Model-driven multi-omic data analysis elucidates metabolic immunomodulators of macrophage activation
Bordbar et al.

Supplemental Table of Contents

Text:

Figures:

Tables:

References

Methods:

1. *Omics-guided reconstruction and functional validation of RAW 264.7 network.*

A murine macrophage (RAW 264.7) metabolic network was reconstructed by mapping gene expression and proteomic data to determine metabolic reaction activities. A full workflow is presented in Figure S1. First, we used a NCBI HomoloGene-based mapping to the global human metabolic reconstruction, Recon 1 (Duarte et al. 2007), as the basal framework for constructing a murine RAW 264.7 cell line metabolic network. A similar approach was taken for a recently published homolog-based mouse reconstruction (Sigurdsson et al. 2010). Given the ~92% homology between the human and mouse genomes, this network approach was used in lieu of a previously reconstructed mouse metabolic network (Sheikh et al. 2005) as the Homologene-based mapping provided: (1) higher genomic coverage, (2) more extensive cellular compartmentalization of reactions, and (3) readily available gene-protein-reaction (GPR) associations for high-throughput data mapping.

We utilized two established algorithms (GIMME and iMAT) and a novel algorithm (GIMMEp) to construct the RAW 264.7 cell line network from high-throughput data. The GIMME algorithm was improved upon its implementation into a transcripto-proteomic approach. GIMMEp constructs separate sub-networks that each satisfies one of the proteome-defined objective functions of the RAW macrophage cell (see Supplementary Data for the list of 382 proteome-mapped metabolic reactions). Objective functions pertaining to biomass growth and nitric oxide synthase activities were also evaluated. The list of reactions (and associated genes) in the separate predicted sub-networks was then combined to generate a full murine macrophage-specific metabolic networkThe "GIMMEp" approach is very similar to a recent tissue building algorithm (MBA (Jerby et al. 2010)). There are however two notable differences, 1) GIMMEp chooses the more biologically active path to link core reactions rather than the shortest path, and 2) GIMMEp does not remove flux carrying reactions that do not have gene association resulting in a larger model than MBA. We also implemented a traditional version of GIMME, optimizing for macrophage biomass (GimmeBM). Exchange inputs were set at arbitrary values of -1 for metabolites defined in DMEM media. Finally, we implemented iMat (Shlomi et al. 2008) to predict a macrophage-specific network based on the flux distribution most consistent with the omics data.

We first compared the content of the metabolic networks to the gene expression and proteome experimental data to verify the accuracy of the predicted networks (Table S1). The GIMMEp metabolic network contained 809 genes, 1026 intracellular reactions, and 552 unique metabolites. The predicted subnetwork that satisfied only biomass growth (GimmeBM) was substantially smaller, containing 446 genes, 424 intracellular reactions, and 304 unique metabolites, and resulted in a substantially lower recall of genes and protein-associated reactions. This indicates that a single objective function was not sufficient in predicting a comprehensive metabolic network using the GIMME approach and additional tested objective functions improved the comprehensiveness of the data-based predictions. The iMat sub-network contains 938 genes, 1,327 intracellular reactions, and 714 unique metabolites and yielded a higher recall in the number of activated genes, an expected result given that the algorithm's objective is to achieve a sub-network optimally correlated with the gene expression data. The reaction recall,

however, was comparable between the two approaches with both accounting for ~73% of the metabolic reactions associated with the RAW 264.7 macrophage proteomic data (Table S1).

The physiological capabilities of the GIMMEp and iMat metabolic networks were also evaluated using nutrient uptake rates derived from *in vitro* experimental data sets for murine macrophages (see Materials and Methods in main text). In order to test the metabolic functionality of the mouse macrophage model, we compared the ability of the network to perform the published 288 metabolic functions of Recon 1 (Duarte et al. 2007). To characterize the metabolic functionality of the three algorithm-derived networks as well as the reconciled model, the ability of the networks to complete the 288 metabolic functions in Recon 1 were assessed. For each test, sink reactions were added for the metabolites in question, and FBA was used to determine whether a feasible non-zero solution could be attained. The tests were performed by utilizing the exchange constraints set from literature and modeling requirements.

The three networks, GimmeBM, GIMMEp, and iMat, performed quite differently (Table S1). GimmeBM's performance represented the basal functionality of the three metabolic models, as the sole objective of the algorithm was to produce a biomass flux (Table S1). The GIMMEp and iMat draft reconstructions have a more comprehensive functionality as shown by passing the majority of the tests. The GIMMEp and iMat models have similar functionality across all subsystems except for lipid metabolism due to the iMat's lack of ability to produce triglycerides and certain glycerophospholipids. The iMat model was also unable generate biomass. The GIMME algorithm conserves functions when mapping expression data while the iMat algorithm conserves reaction numbers, thus it was expected that the GIMMEp reconstruction would have a higher functionality.

The GIMMEp and iMat models were reconciled by manually curating discrepant reactions (based on discrepancies in metabolic functionalities) in the two draft reconstructions, with the resulting reconciled model having more metabolic functionality than either model (Table S1 and Figure S2). The reconciled network was more similar to GIMMEp than iMat in reaction and metabolite number as fewer active reactions from the iMat model were implicated in differential metabolic functionality. The full RAW 264.7 macrophage metabolic model is provided in the supplementary material.

2. *Cell culture and stimulation*

RAW 264.7 (ATTC) murine macrophage-like cell line was grown at 37 ºC and 5% CO_2 atmosphere in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum and Penicillin (100 U/mL)-Streptomycin (100 µg/mL). The samples were prepared in two biological replicates. For each biological replicate, cells were grown in parallel with independent set of media. Cells were seeded at a density of 1.0e6 cells/well in 6-wells plates, grown overnight and stimulated 0, 1, 2, 4 and 24 hours with 100 ng/mL LPS, a TLR4 agonist (from *Salmonella typhimurium*, cat # L6143, Sigma-Aldrich) diluted in fresh medium. After stimulation, cells were washed twice with Dulbecco's PBS, and either treated with RNA Later (Qiagen) for transcriptomic analysis or lysed with 100 mM NH_4HCO_3, 1% CHAPS and 0.1% SDS for proteomic analysis. For each sample, the content of two wells were pooled together to obtain enough material for analysis.

3. *Sample preparation for proteomic analysis*

Cell extracts were sonicated three times for 30 s each, 100% amplitude and 0.5 pulse (UTR200, Hielscher). After sonication, extracts were centrifuged for 20 min at 4 ºC and 16,000 x*g*, and quantified by BCA assay according to the manufacturer recommendations (Thermo Scientific). The supernatants were collected and precipitated for 20 min on an ice bath. The samples were centrifuged (20 min at 4 ºC and 16,000 x*g*), the protein pellets were washed with cold acetone and centrifuged 5 min at 4 ºC and 16,000 x*g*. The resulting pellets were dissolved in a concentration of 5 mg/mL in 100 mM NH_4HCO_3, 7 M urea, 2 M thiourea and 5 mM DTT. After reducing for 15 min at 50 ºC, the cysteine residues were alkylated for 30 min at room temperature in the presence of 10 mM iodoacetamide. The solutions were 5-fold diluted with 25 mM NH_4HCO_3 and the proteins were digested overnight with sequencing-grade trypsin (1/50 enzyme to substrate ratio). The pH of the samples were adjusted to ~2.5 and peptides were purified using 1 mL 100 mg strong cation-exchange (SCX) SPE cartridges using an automated station (GX-274, Gilson). The cartridges were washed with methanol, 10 mM ammonium formate (pH 3.0)/25% acetonitrile (ACN), 500 mM ammonium formate (pH 6.8)/25% ACN and water, before being equilibrated with 10 mM ammonium formate (pH 3.0)/25% ACN. After loading the samples, cartridges were washed with with 10 mM ammonium formate (pH 3.0)/25% ACN, and the peptides were eluted with methanol:water:NH_4OH (80:15:5, v:v:v) and concentrated in the speed vac. Resulting peptides were quantified by BCA assay.

4. *Liquid chromatography-tandem mass spectrometry analysis*

Digested peptide aliquots from cells stimulated with LPS were pooled by stimulus into two samples and each was separated in 24 fractions by high pH reverse chromatography, essentially as described by Wang et al (Wang et al. 2011). Each fraction or unfractionated sample was submitted to liquid chromatography-tandem mass spectrometry analysis. Peptides were loaded into capillary columns (75 µm x 65 cm, Polymicro) packed with C18 beads (3 µm particles, Phenomenex) connected to a custom-made 4-column LC system (Livesay et al. 2008). The elution was performed in an exponential gradient from 0-100% B solvent (solvent A: 0.1% FA; solvent B: 90% ACN/0.1% FA) in 100 min with a constant pressure of 10,000 psi and flow rate of approximately 400 nL/min. Eluting peptides were directly analyzed either on a linear ion-trap (LTQ XL, Thermo Scientific, San Jose, CA) (fractionated samples) or an orbitrap (LTQ Orbitrap XL, Thermo Scientific) (unfractionated samples, run in technical duplicates) mass spectrometer using chemically etched nanospray emitters (Kelly et al. 2006). Full scan mass spectra were collected at 400-2000 m/z range and the top ten (linear trap) or six (orbitrap) most intense ions were submitted to low-resolution CID fragmentation once (35% normalized collision energy), before being dynamically excluded for 60 s.

5. *Data analysis*

Tandem mass spectra were searched with SEQUEST (v27.12) (Yates et al. 1995) against mouse IPI database (v3.52) and common contaminant sequences (all in forward and reversed orientations, total of 111,010 searched sequences), using the following parameters: (i) partial tryptic digestion, (ii) 3 Da of parent mass tolerance for linear ion trap data and 0.1 Da for

orbitrap data, (iii) 0.5 Da for fragment mass tolerance, and (iv) cysteine carbamidomethylation and methione oxidation as fixed and variable modifications, respectively. Quantification was done with the accurate mass and time (AMT) tag (Zimmer et al. 2006). SEQUEST results were filtered with Xcorr $\geq$ 1.9, 2.2 and 3.75 (for singly-, doubly- and triply-charged peptides, respectively); DCn $\geq$ 0.08; a minimum length of 6 amino acid residues; and Peptide Prophet (Keller et al. 2002) $\geq$ 0.5 to build a mass tag (MT) database. High resolution orbitrap runs were used to peak match against the MT database, using the VIPER tool (Monroe et al. 2007). To ensure the quality of peptide matching, all peptides matched to the MT database were filtered with a Statistical Tools for AMT tag Confidence (STAC), an updated version of SMART (Jaitly et al. 2008) score $\geq$ 0.7 and uniqueness probability $\geq$ 0.5. Additionally, peptides were required be present in more than half of the replicates in at least one sample, and proteins were required to have at least 2 peptides and at least one peptide with STAC $\geq$ 0.9 and MS-GF (Jaitly et al. 2008) $\leq$ 1.0e-10. Peptide abundance values were rolled-up into proteins using Qrollup tool, available in DAnTE (Kim et al. 2008). Changes in the abundances comparing treatment and control, and their p-values (by ANOVA test) were calculated using DAnTE.

6. *Microarray Data Generation and Analysis*

Amino-allyl cDNA (aa-cDNA) were generated by adding 2 g of total RNA in a mixture containing 6 g of random hexamers (Invitrogen), 0.01 M dithiothreitol, an aminoallyl-deoxynucleoside triphosphate mixture containing 25 mM each dATP, dCTP, dGTP, and dTTP, reaction buffer, and 400 units of SuperScript III reverse transcriptase (Invitrogen) at 42 °C overnight. The RNA template was then hydrolyzed by adding NaOH and EDTA to final concentrations of 0.2 and 0.1 M, respectively, and incubated at 70 °C for 15 min. Unincorporated aa-dUTP was removed with a Minelute column (Qiagen). The aa-cDNA were eluted with a phosphate elution buffer (4 mM KPO_4, pH 8.5, in ultrapure water), dried, and resuspended in 0.1 M sodium carbonate buffer (pH 9.0). To couple the aa-cDNA with fluorescent labels, aa-cDNA was incubated at room temperature for 1 h with normal human serum-Cy3 or normal human serum-Cy5 (Amersham) for the control or treatment aa-cDNA, respectively; and, the uncoupled dye was removed using the Qiagen Minelute column (Valencia, CA). A mixture of the labeled cDNA for the treatment and the control were hybridized to Agilent Mouse GE 4x44K v2 Microarray (Agilent Technologies) and processed as per manufacture's protocol. Slides were scanned with an Axon GenePix 4000 scanner. Image analysis and intra-chip normalization were performed with Feature Extraction 9.5.3.1 (Agilent). Probes that were saturated, non-uniform, not found, or manually flagged were excluded from further analyses. Probe Ids were mapped to Entrez Gene IDs, and average error-weighted log2 fold-changes were calculated when more than one probe matched to single Entrez Gene ID; combined p-values were calculated using the weighted Z-method (Whitlock 2005). Or expression values were processed with MeV (http://tm4.org/).

7. *Reporter Metabolite Analysis*

Reporter metabolite calculations were performed in a python (http://www.python.org) implementation of the COBRA Toolbox (http://opencobra.sourceforge.net) using SciPy's (http://www.scipy.org) statistical modules. Reporter metabolite calculations (Patil and Nielsen 2005) were performed by converting the network into a bipartate graph where the nodes were the

metabolites and the reactions the edges. From the transcriptome data, for each reaction a combined p-value was calculated from the gene-products comprising the enzyme catalyzing the reaction using the weighted Z-method; in the case that a gene product was not detected a default p-value of 1 was assumed. Then each reaction's p-value was converted into a Z-score using the inverse normal cumulative distribution. Next, based on metabolic network topology, a size-independent aggregate Z-score was calculated for each metabolite from the Z-scores of the surrounding genes. For transcriptome data, the metabolite Z-scores were then background corrected and metabolites were ranked by Z-score – the larger the Z-score the more activity associated with the metabolite. Background correction was not applied for calculations performed with proteome data due to the typical coverage being less than 30% of the network.

8. *Metabolomics preparation and analysis*

For the metabolomics experiment, cells were seeded at a density of 3.0×10^6 cells in 150-mm dishes, grown for 2 days and stimulated for 24 hours with 100 ng/mL LPS diluted in fresh medium. A control culture was run in parallel by incubating for the same period of time with fresh medium only. Two dishes were used for each of the biological triplicates. After stimulation, cells were washed twice with Dulbecco's PBS, scraped out and harvested into 15-mL centrifuge tubes. Cell suspensions were softly centrifuged (230 ×*g* for 5 min) and the most as possible of the buffer was removed. Then, 170 µL of 150 mM ammonium bicarbonate was added to the cell pellet and the cell suspensions were transferred to 2.0-mL micro-centrifuge tubes for extraction. Subsequently, the water soluble metabolites were extracted with four volume of chilled (-20 °C) chloroform/methanol mixture (2:1, v:v). After the vigorously vortexing, the samples were centrifuged (12,000 ×*g* for 5 min) and upper (aqueous) layer containing water soluble metabolites were transferred into glass vials, followed by drying in a vacuum concentrator (SpeedVac; Thermo Scientific, Waltham, MA).

Extracted metabolites were successively derivatized by two chemical reagents to enhance their stability and volatility for the gas chromatography-mass spectrometry (GC-MS) analysis (Kim et al. 2011). Each sample was initially derivatized with 30 µL of 30 mg/mL methoxyamine (cat # 33045, Sigma-Aldrich) in pyridine at 37°C with shaking for 90 min. This step protects carbonyl groups as well as reduces the number of tautomeric peaks. Next, 120 µL of *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS) (cat # 69478, Sigma-Aldrich) was added to each vial, followed by incubation at 37°C with shaking for 30 min to modify hydroxyl and amine groups into trimethylsilyated (TMS) forms. The samples were then allowed to cool down to room temperature and analyzed by GC-MS.

Samples were analyzed in technical duplicates in a GC-MS system (Agilent GC 7890A coupled with a single quadrupole MSD 5975C; Agilent Technologies, Santa Clara, CA) connected to HP-5MS column (30 m × 0.25 mm × 0.25 µm; Agilent Technologies). One microliter of each sample was injected in splitless mode and with a port temperature held at 250°C throughout analysis. The GC oven was initially maintained at 60°C for 1 min and then ramped to 325°C by 10°C/min and kept at this final temperature for 5 min (Kind et al. 2009). After format conversion to netCDF, generated GC-MS raw data files were processed by Metabolite Detector for peak deconvolution (Hiller et al. 2009). Retention indices (RI) were calculated based on the pre-analysis of a mixture of FAMEs (C8 - C30) (Agilent Technologies) as external retention time standards, then their information was subsequently applied to all chromatograms for retention

time alignment. Deconvoluted features were identified by matching to the Fiehn Metabolomics library (Kind et al. 2009) which contains mass spectrum and RI information of approximately 700 metabolites (more than a thousand spectra). The quantification was done by feature abundances and was subjected to a statistical analysis (*t*-test, p-value cut off < 0.05) to determine the significant metabolic changes (Polpitiya et al. 2008).

9. Comparison of RAW 264.7 metabolic network to global Recon 1

To validate the RAW 264.7 metabolic network and show its functional accuracy improvement over Recon 1, we repeated calculated flux rates and sensitivity analysis for Recon 1. We applied the same set of *in vitro* uptake exchange rates on the Recon 1 network as was applied on the RAW 264.7 network. We then calculated the biomass growth rate, ATP production, and NO production (Table S2). Growth rate with Recon 1 was ~1.5x higher than the cell-specific network but was within the physiological range. However, the ATP and NO production rates were substantially higher than the measured experimental rates. The RAW 264.7specific network achieved ATP and NO production rates within 10% of the actual experimentally measured rates. Recon 1's production rates for ATP and NO production, however, were > 3x the experimental rates. A cell-specific model is essential for accuracy as the full non-cell-specific model contains additional reactions that add biologically infeasible functionalities that overestimate production rates.

To further test the difference between the generic and cell-specific networks, we re-calculated the sensitivity results using Recon 1. The results were quite different from the RAW cell network (Figure S4). Generally, the effect of metabolites was weaker on the macrophage activation objectives. In addition, some of the Recon 1 results did not seem biologically accurate. For example, arginine uptake was not activating of nitric oxide production. Glutamine had little effect on any of the macrophage objectives. Glucose and oxygen uptake was much less effective in activating the macrophage functions than previously found. Phosphate was the most suppressive substrate. For the products, we found no metabolite that when produced would activate the macrophage functions. Previously, the RAW network identified nitrogen products as activating (urea, glutamate, ammonia). The sensitivity results obtained from Recon 1 are quite different from the RAW 264.7 cell line specific network and in many cases are inaccurate. The cell-specific network is crucial for accurate modeling of macrophage metabolism.

10. Comparison of in silico sampling predictions and omics data

To validate the RAW 264.7 network's flux sampling predictions, we compared our significantly changed fluxes with three high-throughput data types. Metabolite levels cannot be compared directly as the network only predicts flux changes. The metabolite changes were however consistent from a qualitative comparison (see Main Text). However, the transcriptomics and proteomics analyses can be compared systematically with *in silico* predictions.

To do so, we first determined the significantly changed metabolic reactions during tryptophan and nucleotide uptake versus no uptake across the five objective functions through a Wilcoxon signed-ranked test ($p < 0.05$) for 10,000 randomized fluxes. To avoid bias by the objective function, we required that a reaction be significantly changed across at least three objectives to

be considered significantly changed. The predicted significantly changed reactions were compared to the generated transcriptomic and proteomic data. Transcripts and proteins that were significantly changed ($p < 0.05$, multiple hypotheses t-test) and had an absolute log2 fold change of at least 0.5 were compared to *in silico* predictions.

There were 134 transcript comparisons and 55 protein comparisons across all time points (0h, 2h, 4h, 24h). The accordance of the predictions and omics data was statistically significant ($p < 1e-5$, permutation test).

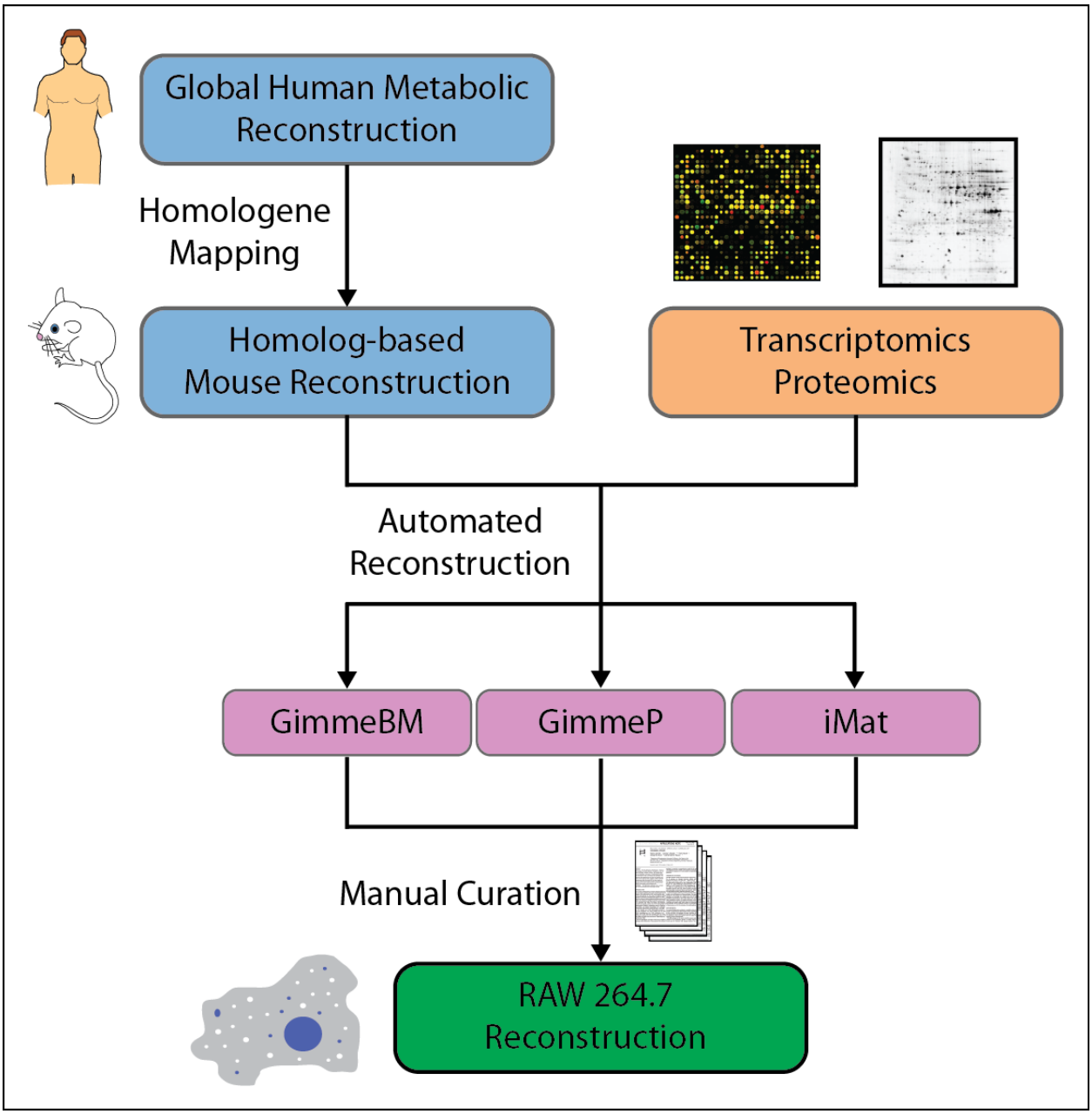

Figure S1: Comprehensive workflow for reconstructing the RAW 264.7 murine macrophage genome-scale metabolic reconstruction. The global human metabolic network, Recon 1, was converted to a draft global mouse network using homologous gene data. Combined with high-throughput data and established algorithms, three draft macrophage reconstructions were built. The three drafts were analyzed, compared, and finally reconciled using existing literature data to build a final RAW 264.7 cell line metabolic reconstruction.

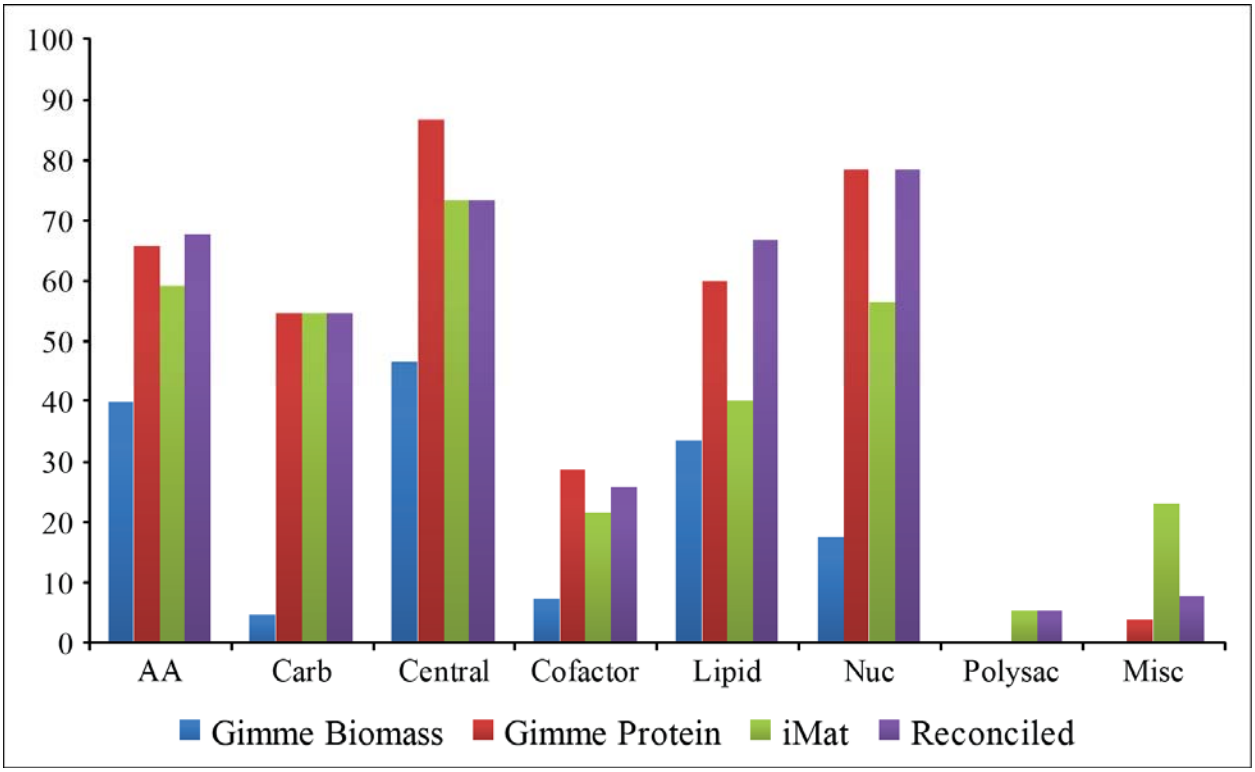

Figure S2: We determined the metabolic functionality of the three draft and the final murine macrophage reconstructions. The results are sorted by metabolic subsystem. GimmeBM represents the basal functionality of all the draft models while GIMMEp has the highest functionality. The Shlomi-NBT-08 draft reconstruction has more functionality in polysaccharide metabolism and miscellaneous, which consists of functions dealing with bile synthesis. Abbreviations: AA – amino acid metabolism, Carb – carbohydrate metabolism, Nuc – nucleotide metabolism, Polysac – polysaccharide metabolism, Misc – miscellaneous functions

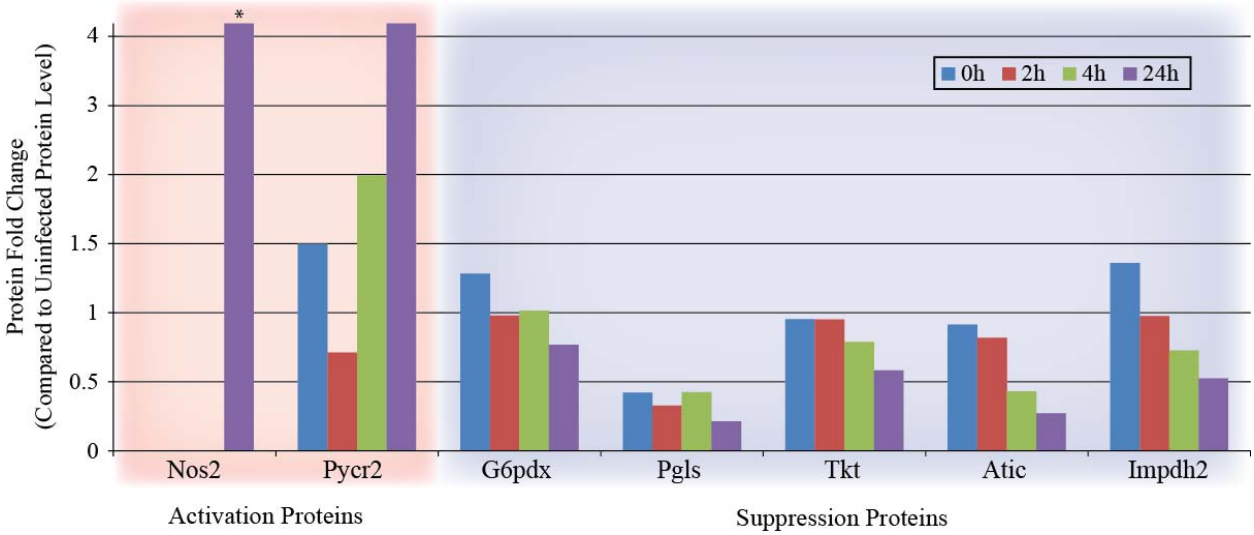

Figure S3: Time course protein fold changes in RAW 264.7 cells infected with *Salmonella typhimurium* were reanalyzed from a previous study (Shi et al. 2009) . Seven proteins dealing with activation (Nos2, Pycr2), pentose phosphate pathway (G6pdx, Pgls, Tkt), and nucleotide biosynthesis (Atic, Impdh2) were detected. The protein changes follow the trends determined from the LPS data suggesting similar metabolic changes during infection. In addition, the time course data shows a clear accumulation of metabolic activation proteins and a continual decrease in pentose phosphate pathway and nucleotide biosynthesis pathways as infection progresses. Note: Nos2 protein was not detected in uninfected protein levels or 0/2/4 h.

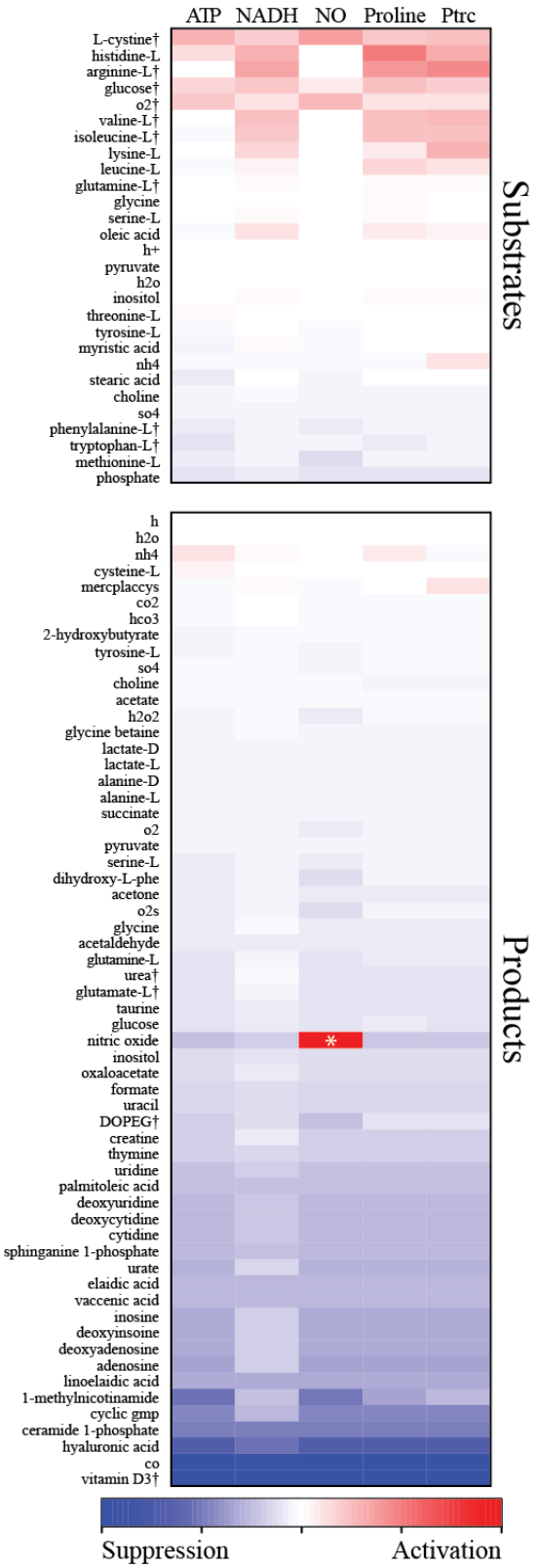

Figure S4: Sensitivity results for substrates and products for the global human metabolic network. The results vary considerably from the ones determined for the RAW 264.7 metabolic network. A full size version is provided in the Supplementary Data.

Table S1: Topological and functional characteristics of the automated draft and final reconstructions

	Gene Recall	Protein Recall	Genes	Unique Metabolites	Intracellular Reactions	Active Reactions	Functional Tests Passed
GimmeBM	0.342	0.369	446	304	424	98.6%	70
GIMMEp	0.642	0.733	809	552	1026	92.0%	139
iMat	0.764	0.736	938	714	1327	80.3%	127
Reconciled	0.658	0.733	820	574	1067	90.1%	146

Table S2: Calculated flux rate comparison of RAW 264.7 and Recon 1

	Experimental	RAW 264.7 Network	Recon 1
Biomass Generation (hr^{-1})	0.0281-0.0630	0.0408	0.0624
ATP Generation (mmol/h/g cell DW)	0.712	0.796	2.40
NO Generation (mmol/h/g cell DW)	0.0365	0.0399	0.132

References:

Duarte N. C., Becker S. A., Jamshidi N., Thiele I., Mo M. L., Vo T. D., Srivas R. and Palsson B. O. (2007). "Global reconstruction of the human metabolic network based on genomic and bibliomic data." Proc Natl Acad Sci U S A **104**(6): 1777-1782.

Hiller K., Hangebrauk J., Jäger C., Spura J., Schreiber K. and Schomburg D. (2009). "MetaboliteDetector: Comprehensive Analysis Tool for Targeted and Nontargeted GC/MS Based Metabolome Analysis." Analytical Chemistry **81**(9): 3429-3439.

Jaitly N., Adkins J. N., Monroe M. E., Norbeck A. D., Mottaz H. H., Dabney A. D., Lipton M. S., Anderson G. A. and Smith R. D. (2008). A Statistical Approach to Quantifying Uncertainties in the Accurate Mass and Time (AMT) tag Pipeline. 55th ASMS Conference on Mass Spectrometry and Allied Topics, Indianopolis, Indiana, USA.

Jerby L., Shlomi T. and Ruppin E. (2010). "Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism." Mol Syst Biol **6**: 401.

Keller A., Nesvizhskii A. I., Kolker E. and Aebersold R. (2002). "Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search." Anal Chem **74**(20): 5383-5392.

Kelly R. T., Page J. S., Luo Q., Moore R. J., Orton D. J., Tang K. and Smith R. D. (2006). "Chemically etched open tubular and monolithic emitters for nanoelectrospray ionization mass spectrometry." Anal Chem **78**(22): 7796-7801.

Kim S., Gupta N. and Pevzner P. A. (2008). "Spectral probabilities and generating functions of tandem mass spectra: a strike against decoy databases." J Proteome Res **7**(8): 3354-3363.

Kim Y.-M., Metz T. O., Hu Z., Wiedner S. D., Kim J.-S., Smith R. D., Morgan W. F. and Zhang Q. (2011). "Formation of dehydroalanine from mimosine and cysteine: artifacts in gas chromatography/mass spectrometry based metabolomics." Rapid Communications in Mass Spectrometry **25**(17): 2561-2564.

Kind T., Wohlgemuth G., Lee D. Y., Lu Y., Palazoglu M., Shahbaz S. and Fiehn O. (2009). "FiehnLib: Mass Spectral and Retention Index Libraries for Metabolomics Based on Quadrupole and Time-of-Flight Gas Chromatography/Mass Spectrometry." Analytical Chemistry **81**(24): 10038-10048.

Livesay E. A., Tang K., Taylor B. K., Buschbach M. A., Hopkins D. F., LaMarche B. L., Zhao R., Shen Y., Orton D. J., Moore R. J., Kelly R. T., Udseth H. R. and Smith R. D. (2008). "Fully automated four-column capillary LC-MS system for maximizing throughput in proteomic analyses." Anal Chem **80**(1): 294-302.

Monroe M. E., Tolic N., Jaitly N., Shaw J. L., Adkins J. N. and Smith R. D. (2007). "VIPER: an advanced software package to support high-throughput LC-MS peptide identification." Bioinformatics **23**(15): 2021-2023.

Patil K. R. and Nielsen J. (2005). "Uncovering transcriptional regulation of metabolism by using metabolic network topology." Proc Natl Acad Sci U S A **102**(8): 2685-2689.

Polpitiya A. D., Qian W.-J., Jaitly N., Petyuk V. A., Adkins J. N., Camp D. G., Anderson G. A. and Smith R. D. (2008). "DAnTE: a statistical tool for quantitative analysis of -omics data." Bioinformatics **24**(13): 1556-1558.

Sheikh K., Forster J. and Nielsen L. K. (2005). "Modeling hybridoma cell metabolism using a generic genome-scale metabolic model of Mus musculus." Biotechnol Prog **21**(1): 112-121.

Shi L., Chowdhury S. M., Smallwood H. S., Yoon H., Mottaz-Brewer H. M., Norbeck A. D., McDermott J. E., Clauss T. R., Heffron F., Smith R. D. and Adkins J. N. (2009). "Proteomic investigation of the time course responses of RAW 264.7 macrophages to infection with Salmonella enterica." Infect Immun **77**(8): 3227-3233.

Shlomi T., Cabili M. N., Herrgard M. J., Palsson B. O. and Ruppin E. (2008). "Network-based Prediction of Human Tissue-specific Metabolism." Nat Biotech **26**(9): 1003-1010.

Sigurdsson M. I., Jamshidi N., Steingrimsson E., Thiele I. and Palsson B. Ø. (2010). "A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1." BMC Syst Biol **4**: 140.

Wang Y., Yang F., Gritsenko M. A., Clauss T., Liu T., Shen Y., Monroe M. E., Lopez-Ferrer D., Reno T., Moore R. J., Klemke R. L., Camp D. G., 2nd and Smith R. D. (2011). "Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells." Proteomics.

Whitlock M. C. (2005). "Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach." J Evol Biol **18**(5): 1368-1373.

Yates J. R., 3rd, Eng J. K., McCormack A. L. and Schieltz D. (1995). "Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database." Anal Chem **67**(8): 1426-1436.

Zimmer J. S., Monroe M. E., Qian W. J. and Smith R. D. (2006). "Advances in proteomics data analysis and display using an accurate mass and time tag approach." Mass Spectrom Rev **25**(3): 450-482.