SUMOylation in the *Drosophila* Innate immune response: Proteomics to immune signaling

A Thesis

Submitted in partial fulfillment of the requirements
Of the degree of
Doctor of Philosophy

By

Mithila Handu

20083022

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE

2014

CERTIFICATE

Certified that the work incorporated in the thesis entitled **"SUMOylation in the *Drosophila* Innate immune response: Proteomics to immune signaling"**, submitted by Mithila Handu was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or Institution.

Dr. G.S. Ratnaparkhi
Supervisor

Date:

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources that have not been properly cited or from whom proper permission has not been taken when needed.

Mithila Handu

20083022

Date:

This work is dedicated to my parents, brother and husband.

ACKNOWLEDGEMENTS

During the course of my PhD, many people have added value to my life both professionally and personally, and this thesis would be incomplete without expressing my sincere gratitude to all of them.

I would like to thank my thesis supervisor Dr. Girish Ratnaparkhi for giving me an opportunity to work with him and be tagged as the “first student”. Interactions and discussions with him over the years has indeed shaped my analytical thinking and scientific approach to experiments and writing.

I would like to thank my committee members for scientific discussions, feedback and encouragement over the years especially Dr. Jomon Joseph who helped me through my early experiments and troubleshooting. Pankhuri, graduate student from Dr. Joseph’s lab, assisted me in my first successful pulldown experiment. I am grateful to Prof. L.S. Shashidhara for being continually involved in my work and progress even though his busy schedule as the Head of Department. I also would like to thank Dr. Richa Rikhy for her interesting insights regarding experimental results during our regular lab meetings and Dr. Aurnab Ghose for his feedback and critique of my experiments and analysis.

Many contributed in materials and their expertise to make my PhD successful. Most of the results in the thesis would not have been possible without 529SU cells and the in-bacto SUMOylation plasmids. My sincere thanks go to Dr. M. Nie and Dr. Albert Courey for providing me with the same. I thank Dr. Harsha Gowda and the proteomics facility at Institute of Bioinformatics in Bangalore for carrying out the iTRAQ mass spectrometry. I thank Dr. Susanna Valanne and Dr. Mika Ramet for providing the FLAG-Relish-GSH_6 construct. I thank Sneha and other members of Dr. Anjan Banerjee’s lab for their help with intitial qRT-PCR standardizations. I thank Dr. Manjunath for frequent and enriching scientific discussions. I thank Dr. Rakesh Mishra, Dr. Shubha Govind and Dr. Girish Deshpande for evaluating and providing critique on my PhD work and reviewing this thesis.

I truly thank all my past and present colleagues from the Ratnaparkhi lab for their support, encouragement and discussions. I would start by thanking my first ever lab mates Vikram, Sakshi, Pooja and Umair who made the lab atmosphere scientifically enriching and

exciting during the intitial days. I thank the other lab members: Vallari, Senthil, Bhagyashree, Kriti, Ramya, Amar, Srija, Vijay, Aanchal, Janavi, Varsha and Lokesh who have extended their support whenever required and made fun working in the lab amidst all the work. I thank Vallari who is a great friend and has been there for me through a series of discussions about work or otherwise and been great support in the lab.

I positively thank all my IISER colleagues and friends without whom these years could have been dull and tiresome. The road to my PhD started with Rini, Payal, Aniruddh, Abhishek S., Ameya, Harsha, Anurag, Reshma, Sneha and Shraddha and I am thankful to them for making it very eventful. I cherish my discussions about biology, life and play with these friends. I look back to the good times in the lab and hostel. I thank my extra-curricular partners at dancing and sporting events and never-ending debates. Rini especially has been a dear friend and my partner in crime for a plethora of things during the last few years. Her coffee and humor would cure my tensions. Long discussions about my project with Rini and Harsha need a special mention here. I thank Vimal and Kanika for introducing me to the world of physics. I thank my senior and friend, Abhishek K. for his doses of guidance and encouragement. I thank my dear friends Nikhil and Abhijit for keeping me alive and enlightened of life and support outside IISER.

I am extremely grateful to the lab manager, instrument and chemical in-charges, and all other supporting staff at IISER without whom working in the lab would have been chaotic. I thank IISER for providing me with the infrastructure to carry out my research and a scientifically stimulating environment. It was a pleasure to be a part of a new and sprouting institute. I thank CSIR for my fellowship and Wellcome DBT for the project funding. I thank BDGP for the plasmids and various fly facilities for Drosophila lines.

I extend my heartfelt thanks to my family who have been an inspiration and driving force to pursue my goals. I am indebted to my parents Nancy and T.K. Handu who have motivated and supported me at every instance through this rollercoaster PhD ride. I am also thankful to my brothers Jeeten and Manish for caring for me on my low and high days. I acclaim my dear husband Jasmeet for being the best friend to share the continued frustrations and joys of PhD, while learning from his PhD anecdotes.

Thanks you all!

TABLE OF CONTENTS

List of figures

List of tables

SYNOPSIS

SUMOylation in the *Drosophila* Innate immune response: Proteomics to immune signaling

Name of the Student	:	**MITHILA HANDU**
Roll number	:	20083022
Name of the thesis advisor	:	Dr. Girish Ratnaparkhi
Date of Registration	:	18th August 2008

Indian Institute of Science Education and Research (IISER), Pune, India

INTRODUCTION

The innate immune response serves as the first line of defense to combat a wide-range of microbial pathogens in multi-cellular organisms. The process involves the recognition of specific pathogen associated molecular patterns (PAMPs) by cellular receptors, which trigger downstream effector responses (Akira et al., 2006). *Drosophila melanogaster*, like most invertebrates lacks most hallmarks of an adaptive immune response, having only the innate components. Its genome codes for immune signaling pathway proteins similar to those involved in mammalian innate immunity (Ferrandon et al., 2007), thus serving as an ideal model system to study innate immunity. *Drosophila* immune response can be broadly classified into three categories mediated through different cell types. These include the cellular responses mediated by the blood cells or hemocytes, melanization and wound repair. The humoral response includes the production of a wide range of anti-microbial peptides (AMPs) primarily by the fat body and also the hemocytes. Infection leads to the activation of the Toll and Immune deficient (IMD) signaling cascades in a pathogen-specific manner, which in turn leads to the translocation of NF-κB transcription factors Dorsal, DIF and Relish into the nucleus. These NF-κB factors act as transcriptional activators for defense genes (Anderson, 2000; De Gregorio et al., 2002; Hetru and Hoffmann, 2009; Tanji et al., 2007). The JNK, JAK-STAT, and Ras/MAPK pathways also play important roles in modulating the immune responses (Agaisse and Perrimon, 2004; Boutros et al., 2002; Chen et al., 2010; Delaney et al., 2006; Ragab et al., 2011).

Reversible post-translational modifications (PTMs) of proteins through addition and removal of molecular moieties prove essential to bring about rapid changes to external stimuli, without affecting transcription, protein synthesis and subsequent mRNA and protein turnover. Recently, covalent modification of proteins by **S**mall **U**biquitin-like **MO**difier (SUMO), has emerged as an important PTM mechanism in regulating transcription, translation, cell cycle, DNA repair and other cellular processes (Geiss-Friedlander and Melchior, 2007; Hay, 2005; Muller et al., 2001). The conjugation of SUMO to its target is a reversible process carried out by E1 activating enzyme, E2 conjugase and E3 ligases as in the case of ubiquitin. SUMO proteases are involved in SUMO protein maturation and deconjugation. The enzymes involved in the SUMOylation pathway are distinct from that used in ubiquitination and unlike ubiquitination, SUMOylation does not directly lead to proteasomal protein degradation. The attachment of SUMO to target proteins usually occurs at a consensus site ψKXE (where ψ is a hydrophobic residue, most often I, L or V) (Rodriguez et al., 2001; Sampson et al., 2001), however other SUMOylated proteins have been identified without this sequence and a few proteins with this sequence do not show any SUMOylation The presence of **S**UMO **I**nteraction **M**otifs (SIMs) on proteins provides a non-covalent binding site for SUMO binding (Kerscher, 2007; Zhu et al., 2008).

A number of proteins in the mammalian and fly immune signaling cascades are regulated by PTM mechanisms like phosphorylation and ubiquitination (Silverman et al., 2000; Zhou et al., 2005). In comparison, very few SUMOylated proteins in these pathways have been identified and their functions known. We chose *Drosophila* as a model system to better understand the role of SUMO modification in regulating different aspects of immune responses within the cell. *Drosophila* has a single SUMO isoform (Smt3), a single E2 (Ubc9) and multiple E3 ligases. An early evidence of the role of SUMO in *Drosophila* immunity came from a study by Govind and colleagues. They showed that in Ubc9 (SUMO conjugating enzyme) mutants there is constitutive activation of the Toll/NF-κB signaling leading to formation of melatonic tumors and increase in production of a few anti-microbial peptides (Chiu et al., 2005). Also, SUMO modification of Dorsal, the NF-κB transcription factor is important for Toll dependent anti-microbial peptide production (Bhaskar et al., 2002).

In this study, we characterize the involvement of SUMO modification in *Drosophila* innate immunity using both the fly and Schneider (S2) cells. Through our study, we show that the absence of SUMO affects the expression profile of NF-κB induced anti-microbial peptide response. We use stringent biochemical techniques in combination with quantitative mass

spectrometry to identify changes in the S2 cell proteome in response to LPS-mediated immune challenge, enriching and identifying SUMO conjugates. Bioinformatic analysis on a confident set of differentially expressed 858 proteins provides further insights into the prospective alterations in different cellular processes in response to immune challenge. We have also identified SUMO modification of two target proteins in the IMD pathway – Caspar and Jra. We also present some initial data on Caspar SUMOylation.

In summary, our study not only confirms few of the previously identified SUMO targets but also provides the first quantitative proteomics data in response to immune challenge in *Drosophila melanogaster*.

OBJECTIVES

We used *Drosophila* third instar larvae and S2 cells for our study of the immune response. S2 cells are polyclonal cell line derived from a hematopoietic lineage and have been extensively used as a model system to study innate immunity. We used high-throughput proteomics approach to identify changes in the SUMOylation states of proteins and pathways using iTRAQ mass spectrometry. A detailed bioinformatic analysis followed by biochemical validations allows us to gain insight into the mechanisms by which SUMO modification may affect *Drosophila* innate immune responses.

The specific aims of the study were as follows:

1. To identify changes in expression profile of the anti-microbial peptides post infection after SUMO depletion.

2. To identify known and unknown proteins with changes in their SUMOylation states in response to infection using quantitative mass spectrometry.

3. To validate SUMO modification of proteins and further identify the SUMO-lysine acceptor sites for further analysis.

4. To molecularly characterize the role of SUMOylation of selected target proteins, in order to understand their role in immune response.

RESULTS AND DISCUSSION

1. To identify changes in expression profile of the anti-microbial peptides post infection in *Drosophila* after SUMO depletion

The production of AMPs in response to immune challenge is a hallmark in *Drosophila* immune response. These AMPs have a distinct or over-lapping anti-microbial spectrum and different expression kinetics (Lemaitre and Hoffmann, 2007). In *Drosophila* larva, fat body is the primary site for AMP production, however hemocytes also contribute to AMP response within the organism. *Drosophila* S2 cells also produce anti-microbial peptides in response to immune challenge triggered either by lipopolysaccharide (LPS), peptidogylcan (PGN) or heat-killed bacteria. We have used AMP transcript levels as a read-out to understand the effects of SUMO knockdown using RNAi on NF-κB induced expression of AMPs.

Previous studies have shown that *Drosophila Ubc9* (SUMO conjugating enzyme) mutant larvae show constitutive Drosomycin response even in the absence of infection (Chiu et al., 2005). Studies in S2 cells have also shown that SUMO pathway components are required for Cecropin A1 and Drosomycin induction (Bhaskar et al., 2002). To more thoroughly understand its effects in a model cell line responsive to LPS treatment, we examined AMP expression in SUMO-knockdown S2 cells. The changes in the time dependent expression kinetics may suggest the role of SUMO in various transcriptional and post-transcriptional mechanisms that govern the stability of the AMPs or its effect on the interplay of the NF-kB pathways and their positive/ negative regulators.

We first checked for ability of S2 cells to produce AMPs in response to addition of crude LPS for variable time periods. Once it was established that these cells are well responsive, dsRNA-mediated knockdown of SUMO was carried out efficiently to approximately 95%. These cells were further infected with crude LPS and the response curve for different AMPs was monitored over 24 hrs using qRT-PCR. It was seen that the AMP expression profile in SUMO knockdown conditions differed from no RNAi control condition. A few AMPs showed an increase in their expression while others showed a decrease or no change. Decrease in global SUMOylation in knockdown conditions as shown in our results, may be affecting a variety of interplaying factors to maintain the AMP response and further studies are required to decipher the exact mechanisms governing such alterations.

2. To identify known and unknown proteins with changes in their SUMOylation states in response to infection using quantitative mass spectrometry

Very few immune related targets are identified to be modified by SUMO in both *Drosophila* and mammals (Bhaskar et al., 2002; Desterro et al., 1998; Gronholm et al., 2010). We performed biochemical pulldowns followed by mass spectrometry to identify immune responsive SUMOylated protein species within cells. We aimed at comparing the protein lists obtained from both *Drosophila* larvae and S2 cells immune responsive SUMO proteomes. We carried out tandem affinity purification using larvae expressing His-FLAG-SUMO-GG under Daughterless-Gal4 driver (Da-Gal4) post 2 hrs of infection. The larvae were pricked with a mixture of gram negative *Salmonella typhi* and gram positive *Micrococcus luteus* bacteria to induce a potent systemic response. The larval lysate was used for Ni-NTA purification in denaturing conditions followed by renaturation and affinity pulldown using anti-FLAG agarose beads to obtain only the SUMOylated proteins. The S2 cells we used were stably transfected with FLAG-SUMO-GG and HA-Ubc9 and we chose a 2.5 hr time point after considering the AMP expression data, post LPS induction for affinity purification using anti-FLAG agarose in native conditions with stringent detergent and elution conditions. This would help identify not only the SUMOylated proteins but also the tightly interacting proteins to obtain complexes that play a part in different immune response mechanisms. One or more proteins in the complex might be SUMOylated and SUMOylation might be essential for complex interactions. It was observed on immunoblotting of the enriched proteins that there is an increase in the levels of SUMOylation within a cell post immune challenge. Thus, we decided to undertake a quantitative mass spectrometry approach – isobaric **T**ags for **R**elative and **A**bsolute **Q**uantitation (iTRAQ) to quantify the levels of the purified proteins.

Pulldown experiments from larval extracts followed by mass spectrometry did not yield reproducible results due to excessive non-specific protein pulldowns in the negative controls. However, pulldown and mass spectrometry experiments from S2 cells were carried out in triplicates to obtain reliable results. ~1850 non-redundent proteins with a high false detection rate (FDR) of 1% were identified. A stringent cut-off was further applied and only proteins with iTRAQ ratios of <0.5 and >2.0 were considered to obtain a confident set of 858 proteins that showed significant changes in their levels post LPS treatment in S2 cells. Gene ontology analysis of these 858 genes showed 5% of the proteins to be involved in immune response related functions while various other cellular functions are well represented in the analysis, which implies that immune challenge induces global changes in the cell in order to

help it combat infection. Further DAVID and PANTHER analysis were used to identify various enriched pathways and cellular functions.

Comparison of the SUMOylation influenced immune proteome with the various genes that are transcriptionally regulated post immune challenge in *Drosophila* cell culture studies show almost overlap suggesting distinct regulation mechanisms controlling immunity (De Gregorio et al., 2001; Irving et al., 2001). Our resulting protein list shows many previously identified SUMOylated proteins and also provides a new Drosophila database of proteins that are potentially SUMOylated in a variety of cellular and developmental processes.

3. To validate SUMO modification of proteins and identify the SUMO-lysine acceptor sites for further analysis

It is well known that a very small fraction of a given protein is SUMOylated within the cell. To show physical SUMOylation of individual proteins has thus always posed to be a challenge in the field. To help increase the probability of detection of SUMO modified form of target proteins we used both in-bacto SUMOylation system which enables high expression of the protein and S2 cell overexpression tagged constructs and Ulp1 depletion.

The in-bacto system was a kind gift from the Courey Lab (Nie et al., 2009). This system has the advantage that bacteria lacks SUMO machinery and hence SUMO deconjugases and that the substrate proteins can be expressed in large amounts with detectable amount of SUMOylated species. We cloned a number of targets into a pGEX vector to aid enrichment of the proteins using GST pulldowns. We could successfully demonstrate SUMOylation of a couple of targets: 14-3-3, cdc42, and Jra to name a few. The SUMOylation was confirmed in immunoblots by detection of SUMO-modified species about +20kDa above the unmodified protein band. SUMO prediction softwares were used to identify SUMO-acceptor lysine sites in these proteins. A few of these proteins had the consensus lysine residue, while others did not show the consensus lysine residue consistent with previous observations that the consensus lysine residue is not a pre-requisite for SUMO modification (Geiss-Friedlander and Melchior, 2007).

We also used tagged protein constructs obtained from DGRC for validations in S2 cells. However, under the present set of conditions being used, we have yet not been successful in demonstrating SUMOylation of proteins in S2 cells. We are still in the process of refining and standardizing our experiments to demonstrate SUMOylation.

4. To molecularly characterize the role of SUMOylation of Caspar and Jra, in order to understand their role in the regulation of the immune response

Drosophila Caspar, homolog of mammalian Fas-associating factor 1 (FAF1) has been shown to negatively regulate IMD mediated immune response by affecting DREDD mediated cleavage of Relish (Kim et al., 2006). In mammals, FAF1 is an important player in regulating various processes along with immune response within the cell (Lee et al., 2012; Lee et al., 2013; Park et al., 2004; Zhang et al., 2012). Flies mutant for *Caspar* gene in *Drosophila* showed a constitutive expression of *Diptericin* AMP transcripts even in the absence of an immune challenge whereas, overexpression of Caspar lead to inactivation of the IMD pathway leading to no AMP production (Kim et al., 2006). *Drosophila* Jra is a transcription factor that is activated via the JNK pathway on pathogen attack and negatively regulates to attenuate the activation of the IMD pathway (Kim et al., 2007). It is the mammalian homolog of cJun and is involved in a variety of cellular processes within the cell.

In-bacto validation showed that both Caspar and Jra proteins are modified by SUMO. SUMO prediction software (SUMOsp) showed strong consensus sites for SUMOylation in both Caspar and Jra. We used protein sequence alignment to identify the lysines that are conserved among their orthologs in other organisms. These lysine residues were initially tested as putative sites for SUMO modification. In substitution mutations, we replaced lysine residues with arginine residues. Mutation of lysine 551 abolished the SUMO modified form of the Caspar. Mution in lysine 190 seems to be the primary site of SUMOylation in Jra, however it does not show complete loss of SUMOylation.

To characterize the role of Caspar in NF-κB response, we monitored Relish cleavage in response to LPS treatment in S2 cells by over-expressing the wild type and mutant forms of Caspar in S2 cells. We observe that there is a difference in the DREDD-mediated Relish cleavage between wild type and SUMO-deficient form of Caspar. Studies are in progress to understand how the cleavage is affected.

SUMOylation appears to be widespread in *Drosophila* proteome, with specific roles in immunity. Our study provides a basic framework in identifying the various SUMO enriched components involved in immunity, however extensive studies need to be done further to get a better understanding of the interplay of all the players in the immune response pathway regulation by SUMO.

BIBLIOGRAPHY

Agaisse, H., and Perrimon, N. (2004).The roles of JAK/STAT signaling in *Drosophila* immune responses. Immunological reviews 198, 72-82.

Akira, S., Uematsu, S., and Takeuchi, O. (2006).Pathogen recognition and innate immunity. Cell 124, 783-801.

Anderson, K.V. (2000). Toll signaling pathways in the innate immune response.Current opinion in immunology 12, 13-19.

Bhaskar, V., Smith, M., and Courey, A.J. (2002). Conjugation of Smt3 to dorsal may potentiate the *Drosophila* immune response. Molecular and cellular biology 22, 492-504.

Boutros, M., Agaisse, H., and Perrimon, N. (2002).Sequential activation of signaling pathways during innate immune responses in *Drosophila*.Developmental cell 3, 711-722.

Chen, J., Xie, C., Tian, L., Hong, L., Wu, X., and Han, J. (2010). Participation of the p38 pathway in *Drosophila* host defense against pathogenic bacteria and fungi. Proceedings of the National Academy of Sciences of the United States of America 107, 20774-20779.

Chiu, H., Ring, B.C., Sorrentino, R.P., Kalamarz, M., Garza, D., and Govind, S. (2005). dUbc9 negatively regulates the Toll-NF-kappa B pathways in larval hematopoiesis and drosomycin activation in *Drosophila*. Developmental biology 288, 60-72.

De Gregorio, E., Spellman, P.T., Rubin, G.M., and Lemaitre, B. (2001).Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. Proceedings of the National Academy of Sciences 98, 12590-12595.

De Gregorio, E., Spellman, P.T., Tzou, P., Rubin, G.M., and Lemaitre, B. (2002). The Toll and IMD pathways are the major regulators of the immune response in *Drosophila*. The EMBO journal 21, 2568-2579.

Delaney, J.R., Stoven, S., Uvell, H., Anderson, K.V., Engstrom, Y., and Mlodzik, M. (2006).Cooperative control of *Drosophila* immune responses by the JNK and NF-kappaB signaling pathways.The EMBO journal 25, 3068-3077.

Desterro, J.M., Rodriguez, M.S., and Hay, R.T. (1998). SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. Molecular cell 2, 233-239.

Ferrandon, D., Imler, J.L., Hetru, C., and Hoffmann, J.A. (2007). The *Drosophila* systemic immune response: sensing and signalling during bacterial and fungal infections. Nature reviews. Immunology 7, 862-874.

Geiss-Friedlander, R., and Melchior, F. (2007). Concepts in SUMOylation: a decade on. Nature reviews. Molecular cell biology 8, 947-956.

Gronholm, J., Ungureanu, D., Vanhatupa, S., Ramet, M., and Silvennoinen, O. (2010).SUMOylation of *Drosophila* transcription factor STAT92E. Journal of innate immunity 2, 618-624.

Hay, R.T. (2005). SUMO: a history of modification. Molecular cell 18, 1-12.

Hetru, C., and Hoffmann, J.A. (2009).NF-kappaB in the immune response of *Drosophila*.Cold Spring Harbor perspectives in biology 1, a000232.

Irving, P., Troxler, L., Heuer, T.S., Belvin, M., Kopczynski, C., Reichhart, J.M., Hoffmann, J.A., and Hetru, C. (2001).A genome-wide analysis of immune responses in *Drosophila*. Proceedings of the National Academy of Sciences of the United States of America 98, 15119-15124.

Kerscher, O. (2007). SUMO junction-what's your function? New insights through SUMO-interacting motifs. EMBO reports 8, 550-555.

Kim, L.K., Choi, U.Y., Cho, H.S., Lee, J.S., Lee, W.B., Kim, J., Jeong, K., Shim, J., Kim-Ha, J., and Kim, Y.J. (2007). Down-regulation of NF-kappaB target genes by the AP-1 and STAT complex during the innate immune response in Drosophila. PLoS biology 5, e238.

Kim, M., Lee, J.H., Lee, S.Y., Kim, E., and Chung, J. (2006).Caspar, a suppressor of antibacterial immunity in *Drosophila*. Proceedings of the National Academy of Sciences of the United States of America 103, 16358-16363.

Lee, J.J., Kim, Y.M., Jeong, J., Bae, D.S., and Lee, K.J. (2012). Ubiquitin-associated (UBA) domain in human Fas associated factor 1 inhibits tumor formation by promoting Hsp70 degradation. PloS one 7, e40361.

Lee, J.J., Park, J.K., Jeong, J., Jeon, H., Yoon, J.B., Kim, E.E., and Lee, K.J. (2013). Complex of Fas-associated factor 1 (FAF1) with valosin-containing protein (VCP)-Npl4-Ufd1 and polyubiquitinated proteins promotes endoplasmic reticulum-associated degradation (ERAD). The Journal of biological chemistry 288, 6998-7011.

Lemaitre, B., and Hoffmann, J. (2007). The host defense of *Drosophila* melanogaster. Annual review of immunology 25, 697-743.

Mabb, A.M., and Miyamoto, S. (2007).SUMO and NF-kappaB ties. Cellular and molecular life sciences : CMLS 64, 1979-1996.

Muller, S., Hoege, C., Pyrowolakis, G., and Jentsch, S. (2001). SUMO, ubiquitin's mysterious cousin. Nature reviews. Molecular cell biology 2, 202-210.

Nie, M., Xie, Y., Loo, J.A., and Courey, A.J. (2009).Genetic and proteomic evidence for roles of *Drosophila* SUMO in cell cycle control, Ras signaling, and early pattern formation.PloS one 4, e5905.

Park, M.Y., Jang, H.D., Lee, S.Y., Lee, K.J., and Kim, E. (2004). Fas-associated factor-1 inhibits nuclear factor-kappaB (NF-kappaB) activity by interfering with nuclear translocation of the RelA (p65) subunit of NF-kappaB. The Journal of biological chemistry 279, 2544-2549.

Ragab, A., Buechling, T., Gesellchen, V., Spirohn, K., Boettcher, A.L., and Boutros, M. (2011).*Drosophila* Ras/MAPK signalling regulates innate immune responses in immune and intestinal stem cells. The EMBO journal 30, 1123-1136.

Rodriguez, M.S., Dargemont, C., and Hay, R.T. (2001). SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. The Journal of biological chemistry 276, 12654-12659.

Sampson, D.A., Wang, M., and Matunis, M.J. (2001). The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. The Journal of biological chemistry 276, 21664-21669.

Silverman, N., Zhou, R., Stoven, S., Pandey, N., Hultmark, D., and Maniatis, T. (2000). A *Drosophila* IkappaB kinase complex required for Relish cleavage and antibacterial immunity. Genes & development 14, 2461-2471.

Tanji, T., Hu, X., Weber, A.N., and Ip, Y.T. (2007). Toll and IMD pathways synergistically activate an innate immune response in *Drosophila* melanogaster. Molecular and cellular biology 27, 4578-4588.

Zhang, L., Zhou, F., Li, Y., Drabsch, Y., Zhang, J., van Dam, H., and ten Dijke, P. (2012). Fas-associated factor 1 is a scaffold protein that promotes beta-transducin repeat-containing protein (beta-TrCP)-mediated beta-catenin ubiquitination and degradation. The Journal of biological chemistry 287, 30701-30710.

Zhou, R., Silverman, N., Hong, M., Liao, D.S., Chung, Y., Chen, Z.J., and Maniatis, T. (2005).The role of ubiquitination in *Drosophila* innate immunity.The Journal of biological chemistry 280, 34048-34055.

Zhu, J., Zhu, S., Guzzo, C.M., Ellis, N.A., Sung, K.S., Choi, C.Y., and Matunis, M.J. (2008). Small Ubiquitin-related Modifier (SUMO) Binding Determines Substrate Recognition and Paralog-selective SUMO Modification. Journal of Biological Chemistry 283, 29405-29415.

CHAPTER 1

INTRODUCTION

1.1 *Drosophila* Innate Immunity

In its life cycle, different stages of *Drosophila*, from embryo to adult, co-exist with micro-organisms in their natural environment. Insect larvae, in particular, feed and grow on decaying organic matter. Thus, for its survival the animal has developed robust methods for pathogen recognition and uses a multitude of immune effector responses to combat pathogenic bacteria, fungi, parasites and viruses. *Drosophila* lacks adaptive immunity and depends on sophisticated innate immune response mechanisms that are evolutionary conserved with vertebrates, making it a good model system to study innate immunity. The study of *Drosophila* immunity started with the first observation published by Hans Boman in 1972, where he proposed the presence of an inducible, non-specific immune response mechanism (Boman et al., 1972). Over the years, various genetic and molecular studies in *Drosophila* have helped understand these complex defense pathways elaborately.

Drosophila immune system relies on responses that are broadly classified into the cellular response and humoral response. The first-line of defense against evading pathogens includes external cuticle and epithelial barriers (gut, trachea). Once the pathogens breach these barriers and reach the hemocoel they encounter cellular defenses mediated by the hemocyte/blood cells. These defenses include clotting, melanization, encapsulation and phagocytosis. The molecular patterns on the pathogens lead to the activation of systemic humoral response through multiple cell signaling cascades and upregulates production of a battery of anti-microbial peptides or lytic peptides mainly by the fat body (Brennan and Anderson, 2004; Hoffmann, 2003; Hultmark, 2003). Interplay of all these robust mechanisms makes *Drosophila* highly resistant to microbes. The different *Drosophila* immune responses are represented in **Figure 1-1**.

1.1.1 Cellular Response

In *Drosophila*, cellular responses against evading pathogens are carried out by free-floating and sessile blood cells called hemocytes present in the hemolymph of the organism. These hemocytes can be divided into following 3 types based on their structural and functional properties: plasmatocytes, lamellocytes and crystal cells (Lanot et al., 2001). Plasmatocytes comprise of more than 90% of the larval hemocyte population. They are involved in phagocytosis of microbial pathogens and apoptotic cells via pathogen recognition and internalization. Lamellocytes are large, flat adherent cells and are responsible for encapsulation of large pathogens like parasitoid wasp eggs that cannot be phagocytosed easily. They are not present in the embryo or adult organisms and are rarely seen in healthy larvae. However, parasitoid wasp infection induce the differentiation of lamellocytes from hemocyte precursors (Rizki and Rizki, 1984; Russo et al., 2001; Sorrentino et al., 2002). The third category of cells is the non-phagocytic crystal cells that comprise about 5% of the total larval hemocyte population. They carry prophenol-oxidase (proPOs) crystals that are realeased in the hemolymph upon activation, causing melanization of the pathogens (Rizki et al., 1985; Soderhall and Cerenius, 1998).

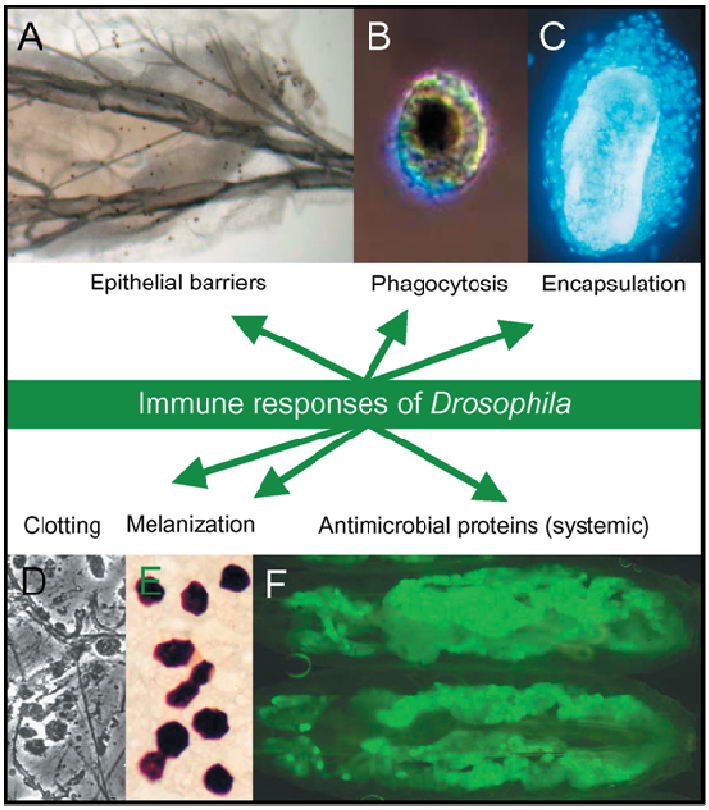


Figure 1-1: Overview of *Drosophila* Immune responses (Copyright PLoS Biology 2004, Govind and Nehm)
Drosophila has both the cellular and humoral immune responses. The cellular responses include phagocytosis (plasmatocytes), encapsulation (lamellocytes and clotting and melanization (crystal cells). The humoral response primarily includes the production of anti-microbial peptides from the fat body.

Phagocytosis

Phagocytosis by plasmatocytes is an immediate and efficient way to eliminate apoptotic cells and microbial infections (bacteria, yeast and viruses). The process of phagocytosis is evolutionary conserved among eukaryotic organisms. It involves recognition and binding of the pathogen to cell surface receptors which triggers a cascade of events; cytoskeleton re-organization, engulfment and internalization via vesicle trafficking, and eventually destruction of the engulfed target within phagosomes. These processes have been extensively studied using *Drosophila* embryonic-hemocyte derived cell line, Schneider (S2) cell line and mbn-2 cell line. These cells are highly phagocytic and an easy system to manipulate via dsRNA mediated knockdown (Clemens et al., 2000; Schneider, 1972). dsRNA mediated knockdown in S2 cells and other studies have shown that there are many similarities in pathogen recognition and engulfment between flies and mammals.

Phagocytosis is triggered by binding of a variety of pathogens and modified self ligands to scavenger receptor proteins. Several of these receptors have been identified and distributed into different classes: 1) Class B scavenger receptor proteins, Croquemort and Peste (Philips et al., 2005; Stuart et al., 2005), and Class C scavenger receptor protein, SR-CI (Ramet et al., 2001); 2) EGF-like repeat containing Nimrod family of proteins Draper, Eater and Nimrod C1 (Kocks et al., 2005; Kurucz et al., 2007); 3) the IgSF-domain protein Dscam (Watson et al., 2005); and 4) PGRP family proteins PGRP-LC and PGRP-SC1a (Garver et al., 2006; Ramet et al., 2002). These receptors are pathogen specific. The binding of the pathogen to the receptors is proposed to be aided by opsonization by thioester-containing proteins (TEPs). TEPs are a family of six secreted proteins, structurally related to the mammalian complement alpha2-macroglobulin family proteins. TEPVI or Mcr (macrophage-complement related) bind and enhance phagocytosis of fungi *C. albicans* while TEP II and TEP III bind and increase phagocytosis of *E. coli* and *S. aureus* respectively (Stroschein-Stevenson et al., 2006).

Various studies have identified the synergistic roles of various pathways for efficient pathogen internalization. These include actin remodeling, vesicle-mediated endocytosis involving clathrin and coat protein complexes (COPI and COPII) amongst other processes (Agaisse et al., 2005; Cheng et al., 2005; Stroschein-Stevenson et al., 2006; Stuart et al., 2007). Finally, the internalized pathogen is destroyed within the phagosomes by a cocktail of lysosomal enzymes, reactive oxygen species (ROS), nitric oxide (NO), and intracellular anti-microbial peptides.

Encapsulation

Encapsulation is a defense mechanism mediated by lamellocytes in which the foreign body is completely surrounded by these cells to insulate it from the host tissues and eventually destroyed. This defense mechanism has developed to protect the host from parasitoid wasps that lay their eggs into *Drosophila* larvae. The detection and recognition of the wasp eggs by plasmatocytes triggers a still poorly characterized signaling network which results in the proliferation and activation of large, flat and adhesive lamellocytes from lymphoid precursors (Rizki and Rizki, 1984; Sorrentino et al., 2002). The egg is eventually killed inside the capsule due to ROS production or toxic intermediates of the melanization cascade. The molecular mechanisms involved in this process are largely unknown.

Melanization and Coagulation

Melanization is a rapid reaction which leads to the de-novo synthesis and deposition of melanin at the site of cuticular injury or on the surface of an encapsulated parasitoid egg. Crystal cells are called so, due to presence of pro-phenol oxidase (pro-PO) crystals that are released on pathogenic mediated activation. On release, pro-PO is cleaved and converted to an active form by a serine protease. An enzymatic process eventually leads to the production of melanin and toxic intermediates that may contribute to the killing of the microbes (Nappi et al., 2009). The PO cascade must be strictly regulated to prevent excess melanization. This inhibition is mediated by Serpin27A, a serine protease inhibitor that prevents pro-PO cleavage (De Gregorio et al., 2002). Apart from crystal cells, *Drosophila* lamellocytes are also involved in melanization. They show the presence of specific Pro-PO3 as compared to pro-PO1 and 2 in crystal cells (Nam et al., 2008).

Cuticular wounding due to pathogen invasion is also repaired by rapid clotting and melanization mediated by crystal cells. The clot formed at the site of injury is composed of fibers trapping hemocytes and helps initiate wound healing (Bidla et al., 2005). Initial clot formation has been demonstrated to be independent of melanization, however melanization may be important in hardening the clot and wound closure. Hemolectin, a large hemocyte-specific protein is the most abundant protein in the clot (Lesch et al., 2007). Proteomic studies have identified a variety of proteins to be involved in clotting (Karlsson et al., 2004; Scherfer et al., 2004). Fondue, a hemolymph protein is not involved in initial clot formation but in cross-linking of the clot fibres (Scherfer et al., 2006), while transglutaminase (TG) aids in connection between bacterial surfaces and clot matrix (Wang et al., 2010).

1.1.2 Humoral response

In addition to the cellular responses, the pathogens entering the hemolymph trigger the rapid and transient expression of small anti-microbial peptides (AMPs), primarily from the fat body. Hemocytes also contribute to the production of these AMPs. AMPs are present at basal non-detectable levels in healthy individuals and their expression is induced on recognition of pathogen associated molecular patterns (PAMPs) by pathogen recognition receptors (PRRs). This triggers the activation of two major signaling cascades; Toll and IMD pathway. Both these pathways lead to the translocation of NF-κB (nuclear factor-κB) Dorsal

(Dl), DIF (dorsal-related immunity factor) and Relish (Rel) into the nucleus, thus controlling the expression of immune-responsive peptides and proteins.

Rel/ NF-κB family of proteins are conserved from *Drosophila* to humans. Rel/ NF-κB homologs have also been found in lower organisms like Cnideria, Porifera and single celled eukaryote *Capsaspora owczarzaki*. However the Toll-like receptors (TLRs) were not found in these single celled organisms. In the course of evolution, Rel/ NF-κB genes are lost in *C. elegans*. Rel/ NF-κB proteins across species are related through the presence of a highly conserved DNA binding/ dimerization domain called the Rel homology domain (RHD) which includes a NLS (nuclear localization signal) at the N-terminal. *Drosophila* Dorsal and DIF are 70 kDa proteins and their RHD is ~45% identical to mammalian proteins c-Rel, RelA and RelB. *Drosophila* Relish is a 100 kDa protein and is similar to mammalian p105 and p100. Dorsal and DIF factors are retained in the cytoplasm by binding to a 54 kDa protein Cactus with ankyrin repeats, a homolog of mammalian IκB factor. However, the other NF-κB protein Relish has a stretch of ankyrin repeats at the C-terminal which retains it in the nucleus. The ankyrin repeats associate with the RHD and mask the NLS to prevent it from going into the nucleus in unstimulated cells.

The NF-κB factors generally bind as homo or hetero dimers to their consensus sites in the promoter/enhancer region of target genes. This short stretch of DNA is called κB site. In *Drosophila*, studies show that DIF preferentially binds a sequence with three Gs followed by 4–5 AT-rich nucleotides [GGGAAA(A/T/G)(C/T)CC] while Relish prefers 4 Gs followed by a shorter AT-rich stretch [GGGGATT(T/C)(T/C)(T/C)]. A DIF/Relish heterodimer binding sequence was identified as GGGA(A/T)TC(C/A)C (Busse et al., 2007; Senger et al., 2004). The type of κB site present upstream an AMP gene thus determines its responsiveness to the Toll or IMD pathway.

Immune activation through NF- κB, leads to production and release of peptides with anti-microbial activity directly into the hemolymph. The AMPs produced, are small cationic, lytic peptides that are categorized into seven families. These include 4 Attacins, 4 Cecropins, 7 Drosomycins, 2 Diptericins, and one Drosocin, Defensin, and Metchnikowin (Imler and Bulet, 2005). The AMPs have distinct or overlapping anti-microbial spectrum and vary in their expression kinetics. Drosomycin and Metchnikowin are anti-fungal in nature (Fehlbaum et al., 1994; Levashina et al., 1995). Defensin is primarily active against Gram-positive

bacteria (Dimarcq et al., 1994), whereas Diptericin, Drosocin and Attacin are very effective against Gram-negative bacteria (Asling et al., 1995; Bulet et al., 1993; Wicker et al., 1990). Cecropins are have a broad range spectrum and eliminate both bacteria and fungi (Ekengren and Hultmark, 1999; Kylsten et al., 1990; Samakovlis et al., 1990). These AMPs are under tight post-transcriptional regulation that along with other factors might be involved in determinining the stability and expression kinetics of these peptides (Lauwers et al., 2009). The exact mechanism of action of these AMPs is still not fully understood.

Activation of target AMP genes by Rel/NF-κB is facilitated by multiple transcriptional co-activators. GATA binding factors may be important for AMP upregulation due to the presence of GATA sites in close proximity with Rel-binding sites upstream few of the response genes. Another such site is the Deaf1 binding site found upstream few of the AMP genes. It has been indicated that binding of DEAF1 (Deformed epidermal autoregulatory factor1) to these sites might affect the DIF/Dorsal mediated AMP gene expression. Other co-activators include TRAP80 that interacts with DIF, to induce transcription of *Drosomycin* and "POU" domain proteins that are potential regulators of DIF mediated *CecropinA* expression. Akirins are potential regulators of the Relish mediated *Diptericin* expression (Ganesan et al., 2011)

Toll Pathway

Toll pathway is an evolutionary conserved signaling pathway present in flies and mammals. The *Toll* gene was initially identified in a genetic screen to be crucial in *Drosophila* embryonic dorso-ventral patterning (Govind and Steward, 1991). Other components of the Toll signaling pathway were also identified through genetic screens for maternal-effect mutants defective in embryonic dorsal-ventral. Later, in mid-1990s two studies in *Drosophila* mbn-2 macrophage like cells and in adult flies demonstrated the role of Toll signaling in fly immunity. In 1995, Rosetto *et al.* first demonstrated the role of the Toll-1 receptor in pathogen-induced transcriptional activation of anti-microbial peptide *CecA1* (Rosetto et al., 1995). Further in 1996, Lemaitre *et al.* demonstrated that the Toll pathway mutants were susceptible to certain fungi but had normal resistance to bacteria, such as the Gram-negative *E. coli* (Lemaitre et al., 1996). These studies were followed by the identification and functional analysis of Toll signaling in mammalian immunity leading to considerable developments in the field of innate immunity. Apart from Toll-1 and Toll-7, seven other Toll-like receptors (TLRs) have been identified in *Drosophila* and are till date

shown to be involved in functions other than immunity. However, Toll-1 and Toll-7 are involved in anti-bacterial and anti-viral immune responses respectively (Nakamoto et al., 2012).

The Toll pathway is activated in response to fungal or Gram positive bacterial infections. In mammals, the Toll receptor directly binds to pathogens or pathogen-derived compounds however, in *Drosophila* these are recognized and bind to pathogen recognition receptors (PRRs). The binding leads to production of active Spatzle ligand that interacts with the Toll receptor to initiate a signaling cascade like in early development. The pathogen recognition receptors are categorized into different classes: peptidoglycan receptor protein-SA (PGRP-SA), PGRP-SD and gram negative binding protein1 (GNBP1) which recognize lysine-type peptidoglycan (PGN), a component of gram-positive bacterial membrane (Wang et al., 2008; Wang et al., 2006). GNBP3 also recognizes yeast (Mishima et al., 2009). On recognition, the PRRs initiate a proteolytic cascade which leads to the cleavage of N-terminal pro-domain converting inactive Spatzle into its active form by Spatzle processing enzyme (SPE). Proteolysis by SPE leads to a conformational change in Spatzle and exposes its determinants required for binding with the Toll receptor. It has been proposed that binding is achieved by two spatzle dimers, each interacting with one of the two Toll receptors, triggering a conformational change in the Toll to activate downstream signaling events (Arnot et al., 2010; Jang et al., 2006; Ligoxygakis et al., 2002; Weber et al., 2007).

There are multiple recognition - based serine protease cascades which lead to activation of SPE, and thus the Toll pathway. Two protease cascades leading to the activation of Gram-positive–specific serine protease (Grass) are initiated by both fungi and Gram-positive bacteria (Lysine-type peptidoglycan), which in turn activates another protease, spirit. Upstream of Grass, a modular serine protease (ModSP) plays an essential role in integrating signals from the recognition molecules GNBP 3 and PGRP-SA to the Grass-SPE-Spatzle cascade (Buchon et al., 2009). In addition, other serine proteases, namely spheroide and sphinx1/2, were identified in response to both fungi and Gram-positive bacteria (El Chamy et al., 2008; Kambris et al., 2006). A third protease cascade is mediated by Persephone, which is proteolytically activated by the secreted fungal virulence factor PR1 and Gram-positive bacteria specific peptidoglycan (El Chamy et al., 2008).

The activated Toll receptor interacts with the adaptor protein MyD88 via its intracellular TIR domains. MyD88 through its death domain (DD) recruits other DD-

containing proteins Tube and Pelle kinase (Sun et al., 2002). The MyD88-Tube-Pelle oligomeric complex by yet unknown mechanisms leads to phosphorylation of Cactus, the *Drosophila* IκB protein. Phosphorylation of Cactus causes the release of NF-κB proteins Dorsal and DIF that translocate into the nucleus and initiate the transcription of their target genes. In the absence of Toll activation, Dorsal and DIF are sequestered in the cytoplasm. Thus, degradation of Cactus induced by its phosphorylation is important for activation of Toll-dependent *Drosomycin* and *Metchnikowin* AMPs and other responsive genes (**Figure 1-2**). DIF and Dorsal seem to be redundant in their response at the larval stage, whereas DIF is sufficient to mediate Toll activation in adults (Manfruelli et al., 1999; Rutschmann et al., 2000).

Regulation of the Toll Pathway

Various regulators of the Toll pathway have been identified through RNAi screens and other studies. Serpin1 (Spn1), Pellino, deformed epidermal auto-regulatory factor-1 (DEAF1) and G Protein-coupled receptor kinase 2 (Gprk2) are positive regulators of Toll signaling. However, their mechanistic roles are still unknown (Fullaondo et al., 2011; Haghayeghi et al., 2010; Kuttenkeuler et al., 2010; Valanne et al., 2010). A recent study demonstrates the importance of Myopic (Mop) and Hepatocyte growth factor- regulated protein tyrosine kinase (Hrs) mediated endocytosis of Toll receptor in Toll pathway activation (Huang et al., 2010).

WntD is a negative regulator of the Toll pathway and represses it by preventing translocation of Dorsal into the nucleus (Gordon et al., 2005). Necrotic (nec), a member of the SERine Protease INhibitor (serpin) family is another negative regulator. *Nec* and *WntD* mutants show constitutive activation of the Toll pathway and expression of Drosomycin. It is believed that Nec might be regulating the Spatzle processing protease pathway. Other serpins have also been shown to regulate both Toll pathway and melanization responses in *Drosophila*.

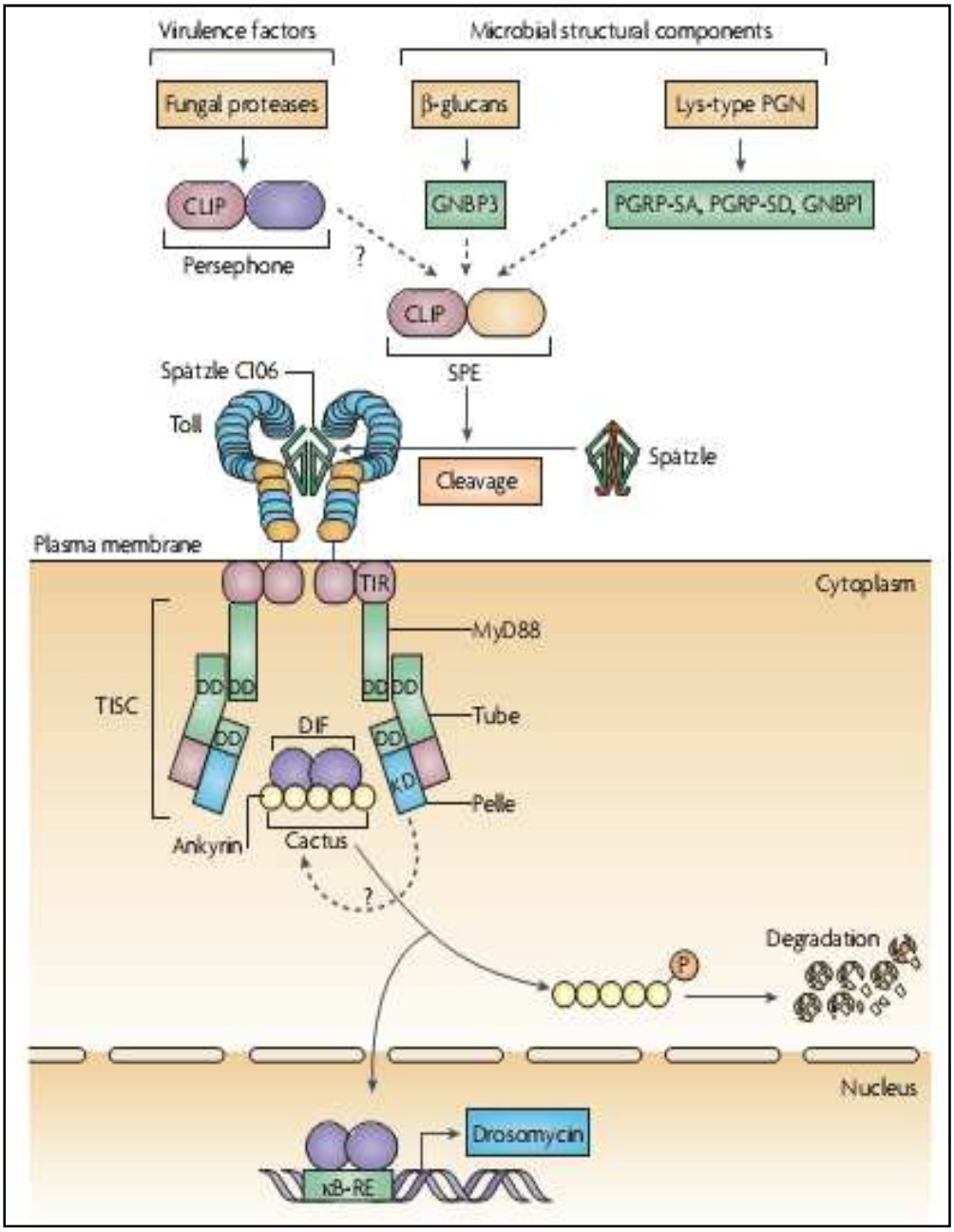


Figure 1-2: *Drosophila* Toll pathway (Copyright npg 2007, Hoffman et. al)
This figure shows the activation of the Toll pathway by a series of molecular players leading to translocation of Dorsal/ DIF into the nucleus and transcriptional upregulation of the immune responsive gene. The recognition of pathogen associated structures by pathogen recognition receptors triggers the proteolytic cleavage of Spaztle into its active form. Spatzle ligand binds to Toll receptor and causes it to dimerize. Further signaling leads to phosphorylation of Cactus which targets it towards ubiquitin-mediated degradation. This allows Dorsal/DIF which are otherwise retained in the cytoplasm by cactus to enter the nucleus and cause upregulation of AMP and other target genes.

Immune deficiency Pathway

Lemaitre et al. 1995 were the first to demonstrate the presence of a pathway alternate to the Toll pathway in immune response. They showed that the mutant *imd* (Black cells) flies had severely impaired production of most AMPs in response to a mixed infection by *E. coli* and *M. luteus*, however the anti-fungal AMP *Drosomycin* remained inducible (Lemaitre et al., 1995). Eventually, other components of the immune deficiency (IMD) pathway were discovered and shown to be activated in response to DAP-type (diaminopimelic acid) PGN. DAP-type PGN can be found in Gram-negative bacteria and also a few Gram-positive bacilli. The recognition proteins in the IMD pathway include the trans-membrane protein PGRP-LC and intracellular PGRP-LE. PGRP-LC encodes three alternative splice variants, PGRP-LCa, PGRP-LCx, and PGRP-LCy. The biological function of PGRP-LCy is yet unknown. All these proteins have similar cytoplasmic and trans-membrane regions but differ in their PGRP domains. PGRP-LCx binds polymeric DAP PGNs and PGRP-LCa binds monomeric PGN, also called the tracheal cytotoxin (TCT). Structural and biochemical studies suggest that PGRP-LCa and PGRP-LCx bind TCT as heterodimers. PGRP-LE encodes a protein without an obvious trans-membrane domain or signal sequence and localizes to the cytoplasm where it recognizes DAP PGN on intracellular pathogens, such as *Listeria*. It is also suggested that, through unknown mechanisms PGRP-LE is processed and released from the cells where it interacts with the extracellular domain of PGRP-LC and recognizes extracellular DAP PGNs like TCT (Akira et al., 2006; Ferrandon et al., 2007; Kleino and Silverman, 2014).

After the initial recognition of the bacterial components, dimerization or multimerization of these receptors propagate the signal via the cytoplasmic domain by binding to IMD. IMD is a death domain-containing protein and homologous to the mammalian receptor interacting protein (RIP1). IMD interacts with *Drosophila* homolog of the Fas-associated death domain (dFADD) via the death domain (DD). dFADD in turn recruits a caspase called Death-related ced-3/Nedd2-like protein (DREDD) to the signaling complex (Hu and Yang, 2000). Dredd is homologous to the mammalian caspase-8, and plays a dual role in IMD pathway activation. First, Dredd cleaves IMD (Paquette et al., 2010) and promotes its activation by exposing its interaction site for dIAP2, an E3 ubiquitin ligase. Second, DREDD is required for cleavage of the transcription factor Relish (Meinander et al., 2012; Stoven et al., 2003). dIAP2 along with E2 ubiquitin conjugating enzymes Uev1a, Bendless (Ubc13) and Effete (Ubc5) bring about IMD poly-ubiquitination. On K63 poly-ubiquination, IMD recruits the Transforming growth factor β-activated protein kinase 1

(TAK1)/ TAK1 binding protein 2 (TBP2) complex. Once activated, TAB2/TAK1 likely phosphorylates and activates the IκB kninase (IKK) complex. The *Drosophila* IKK complex consists of two subunits, the catalytic subunit called immune response deficient 5 (IRD5 or IKKβ) and the regulatory subunit Kenny or IKKγ (Erturk-Hasdemir et al., 2009). This complex plays an important role in phosphorylation and activation of Relish which is further cleaved. The endo-proteolytic cleavage of Relish results in a 68 kDa N-terminal fragment which includes the RHD and a 49 kDa C-terminal fragment containing the ankyrin repeat domain. The N-terminal RHD, translocates into the nucleus and initiates the transcription of its target genes, like AMPs etc. (Stoven et al., 2003). This cleavage of Relish requires the presence of DREDD. dIAP2 brings about the K63 polyubiquitination of DREDD which is required for its activation and probable role in direct cleavage of Relish (Meinander et al., 2012) (**Figure 1-3**)

Regulation of the IMD pathway

The only known positive regulator of the IMD pathway is Akirin. It is proposed to be working downstream of Relish pathway in modulating the immune response however its exact role is still unknown (Goto et al., 2008).

Negative regulation of IMD pathway occurs at various steps:

There are specific PGRPs which degrade PGN into non-stimulatory fragments to block immune activation. Further, PGRP-LF receptor competes with PGRP-LC for dimerization, the resulting dimer being inactive. PIRK is known to remove PGRP-LC from the membrane which blocks further relay of the pathway (Kleino et al., 2008). Various proteins affect the ubiquitination status of other proteins involved in the IMD pathway. These include SKPA/SLMB/DCUL1, dUSP36, CYLD, POSH, DNR-1 (Guntermann et al., 2009; Lee and Ferrandon, 2011; Thevenon et al., 2009; Tsuda et al., 2005). Caspar negatively regulates the IMD pathway by inhibiting Rel cleavage (Kim et al., 2006). The Ras/MAPK and JNK pathways also play an important role in limiting the IMD induced immune response (Ragab et al., 2011) .

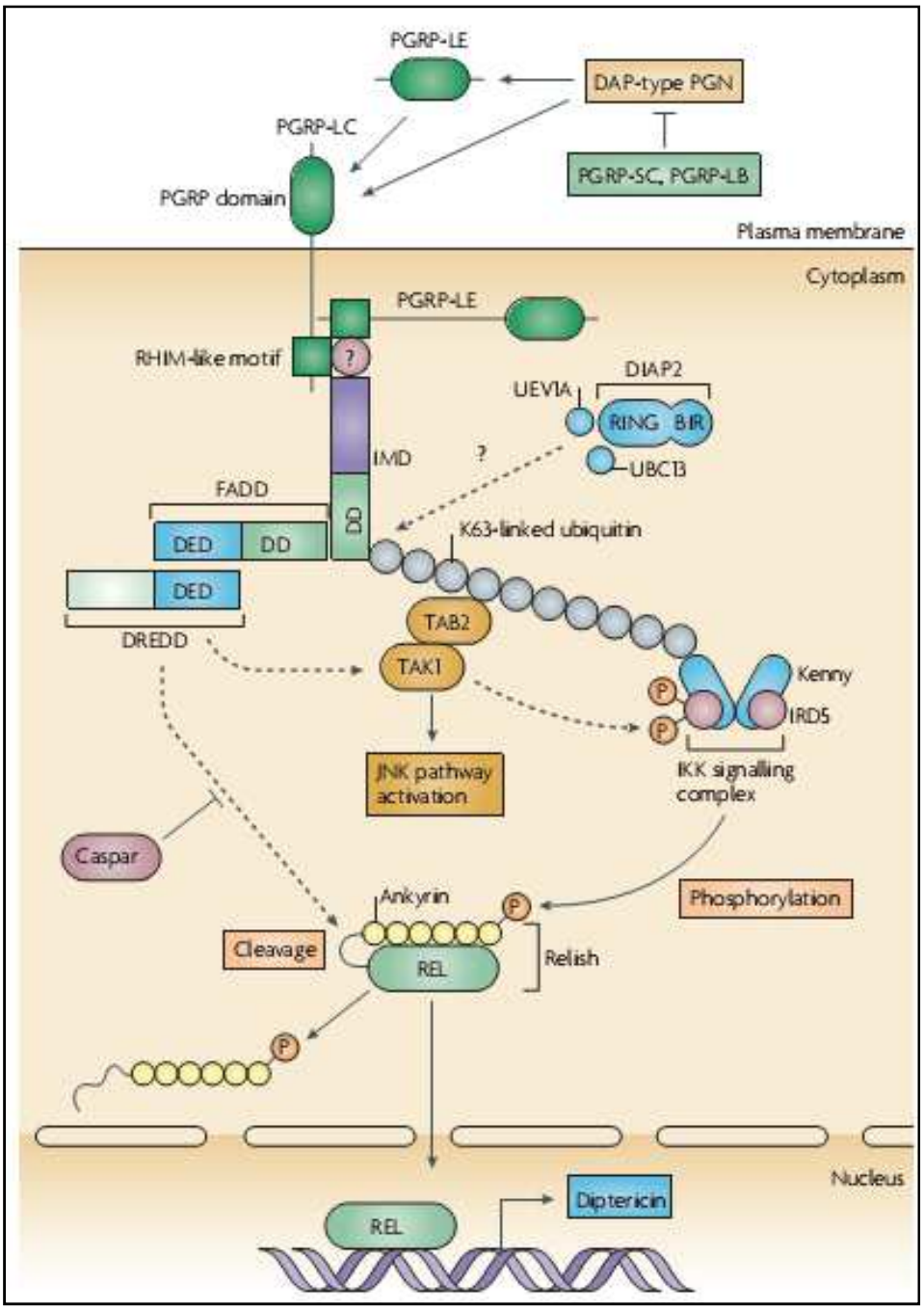


Figure 1-3: The IMD pathway (Copyright npg 2007, Hoffman et. al)

Activation of the IMD pathway by a series of molecular players leads to translocation of Relish into the nucleus and transcriptional upregulation of the immune responsive genes. Upon pathogen recognition by PGRPs, the intracellular domain of the PGRP is associated with IMD which further recruits and interacts with other adaptor proteins. This association eventually, leads to phosphorylation of Relish followed by its cleavage that is mediated through DREDD. The N-terminal RHD domain of Relish enters the nucleus to activate AMPs and other target genes. The IMD pathway also feeds into the JNK pathway which leads to activation of immune responsive genes and eventually also negatively regulates the IMD pathway.

JNK pathway

Apart from its usual function of regulating environmental stress and apoptosis, JNK pathway plays a role in hemocyte activation (Boutros et al., 2002; Silverman et al., 2003) and wound healing (Bidla et al., 2007). The JNK pathway components upstream of TAK1 are Mishappen (msn) and Wengen, which bind the *Drosophila* cytokine Eiger at the plasma membrane. These proteins have not been associated with the humoral immune response. In addition to stress, the JNK pathway can also be activated via the IMD pathway. The JNK arm of the IMD pathway is activated by TAK1-mediated signaling to Hemipterous which further phosphorylates basket (dJNK) and activates AP-1 signaling. Basket then activates the *Drosophila* Jun and Fos homologs; Jun-related antigen (Jra) and Kayak (Kay), which initiate transcription of target genes (Sluss et al., 1996). The targets of JNK pathway include the negative regulator puckered, cytoskeleton remodeling and hemocyte activation genes. In addition, the JNK pathway is involved in upregulation of AMPs during the early response to bacteria in adult flies (Delaney et al., 2006; Kallio et al., 2005), although this is somewhat controversial.

A putative ubiquitin E3 ligase Plenty of SH3 (Posh) regulates the signaling of both the IMD and the JNK pathways (Tsuda et al., 2005), possibly by regulating TAK1 stability and targeting it for proteasomal degradation. In addition, the IMD pathway negative regulator dUSP36 also suppresses JNK signaling (Thevenon et al., 2009) in IMD-dependent manner.

JAK STAT Pathway

JAK-STAT pathway proves to be essential for normal organismal development. In *Drosophila*the ligand that activates the pathway is a secreted molecule unpaired (Upd). Upd binds to the receptor domeless (dome) causing it to dimerize. On receptor dimerization, two molecules of *Drosophila* Janus Kinase hopscotch (Hop), are recruited to intracellular domain of dome where they trans-phosphorylate each other. The phosphorylated Hop further phosphorylates its associated receptor that now recruits STAT92E. STAT92E recruited at the receptor is also phosphorylated by Hop and forms a dimerize that translocates into the nucleus. On immune challenge, STAT92 activates Tep and Tot protein families in larval fat body cells. JAK-STAT pathway is also important in cellular immunity where it is involved in proliferation and differentiation of hemocytes upon infection in larva and adult flies (Agaisse and Perrimon, 2004; Shuai and Liu, 2003).

1.2 SUMO modification

1.2.1 Post-translational modifications

Post-translation modifications involve covalent attachment of small molecular adducts to target proteins in a reversible manner. They serve as a favorable regulatory system for rapid cellular responses to various external and internal factors due to their fast kinetics and localized modification as opposed to regulation at transcriptional and translational level. PTMs occur either by small chemical moieties, like phosphorylation, acetylation, glycosylation or small protein molecules like ubiquitin protein modification which was first discovered in 1975. Over time, a number of ubiquitin-like protein modifiers (ULMs) have been identified based on their similarity to ubiquitin (Herrmann et al., 2007).

ULMs do not share much sequence similarity with ubiquitin, however have a common three dimensional structure, the ubiquitin fold. Ubiquitin and ULMs have a diglycine residue at their C-terminal end, whose carboxyl group attaches to the lysine residue on the substrate via a covalent isopeptide bond. This conjugation is aided by a series of enzymes – activating, conjugating, and ligating enzymes, which are unique to each modification. Modification by ULM, affect localization, stability or interactions of substrates. There are 10 different ULMs identified in mammals most of which are also present in yeast. Each of these ULMs interacts with different target proteins to regulate different cellular processes within the cell (Cajee et al., 2012; Schwartz and Hochstrasser, 2003). These ULMs are listed in **Table 1-1**

Table 1-1: Ubiquitin like modifiers

	Ubiquitin-like modifiers	**Names**	**Ubiquitin sequence homology (%)**	**Major function**
1	**SUMO (Smt3)**	Small Ubiquitin like modifier	16-18	Transcription, cell cycle, DNA repair, chromatin modification

2	**NEDD8 (Rub1)**	Neural Precursor Cell-Expressed Developmentally Downregulated-8	58	Regulation of mitosis, regulator of ubiquitin–protein ligases , proteasomal degradation
3	**FUB1** **(MNSFβ)**	Fau Ubiquitin like protein (monoclonal nonspecific suppressor factor)	37	Immune regulation
4	**ISG15 (UCRP)**	Interferon stimulated gene-15 (ubiquitin cross-reactive protein)	29, 27 (2 domains)	Interferon response, Immune regulation
5	**APG12**	Autophagy defective-12	17	Autophagy
6	**URM1**	Ubiquitin-Related Modifier-1	12	unclear
7	**APG8 (LC3)**	Autophagy defective-8	10	Autophagy
8	**FAT10**	F-Adjacent Transcript-10	29,36 (2 domains)	Proteasomal degradation, apoptosis, interferon response
9	**UBL5 (HUB1)**	Ubiquitin-Like Protein-5 (Homologous to Ubiquitin-1)	25	mRNA splicing
10	**Ufm1**	Ubiquitin Fold modifier-1	16	unclear

1.2.2 SUMO

SUMO or Small Ubiquitin-like modifier is a 11 kDa polypeptide and shares approximately 18% similarity with ubiquitin (Bayer et al., 1998). *Smt3* was first identified in *Saccharomyces cerevisiae* as a suppressor of temperature sensitive mutation in a centromeric protein MIF2 (Meluh and Koshland, 1995). RanGAP1 was the first mammalian protein which was shown to be physically modified by SUMO, thus affecting its localization within the cell (Mahajan et al., 1997; Matunis et al., 1996). Subsequent studies have identified the role of SUMO in a wide range of cellular processes like cell cycle, DNA repair and transcription, etc from yeast to humans. (Zhao, 2007).

SUMO is well conserved among all eukaryotes. The unicellular eukaryote *S. cerevisiae* and lower invertebrates like, *C. elegans* and *Drosophila melanogaster* have only one *SUMO* gene, whereas the higher eukaryotes like plants and vertebrates show multiple SUMO paralogs (Geiss-Friedlander and Melchior, 2007). Both humans and mice have atleast three different SUMO isoforms SUMO1, SUMO2 and SUMO3. SUMO2 and SUMO3 share ~95% sequence similarity with each other and only ~50% similarity with SUMO1. SUMO-4 isoform present in mammals cannot be processed for substrate conjugation in vivo. It is believed that different SUMO isoforms have distinct roles in the cell. SUMO1 is predominantly in the conjugated form with the cell, however, SUMO2/3 are mostly present in the unconjugated free form which conjugates to protein substrates upon cellular stresses. Plants also have different SUMO isoforms, a few of which are specifically conjugated under stress conditions.

Drosophila SUMO (Smt3) was first identified in 1998 following which a number of studies identified other pathway components and their putative roles in the fly (Huang et al., 1998; Talamillo et al., 2008). It shares higher sequence homology with mammalian SUMO 2/3. *Drosophila* SUMO is present throughout development with particularly high levels of maternally contributed SUMO and its conjugation pathway proteins in early *Drosophila* embryos (Talamillo et al., 2008).

1.2.3 SUMO conjugation pathway

Modification by SUMO is a dynamic and reversible process (Hay, 2005). A very small fraction of the substrate protein is SUMOylated at any moment in the cell. SUMOylation process is catalyzed by 3 enzymes 1) E1 activating enzyme, 2) E2 conjugating

enzyme and 3) E3 ligating enzymes. SUMO is synthesized in an immature form with a C-terminal extension after a di-glycine residue. This pro-form is post-translationally processed to expose the di-glycine motif through the hydrolase activity of SUMO protease, Sentrin specific proteases (SENPs). The mature SUMO can then form a thioester bond with the catalytic cysteine of the E1 activating enzyme complex SAE1/SAE2 in an ATP dependent reaction (Desterro et al., 1999; Gong et al., 1999; Johnson et al., 1997). The activation step is followed by conjugation of the activated SUMO to the E2-conjugating enzyme, Ubc9 through another thioester bond. Finally, SUMO is transferred from Ubc9 to the lysine residue of the substrate forming an isopeptide bond between SUMO and the ε-amino group of a lysine residue within the substrate (Desterro et al., 1997; Johnson and Blobel, 1997). E3 ligases are not absolutely essential for this conjugation; however they might be important to provide substrate specificity. This SUMO conjugation can be reverted by isopeptidase activity of the SUMO protease that hydrolyzes the isopeptide bond, causing removal of SUMO from its substrates (Geiss-Friedlander and Melchior, 2007; Hay, 2001; Melchior, 2000) (**Figure 1-4**).

The SUMOylation pathway has only one conjugatisng enzyme, Ubc9 (E2). It is understood that E2 mediates conjugation through binding to both the sustrate via their sumo consensus motif and SUMO, via a thioester bond. This brings both the SUMO and substrate in close proximity aiding in SUMO modification of the target protein (Hay, 2005). In vitro studies have shown that E3 ligases might not be important for efficient SUMO modification of the substrate. However, in vivo Ubc9 mediated interaction might not be sufficient for substrate modification, thus depending on a class of enzymes called E3 ligases. E3 ligases help in accelerating the SUMO modification reaction. Studies have identified four different classes of E3 ligases : 1) Sp RING family including Siz/PIAS family members, 2) nucleoporin Ran binding protein2 (RanBP2) or Nup358 (**Figure 1-5**) , 3) polycomb group protein 2 (Pc2), 4) TOPORs.

More recently, it has been shown that mammalian and yeast Ubc9 which are more than 80% identical are modified by SUMO at lysine 14 and 153 respectively (Knipscheer et al., 2008). Lysine 157 in yeast is another identified minor SUMO site (Ho et al., 2011). The lysine 14 residue is homologous in both yeast and mammals, however, the reason for the discrepency in the SUMO acceptor lysine is still not understood. This modification of Ubc9 helps generate an additional binding surface that recognizes targets through their SUMO

interacting motif (SIM) and brings them in close proximity to SUMO which increases SUMO modifications of proteins like in the case of transcription factor Sp100 (Knipscheer et al., 2008). Thus, the SUMOylation of Ubc9 might prove to be an alternative to E3 ligase activity. However, this is not always true. SUMOylation of Ubc9 leads to a decrease in SUMOylation of RanGAP1 whereas shows no change in SUMOylation of HDAC4, TDG etc (Knipscheer et al., 2008). E3 ligases like RanBP2, Pc2, TOPORs and PIAS proteins are also modified by SUMO probably to aid in SUMO modification of substrate proteins (Wilkinson and Henley, 2010).

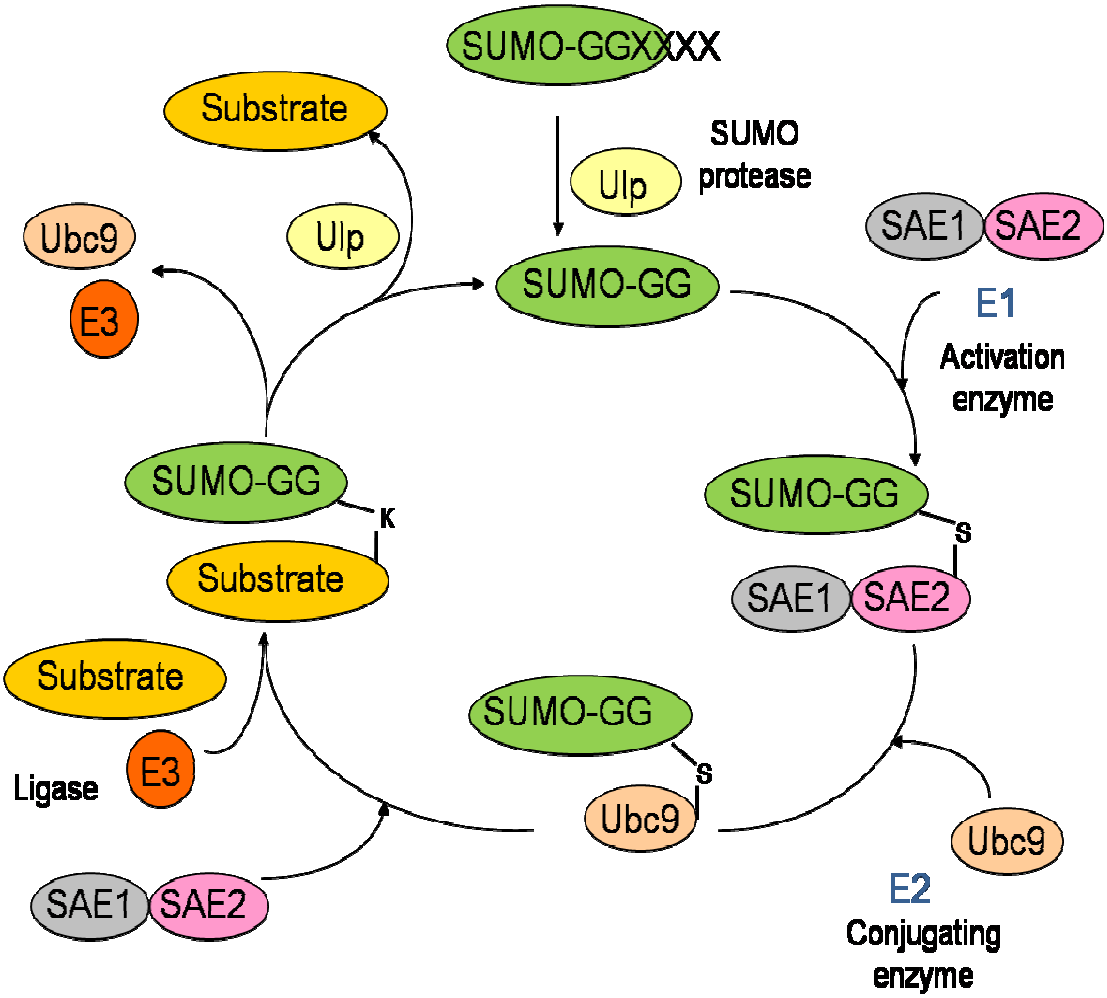


Figure 1-4: SUMO conjugation pathway
The attachment of SUMO to its substrate occurs by a sequence of enzymatic reactions in a reversible manner. The immature form of SUMO is cleaved by the SUMO protease to expose the di-glycine at the C-terminal end. The cleaved mature form of SUMO further transferred to the substrate via E1, E2 and E3 enzymes as described in the text. SUMO proteases can cleave the isopeptide bond releasing free SUMO making the cycle reversible.

A family of SUMO specific cysteine proteases identified in different organisms is involved in deconjugation of SUMO from its substrates. These proteins are Ulp1 and Ulp2 in yeast and *Drosophila*, and six sentrin specific proteases (SENPs) in mammals. These proteases have both isopeptidase activity required for breaking the isopeptide bond formed between SUMO and the substrate specific lysine, and C-terminal hydrolase activity that is needed for the maturation of SUMO proteins. These deSUMOylating enzymes vary in their localization, substrate specificity and their activity in maturation and isopeptide cleavage. Many of the SUMO proteases have shown to be associated with the nuclear periphery as shown in yeast, *Drosophila*, and mammals. In *Drosophila*, SUMO deconjugation function is a more prominent role of Ulp1 than SUMO maturation. Ulp1 is usually associated with the nuclear pore complex, and is believed to deSUMOylate the modified proteins before they enter the nucleus (Smith et al., 2004).

SUMO is usually attached to the target protein at a consensus ψKXE site, where ψ is a large hydrophobic amino acid, K is lysine, X is any amino acid and E is glutamic acid (Rodriguez et al., 2001). The variations to the consensus site include phosphorylation influenced SUMO motifs (PDSM) which include the SUMO consensus motif followed by a phosphorylated serine and proline residue (ΨKXEXXpSP) and negatively charged amino acid- dependent SUMO motifs (NDSMs). Both these extensions contribute additional negative charge next to the SUMO consensus motif which may be facilitating SUMO modification of the substrate. However, not all consensus sites are SUMOylated and SUMOylation of some proteins occurs at alternate acceptor sites (Geiss-Friedlander and Melchior, 2007).

Various observations have shown that only a very small proportion of a given protein is SUMOylated at a given time and this is sufficient for maximal effect on downstream processes. This has been described as “The SUMO Enigma” and various explanations have been provided for this phenomenon (Hay, 2005). SUMO has also been shown to form polySUMO chains in mammals and yeast at consensus sites present within the SUMO protein. In budding yeast the consensus lysines in SUMO are K11, K15 and K19 and in mammals K11. SUMO3 does not show the presence of any consensus site but in vitro it is capable of making polySUMO chains at non-consensus motifs. *Drosophila* SUMO does not show any consensus motifs and has not been shown to undergo polySUMOylation (**Figure 1-**

6). In vitro studies have shown sites other than the consensus sites capable of forming polySUMO chains.

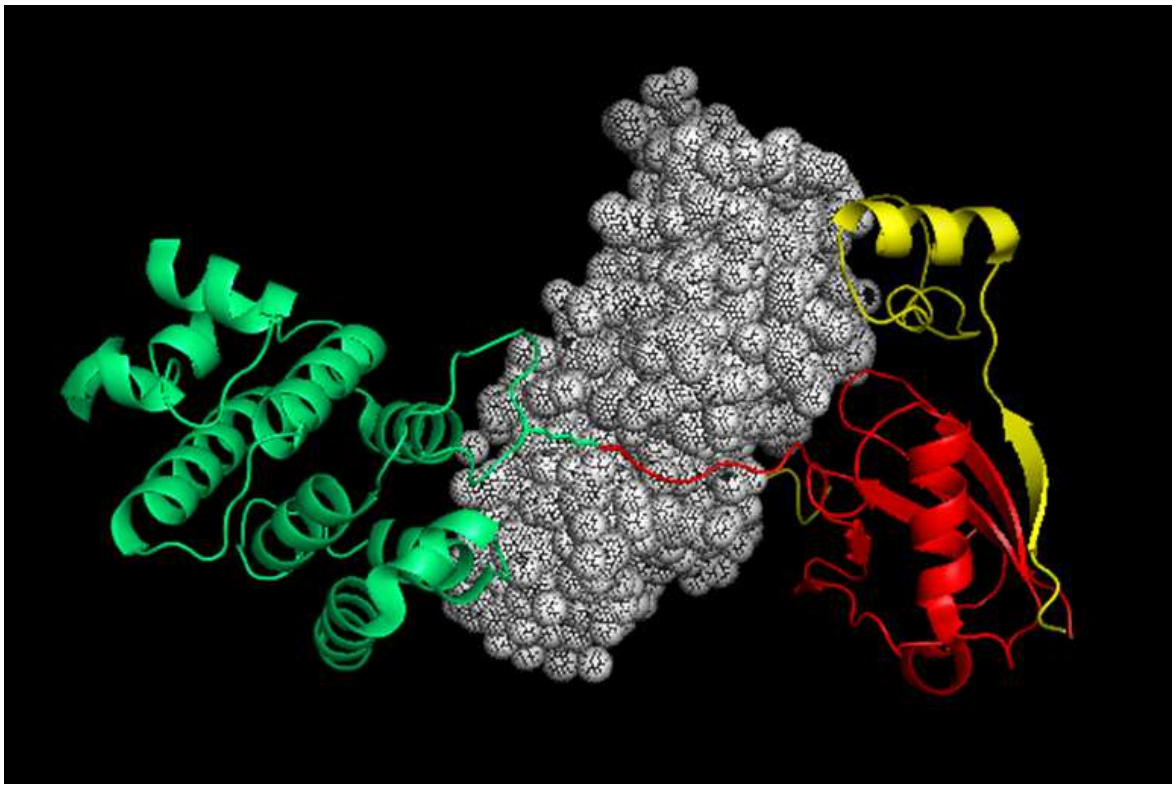


Figure 1-5 : Crystal structure of SUMO conjugated to RANGAP1 by the E2 (Ubc9), with Nup358 acting as an E3 ligase.
The structure shows the interaction between the mammalian proteins RANGAP1 (SUMO substrate) and Ubc9. It shows the cleft/tunnel in Ubc9 that holds the C-terminal GG tail of SUMO for conjugation with the lysine side chain of RANGAP1 (Reverter and Lima, 2005). It can be clearly seen that Ubc9 helps interaction between the substrate and SUMO to bring about SUMO modification. The complex has Ubc9 (Grey), SUMO (Red), RANGAP1 (Green) and Nup358 (yellow). The figure was generated from co-ordinates of 1Z5S, procured from RCSB Protein Data Bank using PyMol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC)

Figure 1-6 : SUMO consensus sequences
Human SUMO1/2 have a consensus K11 lysine acceptor site which is conserved in yeast Smt3p but not seen in SUMO1 allowing the formation of polySUMO1/2 chains. *Drosophila* dSmt3 does not show any consensus lysines. Yeast Smt3p has consensus motif lysines at K11 and K19, other than K15.

PolySUMO chains have been identified in mammals and yeast. The functional importance of polySUMOylation is currently under extensive study. In budding yeast, a mutant defective in forming polySUMO chains is viable and exhibits defects only in sporulation. PolySUMO chains are assumed to be important in meiosis in both yeast and mammals. Another growing importance of polySUMO chains is that it directs the target protein towards ubiquitination and degradation with the help of special class of ubiquitin ligases (Ulrich, 2008).

SUMOylated proteins can bind to additional interacting partners via novel non-covalent attachments. This is aided by the presence of SUMO interaction motif (SIM) in SUMO-recognizing proteins. Different SIM motifs have been identified till date. The SIM motif primarily comprises of a hydrophobic core, consisting of 3–4 aliphatic residues, often juxtaposed to a negatively charged (acidic) cluster of amino acids. These residues help establish a domain that helps in electrostatic interactions with SUMO. One such motif includes a series of hydrophobic residues, mainly composed of valine and isoleucine amino acids, surrounded by acidic and hydrophilic amino acids (V/I-X-V/I-V/I)(Hecker et al., 2006; Song et al., 2004). This motif is present in several proteins along with activating enzyme complex (Uba2/ Sae1) and the E3 ligases PIASX and RanBP2. Sometimes, serine and threonine residues are placed adjacent to the hydrophobic SIM domains. Phosphorylation of these residues helps impart negative charge on the SIMs that then interact with the lysine residues of SUMO. A number of SUMOylated substrates have been shown to contain SIMs, suggesting its importance in mediating SUMOylation by bringing SUMO and the substrate

together before the covalent binding. However, SIMs are also present in non-SUMOylated substrates. This helps them interact with the SUMO modified proteins. This function might be essential in forming active complexes to carry out specific functions (Kerscher, 2007). PolySUMOylation of substrates might help increase the SUMO-SIM interactions thus proving useful in cellular functioning.

1.2.4 Interplay between SUMO and other post translational modifications

There have been instances of crosstalk between different PTMs that may be agonistic or antagonistic to each other. Initially it was believed that SUMO acts to block ubiquitin function and prevents proteins from ubiquitin mediated proteasomal degradation eg. PCNA (proliferating cell nuclear antigen) has been found to be modified by either SUMO or ubiquitin at the same lysine residue, each leading to selection of different pathways in DNA replication stress (Papouli et al., 2005). Later discoveries proved these two modifications to sometimes act sequentially and in synergy with each other, as in the case of NEMO (NF-κB essential modulator), a regulatory subunit of mammalian IKK. SUMOylation of NEMO is required for its phosphorylation-dependent ubiquitination within the nucleus. Ubiquitination of NEMO further leads to its nuclear export and helps formation of an active IKK complex required for signaling (Huang et al., 2003).

Recently, a class of proteins known as SUMO-targeted ubiquitin E3 ligases (STUbLs) was identified to interact with polySUMO chains, subjecting the polySUMOylated target proteins to ubiquitin mediated degradation (Uzunova et al., 2007). The hallmark of these STUbLs is the presence of multiple SIMs for recognition of polySUMOylated targets. This changed the whole SUMO paradigm from competitor of ubiquitin into enhancer of ubiquitin chain formation and promoter of proteasomal degradation in some cases (Perry et al., 2008). Some of the STUbLs identified so far include Slx5/Slx8 heterodimer in *S. cerevisiae* while RNF4 (ring finger protein 4) in mammals. The roles of STUbLs have been shown in DNA damage and arsenic-induced degradation of the PML(Ulrich, 2008).

There have been instances of cross-regulation between ubiquitination and SUMOylation where they modify the components of the other's conjugation pathway. For instance, SUMOylation of E2-25K (E2 ubiquitin enzyme) inhibits its ubiquitin conjugating property and Parkin (ubiquitin E3 ligase) ubiquinates RanBP2 (SUMO E3 ligase) promoting its degradation (Um and Chung, 2006). Certain proteins have been identified to act as both

ubiquitin and SUMO ligases under different conditions. TOPORS (topoisomerase 1 interacting protein) have the ability to both SUMOylate and ubiquitinate p53 (Rajendra et al., 2004; Weger et al., 2005). It has N-terminal RING finger domain which helps recruit ubiquitin E2 enzymes thus acting as ubiquitin ligase and a central region that comprises of a SIM and confers the SUMO ligase activity. Phosphorylation of TOPORs at different serine residues by different kinases activates it SUMO or ubiquitin ligase property (Praefcke et al., 2012).

SUMOylation and acetylation may occur on the same lysine residue of a protein and act antagonistically to each other. A "SUMO-acetyl switch" motif has been identified which targets both SUMO and acetyl group to the substrate protein (Anckar and Sistonen, 2007). This motif consists of a SUMO consensus motif flanked by a C-terminal proline residue. Phosphorylation has been shown to behave as a switch between SUMOylation and acetylation as in the case of transcription factor MEF2A (myocyte-specific enhancer factor-2A). Phosphorylation enhances SUMO modification of MEF2A, and therefore promotes its repressed state whereas dephosphorylation leads to a switch from SUMOylation to acetylation. Certain HDACs which function in removal of the acetyl group from substrates also behave as SUMO E3 ligases. In addition, some HDACs are themselves SUMOylated and some others enhance SUMOylation of their targets post deacetylation (Wilkinson and Henley, 2010). A recent study by Shih and colleagues has shown Ubc9 to be acetylated at lysine 65. Ubc9 acetylation prevents its binding to substrates with a negatively charged amino-acid dependent SUMOylation motif (NDSM) thus inhibiting SUMOylation of those substrates as opposed to substrates with a normal consensus motif (Hsieh et al., 2013).

As mentioned earlier, PSDMs which include phosphorylation of amino acid residues around the basic consensus site facilitate SUMO modification of the protein as seen in HSF-1 (heat shock factor-1). There are number of kinases and phosphatases which are directly regulated by SUMO modification. Various SUMO E3 ligase activities are regulated via their phosphorylation eg. RanBP2, PIAS1etc. SUMO1 itself has been shown to be phosphorylated and this is conserved throughout eukaryotes. The functional relevance of this modification has not been yet reported (Matic et al., 2008).

This complex interplay between different post translational modifications provides a way of fine-tuning the different cellular processes. Also, with very few proteins involved in the SUMO pathway additional modulation of the conjugation pathway components by other

PTMs helps bringing about greater specificity in substrate selection and functional consequences.

1.3 SUMO and immunity

Several studies have implicated the role of SUMO in immunity. SUMO plays an important role in regulating the mammalian Tl/NF-κB pathway. Mammalian IκB is modified by either ubiquitin or SUMO-1 at lysine 21, the latter preventing the protein from ubiquitin-mediated proteolysis (Desterro et al., 1998). Over-expression of SUMO-1 thus inhibits IκB mediated transcriptional activation of immune genes in response to pathogenic attack. In *Drosophila*, the IκB homolog Cactus has not shown to be SUMOylated yet but a yeast two-hybrid screen showed that the *Drosophila* Ubc9 homolog, Lesswright (lwr), physically interacts with Cactus (Bhaskar et al., 2000).

In *Drosophila*, the SUMO conjugation machinery influences anti-microbial response, phagocytosis and hemocyte proliferation. Dorsal, *Drosophila* NF-κB is SUMOylated at lysine residue 382. In *Drosophila* S2 cells, over-expression of SUMO and Ubc9 showed an increase in Dorsal translocation into the nucleus even in the absence of an immune-challenge leading to transcriptional upregulation of anti-microbial peptide *Cecropin A1*(Bhaskar et al., 2002). However, another study shows an increase in Toll mediated *Drosomycin* upregulation in *Ubc9* mutant larvae (Chiu et al., 2005). This discrepancy in outcomes likely reflects the complex differences in the two biological systems with respect to their immune responses. The *Ubc9* mutant larvae show over-proliferation of hemocytes and the presence of melanotic tumors in the hemolymph. These hemocytes showed increased Dorsal transcription factor in the nucleus suggesting that *Ubc9* mutation activates NF-κB signaling (Chiu et al., 2005; Huang et al., 2005). Recent study in adult flies showed that *Ubc9* heterozygous mutants are more susceptible to infection with *E.coli* as compared to wild type control flies. This can be attributed to their inability to clear the bacteria from their body (Fukuyama et al., 2013).

Chapter 2

Aims of the thesis

A number of proteins in the mammalian and fly immune signaling cascades are regulated by post-translational modifications like phosphorylation and ubiquitination. Recent work has indicated that components of the SUMO pathway are involved in modulating immune responses in *Drosophila* and mammals, however not much is understood about the proteins that are SUMOylated and the molecular mechanisms behind SUMO-mediated regulation of these immune related proteins. In our study, we took a genome-wide approach to understand the role of SUMOylation in regulating the different defense strategies in *Drosophila* innate immune response. We used cultured *Drosophila* Schneider (S2) cells, which have been well characterized to respond to LPS which induces a potent immune response in these cells. It is known that there is significant conservation in the mammalian and *Drosophila* innate immune response pathways, and hence through this study we were hoping to acquire some information regarding the SUMO-mediated regulation of the mammalian immune signaling.

The specific aims of the study were as follows:

1. To identify changes in gene expression of the anti-microbial peptides during LPS induction in *Drosophila* Schneider cells after SUMO depletion.
2. To identify known and unknown proteins with changes in their SUMOylation states in response to infection using quantitative mass spectrometry.
3. To validate a subset of targets discovered by quantitative mass spectrometry as SUMO targets.
4. To identify the SUMO-lysine acceptor sites of the proteins those are validated and demonstrate that mutations of target lysines to arginines block SUMOylation.
5. To molecularly characterize the biological role of SUMOylation of the selected target genes, in order to understand their role in the regulation of the immune response

Chapter 3

Drosophila SUMO modulates NF-κB dependent anti-microbial activation

3.1 Summary

In this chapter, we have shown the affect of SUMO knockdown on AMP response in *Drosophila* S2 cells. Knockdown of SUMO using RNAi was achieved to about 90%. This considerable reduction in the amount of SUMO lead to variation in the response curves of the AMPs post LPS induction as compared to control wild type cells. This suggests that SUMO modification regulates the Rel/NF-κB pathways at either single or multiple levels to alter the pathway response.

3.2 Introduction

In *Drosophila* larvae and adults, the humoral response includes the production of a wide range of lytic peptides - AMPs from the fat body. The transcriptional activation of these AMPs occur preferentially in response to activation of either the Toll or the IMD pathway by bacteria or fungi (Lemaitre et al., 1997). The extent and duration of AMP production needs to be tightly regulated to provide a rapid and transient response. Both, the activation of the response and later a feedback loop to culminate the response are important to prevent spread of pathogens in the body and to prevent inhibition of mutualistic bacteria in the gut. The transcriptional upregulation of these AMPs is primarily governed by the nuclear transport and binding of NF-κB transcriptional factors to the NF-κB binding sites upstream to the AMP gene in response to Toll or IMD pathway activation.

In flies, it has been shown that an immune challenge by gram positive/ gram negative bacteria, leads to specific temporal pattern of activation of each signaling pathway (Boutros et al., 2002). Owing to this, the expression of the AMPs produced downstream of these pathways is temporally distinct. AMPs are also known to be influenced by post-transcriptional modifications. For example, the highly expressed *Cecropin*A1 mRNA produced in response to both bacterial and fungal infections has a shorter half-life than the

moderately expressed *Diptericin* mRNA that responds to gram-negative infection. This is suggested to be regulated through AU rich elements (AREs). The mRNA of most of the AMP genes contain ARE which interacts with other factors to affect the stability of the AMP mRNA leading to differential kinetics of these peptides (Lauwers et al., 2009; Wei et al., 2009).

Post-translational modifications like phosphorylation and ubiquitination modulate the immune pathway responses by modifying various players of the pathway (Silverman et al., 2000; Zhou et al., 2005). However, direct modification of the AMPs through any of the post-translational modifiers has not been determined till date. Further, a class of modifiers, called Ubiquitin-like modifiers (UBLs) regulate a variety of processes within the cell including mammalian and *Drosophila* immunity (Herrmann et al., 2007; Oudshoorn et al., 2012). The role of ubiquitin is established in both *Drosophila* and mammalian immunity. Ubiquitylation of IκB leads to its degradation and release of NF-κB, which enters the nucleus to up-regulate transcription of target genes. K63 Polyubiquitination of other payers in the pathway including IMD, DREDD serves as a regulatory modification and mediates interactions with other proteins for proper downstream functioning. UBLs- ISG15, FUB1 and FAT10 are expressed in response to an interferon signaling in mammals. ISG15 plays a role in mammalian anti-viral response and other inflammatory responses. ISG15 deficient mice succumb to influenza, herpes and Sindbis viruses easily than wild type mice. ISGylation of IRF3 prevents it from ubiquitin mediated degradation thus sustaining its activation (Jeon et al., 2010; Skaug and Chen, 2010). FUB1/MNSFβ is a cytokine that can regulate mammalian immune response by both inhibiting the proliferation of T and B cells and regulating cytokine secretion by T cells. Further, MNSFβ inhibits the secretion of interleukin-4 (IL-4) by bone marrow-derived mast cells and a cell line of type 2 helper T cells. It also inhibits production of tumor necrosis factor-alpha (TNFα) in LPS-stimulated macrophages. These studies have helped characterize the anti-inflammatory and anti-proliferative roles of MNSFβ. Thus, mainly on leukocytes (Herrmann et al., 2007). SUMO, another popular UBL has emerged as an important modifier having implications in a number of cellular processes like DNA damage, cell cycle, chromatin modification and stress responses etc. among others (Geiss-Friedlander and Melchior, 2007; Tempe et al., 2008), proving as an interesting candidate for study in light of immunity.

Studies involving the knockdown of SUMO and its conjugation pathway components in different organisms have shown SUMO to be essential for survival. Knockdown of the single *SUMO* or *Ubc9* gene in *C.elegans* leads to post gastrulation embryonic arrest and multiple defects in larval development (Jones et al., 2002). *Ubc9* and *Smt3* deletion mutants in *S. cerevisiae* are mostly lethal due to undivided nuclei and defects in chromosome segregation (Dieckhoff et al., 2004). It was further shown that *Ubc9*-null mouse embryos die at early post-implantation stage due to mitotic defects (Nacerddine et al., 2005). Recently, it was shown that SUMO-1 knockout mice are viable with no visible phenotype thus proving SUMO-1 to be dispensable in normal mouse development. SUMO-1 deficiency is compensated for by SUMO2/3 within the cell (Zhang et al., 2008). *Drosophila Smt3* mutant embryos show defects in various aspects of cell cycle (Nie et al., 2009). *Smt3* and *lesswright* (*lwr*)-*Drosophila* E2 enzyme nulls show embryonic or larval lethality whereas hypomorphs show varied defects.

Despite developmental defects, SUMO has also been shown to influence *Drosophila* immune response. In *Drosophila* larvae, *Ubc9* (SUMO conjugating enzyme) mutants show an increase in hemocyte proliferation and constitutive over-expression of AMPs *Cecropin A and Drosomycin* (Chiu et al., 2005). An RNAi study performed in LPS responsive S2 cells showed a decrease in *Smt3* (SUMO) transcript within the cell substantially reduces the expression of *CecropinA* and *Drosomycin* (Bhaskar et al., 2002).

Drosophila cell lines like Schneider (S2) cells or mbn2 cells are derived from hematopoetic lineages thus making them favorable to study immunity. Several large-scale RNAi based in-vitro screens have been performed to identify the molecular players important in immune responses against a wide array of infectious pathogens like *E.coli, S. aureus,* etc. Many of these studies have helped identify novel genes as part or as regulators of the Toll and IMD pathways. *Drosophila* tissue culture cells are extensively used for RNAi (RNA interference) studies for the following reasons: 1) Easy delivery of dsRNA into the cell. *Drosophila* tissue culture cells can bind and internalize long dsRNA fragments present in the culture medium, by scavenger-receptor mediated endocytosis, 2) *Drosophila* cells have an intact RNA interference machinery. Thus they cleave the long dsRNA into 21-23bp fragments which are small enough to minimize off-target effects, and 3) relatively easy in synthesizing long dsRNA strands in-vitro in large amounts for multiple experiments.

In spite of the extensive study on phagocytosis and AMP production in S2 cells in response to different bacterial challenges, there have been no reports which have looked at these changes in a temporal manner. Knocking down SUMO might have various effects on AMP response including complete abrogation of the response or alteration in the degree of response in a time-dependent manner which might hamper proper clearing of the infection and compromise the immunity of the organism. With this aim, we chose *Drosophila* cell culture model to obtain the temporal profile of various AMPs in response to LPS treatment.. The levels of four peptides were checked over 24 hrs post LPS stimulation and the analyzed using quantitative real time PCR (qRT-PCR).

3.3 Materials and methods

3.3.1 Cell culture

We used 529SU cells for the knockdown experiments which was a kind gift from the Courey Lab, UCLA. This is a *Drosophila* Schneider (S2) stable cell line expressing full length FLAG-*smt3(SUMO),* HA-*ubc9* under the control of a metallothionein promoter (Bhaskar et al., 2002). These cells were maintained in the presence of (300μg/ml) hygromycin and the absence of antibiotics, in *Drosophila* Schneider cell medium (Sigma) complemented with heat inactivated 10% fetal bovine serum (FBS; Gibco) at 25°C. For large-scale experiments, hygromycin was not added to the medium.

3.3.2 LPS induction

529SU cells do not appear to need hormonal supplements for activation and give a robust and reproducible immune response as published (Bhaskar et al., 2002). Crude Lipopolysaccharide (LPS; Sigma 0111:B4, Batch 129K4025) was added to the culture medium at a final concentration of 10μg/ml to induce a comprehensive immune response for the stipulated time.

3.3.3 Targeted dsRNA synthesis

Targeted dsRNAs were generated with gene specific primers using either S2 cell derived cDNA or vector clones containing target genes. The primers were selected through in-silico analysis to make dsRNA fragments with minimal or no off-target effects. Both forward and reverse primers used for amplification of the gene specific fragment, contained

the T7 promoter binding sites: 5'-TAATACGACTCACTATAGGGAGA-3'. The resulting PCR product was used as a template for in-vitro transcription to generate double-stranded RNA (dsRNA) using the T7 Megascript RNA polymerase kit (Ambion) according to the manufacturer's instructions.

Table 3-1: List of primers used for dsRNA synthesis

Gene	dsRNA synthesis primer sequences 5'-3'
Ulp1-F-rnai	TAATACGACTCACTATAGGGAGAAGTCATGTTCCGGTCTGGTGCTCC
Ulp1-R-rnai	TAATACGACTCACTATAGGGAGATTGTGCTTGAACCCCTACTGCCAC
Smt3-F-rnai	TAATACGACTCACTATAGGGAGAGCCACCAGTCTGCTGCTG
Smt3-R-rnai	TAATACGACTCACTATAGGGAGATGACGAAAAGAAGGGAGGTG

3.3.4 RNA knockdown experiment

For RNA interference assay, the cells were split and allowed to grow to 40-50% confluency. The old medium was then completely removed by washing with serum free media. The cells were suspended in serum free media to which 10 µg of relevant dsRNA/ml was added and incubated for 2 hours. The serum free media was then replaced by Serum containing media and allowed to grow. After 72 hours one set of cells was treated with 10µg/ml of LPS for 2-24 hours while sterile water was added to control cells. For immunoblotting experiments, these cells were induced with 500 mM $CuSO_4$ after 24 hrs of adding dsRNA for 48 hrs.

3.3.5 RNA extraction, cDNA synthesis and quantitative RT-PCR

Post infection, the cells were centrifuged, washed with 1X PBS (phosphate buffer saline) and total RNA was isolated using TRIzol (Invitrogen) following the manufacturer's instructions. After degrading any genomic DNA with RNase-free DNase (Promega), the RNA was re-precipitated and subjected to reverse transcription-PCR (RT-PCR) with poly dT primers to obtain cDNA. 2 µg of the RNA was used per 50 µl cDNA reaction using MMLV Reverse transcriptase (Promega) following manufacturer's instructions. qRT-PCR was performed with Realplex EP mastercycler Real-Time PCR System (Eppendorf) in 25 µl

reactions containing SYBR Green Master Mix (Roche) and 300 nM forward and reverse primers. PCR reaction included a 10 min denaturation step at 95°C followed by 40 cycles of 15s denaturation at 95°C, 15s annealing at 60°C, 20s extension at 68°C and 20 min for the melting curve measurement. This cDNA was used to check for the levels of AMPs and SUMO pathway components normalizing it to Rp49 housekeeping control.

Table 3-2: List of primers used for qRT-PCR

Gene	qRT-PCR Primer sequence 5'-3'
Rp49-F	GACGCTTCAAGGGACAGTATC
Rp49-R	AAACGCGGTTCTGCATGAG
Attacin AB-F	GGCCCATGCCAATTTATTCA
Attacin AB-R	CATTGCGCTGGAACTCGAA
CecropinA-F	TCTTCGTTTTCGTCGCTCTC
CecropinA-R	CTTGTTGAGCGATTCCCAGT
Defensin-F	AGGTTCCTTAACCTCCAATC
Defensin-R	CATGACCAGCATTGTTGTAG
Diptericin-F	AGGTGTGGACCAGCGACAA
Diptericin-R	TGCTGTCCATATCCTCCATTCA
Metchnikowin-F	GCTACATCAGTGCTGGCAGA
Metchnikowin-R	AATAAATTGGACCCGGTCT
Drosomycin-F	CGTGAGAACCTTTTCCAATATGATG
Drosomycin-R	TCCCAGGACCACCAGCAT
Drosocin-F	GCACAATGAAGTTCACCATCGT
Drosocin-R	CCACACCCATGGCAAAAAC
Smt3-utr-F	AACCACAAAAGCAAAAACACAAC
Smt3-utr-R	GTTATTTACGCACACAGACGC

3.3.6 Protein isolation and Immunoblotting

The cells were collected after 48 hrs of $CuSO_4$ induction, washed with 1X PBS and lysed in RIPA buffer (50 mM Tris pH 7.4, 1% NP-40, 0.5% Na-deoxycholate, 0.1% SDS, 150 mM NaCl, 1mM EDTA and 0.01% Sodium azide), freshly supplemented with 40 mM N-ethylmaleimide (NEM, Sigma) and Complete Protease Inhibitor Cocktail (1 Tablet per 100 ml; Roche). The NEM added serves as a SUMO protease inhibitor. The suspension was allowed to rest on ice for 20-30 min following which it was lysed in a bioruptor (Diagenode, 130W, 15min total time, 30sec on/off pulse). The lysate was then centrifuged at 25,000g at 4°C for 45 min. The supernatant protein sample collected was quantitated using DC Protein assay kit (Bio-Rad). 20µg protein from each experimental and control set was boiled in 1X SDS Laemmli buffer and loaded onto 10% SDS-PAGE gel. The separated proteins were transferred to PVDF membrane using semi-dry transfer buffer (250 mM Tris, 1.5M glycine with 5% methanol) at constant 350 mA for 1 h 30min. Membrane was incubated in TBS (Tris-buffered saline) with 0.1% Tween 20 (TBS-T) and 5% NFDM for 1 hr. Primary antibodies were diluted in TBS-T/5% milk at 1:1000 for anti-rabbit FLAG (Sigma) for 1h 30min. The blots were washed with TBST and then incubated with HRP-conjugated anti-rabbit secondary antibody (Jackson Immunoresearch) at 1:10,000 in TBS-T/5% milk for 1hr. The blot was washed thoroughly with TBST and was developed using Millipore ECL on LAS 4000 Imager (Fujifilm).

3.4 Results

3.4.1 Knockdown of SUMO pathway components affects the SUMO conjugated protein profile in S2 cells

We used S2 cells stably transfected with FLAG-*smt3 (SUMO)*, HA-*ubc9* (529SU cells) under the control of a metallothionein promoter, to study the dynamics of SUMO conjugation within the cell. These cells were treated with dsRNA to deplete SUMO, Ubc9 and Ulp1 by RNAi as opposed to control cells. Post 72 hrs of dsRNA treatment, equal concentrations of whole cell lysates from untreated and treated cells were used for immunoblotting and probed with anti-FLAG antibody to examine the SUMO conjugated protein profile within the cell. In untreated cells, there is an array of SUMOylated proteins of different molecular weights seen on the blot (lane 1, **Figure 3-1**). In *SUMO* and *Ubc9* gene knockdown conditions, this ladder almost disappears with very few bands seen on the blot.

This indicates that there is significant decrease in the amounts of SUMO and Ubc9 which disrupts SUMO modification of substrates within the cell (lane 2 and 3, **Figure 3-1**). However, knockdown of Ulp1 seems to cause an increase in the SUMO modified species within the cell (lane 4, **Figure 3-1)**. Due to overexpression of the SUMO-GG mature protein, only the Ulp1 deconjugation function is affected in these cells leading to accumulation of SUMOylated substrates. These results are in concert with previously published data which confirm that the knockdowns are working in the given set of conditions (Smith et al., 2004).

3.4.2 Knockdown of SUMO affects the AMP expression profiles

To study the effects of SUMO knockdown in immune response, dsRNA mediated knockdown of *SUMO* was carried out. The dsRNA treatment reduced the SUMO transcript to ~10% as determined using qRT-PCR. We monitored transcripts levels of AMP genes in *SUMO* depleted S2 cells and control untreated cells in response to crude LPS over a time period of 24 hrs i.e 0 hr, 2 hrs, 6 hrs, 12 hrs and 24 hrs (**Figure 3-2**).

A*ttacin*and *Metchnikowin* transcript levels are reduced in SUMOi condition as compared to wild type control without RNAi, however, the response curve in both conditions peak at the same time post LPS stimulation. The other AMP *Drosocin* also shows changes in its expression profile, with an increase in transcript levels in the absence of SUMO. The *Drosomycin* response however, seems to be abolished almost completely in SUMO knockdown conditions. Due to high standard deviation values in the three biological replicates, only one of the set is represented in the figure, however the trends of activation remain the same in each case.

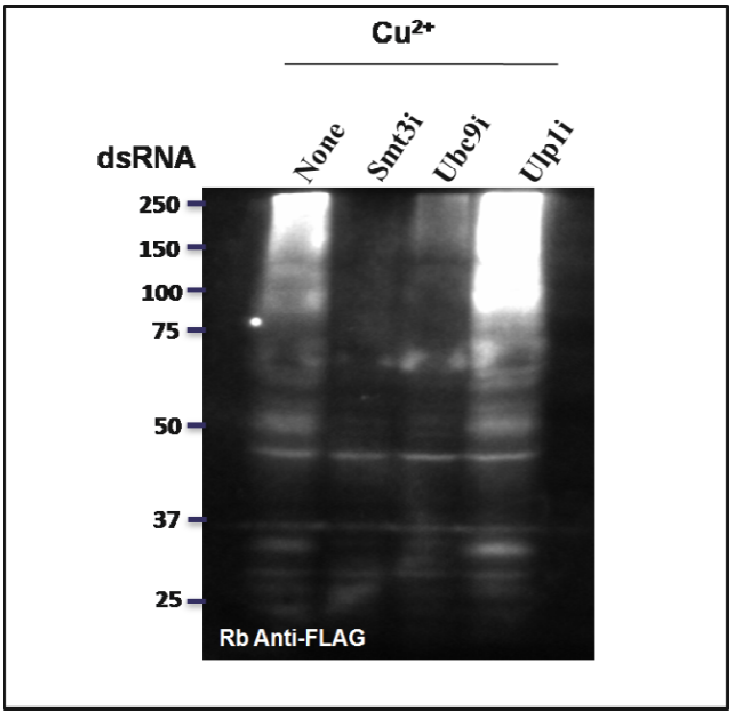


Figure 3-1: Knockdown of SUMO pathway components affects the SUMO substrate profile in S2 cells.

A stable cell line (529SU), expressing FLAG-SUMO^{GG} and HA-Ubc9 under a Copper inducible metallothionein promoter was used to express tagged SUMO cycle components. In control cells without any dsRNA treatment, FLAG-SUMO conjugates to a large number of proteins, generating a 'SUMO' ladder as seen in the blot (lane 1). dsRNA interference of SUMO (Smt3i; lane 2) or Ubc9 (Ubc9i; lane 3) transcripts lead to a global decrease in SUMOylation, while reduction of Ulp1 (Ulp1i, Lane 4) leads to a global increase in SUMOylation as these cells express mature SUMO^{GG} protein and hence affecting only the Ulp1 deconjugation function.

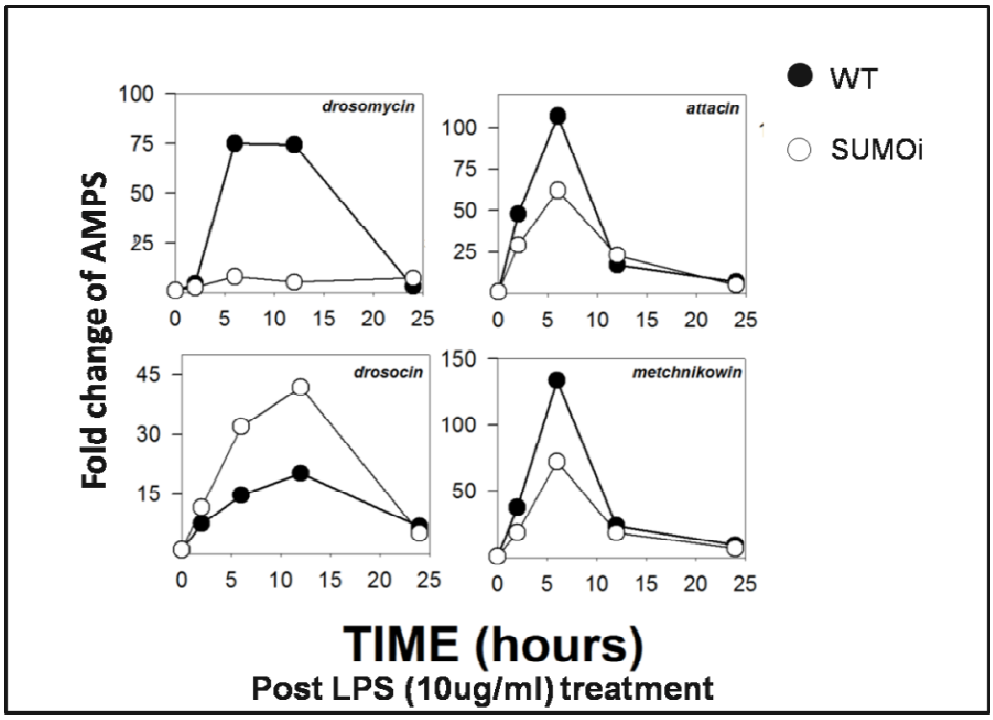


Figure 3-2: SUMO knockdown affects kinetics of anti-microbial peptides
Quantitative real Time PCR (qRT-PCR) was used to measure the expression of defense peptide genes, in response to crude LPS infection, over a time period of 24 hours in Smt3 depleted and control wild type cells. The graph shows transcript levels of AMPs in wild type (filled circles) and in SUMO knockdown (~90%; empty circles) cells. The most dramatic reduction is for *Drosomycin* (drs) whose activation is almost completely suppressed in cells post SUMO knockdown while transcripts levels and kinetics of other NF-κB target genes are moderately affected.

3.5 Discussion

SUMO (Smt3) gene encodes a protein that conjugates with a variety of target proteins affecting a number of cellular processes within the cell. Here, we demonstrate that SUMO function is required to maintain the expression kinetics of the anti-microbial peptides in *Drosophila* S2 cells. In response to SUMO knockdown a few AMPs are up-regulated while a few others are down-regulated post LPS-challenge. The reason for the differential regulation of the AMP genes upon SUMO knockdown is unclear at this point. It may include the possibility that the Toll and IMD pathway targets are differently modulated by SUMO thus affecting the preferential activation of AMPs in response to either pathway. Also, SUMO-mediated control of other pathways like JAK-STAT and JNK which are known modulators of the NF-kB response pathways may lead to these discrepancies.

Thus, inorder to better understand the mechanisms by which SUMOylation affects immune signaling, we need to identify substrates in the immune regulatory networks that are SUMOylated or deSUMOylated in response to infection. This might give us an understanding of how SUMOylation of different players in the network affect the response individually and in concert with each other to fine-tune the system.

Chapter 4

Identification of SUMO modified targets through quantitative mass spectrometry

4.1 Summary

In this chapter, we describe a *list* of proteins that show a change in their SUMOylation states on immune activation by *crude* LPS. The list was generated using immuno-pulldowns followed by quantitative mass spectrometry technique, iTRAQ. The list provides a confident set of 858 proteins which show changes in their SUMOylation states upon LPS addition.

4.2 Introduction

SUMO modification helps in maintaining immune homeostasis. Proteins in both mammalian and *Drosophila* immune pathways have been shown to be modified by SUMO leading to different functional consequences. IκBα was the first SUMO-modified protein implicated in the NF-κB signaling pathway (Desterro et al., 1998). SUMOylation prevents its ubiquitination at the same lysine residue thus preventing ubiquitin mediated proteasomal degradation. IκBα thus continues to sequester NF-κB in the cytoplasm, not allowing its nuclear translocation and activation of the immune response genes. In *Drosophila* Dorsal which is the NF κB transcription factor has been shown to be SUMOylated at K382. SUMOylation of Dorsal affects the AMP response in S2 cells (Bhaskar et al., 2002). STAT92E, component of the JAK-STAT pathway which is a known regulator of the immune response has been shown to be SUMOylated, however the molecular consequence of the modification is yet unknown (Gronholm et al., 2010). These studies have only targeted specific components of the various immune response pathways. It is important to undertake a genome-wide unbiased approach using the new mass spectrometry techniques to identify the wide repertoire of proteins that are involved in eliciting an efficient immune response. Various studies in different organisms including yeast, *C. elegans*, *Drosophila*, mammals and *Arabidopsis*, have identified SUMO modified proteins using both qualitative and quantitative mass spectrometry approaches (Denison et al., 2005; Golebiowski et al., 2009; Kaminsky et al., 2009; Kurepa et al., 2003; Tatham et al., 2011; Wohlschlegel et al., 2004). A number of

studies in mammals have shown the importance of SUMO in heat stress and proteasomal pathways while identifying the proteins that are SUMO modified in these states. These help in identifying SUMOylated targets to further understand the functional relevance of SUMO modification of the proteins in varied contexts.

4.2.1 Mass Spectrometry as a tool for proteomics analysis

Proteomics is the large scale study of proteins which is applied to identify and analyze the entire protein repertoire in a given biological sample. Mass spectrometry is an analytical tool to measure the molecular mass of a sample. It was initially used to identify small chemical moieties and peptides. However, after the invention of the soft protein ionization methods such as electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) techniques which do not lead to fragmentation of large protein molecules, mass spectrometry tool is widely used for protein analysis. Also, the completion of genome sequencing of a number of organisms has contributed to the ease of MS-based proteomics by providing protein databases against which the experimental spectra can be matched to identify proteins in an automated manner. Protein mass spectrometry using LC-MS/MS; a combination of liquid chromatography (LC) and tandem mass spectrometry (MS) is extensively used to study protein mixtures.

Most of the quantitative proteomic studies aim at the comparison of a stressed or disturbed status to an undisturbed reference sample. The most common form of quantitative proteomics was the use of dyes, radioactivity or fluorophores after separating the proteins on the gel using two-dimensional gel electrophoresis (2-DE). This technique is tedious and does not allow direct identification of proteins. The protein band has to be excised from the gel and further analyzed by mass spectrometry. This technique has limitations with respect to low protein resolution obtained on 2-DE gels and under-representation of proteins due to size and pH. However with the advancement in instrumentation and labeling techniques, quantitative mass spectrometry is now a rapidly growing field.

Inherently LC-MS/MS based proteomics is not quantitative because of different physical and chemical nature of the tryptic peptides produced. Peptides from the same protein may differ in charge state, peptide length, amino acid composition, or post-translational modifications resulting in great differences in the ion intensities for the peptides. Thus, for accurate quantitation using ion intensities, the peptides to be compared between different

samples should have the same peptide mass-to-charge ratios (m/z) that were acquired under the same general conditions during LC-MS/MS experiments. To aid accurate determination of quantitative information with MS two different methods are currently in use: 1) label free and 2) stable isotopic labeling approach. Stable isotope labeling methods ensures that the peptide labeled with stable isotopes does not differ in their chemical properties and hence the two peptides behave identically during chromatography and further mass spectrometry analysis. The labeling methods include isotope coded affinity tags (ICAT), stable isotope labeling by amino acids in cell culture (SILAC) and isotope tags for relative and absolute quantification (iTRAQ) which are the most commonly used techniques. The isotope tag is introduced either metabolically i.e., in vivo or chemically or enzymatically i.e., in vitro. The basic outline for all isotopic label based quantitation involves the incorporation of the stable isotope labels into the samples to be compared following which these samples are combined and subjected to separation and analysis by mass spectrometry. Combining the samples throughout the experimental procedure makes sure that the samples to be compared are always together and exposed to identical conditions. This helps minimize sample variation. Each technique has its advantages and limitations and the choice of approach would depend on the experimental setup.

ICAT was the first isotope labeling technique developed by S. Gygi *et al.* (2000) (Gygi et al., 1999). The tag used contains an iodoacetamide group which reacts with reduced cysteine residues, a bridge that contains the isotopic carbon or hydrogen atoms for differential labeling and a biotin group which allows the specific isolation of cysteine containing peptides onto a streptavidin column. This approach ensures that all tryptic cleavage peptides of a protein carry at least one labeled amino acid. The protein is identified based on the fragmentation spectrum of either of the labeled peptides while relative quantification is performed by comparing the intensities of the isotope clusters of differently labeled peptides in MS spectra. ICAT labeling has been used extensively for a number of studies, however the major limitations of the system include: 1) the labeling is cysteine specific and only 80-90% of the proteins in the complete proteome have cysteine. Hence there is incomplete proteome coverage; 2) only a pair-wise comparison of protein samples is possible with no multiplexing allowed; 3) non-specific binding to streptavidin matrix.

The first demonstration of in-vivo labeling using SILAC was provided by Mann and colleagues (Ong et al., 2002). In this technique labeled essential amino acids are added to

amino acid deficient cell culture media. Overtime, the cells grown in this media have the labeled amino acids incorporated into their proteome. Protein identification and relative quantification is carried out as in the case of ICAT in the MS spectra. The major advantage of this technique is that there is minimal sample manipulation. Since the incorporation of the tags is uniform in both control and experimental sets, the two sets can be mixed for sample preparation and other chromatography and mass analysis. This technique is suitable to measure relatively small changes in protein levels and study post translational modifications. However, multiplexing in this technique is difficult due to increase in sample complexity on using multiple isotopes. Also this technique is more feasible in cell lines and difficult in whole tissues from organisms (Bantscheff et al., 2007; Ong et al., 2002).

4.2.2 iTRAQ

Isobaric tags for relative and absolute quantitation (iTRAQ) method is based on chemically labeling the peptides with isobaric tags. The method was developed by Ross *et al.* (2004) and allows for simultaneous relative quantification of upto eight different samples in a single run (Ross et al., 2004).

iTRAQ labels have a total isobaric mass of 145 Da and consist of a reporter group and an amine-reactive group (NHS ester derivative) flanking a balancing group (carbonyl group). The introduction of the iTRAQ reagent occurs at the level of tryptic peptides. While comparing different samples, identical peptides with the different isobaric labels having the same physico-chemical properties co-elute in the same fraction on chromatography and are selected for fragmentation as a single precursor giving a single peak in MS scan. After further fragmentation, different mass tags separate, releasing reporter ions. The intensity ratio of the different reporter ions in the MS/MS spectrum enables relative quantification of the peptide across the different samples (Boehm et al., 2007; Schulze and Usadel, 2010). In their original form the iTRAQ reagents consist of a set of four reagents which release reporter ions of masses 114, 115, 116 and 117 Da. The proteins or whole lysates are reduced, alkylated, tryptically digested and these peptides are labeled with the iTRAQ reagents. The resulting peptides are mixed in equal amounts following which they undergo separation by strong cation exchange chromatography (SCX) and reverse phase HPLC. This fractionation helps remove any free iTRAQ reagents and reduces the sample complexity. The separated fractions are subsequently analyzed by tandem mass spectrometry (MS/MS). The iTRAQ ratios are usually compressed under-represented values of the actual levels of the proteins in the

sample, typically reporting fold changes of less than two orders of magnitude (Ow et al., 2009).

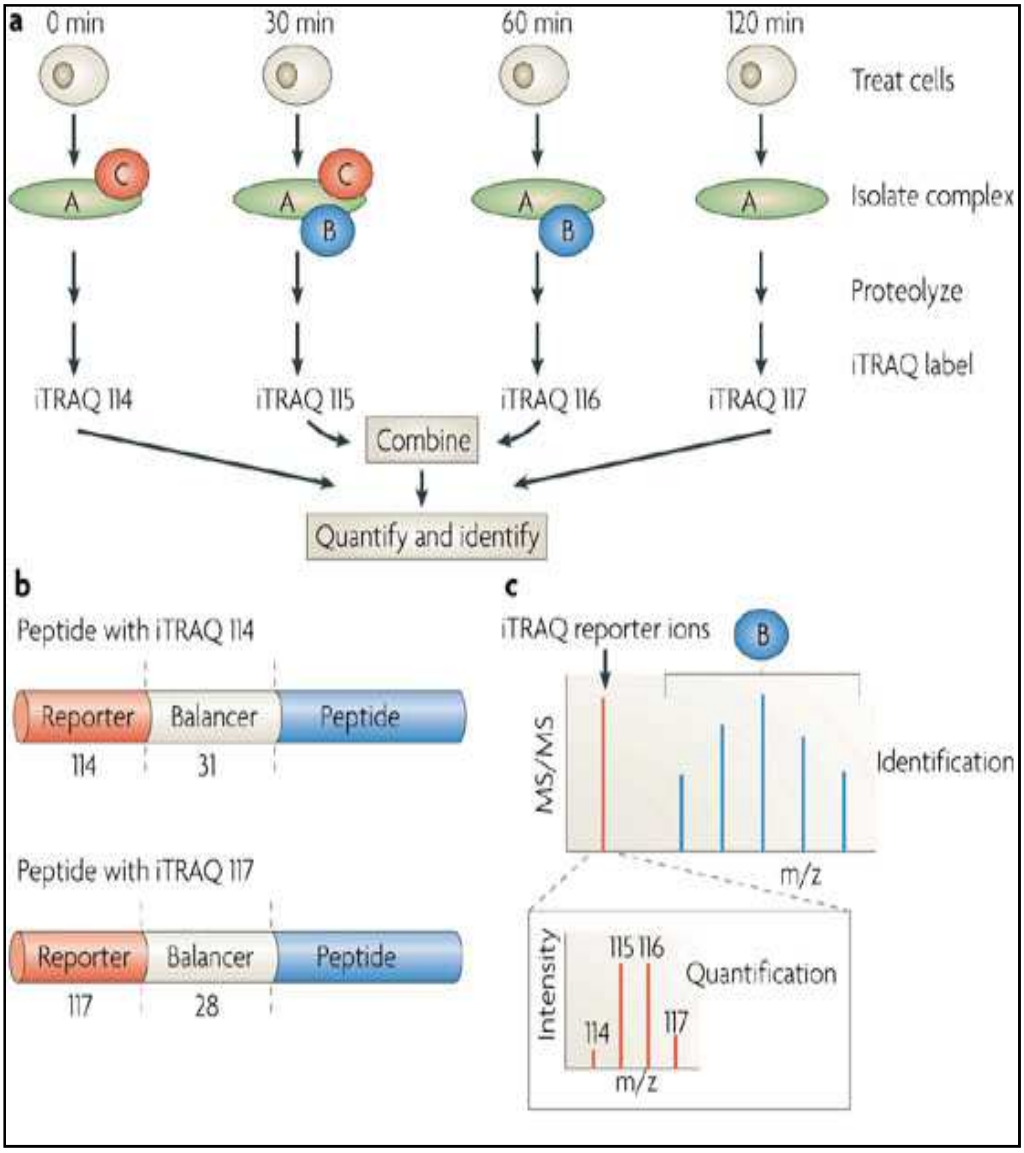


Figure 4-1: iTRAQ method work flow (Copyright of npg 2007, Aebersold *et al.*)
4-plex iTRAQ labeling can be used to multiplex four different samples. Different protein samples are digested with trypsin and then labeled with the iTRAQ labels named iTRAQ 114, 115, 116 and 117 depending on its reporter ion. The labeled peptides are then pooled subjected to LC-MS/MS analysis for protein identification and to determine their relative amounts (Gingras et al., 2007).

iTRAQ method of labeling is more efficient than ICAT with respect to coverage as it reacts with primary amines which are present in all proteins. Thus, a protein can be identified and quantified with more confidence using data from multiple peptides of the same protein often with multiple values per distinct peptide. iTRAQ data is however, shown to be prone to variation which can be tackled at different levels from sample preparation and processing to analysis of the data obtained from mass analysis.

4.2.3 Identification of Ub and UBL acceptor lysines using mass spectrometry

Proteomic approaches are now being applied to identify peptides modified by ubiquitin (Ub) and UBLs. Following ubiquitination of a protein, the modified lysine is not subjected to tryptic digestion. The resulting Ub-conjugated peptide is modified at that lysine residue with a -GG remnant which shifts the mass of the peptide by 114 Da. This predictable mass shift of the precursor ion allows an easy detection of ubiquitinated peptides. However, apart from ubiquitin other UBLs like NEDD8 and ISG15 share the same Gly-Gly mass shift making it difficult to differentiate between the modifications. To combat this problem, researchers have tagged Ub to enrich only the Ub modified proteins using biochemical pulldowns followed by mass analysis.

Mapping of SUMOylation sites poses additional challenges for detection with mass spectrometry due to the sequence of SUMO. Trypsin digestion of SUMO modified proteins results in a longer peptide remnant conjugated to the target which is not the case in ubiquitin where the trypsin cleavage site is close to the GG residue. This tryptic peptide is large and produces many fragment ions during MS/MS fragmentation, making the identification difficult by conventional automated database searching engines. For this SUMO protein with a trypsin cleavage site close to the modified lysine has been engineered and is being used to identify the SUMO acceptor sites.

Another common problem in identifying UBLs is that their abundance in a given biological sample is very low probably due to rapid turnover rates within the cell or rapid deconjugation during cell lysis. This can be avoided to a certain extent by using generic protease inhibitors and SUMO protease inhibitor NEM (N-ethylmaleimide) during sample processing.

Due to the complications and limitations in mass spectrometry-based proteomics of Ub/UBLs, sample purification and enrichment seem to be critical steps to obtain consistent

and desirable results. To increase the abundance of modified proteins it has become a common practice to overexpress Ub/UBLs using non-endogenous tags which allow efficient enrichment of the modified sample. However, further methodological advances are required to improve the applicability of mass spectrometry to Ub/UBL modifications.

In our study we used the iTRAQ technique to identify quantitative changes in response to LPS induced immune challenge in S2 cells. iTRAQ was chosen due to ease in multiplexing and availability of the facility and reagents. As discussed above, we used wild type SUMO protein over expressed with FLAG-tag to increase SUMOylation of the target proteins and mediate easy purification of the tagged proteins. The pulldowns were carried out in native conditions with the presence of NEM to hamper de-SUMOylation and FLAG peptide was used for elution to increase specificity of eluted proteins.

4.3 Materials and methods

4.3.1 Cell culture and LPS treatment

529SU cells expressing FLAG-SUMO as described in section 3.2.1 were used for affinity pulldowns. After characterization of the LPS mediated AMP response kinetics over 24 hrs, we chose the 2.5 hr time point for quantitative proteomic analysis. LPS used was as described in section 3.2.2.

4.3.2 FLAG affinity purification of SUMO conjugates

The starting point for the proteomic experiments was 1800 ml of 529SU cells, at a cell density of ~1 X 10^6 cells/ ml, in thirty, 300 cm^2 culture flasks. The cells were split into three 600 ml aliquots (10 flasks each), with 2 flasks induced with 0.5 mM $CuSO_4$ and the third serving as master control (**Figure 4-2**). Three days after the split (and induction; cell density ~1 X 10^7), half of the induced flasks (10 Flasks) were mock treated with sterile water, while another 10 (induced) Flasks were treated with 10ug/ml LPS for a period of 2.5 hours. The cells were collected from the flasks near the end of the incubation period, centrifuged at 1000g, washed with 1XPBS, lysed with RIPA buffer (50 mM Tris pH 7.4, 1% NP-40, 0.5% Na-deoxycholate, 0.1% SDS, 150 mM NaCl, 1 mM EDTA and 0.01% Sodium azide), freshly supplemented with 40 mM N-ethylmaleimide (NEM, Sigma) and Complete Protease Inhibitor Cocktail (1 Tablet per 100 ml; Roche), exactly at 2.5 hrs. The entire 1800 ml set experiment was repeated thrice for analysis.

The suspension was further lysed in a bioruptor (Diagenode, 130W, 15 min total time, 30 sec on/off pulse) and the lysate was centrifuged at 25,000g at 4°C for 45 min. The supernatant was pre-cleared using protein G sepharose (GE Healthcare) for 1hr at 4°C. Equal concentrations of the pre-cleared lysates from both LPS-induced, and un-induced cells were nutated with RIPA buffer equilibrated with anti-FLAG agarose (1 ml) at 4°C overnight. The next day, the beads were separated from the lysate by centrifugation, washed 3 times with TBST (50 mM Tris, pH 7.4, 150 mM NaCl and 1% Triton X-100) followed by 2 washes in TBS. The bound proteins were eluted from the beads using 10 bead volumes of FLAG peptide (Sigma) at a concentration of 200 µg/ml for 4 hrs. The eluted proteins were concentrated using amicon concentrator (Millipore). A portion of the eluted proteins were separated by SDS-PAGE followed by Western-Blot analysis using Rb anti-FLAG antibody (Sigma) and silver staining. The remaining protein elute was dialyzed against 100 mM NH_4HCO_3 to remove any TBS and lyophilized for further iTRAQ analysis.

4.3.3 Immunoblotting and silver staining of immunoprecipitated samples

The amount of protein in the samples from affinity pulldowns was quantitated using Lowry reagent (Bio-Rad). 20 µg from each experimental and control set was separated on a 10% SDS-PAGE gel followed by western blotting as described in section 3.2.6. For silver staining the SDS-PAGE gels were fixed in 5:4:1 of ethanol: water: acetic acid. After washing the fixed gels with distilled water (D/W), they were sensitized in 0.02% sodium thiosulphate. The gel was further washed briefly with D/W and kept in 0.2% (w/v) silver nitrate solution for 30 min. The gel was thoroughly washed with D/W and developed using sodium carbonate with sodium thiosulphate and formaldehyde until bands were seen clearly following which the reaction was stopped using 6% acetic acid.

4.3.4 iTRAQ labeling and strong cation exchange chromatography (SCX) fractionation (Figure 4-2)

Three biological replicates of each set (un-induced master control, induced untreated, induced LPS treated) were processed in the Mass Spectrometry Facility at the Institute of Bioinformatics, Bangalore. For each set, the lyophilized samples were re-suspended in water and protein was quantitated using Bradford test. Equal concentrations of the proteins was taken from each experimental set after confirmation with silver staining and treated with 2 µl of reducing agent [tris (2-carboxyethyl) phosphine (TCEP)] at 60°C for 1 h and alkylated with cysteine blocking reagent, methyl methanethiosulfonate (MMTS) for 10 min at room

temperature. The samples were digested overnight with sequencing grade trypsin (Promega, Madison, WI) (1:20) at 37°C. Peptides from master control, induced-LPS-untreated and induced-LPS-treated experimental sets were labeled with iTRAQ reagents that would yield reporter ions of m/z 114, 115 and 116, respectively. Labeled peptides from all three conditions were pooled and fractionated by strong cation exchange chromatography on Poly SULFOETHYL A column (PolyLC, Columbia, MD) (100*2.1 mm, 5 µm particles with 300 A pores) using a linear gradient of 5% - 40% Solvent B (350 mM KCl in 10 mM KH_2PO_4, 20% acetonitrile, pH 2.8). Fractionated samples were collected, desalted using stage tips vacuum dried and stored at -80° C until LC-MS/MS analysis. The uninduced master control, containing mock immune-precipitated samples, as expected, did not contain enough protein (1% or less of the main experiment) and was not used for further analysis.

4.3.5 LC-MS/MS

LC-MS/MS of iTRAQ-labeled peptides was carried out on an LTQ-OrbitrapVelos mass spectrometer (Thermo Electron, Bremen, Germany) interfaced with Agilent's 1100 series nanoflow liquid chromatography system (Agilent Technologies, Santa Clara, CA). Peptides from each fraction were enriched and washed on a trap column (75 µm x 2 cm, 5 µm, 120Å, Magic C_{18} AQ Michrom Bioresources), at a flow rate of 3 µl/min and then resolved on an analytical column (75 µm × 10 cm, 5 µm, 120Å, Magic C_{18} AQ Michrom Bioresources) at a flow rate of 300 nl/min using a linear gradient of 5% - 40% solvent B (90% acetonitrile in 0.1% formic acid) over a period of 65 min. The total run time per sample was 85 min. The resolved peptides from analytical column were delivered to mass spectrometer through an emitter tip (8 µm, New Objective, Woburn, MA). LC-MS/MS data was acquired in a data dependent manner in FT-FT mode. MS spectra were acquired with a window of *m*/*z* 350 to 1800. Twenty most abundant precursor ions were selected for fragmentation from each MS scan. Data was acquired at MS resolution of 60,000 (m/z 400) and MS/MS resolution of 15,000. Precursor ion fragmentation was carried out using higher energy collision (HCD) mode with normalized collision energy of 41%. Monoisotopic precursor selection was enabled and the precursor ions that were selected for fragmentation was dynamically excluded for 50 sec.

4.3.6 MS data analysis

The MS data was analyzed using the Proteome Discoverer software (Thermo Scientific, version 1.3.0.339). The data was searched against Flybase (FB2010_04 Dmel

Release 5.27) database containing 43,900 protein sequences along with known contaminants using SEQUEST search algorithm. The parameters used for data analysis included trypsin as a protease (allowed one missed cleavage), iTRAQ labeling at N-terminus and lysine residues, and cysteine modification by methyl methane thiosulfonate (MMTS) as fixed modifications and oxidation of methionine as a variable modification. The precursor and product ion mass error tolerance were fixed at 20 ppm and 0.1 Da respectively. The precursor range was set at 500 to 8,000 Da. The peptide and protein data were extracted using high peptide confidence (1% FDR) and top one peptide rank filters. Relative abundance of proteins across conditions was determined by Proteome Discoverer based on difference in the peak intensity of reporter ions in the MS/MS spectra of each peptide that was ultimately used for quantitating corresponding protein. Differentially regulated proteins were identified as having >2 or <0.5 fold change, calculated by dividing the LPS induced score for each protein by the un-induced value. For downstream analysis, such as for building interaction networks, all the proteins identified in at least two of the three data set with ratio >2 or <0.5 was used.

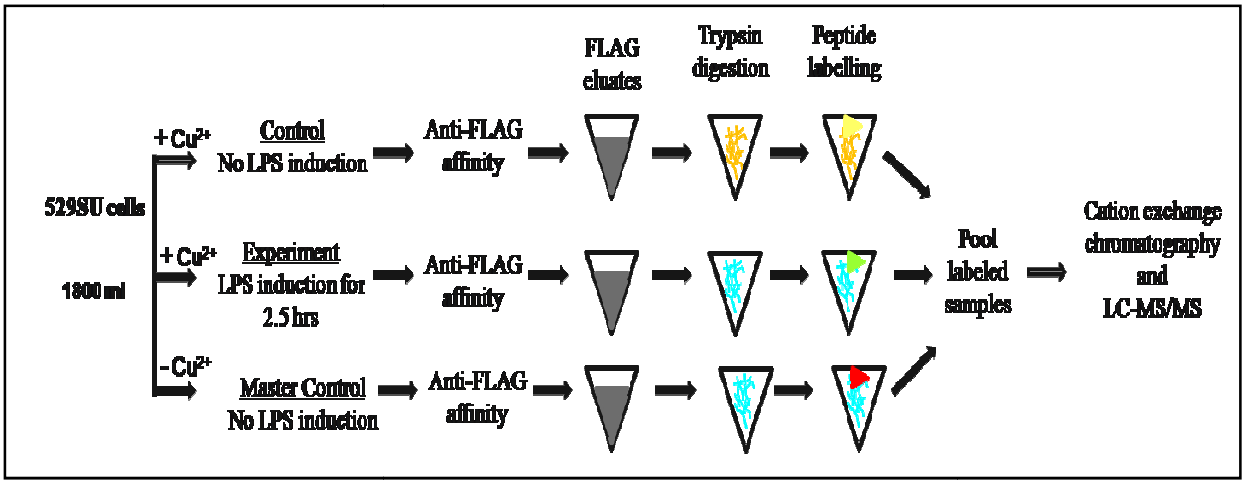


Figure 4-2: Schematic representation of the protocol followed for iTRAQ analysis
529SU cells were without $CuSO_4$ induction was used as a master control to determine any non-specific pulldowns, whereas cells induced with $CuSO_4$ but with or without LPS treatment were the experimental sets. Post 2.5 hrs of LPS induction, the cells were lysed and the supernatant was incubated with FLAG beads. Proteins eluted from the FLAG beads were subject to trypsin digestion and labeling by iTRAQ reagents of same mass. Tryptic digests from various samples were pooled, separated into fractions by cation exchange chromatography and each fraction was analyzed by LC-MS/MS. This protocol was used to collect data for three biological replicates.

4.4 Results

4.4.1 LPS challenge leads to increase in global SUMOylation in *Drosophila* S2 cells

S2 cells stably transfected with FLAG-*smt3 (SUMO)*, HA-*ubc9* (529SU cells) under the control of a metallothionein promoter have been previously shown to be responsive to LPS challenge by Bhaskar *et al*. We used these cells in 3 different conditions: 1) Master negative control without FLAG-SUMO expression and LPS induction, 2) LPS un-induced set expressing FLAG-SUMO and, 3) LPS induced set expressing FLAG-SUMO. The LPS induction was given for 2.5 hrs after which the cells were collected, lysed, affinity purified using anti-FLAG affinity beads and the eluates after FLAG peptide elution were subjected to both silver staining and immunoblotting. The 2.5 hrs time point was chosen based on AMP expression analysis. The master control shows little or no protein in sliver staining (**Figure 4-3**, Lane 1) and does not show reactivity with the Rb-FLAG antibody indicating that there is negligible pull-down of non-FLAG tagged proteins by the anti-FLAG agarose beads under our experimental conditions. The lanes with FLAG-SUMO induction have characteristic SUMOylated species laddered on the gel. The cells treated with LPS seem to show an increase in global SUMOylation, (**Figure 4-3**, Lane 3) more prevalent for proteins > 200 kD. This may be either due to polySUMOylation of a number or target proteins or increased SUMOylation of a specific set of high molecular weight proteins.

4.4.2 iTRAQ analysis to determine LPS induced changes in the SUMO proteome

The immune-purified samples from the three experimental sets discussed above were given for mass analysis as shown in **Table 4-1**. The experimental protocol is shown in **Figure 4-2**. For iTRAQ analysis 80μg of protein from each set was pooled and processed for LC MS/MS based iTRAQ analysis. A total of 1820 unique proteins were identified using a cutoff of 95% probablility in the correct identification of peptides (**Figure 4-4**). Of these 681 (37%) were common to all three sets and 1112 (or 61%) common to atleast 2 sets. The LPS treated versus control set iTRAQ ratios vary from 0.1 to 6. As is well documented, the ratios measured by the iTRAQ experiment are compressed (Karp et al., 2010) and are a relative measure rather than the actual fold values of SUMOylation. To obtain a confident set of proteins showing changes in response to LPS, we have considered proteins present in two or more biological replicates with iTRAQ ratios <0.5 and >2.0 as significant hits. This list of 858 proteins comprises the LPS-induced SUMO proteome and has been used for further analysis.

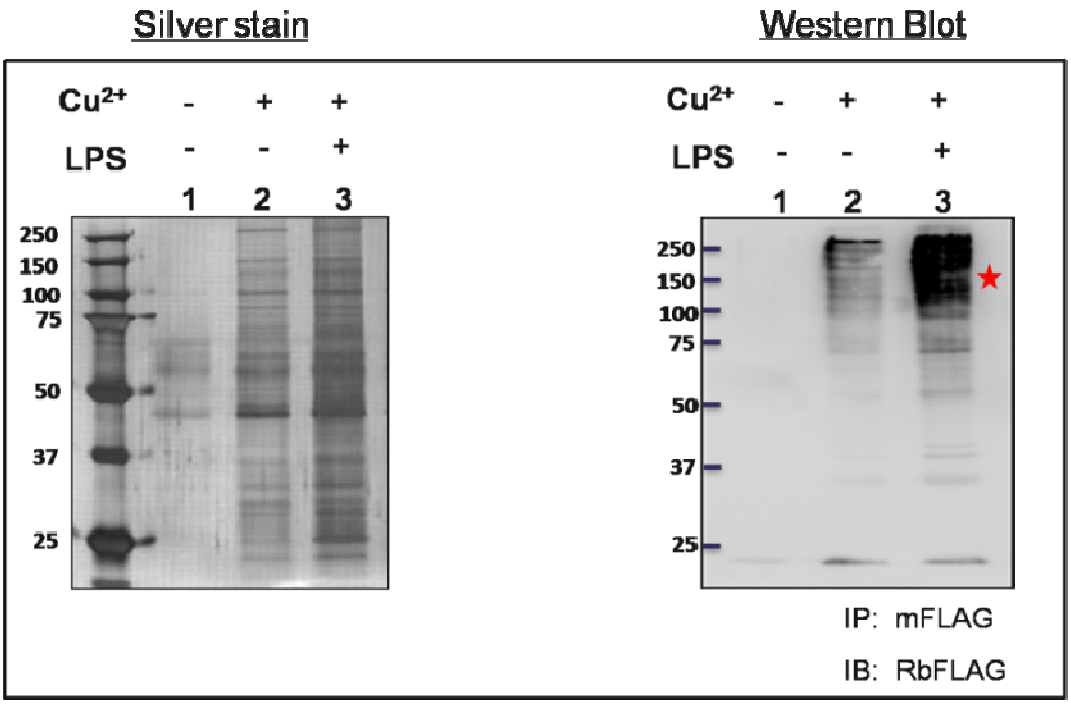


Figure 4-3: Affinity purification of SUMOylated proteins
529SU cells lysates, post LPS treatment were used for affinity purification using mouse anti-FLAG agarose (Sigma). 20 µg of the total, affinity purified FLAG eluate was separated using SDS-PAGE and subjected to silver staining and immunoblotting. Lane 1 has the affinity purified, un-induced master control sample, showing very minimal non-specific pull down, while lanes 2 & 3 are affinity purification of induced cells with (lane 3) and without (lane 2) LPS. There is an increase in the intensity of the higher molecular weight SUMO-purified proteins in response to LPS treatment.

Table 4-1: Raw material for the three biological replicate experiments

	Volume of culture (~10^6cells/ml)	**Concentration of purified protein**	
		uninfected	**infected**
SET1	2 X 600 ml	90 µg	120 µg
SET2	2 X 600 ml	130 µg	200 µg
SET3	2 X 1000 ml	205 µg	380 µg

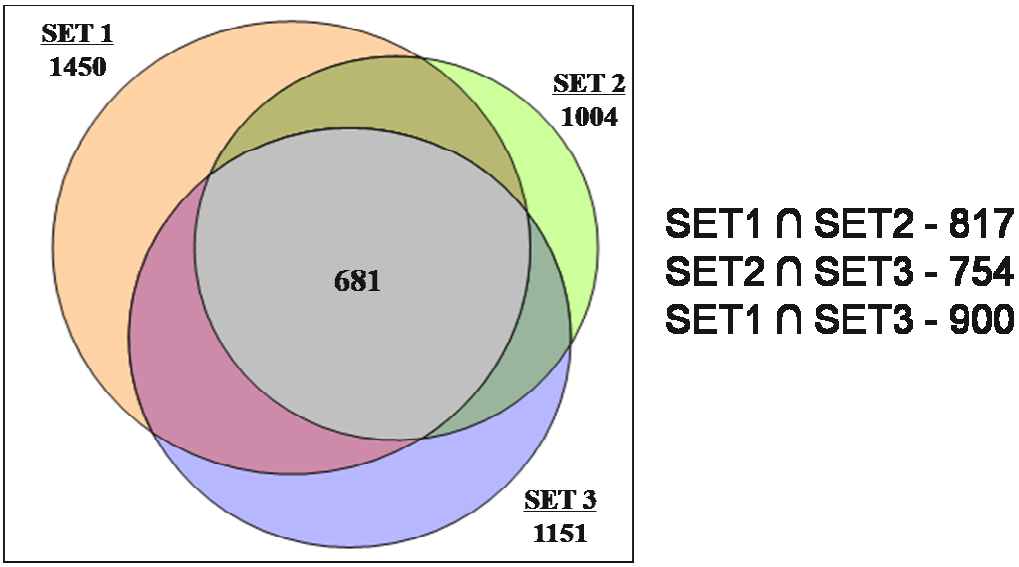


Figure 4-4: Venn diagram depicting the triplicate iTRAQ numbers.
iTRAQ analysis was carried out on three biological replicates, with each replicate set containing a control (no LPS) and an experiment (+LPS). The three data sets identified 1450, 1004, and 1151 proteins with approximately 40% overlap between all three sets.

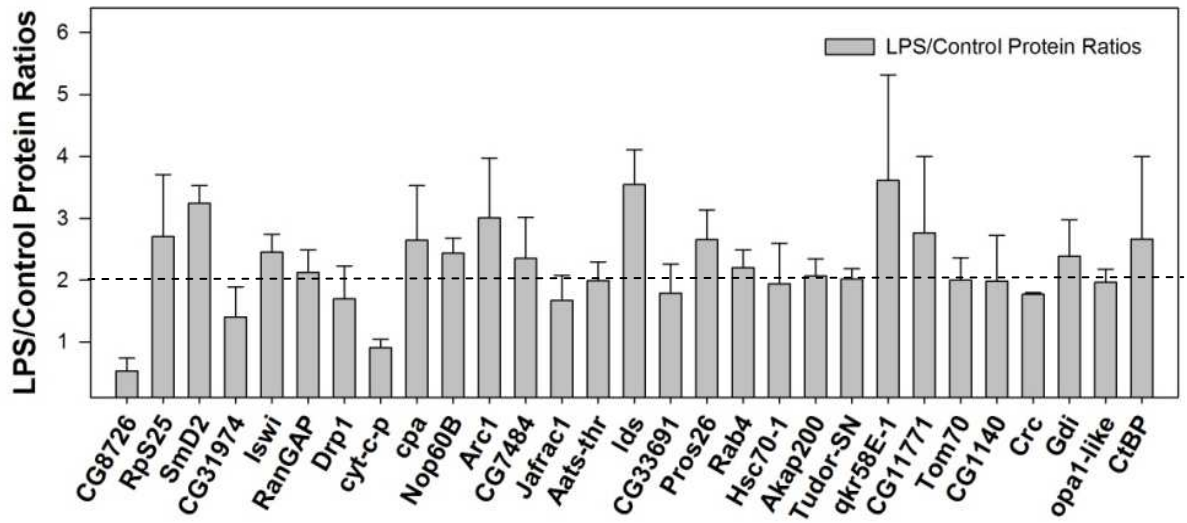


Figure 4-5: iTRAQ ratios of representative proteins
The above graph shows a few representative proteins that show a range of increased SUMOylation states post LPS treatment. These ratios of SUMOylated proteins in LPS treated versus untreated sets differ from 0 to 5 fold. To obtain a confident set of proteins we have considered proteins with fold change values <0.5 and >2 present in all three sets.

4.5 Discussion

We used affinity pulldowns to enrich SUMO modified proteins in response to LPS challenge in S2 cells followed by quantitative mass analysis to obtain a better understanding of the various pathways and complexes that might be under SUMOylation regulation in immune response. The pulldowns were carried out in stringent buffer and detergent conditions to minimize non-specific interactions while, maintaining strong complex interactions in native conditions. This gives an understanding of the complexes which might be regulated by SUMOylated proteins and may be important in immune response pathways. Also there are only upregulated proteins identified in these conditions post LPS treatment which might be because of global increase in SUMOylation post LPS treatment and the downregulated SUMO-substrates might be in very small quantities to be detected by iTRAQ analysis. We applied stringent peptide detection techniques to identify proteins and cut-offs to limit the total number of 1820 proteins to a confident set of 858 proteins, leading to false negatives.

The elimination of the pathogen from the organism occurs either by hemocyte mediated phagocytosis or production of AMPs. The phagocytic response includes processes like internalization of the pathogen through vesicle-mediated endocytosis, actin cytoskeletal reorganization to aid phagocytic movement of the hemocytes etc. The production of AMPs occurs through activation of two major pathways- Toll/NF-κB and the IMD/NF-κB pathways due to translocation of the three NF-κB molecules, Dorsal (Dl), Dorsal like immune factor (DIF) and Relish (Rel) into the nucleus. In addition, these two major pathways are interconnected at various levels and also subject to regulation by other signal transduction pathways. Studies have demonstrated that Ras/MAPK, JNK and JAK-STAT pathways are responsible for negative regulation of the NF-κB pathways or work independently to activate certain effector responses like apoptosis, stress responses and increased hemocyte proliferation. Thus the total proteins identified through mass spectrometry analysis may not be directly involved in an immune response but involves global changes within the cell triggered in response to pathogen attack including changes in general transcription and translation machinery in the cell, cell cycle changes apart from the immune responsive changes discussed above. A thorough analysis of the list will give us a better insight into these changes and is discussed in the next chapter.

Chapter 5

Analysis of the *Drosophila* S2 Cell proteome

5.1 Summary

In this chapter, we have done an extensive analysis of the protein list provided by the iTRAQ experiments to help identify cellular processes, pathways or domains enriched in our study. We also compared our list with the published *Drosophila* protein interactome to identify the probable SUMOylation aided complex enrichments. We further compared our protein list with previously published list of SUMOylated proteins from different organisms and see a considerable overlap between them.

5.2 Introduction

5.2.1 Enrichment of SUMOylated proteins and their analysis in various organisms

Multiple proteomic studies have been carried out to identify SUMOyled substrates in different organisms in a wild type state or other physiological conditions making use of biochemical techniques along with sophisticated mass spectrometry tools. These large scale experiments have helped identify SUMOylated proteins in budding yeast (Denison et al., 2005; Hannich et al., 2005; Panse et al., 2004; Rosas-Acosta et al., 2005; Wohlschlegel et al., 2004; Wykoff and O'Shea, 2005; Zhou et al., 2004), human (Galisson et al., 2011; Golebiowski et al., 2009; Rosas-Acosta et al., 2005; Tatham et al., 2011; Vertegaal et al., 2006; Vertegaal et al., 2004), *C. elegans (Kaminsky et al., 2009)*, and *Drosophila* (Nie et al., 2009) among others. These proteome-wide studies provided great insights into functional role of SUMO and SUMOylation proteins. However, very few of these studies have been able to identify exact SUMOylation sites due to difficulty in mass spectrometry approaches to correctly identify the acceptor lysine residues.

Thus, in-silico identification of SUMO substrates with their respective sites is important to understand the mechanisms of SUMOylation-related regulations. These prediction softwares use published data to create algorithms which help detect SUMO acceptor lysine sites in proteins for functional analysis.

5.2.2 SUMO prediction software

The majority of the SUMOylation sites follow a consensus motif with ψ-K-X-E (ψ is a hydrophobic amino acid). However, the accumulating experimental data shows that a number of SUMOylated proteins do not have the standard motif and include various modifications of the standard consensus or completely different motifs. These unexpected features introduce the difficulties into the SUMOylation proteome analysis making the field of SUMO proteomics a great challenge.

There are a couple of in-silico SUMO site prediction softwares like SUMOsp, SUMOplot, etc. These prediction softwares use experimental datasets for analysis with either one or a combination of parameters including sequence information for motif prediction, secondary structures, and evolutionary conservation to predict SUMO acceptor lysines in proteins. Newer prediction models are also trying to consider the physico-chemical properties of the protein including hydrophobicity, buriality, isoelectric point, hydrophilicity, polarity, bulkiness, and molecular weight of residues as these might affect accessibility of SUMO and SUMO pathway components to lysine acceptor site in the substrate protein. Based on these parameters and thorough statistical evaluations these softwares provide prediction scores to the predicted lysines which might be the site of SUMO modification (Xue et al., 2006).

5.2.3 Gene Ontology databases

With increasing number of high-throughput proteomic experiments being done regularly it is important to develop tools to help analyze the data and help give it functional relevance. Many databases, which are publically available today, include KEGG, Panther, Ensembl, Swiss-Prot, and DAVID.

Gene ontology (GO) provides a common base of characterization of genes and proteins from different organisms into defined classes mainly based on biological process, cellular component and molecular function. This approach not only helps annotate genes and proteins in an organism to GO terms at varying levels of details but also provides a basis for comparison across species which might help in finding new genes and functions. Certain databases also provide tools to analyze the data and visualize it in an enrichment analysis or help in building gene networks. In summary, GO analysis predicts how a gene product may function within the cell.

There are a number of bioinformatic tools to analyze large datasets for gene ontology with a number of statistical tests. Protein analysis through Evolutionary Relationships (PANTHER) classification system is one such tool to mine through the biological data associated with genes and gene products for their accurate classification from experimental datasets and build phylogenetic trees which help extrapolate information from a few other model organisms (Mi et al., 2013; Thomas et al., 2003). It also describes biochemical pathway maps and conserved protein domain architectures, while being linked to a rich source of biological annotation. Database for Annotation, Visualization and Integrated Discovery (DAVID) is another such bioinformatics database (Huang da et al., 2009a, b). Both these tools are user-friendly and open source however, compared with DAVID, PANTHER has a few advantages. First, PANTHER integrates more updated GO curation data to build its network. Second, PANTHER enables users to analyze genome data from a larger number of organisms than DAVID. Third, the phylogenetic trees in PANTHER protein library makes it easier to make more accurate ortholog prediction.

5.2.4 *Drosophila* protein interactome

An extensive study on protein-protein interactions was published by Guruharsha *et al.* in 2011. In this study, protein interactors of nearly 5000 FLAG-HA epitope tagged *Drosophila* proteins were identified using a combination of co-affinity purification and mass spectrometry analysis in $S2R^{+}$ cell line. Using cell lines makes it easy to express tagged proteins and carry out extensive affinity purifications on a large scale. Stringent statistical analysis of this data helped define a complex network of individual protein-protein interactions. This led to the generation of a *Drosophila p*rotein *i*nteraction *m*ap (DPiM) encompassing 556 protein complexes. In addition to validating previously known interactions, it helped define potential novel members for several important protein complexes and assign functional links to 586 protein-coding genes lacking previous experimental annotation (Guruharsha et al., 2011).

5.2.5 Visualization through Cytoscape

One of the most popular softwares to visualize data relationships like overlap and exclusion between data sets is Cytoscape. It is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating with gene expression profiles and other state data (Shannon et al., 2003).

5.3 Materials and methods

5.3.1 GO analysis using DAVID and PANTHER

To investigate the affect of LPS treatment on different biological processes in 529SU cells, enrichment analysis was performed by comparing data using the DAVID Bioinformatics resources 6.7 and PANTHER Classification system 7.2. The GO analysis was performed on the 858 list of differentially expressed proteins in LPS induced versus un-induced control using *Drosophila melanogaster* gene database as the reference list. Statistically overrepresented GO categories were listed and then selected for further analysis.

The DAVID gene ontology analysis was carried out using an ease score of 0.1 and Benjamini correction. The statistical overrepresentation test of PANTHER was used to identify pathways and domains significantly enriched in our dataset using the Bonferroni correction. The DAVID and PANTHER sites used for analysis are listed below:

http://david.abcc.ncifcrf.gov/home.jsp

http://www.pantherdb.org/

5.3.2 Cytoscape analysis

The SUMO proteomic list was compared to the DPiM (Guruharsha et al., 2011)) using a DPiM input file generated by the authors (Supplementary material, Gurusharsha *et. al.* 2011). The *Drosophila* interactome was displayed using the open source bioinformatics software platform Cytoscape (Cline et. al., 2007).

5.3.3 SUMO site prediction

SUMO acceptor lysine and SBM prediction The SUMOylation site prediction was carried out using SUMOsp ver2.0 and SUMOplot™ analysis program provided by Abgent, and the SUMO binding site prediction was carried out using GPS-SBM 1.0.

The prediction software sites are listed below:

http://sumosp.biocuckoo.org

http://www.abgent.com/sumoplot

http://sbm.biocuckoo.org

5.4 Results

5.4.1 Gene Ontology analysis of the LPS-modified SUMO proteome

Gene ontology analysis of the 858 list using DAVID Bioinformatics resource shows 5% of the proteins to be involved in immune response related functions. Various other cellular functions like translation, actin cytoskeletal organization, cell redox homeostasis among others are well represented in the analysis (**Figure 5-1**). This suggests that immune challenge induces global changes within a cell in order to help it combat infection.

5.4.2 Analyzing the SUMO Proteome

Of the proteins we have identified, 102 proteins are common with the published SUMO proteome (a list of 150 proteins) from 0-3 hour *Drosophila* embryos and approximately 12% of the proteins in our list are represented in SUMO proteomes from other organisms.

On analyzing the fold changes it was observed that there is no significant change in the levels of SUMO itself in response to LPS treatment. However, the list indicates enrichment of the SUMO conjugation machinery components including Uba2 a subunit of the activating enzyme, Ubc9 and some of the *Drosophila* ligases like Su(Var)2-10 etc. This kind of enrichment is also observed in previous studies in mammalian cells and Arabidopsis under stress conditions.

The list of 858 proteins was analyzed using online open source programs as described in Materials & Methods. The aim was to mine the list for insights into the SUMO enriched proteome. The primary analysis looks at enrichment of proteins in the list as compared to a dataset of all *Drosophila* proteins. Fold enrichment (log of fractional difference observed vs expected) indicated in **Figure 5-2** that many biological processes had components that were regulated by SUMO. tRNA synthetases were shown to be highly enriched along with vesicle mediated transport processes including endocytosis and exocytosis, nuclear transport and redox homeostasis among others by PANTHER enrichment database tool. iTRAQ ratios of few of the identified proteins in tRNA synthetases and vesicle mediated transport are shown in **Figure 5-3**.

DAVID analysis was used to identify protein domains and pathways that are enriched in our list. The enriched protein domains may indicate a common protein fold as a target for

the SUMOylation machinery or may lead us to a protein domain family that is more prone to SUMOylation as compared to others in *Drosophila* and subsequently in eukaryotes. Our data predicts that amongst others, WD40 domain containing proteins, thioredoxin folds and members of the ATPase AAA+ family are possible targets in eukaryotes. A number of enzymatic/metabolic pathways involved in translation and degradation are also substantially enriched (**Figure 5-4**). Certain KEGG pathways that are enriched in the list are represented in **Figure 5-5**.

5.4.3 Global changes in SUMOylation – The SUMO interactome

A significant proportion of the proteins in our list may be SUMOylated. However, we performed the pulldown in native conditions to maintain very strong protein-protein interactions and hence complexes involved in immune responses. SUMOylation can act by regulating the interactions of the substrate protein with other macromolecules. SUMOylation of a substrate can create a binding site and/or modulate conformation leading to the enhancement or decrease in binding affinity. SUMO interaction motif's (SIM's) have been discovered pointing to a significant role for SUMO in enhancing (or reducing) protein-protein interactions, especially in relation to large protein complexes. Many previously published studies have proposed that multiple proteins can be SUMOylated in a single functional complex.

In order to analyze our list in terms of global protein interactions, we turned to a recent comprehensive study on protein-protein interactions in S2 cells, the same system we used for our studies (Guruharsha et al., 2011). Since our list represents a subset of proteins in S2 cells that is biased towards SUMOylation or SUMO interaction, we mapped our set of 858 proteins, ($1/5^{th}$of the proteins the DPiM network), onto the DPiM network. The comparison is not absolute as the methods used to generate the interactors involved are different, but the analysis leads to interesting findings. First, many of the major clusters/complexes shown in the wild-type, DPiM are missing in the SUMO-interactome (**Figure 5-6A**). This is a striking observation, especially considering the predicted global roles for SUMO as a mediator of protein-protein interaction.

A finer analysis for protein complexes within these clusters indicates that well represented complexes such as the Mediator Complex, SNARE/Syntaxin complexes and Arp/Arc protein complexes are underrepresented in our SUMO proteome. Other known

complexes, such as the tango complex as also the tRNA synthetase (MARS) complexes are however well represented (**Figure 5-6B, Table 5-1**).

Since tRNA synthetases have been shown to be enriched and represented in various different analyses, it would be interesting to study if a few or all of the tRNA syntethases are SUMO modified. Since immune response leads to an increase in SUMOylation of tRNA synthetases, they may be involved only in global upregulation of translation machinery or might have other moon-lighting roles in immunity.

5.4.4 The Immune SUMO proteome

The Immune SUMO Proteome should be a subset of the global SUMO proteome. An extensive literature survey identified 115 proteins, from our list of 1820 proteins, which could be implicated, directly or indirectly to the immune response. **Table 5-2** lists a few these proteins, their molecular function and their current known SUMOylation status. Many of the proteins discovered in our proteomic screen have not been shown to be SUMOylated, though many have putative SUMOylation motifs and/or SUMO interacting motifs. **Figure 5-5** represents a subset (twenty five) of these genes with their iTRAQ ratios. Proteins that have already been demonstrated to be SUMOylated, in flies or any other model organism, have been marked with arrowheads. Previous studies have shown physical SUMOylation of *Drosophila* Dorsal and STAT92E (Bhaskar et al., 2002; Gronholm et al., 2010) and mammalian orthologs of *jra* and *kay* (Bossis et al., 2005).

5.4.5 Comparison of Immune SUMO proteome with published immune transcriptome

The most dramatic effect of the initiation of infection is the transcriptional upregulation or down regulation of about 400 genes, defining the immune transcriptome (De Gregorio et al., 2001; Johansson et al., 2005). This regulation is a result of the activation of the two major immunity pathways, the Toll/NF-κB and the IMD/NF-κB pathways. Since PTMs are postulated to be immediate and dynamic, working on faster time scales to the slower transcriptional response, they should be distinct from the transcriptome especially at the early time points. A comparison of the two data sets, the SUMO Immune Proteome, with the Immune Transcriptome indicates a 1% overlap, even when we take the entire SUMO proteome, rather than the SUMO immune proteome into consideration. This result clearly underscores the differences between regulations by SUMOylation, and resultant changes in the transcriptome. The lack of overlap highlights the distinct spatiotemporal roles for

substrate proteins that are post-translationally modified in signaling. **Figure 5-7** shows iTRAQ ratios a few representative proteins involved in immunity.

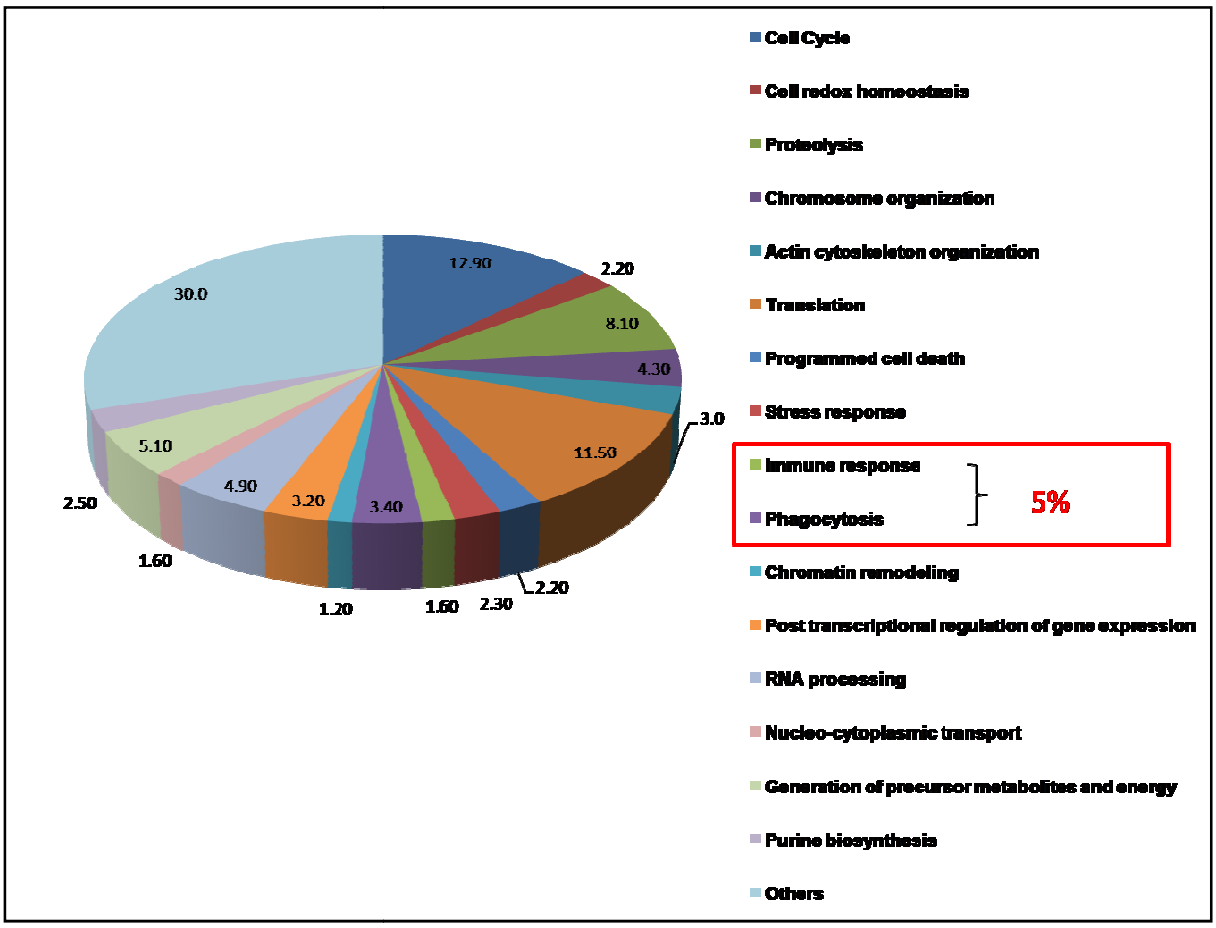


Figure 5-1: Gene Ontology analysis

A confident set of 858 proteins were selected with iTRAQ ratios <0.5 and >2.0 for further analysis. Gene Ontology analysis was performed using David Bioinformatics resource. The proteins identified are classified into various functional groups with the maximum representation in cell cycle and translation. 5% of the total proteins identified could be directly related to a function in the immune response including the regulation of the signaling pathways and phagocytosis.

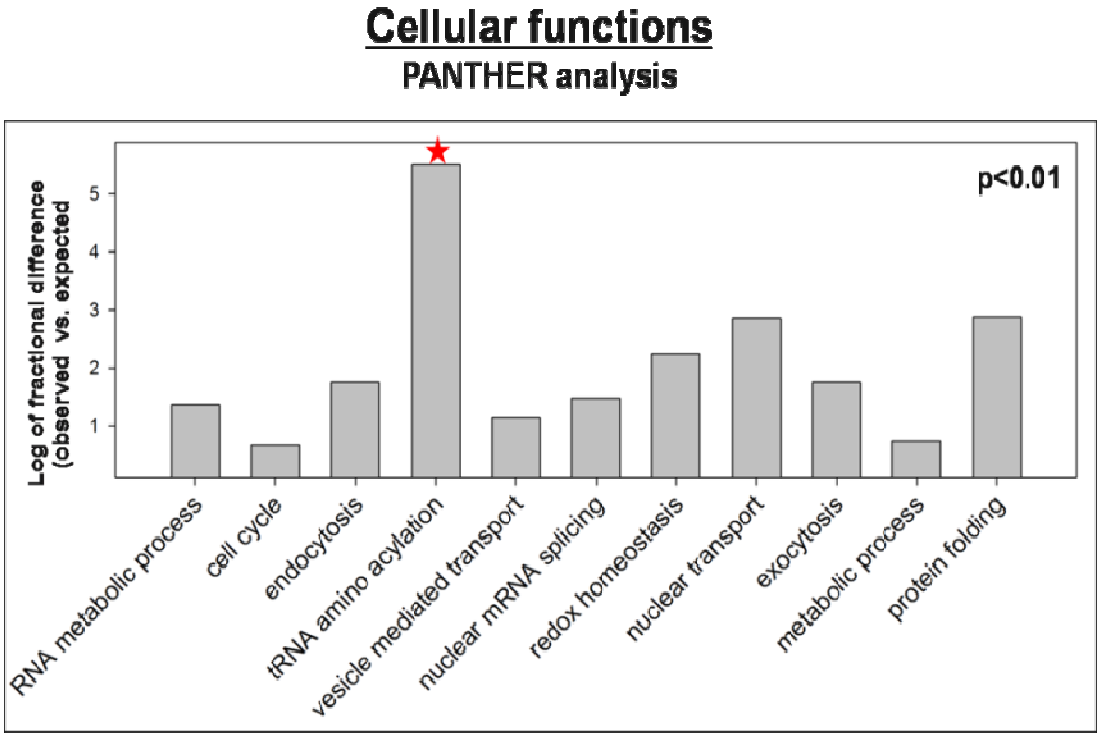


Figure 5-2: Cellular process enrichment using PANTHER

Fold enrichment normalized to a standard *Drosophila* data set, for protein function for few of the processes with a $p<0.05$. tRNA amino acylation is highly enriched in our dataset closely followed by nuclear transport, protein folding and redox homeostasis. The PANTHER resource was used to calculate the above values.

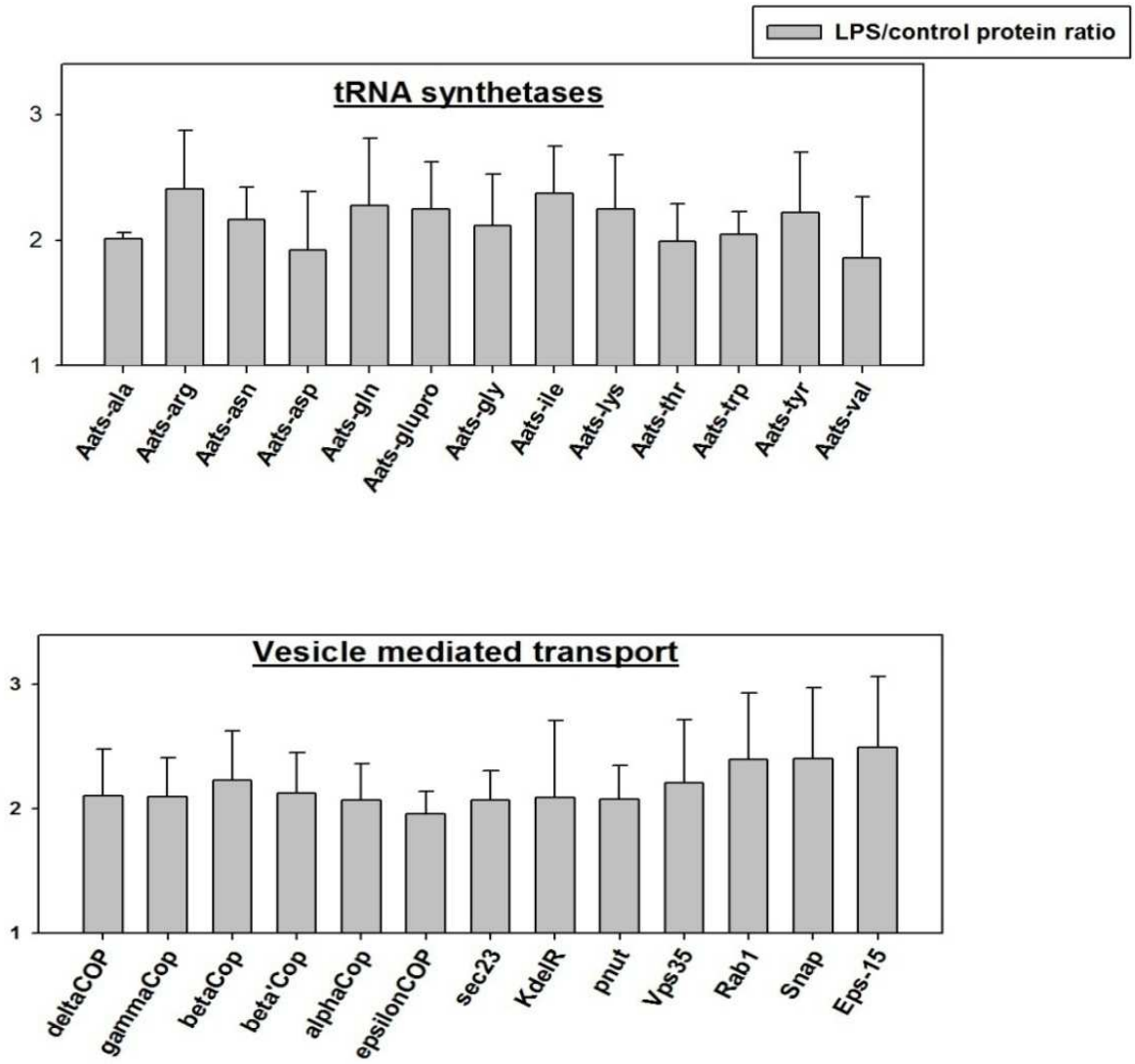


Figure 5-3: Representative proteins in enriched functional groups
LPS/control iTRAQ ratios of representative proteins of the two enriched classes: tRNA synthetases and vesicle mediated transport.

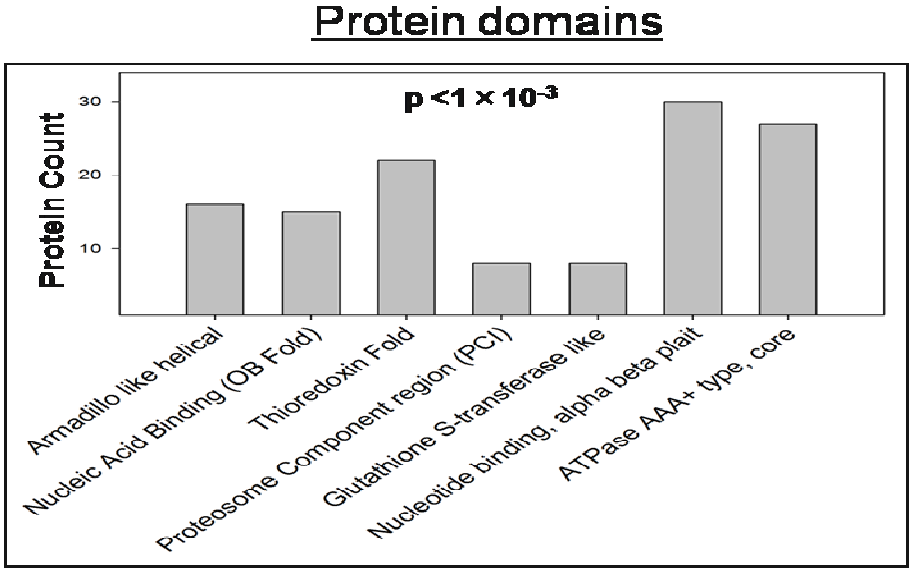


Figure 5-4: Protein domain analysis using DAVID

Number of proteins in the parts-list that show significant enrichment, normalized to a standard *Drosophila* data set, for protein domains as analyzed by the DAVID bioinformatics resource.

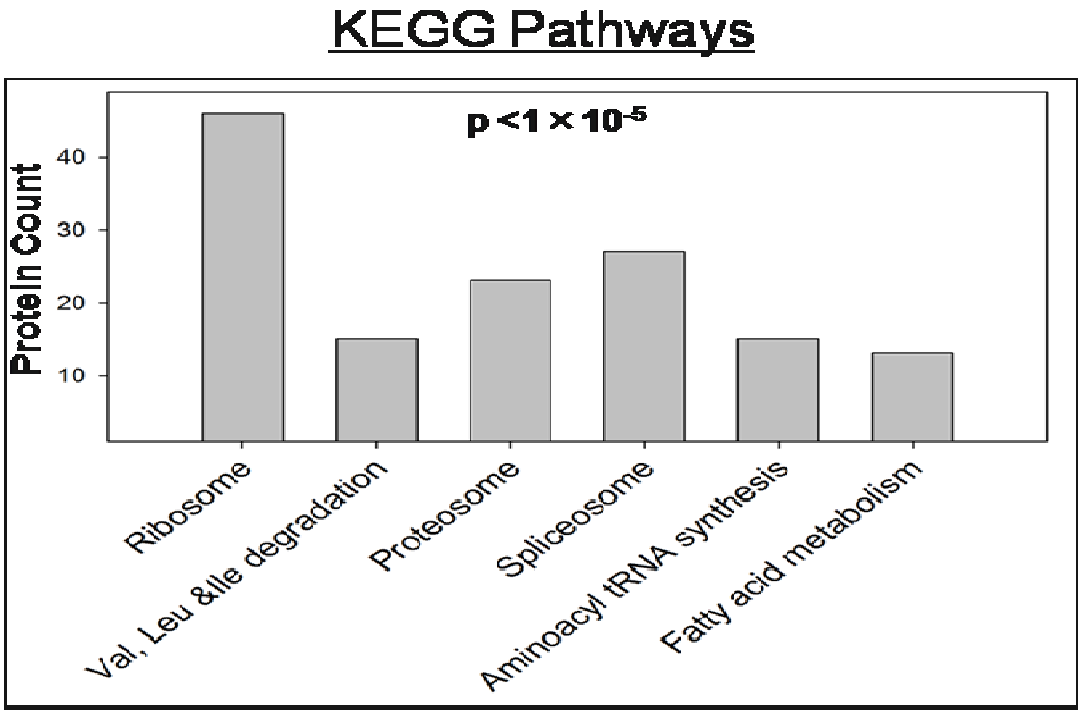


Figure 5-5: Protein domain analysis using DAVID

Number of proteins in the parts-list that show significant enrichment, normalized to a standard *Drosophila* data set, for KEGG pathways as analyzed by the DAVID bioinformatics resource.

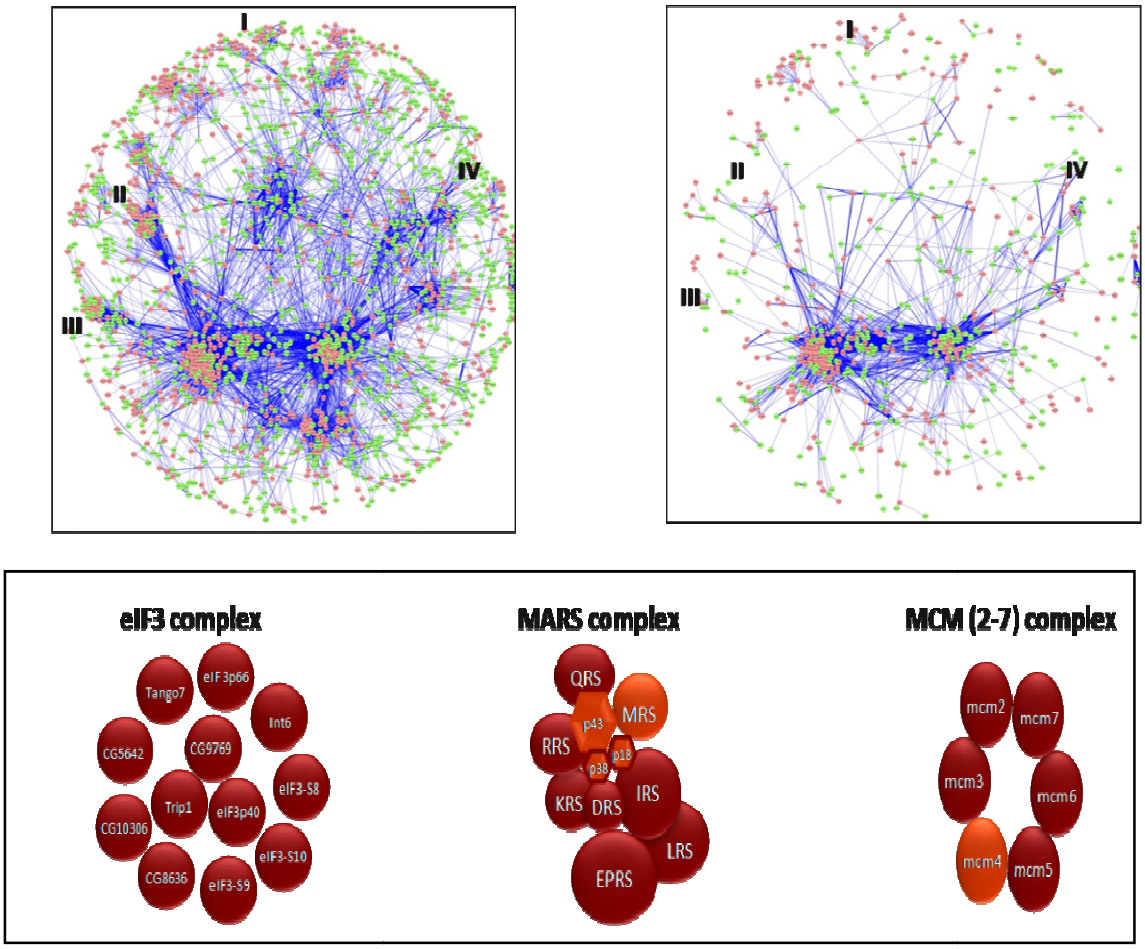


Figure 5-6: SUMO interactome and SUMO proteome

A. Cytoscape representation of a molecular interaction networks for Schneider cells, based on data from Guruharsha et. al, 2011. The figure on the left represents an interaction map of 4500 proteins (DPIM), as discovered by a large scale affinity purification experiment.

B. A combined SUMO enriched proteome of 858 proteins we have generated, mapped onto the DPIM map. Many complexes such as the Histone Acetyl Transferase Complex (I), Mediator Complex (II), the SNARE/Syntaxin Cluster (III) and the Arp/Arc complex (IV) are under-represented in the SUMO enriched network.

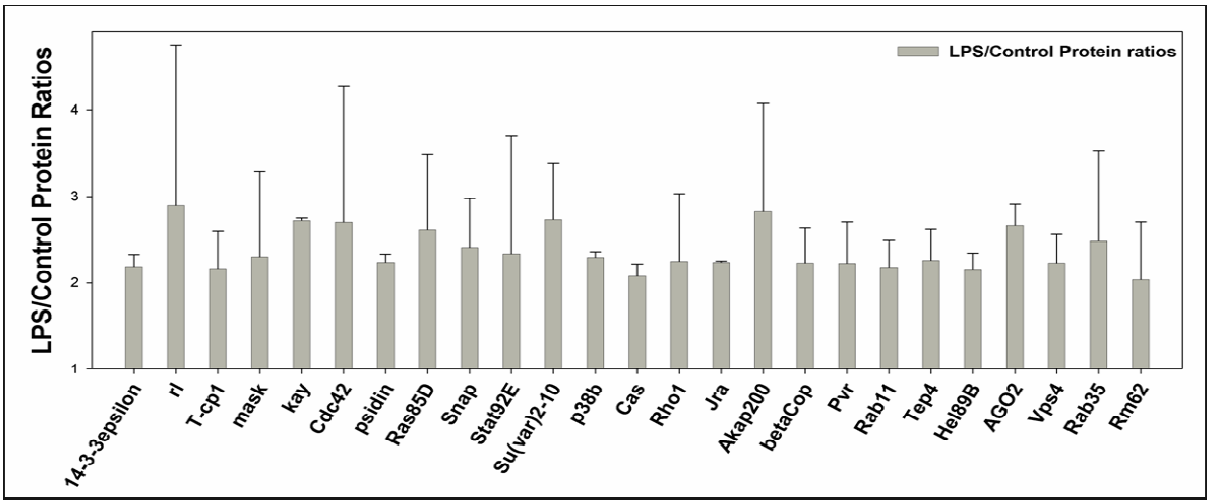


Figure 5-7: SUMO Immune proteome
iTRAQ ratios for a subset of immune related proteins in our parts-list.

Table 5-1: Examples of Protein Complexes in the SUMO proteome

Complex (number of members)	Proteins listed in SUMO Proteome	Not identified in SUMO proteome
RNA processing-Exosome (RNase Complex; 9)	Rrp6, Dis3, Rrp42, Rrp4, RRp40. Rrp46 (6)	Ski6, Mtr3, Csl4 (3)
SNAP-SNARE Complex (31)	Snap, Nsf2 (2)	Syx16, usnp,Syx1A, Slh, Use1, gammaSnap, Slh, membrin, Snap25,Sec22, Syx8, Snap24, Syx5, CG1599, Ykt6, CG2023, Syx13, Syx4, Koko, Syx18, Vti1,Syb, Syx, Bet1,CG6208, Rme-8, AttD, Syx17, n-syb (29)
Proteosome Complex (50)	Rpn12, Uch-L3, Pros54, Rpn9, Rpn7, CG13349, Prosbeta2, Prosalpha7, Rpn5, Mov34, Pros45, Prosbeta5, Pros29, Pros35, Rpn3, Rpn1, Pros26.4, Prosbeta3, Prosbeta7, Rpn2, Rpt3, Rpn6, Tbp-1, Rpn1, CG17331, Pros26, Prosalpha5, Rpt4, REG, Ufd1-like, Rpt1 (31)	CG11885, CG2036, Prosbeta4R2, Prosbeta2R1, pomp, Prosalpha6T, Prosbeta4R1, Prosbeta1, pros28, Pros25, Prosalpha1, rpr, CG12321, CG2046, CG13319, GNBP2, CG11885, CG3812, CG9588 (19)
Escrt Complexes (21)	RAB11, RAB7, TSG101, VPS4, RAB35, RAB8, RAB4, VPS28 (8)	RAB5, HRS, VPS23, VPS37, VPS36, VPS25, VPS20, VPS60,VPS46, VPS24, VPS2, VTA1, SNF7 (13)
Ribosomal Protein Complex (Small Subunit)	Sta, RpS2, RpS3, RpS3A, RpS4, RpS6, RpS7, RpS8, RpS9, RpS10b, RpS11, RpS12, RpS13, RpS14b, RpS15Aa, RpS16, RpS17, RpS18, RpS19a, RpS20, Rps21, RpS23,	RpS5a, RpS5b, RpS10a, RpS14a, RpS15, RpS15Ab, RpS19b, RpSoho23b, RpS27a, RpS28a, RpS30 (11)

(40)	RpS24, RpS25, RpS26, RpS27, RpS28b, RpS29 (29)	
Mitochondrial Ribosomal Protein Complex (Large Subunit) (47)	mRpL2, mRpL17, mRpL19, mRpL41 (4)	mRpL1,mRpL3,mRpL4,mRpL9,mRpL10, mRpL11,mRpL12, mRpL13,mRpL14 mRpL15,mRpL16, mRpL18, mRpL20, mRpL21, mRpL22, mRpL23,mRpL24, mRpL27, mRpL28, mRpL30, mRpL32, mRpL33, mRpL34, mRpL35, mRpL36, mRpL37, mRpL38, mRpL39, mRpL40, mRpL42, mRpL43, mRpL44, mRpL45, mRpL46, mRpL47/Rlc1, mRpL48,mRpL49,mRpL50,mRpL51, mRpL52, mRpL53,mRpL54, mRpL55 (43)

Table 5-2: Representative immunity related hits from the iTRAQ data set, sorted alphabetically

Protein	Molecular function	Lysine residues predicted to be SUMOylated[#]	Predicted SUMO Binding Motifs (SBMs)[€]	Previously demonstrated to be SUMOylated
14-3-3 ε	Involved in signaling and protein transport	-	-	-
AGO2	Endonuclease involved in siRNA mediated silencing	1048, 1193	-	-
Basket	Jun N-terminal kinase	316	-	-
Caspar	Involved in anti-microbial response	436, 551	211-214, 324-327	-
Cdc42	GTP binding protein	-	-	-
Colt	Transport protein involved in phagocytosis	69	-	-
Dos	Adaptor protein in Sevenless signaling	781	-	-
Drk	SH3/SH2 adaptor protein in sevenless signaling	-	178-181	-
Epsilon-Cop	Vesicular transport protein	36, 359, 547	297-300, 826-829	-
Hel89B	ATP-dependent	84, 341, 596,	879-882	-

	DNA helicase activity	600, 606, 624, 964		
Hrs	Involved in endocytosis	263, 747	-	-
IMD	Involved in signal transduction of immune response	-	-	-
Jra	JNK transcription factor	29, 190, 214, 248	-	-
Kay	JNK transcription factor	533	-	-
Listericin	Anti-microbial peptide	-	-	-
Mbo	Nuclear Transport factor	240, 557	-	-
Ntf-2	Nuclear transport factor	-	-	-
p38b	MAP kinase	-	-	-
Psidin	Involved in phagocytosis	236, 909	-	-
Pvf2	vascular endothelial growth factor receptor	282	-	-
Pvr	Transmembrane receptor protein tyrosine kinase	67, 129, 883, 944	277-280	-
Rab11	Rab family GTPase required in endocytic recycling	-	-	-

5.5 Discussion

Protein SUMOylation is an essential cellular process conserved from yeast to mammals and plays an important role in the regulation of intracellular trafficking, cell cycle, DNA repair and replication, cell signaling and stress responses. In our study, we performed high-throughput proteomics to identify LPS induced SUMOylated substrates and their interactions in *Drosophila* S2 cells. On further analysis our study helps reiterate the fact that SUMO modification is important in regulating diverse cellular processes as seen in our gene ontology analyses. The enrichment of translation, actin cytokeleton remodeling and vesicle mediated transport might not seem directly related processes to combat pathogens, however each of these processes might be important in global responses in the cell like translation upregulation for rapid or excess production of certain effector proteins, actin cytoskeletal remodeling and vesicle mediated transport for internalization of pathogens via phagocytic responses by the cell. All these processes being regulated by SUMO in totality bring about effective immune responsiveness.

A recent study by Ezekowitz and colleagues has provided an extensive list of protein-protein interactions associated with *Drosophila* phagosomes (Stuart et al., 2007). Latex-bead-containing phagosomes were isolated from S2 cells to identify the proteins involved in formation of the phagosome by tandem mass spectrometry analysis. A total of 617 proteins were identified and roles of 214 of them were further confirmed through RNAi based studies. Comparison of our list with the 617 list of proteins shows an overlap of 128 proteins. These include a number a coat proteins (alpha COP, beta' COP and delta COP), Rabs (Rab1, Rab7, Rab8, Rab10 and Rab11), actin regulator proteins (14-3-3ε, Actin57B, Actin5C and Ter94) and chaperonin-containing T complex proteins (T-CP1, CCT5, CCT gamma, CG7033 and CG8231) among other vesicle trafficking regulators.

Another study to identify components and regulators of the IMD pathway was performed by Hoffman and colleagues. They used 11 tagged proteins of the IMD pathway to identify novel interactors in heat-killed *E. coli* stimulated *Drosophila* S2 cells using mass spectrometry analysis. The study identified 369 proteins and their corresponding 219 genes representing the "IMD interactome" (Fukuyama et al., 2013). 122 proteins from our list are represented in their study. These include previously identified proteins like KAY (Kayak), PVR (PDGF- and VEGF-receptor related), MASK (multiple ankyrin repeats single KH domain), OST48 (Oligosaccharyl transferase 48kD), EIF-2α (eukaryotic translation Initiation

Factor 2α), and AGO2 (Argonaute 2) etc. Their studies identified novel proteins cindr and RPS3 (Ribosomal protein 2) which have been implicated in immune responses in mammals with as yet unidentified roles in the context of *Drosophila* immunity. They also identified proteins Pontin and Reptin which associated with chromatin remodelling complexes. They are constitutively associated with the IKK complex and their RNAi mediated knockdown in S2 cells lead to a significant reduction in NF-κB reporter activity on stimulation with heat-killed *E. coli*. All these new proteins identified as part of the IMD interactome show high prediction lysine sites for SUMO modification and serve as interesting targets to be studied in the light of immunity.

Small nuclear Ribonuclear proteins (snRNPs) form a protein complex with RNA to form a spliceosome which is involved in processing of pre-mRNA. This entire complex consists of 7 Sm proteins (SmD1, SmD2, SmD3, SmE, SmG and SmF), all of which are present in list and are within top 50 of the LPS/control iTRAQ ratios. Almost all the tRNA synthetases are also represented in our list with high ratios detected for a few of them. Both these protein complexes are essential in protein translation which probably is affected drastically during immune response, however, it would be interesting to validate their roles in regulating the humoral or cellular responses within the cell, if any.

Enrichment analyses showed a number of protein domains to be over-represented in our list. Studies are required to link presence of these enriched domains to increase in probability of SUMO modification of these proteins and thus will provide an interesting insight into SUMO modification.

Keeping into account this extensive analysis, it was further important to validate SUMOylation of interesting protein candidates whose roles have been determined in immunity or other biological processes for further functional studies. This is discussed in the next chapter and Appendix II.

Chapter 6

Validation of immune related targets discovered in the iTRAQ screen

6.1 Summary

In this chapter, we perform validation studies to depict SUMOylation of a few out of the 858 proteins in bacteria and S2 cells. Bacterial validations showed 6 out of 10 proteins tested to be SUMOylated. However, in S2 cells we are yet to standardize conditions to detect SUMOylation.

6.2 Introduction

The SUMOylation process is extremely dynamic and not all real SUMO substrates will be SUMOylated in vivo simultaneously. Only a small fraction of the substrate, about<1%, is SUMOylated in vivo at any given time. Hence to identify SUMOylated proteins through mass spectrometry analysis and further show their physical SUMOylation in-vivo continues to be a challenge in the field.

Previous studies have made use of in-vitro SUMOylation tool to demonstrate SUMOylation of the protein of interest. This includes purifying enzymes of the SUMO conjugation machinery and the target protein separately and then adding them together in a reaction aided with ATP to trigger SUMOylation. Recently, an in-bacto SUMOylation system was developed by Courey and colleagues. Since bacteria lack the SUMO machinery, a plasmid construct with His_6-tagged SUMO, activating enzymes SAE1/SAE2 and Ubc9 is used to overexpress components of the SUMO machinery in bacteria (Nie et al., 2009). When these are expressed along with a protein of interest, the protein gets SUMOylated. Since it is possible to express huge amounts of protein using the bacterial expression system, the chances of obtaining the SUMOylated form of the protein increases. This system might however not work in conditions where certain other modifications present only in-vivo conditions would determine SUMOylation of the target protein.

To increase the likelihood of SUMOylation of a protein in-vivo, usually the protein is overexpressed with an affinity tag which would help in enrichment of the protein and its modified form. Studies have also overexpressed SUMO conjugation pathway components in cultured cells to increase SUMOylation of the protein. We used both bacteria and S2 cells to detect SUMOylation of target proteins under varying conditions.

6.3 Materials and Methods

6.3.1 Cloning

Ten genes were cloned into pGEX-4T1 vector for bacterial expression. 14-3-3ε and cdc42 were cloned into pGEX-4T1 at BamHI and SalI restriction sites. Homologous recombination based cloning was used to clone the remaining genes. Specific CDS were PCR amplified from cDNA obtained from S2 cell RNA or BDGP gold collection. Positive clones were confirmed using colony PCR and further sequenced. The sequencing result matches with Flybase cDNA sequence for both the constructs.

Table 6-1: List of primers used for cloning in pGEX-4T1

Sr. No	Primers	Sequence 5'- 3'
1.	pGEX FP	GCGGCCGCATCGTGACTGACTGACGA
2.	pGEX RP	GAATTCCGGGGATCCACGCGGAACCAG
3.	14-3-3ε FP	GATGGATCCATGACTGAGCGCGAGAACA
4.	14-3-3ε RP	CTTGTCGACTTACGACACGTCCTGATCCTC
5.	cdc42 FP	GATGGATCCATGCAAACCATCAAGTGCGT
6.	Cdc42 RP	CTTGTCGACTTATAAGAATTTGCACTTCCTTTTC
7.	Rolled FP	CTGGTTCCGCGTGGATCCCCGGAATTCATGGAGGAATTTAATTCGAG

		CGGATC
8.	Rolled RP	TCGTCAGTCAGTCACGATGCGGCCGCTTAAGGCGCATTGTCTGGTTG TCG
9.	Basket FP	CTGGTTCCGCGTGGATCCCCGGAATTCATGACGACAGCTCAGCACCA ACA
10.	Basket RP	TCGTCAGTCAGTCACGATGCGGCCGCCTACCGCGTTCTATTATTTGT ATTGTG
11.	mbo FP	CTGGTTCCGCGTGGATCCCCGGAATTCATGTCGCTCACCGATGTCTT GGAAT
12.	mbo RP	TCGTCAGTCAGTCACGATGCGGCCGCTTAGATGCCAACGATTTTATT AATGCGC
13.	Rab11	CTGGTTCCGCGTGGATCCCCGGAATTCATGGGTGCAAGAGAAGACG AGTACGA
14.	Rab11	TCGTCAGTCAGTCACGATGCGGCCGCTCACTGACAGCACTGTTTGCG CAC
15.	p38b FP	CTGGTTCCGCGTGGATCCCCGGAATTCATGTCGCGCAAAATGGCCAA ATTC
16.	p38b RP	TCGTCAGTCAGTCACGATGCGGCCGCTTACTGCTCTTTGGGCAGGAG CTCAG
17.	cpa FP	CTGGTTCCGCGTGGATCCCCGGAATTCATGGAGCAGACACCGATCAC CGATG
18.	cpa RP	TCGTCAGTCAGTCACGATGCGGCCGCTTATTGCGTCTTCAGTTCCTTG CCAA
19.	βCop FP	CTGGTTCCGCGTGGATCCCCGGAATTCATGACGTCGCAAGTGCCGTG CTACACG
20.	βCop RP	TCGTCAGTCAGTCACGATGCGGCCGCCTAGGCCGCCTGCACCGACTG CTTC
21.	Snap FP	CTGGTTCCGCGTGGATCCCCGGAATTCATGGGTGACAACGAACAGA AGGCGC
22.	Snap RP	TCGTCAGTCAGTCACGATGCGGCCGCTTATCGCAGATCGGGATCCTC GTCC

6.3.2 Bacterial expression to check for SUMOylation and GST pulldowns

The proteins were expressed and purified by using *E. coli* BL21 (DE3) strain. For detecting SUMOylation of the proteins, the pGEX-4T1 constructs were co-transformed with either Q^{SUMOGG} or $Q^{SUMO\Delta G}$ vector. The Q vectors are a kind gift from the Courey Lab. These vectors help express all the components of the SUMO machinery; SAE1, SAE2, Ubc9, and His-tagged $SUMO^{GG}$/ $SUMO^{\Delta G}$, sufficient to carry out SUMO modification of proteins in-bacto (Nie et al., 2009). A single, transformed, isolated colony of E. coli BL 21 (DE3) was inoculated in 1 ml LB medium and grown overnight at 37° C with vigorous shaking (200-250 rpm). 0.5 ml inoculum was added to 50ml LB broth and allowed to reach 0.8 O.D.600. Protein expression was induced by adding 1mM isopropyl-D-thiogalactoside (IPTG) growing the cells at 37°C for 3 hours. The cells were lysed by sonication in lysis buffer (50 mM Tris (pH 7.5), 20 mM β-mercaptoethanol (BME), 500 mM NaCl, 0.1% Triton X-100) with protease inhibitor cocktail (Roche). The culture was then centrifuged at 15000 rpm for 30 min at 4° C and the supernatant containing soluble protein was used for further purification.

The supernantant was incubated with 200µl of glutathione beads overnight at 4°C on an end-to-end shaker. The beads were further washed with lysis buffer 5 times to remove contaminants and then boiled at 95°C for 15min with 1X SDS Laemmli buffer. The beads were then centrifuged at 15000rpm at 4°C for 15min and elute was collected.

6.3.3 S2 cell transfections

Good quality DNA preparations were made using Qiagen midi kit to carry out transfections. Cells were split and grown to 50-60% confluency. Transfections were then carried out in 12 well plates using TransIT-2020 transfection reagent (Mirus Bio LLC) following manufacturer's protocol. 0.5µg of each construct was added per ml of cells. These cells were later induced with 500 mM $CuSO_4$ and kept for 48 hrs post which the cells were collected, washed with 1X PBS and boiled in 1X SDS Laemmli buffer.

6.3.4 SDS PAGE and Immunoblotting

Bacterial protein samples as well as S2 cell lysates were separated on a 10% SDS-PAGE gel followed by western blotting as described in section 3.2.6. The blots were probed with anti- His mouse antibody (Santacruz) in 1:1000 dilution for 2 hrs and with anti- GST mouse antibody (Santacruz) in 1:5000 dilution for 1 hr for in-bacto SUMOylation. Anti-FLAG rabbit antibody (Sigma) was used in 1:1000 dilution.

6.4 Results and Discussion

6.4.1 Validation of target proteins using in-bacto SUMOylation

The S2 cell SUMO proteome consists of a large number of proteins. In order to gain confidence on the ability of the screen to identify genuine SUMO substrates, we need to prove that the list contains a large proportion of SUMOylated targets. Literature from the SUMOylation field clearly indicates that a very small proportion of the total substrate is SUMOylated at any given time, indicating that a SUMOylated species may not be detected by our methods, even after enrichment of substrate.

We attempt to validate the targets shown in **Figure 6-1**. In order to maximize our chances for demonstrating SUMO modification, we used the *in-bacto* system (See Materials & Methods) developed by the Courey Lab (Nie et al., 2009). Bacteria lack a SUMO deconjugase and hence this system has the advantage of preserving SUMOylated species. Bacteriaal overexpression system also helps to produce milligram amounts allowing detection of the SUMOylated species Utilization of *in-bacto* SUMOylation enhances our ability to demonstrate physical SUMOylation, however does not guarantee it. We tested fifteen proteins and demonstrated SUMOylationof seven of the proteins, namely 14-3-3, cdc42, Jra, p38b, caspar, rab11 & rolled *in-bacto*. Representative examples are pictured in **Figure 6-1**. Bands for protein expression are seen in the anti-GST western blots, along with faint bands at +20kD indicating a possible SUMOylated species (*). The presence of the SUMOylated species could be confirmed by the presence of a band in the anti-6X-His western at the expected molecular weight. SUMOylation was possible in the presence of mature SUMO (SUMO-GG) but not when a defective version of SUMO (SUMO-ΔGG) was used.

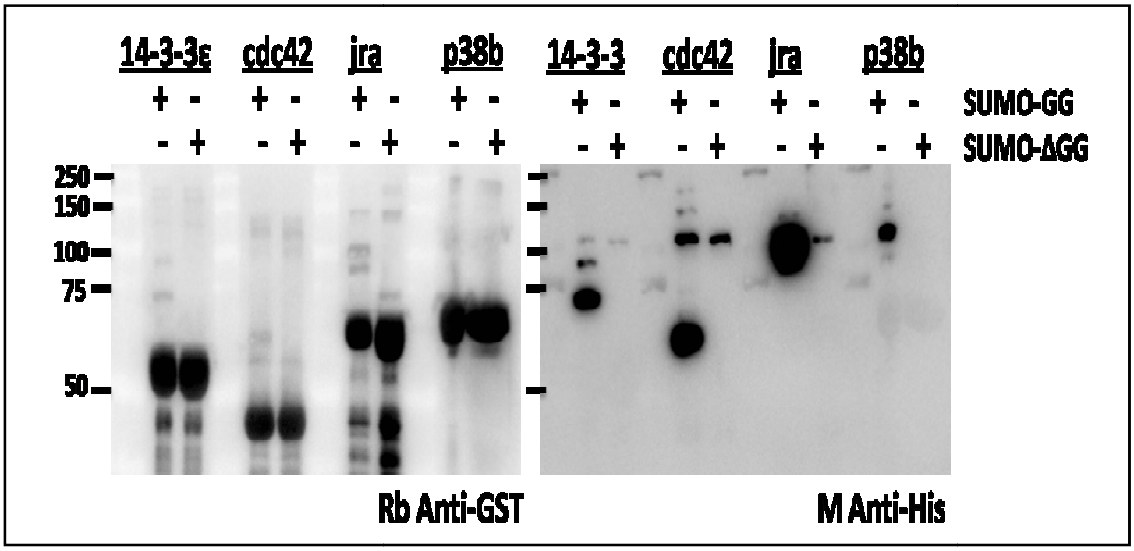
6.4.2 Validation of target proteins using S2 cells

Both S2 and 529SU cells were used to detect SUMOylation of target proteins within the cell. HA-FLAG tagged constructs were obtained from *Drosophila* DGRC collection. These constructs were expressed in S2 cells or 529SU cells (over-expressing the SUMO machinery) to help detect SUMO modified forms of the target protein. The different proteins tested are listed in **Table 6-2**. To aid in detecting SUMO modified forms of the tagged protein different approaches were used: 1) FLAG immune-pulldowns to enrich SUMOylated species, 2) Ulp1 knockdown to reduce deSUMOylation of the target protein, 3) Knockdown

of endogenous protein using UTR specific dsRNA. This would allow SUMO to modify only the tagged construct to enrich it SUMOylation.

However, we are still to standardize ideal conditions under which SUMO modified forms of the target proteins would be detected. A few representative validation conditions are discussed below in **Figures 6-2, 6-3, 6-4 and 6-5**. We did not observe SUMO modified bands in any of the conditions used.

A.



B.

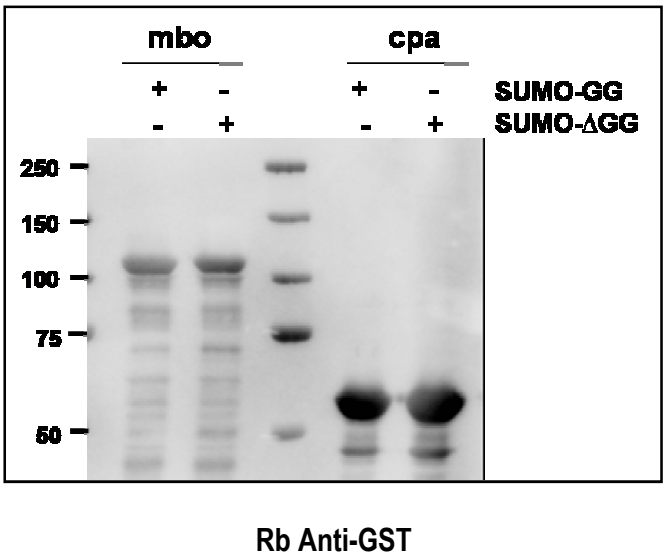


Figure 6-1: Bacterial validations of target proteins

The system used for validation is the in-bacto 'Q' system (Nie et. al., 2010). Proteins to be validated are co-expressed as GST fusions in bacteria along with 6XHis-SUMO-GG (or 6XHis-SUMO-ΔGG), E1 and E2 enzymes. SUMOylated proteins can be identified by the presence of a weak, higher molecular weight band (20 kD or more) that cross-reacts with the Anti-His antibody. A) 14-3-3ε, cdc42, jra and p38b show a SUMO modified band in the SUMO-GG lane as compared to the SUMO-ΔGG lane, hence confirming that these proteins are indeed SUMOylated. B) mbo and cpa proteins do not show any SUMO modified form as seen in the Anti-GST westerns. The Anti-His westerns for these were blank and did not show any bands for SUMO modified proteins.

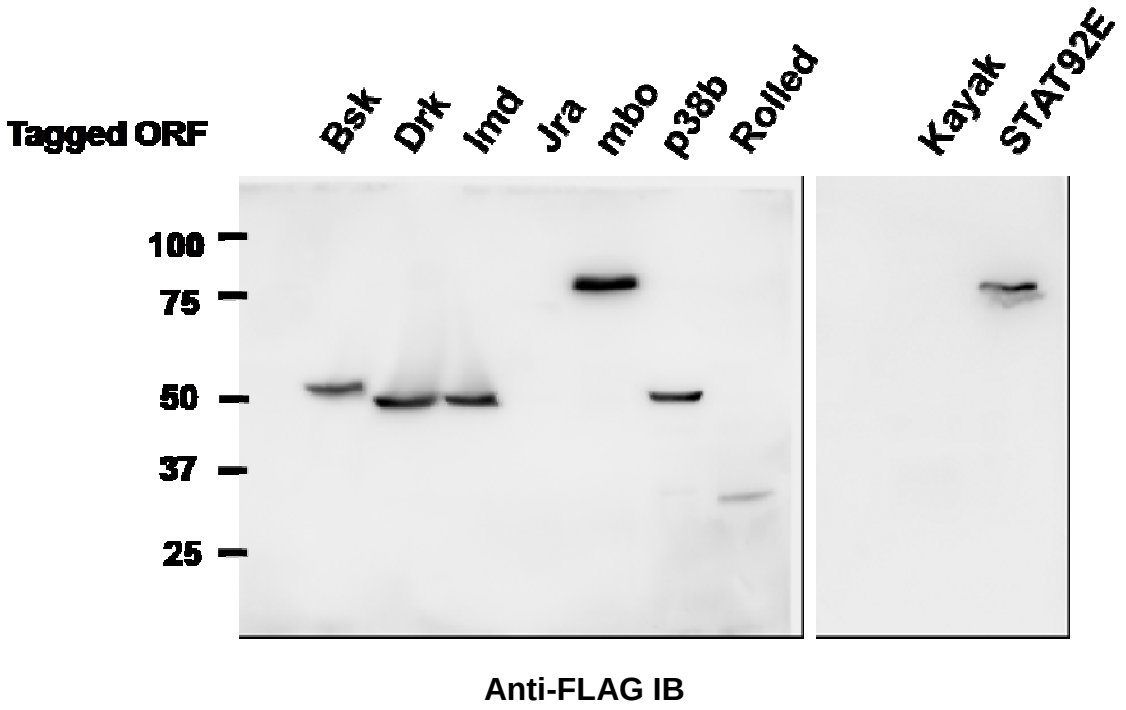


Figure 6-2: S2 validations to show SUMOylation of target proteins by overexpression
S2 cells were transfected with tagged ORF constructs with HA and FLAG tag. These cells were further lysed in 1X SDS Laemmli Buffer and transferred onto a PVDF membrane. The proteins are expressed at correct molecular sizes however, none of the proteins show an additional SUMO modified form as seen using rabbit anti-FLAG antibody.

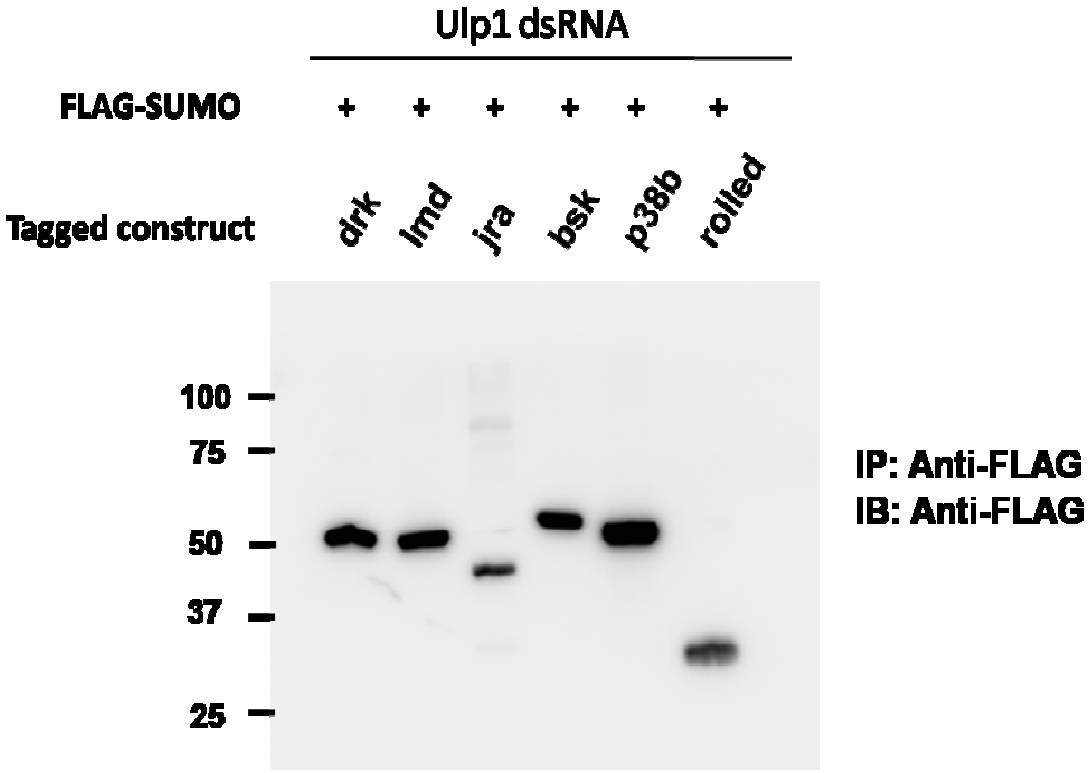


Figure 6-3: S2 validations to show SUMOylation of target proteins using Ulp1 knockdown

Further, S2 cells were transfected with tagged ORFs along with FLAG-SUMO construct. These constructs were expressed in cells post 72 hrs of Ulp1 dsRNA treatment to knockdown SUMO and reduce deSUMOylation of proteins within the cell. Over-expression of SUMO might enhance the SUMOylation of substrate. These cells are further lysed and the supernatant is used to carry out immuno-pulldowns using mouse anti-FLAG agarose (Sigma). These beads were boiled in 1X SDS Laemmli Buffer and the supernatant was transferred onto a PVDF membrane. The proteins are pulled-down as seen in rabbit anti-FLAG immunoblots with no SUMO modified forms of the proteins detected.

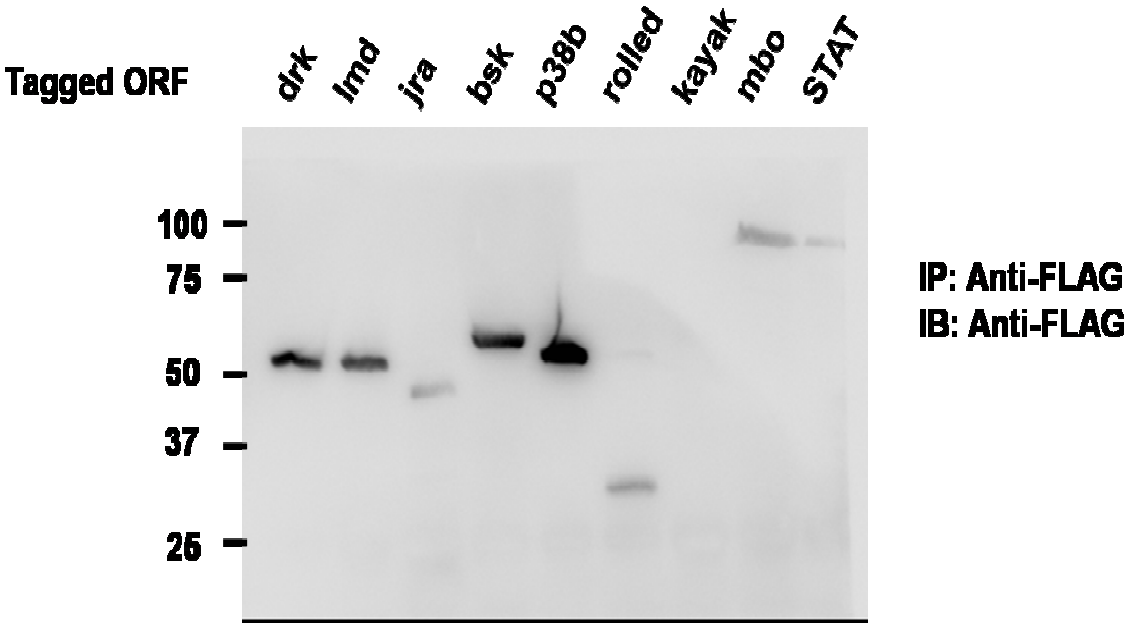


Figure 6-4: Validations to show SUMOylation of target proteins in 529SU cells
529SU cells were transfected with tagged ORFs to increase their likelihood of SUMOylation due to the overexpression of SUMO pathway components (SUMO and Ubc9). These cells are further lysed and the supernatant is used to carry out immuno-pulldowns using mouse anti-FLAG agarose (Sigma). These beads were boiled in 1X SDS Laemmli Buffer and the supernatant was transferred onto a PVDF membrane. The proteins are pulled-down as seen in rabbit anti-FLAG immunoblots with no SUMO modified forms of the proteins detected.

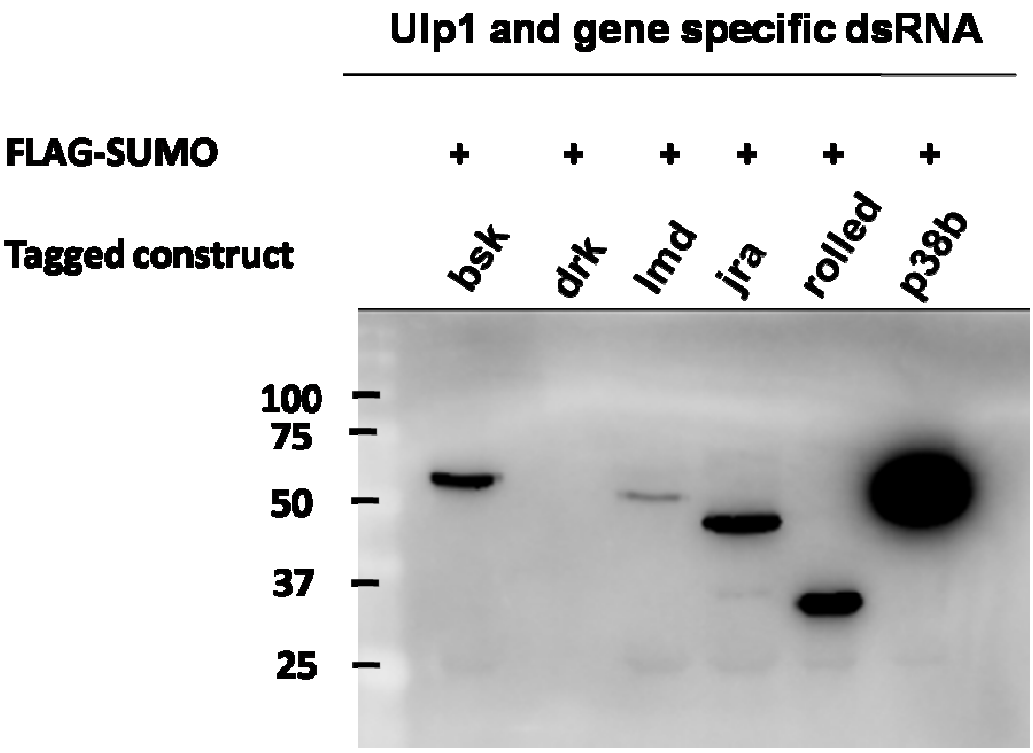


Figure 6-5: S2 cell validation to show SUMOylation of target proteins using Ulp1 and gene specific knockdowns

Further, S2 cells were transfected with tagged ORFs along with FLAG-SUMO construct. These constructs were expressed in cells post 72 hrs of Ulp1 and gene specific UTR dsRNA treatment to knockdown SUMO to reduce deSUMOylation of proteins within the cell and to knockdown endogenous protein to increase the pulldown and detection of tagged constructs. These cells are further lysed and the supernatant is used to carry out immuno-pulldowns using mouse anti-FLAG agarose (Sigma). These beads were boiled in 1X SDS Laemmli Buffer and the supernatant was transferred onto a PVDF membrane. The proteins are pulled-down as seen in rabbit anti-FLAG immunoblots with no SUMO modified forms of the proteins detected.

Table 6-2: List of proteins tested to depict SUMOylation in S2 cells and 529SU cells

Sr. No.	Name	Function
1.	cpa	Actin filament organization
2.	dos	Regulation of Ras pathway
3.	mbo	Nuclear export factor and has roles in immunity
4.	Ntf2	Nuclear transport protein
5.	STAT92E	Transcription factor in JAK-SAT pathway
6.	drk	Component of Ras pathway
7.	Rolled	Drosophila ERK and component of the MAPK pathway
8.	TRAM	Phagocytosis
9.	TM9SF4	Phagocytosis
10.	caspar	Regulation of IMD pathway
11.	SmD3	mRNA splicing
12.	Aos1	SUMO pathway activating enzyme
13.	CtBP	Transcription co-factor involved in a number of processes
14.	CG6084	Unknown
15.	CG3939	Unknown

16.	qkr58E-1	RNA binding protein
17.	lwr	SUMO conjugating enzyme
18.	lola	Implicated in a lot of processes including axon guidance, immunity
19.	RpS6	Ribosomal protein
20.	Rm62	mRNA splicing and implicated in immunity
21.	RpS19	Ribosomal protein
22.	snRNP-U1-70K	mRNA splicing
23.	mms19	unknown
24.	Chd64	Probable actin binding function
25.	CG1171	Involved in neuropeptide signaling pathway
26.	CG13349	Involved in proteasome mediated degradation
27.	p38b	Component of MAPK pathway and involved in immunity
28.	jra	Component of the JNK pathway and negative regulator of IMD pathway
29.	coro	Involved primarily in actin cytoskeleton remodelling
30.	cindr	Involved primarily in actin cytoskeleton remodelling
31.	msp300	Actin filament organization
32.	cortactin	Involved primarily in actin cytoskeleton remodelling

Chapter 7

Does SUMO regulate Imd Signaling?

7.1 Summary

In this chapter, we show that Caspar and jra – negative regulators of the IMD pathway was modified by SUMO. Lysine 551 is identified as the SUMO site for Caspar, and lysine 190 is one of the SUMO acceptor sites for Jra. We made over-expression constructs for Caspar mutant and wild type to look for its affect on Relish cleavage post LPS treatment.

7.2 Introduction

The IMD pathway is one of the two major pathways involved in immune response in *Drosophila*. The pathway upon activation leads to phosphorylation and eventual endo-proteolytic cleavage of Relish mediated by DREDD. The N-terminal 68 kDa fragment of Relish enters the nucleus to up-regulate AMP and other target genes. On LPS treatment in mbn2 cells, full length Relish and its cleaved fragments show a distinct kinetics over time. The 110 kDa Relish band disappears within 30sec of LPS treatment and is detectable after 45 min. The levels of the 68 kDa and 49 kDa bands also decrease and then increase over time as more full length Relish is formed due to transcriptional upregulation (Stoven et al., 2000).

Full length Relish consists of an N-terminal RHD and C-terminal PEST domain and ankyrin repeats. It has been shown by Stoven et. al (2003) that deletion of the PEST domain enhances nuclear localization of Relish and target gene expression. This is also seen on deletion of a serine rich stretch in Relish N-terminal. Hence, these to domains have been proposed to regulate the nuclear translocation of Relish and the mechanisms for this are still unknown (Stoven et al., 2003).

7.2.1 Negative regulators of the IMD pathway

The IMD pathway is regulated at the level of pathogen recognition activated by the enzyme-catalyzed degradation of PGN into smaller subunits by amidases. Members of the PGRP family, namely PGRP-LB, PGRP-SB1, PGRP-SB2, PGRP-SC1, and PGRP-SC2 have amidase activity and differ in their specificities for PGN. PGRP-SC1/2 has been shown to be

able to degrade both DAP-and Lys-type peptidoglycan, while PGRP-LB appears to specifically degrade DAP-type peptidoglycan. PGRP-LF is a trans-membrane protein which blocks IMD pathway by a mechanism other than degradation of PGN (Kleino and Silverman, 2014; Lemaitre and Hoffmann, 2007). PIMS and Rudra (PIRK), identified as negative regulators of IMD pathway are proposed to affect the pathway by interfering with the interaction between PGRP-LC with IMD or endocytosis of the PGRP-LC receptor (Kleino et al., 2008).

Ras/MAPK signaling cascade is also a negative regulator of the IMD pathway. Over-expression of components of the Ras/MAPK pathway inhibits IMD mediated upregulation of immune response genes. Ectopic activation of the PVR/MAPK pathway led to Pirk mediated down-regulation of the immune pathway (Ragab et al., 2011).

Two proteins: the Defense repressor 1 (Dnr1), and Caspar affect the IMD pathway by modulating the DREDD caspase activity. Dnr1 is a RING finger domain protein which has been proposed to down-regulate IMD pathway by suppressing DREDD caspase activity. It physically interacts with DREDD with the actual mechanism of DREDD activity inhibition is not yet understood (Guntermann et al., 2009). Caspar, a homolog of human Fas associated factor 1 (FAF1) was identified in a screen of *Drosophila* mutants for hyper-activated immune responses. Caspar inhibits the DREDD dependent cleavage of Relish leading to constitutive expression of *Diptericin* in Caspar mutant flies (Kim et al., 2006).

A number of proteins involved in the ubiquitin proteosome system have been shown to be important regulators of the IMD pathway. The activated IMD which is polyubiquinated by dIAP2 is de-ubiquitnated by ubiquitin-specific protease dUSP36 or Scrawny thus suppressing the IMD pathway (Thevenon et al., 2009). Cylindromatosis (CYLD), another de-ubiquitinating enzyme has been shown to interact with Kenny and negatively regulate the IMD pathway (Tsichritzis et al., 2007). The components of the Skp1/Cullin/F-box protein (SCF) complex -E3 ubiquitin ligases are important modulators of the IMD pathway. RNAi silencing of skpA or slimb was shown to increase the levels of both full-length and cleaved Relish suggesting that the SCF complex might regulate the stability of Relish and thereby modulate the IMD pathway activity (Khush et al., 2002).

The transcription factors of the JNK and JAK/STAT signaling pathways, AP-1 and STAT92E, have been implicated in curtailing the production of AMPs on activation of the

IMD pathway. It is suggested that in response to continuous immune signaling, levels of dAP1 and Stat92E also increase and these proteins would form a repressor complex with a *Drosophila* High mobility group (HMG) protein called Dorsal switch protein 1 (Dsp1). The complex would replace Relish at the promoter of effector genes and recruit a histone deacetylase to the complex to inhibit transcription of the target genes (Kim et al., 2007; Kim et al., 2005). Recently zinc finger homeodomain 1 (ZFH1) and Akirin, both nuclear proteins were also identified as negative regulator of the IMD pathway (Valanne et al., 2012) (**Figure 7-1**).

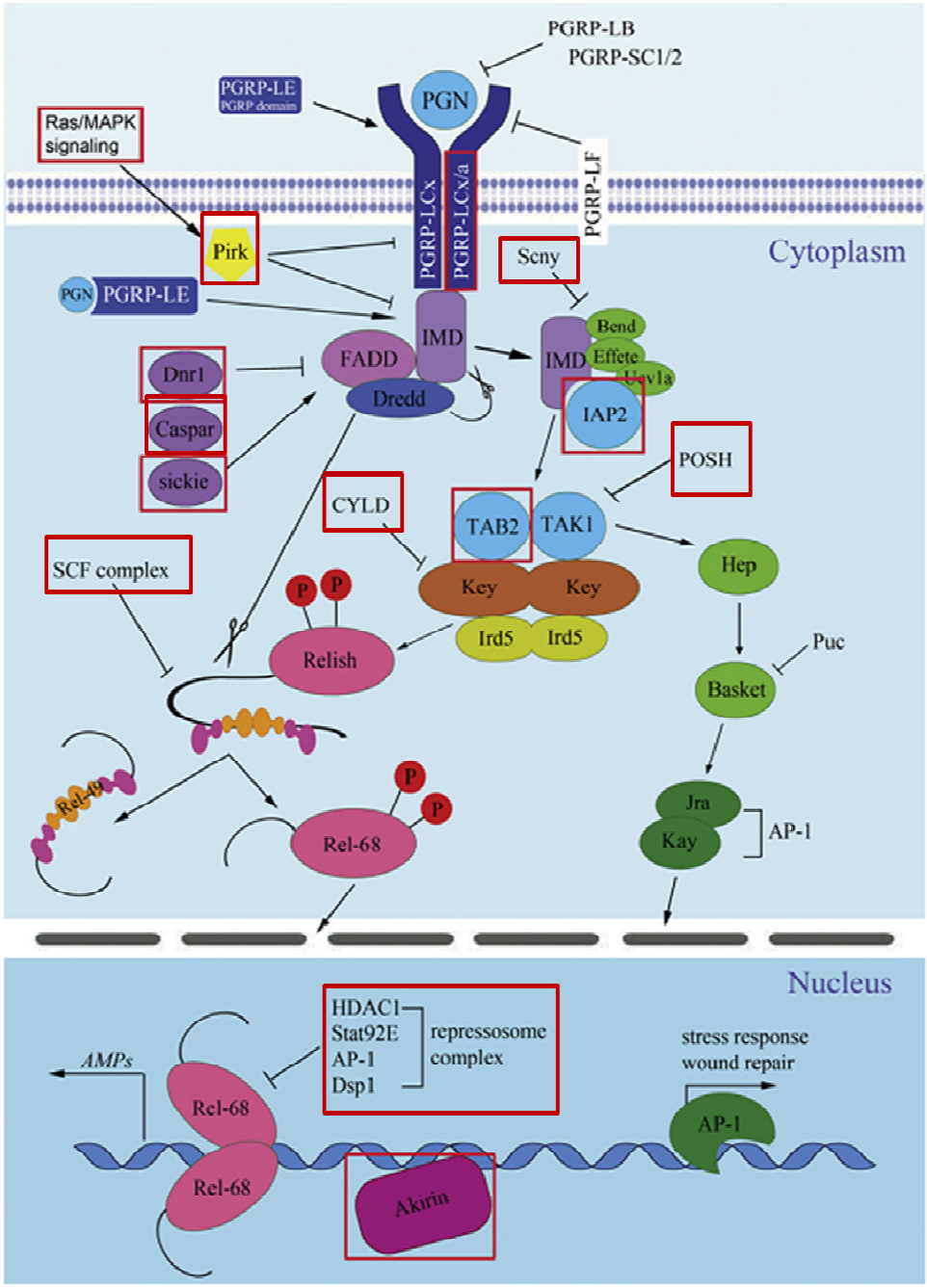


Figure 7-1: Negative regulators of IMD pathway signaling (adapted from S. Valanne *et al*., 2012)

The above picture depicts the canonical IMD pathway and its various negative regulators which include a few PGRPs, Ubiquitin pathway associated proteins, Transcription factors AP-1 and STAT92E among others marked with red boxes. (see text for details)

7.2.2 Caspar and its role in immunity

Drosophila Caspar is a homolog of mammalian Fas-associating factor 1 (FAF1). It is an evolutionarily conserved protein. It was identified in a genetic screen carried out by Kim *et al*. (2006) to identify the suppressors of the *Drosophila* immune pathway. In this study, they demonstrated that loss of Caspar leads to constitutive expression of the IMD pathway AMP, *Diptericin* even in the absence of an immune challenge. Thus, these flies showed improved resistance to bacterial infection. Caspar was identified as a negative regulator of the IMD pathway and it blocked the nuclear translocation of the NF-κB transcription factor probably by interfering with its DREDD-mediated cleavage (Kim et al., 2006).

FAF1 in mammals and other organisms shows the presence of two ubiquitin homologous domains: Ubiquitin associated (UAS) domain and Ubiquitin-like regulatory X (UBX) domain The C-terminal UBX domain is shown to interact with valosin containing protein (Cdc48/VCP) which is involved in various processes like protein degradation , chaperone activity, etc. A few organisms have an additional Ubiquitin associated domain (UBA) which helps it bind to ubiquitinated target proteins and regulate their proteolysis. FAF1 interacts with a lot of proteins and is implicated in a wide variety of cellular functions apart from apoptosis and cell death.

Like in *Drosophila*, FAF1 acts as a negative regulator of the NF-κB pathway induced by TNF-α, lipopolysaccharide or interleukin-1β. FAF1 overexpression was shown to inhibit translocation of RelA into the nucleus and also suppress the IKK activation to affect downstream activation of immune response genes. It achieves so by cytoplasmic retention of Rel65 and interrupting the IKK complex assembly by physically interacting with IKKβ (Παρκ ετ αλ., 2007).

7.2.3 Jun related antigen (Jra) and its role in immunity

Drosophila Jra is a homolog of mammalian cJun. It is an evolutionary conserved transcription factor. It is activated via phosphorylation by either JNK (Jun kinase) or in some cases MAPK (Mitogen activated protein kinase) depending on the signal of activation. In *Drosophila*, Jra forms a dimer with Kayak (homolog of mammalian cFos) to form an active AP-1 (activator protein 1) factor which is involved in a variety of processes within the cell like cell proliferation, different types of cell differentiation, cell migration, apoptosis and immunity.

The Jnk pathway is activated downstream of immune activation and is involved in the production of cytokines and cytoskeletal remodeling needed in phagocytosis. It has been shown that knockdown of *Drosophila*Jnk or AP1 factor (Jra) leads to overexpression of the *Attacin*. Various Relish dependent genes show AP1 binding sites upstream their promoter in close proximity to the Relish binding site. Further, a HMG (high mobility group) protein Dsp1 was found to be the core element which brings a repressor complex together with HDAC (histone deacetylase), AP1 and STAT92E called the repressosome. This repressosome complex is recruited to the promoter region where it causes the Relish target genes to contract via HDAC function thus down-regulating target gene expression (Kim et al., 2007; Kim et al., 2005). Similarly, interaction between cJun/ AP1 and STAT proteins in mammals has been shown to regulate NF-κB dependent immune responses.

7.3 Materials and methods

7.3.1 Cloning

Caspar and Jra were cloned into pGEX-4T1 vector for bacterial expression and into pRM vector for S2 cell expression using homologous recombination based cloning. *Caspar* and *Jra* were PCR amplified from cDNA obtained from S2 cell RNA. This fragment had homologous sites at the two ends which overlapped with PCR amplified vector. Both the vector and insert fragments were transformed into competent cells to obtain clones. Positive clones were confirmed using colony PCR and further sequenced. The sequencing result matches with Flybase cDNA sequence for both the constructs. Further, single lysine mutant constructs were made for Caspar and Jra using mutant site specific primers and homologous recombination. The mutations in Caspar included K551R and K436R. The mutations in Jra included K190R, K214R and K248R. All mutations were confirmed using sequencing.

Table 7-1: List of primers used for cloning and mutagenesis

Sr. No.	Primers	Sequence 5’-3’
1.	Caspar FP	CTGGTTCCGCGTGGATCCCCGGAATTCATGTCAGAGAACAAGGACGAGGCCTTG

2.	Caspar RP	TCGTCAGTCAGTCACGATGCGGCCGCTCATCGCTCCTCCAG GATGACCGT
3.	Jra FP	CTGGTTCCGCGTGGATCCCCGGAATTCATGAAAACCCCCGT TTCCGCTGCTG
4.	Jra RP	TCGTCAGTCAGTCACGATGCGGCCGCTTATTGGTCTGTCGA GTTCGGCGGCA
5.	pGEX FP	GCGGCCGCATCGTGACTGACTGACGA
6.	pGEX RP	GAATTCCGGGGATCCACGCGGAACCAG
7.	CasK436R FP	GTATAATGCTCTCATGTCTCATCAGTTGATCACAG
8.	CasK436R RP	CTGTGATCAACTGATGAGACATGAGAGCATTATAC
9.	CasK551R FP	TGCCCGTGATCAGGTGAGGGCAGAGCAGGACATGG
10.	CasK551R RP	CCATGTCCTGCTCTGCCCTCACCTGATCACGGGCA
11.	JraK190R FP	TTCTCGGTGATTAGGGACGAGCCCGTCA
12.	JraK190R RP	TTGACGGGCTCGTCCCTAATCACCGAGA
13.	JraK214R FP	CAGGAGAAGATCAGGCTGGAGCGCAAGA
14.	JraK214R RP	TCTTGCGCTCCAGCCTGATCTTCTCCTG
15.	JraK248R FP	GTGAAGGTACTTAGGGGCGAGAACGTCG
16.	JraK248R RP	CGACGTTCTCGCCCCTAAGTACCTTCAC
17.	pRM FP	GGATCCTCTAGAGTCGACCTGC
18.	pRM his RP	GTGATGGTGATGATGCATGGTACCGAGCTCGAATTCCCTT
19.	Cas his FP	TGCATCATCACCATCACCATGGAGGCGGAGGCGGAATGTC AGAGAACAAGGACGA
20.	Cas his RP	GCAGGTCGACTCTAGAGGATCCTCATCGCTCCTCCAGGATG ACCGT

7.3.2 Bacterial expression to check for SUMOylation and GST pulldowns

The proteins were expressed and purified by using *E. coli* BL21 (DE3) strain. For detecting SUMOylation of the proteins, the pGEX-4T1 constructs were co-transformed with

either Q^{SUMOGG} or $Q^{SUMO\Delta G}$ vector. The Q vectors are a kind gift from the Courey Lab. These vectors help express all the components of the SUMO machinery,SAE1, SAE2, Ubc9, and His-tagged $SUMO^{GG}$/ $SUMO^{\Delta G}$, sufficient to carry out SUMO modification of proteins in-bacto (Nie et al., 2009). A single, transformed, isolated colony of E. coli BL 21 (DE3) was inoculated in 1 ml LB medium and grown overnight at 37° C with vigorous shaking (200-250 rpm). 0.5 ml inoculum was added to 50ml LB broth and allowed to reach 0.8 O.D.600. Protein expression was induced by adding 1 mM isopropyl-D-thiogalactoside (IPTG) growing the cells at 37°C for 3 hours. The cells were lysed by sonication in lysis buffer (50 mM Tris (pH 7.5), 20 mM β-mercaptoethanol (BME), 500 mM NaCl, 0.1% Triton X-100) with protease inhibitor cocktail (Roche). The culture was then centrifuged at 15000 rpm for 30 min at 4° C and the supernatant containing soluble protein was used for further purification.

The supernantant was incubated with 200µl of glutathione beads overnight at 4°C on an end-to-end shaker. The beads were further washed with lysis buffer 5 times to remove contaminants and then boiled at 95°C for 15min with 1X SDS Laemmli buffer. The beads were then centrifuged at 15000rpm at 4°C for 15min and elute was collected.

7.3.3 Immunoblotting

The GST pulled down proteins from Q^{SUMOGG} and $Q^{SUMO\Delta G}$co-transformations were separated on a 10% SDS-PAGE gel followed by western blotting as described in section 3.2.6. The blots were probed with anti- His mouse antibody (Santacruz) in 1:1000 dilution for 2 hrs to detect SUMO and with anti- GST mouse antibody (Santacruz) in 1:5000 dilution for 1 hr to detect expressed Caspar or Jra proteins.

7.3.4 Determination of SUMO acceptor lysine residues using in-silico prediction softwares

The SUMOylation site prediction was carried out using SUMOsp ver2.0 (http://sumosp.biocuckoo.org), and the SUMO binding site prediction was carried out using GPS-SBM 1.0 (http://sbm.biocuckoo.org).

7.3.5 Multiple Sequence alignment

The Caspar and Jra protein sequences from different organisms - *Drosophila*, *H. sapiens*, *M. musculus, C. elegans* and *R. norvegicus* using NCBI resource. ClustalW2 tool with standard

parameters was used to carry out multiple sequence aligment of all the orthologs of Caspar and Jra in these organisms. (http://www.ebi.ac.uk/Tools/msa/clustalw2/)

7.3.6 S2 cell transfections and LPS induction

DNA preparations were made using Qiagen midipreparation kit to carry out transfections. Cells were split and grown to 50-60% confluency. Transfections were then carried out in 12 well plates using TransIT-2020 transfection reagent (Mirus Bio LLC) following manufacturer's protocol. 0.5µg of each construct was added per ml of cells. These cells were later induced with 500 mM $CuSO_4$ and kept for 48 hrs post which the cells were treated with 10µg LPS/ml cells for Relish cleavage experiment.

7.4 Results and discussion

7.4.1 Lysine 551 is the SUMO acceptor site in Caspar

Caspar protein expressed in the presence of $SUMO^{GG}$ showed both GST tagged caspar band and another slowly migrating band which is the $(His)_6$-SUMO modified Caspar (**Figure 7-2**, Lane 1). This SUMO modified band is absent when Caspar is expressed with $SUMO^{\Delta G}$ (**Figure 7-2**, Lane 2). The SUMO modified bands can be seen in the anti-His western of the GST pulled down Caspar with $SUMO^{GG}$ and is absent in $SUMO^{\Delta G}$.

SUMO prediction software (SUMOsp) showed two strong consensus sites for SUMOylation; K436 and K551 (**Figure 7-3 A**). Of these two residues, K551 is conserved among *Drosophila* and its homologs in humans, zebrafish, and mice (**Figure 7-3 B**). Both these lysines were mutated to arginine to disrupt modification by SUMO, if any. As compared to control, the K551R mutant showed loss of SUMO modified form of the protein whereas the other K436R mutant did not show loss of the SUMO modified Caspar bands as shown in **Figure 7-4**.

7.4.2 Caspar SUMO deficient mutant show alterations in Relish cleavage

In *Drosophila* hemocyte-like mbn-2 cell line on immune challenge with LPS there is rapid endo-proteolytic cleavage of full-length relish. The full-length band of Relish disappears almost completely within 30 seconds of LPS treatment whereas the Rel-49 and Rel-68 cleaved bands are clearly visible. The Relish full length protein reappears only after a

lag of 45 min. In this period the Relish gene is transcriptionally upregulated. The Rel-68 and Rell-49 products decrease overtime post 30 sec and increase after 45min when the Relish full length protein reappears (Stoven et al., 2000).

We tried to repeat this relish kinetic experiment using LPS-stimulated S2 cells using anti-Relish antibody (DSHB). However, the antibody did not work very well with immunobloting. We thus obtained a FLAG-Relish-RGSH6 construct from M. Ramet's lab which was originally made by Hultmark and colleagues (Stoven et al., 2003). Relish was N-terminally tagged with Relish which made it easier to visualize on immunoblots. Thus using anti-FLAG antibody we could see that Full length relish levels decrease on LPS treatment and reappear after a lag of 45 min in S2 cells.

Since Caspar is known to affect Relish cleavage, we tested the affect of overexpressing caspar wildtype along with the caspar SUMO-deficient mutant K551R. On comparing the cleavage of Flag-tagged Relish in the background of caspar wildtype and mutant over-expression there appears to be a difference in the cleavage kinetics (**Figure 7-5**). However these experiments need to be repeated along with other functional studies to conclude the role of caspar SUMOylation in Relish cleavage regulation.

7.4.3 Lysine 190 is one of the SUMO acceptor site in Jra

Jra protein expressed in the presence of SUMOGG showed both GST tagged Jra band and another three other slowly migrating band which is the (His)6-SUMO modified Jra. This SUMO modified band is absent when Jra is expressed with SUMOΔG (**Figure 7-6**). The SUMO modified bands can be seen in the anti-His western of the GST pulled down Jra with SUMOGG and is absent with SUMOΔG.

SUMO prediction software (SUMOsp) showed three strong consensus sites for SUMOylation; K29, K190 and K214 and K248 (**Figure 7-7**). Of these residues, K190, K214 and K248 are conserved among *Drosophila* and its homologs in humans, C. elegans, and mice (**Figure 7-8).** All three lysines were mutated to arginine to disrupt modification by SUMO, if any. As compared to control, the K190R mutant showed loss of 2 of the three SUMO modified form of the protein whereas the other mutants did not show loss of the SUMO modified Jra bands as shown in **Figure 7-9**. The lysine 190 identified as the SUMO acceptor site is homologous to the previously published lysine 229 in mammals. Further lysines need to be mutated to completely abolish SUMOylation of Jra.

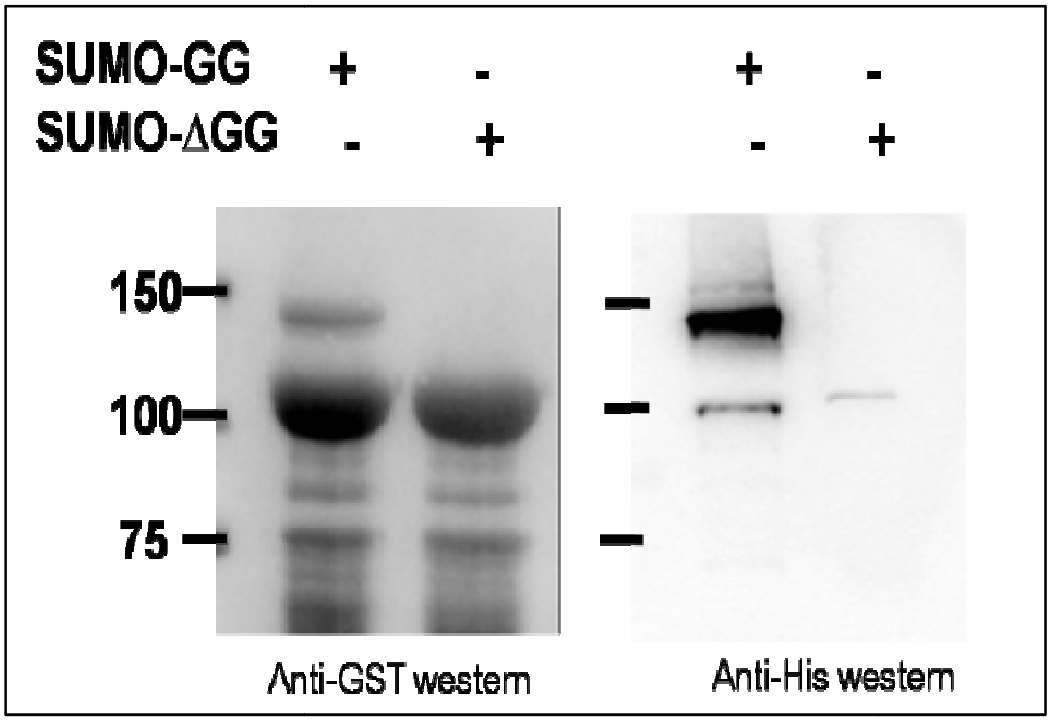


Figure 7-2: Caspar is SUMOylated in-bacto

Caspar is expressed with a GST tag in bacteria co-tranformed with E1, E2 enzymes and 6XHis-SUMO-GG (lane1) or 6XHis-SUMO-ΔGG (lane 2) in bacteria. The SUMOylated form of Caspar is seen as higher molecular weight band in the anti-GST western in lane 1 as compared to lane 2. This band also can be seen using anti-his antibody and confirms SUMOylation.

A.

Position	Peptide	Score	Type
5	SENKDEA	2.897	Non-consensus
436	QLMKHES	1.45	Ψ-K-X-E
484	RNIKLDK	2.691	Non-consensus
551	DQVKAEQ	3.261	Ψ-K-X-E
572	DAAKRQK	2.926	Non-consensus

B.

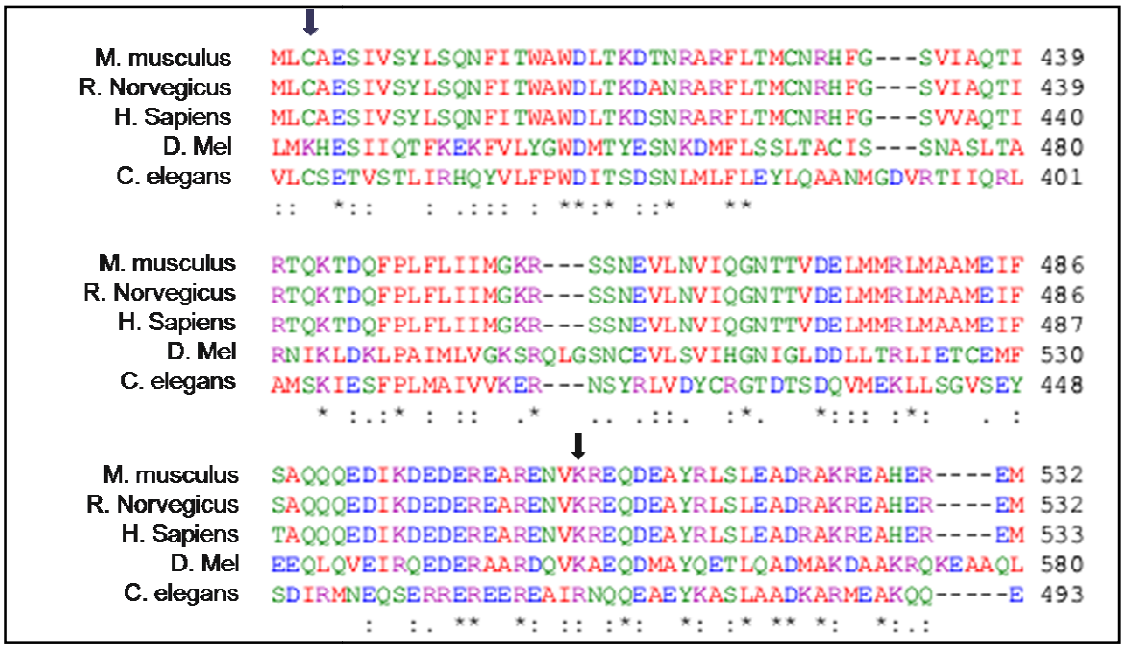


Figure 7-3: SUMO lysine acceptor site prediction for Caspar

A) The table lists out the predicted SUMO acceptor lysine sites using the SUMOsp in-silico prediction software. There are consensus sites at K436 and K551.

B) Amino acid sequence alignment of Caspar homologs from 5 different organism using clustalW2 to determine the conserved lysine residues of the predicted SUMO consensus acceptor lysine. Of the two consensus residues predicted by SUMOsp lysine 551 is conserved across all organisms.

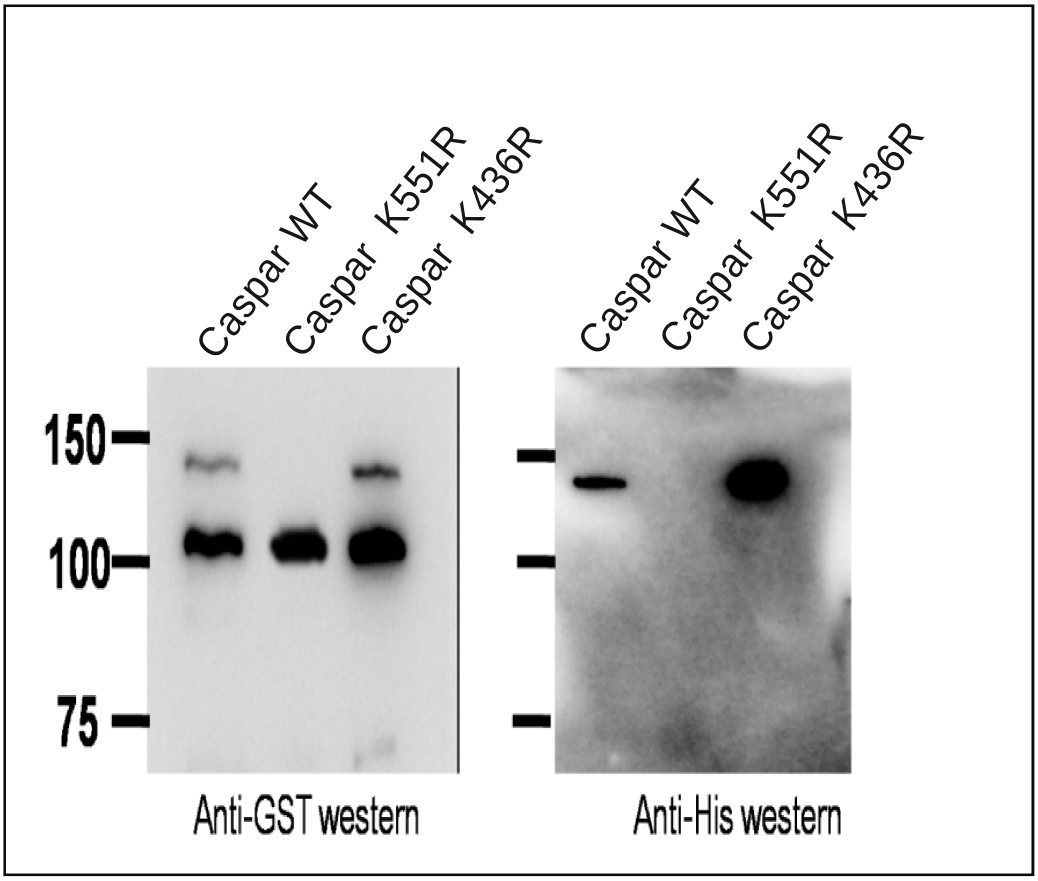


Figure 7-4: K551 is the SUMO acceptor site in Caspar

Caspar WT, Caspar K436R (mutant) and Caspar K551R (mutant) are co-transformed with E1, E2 and 6XHis-SUMO-GG in bacteria. On GST pulldown, the SUMOylated form of Caspar is lost in the K551R mutant hence proving that K551 is the SUMOylation site in Capsar.

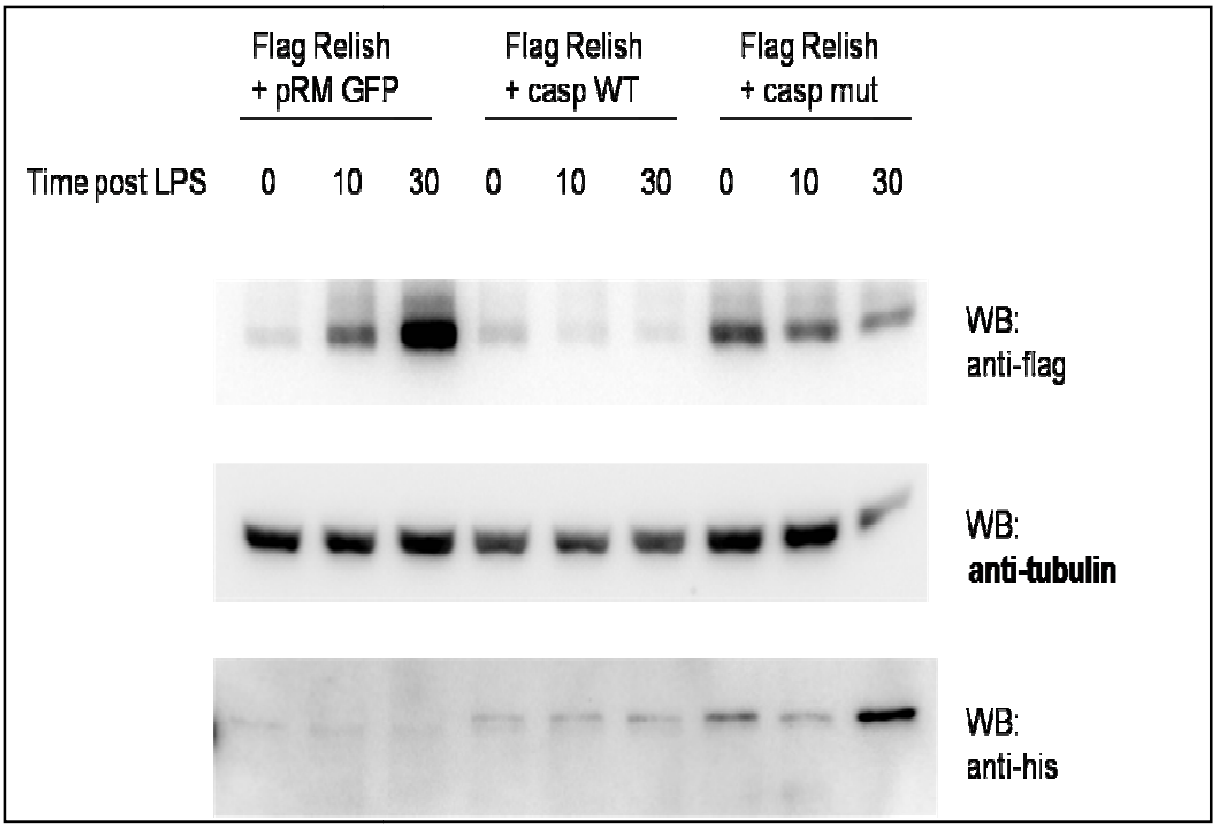


Figure 7-5: SUMOylation of caspar regulates Relish Cleavage post LPS treatment
S2 cells were transfected with N-terminally tagged Relish along with either vector control pRM GFP, Caspar wildtype construct and Caspar K551R mutant construct. Without any caspar overexpression, Relish full length increases from 10 min to 30 min post LPS treatment. However, there is a clear difference in the Relish cleavage kinetics between the wild type and control. The tubulin western shows that there was equal concentration of protein loaded in each well. The his western shows over-expression of wild type and mutant caspar.

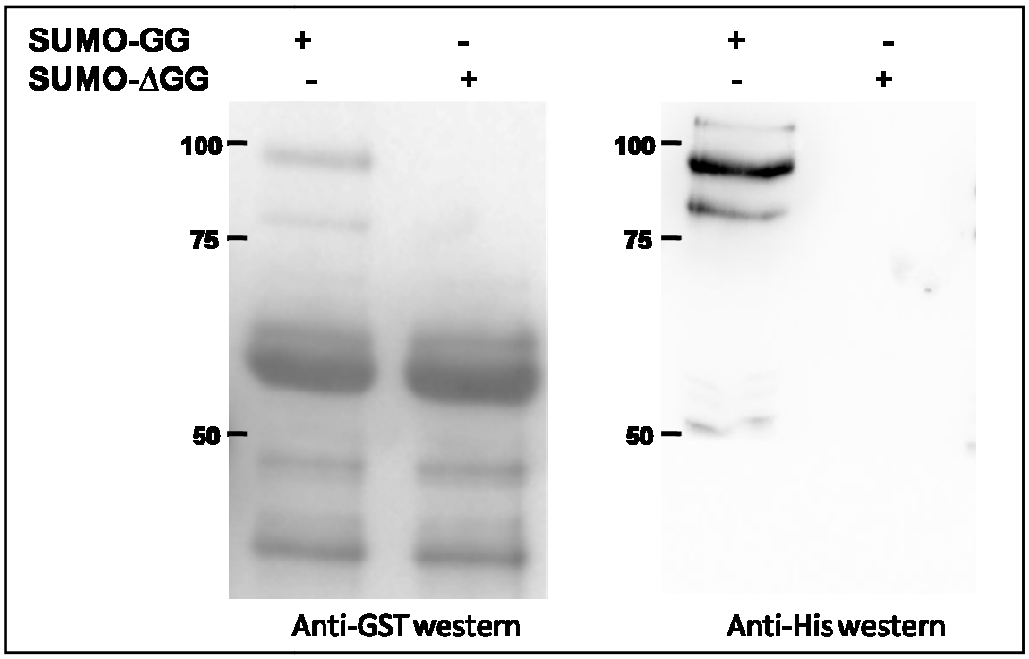


Figure 7-6: In-vitro SUMOylation of Jra

Jra is expressed with a GST tag in bacteria co-tranformed with E1, E2 enzymes and 6XHis-SUMO-GG (lane1) or 6XHis-SUMO-ΔGG (lane 2) in bacteria. The SUMOylated forms of Jra are seen as higher molecular weight bands in the anti-GST western in lane 1 as compared to lane 2. These bands can also be seen using anti-His antibody and confirms SUMOylation.

Position	Peptide	Score	Type
29	SENKDEA	2.897	Ψ-K-X-E
190	QLMKHES	1.45	Ψ-K-X-E
214	RNIKLDK	2.691	Ψ-K-X-E
248	DQVKAEQ	3.261	Ψ-K-X-E
269	DAAKRQK	2.926	Non-consensus

Figure 7-7: SUMO lysine acceptor site prediction for Jra

The table lists out the predicted SUMO acceptor lysine sites using the SUMOsp in-silico prediction software. There are consensus sites at K29, K190, K214 and K248.

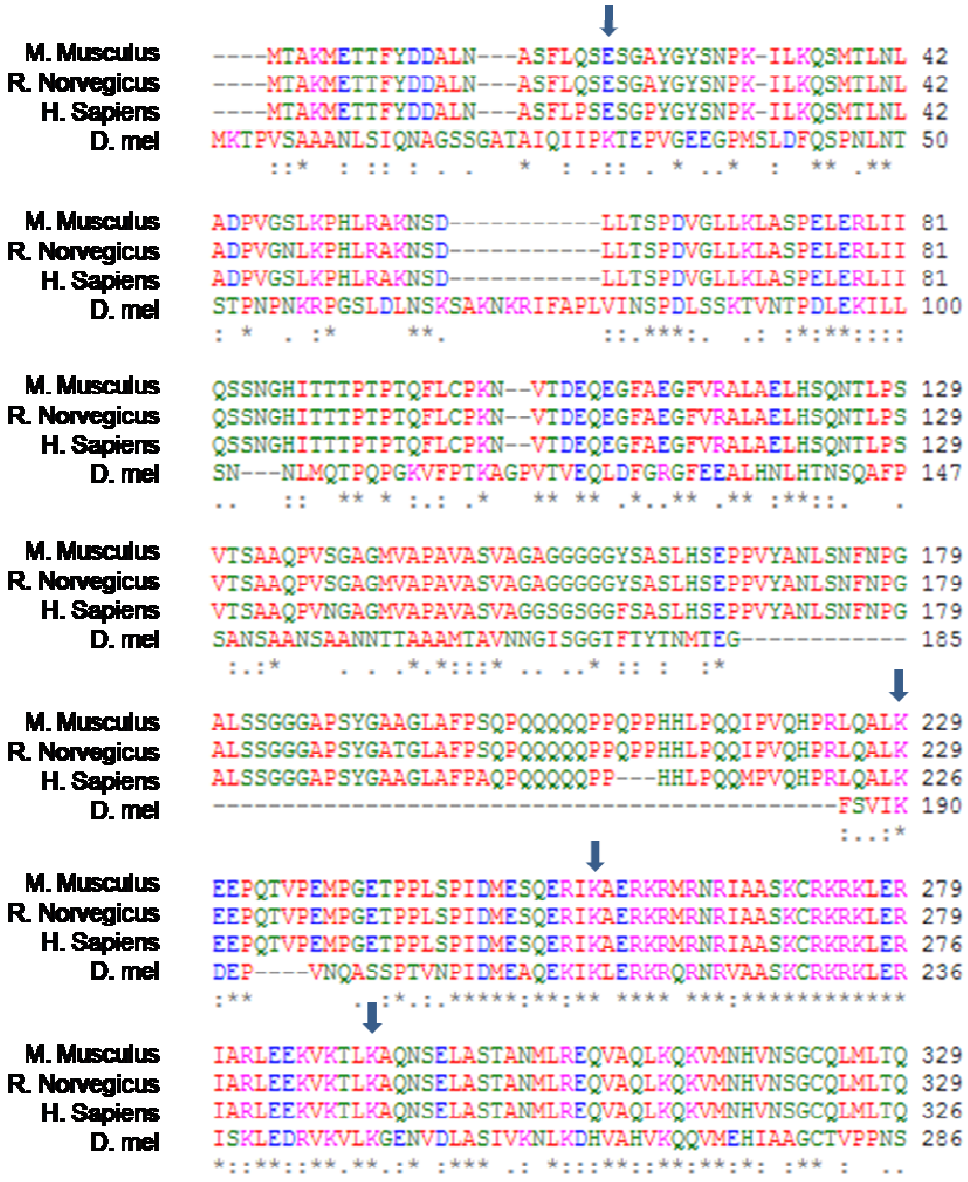


Figure 7-8: Protein sequence alignment for Jra

This figure shows the amino acid sequence alignment of Jra homologs from different organisms using clustalW2. Only a few SUMO acceptor residues are conserved in other orthologs. Of the four consensus residues predicted by SUMOsp lysines190, 214 and 248 are conserved across all organisms.

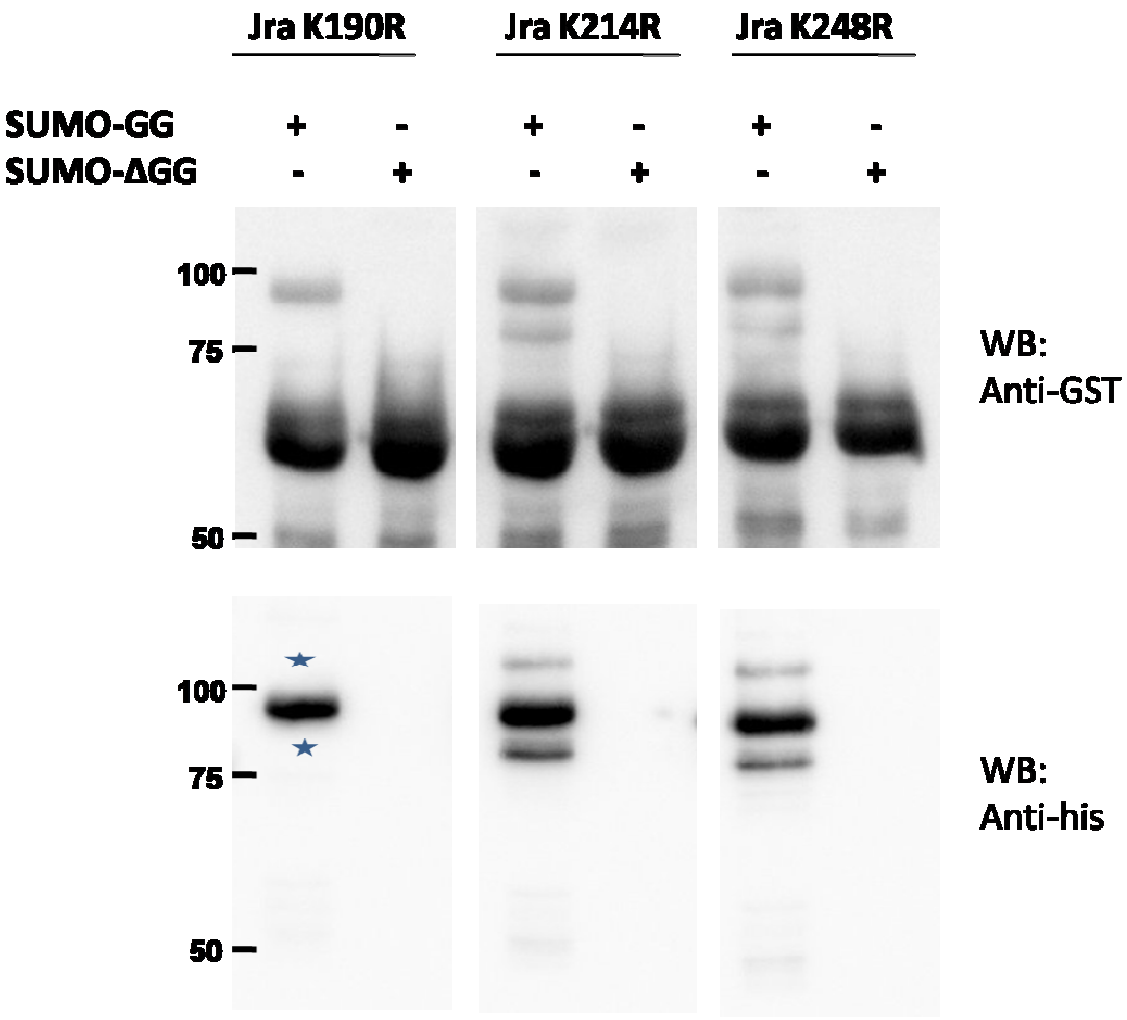


Figure 7-9: K190 is the primary SUMO acceptor site in Jra
Jra mutants K190R, K214R and K248R are co-transformed with E1, E2 and 6XHis-SUMO-GG in bacteria. On GST pulldown, two bands pertaining to the SUMOylated form of Jra are lost in the K190R mutant, which are otherwise retained in the other mutant. Lysine 190 seems to one of the primary sites of SUMOylation for Jra.

Chapter 8

Discussion & future perspectives

The recognition of the pathogen by the host through various pathogen receptors triggers a cascade of inter-dependent events that leads to phagocytosis and production of anti-microbial peptides and melanization. The phagocytic response includes various processes like internalization of the pathogen through vesicle-mediated endocytosis, actin cytoskeleton reorganization to aid phagocytic movement of the hemocytes. The activation of the two major pathways, the Toll/NF-κB and the IMD/NF-κB pathways results in translocation of the three NF-κB molecules, Dorsal (DL), Dorsal like immune factor (DIF) and Relish (REL) into the nucleus thus up-regulating the transcription of various defense and repair genes. In addition, these two major pathways are interconnected at various levels and also subject to regulation by other signal transduction pathways as also positive and negative feedback. Studies have demonstrated that Ras/MAPK, JNK and JAK-STAT pathways are responsible for negative regulation of the NF-κB pathways or work independently to activate certain effector responses like apoptotic response, stress response and increased hemocyte proliferation. As demonstrated in this study, as well as earlier studies, the decrease in *SUMO* and subsequent decrease in SUMOylation levels, modulates the levels of defense genes. There is also increased cell proliferation in the haemopoietic lineage in absence of SUMO. The net effect on the immune response is in all probability an integrated effect of changes of SUMOylation levels of individual molecules involved in the different aspects of immune/healing regulation. At this point, we are far away from a mechanistic understanding of the effect of SUMOylation on immunity.

Our approach of performing the pulldowns in native conditions helps us obtain not only the SUMOylated proteins involved in immune response but also the complexes enriched in the process. Study of these complexes and its role in immunity would help build a comprehensive mechanistic function of SUMOylation in the immune response. Comparison of our data set with published data of humoral and cellular immune response regulators shows a considerable overlap. This, along with the cytoscape analysis which provides an overlap of our list with the DPiM network gives us more belief that SUMO is indeed involved in regulating specific processes in immune responses. A number of proteins as

discussed in earlier chapters serve as interesting targets to study SUMO mediated immune regulation at the level of translation, vesicle mediated transport, chromatin remodeling complexes among the other obvious targets directly involved in NF-κB signaling.

We further show that IMD regulators are modified by SUMO. These proteins namely Caspar and Jra regulate the IMD pathway at different levels. Caspar acts as a block for Relish cleavage and downstream signaling, whereas Jra is activated in response to IMD signaling and acts as a pause signal for IMD pathway. SUMO modification of these proteins may affect both their functioning in a synergistic manner or contradictory to each other. The SUMO acceptor lysine residues have been identified for both these proteins however, in depth characterization of the mutants is required to understand their interactions with other proteins in the pathway and subsequent affect on their immune function.

Figure 8-1 indicates, in a pictorial fashion the proteins known to be SUMOylated in the immune signaling pathways. Each protein that can be cycled through the SUMOylation/deSUMOylation cycle becomes a potential regulatory step for the control of innate immune signaling. A SUMOylation event may work towards increasing or decreasing the signal, modulating the net immune response to a pathogen. If we can generate quantitative data for each and every SUMO substrate in the pathways depicted, we would be able to build an intergrated, quantitative model for predicting the effect of SUMOylation of individual (SUMOylated) parts on the whole system. The model could be used to predict changes in the immune response on perturbation of multiple nodes (e.g., the SUMOylated sites) and used to further improve the model. A refined model, after extensive testing via cycles of prediction and experiments would be close to the actual regulatory model for SUMOylationin the innate immune response.

In summary, our study provides a list of possible SUMOylated proteins in *Drosophila*. These proteins represent ~5% of the genes in *Drosophila*. The hits do not overlap significantly with the immune transcriptome, confirming independent roles for dynamic, post-translational modifications in the early stages of the immune response. This list is a first step in understanding roles for SUMOylation of individual proteins and their effect on the innate immune response. SUMOylation thus appears to be widespread in the *Drosophila* proteome, with specific roles in immunity.

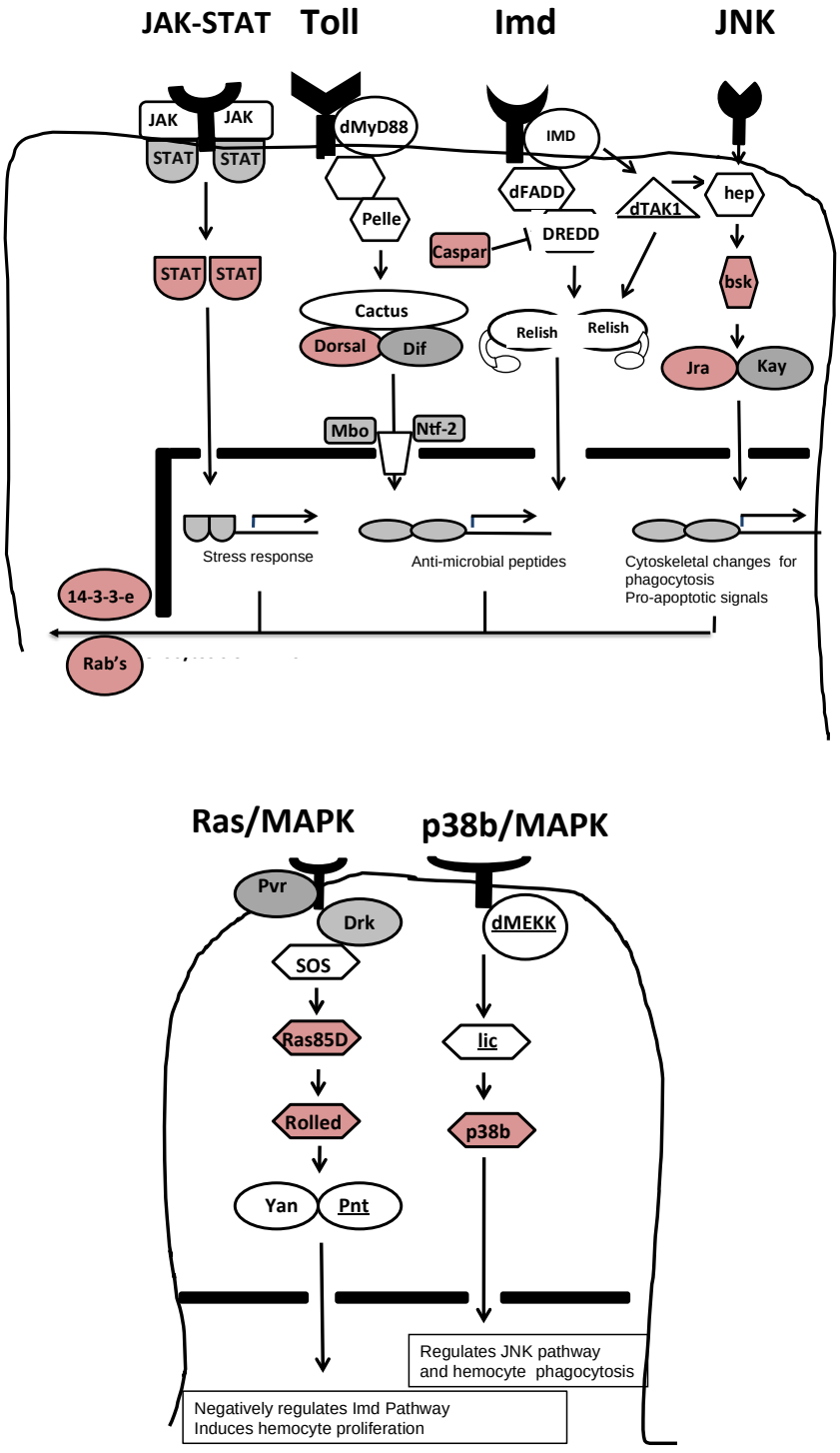


Figure 8-1: SUMO and Signaling pathways in *Drosophila* innate immunity
SUMOylated proteins can be critical regulation points in signaling cascades/ networks. Many important signal transduction pathways in immunity appear to have at least one control point for SUMO mediated regulation. Proteins colored brick red have been demonstrated to be SUMOylated in this and earlier studies, while those in grey are listed in our list, but are yet to be validated.

APPENDIX I

Affinity purification of SUMOylated conjugates from *Drosophila* larvae in denaturing conditions

In this chapter, we describe an attempt to identify SUMOylated proteins from *Drosophila* third instar larvae in response to infection with mixture of gram positive and gram negative bacteria. We used the UAS-Gal4 system commonly used in flies to overexpress His_6-FLAG N-terminally tagged SUMO in the entire larva. We performed tandem affinity purification which showed non-specific bands in silver stained gels in the control. Standardizations are still in process to minimize contaminants to correctly identify SUMO substrates in pathogen stimulated larvae.

Validation of the gal4 and the UAS-SUMO-His_6-FLAG *Drosophila* lines

The UAS-gal4 system in flies is used for targeted gene expression in a spatial and temporal fashion. This system is originally identified in *S. cerevisiae* and involves the transcription factor Gal4 lead to activation of its target genes by binding to four related 17 basepair (bp) sites known as the upstream activating sequence (UAS). Brand and Perrimon used this sytem in flies by making two sets of fly lines. One with the gene of interested under the control of the UAS element and the other with the GaL4 under a tissue specific promoter. To activate the expression of the target genes the UAS and Gal4 flies are mated and their progeny then expresses the target gene in a pattern governed by the respective gal4 driver.

We used this system to drive the expression of $SUMO^{GG}$ (mature SUMO) with two tags i.e. His_6 and FLAG (to aid pull-down of the SUMOylated proteins) within the organism. The UAS SUMO-His_6-FLAG lines were obtained from M. Nie (UCLA, USA). To check the expression of the lines they were crossed to vestigial (Vg) gal4, a wing specific driver. The expression of SUMO-His_6-FLAG in the third larval instar wing disc showing a Vg expression pattern was confirmed using antibody staining with anti-FLAG antibody (**Figure A1-1A**).

To maximize the amount of proteins SUMOylated within the cell for identification we tested the ability of different Gal4 lines to drive expression of the tagged SUMO in the 3rd instar larva in the fly. To help visualize expression easily we crossed: Kruppel-Gal4 (Kr

Gal4, fat body specific driver), Collagen-Gal4 (Cg Gal4, fat body and haemocyte specific) and Daughterless-Gal4 (Da Gal4, whole larval driver) to UAS flies. We discovered that of all these Gal4's tested Da gal4 driver had the strongest expression in almost all the tissues in the 3rd instar larva (**Figure A1-1B**). We used this driver henceforth in our pulldown experiments to identify pathogen induced global changes in SUMOylated substrates.

On validation of the lines, we made a stable line expressing both Da-Gal4 and UAS SUMO-His_6-FLAG in the same fly. The F1 generation 3rd instar larvae were collected from these flies for inducing an immune response by septic injury. The larvae were pricked with a needle dipped in an overnight grown culture of *Mycobacterium spegmatis* (gram positive) and *Salmonella typhii* (gram negative) bacteria previously shown to induce immune responses in fly. These pricked and unpricked control larvae were used for tandem affinity purifications.

Tandem affnitiy purifications to purify SUMOylated substrates

Tandem affinity purification involves the purifications of proteins using affinity tags. We used this system to minimize contaminants in the pulldowns due to purification using two tags in sequence. Initially, the His pulldown was carried out in denaturing conditions. Denaturing conditions would help inhibit action of SUMO deconjugation proteases and other proteases which would help retain SUMO modified proteins. It would also break complexes and help identify only proteins directly conjugated to SUMO. The Elutes from the His pulldown were further used for immune-pulldowns using anti-FLAG agarose. The resulting elute was to be further analyzed using quantitative mass spectrometry.

F1 3rdinstar Da Gal4-UAS - SUMO-His_6-FLAG larvae were initially subjected to singe step anti-FLAG affinity purifications for standardization. The larvae were kept at 37°C for 1 hr to induce a heat shock response as it has been previously shown that heat shock leads to an increase in global SUMOylation in *Drosophila*. Larvae were collected and crushed in liquid nitrogen. Further the lysate was made in RIPA buffer supplemented with protease inhibitors and NEM. After sonication of the sample using 30sec on/off pulse for 15 minutes, the lysates was centrifuged at high temperature. The supernatant was collected and incubated with mouse anti-FLAG sepharose beads (SIGMA) overnight at 4°C. The beads were further washed with RIPA buffer and then boiled in 1X Laemmli buffer. The immune-pulldown samples were subjected to SDS-PAGE electrophoresis followed by immunobloting with

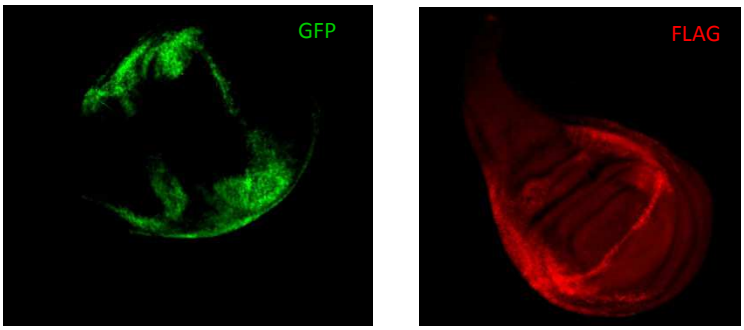
rabbit anti-FLAG antibody. The western showed negligible pulldown with the W1118 larvae and served as a negative control. The heat shock larvae showed increase in SUMOylated substrates as compared to the untreated SUMO-His_6-FLAG expressing larvae (**Figure A1-2**).

Post standardization of anti-FLAG pulldowns, the larvae were used for a two step tandem affinity purification (TAP) using His and FLAG tags. W1118 larvae were used as a negative control as they do not express any tagged construct. The two step His-FLAG purification was chosen with an aim to minimize non-specific pull-downs. In order to prevent any loss of SUMOylated species a cocktail of protease inhibitors was added along with NEM (SUMO protease inhibitor). The larvae were crushed in liquid nitrogen and used for His pulldown in Urea binding buffer (pH 8.0) with 15 mM imidazole. The samples were sonicated using 30sec on/off cycle for 10 min. After centrifugation the lysate was filtered through a mira cloth to remove fat and the collected supernatant was incubated with prewashed Ni-NTA beads (Qiagen) overnight at 4°C. Post-incubation the beads were given sequential washes with urea binding buffer pH 8.0, Urea wash buffer pH 8.0 with 20 mM imidazole and 0.5% triton X, urea wash buffer pH 6.3 with 20 mM imidazole and then eluted with 350 mM imidazole. The elute was further dialysed in 1X TBS (Tris Buffer Saline) with 0.1 mM PMSF (phenylmethylsulfonyl fluoride) for remove excess urea and imidazole which would interfere with anti-FLAG immunoprecipitation. The dialyzed fraction was diluted 5 times in 1X TBS with 1% 1% Triton X-100, 5% glycerol, protease inhibitors and NEM and kept for incubation with prewashed anti-FLAG affinity beads overnight at 4°C. The beads were later washed with 1X TBS supplemented with 1% Triton X-100 thoroughly and the elution is further carried out using FLAG peptide (Sigma). The purified proteins were further analyzed by western blotting and sliver staining.

The anti-FLAG western of the TAP purified samples show a global increase in SUMOylated substrates post bacterial infection as compared to uninfected control. However the silver staining of these samples show protein bands being pulled down in the W1118 negative control. The intensity of these non-specific bands was almost equivalent to that in the Da Gal-SUMO-His_6-FLAG untreated lane. Since the W1118 larvae do not express any tagged SUMO expression, the proteins seen are contaminants from the pull downs (**Figure A1-3**). This leads to the conclusion that under our specific set of experimental conditions there are non-specific contaminants being pulled down which may yield false positive results on further mass spectrometry analysis.

Various modifications of the protocol were tested to minimize these non-specific contaminating proteins - varying imidazole concentrations during washing and elution of the Ni-NTA beads, different dialysis conditions and rigorous washings during FLAG pulldowns. A challenge in using larvae is the large amount of fat deposited in the animal. We believe that this might interfere with the pulldowns increasing the non-specificity. Thus, we tried to precipitate the protein from larval lysate to separate it from the fat and later use this for TAP. We are still in the process of standardizing these conditions to obtain specific pulldowns.

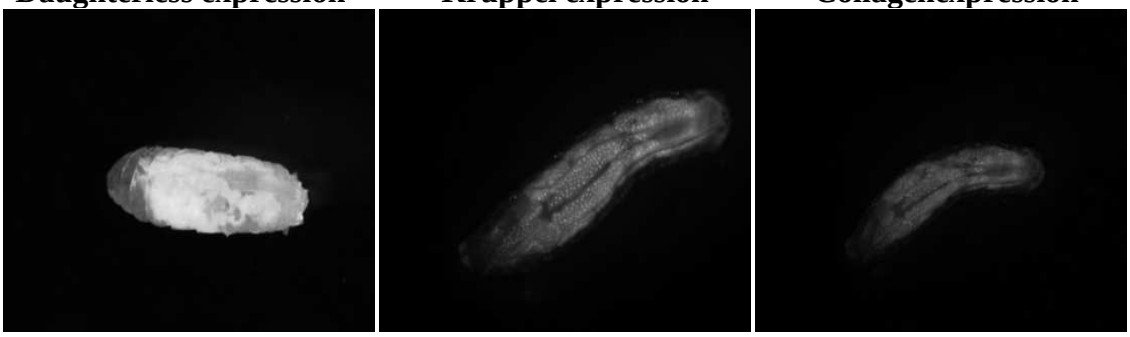
A. **Control** **Experiment**



Vg Gal4×UAS-GFP **Vg Gal4×UAS-SUMO-His$_6$-FLAG**

B.

Daughterless expression **Kruppel expression** **Collagenexpression**



Da Gal4×UAS-GFP **Kr Gal4×UAS-GFP** **Cg Gal4×UAS-GFP**

Figure A1-1: Validation of expression of the UAS and Gal4 *Drosophila* lines
(A) Vestigial expression pattern in3rd instar larva wing disc is seen in the first panel with GFP. The Anti-FLAG immuno-staining pattern in the experiment is similar to the control, confirming the expression of the UAS-SUMO-His$_6$-FLAG.
(B) The crosses with the different gal4 drivers daughterless (Da), Kruppel (Kr) and collagen (Cg) show differential GFP expression pattern within the 3rd instar larva. Da Gal4 driver showed the strongest expression in these larvae.

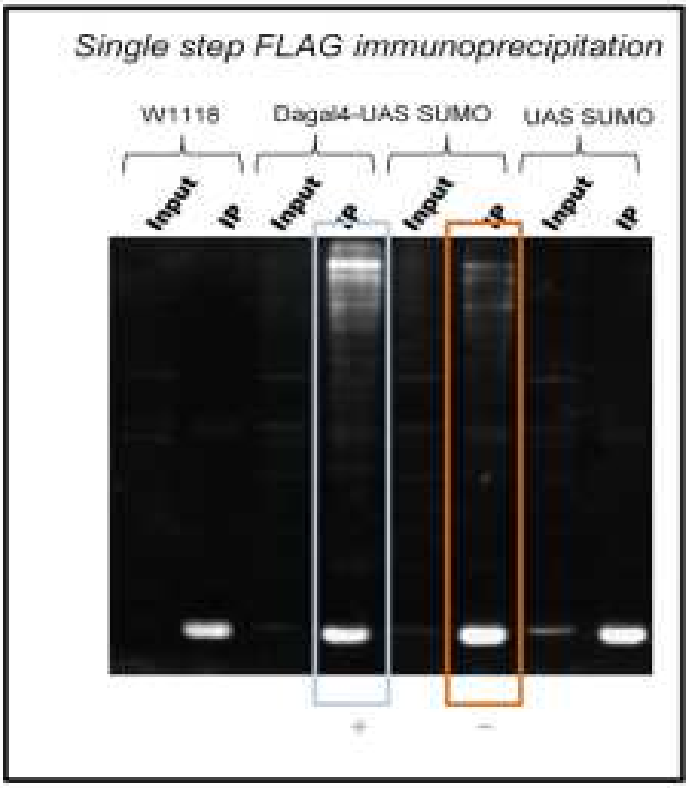


Figure A1-2: Standardization of single step FLAG pulldowns

3rd instar Da Gal4 - UAS-SUMO-His_6-FLAG larvae were used for purifications. Single step FLAG pull-down of SUMOylated proteins from two different sets of larvae with and without septic injury. The immune challenge lane shows more number of SUMOylated proteins in comparison to the non heat shock lane.

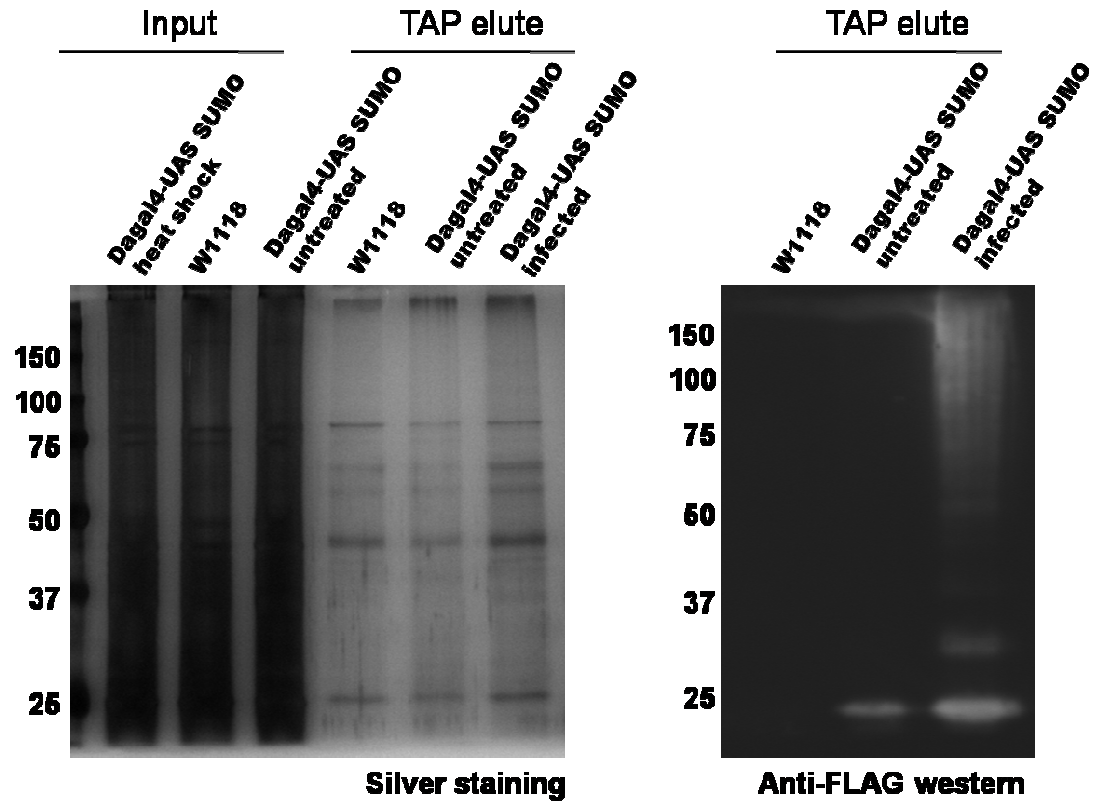


Figure A1-2: Tandem affinity purification

Two- step pulldown of SUMOylated proteins using Ni-NTA purification followed by anti-FLAG immunoprecipitation was performed using 3rd instar larvae expressing SUMO-His$_6$-FLAG. The western shows the immune challenged lane with more SUMOylated protein bands as compared to the non treated control. The western does not show any non-specific bands in the W1118 negative control. However, silver staining shows protein being pulled down even in W1118 control at almost same intensity as in untreated cells. The immune challenged lane shows a slight increase in the intensity of bands.

APPENDIX II

A large scale in-bacto screen to identify SUMOylated proteins

We compared our list of proteins with the published mammalian SUMO proteome. We tested a few of the proteins from that list and other interesting candidate proteins to depict SUMOylation of the proteins using the in-bacto SUMOylation system. This elaborate screening was done in collaboration with other members of the lab. The proteins tested are listed in the table below in **Table A2-1**

Table A2-1: List of proteins validated to show SUMoylation using in-bacto SUMOylation system

Sr. No.	Name	Function
1.	Prosβ4	Proteasome mediated ubiquitin degradation
2	pUf68	mRNA splicing
3	Rept	DNA helicase
4	Rpa70	DNA replication
5	Rpl31	Ribosomal protein
6	SmD3	mRNA splicing via spliceosome
7	Pros29	Proteasome mediated ubiquitin degradation
8	Rpn11	Proteasome mediated ubiquitin degradation
9	CG6724	unclear
10	CG6888	DNA damage response
11	Hsp60d	Heat shock factor
12	RpS10b	Ribosomal protein
13	Lwr	SUMO conjugating enzyme
14	CG3708	Nucleosome assemly
15	CG6523	Cell redox homeostasis

16	CG8636	Cell cycle regulation
17	CG3412	Varied functions in early development, TOR pathway etc.
18	CG5330	Nucleosome assemly
19	CG8415	Centrosome organization
20	tango7	Golgi organization
21	14-3-3 zeta	Regulator of Ras pathway
22	SNAP	SNARE binding protein
23	TRAM	phagocytosis
24	β-cop	Vesicle mediated transport
25	dos	Regulator of the Ras pathway
26	STAT92E	Transcription factor of JAK-STAT pathway involved in diverse functions
27	colt	Mitochondrial function
28	Ntf-2	Nuclear transport protein
29	TRAP1	Oxidative stress response
30	psidin	phagocytosis

REFERENCES

Agaisse, H., Burrack, L.S., Philips, J.A., Rubin, E.J., Perrimon, N., and Higgins, D.E. (2005). Genome-wide RNAi screen for host factors required for intracellular bacterial infection. Science *309*, 1248-1251.

Agaisse, H., and Perrimon, N. (2004). The roles of JAK/STAT signaling in Drosophila immune responses. Immunological reviews *198*, 72-82.

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen recognition and innate immunity. Cell *124*, 783-801.

Anckar, J., and Sistonen, L. (2007). SUMO: getting it on. Biochemical Society transactions *35*, 1409-1413.

Arnot, C.J., Gay, N.J., and Gangloff, M. (2010). Molecular mechanism that induces activation of Spatzle, the ligand for the Drosophila Toll receptor. The Journal of biological chemistry *285*, 19502-19509.

Asling, B., Dushay, M.S., and Hultmark, D. (1995). Identification of early genes in the Drosophila immune response by PCR-based differential display: the Attacin A gene and the evolution of attacin-like proteins. Insect biochemistry and molecular biology *25*, 511-518.

Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007). Quantitative mass spectrometry in proteomics: a critical review. Analytical and bioanalytical chemistry *389*, 1017-1031.

Bayer, P., Arndt, A., Metzger, S., Mahajan, R., Melchior, F., Jaenicke, R., and Becker, J. (1998). Structure determination of the small ubiquitin-related modifier SUMO-1. Journal of molecular biology *280*, 275-286.

Bhaskar, V., Smith, M., and Courey, A.J. (2002). Conjugation of Smt3 to dorsal may potentiate the Drosophila immune response. Molecular and cellular biology *22*, 492-504.

Bhaskar, V., Valentine, S.A., and Courey, A.J. (2000). A functional interaction between dorsal and components of the Smt3 conjugation machinery. The Journal of biological chemistry *275*, 4033-4040.

Bidla, G., Dushay, M.S., and Theopold, U. (2007). Crystal cell rupture after injury in Drosophila requires the JNK pathway, small GTPases and the TNF homolog Eiger. Journal of cell science *120*, 1209-1215.

Bidla, G., Lindgren, M., Theopold, U., and Dushay, M.S. (2005). Hemolymph coagulation and phenoloxidase in Drosophila larvae. Developmental and comparative immunology *29*, 669-679.

Boehm, A.M., Putz, S., Altenhofer, D., Sickmann, A., and Falk, M. (2007). Precise protein quantification based on peptide quantification using iTRAQ. BMC bioinformatics *8*, 214.

Boman, H.G., Nilsson, I., and Rasmuson, B. (1972). Inducible antibacterial defence system in Drosophila. Nature *237*, 232-235.

Bossis, G., Malnou, C.E., Farras, R., Andermarcher, E., Hipskind, R., Rodriguez, M., Schmidt, D., Muller, S., Jariel-Encontre, I., and Piechaczyk, M. (2005). Down-regulation of c-Fos/c-Jun AP-1 dimer activity by sumoylation. Molecular and cellular biology *25*, 6964-6979.

Boutros, M., Agaisse, H., and Perrimon, N. (2002). Sequential activation of signaling pathways during innate immune responses in Drosophila. Developmental cell *3*, 711-722.

Brennan, C.A., and Anderson, K.V. (2004). Drosophila: the genetics of innate immune recognition and response. Annual review of immunology *22*, 457-483.

Buchon, N., Poidevin, M., Kwon, H.M., Guillou, A., Sottas, V., Lee, B.L., and Lemaitre, B. (2009). A single modular serine protease integrates signals from pattern-recognition receptors upstream of the Drosophila Toll pathway. Proceedings of the National Academy of Sciences of the United States of America *106*, 12442-12447.

Bulet, P., Dimarcq, J.L., Hetru, C., Lagueux, M., Charlet, M., Hegy, G., Van Dorsselaer, A., and Hoffmann, J.A. (1993). A novel inducible antibacterial peptide of Drosophila carries an O-glycosylated substitution. The Journal of biological chemistry *268*, 14893-14897.

Busse, M.S., Arnold, C.P., Towb, P., Katrivesis, J., and Wasserman, S.A. (2007). A kappaB sequence code for pathway-specific innate immune responses. The EMBO journal *26*, 3826-3835.

Cajee, U.F., Hull, R., and Ntwasa, M. (2012). Modification by ubiquitin-like proteins: significance in apoptosis and autophagy pathways. International journal of molecular sciences *13*, 11804-11831.

Cheng, L.W., Viala, J.P., Stuurman, N., Wiedemann, U., Vale, R.D., and Portnoy, D.A. (2005). Use of RNA interference in Drosophila S2 cells to identify host pathways controlling compartmentalization of an intracellular pathogen. Proceedings of the National Academy of Sciences of the United States of America *102*, 13646-13651.

Chiu, H., Ring, B.C., Sorrentino, R.P., Kalamarz, M., Garza, D., and Govind, S. (2005). dUbc9 negatively regulates the Toll-NF-kappa B pathways in larval hematopoiesis and drosomycin activation in Drosophila. Developmental biology *288*, 60-72.

Clemens, J.C., Worby, C.A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B.A., and Dixon, J.E. (2000). Use of double-stranded RNA interference in Drosophila cell lines to dissect signal transduction pathways. Proceedings of the National Academy of Sciences of the United States of America *97*, 6499-6503.

De Gregorio, E., Han, S.J., Lee, W.J., Baek, M.J., Osaki, T., Kawabata, S., Lee, B.L., Iwanaga, S., Lemaitre, B., and Brey, P.T. (2002). An immune-responsive Serpin regulates the melanization cascade in Drosophila. Developmental cell *3*, 581-592.

De Gregorio, E., Spellman, P.T., Rubin, G.M., and Lemaitre, B. (2001). Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays. Proceedings of the National Academy of Sciences *98*, 12590-12595.

Delaney, J.R., Stoven, S., Uvell, H., Anderson, K.V., Engstrom, Y., and Mlodzik, M. (2006). Cooperative control of Drosophila immune responses by the JNK and NF-kappaB signaling pathways. The EMBO journal *25*, 3068-3077.

Denison, C., Rudner, A.D., Gerber, S.A., Bakalarski, C.E., Moazed, D., and Gygi, S.P. (2005). A proteomic strategy for gaining insights into protein sumoylation in yeast. Molecular & cellular proteomics : MCP *4*, 246-254.

Desterro, J.M., Rodriguez, M.S., and Hay, R.T. (1998). SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. Molecular cell *2*, 233-239.

Desterro, J.M., Rodriguez, M.S., Kemp, G.D., and Hay, R.T. (1999). Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1. The Journal of biological chemistry *274*, 10618-10624.

Desterro, J.M., Thomson, J., and Hay, R.T. (1997). Ubch9 conjugates SUMO but not ubiquitin. FEBS letters *417*, 297-300.

Dieckhoff, P., Bolte, M., Sancak, Y., Braus, G.H., and Irniger, S. (2004). Smt3/SUMO and Ubc9 are required for efficient APC/C-mediated proteolysis in budding yeast. Molecular microbiology *51*, 1375-1387.

Dimarcq, J.L., Hoffmann, D., Meister, M., Bulet, P., Lanot, R., Reichhart, J.M., and Hoffmann, J.A. (1994). Characterization and transcriptional profiles of a Drosophila gene encoding an insect defensin. A study in insect immunity. European journal of biochemistry / FEBS *221*, 201-209.

Ekengren, S., and Hultmark, D. (1999). Drosophila cecropin as an antifungal agent. Insect biochemistry and molecular biology *29*, 965-972.

El Chamy, L., Leclerc, V., Caldelari, I., and Reichhart, J.M. (2008). Sensing of 'danger signals' and pathogen-associated molecular patterns defines binary signaling pathways 'upstream' of Toll. Nature immunology *9*, 1165-1170.

Erturk-Hasdemir, D., Broemer, M., Leulier, F., Lane, W.S., Paquette, N., Hwang, D., Kim, C.H., Stoven, S., Meier, P., and Silverman, N. (2009). Two roles for the Drosophila IKK complex in the activation of Relish and the induction of antimicrobial peptide genes. Proceedings of the National Academy of Sciences of the United States of America *106*, 9779-9784.

Fehlbaum, P., Bulet, P., Michaut, L., Lagueux, M., Broekaert, W.F., Hetru, C., and Hoffmann, J.A. (1994). Insect immunity. Septic injury of Drosophila induces the synthesis of a potent antifungal peptide with sequence homology to plant antifungal peptides. The Journal of biological chemistry *269*, 33159-33163.

Ferrandon, D., Imler, J.L., Hetru, C., and Hoffmann, J.A. (2007). The Drosophila systemic immune response: sensing and signalling during bacterial and fungal infections. Nature reviews. Immunology *7*, 862-874.

Fukuyama, H., Verdier, Y., Guan, Y., Makino-Okamura, C., Shilova, V., Liu, X., Maksoud, E., Matsubayashi, J., Haddad, I., Spirohn, K., *et al.* (2013). Landscape of protein-protein interactions in Drosophila immune deficiency signaling during bacterial challenge. Proceedings of the National Academy of Sciences of the United States of America *110*, 10717-10722.

Fullaondo, A., Garcia-Sanchez, S., Sanz-Parra, A., Recio, E., Lee, S.Y., and Gubb, D. (2011). Spn1 regulates the GNBP3-dependent Toll signaling pathway in Drosophila melanogaster. Molecular and cellular biology *31*, 2960-2972.

Galisson, F., Mahrouche, L., Courcelles, M., Bonneil, E., Meloche, S., Chelbi-Alix, M.K., and Thibault, P. (2011). A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells. Molecular & cellular proteomics : MCP *10*, M110 004796.

Ganesan, S., Aggarwal, K., Paquette, N., and Silverman, N. (2011). NF-kappaB/Rel proteins and the humoral immune responses of Drosophila melanogaster. Current topics in microbiology and immunology *349*, 25-60.

Garver, L.S., Wu, J., and Wu, L.P. (2006). The peptidoglycan recognition protein PGRP-SC1a is essential for Toll signaling and phagocytosis of Staphylococcus aureus in Drosophila. Proceedings of the National Academy of Sciences of the United States of America *103*, 660-665.

Geiss-Friedlander, R., and Melchior, F. (2007). Concepts in sumoylation: a decade on. Nature reviews. Molecular cell biology *8*, 947-956.

Gingras, A.C., Gstaiger, M., Raught, B., and Aebersold, R. (2007). Analysis of protein complexes using mass spectrometry. Nature reviews. Molecular cell biology *8*, 645-654.

Golebiowski, F., Matic, I., Tatham, M.H., Cole, C., Yin, Y., Nakamura, A., Cox, J., Barton, G.J., Mann, M., and Hay, R.T. (2009). System-wide changes to SUMO modifications in response to heat shock. Science signaling *2*, ra24.

Gong, L., Li, B., Millas, S., and Yeh, E.T. (1999). Molecular cloning and characterization of human AOS1 and UBA2, components of the sentrin-activating enzyme complex. FEBS letters *448*, 185-189.

Gordon, M.D., Dionne, M.S., Schneider, D.S., and Nusse, R. (2005). WntD is a feedback inhibitor of Dorsal/NF-kappaB in Drosophila development and immunity. Nature *437*, 746-749.

Goto, A., Matsushita, K., Gesellchen, V., El Chamy, L., Kuttenkeuler, D., Takeuchi, O., Hoffmann, J.A., Akira, S., Boutros, M., and Reichhart, J.M. (2008). Akirins are highly conserved nuclear proteins required for NF-kappaB-dependent gene expression in drosophila and mice. Nature immunology *9*, 97-104.

Govind, S., and Steward, R. (1991). Dorsoventral pattern formation in Drosophila: signal transduction and nuclear targeting. Trends in genetics : TIG *7*, 119-125.

Gronholm, J., Ungureanu, D., Vanhatupa, S., Ramet, M., and Silvennoinen, O. (2010). Sumoylation of Drosophila transcription factor STAT92E. Journal of innate immunity *2*, 618-624.

Guntermann, S., Primrose, D.A., and Foley, E. (2009). Dnr1-dependent regulation of the Drosophila immune deficiency signaling pathway. Developmental and comparative immunology *33*, 127-134.

Guruharsha, K.G., Rual, J.F., Zhai, B., Mintseris, J., Vaidya, P., Vaidya, N., Beekman, C., Wong, C., Rhee, D.Y., Cenaj, O., *et al.* (2011). A protein complex network of Drosophila melanogaster. Cell *147*, 690-703.

Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature biotechnology *17*, 994-999.

Haghayeghi, A., Sarac, A., Czerniecki, S., Grosshans, J., and Schock, F. (2010). Pellino enhances innate immunity in Drosophila. Mechanisms of development *127*, 301-307.

Hannich, J.T., Lewis, A., Kroetz, M.B., Li, S.J., Heide, H., Emili, A., and Hochstrasser, M. (2005). Defining the SUMO-modified proteome by multiple approaches in Saccharomyces cerevisiae. The Journal of biological chemistry *280*, 4102-4110.

Hay, R.T. (2001). Protein modification by SUMO. Trends in biochemical sciences *26*, 332-333.

Hay, R.T. (2005). SUMO: a history of modification. Molecular cell *18*, 1-12.

Hecker, C.M., Rabiller, M., Haglund, K., Bayer, P., and Dikic, I. (2006). Specification of SUMO1- and SUMO2-interacting motifs. The Journal of biological chemistry *281*, 16117-16127.

Herrmann, J., Lerman, L.O., and Lerman, A. (2007). Ubiquitin and ubiquitin-like proteins in protein regulation. Circulation research *100*, 1276-1291.

Ho, C.W., Chen, H.T., and Hwang, J. (2011). UBC9 autosumoylation negatively regulates sumoylation of septins in Saccharomyces cerevisiae. The Journal of biological chemistry *286*, 21826-21834.

Hoffmann, J.A. (2003). The immune response of Drosophila. Nature *426*, 33-38.

Hsieh, Y.L., Kuo, H.Y., Chang, C.C., Naik, M.T., Liao, P.H., Ho, C.C., Huang, T.C., Jeng, J.C., Hsu, P.H., Tsai, M.D., *et al.* (2013). Ubc9 acetylation modulates distinct SUMO target modification and hypoxia response. The EMBO journal *32*, 791-804.

Hu, S., and Yang, X. (2000). dFADD, a novel death domain-containing adapter protein for the Drosophila caspase DREDD. The Journal of biological chemistry *275*, 30761-30764.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research *37*, 1-13.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols *4*, 44-57.

Huang, H.R., Chen, Z.J., Kunes, S., Chang, G.D., and Maniatis, T. (2010). Endocytic pathway is required for Drosophila Toll innate immune signaling. Proceedings of the National Academy of Sciences of the United States of America *107*, 8322-8327.

Huang, H.W., Tsoi, S.C., Sun, Y.H., and Li, S.S. (1998). Identification and characterization of the SMT3 cDNA and gene encoding ubiquitin-like protein from Drosophila melanogaster. Biochemistry and molecular biology international *46*, 775-785.

Huang, L., Ohsako, S., and Tanda, S. (2005). The lesswright mutation activates Rel-related proteins, leading to overproduction of larval hemocytes in Drosophila melanogaster. Developmental biology *280*, 407-420.

Huang, T.T., Wuerzberger-Davis, S.M., Wu, Z.H., and Miyamoto, S. (2003). Sequential modification of NEMO/IKKgamma by SUMO-1 and ubiquitin mediates NF-kappaB activation by genotoxic stress. Cell *115*, 565-576.

Hultmark, D. (2003). Drosophila immunity: paths and patterns. Current opinion in immunology *15*, 12-19.

Imler, J.L., and Bulet, P. (2005). Antimicrobial peptides in Drosophila: structures, activities and gene regulation. Chemical immunology and allergy *86*, 1-21.

Jang, I.H., Chosa, N., Kim, S.H., Nam, H.J., Lemaitre, B., Ochiai, M., Kambris, Z., Brun, S., Hashimoto, C., Ashida, M., *et al.* (2006). A Spatzle-processing enzyme required for toll signaling activation in Drosophila innate immunity. Developmental cell *10*, 45-55.

Jeon, Y.J., Yoo, H.M., and Chung, C.H. (2010). ISG15 and immune diseases. Biochimica et biophysica acta *1802*, 485-496.

Johansson, K.C., Metzendorf, C., and Soderhall, K. (2005). Microarray analysis of immune challenged Drosophila hemocytes. Experimental cell research *305*, 145-155.

Johnson, E.S., and Blobel, G. (1997). Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p. The Journal of biological chemistry *272*, 26799-26802.

Johnson, E.S., Schwienhorst, I., Dohmen, R.J., and Blobel, G. (1997). The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. The EMBO journal *16*, 5509-5519.

Jones, D., Crowe, E., Stevens, T.A., and Candido, E.P. (2002). Functional and phylogenetic analysis of the ubiquitylation system in Caenorhabditis elegans: ubiquitin-conjugating enzymes, ubiquitin-activating enzymes, and ubiquitin-like proteins. Genome biology *3*, RESEARCH0002.

Kallio, J., Leinonen, A., Ulvila, J., Valanne, S., Ezekowitz, R.A., and Ramet, M. (2005). Functional analysis of immune response genes in Drosophila identifies JNK pathway as a regulator of antimicrobial peptide gene expression in S2 cells. Microbes and infection / Institut Pasteur *7*, 811-819.

Kambris, Z., Brun, S., Jang, I.H., Nam, H.J., Romeo, Y., Takahashi, K., Lee, W.J., Ueda, R., and Lemaitre, B. (2006). Drosophila immunity: a large-scale in vivo RNAi screen identifies five serine proteases required for Toll activation. Current biology : CB *16*, 808-813.

Kaminsky, R., Denison, C., Bening-Abu-Shach, U., Chisholm, A.D., Gygi, S.P., and Broday, L. (2009). SUMO regulates the assembly and function of a cytoplasmic intermediate filament protein in C. elegans. Developmental cell *17*, 724-735.

Karlsson, C., Korayem, A.M., Scherfer, C., Loseva, O., Dushay, M.S., and Theopold, U. (2004). Proteomic analysis of the Drosophila larval hemolymph clot. The Journal of biological chemistry *279*, 52033-52041.

Karp, N.A., Huber, W., Sadowski, P.G., Charles, P.D., Hester, S.V., and Lilley, K.S. (2010). Addressing accuracy and precision issues in iTRAQ quantitation. Molecular & cellular proteomics : MCP *9*, 1885-1897.

Kerscher, O. (2007). SUMO junction-what's your function? New insights through SUMO-interacting motifs. EMBO reports *8*, 550-555.

Khush, R.S., Cornwell, W.D., Uram, J.N., and Lemaitre, B. (2002). A ubiquitin-proteasome pathway represses the Drosophila immune deficiency signaling cascade. Current biology : CB *12*, 1728-1737.

Kim, L.K., Choi, U.Y., Cho, H.S., Lee, J.S., Lee, W.B., Kim, J., Jeong, K., Shim, J., Kim-Ha, J., and Kim, Y.J. (2007). Down-regulation of NF-kappaB target genes by the AP-1 and STAT complex during the innate immune response in Drosophila. PLoS biology *5*, e238.

Kim, M., Lee, J.H., Lee, S.Y., Kim, E., and Chung, J. (2006). Caspar, a suppressor of antibacterial immunity in Drosophila. Proceedings of the National Academy of Sciences of the United States of America *103*, 16358-16363.

Kim, T., Yoon, J., Cho, H., Lee, W.B., Kim, J., Song, Y.H., Kim, S.N., Yoon, J.H., Kim-Ha, J., and Kim, Y.J. (2005). Downregulation of lipopolysaccharide response in Drosophila by negative crosstalk between the AP1 and NF-kappaB signaling modules. Nature immunology *6*, 211-218.

Kleino, A., Myllymaki, H., Kallio, J., Vanha-aho, L.M., Oksanen, K., Ulvila, J., Hultmark, D., Valanne, S., and Ramet, M. (2008). Pirk is a negative regulator of the Drosophila Imd pathway. Journal of immunology *180*, 5413-5422.

Kleino, A., and Silverman, N. (2014). The Drosophila IMD pathway in the activation of the humoral immune response. Developmental and comparative immunology *42*, 25-35.

Knipscheer, P., Flotho, A., Klug, H., Olsen, J.V., van Dijk, W.J., Fish, A., Johnson, E.S., Mann, M., Sixma, T.K., and Pichler, A. (2008). Ubc9 sumoylation regulates SUMO target discrimination. Molecular cell *31*, 371-382.

Kocks, C., Cho, J.H., Nehme, N., Ulvila, J., Pearson, A.M., Meister, M., Strom, C., Conto, S.L., Hetru, C., Stuart, L.M., *et al.* (2005). Eater, a transmembrane protein mediating phagocytosis of bacterial pathogens in Drosophila. Cell *123*, 335-346.

Kurepa, J., Walker, J.M., Smalle, J., Gosink, M.M., Davis, S.J., Durham, T.L., Sung, D.Y., and Vierstra, R.D. (2003). The small ubiquitin-like modifier (SUMO) protein modification system in Arabidopsis. Accumulation of SUMO1 and -2 conjugates is increased by stress. The Journal of biological chemistry *278*, 6862-6872.

Kurucz, E., Markus, R., Zsamboki, J., Folkl-Medzihradszky, K., Darula, Z., Vilmos, P., Udvardy, A., Krausz, I., Lukacsovich, T., Gateff, E., *et al.* (2007). Nimrod, a putative phagocytosis receptor with EGF repeats in Drosophila plasmatocytes. Current biology : CB *17*, 649-654.

Kuttenkeuler, D., Pelte, N., Ragab, A., Gesellchen, V., Schneider, L., Blass, C., Axelsson, E., Huber, W., and Boutros, M. (2010). A large-scale RNAi screen identifies Deaf1 as a regulator of innate immune responses in Drosophila. Journal of innate immunity *2*, 181-194.

Kylsten, P., Samakovlis, C., and Hultmark, D. (1990). The cecropin locus in Drosophila; a compact gene cluster involved in the response to infection. The EMBO journal *9*, 217-224.

Lanot, R., Zachary, D., Holder, F., and Meister, M. (2001). Postembryonic hematopoiesis in Drosophila. Developmental biology *230*, 243-257.

Lauwers, A., Twyffels, L., Soin, R., Wauquier, C., Kruys, V., and Gueydan, C. (2009). Post-transcriptional Regulation of Genes Encoding Anti-microbial Peptides in Drosophila. The Journal of biological chemistry *284*, 8973-8983.

Lee, K.Z., and Ferrandon, D. (2011). Negative regulation of immune responses on the fly. The EMBO journal *30*, 988-990.

Lemaitre, B., and Hoffmann, J. (2007). The host defense of Drosophila melanogaster. Annual review of immunology *25*, 697-743.

Lemaitre, B., Kromer-Metzger, E., Michaut, L., Nicolas, E., Meister, M., Georgel, P., Reichhart, J.M., and Hoffmann, J.A. (1995). A recessive mutation, immune deficiency (imd), defines two distinct control pathways in the Drosophila host defense. Proceedings of the National Academy of Sciences of the United States of America *92*, 9465-9469.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M., and Hoffmann, J.A. (1996). The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults. Cell *86*, 973-983.

Lemaitre, B., Reichhart, J.M., and Hoffmann, J.A. (1997). Drosophila host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. Proceedings of the National Academy of Sciences of the United States of America *94*, 14614-14619.

Lesch, C., Goto, A., Lindgren, M., Bidla, G., Dushay, M.S., and Theopold, U. (2007). A role for Hemolectin in coagulation and immunity in Drosophila melanogaster. Developmental and comparative immunology *31*, 1255-1263.

Levashina, E.A., Ohresser, S., Bulet, P., Reichhart, J.M., Hetru, C., and Hoffmann, J.A. (1995). Metchnikowin, a novel immune-inducible proline-rich peptide from Drosophila with antibacterial and antifungal properties. European journal of biochemistry / FEBS *233*, 694-700.

Ligoxygakis, P., Pelte, N., Hoffmann, J.A., and Reichhart, J.M. (2002). Activation of Drosophila Toll during fungal infection by a blood serine protease. Science *297*, 114-116.

Mahajan, R., Delphin, C., Guan, T., Gerace, L., and Melchior, F. (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. Cell *88*, 97-107.

Manfruelli, P., Reichhart, J.M., Steward, R., Hoffmann, J.A., and Lemaitre, B. (1999). A mosaic analysis in Drosophila fat body cells of the control of antimicrobial peptide genes by the Rel proteins Dorsal and DIF. The EMBO journal *18*, 3380-3391.

Matic, I., Macek, B., Hilger, M., Walther, T.C., and Mann, M. (2008). Phosphorylation of SUMO-1 occurs in vivo and is conserved through evolution. Journal of proteome research *7*, 4050-4057.

Matunis, M.J., Coutavas, E., and Blobel, G. (1996). A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. The Journal of cell biology *135*, 1457-1470.

Meinander, A., Runchel, C., Tenev, T., Chen, L., Kim, C.H., Ribeiro, P.S., Broemer, M., Leulier, F., Zvelebil, M., Silverman, N., *et al.* (2012). Ubiquitylation of the initiator caspase DREDD is required for innate immune signalling. The EMBO journal *31*, 2770-2783.

Melchior, F. (2000). SUMO--nonclassical ubiquitin. Annual review of cell and developmental biology *16*, 591-626.

Meluh, P.B., and Koshland, D. (1995). Evidence that the MIF2 gene of Saccharomyces cerevisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. Molecular biology of the cell *6*, 793-807.

Mi, H., Muruganujan, A., and Thomas, P.D. (2013). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic acids research *41*, D377-386.

Mishima, Y., Quintin, J., Aimanianda, V., Kellenberger, C., Coste, F., Clavaud, C., Hetru, C., Hoffmann, J.A., Latge, J.P., Ferrandon, D., *et al.* (2009). The N-terminal domain of Drosophila Gram-negative binding protein 3 (GNBP3) defines a novel family of fungal pattern recognition receptors. The Journal of biological chemistry *284*, 28687-28697.

Nacerddine, K., Lehembre, F., Bhaumik, M., Artus, J., Cohen-Tannoudji, M., Babinet, C., Pandolfi, P.P., and Dejean, A. (2005). The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. Developmental cell *9*, 769-779.

Nakamoto, M., Moy, R.H., Xu, J., Bambina, S., Yasunaga, A., Shelly, S.S., Gold, B., and Cherry, S. (2012). Virus recognition by Toll-7 activates antiviral autophagy in Drosophila. Immunity *36*, 658-667.

Nam, H.J., Jang, I.H., Asano, T., and Lee, W.J. (2008). Involvement of pro-phenoloxidase 3 in lamellocyte-mediated spontaneous melanization in Drosophila. Molecules and cells *26*, 606-610.

Nappi, A., Poirie, M., and Carton, Y. (2009). The role of melanization and cytotoxic by-products in the cellular immune responses of Drosophila against parasitic wasps. Advances in parasitology *70*, 99-121.

Nie, M., Xie, Y., Loo, J.A., and Courey, A.J. (2009). Genetic and proteomic evidence for roles of Drosophila SUMO in cell cycle control, Ras signaling, and early pattern formation. PloS one *4*, e5905.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and

accurate approach to expression proteomics. Molecular & cellular proteomics : MCP *1*, 376-386.

Oudshoorn, D., Versteeg, G.A., and Kikkert, M. (2012). Regulation of the innate immune system by ubiquitin and ubiquitin-like modifiers. Cytokine & growth factor reviews *23*, 273-282.

Ow, S.Y., Salim, M., Noirel, J., Evans, C., Rehman, I., and Wright, P.C. (2009). iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly". Journal of proteome research *8*, 5347-5355.

Panse, V.G., Hardeland, U., Werner, T., Kuster, B., and Hurt, E. (2004). A proteome-wide approach identifies sumoylated substrate proteins in yeast. The Journal of biological chemistry *279*, 41346-41351.

Papouli, E., Chen, S., Davies, A.A., Huttner, D., Krejci, L., Sung, P., and Ulrich, H.D. (2005). Crosstalk between SUMO and ubiquitin on PCNA is mediated by recruitment of the helicase Srs2p. Molecular cell *19*, 123-133.

Paquette, N., Broemer, M., Aggarwal, K., Chen, L., Husson, M., Erturk-Hasdemir, D., Reichhart, J.M., Meier, P., and Silverman, N. (2010). Caspase-mediated cleavage, IAP binding, and ubiquitination: linking three mechanisms crucial for Drosophila NF-kappaB signaling. Molecular cell *37*, 172-182.

Park, M.Y., Moon, J.H., Lee, K.S., Choi, H.I., Chung, J., Hong, H.J., and Kim, E. (2007). FAF1 suppresses IkappaB kinase (IKK) activation by disrupting the IKK complex assembly. The Journal of biological chemistry *282*, 27572-27577.

Perry, J.J., Tainer, J.A., and Boddy, M.N. (2008). A SIM-ultaneous role for SUMO and ubiquitin. Trends in biochemical sciences *33*, 201-208.

Philips, J.A., Rubin, E.J., and Perrimon, N. (2005). Drosophila RNAi screen reveals CD36 family member required for mycobacterial infection. Science *309*, 1251-1253.

Praefcke, G.J., Hofmann, K., and Dohmen, R.J. (2012). SUMO playing tag with ubiquitin. Trends in biochemical sciences *37*, 23-31.

Ragab, A., Buechling, T., Gesellchen, V., Spirohn, K., Boettcher, A.L., and Boutros, M. (2011). Drosophila Ras/MAPK signalling regulates innate immune responses in immune and intestinal stem cells. The EMBO journal *30*, 1123-1136.

Rajendra, R., Malegaonkar, D., Pungaliya, P., Marshall, H., Rasheed, Z., Brownell, J., Liu, L.F., Lutzker, S., Saleem, A., and Rubin, E.H. (2004). Topors functions as an E3 ubiquitin ligase with specific E2 enzymes and ubiquitinates p53. The Journal of biological chemistry *279*, 36440-36444.

Ramet, M., Manfruelli, P., Pearson, A., Mathey-Prevot, B., and Ezekowitz, R.A. (2002). Functional genomic analysis of phagocytosis and identification of a Drosophila receptor for E. coli. Nature *416*, 644-648.

Ramet, M., Pearson, A., Manfruelli, P., Li, X., Koziel, H., Gobel, V., Chung, E., Krieger, M., and Ezekowitz, R.A. (2001). Drosophila scavenger receptor CI is a pattern recognition receptor for bacteria. Immunity *15*, 1027-1038.

Reverter, D., and Lima, C.D. (2005). Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. Nature *435*, 687-692.

Rizki, R.M., and Rizki, T.M. (1984). Selective destruction of a host blood cell type by a parasitoid wasp. Proceedings of the National Academy of Sciences of the United States of America *81*, 6154-6158.

Rizki, T.M., Rizki, R.M., and Bellotti, R.A. (1985). Genetics of a Drosophila phenoloxidase. Molecular & general genetics : MGG *201*, 7-13.

Rodriguez, M.S., Dargemont, C., and Hay, R.T. (2001). SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. The Journal of biological chemistry *276*, 12654-12659.

Rosas-Acosta, G., Russell, W.K., Deyrieux, A., Russell, D.H., and Wilson, V.G. (2005). A universal strategy for proteomic studies of SUMO and other ubiquitin-like modifiers. Molecular & cellular proteomics : MCP *4*, 56-72.

Rosetto, M., Engstrom, Y., Baldari, C.T., Telford, J.L., and Hultmark, D. (1995). Signals from the IL-1 receptor homolog, Toll, can activate an immune response in a Drosophila hemocyte cell line. Biochemical and biophysical research communications *209*, 111-116.

Ross, P.L., Huang, Y.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., *et al.* (2004). Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Molecular & cellular proteomics : MCP *3*, 1154-1169.

Russo, J., Brehelin, M., and Carton, Y. (2001). Haemocyte changes in resistant and susceptible strains of D. melanogaster caused by virulent and avirulent strains of the parasitic wasp Leptopilina boulardi. Journal of insect physiology *47*, 167-172.

Rutschmann, S., Jung, A.C., Hetru, C., Reichhart, J.M., Hoffmann, J.A., and Ferrandon, D. (2000). The Rel protein DIF mediates the antifungal but not the antibacterial host defense in Drosophila. Immunity *12*, 569-580.

Samakovlis, C., Kimbrell, D.A., Kylsten, P., Engstrom, A., and Hultmark, D. (1990). The immune response in Drosophila: pattern of cecropin expression and biological activity. The EMBO journal *9*, 2969-2976.

Scherfer, C., Karlsson, C., Loseva, O., Bidla, G., Goto, A., Havemann, J., Dushay, M.S., and Theopold, U. (2004). Isolation and characterization of hemolymph clotting factors in Drosophila melanogaster by a pullout method. Current biology : CB *14*, 625-629.

Scherfer, C., Qazi, M.R., Takahashi, K., Ueda, R., Dushay, M.S., Theopold, U., and Lemaitre, B. (2006). The Toll immune-regulated Drosophila protein Fondue is involved in hemolymph clotting and puparium formation. Developmental biology *295*, 156-163.

Schneider, I. (1972). Cell lines derived from late embryonic stages of Drosophila melanogaster. Journal of embryology and experimental morphology *27*, 353-365.

Schulze, W.X., and Usadel, B. (2010). Quantitation in mass-spectrometry-based proteomics. Annual review of plant biology *61*, 491-516.

Schwartz, D.C., and Hochstrasser, M. (2003). A superfamily of protein tags: ubiquitin, SUMO and related modifiers. Trends in biochemical sciences *28*, 321-328.

Senger, K., Armstrong, G.W., Rowell, W.J., Kwan, J.M., Markstein, M., and Levine, M. (2004). Immunity regulatory DNAs share common organizational features in Drosophila. Molecular cell *13*, 19-32.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research *13*, 2498-2504.

Shuai, K., and Liu, B. (2003). Regulation of JAK-STAT signalling in the immune system. Nature reviews. Immunology *3*, 900-911.

Silverman, N., Zhou, R., Erlich, R.L., Hunter, M., Bernstein, E., Schneider, D., and Maniatis, T. (2003). Immune activation of NF-kappaB and JNK requires Drosophila TAK1. The Journal of biological chemistry *278*, 48928-48934.

Silverman, N., Zhou, R., Stoven, S., Pandey, N., Hultmark, D., and Maniatis, T. (2000). A Drosophila IkappaB kinase complex required for Relish cleavage and antibacterial immunity. Genes & development *14*, 2461-2471.

Skaug, B., and Chen, Z.J. (2010). Emerging role of ISG15 in antiviral immunity. Cell *143*, 187-190.

Sluss, H.K., Han, Z., Barrett, T., Goberdhan, D.C., Wilson, C., Davis, R.J., and Ip, Y.T. (1996). A JNK signal transduction pathway that mediates morphogenesis and an immune response in Drosophila. Genes & development *10*, 2745-2758.

Smith, M., Bhaskar, V., Fernandez, J., and Courey, A.J. (2004). Drosophila Ulp1, a nuclear pore-associated SUMO protease, prevents accumulation of cytoplasmic SUMO conjugates. The Journal of biological chemistry *279*, 43805-43814.

Soderhall, K., and Cerenius, L. (1998). Role of the prophenoloxidase-activating system in invertebrate immunity. Current opinion in immunology *10*, 23-28.

Song, J., Durrin, L.K., Wilkinson, T.A., Krontiris, T.G., and Chen, Y. (2004). Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. Proceedings of the National Academy of Sciences of the United States of America *101*, 14373-14378.

Sorrentino, R.P., Carton, Y., and Govind, S. (2002). Cellular immune response to parasite infection in the Drosophila lymph gland is developmentally regulated. Developmental biology *243*, 65-80.

Stoven, S., Ando, I., Kadalayil, L., Engstrom, Y., and Hultmark, D. (2000). Activation of the Drosophila NF-kappaB factor Relish by rapid endoproteolytic cleavage. EMBO reports *1*, 347-352.

Stoven, S., Silverman, N., Junell, A., Hedengren-Olcott, M., Erturk, D., Engstrom, Y., Maniatis, T., and Hultmark, D. (2003). Caspase-mediated processing of the Drosophila NF-kappaB factor Relish. Proceedings of the National Academy of Sciences of the United States of America *100*, 5991-5996.

Stroschein-Stevenson, S.L., Foley, E., O'Farrell, P.H., and Johnson, A.D. (2006). Identification of Drosophila gene products required for phagocytosis of Candida albicans. PLoS biology *4*, e4.

Stuart, L.M., Boulais, J., Charriere, G.M., Hennessy, E.J., Brunet, S., Jutras, I., Goyette, G., Rondeau, C., Letarte, S., Huang, H., *et al.* (2007). A systems biology analysis of the Drosophila phagosome. Nature *445*, 95-101.

Stuart, L.M., Deng, J., Silver, J.M., Takahashi, K., Tseng, A.A., Hennessy, E.J., Ezekowitz, R.A., and Moore, K.J. (2005). Response to Staphylococcus aureus requires CD36-mediated phagocytosis triggered by the COOH-terminal cytoplasmic domain. The Journal of cell biology *170*, 477-485.

Sun, H., Bristow, B.N., Qu, G., and Wasserman, S.A. (2002). A heterotrimeric death domain complex in Toll signaling. Proceedings of the National Academy of Sciences of the United States of America *99*, 12871-12876.

Talamillo, A., Sanchez, J., and Barrio, R. (2008). Functional analysis of the SUMOylation pathway in Drosophila. Biochemical Society transactions *36*, 868-873.

Tatham, M.H., Matic, I., Mann, M., and Hay, R.T. (2011). Comparative proteomic analysis identifies a role for SUMO in protein quality control. Science signaling *4*, rs4.

Tempe, D., Piechaczyk, M., and Bossis, G. (2008). SUMO under stress. Biochemical Society transactions *36*, 874-878.

Thevenon, D., Engel, E., Avet-Rochex, A., Gottar, M., Bergeret, E., Tricoire, H., Benaud, C., Baudier, J., Taillebourg, E., and Fauvarque, M.O. (2009). The Drosophila ubiquitin-specific protease dUSP36/Scny targets IMD to prevent constitutive immune signaling. Cell host & microbe *6*, 309-320.

Thomas, P.D., Kejariwal, A., Campbell, M.J., Mi, H., Diemer, K., Guo, N., Ladunga, I., Ulitsky-Lazareva, B., Muruganujan, A., Rabkin, S., *et al.* (2003). PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. Nucleic acids research *31*, 334-341.

Tsichritzis, T., Gaentzsch, P.C., Kosmidis, S., Brown, A.E., Skoulakis, E.M., Ligoxygakis, P., and Mosialos, G. (2007). A Drosophila ortholog of the human cylindromatosis tumor suppressor gene regulates triglyceride content and antibacterial defense. Development *134*, 2605-2614.

Tsuda, M., Langmann, C., Harden, N., and Aigaki, T. (2005). The RING-finger scaffold protein Plenty of SH3s targets TAK1 to control immunity signalling in Drosophila. EMBO reports *6*, 1082-1087.

Ulrich, H.D. (2008). The fast-growing business of SUMO chains. Molecular cell *32*, 301-305.

Um, J.W., and Chung, K.C. (2006). Functional modulation of parkin through physical interaction with SUMO-1. Journal of neuroscience research *84*, 1543-1554.

Uzunova, K., Gottsche, K., Miteva, M., Weisshaar, S.R., Glanemann, C., Schnellhardt, M., Niessen, M., Scheel, H., Hofmann, K., Johnson, E.S., *et al.* (2007). Ubiquitin-dependent proteolytic control of SUMO conjugates. The Journal of biological chemistry *282*, 34167-34175.

Valanne, S., Kallio, J., Kleino, A., and Ramet, M. (2012). Large-scale RNAi screens add both clarity and complexity to Drosophila NF-kappaB signaling. Developmental and comparative immunology *37*, 9-18.

Valanne, S., Myllymaki, H., Kallio, J., Schmid, M.R., Kleino, A., Murumagi, A., Airaksinen, L., Kotipelto, T., Kaustio, M., Ulvila, J., *et al.* (2010). Genome-wide RNA interference in Drosophila cells identifies G protein-coupled receptor kinase 2 as a conserved regulator of NF-kappaB signaling. Journal of immunology *184*, 6188-6198.

Vertegaal, A.C., Andersen, J.S., Ogg, S.C., Hay, R.T., Mann, M., and Lamond, A.I. (2006). Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. Molecular & cellular proteomics : MCP *5*, 2298-2310.

Vertegaal, A.C., Ogg, S.C., Jaffray, E., Rodriguez, M.S., Hay, R.T., Andersen, J.S., Mann, M., and Lamond, A.I. (2004). A proteomic study of SUMO-2 target proteins. The Journal of biological chemistry *279*, 33791-33798.

Wang, L., Gilbert, R.J., Atilano, M.L., Filipe, S.R., Gay, N.J., and Ligoxygakis, P. (2008). Peptidoglycan recognition protein-SD provides versatility of receptor formation in Drosophila immunity. Proceedings of the National Academy of Sciences of the United States of America *105*, 11881-11886.

Wang, L., Weber, A.N., Atilano, M.L., Filipe, S.R., Gay, N.J., and Ligoxygakis, P. (2006). Sensing of Gram-positive bacteria in Drosophila: GNBP1 is needed to process and present peptidoglycan to PGRP-SA. The EMBO journal *25*, 5005-5014.

Wang, Z., Wilhelmsson, C., Hyrsl, P., Loof, T.G., Dobes, P., Klupp, M., Loseva, O., Morgelin, M., Ikle, J., Cripps, R.M., *et al.* (2010). Pathogen entrapment by transglutaminase--a conserved early innate immune mechanism. PLoS pathogens *6*, e1000763.

Watson, F.L., Puttmann-Holgado, R., Thomas, F., Lamar, D.L., Hughes, M., Kondo, M., Rebel, V.I., and Schmucker, D. (2005). Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science *309*, 1874-1878.

Weber, A.N., Gangloff, M., Moncrieffe, M.C., Hyvert, Y., Imler, J.L., and Gay, N.J. (2007). Role of the Spatzle Pro-domain in the generation of an active toll receptor ligand. The Journal of biological chemistry *282*, 13522-13531.

Weger, S., Hammer, E., and Heilbronn, R. (2005). Topors acts as a SUMO-1 E3 ligase for p53 in vitro and in vivo. FEBS letters *579*, 5007-5012.

Wei, Y., Xiao, Q., Zhang, T., Mou, Z., You, J., and Ma, W.J. (2009). Differential regulation of mRNA stability controls the transient expression of genes encoding Drosophila antimicrobial peptide with distinct immune response characteristics. Nucleic acids research *37*, 6550-6561.

Wicker, C., Reichhart, J.M., Hoffmann, D., Hultmark, D., Samakovlis, C., and Hoffmann, J.A. (1990). Insect immunity. Characterization of a Drosophila cDNA encoding a novel member of the diptericin family of immune peptides. The Journal of biological chemistry *265*, 22493-22498.

Wilkinson, K.A., and Henley, J.M. (2010). Mechanisms, regulation and consequences of protein SUMOylation. The Biochemical journal *428*, 133-145.

Wohlschlegel, J.A., Johnson, E.S., Reed, S.I., and Yates, J.R., 3rd (2004). Global analysis of protein sumoylation in Saccharomyces cerevisiae. The Journal of biological chemistry *279*, 45662-45668.

Wykoff, D.D., and O'Shea, E.K. (2005). Identification of sumoylated proteins by systematic immunoprecipitation of the budding yeast proteome. Molecular & cellular proteomics : MCP *4*, 73-83.

Xue, Y., Zhou, F., Fu, C., Xu, Y., and Yao, X. (2006). SUMOsp: a web server for sumoylation site prediction. Nucleic acids research *34*, W254-257.

Zhang, F.P., Mikkonen, L., Toppari, J., Palvimo, J.J., Thesleff, I., and Janne, O.A. (2008). Sumo-1 function is dispensable in normal mouse development. Molecular and cellular biology *28*, 5381-5390.

Zhao, J. (2007). Sumoylation regulates diverse biological processes. Cellular and molecular life sciences : CMLS *64*, 3017-3033.

Zhou, R., Silverman, N., Hong, M., Liao, D.S., Chung, Y., Chen, Z.J., and Maniatis, T. (2005). The role of ubiquitination in Drosophila innate immunity. The Journal of biological chemistry *280*, 34048-34055.

Zhou, W., Ryan, J.J., and Zhou, H. (2004). Global analyses of sumoylated proteins in Saccharomyces cerevisiae. Induction of protein sumoylation by cellular stresses. The Journal of biological chemistry *279*, 32262-32268.