Equilibrium Model Selection

Tomas Radivoyevitch, Department of Epidemiology and Biostatistics,
Case Western Reserve University, Cleveland, OH 44106 USA; txr24@case.edu

Ribonucleotide reductase (RNR) is precisely controlled to meet the dNTP demands of scheduled (replication driven) and unscheduled (repair driven) DNA synthesis. It has a small subunit R2 that exists almost exclusively as a dimer, and a large subunit R1 (R) that dimerizes when dTTP (t), dGTP, dATP, or ATP binds to its specificity site, and hexamerizes when dATP or ATP binds to its activity site. In general, RNR is modeled as a pre-equilibrium of proteins, ligands, and substrates whose parameters of interest are dissociation constants *K*, and a set of turnover rate parameters *k* that map distributions of active enzyme complexes into expected *k* measurements of mixtures. Because the masses of R1 and R2 are known, it is logical to focus first on *K* estimation from protein oligomer mass measurements, and later on *k* estimation from enzyme activity measurements. Further, it is also logical to begin with the simplicity of ligand-induced R1 dimerization.

The total concentration constraint full model for dTTP-induced R1 dimerization is

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

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

where the subscript T denotes totals (note that $[R]^2[t]^2/K_{RRtt} = [RRtt]$) and the probability that an R molecule is undamaged and capable of dimerizing is *p*. This full model generates 58 *a priori* plausible equilibrium models/hypotheses as follows. Firstly, $K=\infty$ assumptions are used 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*, nine additional models that allege that some *K*s equal others also arise; these nine models correspond to hypotheses of independence between the R and t binding sites on R. 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 to 58.

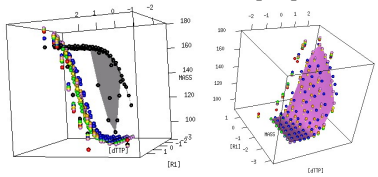
Using the following average mass output model

$$90\frac{[R]+[R_T](1-p)}{[R_T]} + 180\frac{2[RR]+2[RRt]+2[RRtt]}{[R_T]}$$

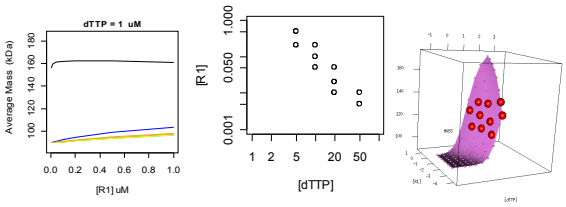
the 58 models were fitted to available data. The top 6 models based on AIC_c are

Model	Parameter	Initial value	Final value	CI
3M	RRtt	1.000	18.697	(4.807,72.966)
violet	p	1.000	1.000	fixed
	SSE	0.064	0.034	
	AIC_c	-48.066	-54.448	
3Mp	RRtt	1.000	5.558	(0.370,84)
blue	p	1.000	0.907	(0.787,1.044)
	SSE	0.064	0.027	
	AIC_c	-44.852	-53.308	
3Rp	p	1.000	0.822	(0.736,0.918)
black	RRtt	0.000	0.000	fixed
	SSE	0.106	0.041	
	AIC_c	-42.954	-52.590	
3I	RRt	1.000	49.568	(5.755,428)
green	RRtt	1.000	37.930	(5.003,290)
	p	1.000	1.000	fixed
	SSE	0.165	0.030	
	AIC_c	-35.303	-52.218	
2M	R_R	75.000	685.986	(2.801,162755)
yellow	RR_t	0.550	0.142	(0.005,3.975)
	RRt_t	0.550	0.142	constrained
	p	1.000	1.000	fixed
	SSE	0.041	0.032	
	AIC_c	-49.222	-51.815	
3F	Rt	1.000	91.059	(1.557,5324)
orange	RRtt	1.000	14.612	(2.545,84)
	p	1.000	1.000	fixed
	SSE	0.221	0.032	
	AIC_c	-32.422	-51.627	

Of these, model 3Rp differs substantially from the other five in its predictions over physiological values of $[t_T]$=.1 to 50 μM and $[R_T]$=.005 to 1 μM.



If 3Rp is rejected by a measurement of 95 kDa at $[t_T] = 1$ μM and $[R_T] = 0.2$ μM, the best next 10 measurements for discrimination between the remaining 5 models are the following points on the hill where their predictions differ most.



The methods, data, R functions and R scripts used to fit the model space can be found in

Radivoyevitch T: Equilibrium model selection: dTTP induced R1 dimerization. *BMC Systems Biology* 2008, **2** (1):15.