dTTP-Induced R1 Dimerization: Predictions and Experimental Design

Tomas Radivoyevitch

Department of Epidemiology and Biostatistics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA

Email: txr24@case.edu

Summary

Of 46 models of dTTP-induced ribonucleotide reductase R1 dimerization fitted to data, the top 6 models differed from the best model by less than 3 corrected Akaike's Information Criterion units. Over a physiologically relevant grid of total [dTTP] and [R1], 5 of the top 6 models have similar reaction surfaces; the outlier assumes infinitely tight binding. It is likely that this outlier model will be rejected by future tests within this relevant grid, and one protein mass measurement alone at total concentrations of [R1] =.2 uM and [dTTP] = 1uM could prove or disprove this hypothesis. Once rejected, new experiments designed to discriminate between the remaining top 5 models would focus on the inclined portions of the surfaces as this is where these models differ most.

Model Space

Ribonucleotide reductase has a 90kDa large subunit R1 (R) that dimerizes when dTTP (t) binds to its specificity site [1-3]. The extent to which data in Fig. 1 of [1] and [2] yields estimates of the dissociation constants K of the total concentration constraint equilibrium system

$$0 = p[R_T] - [R] - \frac{[R][t]}{K_{Rt}} - 2\frac{[R]^2}{K_{RR}} - 2\frac{[R]^2[t]}{K_{RRt}} - 2\frac{[R]^2[t]^2}{K_{RRtt}}$$

$$0 = [t_T] - [t] - \frac{[R][t]}{K_{Rt}} - \frac{[R]^2[t]}{K_{RRt}} - 2\frac{[R]^2[t]^2}{K_{RRtt}}$$

(1)

where the subscript T denotes totals (system inputs) and the probability that an R molecule is undamaged and capable of dimerizing is p, has been described [4]. Briefly, models/hypotheses arise from these equations as follows. Firstly, $K = \infty$ assumptions are applied to remove specific terms one at time, two at a time, and so on, to yield 2^4=16 models, each hypothesizing that the deleted complexes are not detectable above noise. Secondly, of these models, the 4 single K models yield 4 additional models via K=0 assumptions, each alleging that the free concentration of the reactant that is not in excess is indistinguishable from zero. Thirdly, after expanding K into products of strictly binary K, models that allege that some Ks equal others also arise; these models correspond to hypotheses of independence between the R and t binding sites. Finally, for each model it can be hypothesized that the data are not rich enough to discriminate p close to one from p=1, and this expands the model space by an additional factor of two. The 40 spur graph model s and 6 grid graph models that result from this are summarized in the two dataframes printed below.

```
library(ccems)
topology <- list(
      heads=c("R1t0","R2t0"),
      sites=list(
            s=list(                        # s-site     thread #
                  m=c("R1t1"),             # monomer       1
```

```
                    d=c("R2t1","R2t2")  # dimer         2
                )
            )
)
g <- mkg(topology,TCC=TRUE)
sp <- mkSpurs(g,pRows=TRUE,doTights=TRUE)
print(sp$chunk)

gridL <- mkGrids(g,pRows=TRUE)
print(gridL$chunk)
```

The last 4 lines show the 46 models.   The intermediate g produced before them is passed the function ems to generate and fit up to the two parameter models of this model space in the following code.

```
data(RNR)
dd <- subset(RNR, (year==2006)|((year==2001)&(fg==1)&(G==0)&(t>0))
            , select=c(R,t,m,year))
names(dd)[1:2] <- paste(strsplit(g$id,split="")[[1]],"T",sep="")
# The next line produces Table 1 as an html file in the results folder
top <- ems(dd,g,pRows=TRUE,doTights=TRUE) # the output file is Rtglob.html
```

Current Best Models

As is readily seen form the summary that it prints to the screen:

```
 1 Model  17; nbp= 1; id=              IIIJ; AIC=-23.2896
 2 Model  19; nbp= 1; id=             III0p; AIC=-21.3520
 3 Model  20; nbp= 2; id=             IIIJp; AIC=-21.1582
 4 Model  31; nbp= 2; id=              IIJJ; AIC=-20.0722
 5 Model  29; nbp= 2; id=              IJIJ; AIC=-19.4775
 6 Model   2; nbp= 2; id=              HIEE; AIC=-19.2136
 7 Model  25; nbp= 2; id=              JIIJ; AIC=-19.0039
 8 Model  18; nbp= 0; id=              III0; AIC=-14.7873
 9 Model  13; nbp= 1; id=              IIJI; AIC=-12.3019
10 Model  16; nbp= 2; id=             IIJIp; AIC= -8.2696
```

the best model wins by a margin before it is met by two competing models. The output of these top 3 models given in the corresponding html output file is shown below as Table 1.

Table 1. HTML output file

Model	Parameter	Initial Value	Optimal Value	Confidence Interval
1 IIIJ.17	R2t2	1.000	2.685^3	(1.722^3, 4.193^3)
	SSE	0.062	0.032	
	AIC	-16.719	-23.290	
	Cpu	0.000	0.616	fit succeeded
2 III0p.19	P	1.000	0.826	(0.742, 0.919)
	SSE	0.104	0.039	
	AIC	-11.573	-21.352	

	Cpu	0.000	0.060	fit succeeded
3 IIIJp.20	R2t2	1.000	1.780^3	(0.729^3, 4.349^3)
	P	1.000	0.908	(0.794, 1.039)
	SSE	0.062	0.026	
	AIC	-12.433	-21.158	
	Cpu	0.000	0.276	fit succeeded

Physiologically Relevant Predictions of Current Best Models

The K_{RRtt} = 0 model III0p in Table 1 (black in the right panel of Fig. 1) differs substantially from the other two models (violet and orange) in its reaction surface predictions over a physiological grid of G_0 to S-phase concentrations defined by $[t_T]$ ε (.1, .2, .5, 1, 2, 5, 10, 20, 50) μM [5, 6] and $[R_T]$ ε (.005, .01, .02, .05, .1, .2, .5, 1) μM [7]. In part, this is because available data [1, 2] are outside this physiological grid of interest. The degree to which the other two models band together across the grid is seen on the right in Fig. 1. That they coalesce at small $[t_T]$ and $[R_T]$ to monomers can also be seen. Clearly, III0p (black) is an outlier. It is hypothesized here that this model will not withstand future experiments within this grid.
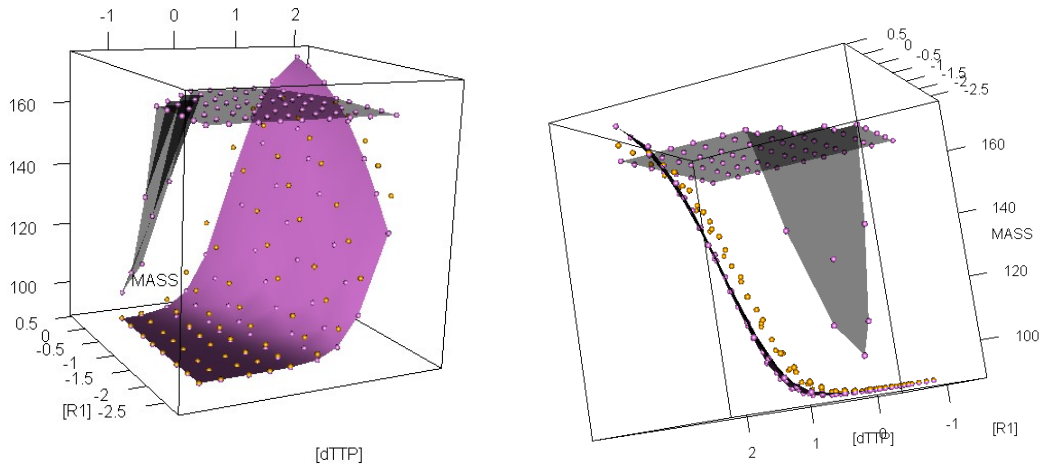

Figure 1. Top three models are IIIJ (violot), III0p (black) and IIIJp (orange). Units are kDa and $\log_{10}$ of μM.

```
library(ccems)    # the following code creates Figure 1
library(rgl)
load("results/Rtglob.RData") # top models were placed in results by ems
top=globalTopN
phR <- c(.005,.01,.02,.05,.1,.2,.5,1)
pht <- c(.1,.2,.5,1,2,5,10,20, 50)
physio=data.frame(RT=rep(phR,each=length(pht)),tT=rep(pht,length(phR)))
df=rbind(rbind(df1 <- simulateData(top[[1]],predict=physio)$predict,
               df2 <- simulateData(top[[2]],predict=physio)$predict),
                    simulateData(top[[3]],predict=physio)$predict)
n=dim(physio)[1]
plot3d(log10(df$tT),log10(df$RT),df$EY,col=c(rep("violet",n),rep("black ",n),
rep("orange",n)),
```

```
type="s",radius=c(rep(.5,3*n)),ylab="[R1]",xlab="[dTTP]",zlab="MASS")
material3d(alpha=1)
surface3d(log10(pht),log10(phR),matrix(df1$EY,ncol=length(pht)),col="violet")
material3d(alpha=.5)
surface3d(log10(pht),log10(phR),matrix(df2$EY,ncol=length(pht)),col="black")
# the snipping tool in vista was used to capture the images in Figure 1
```

Next Experiments

Gas-phase Electrophoretic Macromolecular Mobility Analysis (GEMMA; TSI Corp) is a measurement method that is sensitive enough to detect R1 dimer masses at low $[R_T]$ (≤ 1 μM) [2]. In Fig. 2 A shows the greatest difference between III0p and the other two models, suggesting that $[t_T]$=1 μM and $[R_T]$=.2 μM is a good next experiment. Similarly, C is best for subsequent experiments to discriminate between the remaining two models.
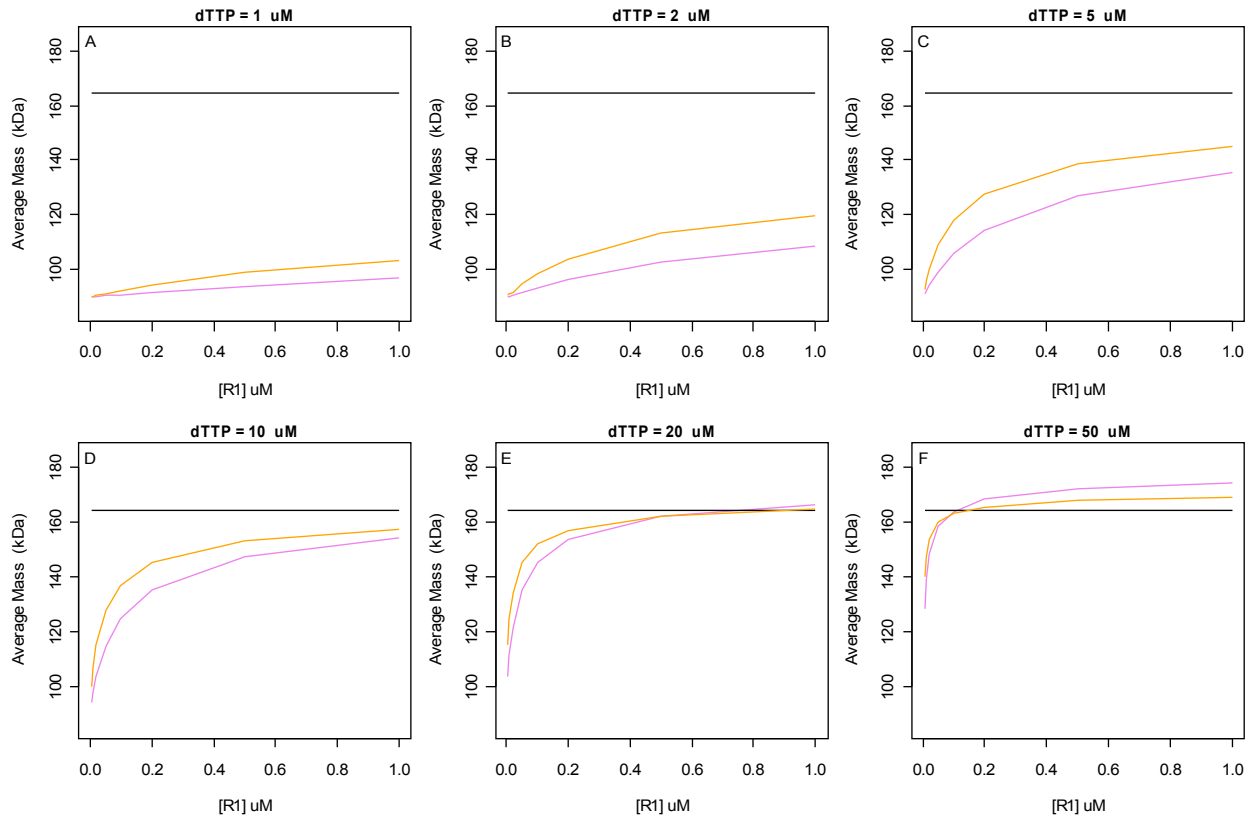

Figure 2. Predictions over the physiological grid.

```
windows(width = 9, height = 6);    # the following code creates Fig. 2
par(mfrow=c(2,3),mar=c(5.1,4.1,1.5,0.5))
dTTP= c(1,2,5,10,20, 50)
df=data.frame(df,mdl=rep(1:3,each=72))
for (i in 1:6){
    df1=subset(df,(tT==dTTP[i])&(mdl==1))
    plot(df1$RT,df1[,3],col="violet",type="l",main=paste("dTTP=",dTTP[i],"uM"),
    xlab="[R1] uM",ylab="Average Mass  (kDa)",ylim=c(85,185))
    text(0,184,LETTERS[i])
    df1=subset(df,(tT==dTTP[i])&(mdl==2))
    lines(df1$RT,df1$EY,col="black")
    df1=subset(df,(tT==dTTP[i])&(mdl==3))
    lines(df1$RT,df1$EY,col="orange")  }
```

Assuming a new measurement of 95 kDa at $[t_T]$ = 1 μM and $[R_T]$ = .2 μM and refitting, model prediction variances of the top five models evaluated at each of the 72 grid points, rank ordered from largest to lowest, yields, as an optimal experimental design, the best next 10 grid measurement points shown in Fig. 3.
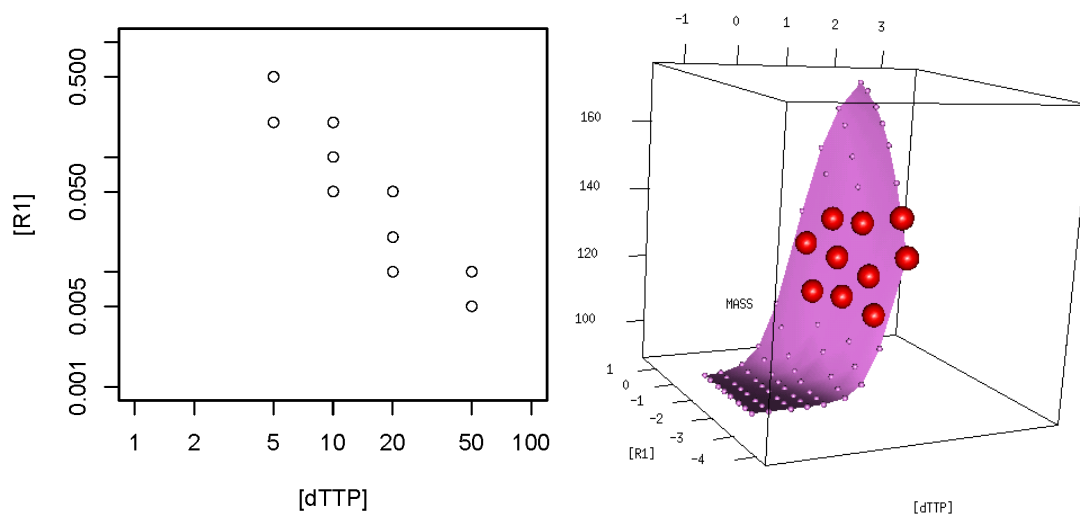

Figure 3. Best next 10 measurements. Units are μM (left) and kDa and $\log_{10}$ μM (right).

```
library(ccems)      # the following code creates Fig. 3
topology <- list(
        heads=c("R1t0","R2t0"),
        sites=list(
                s=list(                             # s-site      thread #
                        m=c("R1t1"),                # monomer         1
                        d=c("R2t1","R2t2")          # dimer           2
                )
        )
)
g <- mkg(topology,TCC=TRUE)
data(RNR)
dd <- subset(RNR, (year==2006)|((year==2001)&(fg==1)&(G==0)&(t>0))
                , select=c(R,t,m,year))
names(dd)[1:2] <- paste(strsplit(g$id,split="")[[1]],"T",sep="")
dd=rbind(dd,c(.2,1,95,2009))  # add in a fake data point
top <- ems(dd,g,pRows=TRUE,doTights=TRUE,topN=5)
phR<-c(.005,.01,.02,.05,.1,.2,.5,1)
pht<-c(.1,.2,.5,1,2,5,10,20, 50)
physio=data.frame(RT=rep(phR,each=length(pht)),tT=rep(pht,length(phR)))
for (i in 1:5)
   physio=cbind(physio,simulateData(top[[i]],predict=physio)$predict$EY)
names(physio)[3:7]=paste("m",1:5,sep="")
mat=as.matrix(physio[,3:7])
sig2=apply(mat,1,var)
physio=cbind(physio,sig2=sig2)
I=order(physio$sig2,decreasing=TRUE)
sdf=physio[I,]
b10=sdf[1:10,]
par(mfrow=c(1,1),mar=c(5.1,4.1,1.5,0.5))
plot(b10$tT,b10$RT,type="p",
        ylab="[R1]",xlab="[dTTP]",log="xy",ylim=c(.001,1),xlim=c(1,100))

library(rgl)
ndf=rbind(b10,physio) # to be big spheres and surface mesh points, respectively
```

```
plot3d(log10(ndf$tT),log10(ndf$RT),ndf$m1,col=c(rep("red",10),rep("violet",72)),
type="s",radius=c(rep(2,10),rep(.5,72)),ylab="[R1]",xlab="[dTTP]",zlab="MASS")
material3d(alpha=1)
surface3d(log10(pht),log10(phR),matrix(physio[,3],ncol=length(pht)),col="violet")
```

References
1.   Scott CP, Kashlan OB, Lear JD, Cooperman BS: **A quantitative model for allosteric control of purine reduction by murine ribonucleotide reductase**. *Biochemistry* 2001, **40**(6):1651-1661.
2.   Rofougaran R, Vodnala M, Hofer A: **Enzymatically active mammalian ribonucleotide reductase exists primarily as an alpha6beta2 octamer**. *J Biol Chem* 2006, **281**(38):27705-27711.
3.   Ingemarson R, Thelander L: **A kinetic study on the influence of nucleoside triphosphate effectors on subunit interaction in mouse ribonucleotide reductase**. *Biochemistry* 1996, **35**(26):8603-8609.
4.   Radivoyevitch T: **Equilibrium model selection: dTTP induced R1 dimerization**. *BMC Syst Biol* 2008, **2**(1):15.
5.   Jackson RC: **The Theoretical Foundations of Cancer Chemotherapy Introduced by Computer Models**. New York: Academic Press; 1992.
6.   Bradshaw PC, Samuels DC: **A computational model of mitochondrial deoxynucleotide metabolism and DNA replication**. *Am J Physiol Cell Physiol* 2005, **288**(5):C989-1002.
7.   Hakansson P, Hofer A, Thelander L: **Regulation of mammalian ribonucleotide reduction and dNTP pools after DNA damage and in resting cells**. *J Biol Chem* 2006, **281**(12):7834-7841.