

Supporting Information

Troutman et al. 10.1073/pnas.1118579109

SI Results

B-Cell Adapter for PI3K Toll-IL-1 Receptor Domain Prediction and Modeling. The presence of an N-terminal Toll-IL-1 receptor (TIR) domain in B-cell adapter for PI3K (BCAP) was first revealed by a fast sequence-based search tool, PsiBlast (1), that starts to collect TIR domains from adapter proteins [like myeloid differentiation primary response (Myd88), TIRAP (TIR-domain-containing adapter protein), TRIF (TIR domain-containing adapter-inducing IFN- β), TRAM (TRIF-related adapter molecule), and SARM (sterile α and HEAT-Armadillo motifs)] (2) at the third iteration with highly significant E values of 4×10^{-24} , and then aligns to the greater family of receptor TIR domains [from Toll-like receptors (TLRs) and IL1Rs, and plant-resistance proteins] in the fourth iteration with an E value of 1×10^{-7} . A more sensitive and recent PsiBLAST variant, CS-Blast (3), similarly encounters the BCAP N-terminal domain similarity to both adapter and receptor TIR domains by the third iteration; HH-Blits, a very sensitive iterative search program that relies on HMM-HMM comparison (4), also clusters BCAP homologs with TIR domain adapters and receptors by the third iteration, with a highly significant E value of 1.2×10^{-7} . Compass (5), a program that uses PsiBLAST-derived profiles to comb structural databases, initially corroborated the sequence-based family searches by linking the BCAP N-terminal domain to a number of TIR domain folds from the PDB, at highly significant E values of 2.96×10^{-4} – 4.96×10^{-6} .

The sequence-based links between the BCAP N-terminal domain and TIR modules are clearly at a low level of similarity, but are akin to the relationships established between divergent adapter, plant, and bacterial TIRs. For example, drawing from the PsiBLAST hits, the human BCAP alignment with the human TIRAP TIR domain shows 18% identity (34% similarity, by BLOSUM62 matrix), and with the *Drosophila* MyD88 TIR domain, 16% identity (38% similarity). To place this into perspective, the acknowledged *Drosophila* ortholog of BCAP, DOF, has an 18% identity (37% similarity) to the same N-terminal domain of human BCAP. This degree of identity between BCAP and other TIR domain adapters is very similar to the sequence relationship of the most recent addition to the human adapter repertoire, SARM (6, 7), and the other four adapters: SARM and TIRAP match up at 19% identity (37% similarity), but SARM and *Drosophila* MyD88 display 16% identity (41% similarity).

The TIR domain superfamily has considerably grown since the first descriptions of this domain (8, 9), and now encompasses a large number of receptors and adapters in vertebrates (10, 11), and across the phylogenetic spectrum, including an assortment of bacterial homologs (12, 13). Sequence alignments that bridge this wide evolutionary gulf show a low level of similarity, with no columns of enforced or invariant identity. Regions of the alignment that were initially described as signature motifs in the TIR domains of human TLRs and IL1Rs (Boxes 1–3, for example), do not display the same degree of identity once the more divergent TIR members are included, like SARM adapters, or plant and bacterial TIRs. Still, X-ray and NMR structural investigations have convincingly shown that the characteristic TIR fold first glimpsed in the structure of TLR2 (PDB 1FYX), or later in MyD88 (2JS7), is equally well preserved in plant (14, 15) and bacterial (16) TIRs. The microbial TIR domain proteins show the lowest degree of identity with their human homologs, yet have been shown convincingly to be functionally capable of interacting with the human innate immune signaling machinery. TIR domain virulence factors from pathogenic bacteria subvert

host defenses by sequestering adapter (MyD88 and TRIF) TIR domains (17–19), which is a striking demonstration that fold and interaction function are much better conserved than sequence in the TIR superfamily.

To further confirm the structural relationship of the predicted BCAP TIR domain to well-defined TIR folds, we used sensitive fold-recognition algorithms that can powerfully combine the evolutionary information in deep sequence profiles with accurate structure prediction tools (4, 20), to search the PDB for statistically significant matches. Here we relied on a number of programs that have displayed great success in past CASP (critical assessment of structure prediction) contests—most recently the ninth iteration (21)—like HHPred (22), I-TASSER (23), Raptor-X (24), Phyre2 (25), and IntFOLD (26). For greater consistency, we performed these fold-recognition searches with both the predicted human BCAP and *Drosophila* DOF TIR domains, which as described earlier, share only 18% sequence identity. HHPred, which relies on HMM-HMM comparisons of both sequence and structural parameters (22), clearly shows that the human BCAP TIR domain retrieves the eight available mammalian, plant, and bacterial TIR domain structures as the top matches with very high E values (7×10^{-5} to 2.5×10^{-1}) and *P* values (2.6×10^{-9} to 9.3×10^{-6}), with corresponding confidence probabilities (of the matches being true positives) ranging from 97% to 79.5%. In comparison, the predicted *Drosophila* DOF TIR domain elicited the same top matches (albeit in a bit different rank order) with E values of 5×10^{-4} to 1.3×10^{-1} , *P* values of 1.9×10^{-8} to 3.8×10^{-7} , and probabilities of 96.1% to 88.7%. Summarizing the remaining routines (which use different computational approaches, and perhaps search distinct “cuts” of the PDB), I-TASSER, Raptor-X, Phyre2, and IntFOLD all converge on a similar answer for the BCAP and DOF TIR domains by mapping these chains with very high reliability and statistical significance to extant TIR structures. Notably, I-TASSER (23), the highest-rated fold-recognition program at the recent CASP9 competition (see evaluation at http://predictioncenter.org/casp9/groups_analysis.cgi?type=server&tbm=on&tbfm=on&fm=on&submit=Filter) scores the predicted human BCAP and *Drosophila* DOF TIR domains with top confidence or C-scores of -0.726 and -1.474 , respectively, strongly indicative of structural homology to solved TIR folds. We have successfully used this fold-recognition approach in an earlier work to show that the Frizzled domain family (operating in Wnt and Hedgehog signaling pathways) was distantly related at a structural and functional level with a clan of sterol and folate-binding proteins (27).

To further test the fit of the predicted BCAP and DOF TIR modules to bona fide TIR domains, we proceeded to generate full atomic, comparative models of their protein structures—drawing from the various fold-recognition analyses and template alignments—to assess the quality of the fold predictions (28). The low but significant degree of identity between BCAP/DOF TIRs and TIR domains of known structure in the PDB allowed very reliable threading by the fold-recognition programs of our query sequences with all of the individual or superposed combinations of TIR domains in the PDB. This template-based modeling approach (20, 28) produced a series of well-ordered models (coordinates available upon request) of both human BCAP and *Drosophila* DOF TIR domains (Fig. S1). We focused our attention on the top comparative models drawn from highly-ranked HHPred (22) that are generated by MODELER (29), I-TASSER (23), Raptor-X (24), Phyre2 (25), and IntFOLD (26). These models are ranked by their respective servers using dis-

tinct confidence scores; for example, for the human BCAP TIR, the first HHpred model has a probability score of 97%, from I-TASSER, a C-score of -0.726 , from Raptor-X, a global distance test (GDT) score of 46, from Phyre2, a confidence score of 96.7%, and from IntFOLD, a P value of 3.96×10^{-3} . In turn, these high-confidence models were then subjected to model quality assessment programs to see how close they approximate “real” (or X-ray-derived) protein structures, by looking at measures like stereochemistry, hydrophobic core packing, electrostatics, self-threading, and free energy (30), that also help to pinpoint regions of the model structures (like surface exposed loops) that might bear refinement. For example, the human BCAP TIR domain I-TASSER model 1 shows a PROSA (31) self-threading z -score of -7.46 , squarely within the normal range for X-ray and NMR structures; ModFOLD (32), in turn, issues a high confidence score of 2.39×10^{-3} ; ProQ (33) gives an extremely good LG score of 4.219; and finally, the total dDFIRE (34) conformation free-energy score is -336 , well in line with solved structures. The *Drosophila* DOF TIR domain models fare quite similarly, also suggesting that the comparative models are of high quality.

Although the BCAP and DOF TIR modules are distantly related to the same measure that they can be respectively linked to the greater family of TIR domain sequences, we furthermore conclude with high confidence that they adopt the distinctive fold-architecture captured in solved TIR domain structures. That the TIR domain fold can be uniquely specified for a range of bacterial, plant, insect and animal proteins—without retaining any invariant residues—is a prime example of how protein folds are far better conserved than sequences, particularly when structural and functional constraints are imposed on their evolution (35). The delineation of TIR domains in the chains of BCAP orthologs followed an iterative process (at both sequence and predicted structural levels) that progressively builds the case that they can be confidently linked to the TIR domain superfamily, and that known TIR domain structures provide the best templates for their 3D folds (27, 36).

SI Materials and Methods

Antibodies and Reagents. Antibodies used were: purified anti-IL-6, biotin-conjugated anti-IL-6, purified anti-IL-12 p40/p70, biotin-conjugated anti-IL-12 p40/p70, anti-IL-17A, biotin-conjugated anti-IL-17A, purified anti-IFN- γ , biotin-conjugated anti-IFN- γ , purified anti-TNF- α , biotin-conjugated anti-TNF- α (BD Biosciences), antiphosphorylated Akt (Cell Signaling #4051 and #4060) anti-Akt (Cell Signaling #9272), antiphosphorylated ERK1/2 (Cell Signaling #9101), anti-ERK1/2 (Cell Signaling #9102), antiphosphorylated I κ B α (Cell Signaling #2859), anti-I κ B α (Cell Signaling #9242), antiphosphorylated JNK (Cell Signaling #9251), anti-JNK (Cell Signaling #9252), anti-MyD88 (R&D #AF3109), antiphosphorylated PI3K p85 (Cell Signaling #4228), anti-PI3K p85 (Cell Signaling #4257), polyclonal anti-BCAP (R&D Systems #AF4857), anti-Flag M2 (Sigma Aldrich), anti-HA (Santa Cruz; clone 12CA5), anti-HA (Covance; clone 16B12), FITC-conjugated anti-TCR β , FITC-conjugated anti-TNF- α , PE-conjugated anti-CD11b, PE-conjugated anti-CD44, PerCP-conjugated anti-CD4, biotin-conjugated anti-CD62L (BD Biosciences), and biotin-conjugated anti-Ly-6c (eBioscience). Secondary antibodies and reagents used were HRP-conjugated streptavidin (Jackson ImmunoResearch), Pacific Blue-conjugated streptavidin (Invitrogen), APC-conjugated streptavidin (eBioscience), HRP-conjugated goat anti-mouse, HRP-conjugated donkey anti-goat, and HRP-conjugated donkey anti-rabbit (Jackson ImmunoResearch). TLR ligands used were Ultrapure LPS, MALP-2, Pam3CSK4 (Invivogen), and CpG ODN 1826, TCCATGACGTTCTGACGTT, with phosphorothioate linkages (W. M. Keck Facility).

Bone Marrow-Derived Macrophages. Bone marrow-derived macrophages (BMDM) were obtained using standard techniques. In brief, marrow cells were isolated from the femurs and tibias of mice. Cells were cultured overnight in tissue culture-treated dishes to remove stromal cells in the presence of conditioned L929 culture supernatant as a source of M-CSF. Nonadherent cells were collected and two to three million cells cultured for an additional 5 to 7 d with conditioned L929 supernatant in non-tissue culture-treated dishes. Following differentiation, BMDM were collected and plated for use in experiments. BMDM were stimulated, unless indicated otherwise, for 20 h with 100 ng/mL LPS, 100 ng/mL Pam3CSK4, 100 ng/mL MALP-2, or 1 μ M CpG.

Cloning. BCAP and mutants were cloned into host vectors following PCR amplification using an Integrated Molecular Analysis of Genomes and their Expression (I.M.A.G.E.) Consortium clone as the template (clone ID 40047744). The BCAP-TIR mutant corresponded with amino acids 1–321 and the BCAP Δ TIR mutant corresponded with amino acids 180–812. Clones were verified by sequencing.

ELISA. Cytokines were quantified by the sandwich ELISA method using coating and detecting antibodies following standard techniques. HRP-conjugated streptavidin was used to detect bound biotin conjugated antibodies.

Flow Cytometry and Intracellular Staining. Erythrocyte lysed cell suspensions were stained with antibodies and samples run on a BD LSR II. In some experiments, samples were run on a BD FACSCalibur. Intracellular staining was performed following *ex vivo* restimulation for 5 h with heat-killed *Salmonella typhimurium* in the presence of brefeldin A. Following stimulation, cells were surface stained, then intracellularly stained using a cell fixation and permeabilization buffer set according to the manufacturer's protocol (Biolegend #421403). Samples were run on a BD FACSCalibur. Flow cytometry data were analyzed using FlowJo software (Tree Star).

In Vitro Priming of OT-II T Cells. Transgenic OT-II CD4 T cells were purified from the spleen and lymph nodes by negative selection using the following hybridoma supernatants: anti-CD8 (TIB-105 and TIB-150), anti-CD11b (TIB-128), anti-B220 (TIB-146 and TIB-164), anti-NK1.1 (HB191), and anti-MHC-II (Y3JP). Cell suspensions were depleted of CD8 T cells, B cells, NK cells, and myeloid cells by coupling antibody labeled cells to magnetic beads bound to goat anti-rat IgG, goat anti-mouse IgG, and goat anti-mouse IgM (Qiagen). Splenic dendritic cells were enriched from B6 or BCAP KO mice bearing B16 melanoma tumors secreting Flt3l by negative selection. In brief, splenocytes were labeled with the hybridoma supernatants anti-CD90 (Y19) and anti-NK1.1 (HB191). Cell suspensions were depleted of T cells, B cells, and NK cells by coupling labeled cells to magnetic beads bound to goat anti-rat IgG, goat anti-mouse IgG, and goat anti-mouse IgM. For *in vitro* priming, enriched OT-II CD4 T cells were cultured with splenic dendritic cells at a 5:1 ratio for 72 h with 100 ng/mL LPS, 100 ng/mL Pam3CSK4, or 1 μ M CpG and titrating doses of ovalbumin. After 72 h cells were pulsed with 0.5 μ Ci 3H-thymidine (Perkin-Elmer) for an additional 16 h and thymidine uptake was monitored as a readout of cell proliferation using a MicroBeta liquid scintillation counter (Perkin-Elmer). For analysis of secreted cytokines, T cells and dendritic cells were cultured as above with 10 μ g/mL ovalbumin for 72 h and cytokines quantified by ELISA.

shRNA Knockdown. BCAP specific shRNA constructs were from Sigma Aldrich (clones used were: NM_031376.1–880s1c1,

NM_031376.1–1464s1c1, NM_031376.1–1538s1c1 and NM_031376.1–1743s1c1) and lentiviruses were packaged using 293T cells. RAW264.7 macrophages were infected with the viral particles and selected for resistance to 2 $\mu\text{g}/\text{mL}$ puromycin. Efficiency of BCAP silencing was monitored by immunoblotting.

Cells were stimulated, unless indicated otherwise, for 20 h with 100 ng/mL LPS, 100 ng/mL Pam3CSK4, or 1 μM CpG.

Statistical Analysis. Data are presented as mean \pm SD and statistical analyses were performed using Prism software.

- Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
- O'Neill LA, Bowie AG (2007) The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat Rev Immunol* 7:353–364.
- Biegert A, Söding J (2009) Sequence context-specific profiles for homology searching. *Proc Natl Acad Sci USA* 106:3770–3775.
- Söding J, Remmert M (2011) Protein sequence comparison and fold recognition: Progress and good-practice benchmarking. *Curr Opin Struct Biol* 21:404–411.
- Sadreyev RI, Tang M, Kim BH, Grishin NV (2009) COMPASS server for homology detection: improved statistical accuracy, speed and functionality. *Nucleic Acids Res* 37(Web Server issue):W90–W94.
- Couillault C, et al. (2004) TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. *Nat Immunol* 5:488–494.
- Carty M, et al. (2006) The human adaptor SARM negatively regulates adaptor protein TRIF-dependent Toll-like receptor signaling. *Nat Immunol* 7:1074–1081.
- Gay NJ, Keith FJ (1991) *Drosophila* Toll and IL-1 receptor. *Nature* 351:355–356.
- Rock FL, Hardiman G, Timans JC, Kastelein RA, Bazan JF (1998) A family of human receptors structurally related to *Drosophila* Toll. *Proc Natl Acad Sci USA* 95:588–593.
- Beutler B, Rehli M (2002) Evolution of the TIR, tolls and TLRs: Functional inferences from computational biology. *Curr Top Microbiol Immunol* 270:1–21.
- Roach JC, et al. (2005) The evolution of vertebrate Toll-like receptors. *Proc Natl Acad Sci USA* 102:9577–9582.
- Koonin EV, Aravind L (2002) Origin and evolution of eukaryotic apoptosis: The bacterial connection. *Cell Death Differ* 9:394–404.
- Turner JD (2003) A bioinformatic approach to the identification of bacterial proteins interacting with Toll-interleukin 1 receptor-resistance (TIR) homology domains. *FEMS Immunol Med Microbiol* 37:13–21.
- Chan SL, Mukasa T, Santelli E, Low LY, Pascual J (2010) The crystal structure of a TIR domain from *Arabidopsis thaliana* reveals a conserved helical region unique to plants. *Protein Sci* 19:155–161.
- Bernoux M, et al. (2011) Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and auto-regulation. *Cell Host Microbe* 9:200–211.
- Chan SL, et al. (2009) Molecular mimicry in innate immunity: Crystal structure of a bacterial TIR domain. *J Biol Chem* 284:21386–21392.
- Newman RM, Salunkhe P, Godzik A, Reed JC (2006) Identification and characterization of a novel bacterial virulence factor that shares homology with mammalian Toll/interleukin-1 receptor family proteins. *Infect Immun* 74:594–601.
- Cirl C, et al. (2008) Subversion of Toll-like receptor signaling by a unique family of bacterial Toll/interleukin-1 receptor domain-containing proteins. *Nat Med* 14:399–406.
- Yadav M, et al. (2010) Inhibition of TIR domain signaling by TcpC: MyD88-dependent and independent effects on *Escherichia coli* virulence. *PLoS Pathog* 6:e1001120.
- Zhang Y (2008) Progress and challenges in protein structure prediction. *Curr Opin Struct Biol* 18:342–348.
- Moult J, Fidelis K, Kryshchukovych A, Tramontano A (2011) Critical assessment of methods of protein structure prediction (CASP)-round IX. *Proteins* 10:1–5.
- Söding J (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21:951–960.
- Roy A, Kucukural A, Zhang Y (2010) I-TASSER: A unified platform for automated protein structure and function prediction. *Nat Protoc* 5:725–738.
- Peng J, Xu J (2011) Raptorx: Exploiting structure information for protein alignment by statistical inference. *Proteins* 10:161–171.
- Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: A case study using the Phyre server. *Nat Protoc* 4:363–371.
- Roche DB, Buenavista MT, Tetchner SJ, McGuffin LJ (2011) The IntFOLD server: An integrated web resource for protein fold recognition, 3D model quality assessment, intrinsic disorder prediction, domain prediction and ligand binding site prediction. *Nucleic Acids Res* 39(Web Server issue):W171–W176.
- Bazan JF, de Sauvage FJ (2009) Structural ties between cholesterol transport and morphogen signaling. *Cell* 138:1055–1056.
- Ginalski K (2006) Comparative modeling for protein structure prediction. *Curr Opin Struct Biol* 16:172–177.
- Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815.
- Pawlowski M, Gajda MJ, Matlak R, Bujnicki JM (2008) MetaMQAP: A meta-server for the quality assessment of protein models. *BMC Bioinformatics* 9:403.
- Wiederstein M, Sippl MJ (2007) ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35(Web Server issue):W407–W410.
- McGuffin LJ, Roche DB (2010) Rapid model quality assessment for protein structure predictions using the comparison of multiple models without structural alignments. *Bioinformatics* 26:182–188.
- Cristobal S, Zemla A, Fischer D, Rychlewski L, Elofsson A (2001) A study of quality measures for protein threading models. *BMC Bioinformatics* 2:5.
- Yang Y, Zhou Y (2008) Specific interactions for ab initio folding of protein terminal regions with secondary structures. *Proteins* 72:793–803.
- Worth CL, Gong S, Blundell TL (2009) Structural and functional constraints in the evolution of protein families. *Nat Rev Mol Cell Biol* 10:709–720.
- Shah PK, Aloy P, Bork P, Russell RB (2005) Structural similarity to bridge sequence space: Finding new families on the bridges. *Protein Sci* 14:1305–1314.

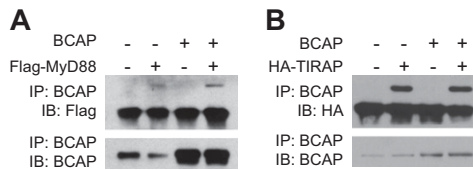


Fig. S3. Full-length BCAP associates with MyD88 and TIRAP. Lysates from 293T cells transfected as indicated with full-length *BCAP* and (A) *Flag-MyD88* or (B) *HA-TIRAP* as indicated were immunoprecipitated using a polyclonal anti-BCAP antibody followed by coupling to protein G beads. Precipitates were run on SDS/PAGE gels, transferred to PVDF membranes, and coprecipitation of MyD88 or TIRAP was assayed by immunoblotting with anti-Flag or anti-HA antibodies. Data represent two independent experiments. Untransfected 293T cells have endogenous BCAP that also associates with transfected MyD88 and TIRAP.

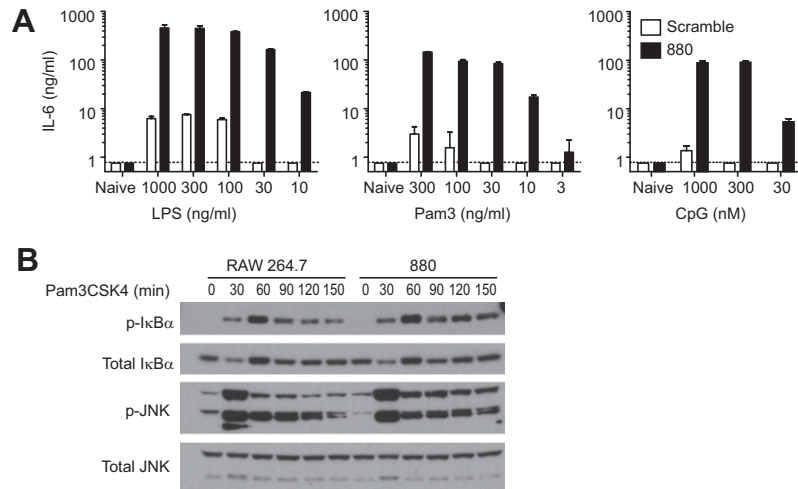


Fig. 54. BCAP represses cytokine production and signaling in RAW264.7 cells. (A) BCAP-silenced RAW264.7 cells were stimulated for 20 h with LPS, Pam3CSK4 (Pam3), or CpG at the indicated concentrations and secreted IL-6 was quantified by ELISA. (B) Control and BCAP-silenced RAW264.7 cells were stimulated with 100 ng/mL Pam3CSK4 and cell lysates were immunoblotted for phosphorylation of $\text{IkB}\alpha$ and JNK.

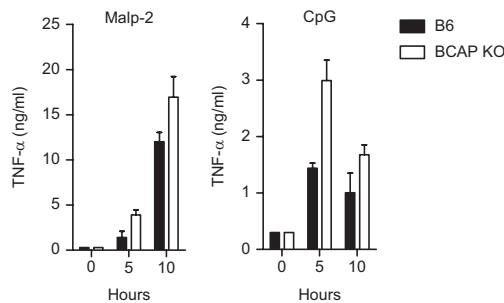


Fig. 55. BCAP KO macrophages secrete more TNF- α in response to TLR stimulation. BMDMs were stimulated as indicated with 100 ng/mL MALP-2 or 1 μM CpG as indicated. Secreted TNF- α was quantified by ELISA.

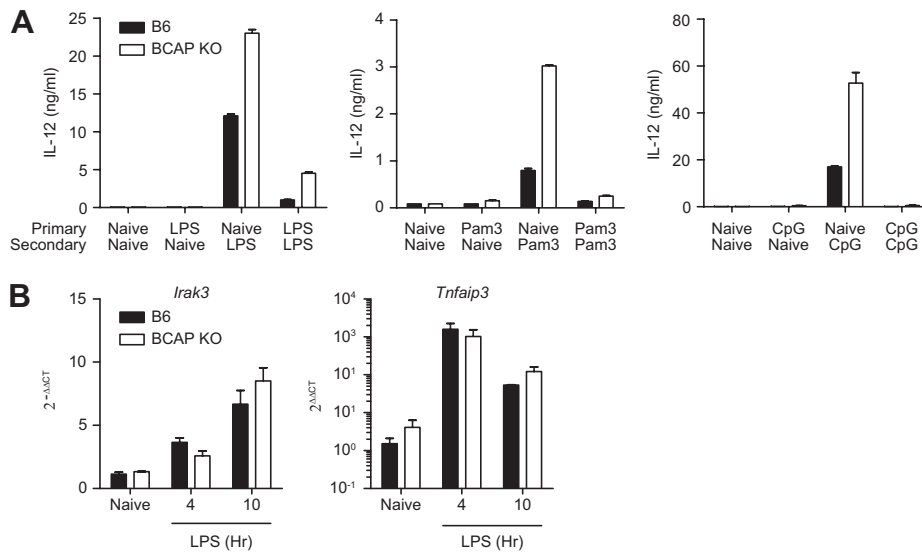


Fig. 56. BCAP is not involved in inflammatory tolerance in bone marrow macrophages. (A) BMDMs were stimulated for 24 h with 100 ng/mL LPS, 100 ng/mL Pam3CSK4 (Pam3), or 1 μ M CpG, washed twice with PBS, and restimulated for 20 h with LPS, Pam3CSK4, or CpG as indicated. Secreted IL-12 p40/p70 after secondary stimulation was quantified by ELISA. Data represent three independent experiments and are presented as the mean \pm SD. (B) Expression of *Irak3* or *Tnfaip3* relative to *Actin* as measured by quantitative rtPCR on cDNA from BMDM stimulated as indicated with 100 ng/mL LPS. Data represent the mean \pm SD of samples from macrophage preparations of two independent mice.

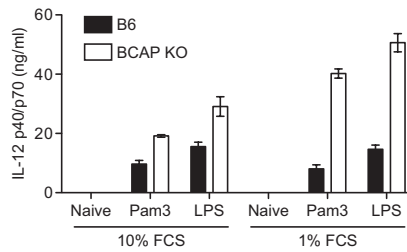


Fig. 57. Serum starvation enhances cytokine secretion by TLR stimulated BCAP KO macrophages. Macrophages were cultured in the presence of 1% or 10% FCS for 8 h, then stimulated for an additional 16 h with 100 ng/mL Pam3CSK4 (Pam3) or 100 ng/mL LPS. Cell-free supernatants were assayed for secreted IL12p40/p70 by ELISA. Data represent the mean \pm SD of duplicate cultures.

