

## Max Mass Objective

The MaxMass algorithm contains the following objective function:

$$\max \sum_{k \in K} y_k \sum_{i=1}^k m_i - \sum_{k \in K} x_k \sum_{i=1}^k m_i$$

where the variables are

$$x_k = \begin{cases} 1, & \text{if gene or promoter } k \text{ is the first gene or promoter within the deleted segment} \\ 0, & \text{otherwise} \end{cases}$$

and

$$y_k = \begin{cases} 1, & \text{if gene or promoter } k \text{ is immediately after the end of the deleted segment} \\ 0, & \text{otherwise} \end{cases}$$

where

$$\sum_{k \in K} x_k = 1$$

and

$$\sum_{k \in K} y_k = 1$$

and the parameters are

- $d_k$ : Position of the first nucleotide of the deleted sequence starting from the origin of replication when gene or promoter  $k$  is selected to be deleted in the beginning of the stretch. Note that  $d_k$  is not always the start site of a gene or promoter. It is the first nucleotide of the nonoverlapped region between the gene/promoter  $k$  and gene/promoter  $k - 1$
- $d'_k$ : Position of the first nucleotide of the gene or promoter  $k$  immediately after the deleted sequence
- $m_i$  is the measured protein mass of gene  $i$  and  $\sum_{i=1}^k m_i$  is the cumulative protein mass of all genes from the origin of replication to the  $k$ th gene. By subtracting the cumulative protein mass of the start gene ( $x_k$ ) from the cumulative protein mass of the end gene ( $y_k$ ), we obtain the cumulative protein mass of the interval for the optimal solution

We maximize the mass that is knocked out based on absolute quantitative proteomics [Schmidt et al 2015](http://www.nature.com/articles/nbt.3418) (<http://www.nature.com/articles/nbt.3418>) in (fg/cell)

## MaxMass constraints

$\sum_{j \in J} S_{i,j} v_j = 0$	$i = 1, \dots, N$	Steady state conservation of metabolite $i$ requirement
$\sum_{k \in K} y_k = 1$	$\sum_{k \in K} x_k = 1$	Ensure only one gene knockout per iteration
$\sum_{j=1}^k x_j - \sum_{j=1}^k y_j = z_k$	$k = 1, \dots, K$	Ensures all genes between start and end area also knocked out
$v_{\text{biomass}} \geq f \cdot v_{\text{biomass, max}}$		Requires that the KO must be capable of growing at the specif
$z_g = \prod_{p \in p^s} z_p$	$p = 1, \dots, P$	Ensures that a gene is not functional if all of its promoters are

## Gene Protein Reaction constraints

**Single gene catalyzes reaction:**  $k \rightarrow v_j$

$$(1 - z_k) \cdot LB \leq v_j \leq (1 - z_k) \cdot UB$$

**Enzyme dimer complex catalyzes reaction:**  $(k_1 \text{ AND } k_2) \rightarrow v_j$

$$(1 - z_{k_1}) \cdot LB \leq v_j \leq (1 - z_{k_1}) \cdot UB$$

$$(1 - z_{k_2}) \cdot LB \leq v_j \leq (1 - z_{k_2}) \cdot UB$$

**Two Isozymes catalyze reaction:**  $(k_1 \text{ OR } k_2) \rightarrow v_j$

$$(2 - z_{k_1} - z_{k_2}) \cdot LB \leq v_j \leq (2 - z_{k_1} - z_{k_2}) \cdot UB$$

$$LB \leq v_j \leq UB$$

**Complex Gene Protein Reaction:**  $(k_1 \text{ AND } k_2) \text{ OR } (k_1 \text{ AND } k_3) \rightarrow v_j$

$$(2 - z_{k_2} - z_{k_3}) \cdot LB \leq v_j \leq (2 - z_{k_2} - z_{k_3}) \cdot UB$$

$$(1 - z_{k_1}) \cdot LB \leq v_j \leq (1 - z_{k_1}) \cdot UB$$