MINVAL - MINimal VALidation for Stoichiometric Reactions

Daniel Osorio, Janneth Gonzalez PhD. and Andrés Pinzón-Velasco PhD.

2016-10-11

Introduction to MINVAL

The **MINVAL** package was designed as a tool to identify orphan metabolites and evaluate the mass and charge balance of stoichometric reactions. MINVAL also includes functions to write models in TSV and SBML formats, extract all reactants, products, metabolite names and compartments from a metabolic reconstruction.

Getting started

```
library(minval)
library(sybilSBML)
```

```
glycolysis <- read.csv2(system.file("extdata", "glycolysisKEGG.csv", package = "minval"),
                        stringsAsFactors = FALSE)
colnames(glycolysis)
```

```
## [1] "ID"          "DESCRIPTION" "REACTION"    "GPR"         "LOWER.BOUND"
## [6] "UPPER.BOUND" "OBJECTIVE"
```

```
head(glycolysis$REACTION)
```

```
## [1] "NADH[c] <=> NADH[b]"
## [2] "H+[c] <=> H+[b]"
## [3] "Pyruvate[c] <=> Pyruvate[b]"
## [4] "alpha-D-Glucose[c] <=> alpha-D-Glucose[b]"
## [5] "NAD+[c] <=> NAD+[b]"
## [6] "Orthophosphate[c] <=> Orthophosphate[b]"
```

```
valid <- isValidSyntax(glycolysis$REACTION)
```

```
## Warning: Reaction 12: Invalid coefficients. Metabolites should have just
## one coefficient.
```

```
## Warning: Reaction 16: Invalid directionality symbols. Arrow symbols should
## be between blank spaces.
```

```
## Warning: Reaction 16: Invalid directionality symbols. Please use <=> or =>
## instead of <-> or -> or -->.
```

```
valid
```

```
##  [1]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
## [12] FALSE  TRUE  TRUE  TRUE FALSE  TRUE  TRUE  TRUE
```

```
mapReactions(reactionList = valid,
             referenceData = glycolysis,
             by = "bool",
             inverse = TRUE)
```

```
##        ID                                          DESCRIPTION
## 12 R01015 D-glyceraldehyde-3-phosphate aldose-ketose-isomerase
## 16 R01518                D-phosphoglycerate 2,3-phosphomutase
##                                                       REACTION
## 12 2 2 D-Glyceraldehyde 3-phosphate[c] <=> Glycerone phosphate[c]
## 16         2-Phospho-D-glycerate[c] <--> 3-Phospho-D-glycerate[c]
##                  GPR LOWER.BOUND UPPER.BOUND OBJECTIVE
## 12 ( 5232 and 5230 )       -1000        1000         0
## 16              7176       -1000        1000         0
```

```
glycolysis$REACTION[12] <- "D-Glyceraldehyde 3-phosphate[c] <=> Glycerone phosphate[c]"
glycolysis$REACTION[16] <- "2-Phospho-D-glycerate[c] <=> 3-Phospho-D-glycerate[c]"
isValidSyntax(glycolysis$REACTION)
```

```
##  [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [15] TRUE TRUE TRUE TRUE TRUE
```

```
convert2sbml(glycolysis,"glycolysis.xml")
glycoModel <- readSBMLmod("glycolysis.xml")
glycoModel
```

```
## model name:             model
## number of compartments  2
##                         c
##                         b
## number of reactions:    19
## number of metabolites:  18
## number of unique genes: 26
## objective function:     +1 R00200
```

```
optimizeProb(glycoModel)
```

```
## solver:                                  glpkAPI
## method:                                  simplex
## algorithm:                               fba
## number of variables:                     19
## number of constraints:                   18
## return value of solver:                  solution process was successful
## solution status:                         solution is optimal
## value of objective function (fba):       10.000000
## value of objective function (model):     10.000000
```

```
chemicalData <- read.csv2(system.file("extdata", "chemData.csv", package = "minval"))
head(chemicalData)
```

```
##                             NAME       FORMULA     MASS CHARGE
## 1                            H2O           H2O  18.0106      0
## 2                             H+             H   1.0078      1
## 3                            ATP C10H16N5O13P3 506.9957      0
## 4                           NAD+ C21H28N7O14P2 664.1169      1
## 5 3-Phospho-D-glyceroyl phosphate     C3H8O10P2 265.9593      0
## 6     beta-D-Fructose 6-phosphate      C6H13O9P 260.0297      0
```

```
balanced <- isBalanced(reactionList = glycolysis$REACTION,
                       referenceData = chemicalData,
                       ids = "NAME",
                       mFormula = "FORMULA")
```

```
balanced
```

```
##  [1]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
## [12]  TRUE FALSE  TRUE  TRUE  TRUE FALSE  TRUE  TRUE
```

```
mapReactions(reactionList = balanced,
             referenceData = glycolysis,
             by = "bool",
             inverse = TRUE)
```

```
##        ID
## 13 R01061
## 17 R01786
##                                                          DESCRIPTION
## 13 D-glyceraldehyde-3-phosphate:NAD+ oxidoreductase (phosphorylating)
## 17                            ATP:alpha-D-glucose 6-phosphotransferase
##
## 13 D-Glyceraldehyde 3-phosphate[c] + Orthophosphate[c] + NAD+[c] => 3-Phospho-D-glyceroyl phosphate[
## 17                                   alpha-D-Glucose[c] + Orthophosphate[c] => 5 alpha-D-Glucose 6-
##                         GPR LOWER.BOUND UPPER.BOUND OBJECTIVE
## 13 ( 230 and 226 and 229 )           0        1000         0
## 17   ( 2597 ) or ( 26330 )           0        1000         0
```

```
glycolysis$REACTION[13] <- "D-Glyceraldehyde 3-phosphate[c] + Orthophosphate[c] + NAD+[c] => 3-Phospho-
glycolysis$REACTION[17] <- "alpha-D-Glucose[c] + Orthophosphate[c] => alpha-D-Glucose 6-phosphate[c] + 
isBalanced(reactionList = glycolysis$REACTION,
                       referenceData = chemicalData,
                       ids = "NAME",
                       mFormula = "FORMULA")
```

```
##  [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [15] TRUE TRUE TRUE TRUE TRUE
```

```
convert2sbml(glycolysis,"glycolysisCurated.xml")
glycoModel <- readSBMLmod("glycolysisCurated.xml")
glycoModel
```

```
## model name:             model
## number of compartments  2
##                         c
##                         b
## number of reactions:    19
## number of metabolites:  18
## number of unique genes: 26
## objective function:     +1 R00200
```

```
optimizeProb(glycoModel)
```

```
## solver:                                  glpkAPI
## method:                                  simplex
## algorithm:                               fba
## number of variables:                     19
## number of constraints:                   18
## return value of solver:                  solution process was successful
## solution status:                         solution is optimal
## value of objective function (fba):       2.000000
## value of objective function (model):     2.000000
```