

NH₂-Terminal Sequence of Macrophage-expressed Natural Resistance-associated Macrophage Protein (*Nramp*) Encodes a Proline/Serine-rich Putative Src Homology 3-binding Domain

By C. Howard Barton, Jacqueline K. White, Tamara I. A. Roach, and Jenefer M. Blackwell

From the Department of Medicine, University of Cambridge Clinical School, Addenbrooke's Hospital, Cambridge CB22QQ, United Kingdom

Summary

The *Lsh/Ity/Bcg* locus on mouse chromosome 1 regulates macrophage (m ϕ) priming/activation for antimicrobial activity against intracellular pathogens. A candidate *Bcg* gene, designated natural resistance-associated m ϕ protein (*Nramp*), recently isolated from a pre-B cell cDNA library encodes a polytopic integral membrane protein with structural features common to prokaryotic and eukaryotic transporters. In the present study, an activated m ϕ cDNA library yielded new *Nramp* clones that differ in the 5' region from the published pre-B cell-derived clone sequence, resulting in addition of 64 amino acids at the NH₂ terminus of the predicted protein. This new domain is rich in proline, serine, and basic amino acids, and includes three protein kinase C phosphorylation sites and a putative Src homology 3 binding domain. RNAs containing this domain are the only form found in the m ϕ . Hence, the protein encoded by this RNA is the candidate molecule mediating natural resistance to intra-m ϕ pathogens.

The murine macrophage (m ϕ) resistance gene *Lsh/Ity/Bcg* regulates the priming/activation of m ϕ for antimicrobial activity against *Leishmania donovani*, *Salmonella typhimurium*, and *Mycobacterium spp.* (1–5). Recently Vidal et al. (6) isolated a candidate gene, designated *Nramp* (natural resistance-associated m ϕ protein), that encodes a polytopic integral membrane protein that has structural features common to prokaryotic and eukaryotic transporters. The presence of a small consensus motif showing sequence identity with nitrate transporters led these workers to hypothesize that *Nramp* might function as a nitrate concentrator in the phagolysosome membrane of the infected m ϕ , the acid environment of the phagolysosome mediating conversion via nitrite to toxic nitric oxide (NO). However, this hypothesis fails to take account of many studies demonstrating that *Lsh/Ity/Bcg* regulates m ϕ activation (7–10) leading to the TNF- α -dependent production of antimicrobial NO (11). The gene also has many pleiotropic effects including: (a) downregulation of 5' nucleotidase (9, 10); (b) upregulation of MHC class II, TNF- α production, IL-1 β expression, AcM.1 antigen expression, oxidative burst, and tumouricidal activity (7–10); (c) rapid (within 30 min) upregulation of the early gene KC, a neutrophil-specific chemoattractant of the IL-8-related C-X-C family of small peptide cytokines (12); and (d) integrin-mediated upregulation of TNF- α production by plating m ϕ onto the extracellular matrix proteins fibrinogen and fibronectin (13). These observations suggest that *Lsh/Ity/Bcg* must encode a mole-

cule that plays a role early in the m ϕ activation pathway leading to antimicrobial activity. A study of m ϕ -expressed *Nramp* was therefore undertaken to determine how gene structure might relate to these functional observations and whether *Nramp* is a valid *Lsh/Ity/Bcg* gene candidate.

Materials and Methods

Sequence Analysis of *Nramp* Clones from m ϕ cDNA Library. M ϕ *Nramp* clones were isolated from an activated (4 h stimulation; 25 U/ml interferon- γ , 10 ng/ml *S. typhimurium* LPS) mouse (B10.L-*Lsh*⁺) m ϕ cDNA library prepared in λ UniZap (Stratagene, La Jolla, CA). Clones were isolated by filter plaque hybridization using a probe generated by reverse transcriptase-PCR corresponding to nucleotides 1410–1812 bp of the published (6) sequence. After plaque purification, 35 clones from 10⁶ recombinants were analyzed by PCR using sense and antisense *Nramp* primers in combination with T3 and T7 vector arm primers. This allowed the mapping of clones with respect to the published sequence. 20 out of 30 were 1.0–1.5 kb and were not analyzed further. The remainder were 2.1–2.3 kb and potentially full-length *Nramp* clones. These were restriction mapped and four were selected for sequencing (Sequenase II) including the longest clone λ 8.1.

Genomic Sequencing. From the m ϕ sequence, PCR primers (CCT GGT GAC CAC ACA CAG and CAC CTT GGG GTA GAG ATG) were generated to amplify a 2-kb region from both yeast artificial chromosome (YAC clone C9C28; Princeton library, Howard Hughes Medical Institute, Princeton University, Princeton, NJ) and mouse genomic DNAs. The products were cloned in the

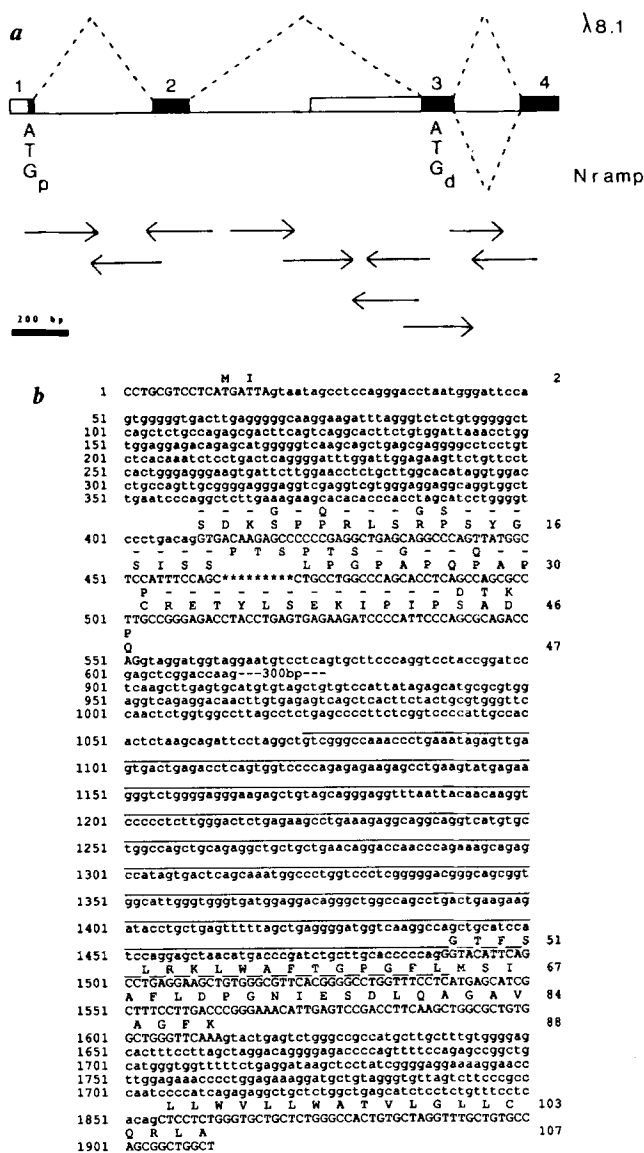


Figure 2. Map and genomic sequence for the 5' region of *Nramp*. (a) Map demonstrating that the additional sequence of $m\phi$ *Nramp* is encoded by two unique exons (1 and 2; solid bars) contiguous in the cDNA sequence (Fig. 1) with exons (3 and 4; solid bars) common to both $\lambda 8.1$ and published *Nramp*. Contiguous with and 5' of exon 3 is the putative 5' UTR (open bar) found earlier (6). Predicted splicing patterns (dotted lines) are indicated above ($\lambda 8.1$) and below (published pre-B cell clone) the map. Arrows indicate primers for sequencing gel reads. (b) Genomic sequence corresponding to nucleotides 31-456 of $\lambda 8.1$ spanning the point of divergence with the published (6) sequence. Exonic nucleotide sequence is in capitals, with the predicted amino acid sequence above in single letter format. Intron sequence is in small letters. The region of 5' UTR from the published clone, contiguous with exon 3, is overlined. The codon (ATG = Met) where this terminates indicates the initiation codon of published *Nramp*. The predicted amino acid sequence for human NRAMP exon 2 is shown above the murine sequence. (-) amino acid identity. (asterisk) Gap introduced for alignment of human and murine sequences.

lished (6) *Nramp* cDNA. The 5' UTR sequence from the published clone was found in the 900-bp intron contiguous with and including part of exon 3, indicating that a single gene encodes both forms. Exon 3 is particularly unusual in that it encodes protein sequence in $\lambda 8.1$, whereas for published *Nramp* it contains both coding and noncoding sequence. Although a complex mechanism involving alternative splicing with an internal splice acceptor site and dual promoter control could be formulated to account for both forms, it seems more likely that the published (6) cDNA clone contains a fragment of the 900-bp intron at its proximal end. This is consistent with the observation that a number of the $m\phi$ -derived *Nramp* clones isolated here were found to contain sequence that exhibited identity with the first *Nramp* intron identified in genomic DNA (not shown). The human genomic sequence corresponding to base pairs 408-455 of the murine genomic sequence (Fig. 2) shows 77% identity with mouse at the nucleotide level (not shown) and 68% (82% with conserved substitutions) for the predicted amino acid sequence (Fig. 2). The predicted amino acid sequence encodes a similar proline/serine-rich domain containing an additional 3 amino acid insert at position 21 of the murine sequence.

Only One Form of Nramp Is Expressed in mφ. To confirm that RNA encoding the longer polypeptide is the form expressed in $m\phi$, a number of experimental approaches were adopted (Fig. 3). Using $m\phi$ RNA as template, primer extension with an oligonucleotide unique to the 5' region of $\lambda 8.1$ yielded products in both susceptible and resistant mice mapping to nucleotides 50/51 of the cDNA sequence (Fig. 1b). In contrast, no products were generated using an oligonucleotide within the putative 5' UTR of published (6) *Nramp*. These experiments confirmed that RNA transcripts bearing the putative 5' untranslated region of the published cDNA are not present in resting (not shown) or activated $m\phi$, whereas transcripts corresponding to the $\lambda 8.1$ sequence were identified with transcriptional initiation sites mapping 21 and 22 bp (doublet) 5' of the proximal ATG codon. A probe covering the 5' region unique to $\lambda 8.1$ also hybridized well to Northern and slot blots of macrophage RNA from susceptible and resistant mice, whereas a probe covering the putative 5' UTR of the published clone showed no hybridization. Hence, the only form of RNA transcript present in $m\phi$ is that which conforms to the $\lambda 8.1$ predicted polypeptide sequence. It is therefore this polypeptide that is the candidate natural resistance protein.

Predicted Structure and Sequence Identities Across the NH₂-terminus of mφ-expressed Nramp. To determine how the additional domains of $m\phi$ -expressed *Nramp* influence the structure of the molecule, hydrophathy (14) plots, and amino acid database searches were undertaken over the 64-amino acid domain. The former (data not shown) demonstrated that the new sequence is hydrophilic, and forms an extension to the NH₂-terminal cytoplasmic domain. The amino acid database search (15) over this proline, serine and basic rich 64-amino acid domain identified three PKC phosphorylation sites (in addition to two identified in published *Nramp*), and a number of matches with unrelated proteins (Fig. 4). The most in-

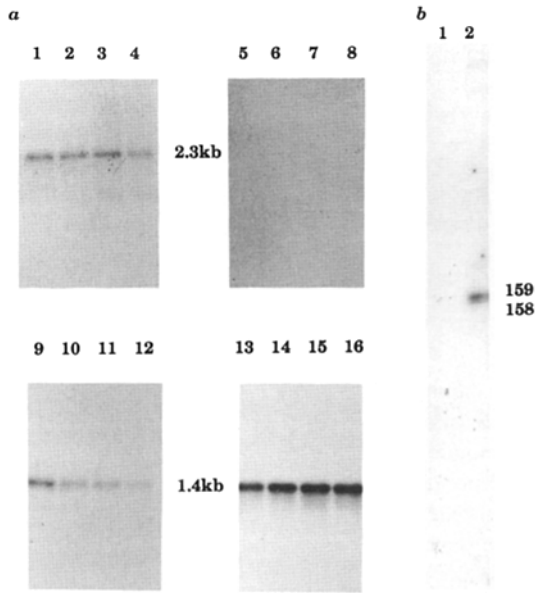


Figure 3. Northern blot (a) and primer extension (b) analysis demonstrating that *Nramp* transcripts encoding the additional 64 amino acids are the only form of *Nramp* expressed in m ϕ . (a) Northern blots from resting m ϕ (lanes 1, 5, 9, and 13), or m ϕ activated with interferon- γ (lanes 2, 6, 10, and 14), LPS (lanes 3, 7, 11, and 15), or interferon- γ plus LPS (lanes 4, 8, 12, and 16) were hybridized with probes specific for unique λ 8.1 5' sequence (lanes 1-4), or for the more distal putative 5'UTR of the published (6) sequence (lanes 5-8). Lanes 9-16 show the same blots reprobed with GAPDH. Twice as much RNA was loaded on blots hybridized with the distal probe. Results are shown for RNA extracted from bone marrow-derived m ϕ from C57BL/10ScSn mice. Slot blot analysis (not shown) confirmed that the λ 8.1-specific probe hybridized to RNA from both susceptible and resistant macrophages. (b) Primer extension products obtained with 10 μ g of total RNA from B10.L-*Lsh*^r m ϕ using oligonucleotides specific to the putative 5' untranslated region of the published (6) sequence (lane 1), or to 5' sequence unique to λ 8.1 (lane 2). The numbers of nucleotides from the 5' end of the primer are shown. Control reactions with tRNA gave no products with either primer (not shown). Similar results were obtained using RNA from C57BL/10ScSn (*Lsh*^s) m ϕ as template.

triguing matches were with: (a) *Drosophila* dynamin (18) (55% identity over 20 residues), related to the synaptic phosphoprotein dephosphin in rat brain (19); (b) the proline-rich third cytoplasmic domain of the adenylate cyclase stimulatory and G protein coupled β 1-adrenergic receptor (20) (57% over 21 amino acids); and (c) focal adhesion kinase (21) (50% over 26 amino acids) modulated by integrin-dependent phosphorylation (22). The region of identity with dynamin has been implicated (23) in binding anionic phospholipids, microtubules and Src homology 3 (SH3) domains. SH3 domains (24, 25), identified as related sequences in tyrosine kinases (TK), are modular and found in proteins such as non-receptor TKs, phospholipase C- γ and structural proteins of the cytoskeleton. Whilst the function of SH3 domains (24) is not as well characterized as the SH2 counterpart, it is believed they mediate specific protein-protein interactions obligatory for signal transduction (25). Members of the Src family of membrane-associated TKs, including *Hck* and *Fgr* which have SH3 domains, are also found in m ϕ (26). Both exhibit

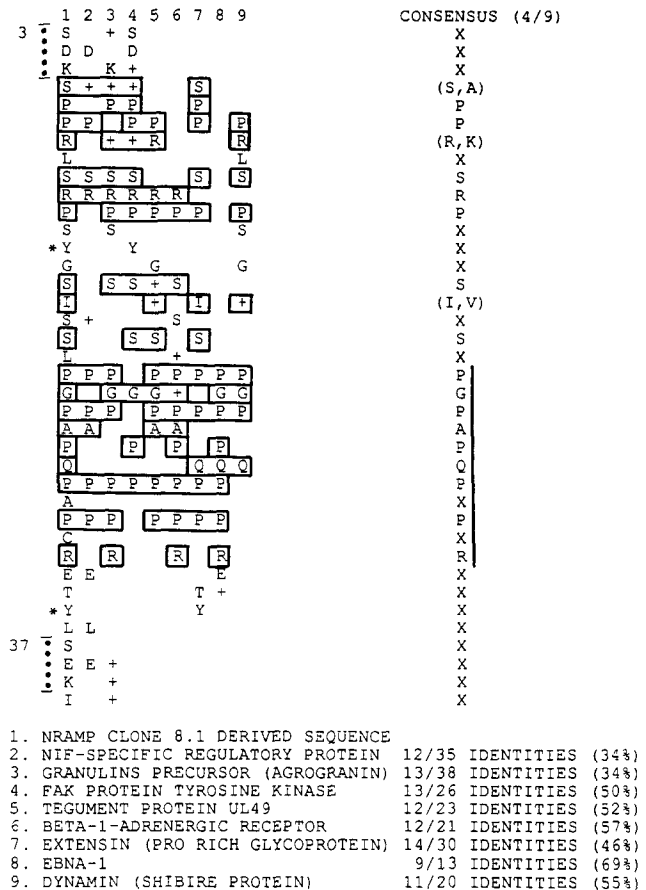


Figure 4. Results of amino acid database searches with the 64 amino acid sequence unique to macrophage-expressed *Nramp* identifying a number of sequence matches particularly with the proline rich sequence; + represents a conserved amino acid. Multiple sequence alignments allowed for the generation of a consensus motif over this region: PGPAPQXPXPR (solid vertical bar). The nine best matches were aligned with each other and residues boxed where four or more exhibited identities. Also shown are the two PKC sites (hatched vertical bars) on S3 and S37 which flank the region exhibiting sequence identity. Tyrosine residues (asterisk) occur on either side of the consensus motif.

differential kinetics in response to activation signals and could be implicated in *Nramp*-mediated signal transduction. *Hck*, in particular, is involved in signal transduction for TNF- α production in murine m ϕ (27), which is crucial for NO production and antimicrobial activity in *Lsh*-resistant m ϕ (11). Early gene *KC* expression in *Lsh*-resistant m ϕ also involves a Ca^{2+} -dependent, NO-mediated, cyclic GMP-dependent kinase pathway (12). Phosphorylation of the *Nramp* SH3-binding domain on tyrosines might itself regulate transport of important substrates (e.g., L-arginine) required initially for generation of NO mediating signal transduction (12) and subsequently for antimicrobial activity (11). Bogle et al. (28) have shown that L-arginine transport by m ϕ is stimulated upon activation with interferon- γ and LPS, providing a mechanism for sustained substrate supply during the generation of NO. Control of transport function through interaction of TKs with the putative SH3 binding domain of *Nramp* may

provide a link between its structural identity with other transporters (6) and the decade of functional analysis demonstrating a role for the *Lsh/Ity/Bcg* gene in m ϕ activation (7–13). The pleiotropic effects of *Nramp* would thus be due to its role in providing the substrate essential for generation of NO in-

involved in signal transduction and as the final effector for antimicrobial activity. Further analysis of m ϕ -expressed *Nramp* provides an exciting basis to future research aimed at understanding m ϕ activation at the molecular level.

This work was supported by The Wellcome Trust.

Address correspondence to Dr. C. Howard Barton, University of Cambridge, Clinical School, Department of Medicine, Level 5, Addenbrooke's Hospital, Hills Road, Cambridge CB22QQ, UK.

Received for publication 29 November 1993 and in revised form 11 February 1994.

References

1. Plant, J., and A.A. Glynn. 1974. Natural resistance to *Salmonella* infection, delayed hypersensitivity and Ir genes in different strains of mice. *Nature (Lond.)* 248:345.
2. Bradley, D.J. 1974. Genetic control of natural resistance to *Leishmania donovani*. *Nature (Lond.)* 250:353.
3. Gros, P., E. Skamene, and A. I. Forget. Genetic control of natural resistance to *Mycobacterium bovis* (BCG) in mice. 1981. *J. Immunol.* 127:2417.
4. Skamene, E., P. Gros, A.J. Forget, P.A.L. Kongshavn, C. St. Charles, and B.A. Taylor. 1982. Genetic regulation of resistance to intracellular pathogens. *Nature (Lond.)* 297:506.
5. Plant, J.E., J.M. Blackwell, A.D. O'Brien, D.J. Bradley, and A.A. Glynn. 1982. Are the *Lsh* and *Ity* disease resistance genes at one locus on mouse chromosome 1? *Nature (Lond.)* 297:510.
6. Vidal, S.M., D. Malo, K. Vogan, E. Skamene, and P. Gros. 1993. Natural resistance to infection with intracellular parasites: isolation of a candidate for *Bcg*. *Cell* 73:469.
7. Blackwell, J.M. 1989. The m ϕ resistance gene *Lsh/Ity/Bcg*. *Res. Immunol.* 140:767.
8. Blackwell, J.M., T.I.A. Roach, S.E. Atkinson, J.W. Ajioka, C.H. Barton, and M.-A. Shaw. 1991. Genetic regulation of m ϕ priming/activation: the *Lsh* gene story. *Immunol. Lett.* 30:241.
9. Schurr, E., D. Radzioch, D. Malo, P. Gros, and E. Skamene. 1991. Molecular genetics of inherited susceptibility to intracellular parasites. *Behring Inst. Mitt.* 88:1.
10. Schurr, E., D. Malo, D. Radzioch, E. Buschman, K. Morgan, P. Gros, and E. Skamene. 1991. Genetic control of innate resistance to mycobacterial infections. *Immunol. Today* 12:A42.
11. Roach, T.I., A.F. Kiderlen, and J.M. Blackwell. 1991. Role of inorganic nitrogen oxides and tumor necrosis factor alpha in killing *Leishmania donovani* amastigotes in gamma interferon-lipopolysaccharide-activated m ϕ from *Lsh^s* and *Lsh^r* congenic mouse strains. *Infect. Immun.* 59:3935.
12. Roach, T.I.A., D. Chatterjee, and J.M. Blackwell. 1994. Induction of early response genes KC and JE by mycobacterial LAMs: regulation of KC expression in murine m ϕ by *Lsh/Ity/Bcg(Nramp)*. *Infect. Immun.* In press.
13. Formica, S., T.I.A. Roach, and J.M. Blackwell. 1994. Interaction with extracellular matrix proteins influences *Lsh/Ity/Bcg* (candidate *Nramp*) gene regulation of m ϕ priming/activation for TNF- α release. *Immunology*. In press.
14. Kyte, J., and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105.
15. Devereaux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acid Res.* 12:387.
16. Kozak, M. 1984. Point mutations close to the AUG initiator codon affect the efficiency of translation of rat pre-proinsulin *in vivo*. *Nature (Lond.)* 308:241.
17. Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44:283.
18. Chen, M.S., R.A. Obar, C.C. Schroeder, T.W. Austin, C.A. Poodry, S.C. Wadsworth, and R.B. Vallee. 1991. Multiple forms of dynamin are encoded by shibire, a *Drosophila* gene involved in endocytosis. *Nature (Lond.)* 351:583.
19. Robinson, P.J., J.-M. Sontag, J.-P. Liu, E.M. Fykse, C. Slaughter, H. McMahon, and T.C. Südhof. 1993. Dynamin GTPase regulated by protein kinase C phosphorylation in nerve terminals. *Nature (Lond.)* 365:163.
20. Friele, T., S. Collins, K.W. Daniel, M.G. Caron, R.J. Lefkowitz, and B.K. Kobilka. 1987. Cloning of the cDNA for the human beta 1-adrenergic receptor. *Proc. Natl. Acad. Sci. USA* 84:7920.
21. Schaller, M.D., C.A. Borgman, B.S. Cobb, R.R. Vines, A.B. Reynolds, and J.T. Parsons. 1992. pp125^{FAK}, a structurally distinctive protein-tyrosine kinase associated with focal adhesion. *Proc. Natl. Acad. Sci. USA* 89:5192.
22. Lipfert, L., B. Haimovich, M.D. Schaller, B.S. Cobb, J.T. Parsons, and J.S. Brugge. 1992. Integrin-dependent phosphorylation and activation of the protein tyrosine kinase pp125^{FAK} in platelets. *J. Cell Biol.* 119:905.
23. Vallee, R.B., and H.S. Shpetner. 1993. Dynamin in synaptic dynamics. *Nature (Lond.)* 365:107.
24. Mayer, B.J., and D. Baltimore. 1993. Signalling through SH2 and SH3 domains. *Trends Cell Biol.* 3:8.
25. Booker, G.W., I. Gout, A.K. Downing, P.C. Driscoll, J. Boyd, M.D. Waterfield, and I.D. Campbell. 1993. Solution structure and ligand-binding site of the SH3 domain of the p85 α subunit of phosphatidylinositol 3-kinase. *Cell* 73:813.
26. Ziegler, S.F. C.B. Wilson, and R.M. Perlmutter. 1988. Augmented expression of a myeloid-specific protein tyrosine kinase gene (*hck*) after macrophage activation. *J. Exp. Med.* 168:1801.
27. English, B.K., J.N. Ihle, A. Myracle, and T. Yi. 1993. *Hck* tyrosine kinase activity modulates tumor necrosis factor production by murine macrophages. *J. Exp. Med.* 178:1017.
28. Bogle, E.G., A.R. Baydoun, J.D. Pearson, S. Moncada, and G.E. Mann. 1992. L-Arginine transport is increased in m ϕ generating nitric oxide. *Biochem. J.* 284:15.