlamydia trachomatis* developmentally regulated protein is homologous to eukaryotic histone H1. *Proc. Natl. Acad. Sci. USA* **88**, 3937–3941.

Koehler, J. E., R. R. Burgess, N. E. Thompson and R. S. Stephens (1990). Chlamydia trachomatis RNA polymerase major sigma subunit: sequence and structural comparison of conserved and unique regions with Escherichia coli s^{70} and Bacillus subtilis s^{43}. *J. Biol. Chem.* **265**, 13206–13214.

Lundemose, A. G., S. Birkelund, P. M. Larson, S. J. Fry and G. Christiansen (1990). Characterization and identification of early proteins in Chlamydia trachomatis serovar L2 by two-dimensional gel electrophoresis. *Infect. Immunol.* **58**, 2478–2486.

Mathews, S. A., K. M. Volp and P. Timms (1999). Development of a quantitative gene expression assay for *Chlamydia trachomatis* identified temporal expression of σ factors. *FEBS Lett.* **458**, 354–358.

McClarty, G. (1994). Chlamydiae and the biochemistry of intracellular parasitism. *Trends Microbiol.* **2**, 157–164.

McSharry, P. E. and L. A. Smith (1999). Better nonlinear models from noisy data: attractors with maximum likelihood. *Phys. Rev. Lett.* **83**, 4285–4288.

Nicholson, T. L., L. Olinger, K. Chong, G. Schoolnik and R. S. Stephens (2003). Global stage-specific gene regulation during the developmental cycle of *Chlamydia trachomatis*. *J. Bacteriol.* **185**, 3179–3189.

Perara, E., D. Ganem and J. N. Engel (1992). A developmentally regulated chlamydial gene with apparent homology to eukaryotic histone H1. *Proc. Natl. Acad. Sci. USA* **89**, 2125–2129.

Rockey, D. D. and A. Matsumoto (1999). The chlamydial developmental cycle, in *Prokaryotic Development*, Y. V. Brun and L. J. Shimkets (Eds), Washington, DC: ASM Press, pp. 403–425.

Shaw, E. I., C. A. Dooley, E. R. Fischer, M. A. Scidmore, K. A. Fields and T. Hackstadt (2000). Three temporal classes of gene expression during the *Chlamydia trachomatis* developmental cycle. *Mol. Microbiol.* **37**, 913–925.

Timms, P. and S. Mathews (2002). Molecular biology of Chlamydia, in *Proceedings of the 10th International Symposium on Human Chlamydial Infections*, International Chlamydia Symposium, San Francisco, ISBN 0-9664383-1-0, pp. 585–594.

Ward, M. E. (2003). Type three secretion in chlamydiae. http://www.chlamydiae.com/docs/biology/biol_tts.htm.

Wilson, D. P., P. Timms and D. L. S. McElwain (2003). A mathematical model for the investigation of the Th1 immune response to *Chlamydia trachomatis*. *Math. Biosci.* **182**, 27–44.

Zhang, L., A. L. Douglas and T. P. Hatch (1998). Characterization of a *Chlamydia psittaci* DNA binding protein (EUO) synthesized during the early and middle phases of the developmental cycle. *Infect. Immunol.* **66**, 1167–1173.

Received 14 February 2003 and accepted 2 September 2003