**A REPORT**

**ON**

# STRUCTURE BASED PREDICTION OF INTERACTING PARTNERS OF WDR13

BY


| | | |
|---|---|---|
| Ashish Baghudana | 2011B1A7575G | MSc. (Hons) Biological Science<br>BE (Hons) Computer Science |


AT

**Medical Research Foundation (Sankara Nethralaya), Chennai**

A PRACTICE SCHOOL–I STATION OF


**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI**
**K. K. BIRLA GOA CAMPUS**

JULY, 2013

**A REPORT**

**ON**

# STRUCTURE BASED PREDICTION OF INTERACTING PARTNERS OF WDR13

BY


| | | |
|---|---|---|
| Ashish Baghudana | 2011B1A7575G | MSc. (Hons) Biological Science<br>BE (Hons) Computer Science |


Prepared in partial fulfillment of the
Practice School-I Course
BITS C221/ BITS C231/ BITS GC221/ BITS 241/BITS GC231

AT

**Medical Research Foundation (Sankara Nethralaya), Chennai**

A PRACTICE SCHOOL–I STATION OF



**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI**
**K. K. BIRLA GOA CAMPUS**


JULY, 2013

**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE**
**PILANI (RAJASTHAN)**

**Practice School Division**

| | | |
|---|---|---|
| **Station** | **:** | **Medical Research Foundation** |
| **Centre** | **:** | Chennai |
| **Duration** | **:** | 7 Weeks |
| **Date of Start** | **:** | May 22nd, 2013 |
| **Date of Submission** | **:** | July 10th, 2013 |
| **Name/ ID No.** | **:** | Ashish Baghudana [ID No. 2011B1A7575G] |
| **Discipline of the Student** | **:** | MSc. (Hons) Biological Science<br>B.E. (Hons) Computer Science |
| **Project Title** | **:** | Structure Based Prediction of Interacting Partners of WDR13 |
| **Experts' Name and Designation** | **:** | **Dr. V Umashankar**<br>*HOD, Bioinformatics*<br>**Mr. Mohammed Al-Ameen**<br>*Junior Scientist, Bioinformatics* |
| **Name of the PS Faculty** | **:** | **Dr. P. R. Deepa** |
| **Keywords** | **:** | WDR13, WD40 Domains, Calcium, MODELLER, docking |
| **Project Areas** | **:** | Bioinformatics, Prediction of Interacting Partners |

**ABSTRACT**

WDR13 is a highly conserved protein in eukaryotes. Its function has been implicated in both diabetes and cancer. It is viewed as a potential drug target for curing diabetes. Recent studies suggest a role in regulating repair of neuronal injury through a Calcium dependent pathway. We have predicted the structure of WDR13 using homology modelling and fold recognition. The structure has been energy minimized using GROMACS. The resultant structure was very stable in a molecular dynamics simulation with less than 3 angstrom. This structure was docked with 12 proteins, selected through extensive literature survey, colocalization and coexpression studies and database predictions, using 5 different servers: ClusPro, GrammX, HexDock, PatchDock and ZDock. We found three proteins FKBP1A, PIM1 and TGFβR1 as potential interacting partners with WDR13 from our study. These proteins match closely with earlier studies on WDR13 implicating it in p21 control and its association with Calcineurin. Further validation using electrostatic studies, heat maps and finally co-immunoprecipitation studies would be useful in ascertaining this result.

**Signature of Student** **Signature of PS Faculty**

**Date** **Date**

**ACKNOWLEDGEMENTS**

First and foremost, I would like to thank Dr. V Umashankar, who gave me the opportunity to work in his lab. His instruction and guidance have been extremely valuable in my understanding of the project. I also want to thank Mr. Mohammed Al-Ameen, who has been overseeing my work on a daily basis and ensuring I have had all the help I needed to work my way through. Without their support and input, I would never have made any headway in my work. I would also like to offer my sincere thanks to Dr. P R Deepa, our PS instructor for supervising my work and giving me useful tips at all times that have helped me not only plan my project but also ensured that I learn research methodology thoroughly.

This project would not have been possible without the tremendous efforts of Professor B N Jain (Vice-Chancellor of BITS-Pilani), Professor K E Raman (Director, BITS-Pilani K K Birla Goa Campus), Professor G Sundar (Deputy Director, BITS-Pilani Off-Campus Programmes), Professor Niranjan Swain (Dean, PSD, BITS-Pilani) and Professor Sutapa Roy Ramanan (Faculty in-charge, PSD, BITS-Pilani Goa Campus) in organizing the PS-I course for us and choosing some of the best labs in the country for students to get first-hand practical experience.

I also take this opportunity to thank Dr. S S Badrinath (Founder, Medical Research Foundation), Dr. H N Madhavan (President, Vision Research Foundation), Dr. Lily Therese (Coordinator, VIBS), Ms. Amurthavalli (Secretary, VIBS) and all the other faculty and students I have interacted with in Sankara Nethralaya and VIBS who have made me feel welcome here. A special mention must be given to the HR department of Sankara Nethralaya who have ensured a comfortable stay at Sankara Nethralaya for us.

# Table of Contents

# LIST OF ABBREVIATIONS

| | |
|---|---|
| 1. PS-I | Practice School I |
| 2. SN | Sankara Nethralaya |
| 3. CSR | Corporate Social Responsibility |
| 4. WD40 | Protein Domain consisting of about 40 amino acids generally ending with Tryptophan and Aspartic Acid |
| 5. ERα/β | Estrogen Receptors α/β |
| 6. HDAC | Histone Deacetylation |
| 7. PHIP | Pleckstrin Homology Domain Interaction Protein |
| 8. AAAS | Triple-A (Allgrove) Syndrome, a neurological disorder that may be caused by deletion of Wdr13. |
| 9. PPP3CA | Calcineurin |
| 10. TGFβ | Transforming Growth Factor β |
| 11. PIM1 | Proto-oncogene serine/threonine-protein kinase |

## 1. PS-I ORGANIZATIONAL PROFILE

Sankara Nethralaya is a not-for-profit tertiary eye care center in Chennai, India. It was founded by Dr. Sengamedu Srinivasa Badrinath in 1978 under the guidance of Sri Jayendra Saraswathi. The hospital is run on missionary principles serving over 1000 patients every day. It has close to 1000 employees across 9 campuses in Chennai, Rameshwaram, Kolkata, Bengaluru and Tirupati.

Sankara Nethralaya is extremely committed to community service. In its effort to serve the poorer sections of our population, Sankara Nethralaya has established Jaslok Community Ophthalmic Centre that offers free services to all economically weaker people. In fact, majority of the funding for the hospital is achieved through voluntary donations from companies and individuals alike. Nani Palkhivala, a prominent Indian jurist and economist, pledged his entire life savings to Sankara Nethralaya for the benefit of the patients. It is also extremely appreciable that all buildings and labs functioning in Sankara Nethralaya are almost fully funded by corporations as part of their CSR.

Sankara Nethralaya also has an education wing called Vidyasagar Institute of Biomedical Technology and Science which collaborates with BITS-Pilani for higher education programmes such as MS-MLT (Medical Laboratory Techniques). It also takes in students for short research projects and has been a PS station of BITS-Pilani for 18 years.

Research facilities at Sankara Nethralaya are top notch as well. A multitude of labs work in the fields of Biochemistry, Microbiology, Nano-biotechnology, Ocular Pathology, Genetics and Bioinformatics. The Bioinformatics Lab (where I have been working) is well equipped with good computers equipped with very good configurations of Intel Xeon Quad Core Processor, 32 GB RAM and 2 GB Graphics Processing RAM. High quality commercial software suite (Schrodinger) is used for computation

## 2. INTRODUCTION

WDR13 belongs to a family of WD40 repeat proteins, reported by L. Singh *et al* in 2003. It contains 9 exons and 8 introns and was mapped to the X chromosome by fluorescent in-situ hybridization and *in silico* mapping. It is encoded by a 485 amino acid sequence that contains 6 WD repeats. Basic knowledge of the WD domain is critical to the understanding of the WDR13 and predicting its structure and function.

### 2.1 WD40 DOMAINS: STRUCTURE AND FUNCTION

A WD repeat is a conserved motif of about 40 amino acids that usually end with Tryptophan (W) and Aspartic Acid (D), hence giving the family its name WD[1]. This motif was first discovered in the β subunit of G-heterotrimeric proteins that are involved in cell signaling. The 7 repeats in $G_\beta$ form a highly symmetrical propeller-like structure. This has been found characteristic of many WD repeat proteins now. It is believed that such a structure is the key to many protein-protein interactions. Interestingly, WD repeat motif has not been found in prokaryotic organisms except in a few cases. However, WD40 domain is among the top ten most abundant domains in eukaryotic genomes. The diversity in both structure and function of WD40 proteins in remarkable.

WD Repeat proteins have a multitude of functions in eukaryotes. These include signal transduction, transcriptional regulation, cell cycle regulation and cytoskeletal organization to mention a few[2]. The beta propeller formed from the WD repeats have two characteristic features:

- WD repeat folds play a significant role in the formation of protein complexes and thereby mediate many protein-protein interactions.
- WD repeats do not have any catalytic function[3].

[1] Smith, 2008
[2] C.U. Stirnimann, 2010
[3] C Xu, 2011

### 2.1.1 STRUCTURE OF WD REPEATS: β Propeller and Velcro Closure

While there is no single amino acid conserved in the WD repeat motif, the general structure of the fold is conserved. This makes identification of WD motifs a little hard. However, despite the differences in composition of amino acids in the repeat, the basic pattern remains conserved. A typical WD protein would look like the following.

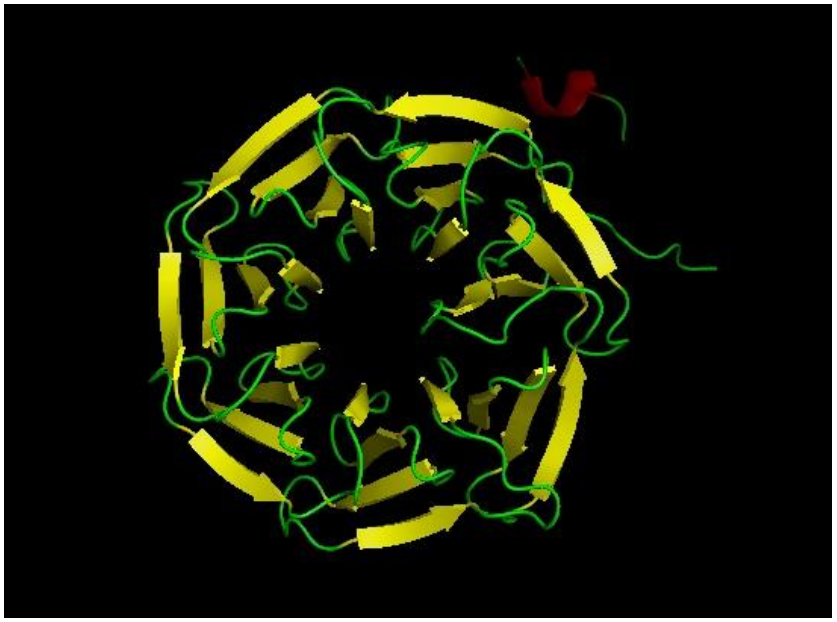


*Figure 1 Structure of WDR5, a WD repeat protein involed in histone modification. This image was visualized using PyMOL from the PDB structure of WDR5.*

Each blade of the propeller-like structure consists of 4 anti-parallel strands. Of the four strands, the first three strands belong to one blade, while the fourth actually belongs to an adjacent blade. Such an overlap between two blades provides a "Velcro Closure" that stabilizes the overall fold of the protein. Some studies however find Velcro Closure not a necessity for the formation of the propeller, but only as incidental. The symmetry of the propeller is explained by the uniform height and width of each blade within a protein. However, the number of amino acids in each blade can vary from 4 to 10 in different proteins. Equally interestingly, a few WD40 proteins do not, by themselves, have enough repeats to form a stable beta propeller. They borrow a blade from their interacting partner to form a

stable complex. Such a phenomenon is seen in Sec13-Nup145C and Seh1-Nup85 nucleoporin pair complex.

Given its ability to facilitate the formation of protein complexes, researchers have found that the beta propeller interacts predominantly through its smaller top surface, but also through its bottom surface and sides[3].

The general sequence of WD repeats is as follows:

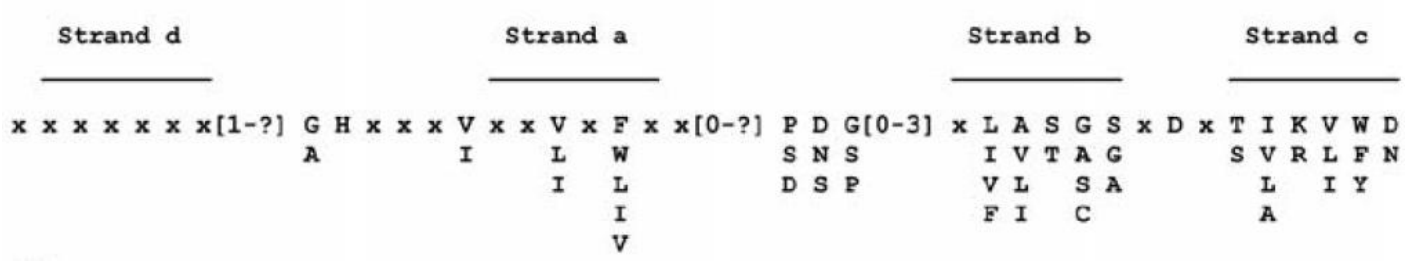



*Figure 2 The sequence of WD repeats*

## 2.2 WDR13: OBSERVATIONS

WDR13 was mapped the X chromosome at the locus Xp11.23 by *in silico* mapping[4]. While this protein is expressed across all tissues, a relatively high expression was found in pancreas, brain, testis and ovaries. The deletion of this gene resulted in mental retardation, obesity and xeroderma. Sathish Kumar *et al* found that knocking out the WDR13 resulted in higher serum insulin levels and increased pancreatic islet mass as a result of enhanced beta cell proliferation[5]. This also led to the downregulation of p21, a cell cycle inhibitor. Conversely, over expression of WDR13 led to the upregulation of p21. In a patent published in 2012, Sathish Kumar & Vijay Pratap Singh reported that WDR13 is a novel negative regulator of beta cell proliferation and acts as an inhibitor of estrogen receptors mediated transcription. They also identified interacting partners of WDR13 as ERα, ERβ, PHIP1, HDAC1, HDAC3 and HDAC7, suggesting a role for WDR13 in chromatin regulation.

[4] L. Singh, 2002
[5] S. Kumar, 2012

Furthermore, ERα and HDAC1 have been reported as a regulator of p21 reconciling the seemingly disconnected upregulation of p21 during over expression of WDR13.

WDR13 has a sequence of:

>sp|Q9H1Z4|WDR13_HUMAN WD repeat-containing protein 13 OS=Homo sapiens GN=WDR13 PE=1 SV=2
MAAVWQQVLAVDARYNAYRTPTFPQFRTQYIRRRSQLLRENAKAGHPPALRRQYLRLRGQ
LLGQRYGPLSEPGSARAYSNSIVRSSRTTLDRMEDFEDDPRALGARGHRRSVSRGSYQLQ
AQMNRAVYEDRPPGSVVPTSAAEASRAMAGDTSLSENYAFAGMYHVFDQ**HVDEAVPRVRF**
**ANDDRHRLACCSLDGSISLCQLVPAPPTVL**RVLR**GHTRGVSDFAWSLSNDILVSTSLDAT**
**MRIWASEDGRCIRE**IPDPDSAELLCCTFQPVNNNLTVVGNAKHNVHVMNISTGKKVKGGS
S**KLTGRVLALSFDAPGRLLWAGDDHGSVFSFLFDMATGKLT**KAKRLVVHEGSPVTSISAR
SWVSREARDPSLLINACLNKLLLYRVVDNEGTLQLKRSFPIEQSS**HPVRSIFCPLMSFRQ**
**GACVVTGSEDMCVHFFDVERAAKAAV**NKLQ**GHSAPVLDVSFNCDESLLASSDASGMVIVW**
**RREQ**K

The WD repeat domains are highlighted in the FASTA sequence above[6]. Apart from the WD repeats, it also contains the *LxxLL* motif[7] that it might use to interact with the Estrogen Receptors and HDAC proteins. This motif is also highlighted in the FASTA sequence above.

WDR13 has also been observed at high concentrations in the brain, specifically the hippocampus. Price *et al* (2003) reported high expression level of the then novel gene WDR13 as part of neuronal response to injury[8]. Synaptic Plasticity, the ability of the synapse to change its strength in response to either use or disuse of transmission over synaptic pathways, is important for the formation of long term memories and in re-establishment of function post injury. The 7 genes identified to have been high expressed in the hippocampus post injury are: Calcineurin, Astrotactin, two G-proteins, two proteins of unknown functions and WDR13. They propose that WDR13 might have an important role in regulation of response to neurodegeneration and reactive synaptogenesis, given that WDR13 mutation is implicated in the Triple-A syndrome (AAAS).

[6] UniProt: Q9H1Z4
[7] S. Kumar, 2012
[8] Price, 2003

## 3. SCOPE AND AIM OF THE PROJECT

My project titled "Structure-based Prediction of Interacting Partners of WDR13" has the following objectives:

i. To study WD40 domains and understand their role in different proteins.
ii. To study specifically, WDR13, a WD40 protein that is highly conserved among higher eukaryotes.
iii. To predict the structure of WDR13 and to understand its structure-function relationship.
iv. To predict potential interacting partners of WDR13 on the basis of its structure.

The structure of WDR13 is still not known through X-Ray Crystallography or NMR studies. Given the large amount of protein sequences generated on an everyday basis, *in silico* methods have gained popularity to predict protein structures. The algorithms used are fairly accurate and reliable. We use these algorithms to predict the structure of WDR13. On the basis of its predicted structure, we hope to identify its interacting partners using physical docking studies. Recent studies have claimed that true and false interacting partners can be distinguished using docking scores. We hope to validate this hypothesis and standardize the protocol to identify genome-wide binary interactions.

This project, hence, serves multiple purposes:

a. Study WDR13 as a protein involved in neuronal regrowth
b. Understand the mechanism in which WDR13 regulates p21
c. Understand its association with Calcineurin and the possible role it plays in the hippocampus
d. Set a standard protocol to identify interacting partners on the basis of physical docking

## 4. METHODOLOGY

WDR13 is a 485 amino-acid protein that has the first WD40 domain from the 161st residue. It has a putative 6 WD40 domains forming a beta-propeller. The raw sequence of WDR13 was given to PSIPRED to predict the secondary structure.

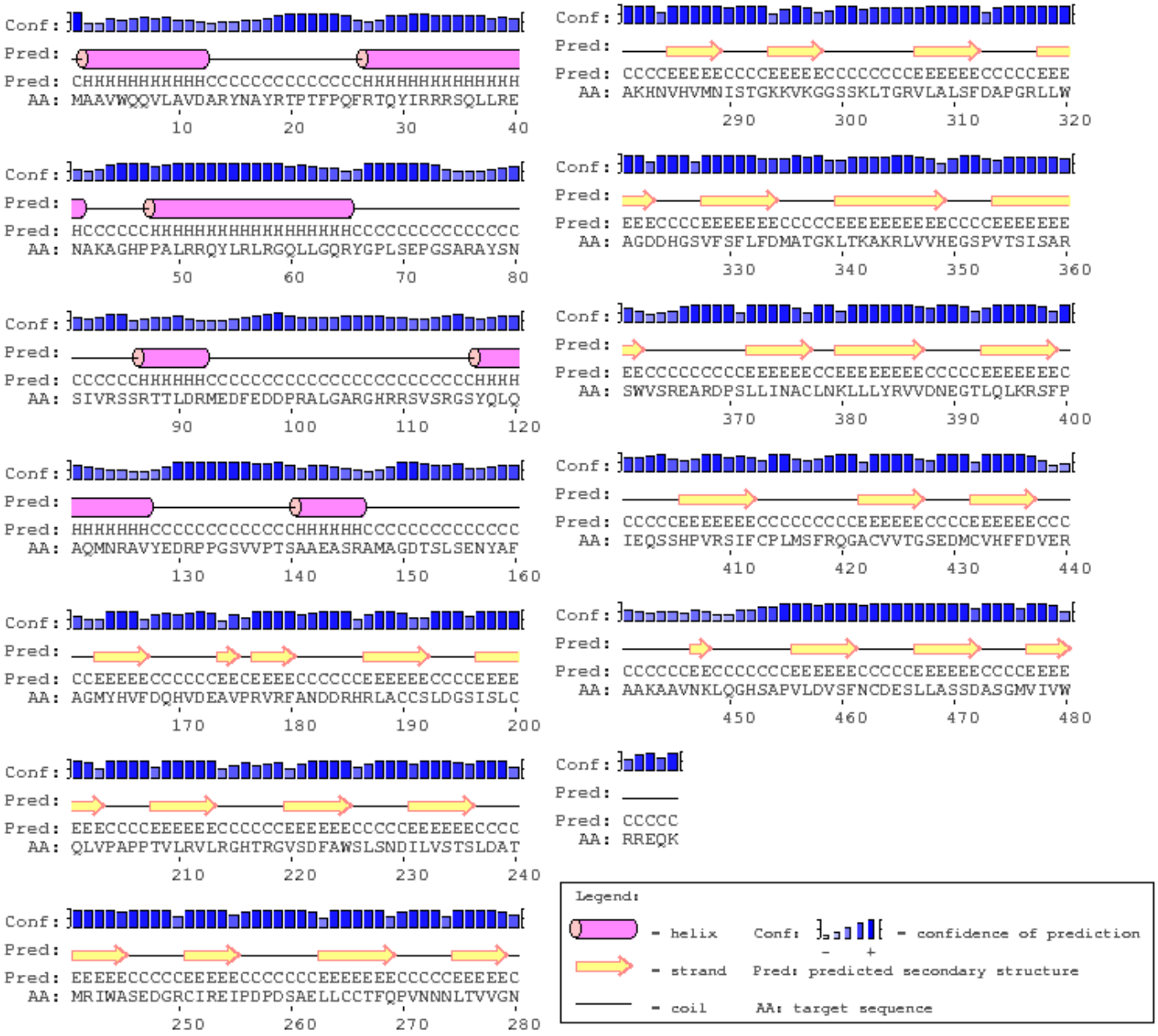



*Figure 3 PSIPRED Prediction of the Secondary Structure of WDR13*

## 4.1 STRUCTURE PREDICTION

We use this prediction as our base to model the protein. The raw sequence was given to Protein Model Portal. This portal gives the sequence to the following servers:

- M4T
- Phyre2
- I-TASSER
- ModWeb
- Swiss-Model

### 4.1.1 HOMOLOGY MODELLING

M4T, ModWeb and Swiss-Model use homology modelling methods, I-TASSER uses threading methods and Phyre2 uses Remote Homology Modelling Technique. All of them use different algorithms to model the given protein.

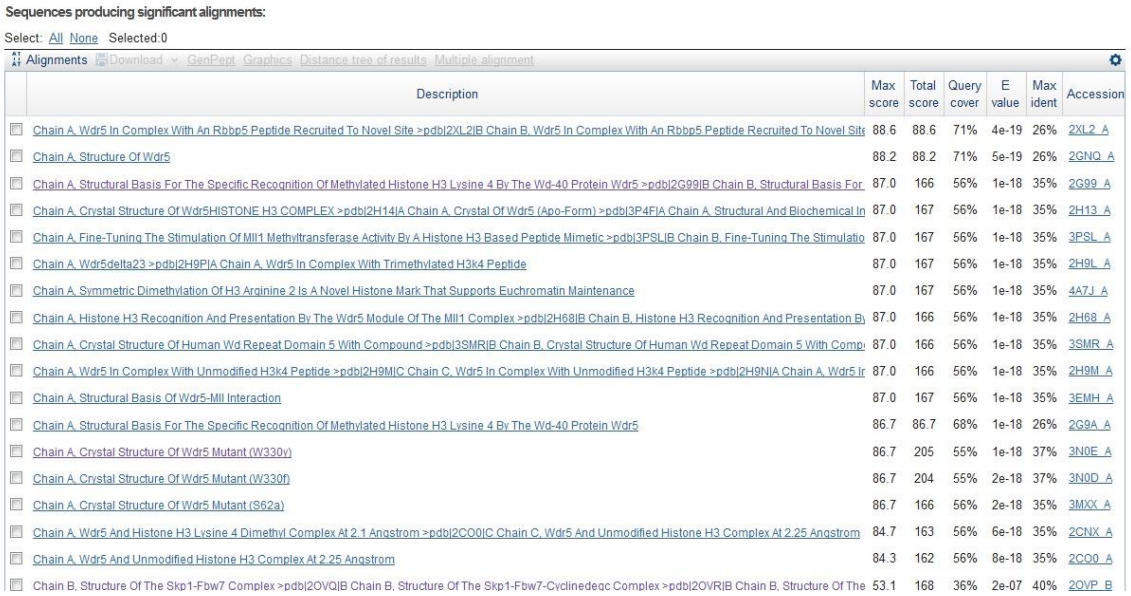
Of the five tools used, only two – M4T and Phyre2 predicted a whole β-propeller structure. While the others did predict strands, they could not form the β-propeller and hence were discarded. Between M4T and Phyre2, M4T model gave a better Ramachandran plot statistic and it was chosen over Phyre2. However, even the M4T and Phyre2 models were not completely formed. They had two flaws:

1. They did not cover the entire sequence of WDR13; they tended to miss the first 140-150 amino acids.
2. One of the blades of the propeller was incorrectly formed to have only 3 strands instead of the usual 4. The missing strand corresponded to the beta sheet in the 340-350 region according to PSIPRED's prediction.

### 4.1.2 FOLD RECOGNITION (THREADING)

To overcome the first difficulty, we gave the first 160 amino acids alone to I-TASSER for automatic fold recognition. Threading method was chosen over homology modelling because no homologue for the N-Terminal of WDR13 was found. The closest homologue to

WDR13 is WDR5 which aligns from the 137th residue onwards. WDR5 has a maximum identity of 26% with WDR13, which by itself isn't particularly high, but can be used as a template to predict the structure. I-TASSER gave an extremely good model for WDR13's N-Terminal which matched extremely closely with PSIPRED's predictions. Hence, we will use I-TASSER's model for N-Terminal and M4T's model for C-Terminal to form the complete protein using MODELLER.



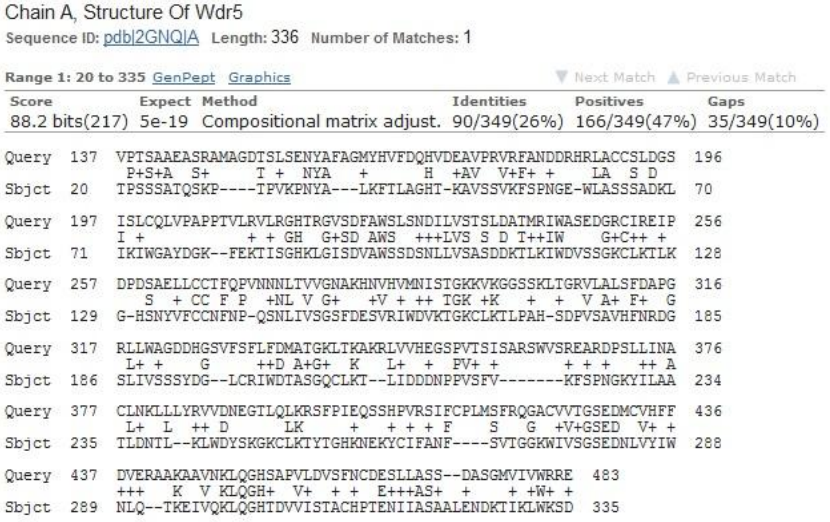
**Sequences producing significant alignments:**

Select: All None Selected:0

Alignments | Download | GenPept Graphics Distance tree of results Multiple alignment

| Description | Max score | Total score | Query cover | E value | Max ident | Accession |
|---|---|---|---|---|---|---|
| Chain A, Wdr5 In Complex With An Rbbp5 Peptide Recruited To Novel Site >pdb\|2XL2\|B Chain B, Wdr5 In Complex With An Rbbp5 Peptide Recruited To Novel Site | 88.6 | 88.6 | 71% | 4e-19 | 26% | 2XL2_A |
| Chain A, Structure Of Wdr5 | 88.2 | 88.2 | 71% | 5e-19 | 26% | 2GNQ_A |
| Chain A, Structural Basis For The Specific Recognition Of Methylated Histone H3 Lysine 4 By The Wd-40 Protein Wdr5 >pdb\|2G99\|B Chain B, Structural Basis For | 87.0 | 166 | 56% | 1e-18 | 35% | 2G99_A |
| Chain A, Crystal Structure Of Wdr5HISTONE H3 COMPLEX >pdb\|2H14\|A Chain A, Crystal Of Wdr5 (Apo-Form) >pdb\|3P4F\|A Chain A, Structural And Biochemical In | 87.0 | 167 | 56% | 1e-18 | 35% | 2H13_A |
| Chain A, Fine-Tuning The Stimulation Of Mll1 Methyltransferase Activity By A Histone H3 Based Peptide Mimetic >pdb\|3PSL\|B Chain B, Fine-Tuning The Stimulatio | 87.0 | 167 | 56% | 1e-18 | 35% | 3PSL_A |
| Chain A, Wdr5delta23 >pdb\|2H9P\|A Chain A, Wdr5 In Complex With Trimethylated H3k4 Peptide | 87.0 | 167 | 56% | 1e-18 | 35% | 2H9L_A |
| Chain A, Symmetric Dimethylation Of H3 Arginine 2 Is A Novel Histone Mark That Supports Euchromatin Maintenance | 87.0 | 167 | 56% | 1e-18 | 35% | 4A7J_A |
| Chain A, Histone H3 Recognition And Presentation By The Wdr5 Module Of The Mll1 Complex >pdb\|2H68\|B Chain B, Histone H3 Recognition And Presentation By | 87.0 | 166 | 56% | 1e-18 | 35% | 2H68_A |
| Chain A, Crystal Structure Of Human Wd Repeat Domain 5 With Compound >pdb\|3SMR\|B Chain B, Crystal Structure Of Human Wd Repeat Domain 5 With Compo | 87.0 | 166 | 56% | 1e-18 | 35% | 3SMR_A |
| Chain A, Wdr5 In Complex With Unmodified H3k4 Peptide >pdb\|2H9M\|C Chain C, Wdr5 In Complex With Unmodified H3k4 Peptide >pdb\|2H9N\|A Chain A, Wdr5 In | 87.0 | 166 | 56% | 1e-18 | 35% | 2H9M_A |
| Chain A, Structural Basis Of Wdr5-Mll Interaction | 87.0 | 167 | 56% | 1e-18 | 35% | 3EMH_A |
| Chain A, Structural Basis For The Specific Recognition Of Methylated Histone H3 Lysine 4 By The Wd-40 Protein Wdr5 | 86.7 | 86.7 | 68% | 1e-18 | 26% | 2G9A_A |
| Chain A, Crystal Structure Of Wdr5 Mutant (W330y) | 86.7 | 205 | 55% | 1e-18 | 37% | 3N0E_A |
| Chain A, Crystal Structure Of Wdr5 Mutant (W330f) | 86.7 | 204 | 55% | 2e-18 | 37% | 3N0D_A |
| Chain A, Crystal Structure Of Wdr5 Mutant (S62a) | 86.7 | 166 | 56% | 2e-18 | 35% | 3MXX_A |
| Chain A, Wdr5 And Histone H3 Lysine 4 Dimethyl Complex At 2.1 Angstrom >pdb\|2CO0\|C Chain C, Wdr5 And Unmodified Histone H3 Complex At 2.25 Angstrom | 84.7 | 163 | 56% | 6e-18 | 35% | 2CNX_A |
| Chain A, Wdr5 And Unmodified Histone H3 Complex At 2.25 Angstrom | 84.3 | 162 | 56% | 8e-18 | 35% | 2CO0_A |
| Chain B, Structure Of The Skp1-Fbw7 Complex >pdb\|2OVQ\|B Chain B, Structure Of The Skp1-Fbw7-Cyclinedegc Complex >pdb\|2OVR\|B Chain B, Structure Of The | 53.1 | 168 | 36% | 2e-07 | 40% | 2OVP_B |

*Figure 4 BLAST results against PDB for WDR13*



Chain A, Structure Of Wdr5

**Sequence ID:** pdb|2GNQ|A **Length:** 336 **Number of Matches:** 1

**Range 1: 20 to 335** GenPept Graphics ▼ Next Match ▲ Previous Match

| Score | Expect | Method | Identities | Positives | Gaps |
|---|---|---|---|---|---|
| 88.2 bits(217) | 5e-19 | Compositional matrix adjust. | 90/349(26%) | 166/349(47%) | 35/349(10%) |

```
Query  137  VPTSAAEASRAMAGDTISLSENYAFAGMYHVFDQHVDEAVPRVRFANDDRHRLACCSLDGS  196
            P+S+A  S+     T +  NYA   +      H  +AV  V+F+ +    LA  S D
Sbjct  20   TPSSSATQSKP----TPVKPNYA---LKFTLAGHT-KAVSSVKFSPNGE-WLASSSADKL  70

Query  197  ISLCQLVPAPPTVLRVLRGHTRGVSDFAWSLSNDILVSTSLDATMRIWASEDGRCIREIP  256
            I +           + + GH  G+SD AWS  +++LVS S D T++IW    G+C++ +
Sbjct  71   IKIWGAYDGK--FEKTISGHKLGISDVAWSSDSNLLVSASDDKTLKIWDVSSGKCLKTLK  128

Query  257  DPDSAELLCCTFQPVNNNLTVVGNAKHNVHVMNISTGKKVKGGSSKLTGRVLALSFDAPG  316
               S  + CC F P  +NL V G+   +V + ++ TGK +K   +  +  V A+ F+  G
Sbjct  129  G-HSNYVFCCNFNP-QSNLIVSGSFDESVRIWDVKTGKCLKTLPAH-SDPVSAVHFNRDG  185

Query  317  RLLWAGDDHGSVFSFLFDMATGKLTKAKRLVVHEGSPVTSISARSWVSREARDPSLLINA  376
             L+ +    G     ++D A+G+  K   L+  +  PV+ +       + + +   ++ A
Sbjct  186  SLIVSSSYDG--LCRIWDTASGQCLKT--LIDDDNPPVSFV-------KFSPNGKYILAA  234

Query  377  CLNKLLLYRVVDNEGTLQLKRSFPIEQSSHPVRSIFCPLMSFRQGACVVTGSEDMCVHFF  436
             L+  L  ++ D      LK     +   + + + F    S   G  +V+GSED  V+ +
Sbjct  235  TLDNTL--KLWDYSKGKCLKTYTGHKNEKYCIFANF----SVTGGKWIVSGSEDNLVYIW  288

Query  437  DVERAAKAAVNKLQGHSAPVLDVSFNCDESLLASS--DASGMVIVWRRE  483
            +++   K  V KLQGH+  V+  + +  E+++AS+  +     + +W+ +
Sbjct  289  NLQ--TKEIVQKLQGHTDVVISTACHPTENIIASAALENDKTIKLWKSD  335
```

*Figure 5 Alignment of WDR13 with WDR5*

Once both the C-terminal and N-terminal ends of the protein were modelled separately, we used them both as templates to prepare the final structure. This was again performed using MODELLER.

### 4.2 ENERGY MINIMIZATION AND MOLECULAR DYNAMICS

This refined structure was then given to GROMACS for Energy Minimization and Molecular Dynamics. The protocol followed for energy minimization and dynamics was taken from Bevan Lab tutorial authored by Justin Lemkul listed at:
http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/lysozyme/index.html

GROMACS uses a pre-defined force field for its energy minimization and molecular dynamics calculations. A force field refers to the form and parameters of mathematical functions used to describe the potential energy of a system of particles. We used the OPLS-AA/L all-atom force field.

NVT and NPT equilibration was done for 1 ns each. This was followed by an MD run of 1 ns. In its analysis, we found the resultant structure to be very stable at the given temperature, pressure and density. Even the RMSD values are fairly steady.

Post analysis, we chose the structure with the lowest energy in the 1 ns timeframe for the next part of the project.

### 4.3 DATASET COLLECTION

As part of the project also involves predicting interacting partners, we simultaneously started on building a dataset for interacting partners on four criteria:

- Literature Survey
- Colocalization
- Coexpression

- Database Predictions: STRING and GeneMania

The dataset currently has the following proteins:

| **Literature Survey** | **Colocalization** | **Coexpression** | **Database Predictions** |
|---|---|---|---|
| Estrogen Receptors: ER-alpha, ER-beta<br><br>Histone Deacetylase: HDAC1, HDAC3, HDAC7<br><br>PHIP1: Pleckstrin Domain Homology Interacting Protein | Calcineurin: PPP3CA, PPP3CB, PPP3CC<br><br>Serine-Threonine Protein Phosphatase 2A<br><br>Hippocalcin<br><br>Astrotactin<br><br>AKAP5: A kinase (PRKA) anchor protein 5 | p21 (CDKN1a) | Transforming Growth Factor β (TGFβ)<br><br>TGFβ Receptor 1 (TGFβR1)<br><br>NF-κB: p50 and p65<br><br>Jun<br><br>FtsJ1<br><br>WDR45<br><br>FKBP1A<br><br>PIM1<br><br>CCND3 |

*Table 1 Dataset of Potential Interacting Partners of WDR13*

Depending on the availability of the structures of these molecules, we further narrowed down our dataset to the following proteins: Estrogen Receptors ER-alpha and ER-beta, Histone Deacetylase HDAC3 and HDAC7, Calcineurin (PPP3CA) also known as Protein Phosphatase 3 catalytic subunit alpha, Transforming Growth Factor β Receptor 1 (TGFβR1),

Nuclear Factor Kappa-light-chain-enhancer of Activated B cells (NF-κB), Peptidyl-prolyl cis-trans isomerase (FKBP1A), Proto-oncogene serine/threonine-protein kinase (PIM1) and Cyclin D3 (CCND3).

The structures of these molecules were taken from PDB. They were edited to remove ligands and other molecules that were complexed to the proteins.

### 4.4 PHYSICAL DOCKING AND POST DOCKING ANALYSIS

Docking studies were performed between WDR13 and these proteins in 5 separate servers that all use different algorithms. The servers used were ClusPro[9], GrammX[10], HexDock[11], PatchDock[12] and ZDock[13]. All the servers used blind docking methods, as specific regions of interaction were not known. The best prediction from each of the servers was analyzed for their protein-protein interface and compared to find similarity in structure. The protein-protein interface was analyzed using PyMOL and DIMPLOT. Key residues involved in hydrogen bonding and hydrophobic interactions were identified.

Online server 2P2IDB[14] was also used for classification of interacting region type of both the receptor and the ligand into alpha, beta, coil or alpha-beta.

[9] http://cluspro.bu.edu/
[10] http://vakser.bioinformatics.ku.edu/resources/gramm/grammx/
[11] http://hex.loria.fr/
[12] http://bioinfo3d.cs.tau.ac.il/PatchDock/
[13] http://zdock.umassmed.edu/
[14] http://2p2idb.cnrs-mrs.fr/2p2i_inspector.html

# 5. RESULTS AND DISCUSSION

## 5.1 WDR13 STRUCTURE

The models that we got from M4T and I-TASSER are shown below. M4T model, as already mentioned, depicts only the C-Terminal whereas the I-TASSER model shows the N-Terminal.

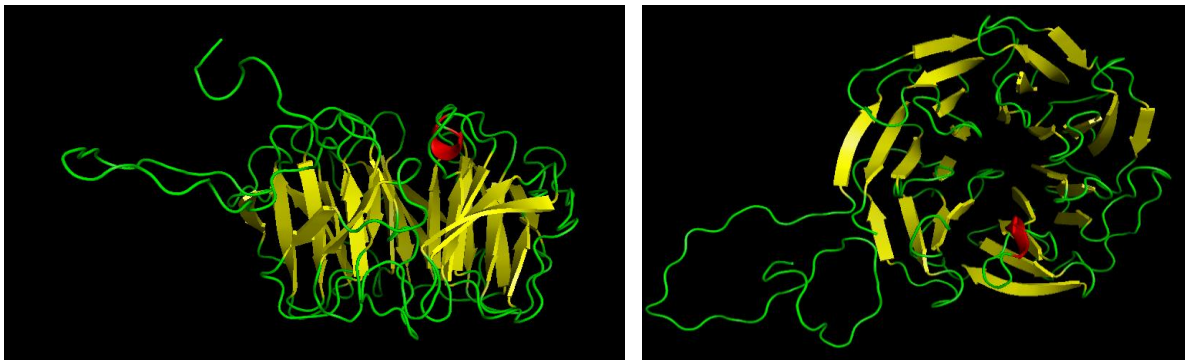


*Figure 6: The Characteristic Beta Propeller of WDR13. From the side, it looks like a funnel structure which might be involved in protein complex formation*

The beta-propeller forms a characteristic funnel like structure from the side and a propeller from its top view.

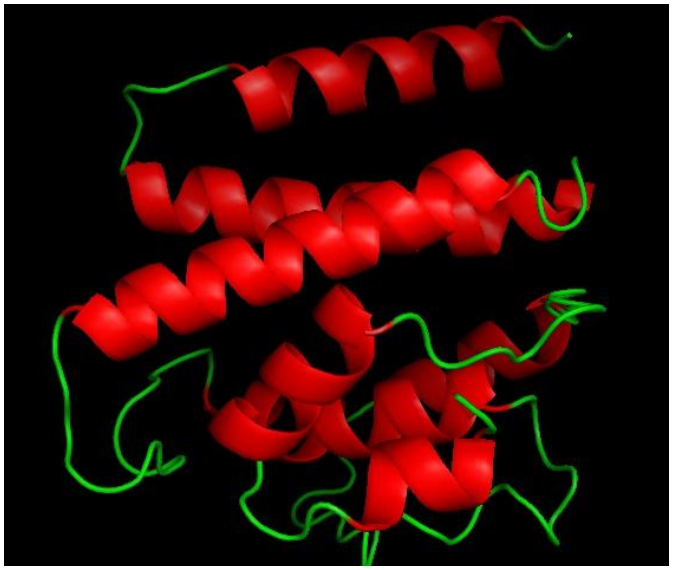


*Figure 7 I-TASSER Prediction for the N-Terminal of WDR13*

The figure at the side shows the N-Terminal alone which was separately modelled using I-TASSER. The alpha helices formed in this structure of particular importance because we suspect some of the specific functions of WDR13 reside in this region.

Using these two models as our templates, we aligned the target sequence manually (MODELLER's salign.py tried to align the entire sequence resulting in gaps at unnecessary places). Through multiple iterations of loop refinement, the model that we currently have

consists of both the N-Terminal and C-Terminal with over 85% residues in the favored regions. The model is shown below:

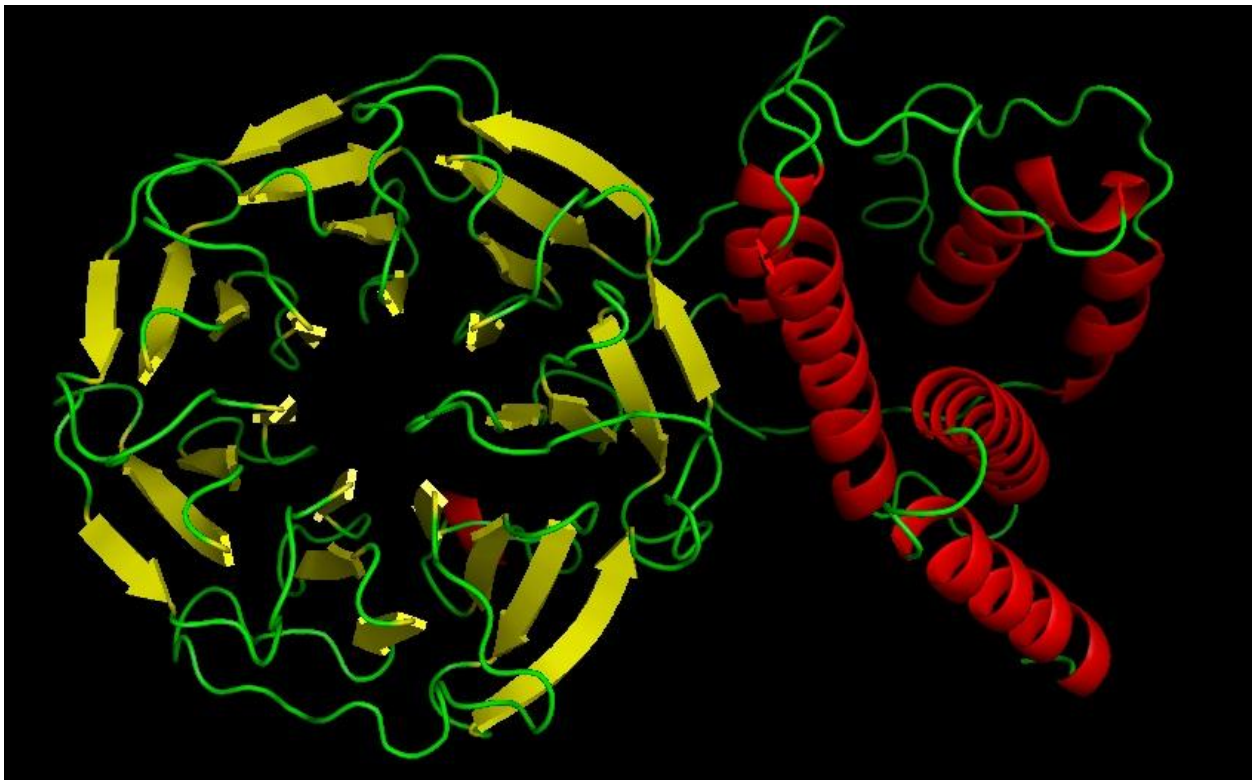


*Figure 8 The complete model for WDR13 using I-TASSER's output and MODELLER's output as templates*

## 5.2 STRUCTURE REFINEMENT

The structure obtained from MODELLER was refined iteratively, till most of its residues were in the allowed region in the Ramachandran Plot. We had close to 87.9% of the residues in the most favored regions, 11.4% in the additionally allowed regions and 0.7% in the generously allowed regions.

We also reduced the number of labelled residues to just 11, out of which 9 were Glycine and Proline. The Verify3D score was 82.10% and the model passed the test.

## Plot statistics

| | | |
|---|---|---|
| Residues in most favoured regions [A,B,L] | 377 | 87.9% |
| Residues in additional allowed regions [a,b,l,p] | 49 | 11.4% |
| Residues in generously allowed regions [~a,~b,~l,~p] | 3 | 0.7% |
| Residues in disallowed regions | 0 | 0.0% |
| Number of non-glycine and non-proline residues | 429 | 100.0% |
| Number of end-residues (excl. Gly and Pro) | 2 | |
| Number of glycine residues (shown as triangles) | 30 | |
| Number of proline residues | 24 | |
| Total number of residues | 485 | |

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

*Figure 9 Ramachandran Plot Characteristics after Loop Refinement*

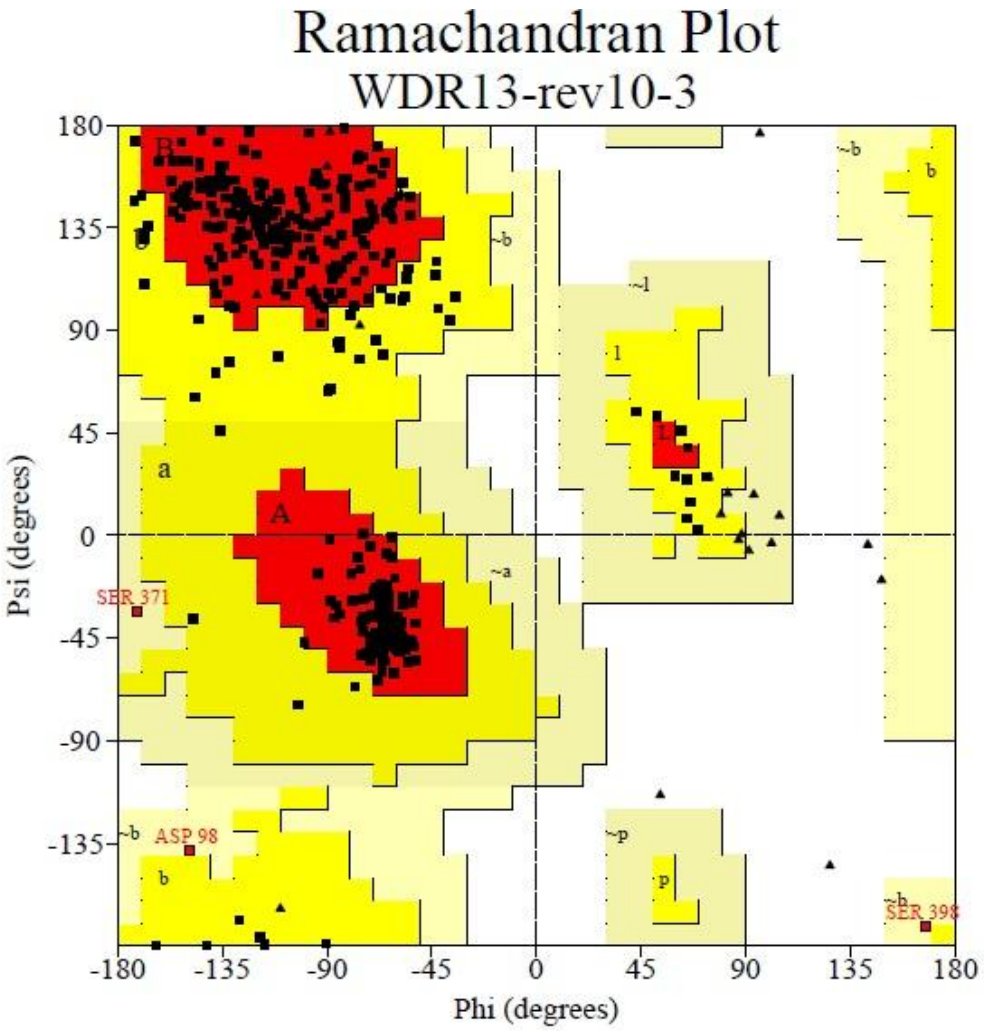



*Figure 10 Ramachandran Plot for WDR13 after Loop Refinement*

## 5.3 ENERGY MINIMIZATION AND MOLECULAR DYNAMICS

This structure was given to GROMACS for energy minimization. The energy was reduced as shown in the graph.

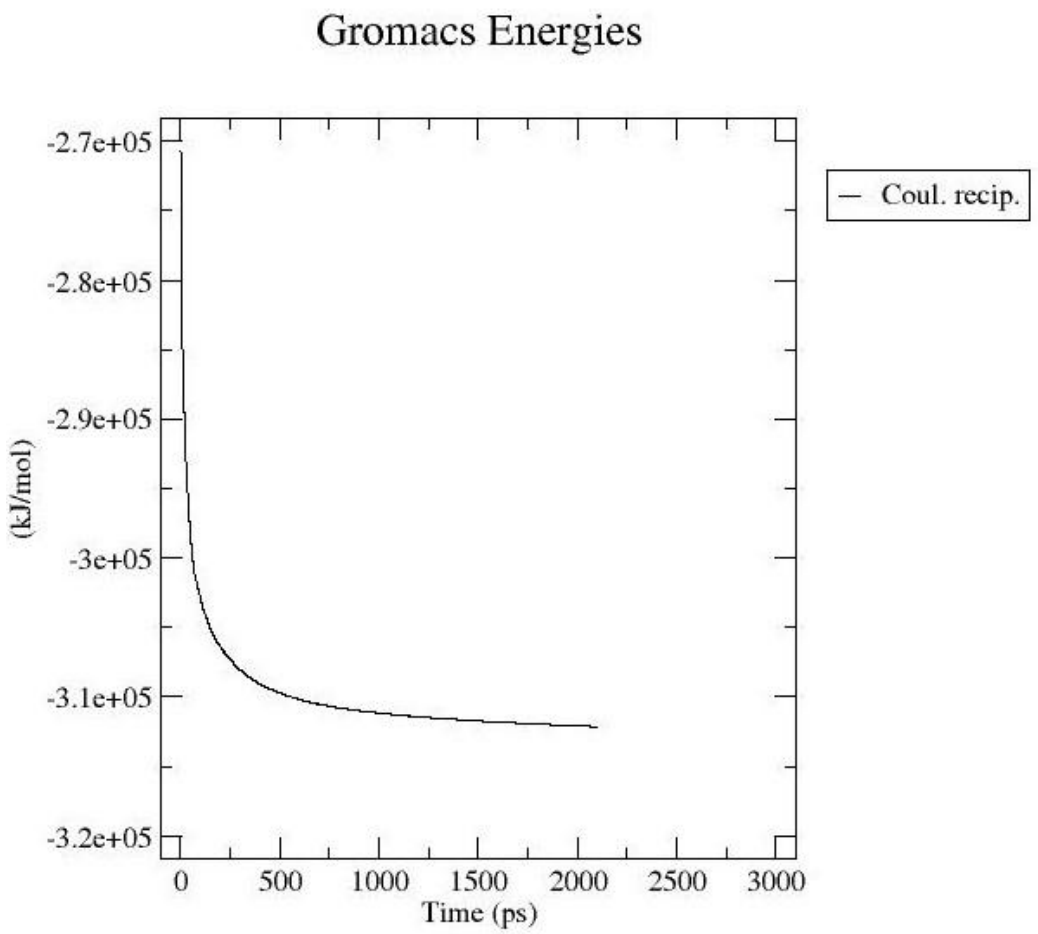



*Figure 11 Energy Minimization of WDR13 using GROMACS*

The structure was very stable over time. This was indicated by fairly stable plots of temperature, pressure and density shown below.

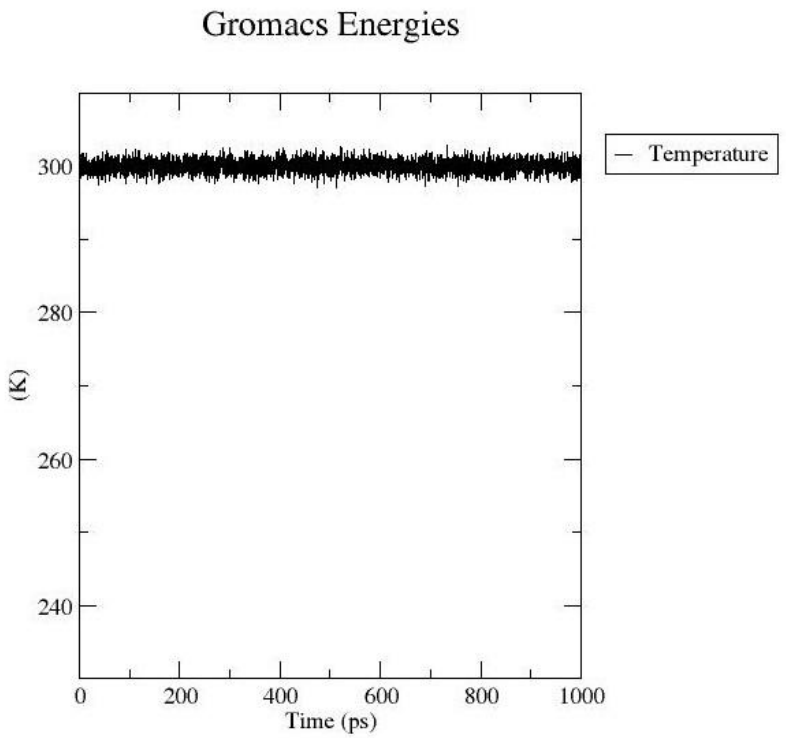



*Figure 12 Temperature Stability of WDR13 over a simulation duration of 1 ns*

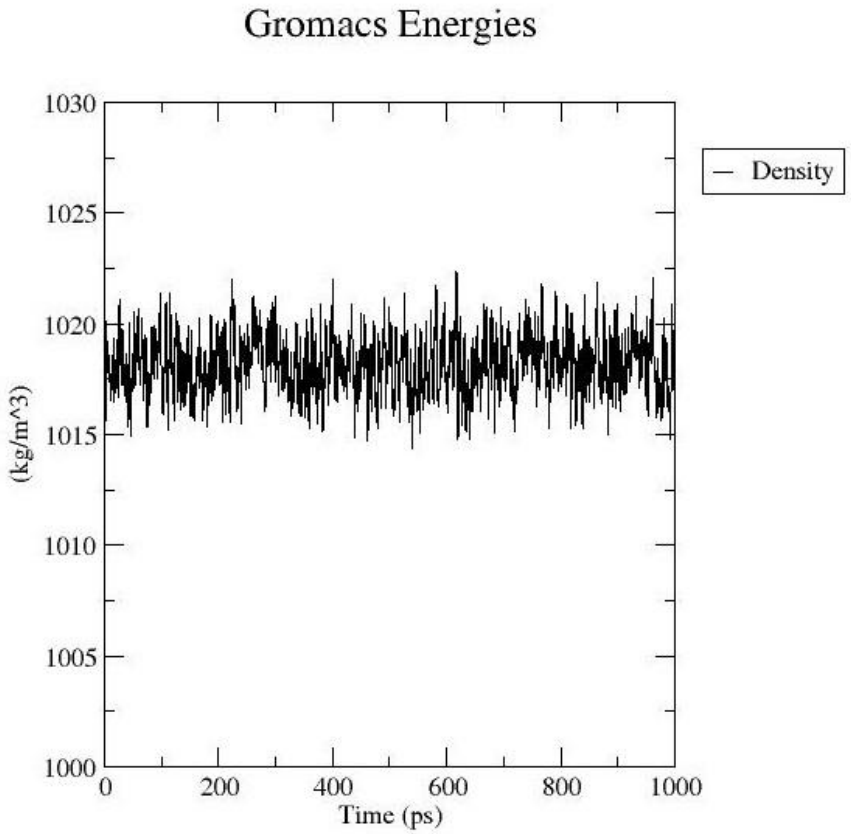



*Figure 13 Density Stability of WDR13 over a simulation duration of 1 ns*

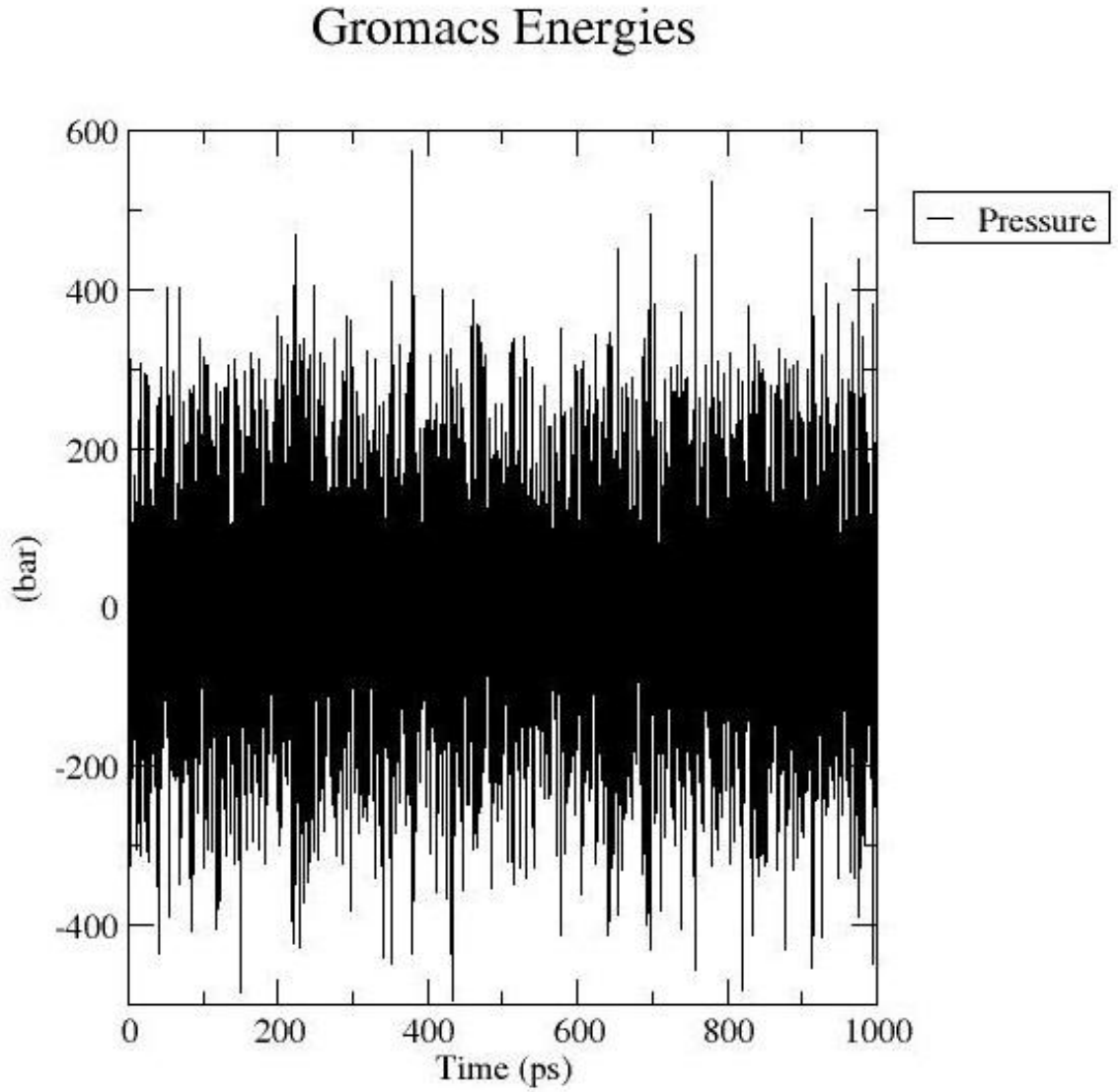



*Figure 14 Pressure Graph of WDR13 over the simulation duration of 1 ns.*

Even though the pressure seems to fluctuate, this behavior is considered normal for most proteins.
(http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/lysozyme/07_equil2.html)

We also performed a molecular dynamics simulation to test the stability of the structure over time. The RMSD plot is also given below. Also plotted is the potential energy during the course of the simulation.

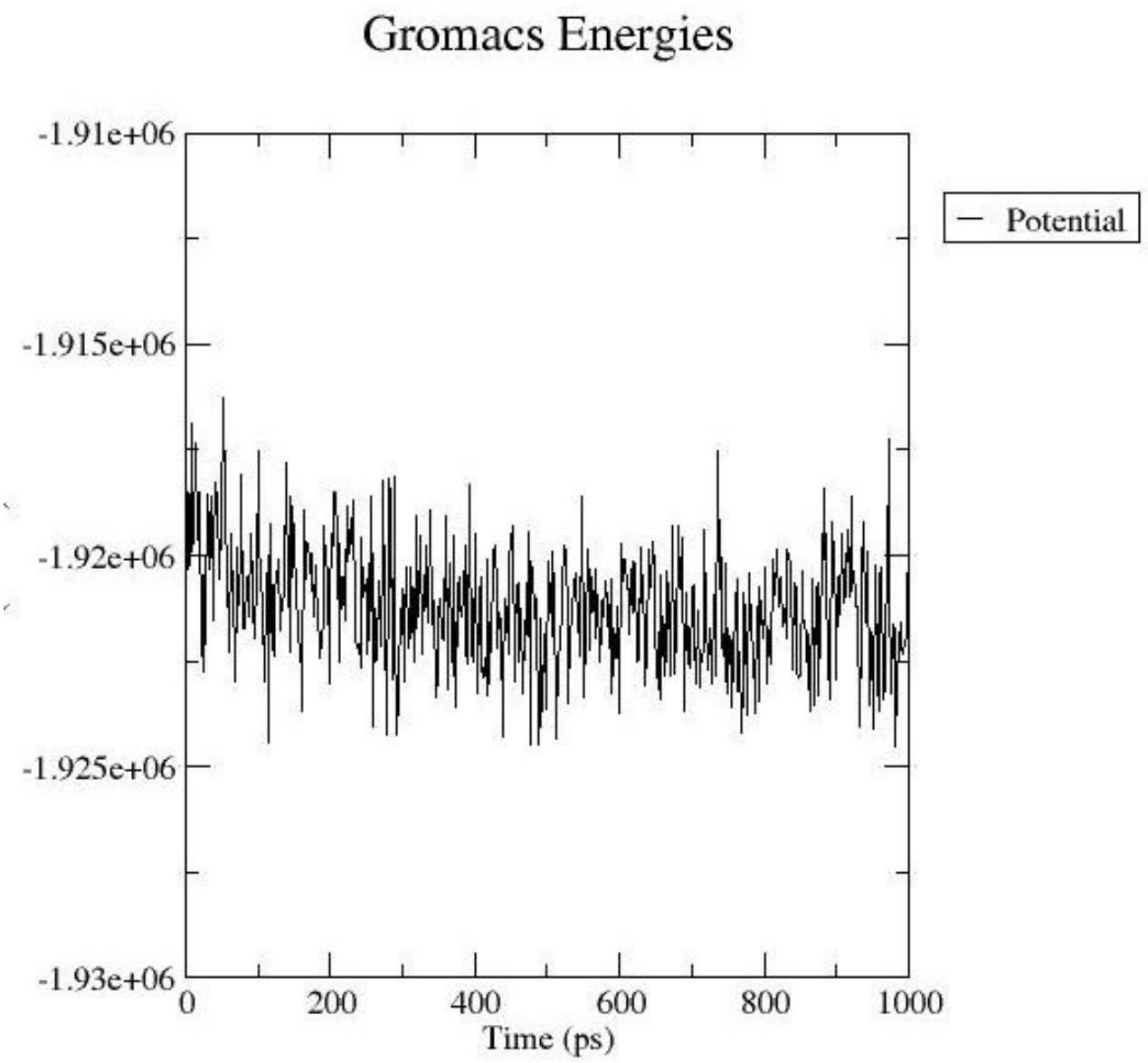



*Figure 15 The Potential Energy Plot of WDR13 during the course of simulation of 1 ns*

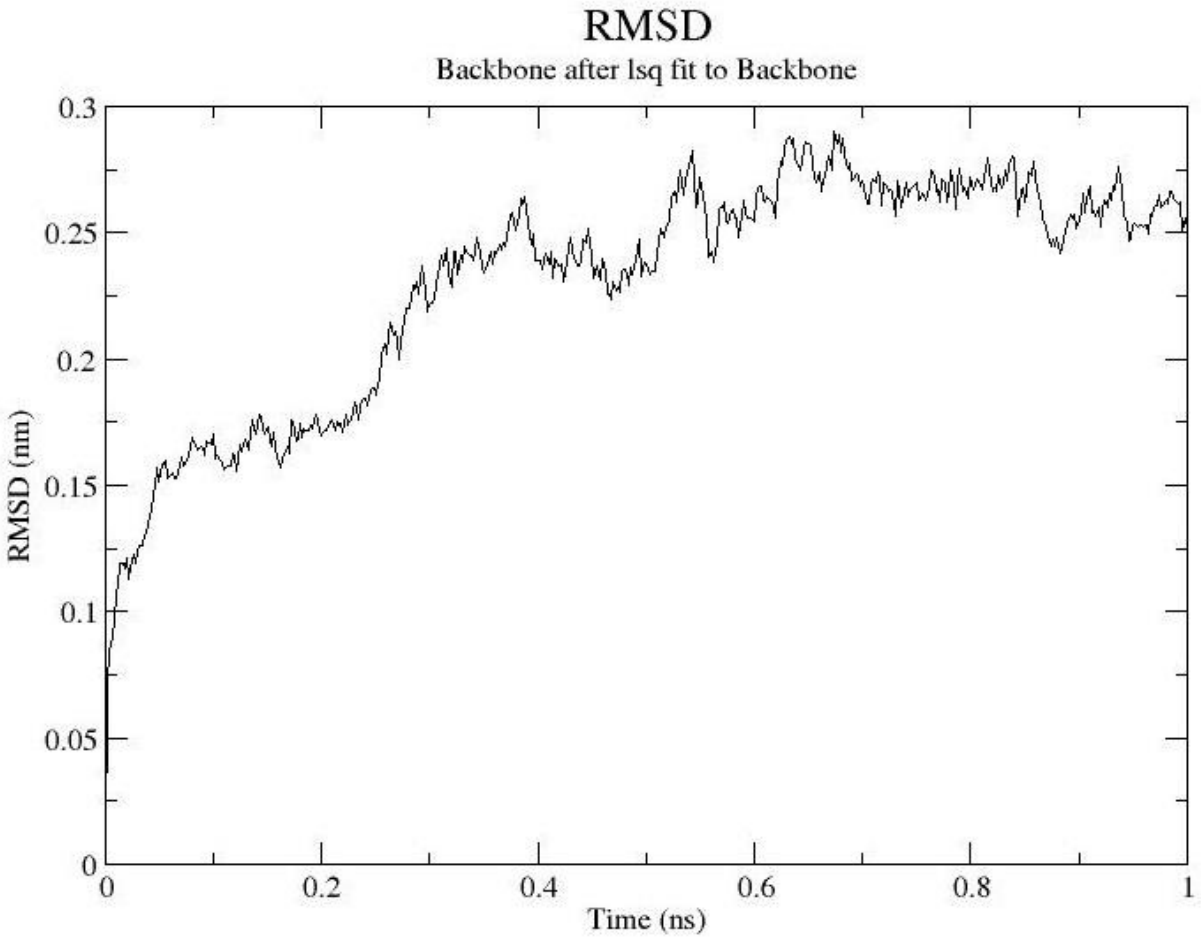



*Figure 16 The RMSD plot of WDR13 structure during the course of simulation of 1 ns*

As shown, the RMSD is always less than 0.3 nm (3 angstrom), which indicates a very stable structure. We chose the structure with the least energy from the Potential Energy plot and further used this in the docking studies.

### 5.4 PHYSICAL DOCKING AND POST DOCKING ANALYSIS

This structure was used as a receptor and submitted to 5 different servers: ClusPro, HexDock, GrammX, PatchDock and ZDock. Altogether 12 proteins were submitted along with WDR13 for docking. The best result from each server was taken and analyzed further. This resulted in the generation of 60 complexes. An example complex (of Estrogen Receptor α, a known interacting partner) is given below.

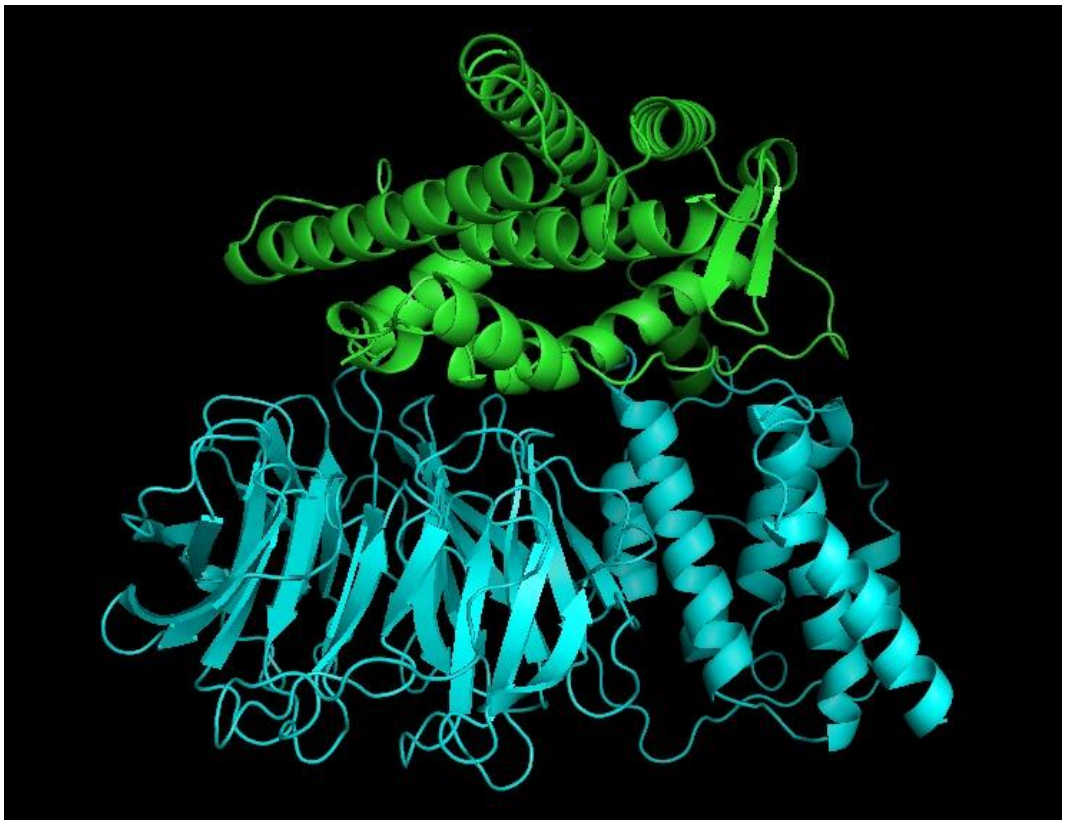


*Figure 17 Interaction of WDR13 with ER-alpha - Cartoon View*

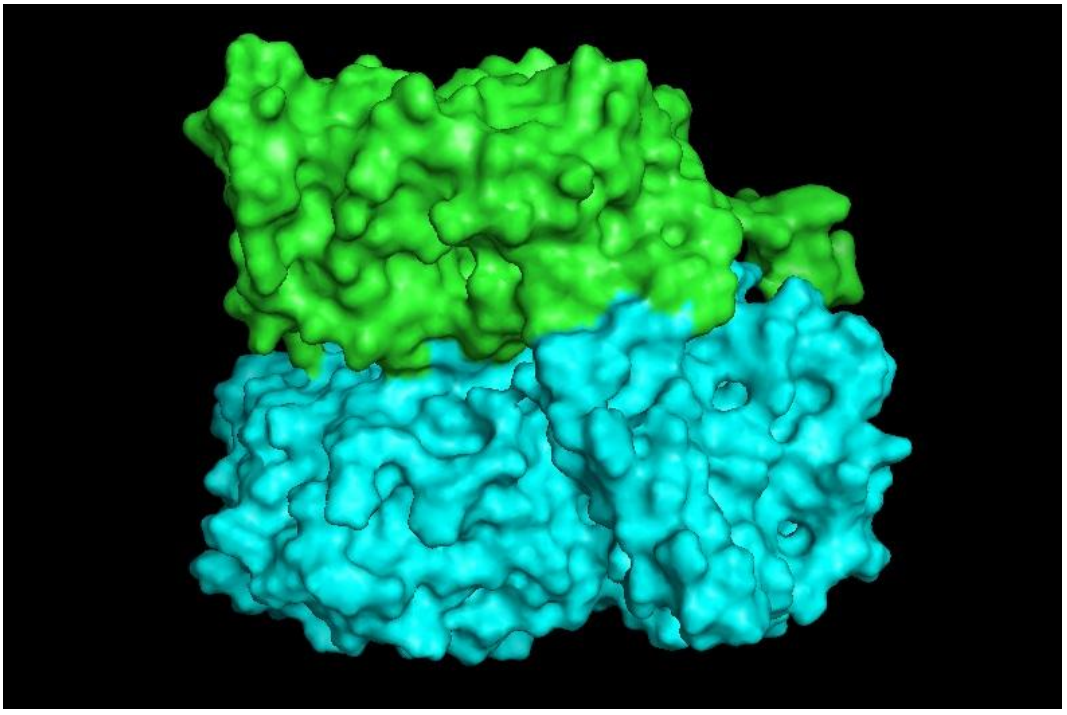
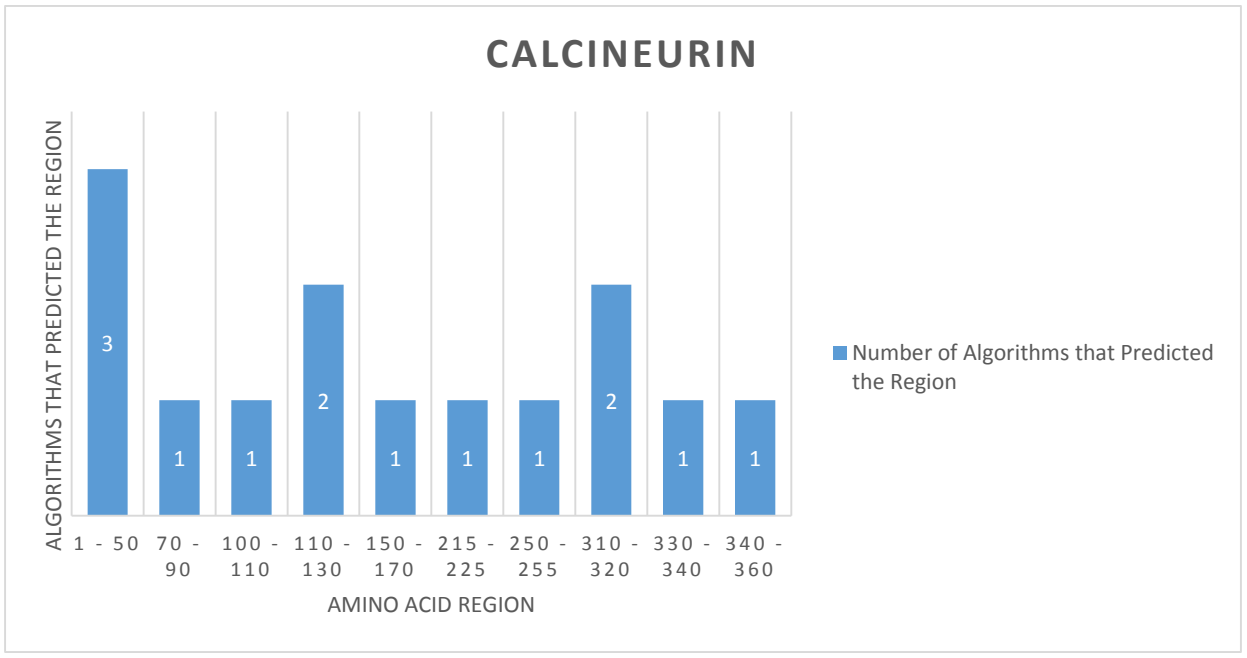
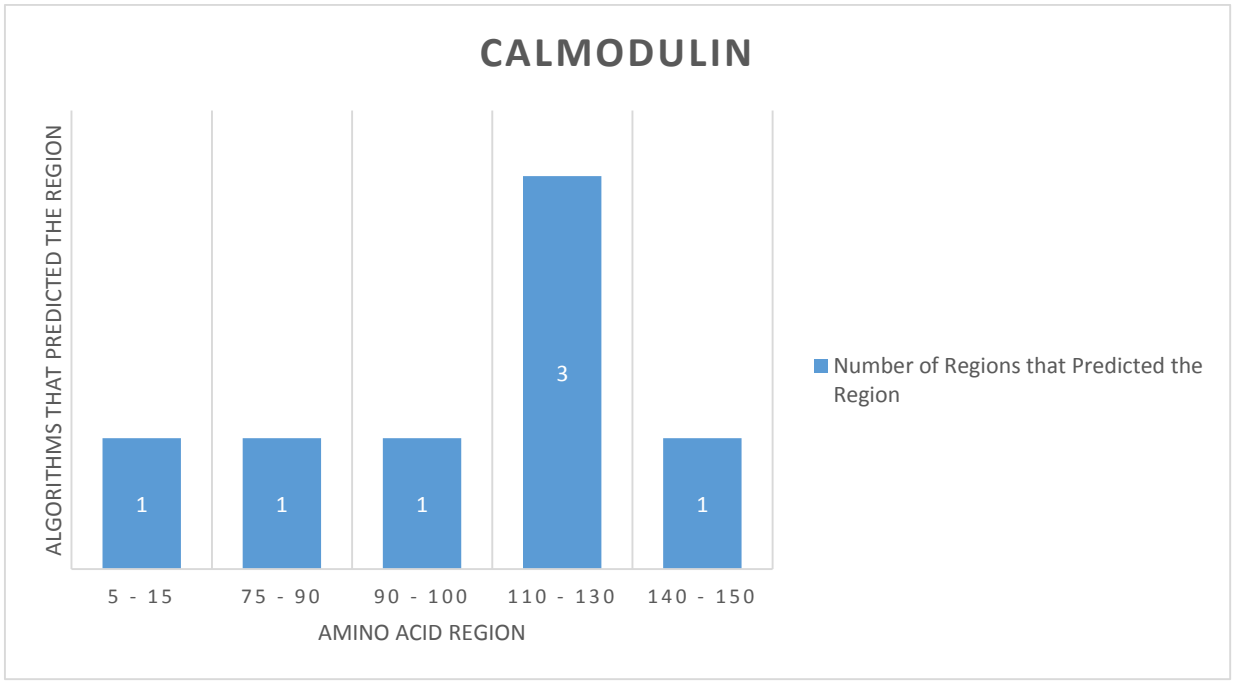


*Figure 18 Interaction of WDR13 with ER-alpha - Surface View*

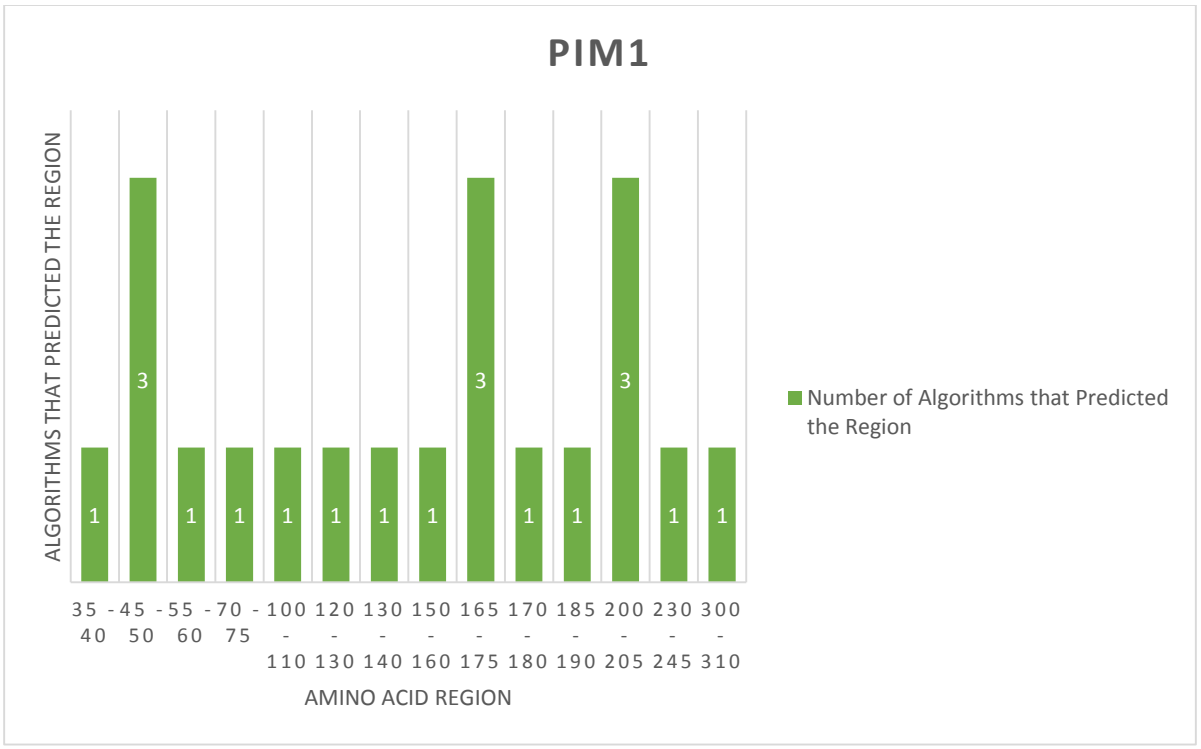
Each complex was analyzed using DIMPLOT for the interface. This was compared with the other 4 results of the same complex. Regions of similar binding were identified. The occurrence of each interacting region of the ligand is plotted in the following graphs



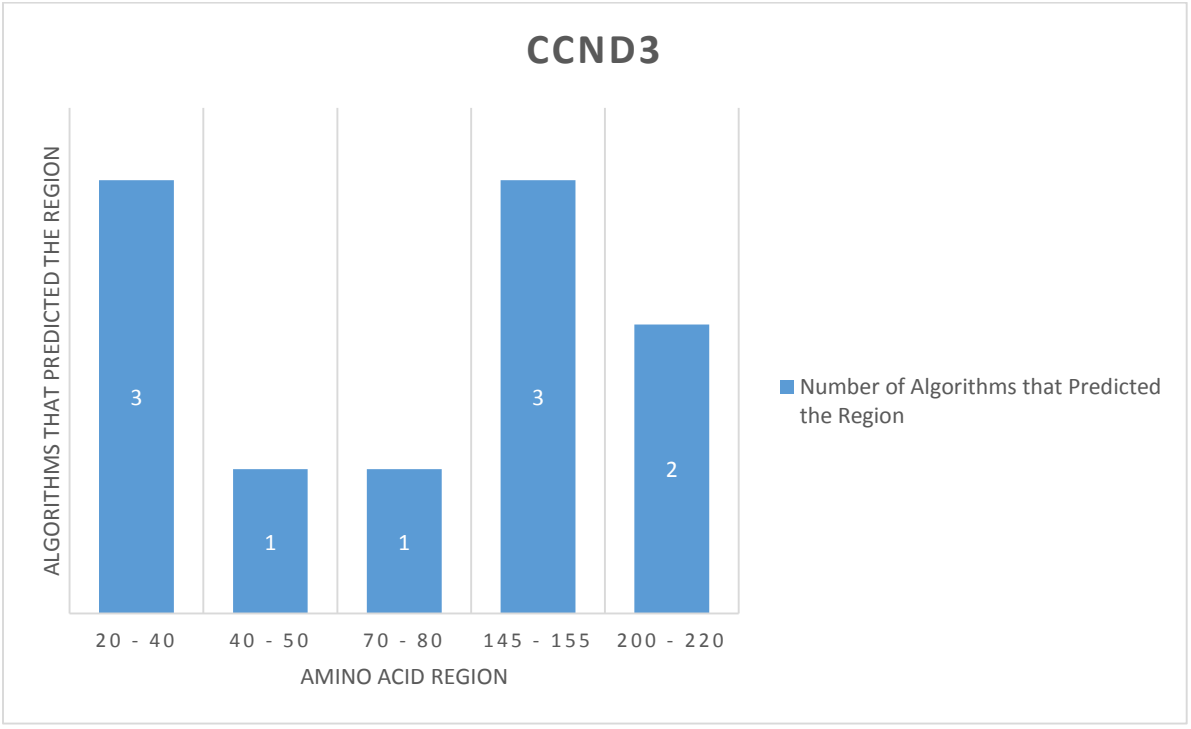

*Graph 1 Occurrence of Specific Regions in all complexes of Calcineurin*



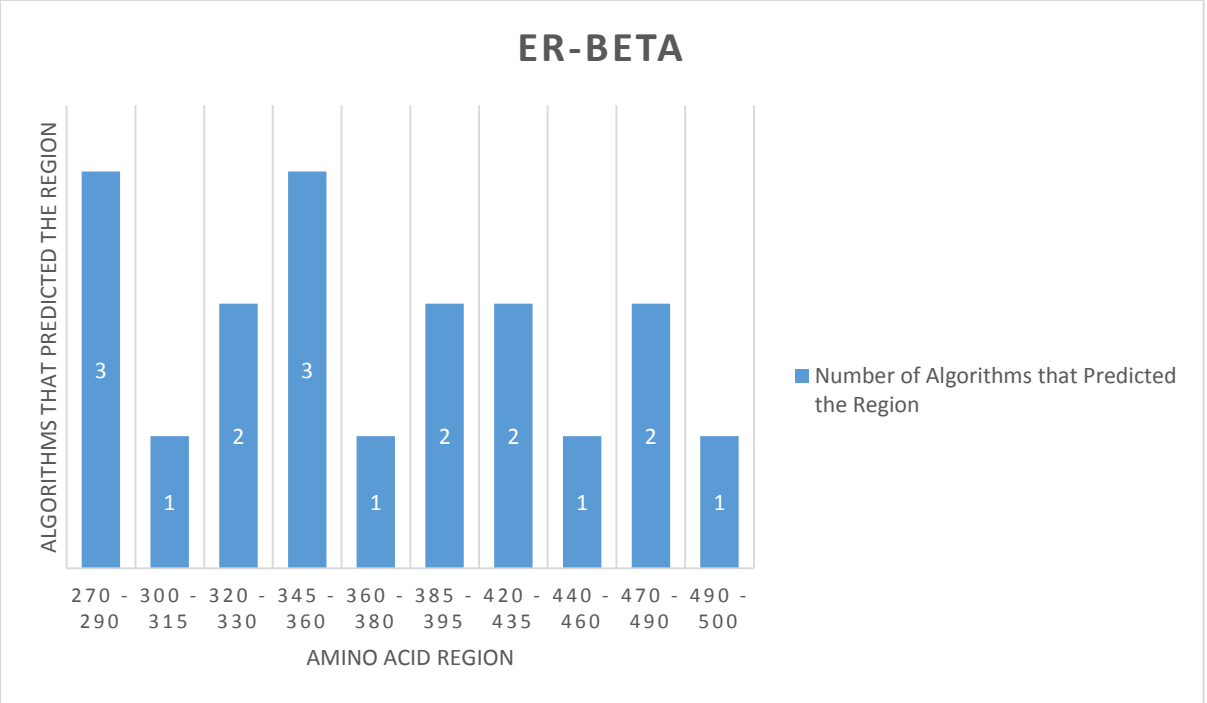

*Graph 2 Occurrence of Specific Regions in all complexes of Calmodulin*



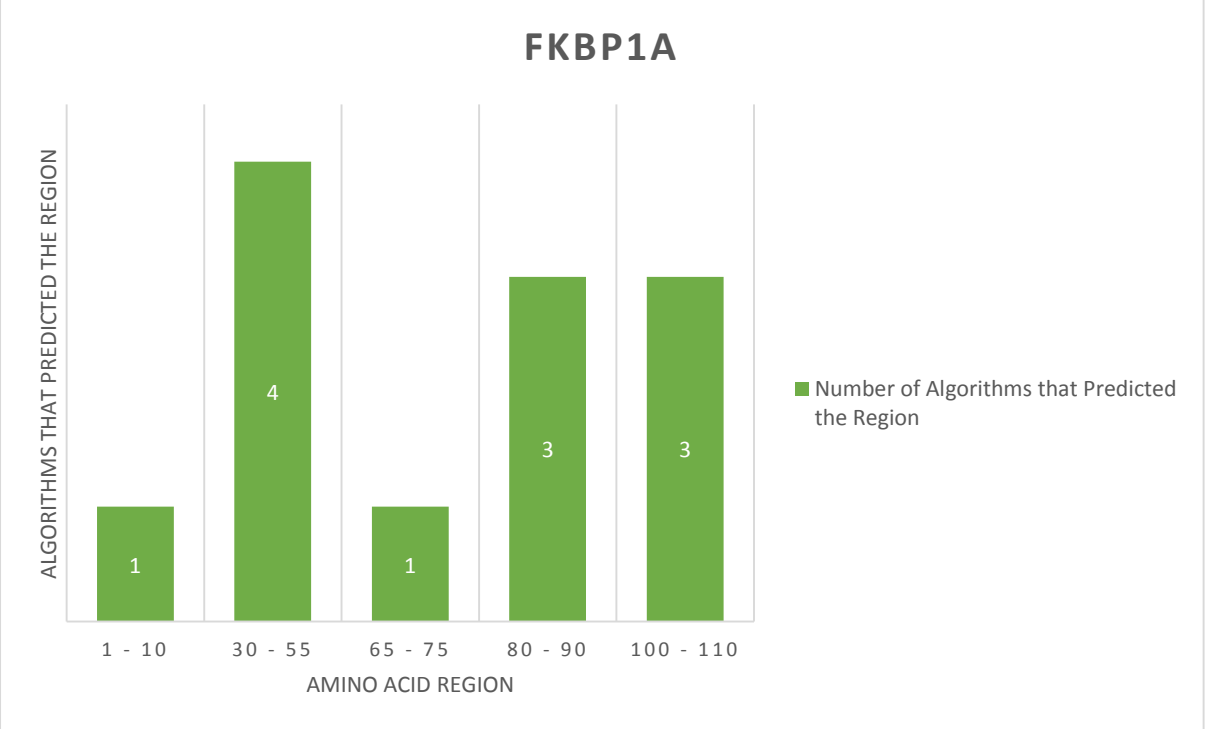

*Graph 3 Occurrence of Specific Regions in all complexes of PIM1*



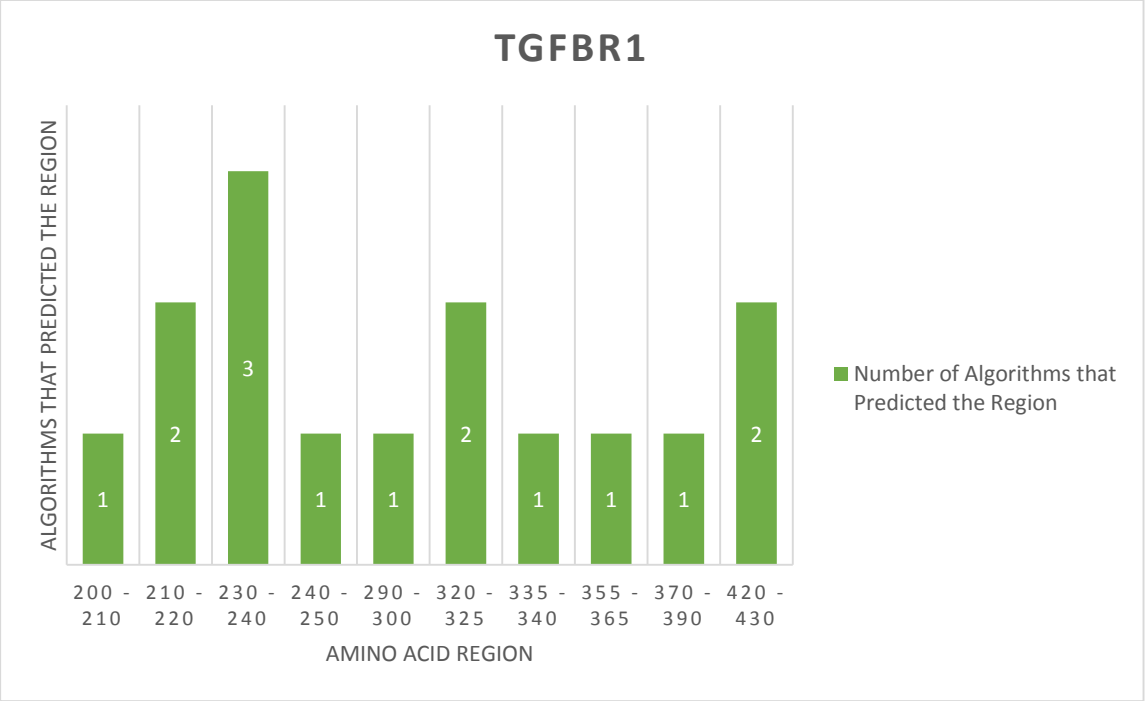

*Graph 4 Occurrence of Specific Regions in all complexes of CCND3*




*Graph 5 Occurrence of Specific Regions in all complexes of ER-Beta*




*Graph 6 Occurrence of Specific Regions in all complexes of FKBP1A*




*Graph 7 Occurrence of Specific Regions in all complexes of TGFβR1*

If a particular region is predicted as an interacting region by multiple algorithms, we take it to be a consensus region. By extension, if a protein complex has multiple such regions it, we consider it to be a true interacting partner of WDR13. The numbers in the bar represent how many algorithms predicted that particular region.

We found FKBP1A, PIM1 and TFGβR1 to have maximum consensus regions between multiple algorithms. In FKBP1A, region 30 to 55 was predicted by 4 out of 5 docking servers: ClusPro, GrammX, HexDock and ZDock. Furthermore, regions 80 to 90 and 100 to 110 were predicted by 3 out of 5 algorithms. In PIM1, regions 45 to 50, 165 to 175 and 200 to 205 were predicted by 3 algorithms. In TGFβR1, region 230 to 240 was predicted by 3 algorithms whereas regions 210 to 220, 320 to 325 and 420 to 430 were predicted by 2 algorithms each.

We also gave a negative dataset, proteins that definitely would not interact with WDR13 to all the docking servers. When we analyzed their DIMPLOT results, we got an even spread of interacting regions. Multiple algorithms did not predict the same regions as the interacting zone, thereby giving a preliminary validation to our method.

We also tried to validate our result by giving the protein complexes to 2P2IDB server. This server calculates protein-protein interface characteristics. We focused our analysis specifically on the secondary structure type at each interface for both the receptor and the ligand. Our results are tabulated below.

## Secondary Structure at Interface for Ligands

| | *ClusPro* | *GrammX* | *HexDock* | *PatchDock* | *ZDock* |
|---|---|---|---|---|---|
| *Calcineurin* | Alpha | Alpha | Alpha | Alpha | Alpha |
| *Calmodulin* | Alpha | Alpha | Alpha | Coil | Coil |
| *CCND3* | Alpha | Alpha | Coil | Coil | Alpha |
| *ER-beta* | Alpha | Alpha | Coil | Alpha | Alpha |
| *FKBP1A* | Coil | Coil | Coil | Coil | Beta |
| *PIM1* | Alpha | Alpha | Coil | Alpha | Alpha |
| *TGFBR1* | Coil | Beta | Coil | Coil | Coil |

*Table 2 Secondary Structure at Interface for different Ligands*

If the secondary structure at the interface is same across at least 4 servers, the result is highlighted. As shown, FKBP1A, PIM1 and TGFβR1 have all got similar secondary structure at interface.

Interestingly, interaction with FKBP1A and TGFβR1 seems to provide a link to the control of p21 (CDKN1a) shown in β-cell regulation. TGFβR1 and FKBP1A are known interacting partners. Hence, WDR13 might act through TGFβR1 and FKBP1A to act as a switch during β-cell proliferation. PIM1 is involved in controlling NFAT transcription factors. As indicated in a study by JJ Heit (2006), Calcineurin/NFAT signaling regulates beta cell proliferation as well. This furthers our claim that WDR13 is a central protein to beta cell regulation. Another paper

by Santini MP, Talora C, Seki T, Bolgan L and Dotto GP discusses cross-talk between Calcineurin, NFAT and p21.

## 5.5 POSSIBLE INTERACTION PATHWAY

A possible interaction pathway is shown below. The image was generated using GeneMania, an online tool that predicts pathways between multiple genes using coexpression, physical interactions, genetic interactions, pathway information, colocalization and shared protein domains.

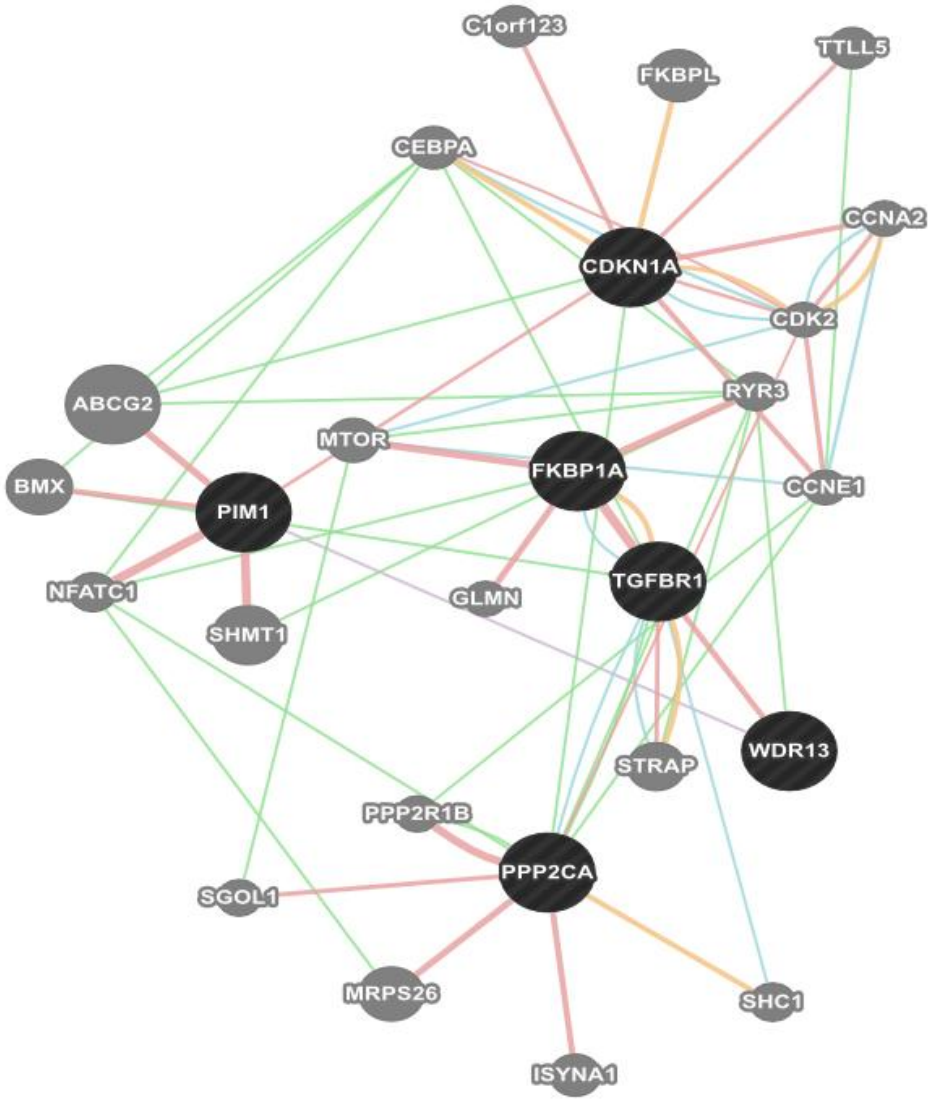



*Figure 19 GeneMania plot of possible pathways for WDR13*

We also found WDR13 might have a role in Wnt signaling pathway. We submitted the structure of WDR13 for function prediction to an online server 3d2GO. 3d2GO is an Imperial College, London initiative that uses a machine learning technique called Support Vector Machine. Our results our tabulated below.

| GO Term | Description | Confidence |
|---|---|---|
| GO:0005515 | Protein Binding | 0.83 |
| GO:0005737 | Cytoplasm | 0.81 |
| GO:0005634 | Nucleus | 0.55 |
| GO:0007165 | Signal Transduction | 0.48 |
| GO:0005488 | Binding | 0.29 |
| GO:0016055 | Wnt Receptor Signaling Pathway | 0.28 |

*Table 3 Predicted Functions of WDR13 by 3d2GO*

To assess the confidence level more reliably, we also gave a known protein, G-β subunit to 3d2GO. The results are given below.

| GO Term | Description | Confidence |
|---|---|---|
| GO:0005515 | protein binding | 0.78 |
| GO:0007165 | signal transduction | 0.48 |
| GO:0016055 | Wnt receptor signaling pathway | 0.28 |
| GO:0006350 | transcription | 0.27 |
| GO:0005737 | cytoplasm | 0.21 |
| GO:0005856 | cytoskeleton | 0.20 |

*Table 4 Predicted Functions of G-beta subunit by 3d2GO*

G-β has known functions in both signal binding as well and the Wnt receptor signaling pathway. Given that WDR13 gets a similar confidence for both these functions, it has a very probable role in the Wnt pathway. Moreover, this property of WDR13 is further cross-referenced by another review article () where the author discusses how Wnt/$Ca^{2+}$ controls NFAT transcription factors through Calcineurin. In this manner, we find multiple connections between p21 (CDKN1a), Calcineurin (PPP3CA), Calcineurin regulator (PPP2CA) and the Wnt signaling pathway.

## 6. CONCLUSION

WDR13 is a characteristic WD repeat protein forming a beta propeller with 7 blades. The blades of the propeller are highly symmetric and extremely stable in GROMACS simulation. This, like other WD proteins, must be a result of Velcro Closure.

Unlike its close homologue WDR5, WDR13 has an N-Terminal which consists of many alpha helices. The role of these alpha helices seems fairly important as many of the complexes formed from docking studies indicated this as the region of interaction. In our docking studies, we found FKBP1A, PIM1 and TGFβR1 as strong candidates for WDR13 interacting partners. This is demonstrated by the inter-connectivity of these proteins in the control of p21, Calcineurin and Wnt signaling pathway.

FKBP1A and TGFβR1 are themselves known interacting partners with each other. They are both involved in the control of p21. PIM1 is an activator of the Nuclear Factor of Activated T-Cells (NFAT) transcription factors, specifically NFATC1. Hence we establish that WDR13 has a central role to multiple pathways. It probably acts as a switch in many situations and the control of WDR13 expression in specific tissues can cause a differential growth pattern.

The expression of WDR13 is not only high in the pancreas, ovaries and testes, but also in the hippocampus of the brain. It generally occurs post neuronal injury and is possibly responsible for neuronal regeneration and plasticity. It expresses along with Calcineurin, a protein that in association with NFAT is known to signal cell proliferation in other types of cells. Our project gives a possible explanation to the association of Calcineurin and WDR13 through PIM1, as PIM1 is an activator of NFAT transcription factors. NFAT along with Calcineurin/NFAT has a role not only in β-cell proliferation, but also in lung tissue maturation and other developmental roles.

## 7. RECOMMENDATIONS

This project demonstrates the use of *in silico* approaches in finding solutions to the many problems of modern biology. Bioinformatics tools are abundant on the net. These are either available for free download or are accessible as online servers. This project outlines a possible approach for development of protein structures and identification of interacting partners through such bioinformatics tools.

At the end of 2012, PDB hosted an overall of 87,089 protein structures. UniProtKB, however, hosts sequences for about 39870577 proteins. This means we know protein structures for hardly 0.22% of the sequences in UniProtKB. As our project illustrates, homology modeling and fold recognition are extremely robust methods of predicting structures for proteins. *De novo* or *ab initio* methods that are being implemented for protein structure prediction are not completely successful currently. Their predictions go awry once the length of the protein exceeds 100 amino acids. Further development of the *de novo* method of protein structure prediction would be a big leap in understanding how proteins actually fold in their native environments.

While the focus of my project has been on WDR13, the approach used to predict interacting partners can be implemented on a genome wide scale. The use of physical docking to score interacting partners is fairly new and requires better algorithms too. Some methods of validating true interacting partners can include generating heat maps and studying electrostatic interactions, along with finding a consensus binding region. Studies of protein-protein interface characteristics will also help in analyzing whether the same type of interface forms across different algorithms. Once the binding region is identified, docking can be re-performed by specifying the hotspot residues. Such an iterative process will yield very good complexes and eliminate false positives.

WDR13, a WD protein implicated in diabetes also has strong roots to neural plasticity and neuronal regeneration in the hippocampus. It shares a calcium binding motif with hippocalcin, also a hippocampal protein. This suggests that WDR13 might be involved in a calcium dependent pathway. This requires further study and will be need to evaluated in a comprehensive manner.

This project brought out the inter-connectivity of WDR13, Wnt signaling pathway, p21 and Calcineurin. While most signs point to WDR13's central role in controlling all three pathways, further research and time would be required to elucidate the mechanism of control.

At this juncture, we would like to iterate WDR13's significance in the human. While we have given our attention to its role in the pancreas and brain, it is also expressed in retinal cells. It might be implicated in many retinal ailments. Thus, we feel there is substantial incentive to study this protein in depth and understand the mechanisms in which it controls different pathways.

Finally, FKBP1A, TGFβR1 and PIM1 as interacting partners of WDR13 is an *in silico* experimental result. The only way of truly knowing whether these actually interact is by performing co-immunoprecipitation. Bioinformatics can only help narrow down a large dataset into few specific proteins of interest. It cannot replace wet lab techniques in proving results, at least not until bioinformatics algorithms are extremely precise.

## 8. REFERENCES


1. Chao Xu & Jinrong Min, *Structure and function of WD40 domain proteins*, Protein and Cell 2011 Mar 2(3):202-14

2. Smith, Temple, *Diversity of WD-Repeat Proteins: Chapter 2*, 2008, Landes Bioscience

3. Christian U. Stirnimann, Evangelia Petsalaki, Robert B. Russell and Christoph W. Muller, *WD40 proteins propel cellular networks*, Trends Biochem Sci 2010 Oct 35(10):565-74

4. Steven van Nocker and Philip Ludwig, *The WD-repeat protein superfamily in Arabidopsis: conservation and divergence in structure and function*, BMC Genomics 2003 Dec 12 4(1):50

5. Bhupendra N. Singh, Amritha Suresh, Gogineni UmaPrasad, Subbaya Subramanian, Mehar Sultana, Sandeep Goel, Satish Kumar, and Lalji Singh 2002, *A highly conserved human gene encoding a novel member of WD-repeat family of proteins (WDR13)*, Genomics 2002 Mar 81(3):315-28

6. Vijay Pratap Singh, B. Jyothi Lakshmi, Shalu Singh, Vanya Shah, Sandeep Goel, D. Partha Sarathi, Satish Kumar, *Lack of Wdr13 Gene in Mice Leads to Enhanced Pancreatic Beta Cell Proliferation, Hyperinsulinemia and Mild Obesity*, PLOS One 2012 7(6): e38685. doi:10.1371/journal.pone.0038685

7. Mitch Price, Molly G. Lang, Ami T. Frank, M. Paula Goetting-Minesky, Samip P. Patel , Matthew L. Silviera, J. Kyle Krady, Robert J. Milner, Andrew G. Ewingb, Jonathan R. Day, *Seven cDNAs enriched following hippocampal lesion: possible roles in neuronal responses to injury*, Molecular Brain Research 2003 Sep 10;117(1):58-67

8. Yang Zhang, *I-TASSER server for protein 3D structure prediction*, BMC Bioinformatics 2008 Jan 23;9:40

9. James D Watson, Roman A Laskowski and Janet M Thornton, *Predicting protein function from sequence and structural data*, Current Opinion in Structural Biology. 2005 Jun;15(3):275-84

10. Marco Punta, Yanay Ofran, *The Rough Guide to In Silico Function Prediction, or How To Use Sequence and Structure Information To Predict Protein Function*, PLoS Computational Biology. 2008 Oct;4(10):e1000160

11. Ambrish Roy, Dong Xu, Jonathan Poisson, Yang Zhang, *A Protocol for Computer-Based Protein Structure and Function Prediction*, Journal of Visualized Experiments, 2011 Nov 3;(57):e3259

12. Qiangfeng Cliff Zhang, Donald Petrey, Jose´ Ignacio Garzo, Lei Deng and Barry Honig, *PrePPI: a structure-informed database of protein–protein interactions*, Nucleic Acids Research 2013 Jan;41(Database issue):D828-33

13. Mark Nicholas Wass, Gloria Fuentes, Carles Pons, Florencio Pazos5 and Alfonso Valencia, *Towards the prediction of protein interaction partners using physical docking*, Molecular Systems Biology 2011 Feb 15;7:469

14. Shujun Yuan, Maya Topf, Thomas F. Reubold, Susanne Eschenburg, and Christopher W. Akey, *Changes in Apaf‑1 Conformation That Drive Apoptosome Assembly*, Biochemistry 2013 Apr 2;52(13):2319-27

15. N. Eswar, M. A. Marti-Renom, B. Webb, M. S. Madhusudhan, D. Eramian, M. Shen, U. Pieper, A. Sali, *Comparative Protein Structure Modeling With MODELLER,* Current Protocols in Bioinformatics, John Wiley & Sons, Inc., Supplement 15, 5.6.1-5.6.30, 2006

16. Pronk, et al*. 2013* Bioinformatics ***29*** *845-854.*

17. Bevan Lab GROMACS Tutorial (Lysozyme in Water) *http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/lysozyme/index.html* (fetched 29.06.13)

18. 3d2GO Online Server for Function Prediction of Proteins based on Structure (http://www.sbg.bio.ic.ac.uk/~phyre/pfd/)

19. Chen R, Li L, Weng Z, *ZDock: An initial-stage protein docking algorithm*, Proteins. 2003 Jul 1;52(1):80-7

20. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ, *PatchDock and SymmDock: servers for rigid and symmetric docking*, Nucleic Acids Research, 2005 Jul 1;33(Web Server issue):W363-7

21. Comeau SR, Gatchell DW, Vajda S, Camacho CJ, *ClusPro: A fully automated algorithm for protein-protein docking¸* Nucleic Acids Research, 2004 Jul 1;32(Web Server issue):W96-9

22. Tovchigrechko A, Vakser IA, *GRAMM-X public web server for protein-protein docking*, Nucleic Acids Research, 2006 Jul 1;34(Web Server issue):W310-4

23. Gary Macindoe, Lazaros Mavridis, Vishwesh Venkatraman, Marie-Dominique Devignes, and David W. Ritchie, *HexServer: An FFT based protein docking server powered by graphics processors*, Nucleic Acids Research, Nucleic Acids Res. 2010 July 1; 38(Web Server issue)

24. Santini MP, Talora C, Seki T, Bolgan L, Dotto GP, *Cross talk among calcineurin, Sp1/Sp3, and NFAT in control of p21(WAF1/CIP1) expression in keratinocyte differentiation*, Proceedings of the National Academy of Sciences USA 2001 Aug 14;98(17):9575-80

25. T P Rao, Michael Kühl *An Updated Overview of Wnt Signaling Pathways: a prelude for more*, Circulation Research, 2010 Jun 25;106(12):1798-806

26. Surendra S. Negi, Werner Braun, *Statistical Analysis of Physical-Chemical Properties and Prediction of Protein-Protein Interfaces*, Journal of Molecular Modeling, 2007 November; 13(11): 1157–1167

27. K G Tina, R Bhadra, N Srinivasan, *PIC: Protein Interactions Calculator*, Nucleic Acids Research, 2007 July; 35(Web Server issue): W473–W47

28. P. Christoph Champ, Carlos J. Camacho, *FastContact: A Free Energy Scoring Tool for Protein-Protein Complex Structures*, Nucleic Acids Research, 2007 July; 35(Web Server issue): W556–W560

29. Vakser IA, Kundrotas P, *Predicting 3D Structures of Protein-Protein Complexes*, Current Pharmaceutical Biotechnology, 2008 Apr;9(2):57-66

30. Murakami Y, Jones S, *SHARP2: protein-protein interaction predictions using patch analysis,* Bioinformatics, 2006 Jul 15;22(14):1794-5

31. Reynolds C, Damerell D, Jones S, *ProtorP: a protein-protein interaction analysis server*, Bioinformatics, 2009 Feb 1;25(3):413-4

32. Nan Zhao, Bin Pang, Chi-Ren Shyu, Dmitry Korkin, *Structural Similarity and Classification of Protein Interaction Interfaces*, PLoS One, 2011; 6(5); e19554

33. Johansson JU, Lilja L, Chen XL, Higashida H, Meister B, Noda M, Zhong ZG, Yokoyama S, Berggren PO, Bark C, *Cyclin-dependent kinase 5 activators p35 and p39 interact with α-subunit of $Ca^{2+}$/Calmodulin depend kinase II and α-actinin-1 in a Calcium dependent manner,* Brain Research/Molecular Brain Research, 2005 Aug 18;138(2):215-27

34. Ofran Y, Rost B, *Protein-protein Interaction Hotspots Carved into Sequences*, PLoS Computational Biology, 2007 Jul;3(7):e119

35. Rulifson IC, Karnik SK, Heiser PW, ten Berge D, Chen H, Gu X, Taketo MM, Nusse R, Hebrok M, Kim SK, *Wnt signaling regulates pancreatic beta cell proliferation,* Proceedings of the National Academy of Sciences USA, 2007 Apr 10;104(15):6247-52

36. Heit JJ, Apelqvist AA, Gu X, Winslow MM, Neilson JR, Crabtree GR, Kim SK, *Calcineurin/NFAT signaling regulates pancreatic beta-cell growth and function*, Nature, 2006 Sep 21;443(7109):345-9

37. UniProt: WDR13_Human (http://www.uniprot.org/uniprot/Q9H1Z4) fetched 14.06.2013