Chapter 5

Spontaneous cancer regression as reversion of spontaneous cancer progression: Molecular targets and pharmacological implications.

CHAPTER 5: SPONTANEOUS CANCER REGRESSION AS REVERSION OF SPONTANEOUS CANCER PROGRESSION: MOLECULAR TARGETS AND PHARMACOLOGICAL IMPLICATIONS

5. Outline

We now endeavour to obtain more rigorous and deeper dynamic insights into the nature of the spontaneous cancer regression process. For this we contrast and compare the process of cancer progression with cancer regression, and their corresponding signal transduction and gene basis. Thereby we focus on to target two main signaling pathways of melanoma progression i.e. MAPK and PI3K/AKT signaling pathways, whose activity we try to alter by inhibiting two melanoma causing genes (BRAF and NRAS), thereby halting the melanoma progression process, and enabling the duplication of the melanoma regression process. For this methodology, we have taken microarray date of spontaneous melanoma regression process, and thereby confirm that by dealing or blocking with MAPK and PI3K/AKT signaling pathways of the tumour progression process, we may induce tumour regression, and thus explore an enhanced therapeutic framework for melanoma patients. To achieve this approach, we further investigated the time-wise biopsy and microarray dynamics of the aforesaid spontaneously regressing melanoma tumours, which in the initial time points display the tumour progression behaviour, before the tumours start to undergo spontaneous regression. We analysed the differentially-expressed genes (DEGs), and performed Gene-ontology investigation, and bioinformatics network pharmacology of this melanoma regression data.

In the initial progression phase of those melanoma tumours, we found that most of the DEGs were involved in either MAPK or PI3K/AKT signaling pathways or both. Additionally, we found out that BRAF and NRAS, the prospective oncogenes, are mostly

involved in melanoma progression. To further elucidate the mechanism, molecular docking was performed to confirm that inhibition of BRAF oncogene (PDB ID: 5C9C) and NRAS oncogene (PDB ID: 6ZIZ), this inhibition would actuate spontaneous tumour regression. We identified 3 small molecules (Alpelisib, Cetuximab, and Obatoclax) which targets both these oncogenes, and none of these molecules have been earlier used therapeutically on melanoma. Out of these agents, Alpelisib and Cetuximab have been selected as potential repurposed drugs which inhibit both genes together efficiently, as we found from drug docking and molecular dynamics analysis. It may be mentioned that Our current investigations point out that Alpelisib and Cetuximab can be potential repurpose drugs which target both the BRAF and NRAS pathways together, instead of targeting one gene at a time, and this conjoint activity may enable therapeutically replicating the process of spontaneous regression phenomenon on malignant melanoma patients.

5.1 Introduction

Spontaneous cancer regression occurs when a tumour which initially grows spontaneously, undergoes retardation and thence reversion, whereby the tumour gradually undergoes involution and finally elimination. Thus in the initial phase there is spontaneous tumour progression which will be actuated by genes, including oncogenes, which produces cell proliferation of these proliferated cell, i.e. these genes will have upregulation in this phase. Furthermore, in this phase, other genes (e.g. cell stability genes) that inhibits cell proliferation genes or oncogenes, will be downregulated, thereby leading to the progression of the tumour. Indeed, the process of tumour reversion, i.e. conversion from progression to regression behaviour of a tumour is well-documented by cytologists and tumour biologist [116].

When tumour reversion occurs, the tumour undergoes elimination, thus there will be reversal of upregulation of the cell proliferation genes and the reversal of the cell stability genes become vice versa. That is during regression, the oncogenes or cell proliferation related genes become downregulated, while the oncogene-inhibitory gene (or cell-stability genes) become upregulated. Indeed, one can perform incisive analysis of the gene-profile in spontaneous tumour progression phase vis-a-vis that in spontaneous tumour regression phase. This analysis may have potential to furnish novel molecular target with insightful prospect of innovative pharmacological approach. Hence in this chapter we give attention to the spontaneous tumour progression phase, and thereby we hope to obtain more molecular biological and therapeutic insights and mechanistic implications.

As has been delineated in chapter 3 regarding the case of spontaneous cancer regression of melanoma, the tumour is under full progression in time t_1; then (at time t_2) the tumour progression has halted for an appreciable time, while the regression process has started and is ongoing. The regression process continues to go on through the next time point t_3; thereafter at the last time point t_4, the regression process reaches full efficiency producing complete regression of the tumour. Thus the temporal regime t_1 pertains to the progression phase, while the temporal regime t_2-t_3-t_4 pertains to the regression phase. Thus in the regression phase, the expression of the genes associated with regression will have positive value (upregulated) in all the three time points t_2, t_3 and t_4 (if the genes inhibit oncogenes or blocks cell-proliferation signalling pathways). On the other hand, if the genes are themselves oncogenes or activates cell- proliferation signalling pathways, then these genes will be downregulated (negative values) during the regression phase t_2-t_3-t_4. This approach was taken in the last chapter.

Coming to the present chapter, we here deal with comparison of the full progression point (t_1) and the full regression point (t_4). As per the elucidation in the paragraphs above,

we may assume that the genes (that will be relevant to the tumour reversion process) shall have opposite values of expression at time t_1 and t_4, i.e either upregulation at t_1 and downregulation at t_4, or vice-versa. Let us now first analyze the signaling pathways involved in the progression of melanoma. In recent years, much progress has been made in understanding the regulatory and genomic processes involved in melanoma initiation and progression. The genetic alteration in mitogen activated protein kinase (MAPK) and the phosphoinositide 3-kinase (PI3K) signaling pathways are the main cause of melanoma progression. MAPK and PI3K signaling pathways helps melanoma in proliferation, invasion, survival and angiogenesis [38], both the pathways are activated by Ras protein mutation. The v-Raf murine sarcoma viral oncogene homolog B (BRAF) is by far the most crucial mutated oncogene in melanoma, with the BRAF V600E mutation being the most significant mutation, accounting for 40–50% of all mutated melanomas and 80% of BRAF-mutated tumours [39]. The neuroblastoma RAS viral oncogene homolog (NRAS), which is mutated in over 30% of all melanomas, is the second most significant mutant oncogene [39].

In our previous chapter, we investigated the regression genes involved in the melanoma spontaneous regression phenomenon by considering the genes whose expression was consistent from t_2 to t_4. While going deep into this approach, we found that the Top2a muted gene mutation is more frequent and common in the case of melanoma. Hence, we targeted the Top2a gene and found the candidate drugs to inhibit this gene, so as to replicate the regression phenomenon. However, in the present chapter, we have taken another collateral approach to selecting the melanoma regression genes. As mentioned earlier, we have here taken those genes whose expression levels differ between time points t_1 (progression) and t_4 (regression); these genes are mainly significant in MAPK and PI3K

signaling pathways. Furthermore, we have combined both approaches of the last chapter and this chapter to obtain the most effective pathways and potent drugs.

In this study, we systemically probe the signal pathways driving melanoma pathogenesis, including the main mutated driver genes and the associated targets and possible anti-melanoma drugs thereby implied. Furthermore, we have performed the docking study and molecular dynamic analysis to validate our findings and predict the drug which would thereby mimic or replicate the spontaneous tumour regression process on the melanoma growth of the patients.

5.2 Methods and Materials:

5.2.1 Data Collection and Investigation:

Here we have analyzed the raw microarray information of the spontaneous tumour regression model of melanoblastoma bearing Libechov minipigs (MeLiM) from the ArrayExpress platform (https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MEXP-1152?query=E-MEXP-1152) using R-studio platform [117]. Here, the tumour spreads and kills the majority of the pigs, but in the minority of the animals, the tumour develops up to a certain point before spontaneously shrinking and healing, and these minority animals remain healthy.

5.2.2 Identification of differentially expressed genes (DEGs) of spontaneous melanoma regression data:

After applying the two primary cutoff criteria ($-2 > FC > +2$ and p-value < 0.05), we have 70, 322, 1147, and 1349 DEGs at time points t_1, t_2, t_3, and t_4. For our current study, we have taken two types of approaches:

a) As in our prior study, we have described how the regression started from time point t_2; hence we only included genes whose expression was consistent from t_2 to t_4.

b) The other gene sets are those genes whose expression levels differ between time points t_1 and t_4; for example, a gene that is upregulated at t_1 should be downregulated at t_4.

We also consider the protein-protein interaction (PPI) networks which are protein complexes arranged in networks that are created by electrostatic or biochemical forces [99]. The PPI network is essential for molecular functions, and the String database is used to generate the protein-protein interaction networks for selected DEGs.

5.2.3 Gene Ontology Analysis:

The Gene enrichment analysis of DEGs is done by DAVID database [107] (https://david.ncifcrf.gov/), FunRich software (version: 3.1.3) [118] and ClueG0-v2.5.5 plugin from Cytoscape [108]. Significant terms for cellular component (CC), biological process (BP), and molecular function (MF) in Gene Ontology (GO) results were graded by p-value and displayed as bar charts. As a result, in this study, we created and compared a network of functionally related GO term using the standard adjusted kappa statistic of $\kappa > 0.4$. FunRich was employed in this study to examine the biological pathways involved in DEGs. The top 5 biological pathways for genes that were upregulated and downregulated, respectively, were then displayed as bar charts. Statistics were considered significant at $p < 0.05$.

5.2.4 Identification of Melanoma regression genes:

Now, we need to identify the most effective melanoma regressing driver genes from the aforesaid two type of approaches of selecting DEGs. For the first case we have used cytohubba plugin in cytoscape, already explained in our previous work [117]. For the

second case we have selected the common gene among (i) the pig spontaneous melanoma microarray data genes (those genes whose expression levels differ between time points t_1 and t_4), (ii) the human melanoma gene from NCBI database and the human melanoma gene from Gene card database.

5.2.5 Identification of target proteins

Targeted molecular therapies have mostly been recognized as a novel therapeutic procedure in most malignancy types. The signaling pathways involved in malignant melanoma targeted by the molecular therapies are RAS–RAF–MEK–ERK and PI3K–AKT pathways; for this purpose, small-molecule inhibitors play a crucial role. Recent evidence shows that tumour growth cannot be blocked or stopped by only inhibiting a single effector protein of the signal transduction cascades which play a role in melanoma pathogenesis, but a multimodal molecular modification and inhibition is necessary for the most effective targeted melanoma treatment[119]. We elucidate the latter aspect below.

(i) Melanoma signaling pathway

Recent developments in molecular oncology have produced new treatment approaches that specifically target (i) immune regulatory molecules involved in suppressing the antitumour immune response, such as T-lymphocyte-associated antigen 4 (CTLA4), programed cell death 1 receptor (PD-1), and its ligand (PD-L1), (ii) key effectors of the pathways such as mutations in BRAF, NRAS or cKIT genes, found to play an important role in the pathogenesis of melanoma[120].

(ii) MAP Kinase-ERK Dependent Pathway

The NRAS and BRAF proteins are part of the mitogen-activated protein kinase (MAPK) signal transduction system, which controls cell growth, survival, and proliferation and mediates the response of cells to mitotic external stimuli. Three tissue-specific isoforms

of the small proteins produced by the RAS gene family—HRAS, KRAS, and NRAS—are linked to the cytoplasmic membrane. Indeed most NRAS mutations have been found in melanomas [40, 41]. RAF and phosphatidylinositol 3 kinase (PI3K) are two downstream cytoplasmic proteins that can be activated by NRAS. The three members of the RAF kinase family—ARAF, BRAF, and CRAF—are proteins whose activation depends on the formation of complexes by these various isoforms [42]. All three proteins contribute to the signal transduction of MAPK pathway's. BRAF causes the activation of MEK kinase in melanocytes, which then triggers the activation of ERK, the last enzyme in the MAPK cascade (Figure 5.1). In 40–60% of melanoma cases, the BRAF gene is altered; the most frequent mutation (occurring in around 90% of cases) is represented by the substitution of valine for glutamic acid at codon 600 (BRAFV600E) [43]. The BRAFV600E variation provides ongoing stimulation of cell proliferation and tumour formation through activating phosphorylation of ERK, as do the remaining mutations in the BRAF kinase domain. However, the finding that BRAF is even mutated in common nevi implies that melanoma development requires the activation of BRAF's oncogenic pathway but is not a prerequisite for it [44].

(iii) PI3Kinase-AKT Dependent Pathway

The second method that uses RAS to regulate cell proliferation is made up of the signal transduction pathway: PTEN-PI3K-AKT [45]. The activation of PIK3 and the activity of the phosphatase PTEN protein both affect intracellular levels of PIP2 and PIP3 phosphoinositols in physiological settings [46]. High PIP3 levels modulate the synthesis of proteins essential in cell growth and survival as well as apoptosis by progressively activating downstream AKT (mostly AKT3 in melanoma) and its substrate mTOR. The activation of AKT leads to inhibiting apoptosis by inactivating numerous pro-apoptotic

proteins, such as BAD (BCL-2 antagonist of cell death) and MDM2 (which results in the degradation of p53), as well as promoting cell proliferation through the induction and stabilization of CCND1and through the inhibition of apoptosis [40] (Figure 5.1). In conclusion, PTEN inactivation and PI3K-AKT stimulation result in the abnormal proliferation of cancer cells and development of apoptosis resistance. Around 70% of all melanomas have dysregulated activation of the AKT pathway, which is caused by AKT3 amplification and PTEN loss due to epigenetic silencing or deletion, as originally described [47, 48].

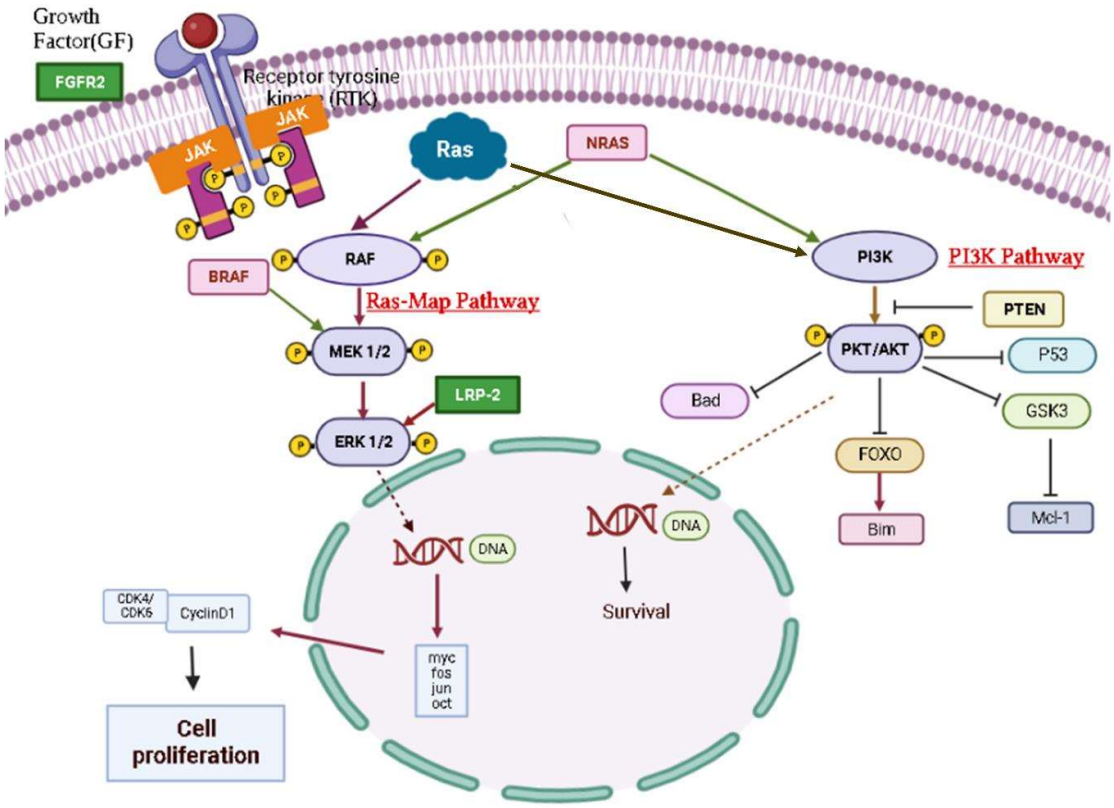

Figure 5.1 The MAP/ERK and PI3K/AKT signaling pathway for melanoma tumour proliferation.

5.2.6 Drug Identification

Here with the help of DGIdb (The Drug Gene Interaction) platform (https://www.dgidb.org/), we have selected only the FDA approved drugs which are targeting both the signaling pathways i.e. BRAF and NRAS. Furthermore, we have taken those drugs which effect are common in both pathways, and we select those drugs on which no study has been performed in melanoma. Also we consider those drugs that are used to target both the pathways in other cancer but have not been used in melanoma.

5.2.7 Validation of Drug using Molecular docking

5.2.7.1 Molecular docking with BRAF and NRAS:

The protein sequence of the human BRAF gene [UniProt ID: P15056 (https://www.uniprot.org/uniprotkb/P15056/entry); 766 amino acids] and NRAS genes [UniProt ID: P01111 (https://www.uniprot.org/uniprotkb/P01111/entry); 189 amino acids] were retrieved from the UniProt database. The 3D PDB structure of protein were downloaded from PDB (Protein Data Bank (www.rcsb.org)) platform having PDB ID: 5C9C (for BRAF, with a resolution of 2.70 Å) [121] and PDB ID: 6ZIZ (for NRAS, with a resolution of 1.78 Å) [122]. The reason we selected 6ZIZ structure of NRAS from PDB database is because it had Q61R mutation which is well known in Melanoma. Similarly, 5C9C structure of BRAF was chosen because it has V600E mutation which is again a crucial mutation found in melanoma patients. The 3D structures of the drugs thus obtained, from PubChem database in '.sdf' format, the drugs having are - Alpelisib (PubChem CID: 56649450), Obatoclax (PubChem CID: 11404337) and the third drug used is a monoclonal Antibody, Cetuximab, thus its structure was retrieved from PDB since it is a protein (PDB ID: 1YY8). The Discovery Studio platform provides a vast range of features for not only

protein-ligand docking but also for protein-protein docking. Since our 6ZIZ protein had GTP, Mg^{2+} ion and a compound EZZ (3~{S})-3-[2-[(dimethylamino)methyl]-1~{H}-indol-3-yl]-5-oxidanyl-2,3-dihydroisoindol-1-one**)** bound to it, we removed these and used the GTP binding sites for docking calculation. Similarly, 5C9C is complexed with Cl^{-1} ion and compound LY3009120 or 4Z5 [**(**1-(3,3-dimethylbutyl)-3-{2-fluoro-4-methyl-5-[7-methyl-2-(methylamino) pyrido[2,3-d]pyrimidin-6-yl]phenyl}urea], which were removed and the 4Z5 binding was used for docking studies. The remaining parameters were kept as default.

The protein and ligand preparation was performed on BIOVIA Discovery Studio platform. The protein (macromolecule) preparation is done as per the guidelines provided in the Discovery Studio manual by the following steps: The protein was pre-processed by cleaning the protein first, the removal of hetero atoms and water molecules, the addition of polar hydrogens undertaken and the then defining the binding site from the current selection option. The prepared protein saved has a '.pdb' format and its ready to dock. The ligand structure was gathered from PubChem (https://pubchem.ncbi.nlm.nih.gov/) in sdf format and ligand preparation (e.g removal of charges, ionization method which is pH based etc) was also performed using Discovery Studio. Protein-ligand docking is performed using LibDock tool under the receptor-ligand interaction menu on Discovery Studio[123]. The algorithm aligns the ligand conformation to hotspots (polar and apolar) of receptor interaction sites and retains best scoring poses of ligands [124].

Further, CDOCKER, a grid-based molecular docking programme that makes use of the CHARMm force field, was used to form a stable complex inside the active site of protein [125]. Ligands with lowest CDOCKER energy and highest LibDock score were shortlisted for further Molecular dynamic simulations. Under Macromolecules section, we used Dock and Analyze Protein Complexes and docked the Receptor proteins (NRAS & BRAF) with

the monoclonal antibody (Cetuximab) and further processed the poses. The Z-Dock protocol provides rigid body docking of two protein structures using ZDock algorithm [126] as well as clustering the poses according to the ligand position. Processing poses allows selection of a subset from a set of docked protein poses generated by ZDock protocol. After this, we refine the docked poses and finally Analyze the interactions.

C-Docker study process:

proposed ligands were subjected to molecular docking in order to find the hits that were biologically active. The conformation with the lowest binding energy is taken to form a stable complex inside the active site of NRAS and BRAF. CDOCKER, a grid-based molecular docking programme that makes use of the CHARMm force field, was used to carry out this investigation. A greater value denotes a more advantageous binding, and the CDOCKER score is expressed as a negative value (i.e., -CDOCKER_ENERGY). The H-bonds, van der Waals, and electrostatic interactions between the target protein and the ligand were used to compute the CDOCKER energy [127]. Based on the crystal data of the template protein, the binding site of the modelled protein was fixed. To find the local minima (lowest energy conformation) of the modelled NRAS and BRAF with an energy gradient of 0.1 kcal mol^{-1} $Å^{-1}$, the smart minimizer technique was used in conjunction with the CHARMm force field.

5.2.7.2 Molecular dynamic simulation:

The Desmond platform was used for molecular dynamic (MD) simulation study for protein-ligand interaction stability analysis. The MD simulation systems were built using system builder module of Desmond. Solvation of system were carried out using TIP3P water model. Periodic boundary conditions for each system were established using the buffer method and orthorhombic box shape with a 10 Å distance. Counter ions were added

to neutralize the system as required. The systems used for the MD simulations were prepared using OPLS-2005 force field [128].

The MD simulation was performed for 100 ns for each system. The energy was recorded at every 1.2 ps, and the trajectory was saved at every 100 ps. The volume of the dynamical box was equilibrated with the NPT (constant-temperature, constant-pressure) ensemble at 300 K and 1.01325 bar pressure. Before selecting the simulation option, the systems were relaxed using Desmond's default relaxation protocol, i.e. Relax model system [129]. All MD simulation trajectory analyses were carried out using Desmond's simulation event analysis tool, which was used to calculate various properties and information.

The Root Mean Square Deviation (RMSD) parameter provides information on a protein's structural features during the simulation. If the simulation has reached equilibrium, it can be revealed by this parameter. On the other hand, the root mean square fluctuation (RMSF) is calculated to describe the local alterations or variations near the amino acids in the protein chain. Furthermore, the radius of gyration (Rg) reveals the compactness of the protein and its folding. Since, intermolecular interactions between the protein-ligand complex are crucial measurements to infer the binding stability of the complex, we measured the various types of intermolecular interactions (Hydrogen bonds, salt bridges, and pi interactions). Figures were generated using Biovia Discovery Studio and Schrodinger's Maestro.

5.3 Result

The systems biology approach for permanent tumour elimination, which we have developed in our previous study enabled us to find the signaling pathways related to permanent tumour elimination. It inspires us to feather investigate the driving genes and downstream regulated pathway for replicating this elimination phenomenon in melanoma.

Further, we have also identified and clarified the potential drugs which could block this melanoma progression pathway and thus enables the elimination of malignant melanoma. Finally, to validate our findings, molecular docking and molecular dynamic simulation were carried out on the targets and drugs.

5.3.1 Microarray data analysis and DEGs identification

The DEGs were identified based on fold change value (-2 >FC > +2) and p-value (p-value < 0.05). We have taken two types of approaches to investigate the spontaneous regression driving genes as discussed in method section 2.3. From the first approach, we have 176 upregulated and 116 downregulated DEGs whose expression values were consistent from time point t_2 and t_3. Further, and from the second approach, we have 213 DEGs whose expressions values are different at time point t_1 and t_4. After combining both the approaches and removing the common genes we have obtained 344 upregulated and 160 downregulated DEGs. Figure 5.2 provides the string network for these DEGs that are respectively upregulated and downregulated.

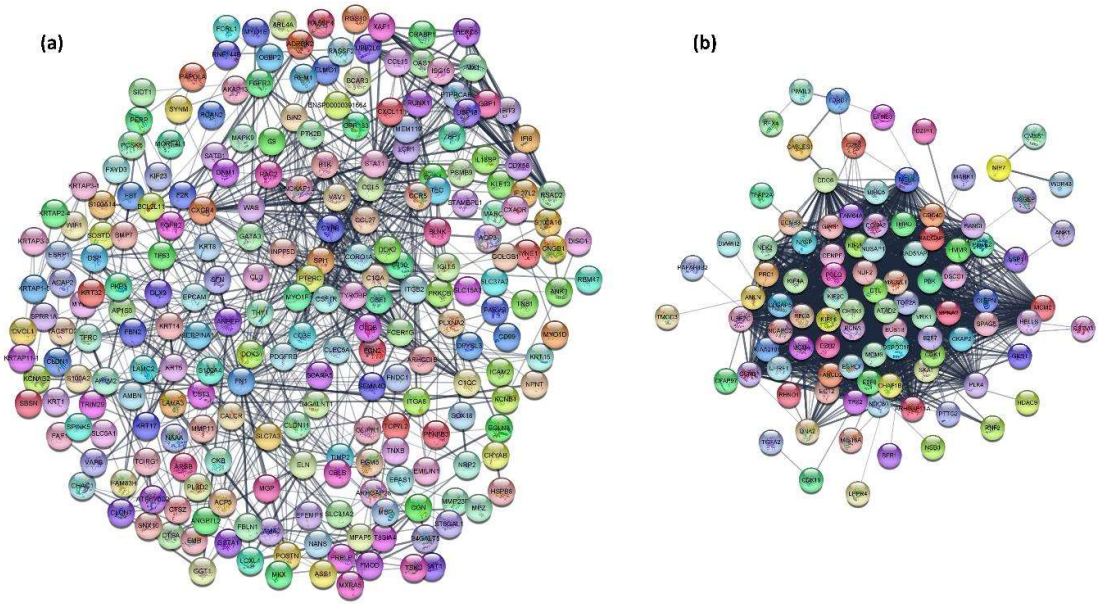

Figure 5.2 The Protein-Protein interaction (PPI) network: (a) Upregulated differentially expressed genes (DEGs); (b) Downregulated differentially expressed genes (DEGs).

5.3.2 Gene enrichment analysis of DEGs

With $p < 0.05$ used as the cutoff, the gene ontology (GO) and pathway enrichment were examined using a multiple databases, including DAVID [107], KEGG pathway [106], and FunRich platform [118]. The DEGs divided into three functional groups by GO analysis: The Cellular Component (CC), Biological Process (BP), and Molecular function (MF) groups. The *upregulated* genes in Cellular Component (CC) group were mainly enriched in extracellular, extracellular matrix, proteinaceous extracellular matrix, exosome and intermediate filament (Figure 5.3 (a)), whereas the *downregulated* genes were enriched in nucleoplasm, microtubules, spindle microtubules, microtubule cytoskeleton and chromosome (Figure 5.3 (b)). On the other hand, the *upregulated* genes in Biological Process (BP) group were enriched in cell growth and/or maintenance, immune response, signal transduction, cell communication and chemosensory behavior, whereas the *downregulated* genes were mainly enriched in cell communication, signal transduction, cell growth and/or maintenance, nuclear organization and biogenesis and regulation of gene expression, and epigenetic functioning. Furthermore, as for Molecular function (MF), the *upregulated* genes were related to extracellular matrix, structural constituent, structural molecule activity, complement activity, receptor signaling complex scaffold activity and chemokine activity whereas the *downregulated* genes were mainly associated with kinase binding, protein serine/threonine kinase activity, DNA binding, DNA-directed DNA polymerase activity and cytoskeletal protein binding.

Moreover, from Figure A1, A2 and A3, of Appendix-III(B), we observed that the *upregulated* genes were significantly connected to the regulation of immune response and its activation, while the *downregulated* genes were associated with mitotic cell cycle process, DNA replication, cyclin depended protein serine/threonine kinase activity and DNA directed DNA polymerase activity. The gene ontology GO term analysis showed that

the majority of the *upregulated* genes were mainly involved in regulating and activating the immune response, while the *downregulated* genes were related to mitotic cell cycle process, spindle microtubules formation and DNA replication; indeed, the *downregulated* of these genes is essential for any malignant tumour in a host body to go under permanent regression process.

Furthermore, from the Biological pathway analysis, the *upregulated* genes were enriched in mesenchymal-to-epithelial transition, PAR1-mediated thrombin signaling events, thrombin/protease-activation receptor (PAR) pathway, integrin family cell surface interactions and beta1 integrin cell surface interactions (Figure 5.4 (a)). Also, the *downregulated* genes were mainly enriched in cell cycle, mitotic, DNA replication, G2/M checkpoints, G2/M DNA Damage checkpoints and cell cycle checkpoints (Figure 5.4 (b)).

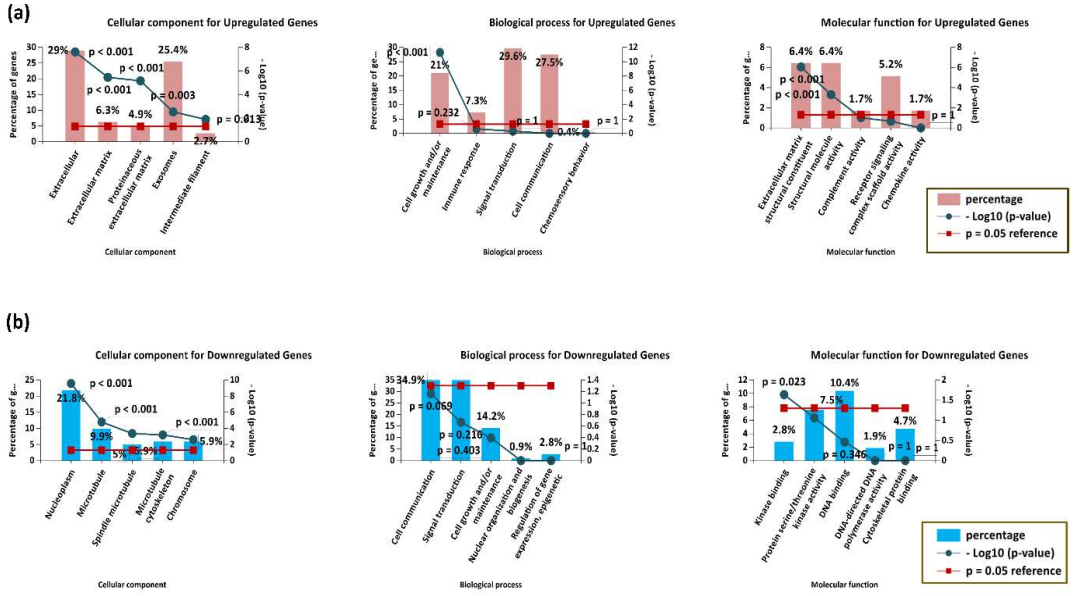

Figure 5.3 Gene Ontology (GO) term analysis of significantly enriched DEGs, classified in 3 groups (Cellular component, Biological process and Molecular function): (a) Upregulated DEGs; (b) Downregulated DEGs

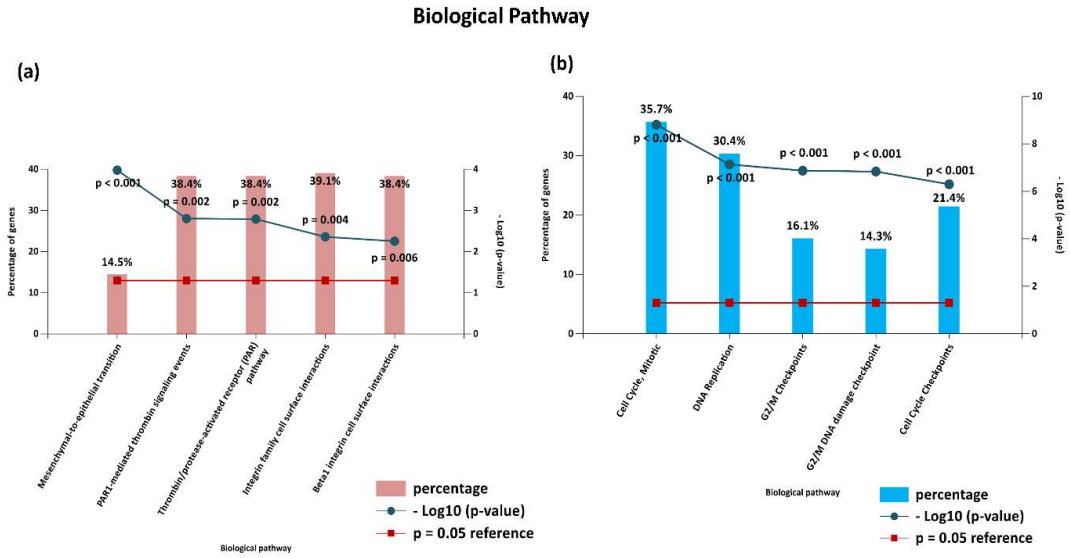

Figure 5.4 Significant Biological Pathway enrichment analysis of DEGs: (a) Biological pathway of upregulated DEGs; (b) Biological pathway of downregulated DEGs

5.3.3 Identification of melanoma tumour regression driving genes

The top 10 hub genes were identified using cytohubba for the first group of genes (Table 5.1) and 17 common genes were identified by taking common genes between microarray data, NCBI database genes and gene card genes (Table 5.2) (Figure 5.5). All these total 27 melanoma driving genes are either involved in cell cycle regulation, cell cycle division or signal pathway activation mainly in MAPK and PI3K/AKT signaling pathways activation. Both the MAPK/ERK (mitogen-activated protein kinase/extracellular signal-regulated kinase signalling pathway) and the PI3K/Akt (lipid kinase phoshoinositide-3-kinase signalling pathway) play a crucial role in the transmission of cell signals via transduction systems (ligands, transmembrane receptors, and cytoplasmic secondary messengers) to the cell nucleus, where they affect the expression of genes that control crucial biological functions like cell growth, proliferation, differentiation and apoptosis [130, 131]. The enrichment analysis of 27 driving gene also shows that the significant biological pathway was enriched in (i) Chk1/Chk2 (cds1) - mediated inactivation of cyclin

B: Cdk1complex, (ii) cell cycle, Mitotic, (iii) Cyclin a/B1 associated event during G2/M transition, (iv) cell cycle checkpoints and (v) phosphorylation of protein involved in the G2/M transition by cyclin A: Cdc2 complexes (Figure 5.6). The information is consistent with the notion that abnormalities in cell cycle checkpoints and cell growth regulators are the primary contributors to tumourigenesis [132, 133].

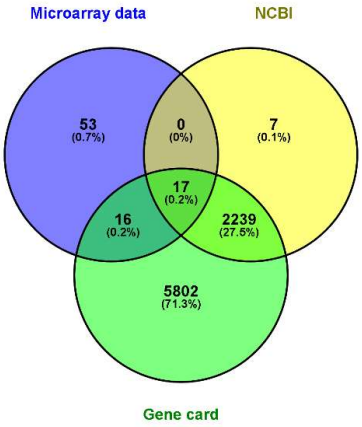

Figure 5.5 Venn Diagram representing the common genes between microarray data, NCBI database and Gene Card database.

Table 5.1 The top 10 hub genes with their fold change value for melanoma regression.

Gene name	log FC value (t_1)	log FC value (t_2)	log FC value (t_3)	log FC value (t_4)
TOP2A	-0.459517001	-1.810469734	-2.25435386	-1.672954085
KIF20A	-0.211683269	-1.807779605	-2.478033628	-2.293056648
KIF23	0.827249425	1.98418174	3.664080664	3.065167151
CDK1	-1.09722215	-2.415204476	-3.692431714	-2.26071039
CCNB1	-0.636137771	-1.936436546	-2.738281085	-2.068221461
RAD51AP1	-0.866296535	-2.212584635	-3.289193264	-2.476289109
BUB1B	-0.205789146	-1.608319671	-2.296878874	-1.649891329
CHEK1	-0.510946563	-1.22006972	-1.519230504	-1.091561664
CCNA2	-1.338388721	-2.449128131	-2.607963513	-1.924035845
KIF11	-0.896395986	-2.28949233	-2.7775018	-1.959737739

Table 5.2 The 17 genes common between microarray data, NCBI database and Gene Card database with their fold change values.

Gene name	log FC value (t_1)	log FC value (t_4)
PTGDS	1.306166618	-2.738621353
POSTN	-1.025105746	3.437919984
HOXD3	1.640127673	-2.164557431
F2R	-1.092731016	1.584284104
HOXD9	1.568937642	-1.058619396
LRP2	1.06431968	-3.002865195
TFAP2A	1.616290548	-2.561774844
IFIT3	-1.04982098	2.207899373
STAT1	-1.279061168	1.224467396
PSMB9	-1.264566206	2.860071669
PCSK6	-1.205405347	1.346377506
FGFR2	-1.166316476	1.677347703
FBLN1	-1.386518814	1.612005154
HSPB8	-1.143988474	1.887807014
DEFB1	-1.105396763	2.654211797
CLDN11	-1.051796694	2.881233246
SOX5	1.251525957	-1.344268009

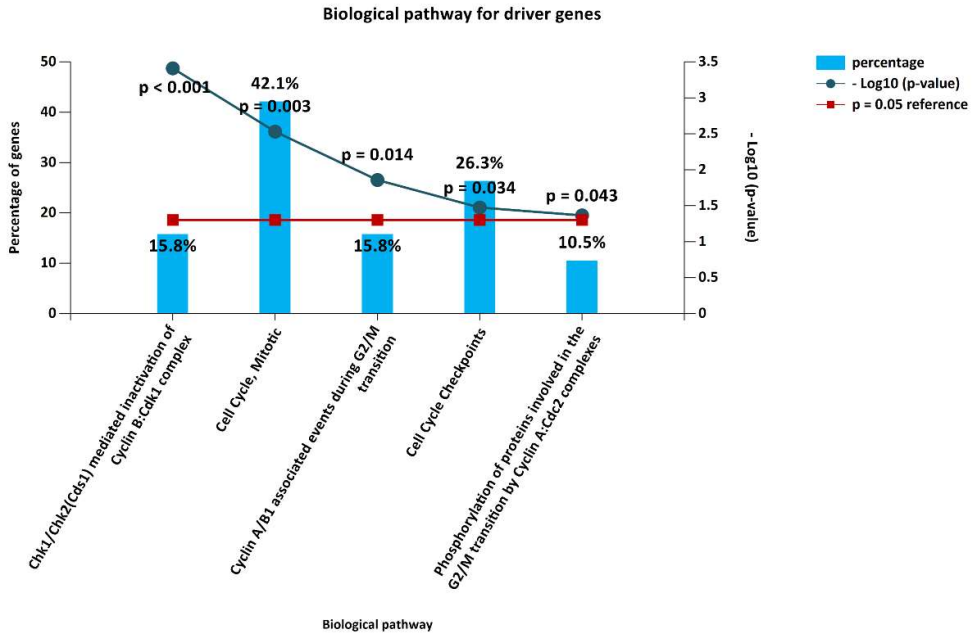

Figure 5.6 Enrichment analysis of significant Biological Pathway analysis of melanoma driving genes.

5.3.4 Downstream signaling pathway of Melanoma and its target protein

The mutually exclusive BRAF or NRAS mutations are the most frequent mutations that present in melanoma. These mutations appear early in melanoma progression and continue throughout the disease's progression[134]. Moreover, the NRAS and BRAF proteins help melanoma in progression through downstream activation of two signaling pathways: RAS-RAF-MEK-ERK pathway and PI3K-AKT pathway (Figure 5.1) which control cell growth, survival, and proliferation [119]. The pathways mediate the response of cells to mitotic extracellular stimuli. In RAS-MAP pathway, the BRAF mutated gene acts on activated MEK, then triggers the activation of the ERK downstream protein, which act on DNA and enhances cell proliferation. Similarly, the mutated NRAS gene act on PI3K and its activated PI3K pathway downstream. Thus activated phosphoinositide 3-kinase (PI3K) is typically the first step in the activation of the AKT pathway after exogenous

growth factors have stimulated it. This is followed by increased production of the second messenger phosphatidylinositol-3,4,5-trisphosphate (PIP3), which can facilitate the translocation of AKT to the plasma membrane for its subsequent phosphorylation and activation. Since the phosphatase PTEN protein controls the intracellular level of PIP3, then PTEN's functional deficit can cause the overexpression of PIP3 and induce AKT activation [135]. It is known that AKT3 is the most prevalent AKT isoform in melanoma. Furthermore, SiRNA transfection directed against AKT3 or overexpression of PTEN are efficient ways to inhibit AKT3 activity and reduce the ability of melanoma cells to become tumourigenic [48]. As a result, the hyper activation of the AKT pathway is a crucial oncogenic event for the development of melanoma.

5.3.5 Identification of Drugs

To screen out and select the drugs we considered three characteristics to select the drugs. Firstly, the drugs should interact with our proteins of interest viz., NRAS and BRAF, as analysed through DGIdb (Drug-gene Interaction database). It may be recollected that the drugs selected are Alpelisib, Obatoclax and Cetuximab (section 5.2.6). Thereby we found that Alpelisib has an interaction score of 0.14 with BRAF and 0.08 with NRAS (higher than other isoforms of RAS), and Obatoclax has an interaction score of 0.53 with NRAS and 0.25 with BRAF, while Cetuximab has an interaction score of 1.6 with BRAF and 0.4 with NRAS. Secondly, we have selected only those drugs which are FDA approved but have not been used for melanoma treatment. Thirdly, we shortlisted our final drugs with LibDock score.

5.3.6 Molecular Docking with BRAF and NRAS

The exploration of the binding of small molecules inside protein-ligand complexes as well as structure-based virtual screening have both made extensive use of molecular

docking. Docking of selected drug molecules against the BRAF and NRAS gives several poses and the best pose for the individual docking complex were selected based on the best -CDOCKER score (lowest CDOCKER energy kcal/mol) in case of protein-ligand docking, while in case of protein-protein docking the pose is selected based on the best E_RDock value (lowest one is the best). Results from the docking are shown in the Table 5.3. Further analysis of docking conformation within the binding pocket were analyzed and 3D diagram were generated along with 2D interaction diagram to visualize the interacting residues.

Table 5.3 Docking results of selected compounds with BRAF and NRAS receptors.

Receptor	**Ligand Name**	**CDOCKER Energy (kcal/mol)**	**CDOCKER Interaction Energy (kcal/mol)**	**LibDock Score**
BRAF (PDB ID: 5C9C_A)	Alpselisib	-5.689	-39.2787	90.2357
	Obatoclax	-15.6203	-41.5001	90.9137
NRAS (PDB ID: 6ZIZ_A)	Alpselisib	-18.6548	-66.7291	84.5067
	Obatoclax	-14.4835	-50.8704	123.64
Receptor	**Antibody**	**ZDock Score**	**Z Rank Score**	**E_RDock (kcal/mol)**
BRAF (PDB ID: 5C9C_A)	Cetuximab	17.24	13.172	-12.2086
NRAS (PDB ID: 6ZIZ_A)	Cetuximab	15.84	-47.903	-10.8281

Interaction study:

The 3D representation of intramolecular interactions for both the protein complexes (BRAF and NRAS) with Alpelisib and Obatoclax drug are given in Figure 5.7 (a, b). Additionally, the 2D interaction were shown for the individual drugs with receptors, and were made to visualize the types of interactions between the receptor residues and ligand (Figure 5.7). Furthermore, the 3D complex of protein-ligand structure is also represented in Figure 5.8 (a, b, c, and d) to show the surface representation of the protein. The purple coloured solid surface is a protein, the green colour shows the non-polar interaction, while the blue colour represents the polar interaction, and the red denoted the ligand in 3D interaction diagram. Similarly, for our third drug (Cetuximab), the 3D interaction for BRAF and NRAS protein is illustrated in Figure 5.9 (a, b). Here in Figure 5.9 (a, b), the blue and purple colour ribbons are the proteins (BRAF and NRAS), while the grey and light brown ribbons are the two chains (chain A and chain B) of the monoclonal antibody i.e. Cetuximab. The residues involved in the protein-protein docking at the interface are shown in the zoomed out images of Figure 5.9.

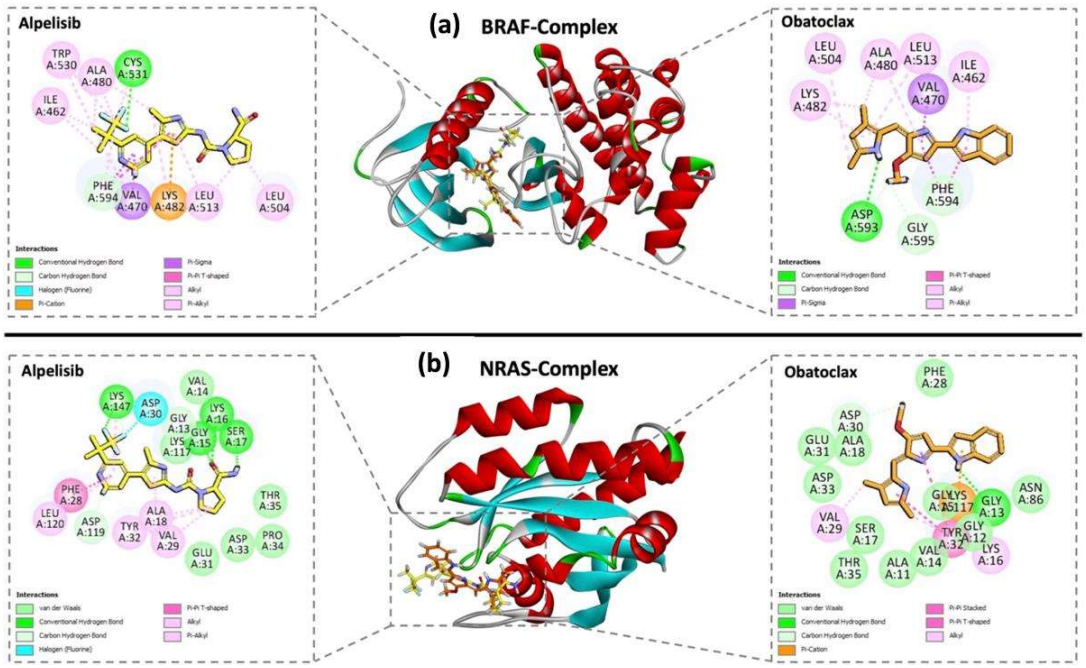

Figure 5.7 The 3D representation of the target receptors docked with ligands, illustrating the interacting residues in 2D interaction diagram. (a) BRAF complex with Alpelisib and Obatoclax, (b) NRAS complex with Alpelisib and Obatoclax.

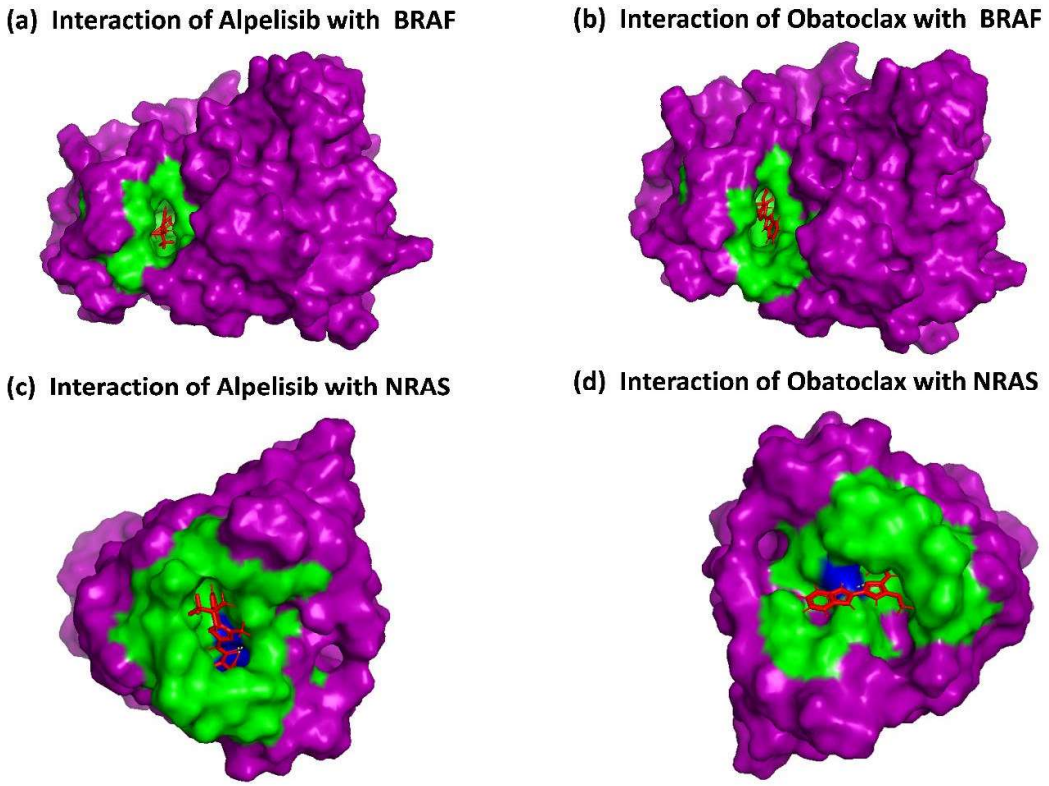

Figure 5.8 The 3D representation of Protein-Ligand complex: (a) Alpelisib with BRAF(5C9C); (b) Obatoclax with BRAF protein (5C9C); (c) Alpelisib with NRAS protein (6ZIZ); (d) Obatoclax with NRAS(6ZIZ).

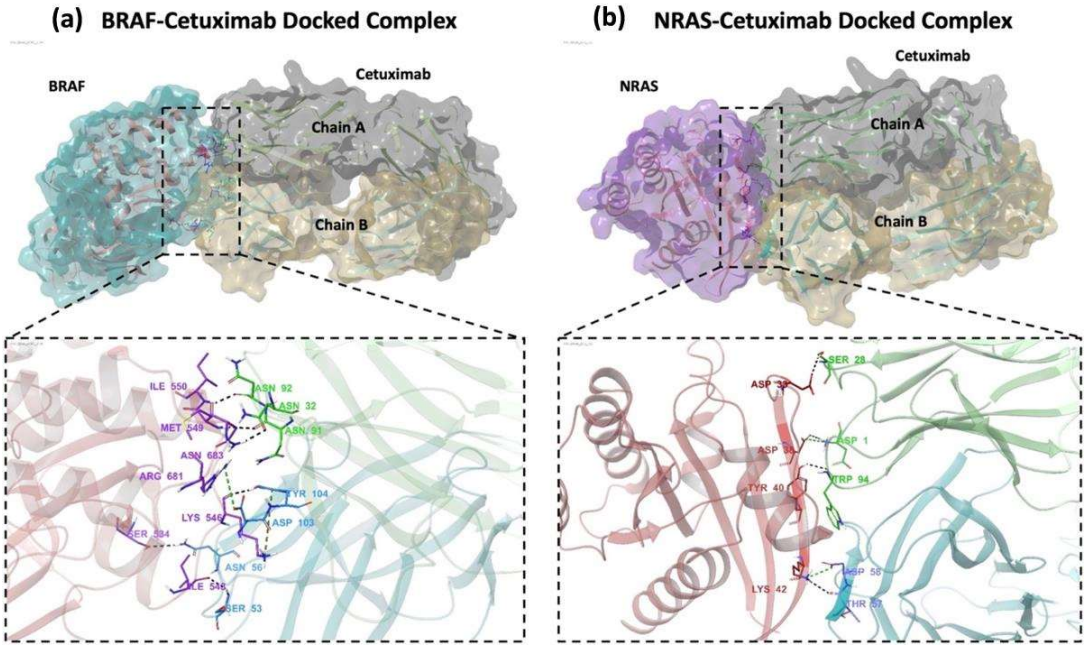

Figure 5.9 The 3D representation of the target receptors docked with Cetuximab, illustrating the interacting residues. (a) BRAF-Cetuximab docked complex, (b) NRAS-Cetuximab docked complex. Color scheme used for dashed line showing interactions are as follows: hydrogen bond (black), salt-bridge (green), and pi-cation (dark green).

5.3.7 Molecular dynamics simulation investigation

The study of molecular dynamics simulation enables us to delineate the atomic structures of the protein as well as its dynamic and structural characteristic changes occurring in real time [136]. The stability effect of the small molecules and monoclonal antibody were analysed and studied for 100ns, for each of the NRAS and BRAF proteins. This analysis was done based on RMSD, RMSF, radius of gyration, and intermolecular interaction (hydrogen bonds, salt bridges, pi-cation and pi-pi stacking).

In order to analyse protein stability, RMSD analysis was used to examine the conformational changes of protein atoms in relation to the reference frame. As it can be seen from the graphs, the conformation stability for Alpelisib shows that the ligand stabilizes after approximately 30 ns with each of the BRAF (5C9C) and NRAS (6ZIZ) respectively (Figure 5.10 (a, b), first row). Further, in case of Obatoclax, there is

stabilization with BRAF after approximately 30 ns, but there is rigorous deviation of ligand with NRAS (Figure 5.10 (a, b), second row). For Cetuximab, the deviation of the ligand is observed to be same as the protein within the range of 2Å to 3.5Å (for BRAF Figure 5.10 (a), third row) and 2Å to 3.8Å (for NRAS Figure 5.10 (b), third row); these findings again demonstrate the stabilization of protein-ligand complex conformation.

The RMSF analysis helps characterize the local fluctuation along the protein chain. It enables us to study the contributions of individual amino acids, if they are stable or fluctuating during the simulation. The terminal and loop regions of the protein are extremely mobile. Here, the amino acid fluctuations in each protein-BRAF and protein-NRAS bound with the drug have been studied which demonstrate the protein residues fluctuation (Figure 5.11 (a, b)). Thus, as represented by the graphs, it can be seen that Alpelisib bound BRAF and NRAS shows less RMSF values or fluctuations as compared to Obatoclax and Cetuximab bound proteins (Figure 5.11 (a, b)).

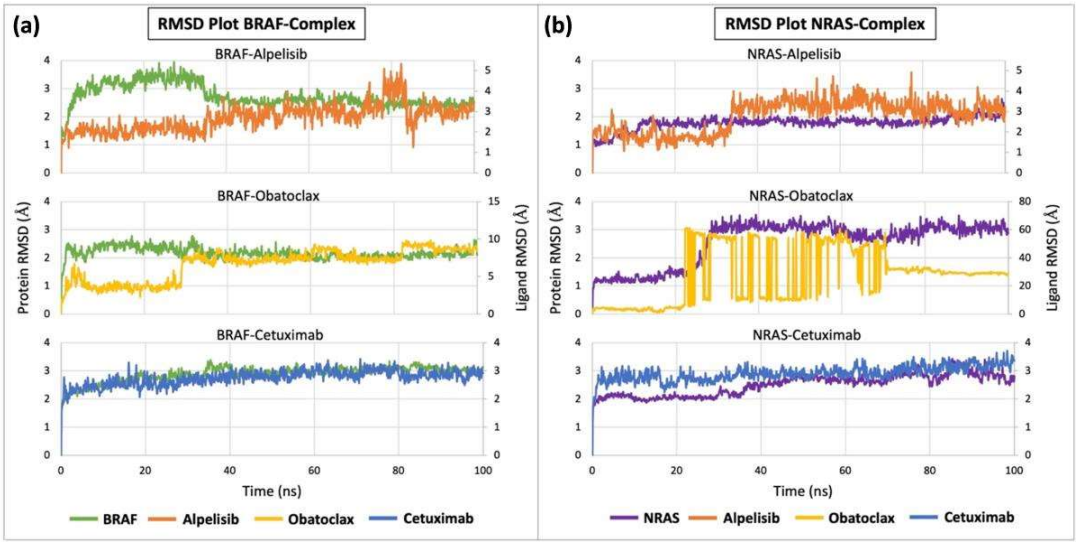

Figure 5.10 RMSD analysis of protein-ligand complexes throughout the simulation trajectories. The RMSD of the protein is plotted on the left Y-axis, and the RMSD of the ligand is plotted on the right Y-axis, while simulation time (ns) is on the X-axis. (a) The RMSD of BRAF protein along with three selected ligands (Alpelisib, Obatoclax, and Cetuximab). (b) RMSD of NRAS protein along with three selected ligands (Alpelisib, Obatoclax, and Cetuximab).

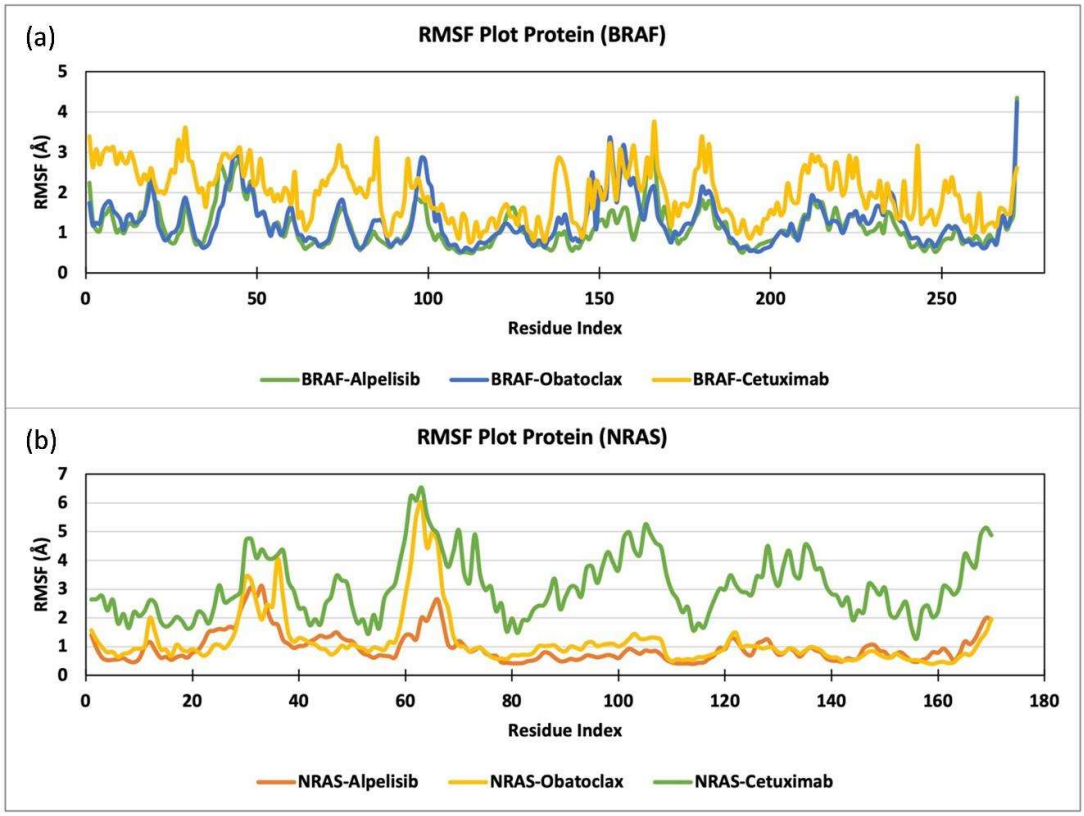

Figure 5.11 RMSF analysis of protein C-alpha atoms for each residue throughout the simulation time, depicting the flexibility of the protein residues in the complexes. (a) RMSF plot of the BRAF complex, and (b) RMSF plot of the NRAS complex.

Intermolecular interactions such as hydrogen bonds (H-bonds) and salt bridges enable proteins to maintain their conformation with the ligand and therefore, indicates the stability of the ligand-protein complex. Also, pi-cation and pi-pi stacking contribute the binding of two molecules. We measured these interactions between the protein-ligand and protein-protein complexes in this study, to see the binding stability of complexes throughout the simulation trajectories of 100ns each. It can be seen in the Figure 5.12 (a, d) that Alpelisib shows more hydrogen bond interaction with NRAS as compared to BRAF, while Obatoclax shows very minimal hydrogen bond interactions with either of the proteins (Figure 5.12 (b, e)). Also, in case of BRAF Alpelisib and Obatoclax shows mostly pi-pi stacking types of interaction as compared to NRAS. We also observe that, Cetuximab also

shows continuous stable hydrogen bonding throughout simulation with both BRAF and NRAS shown in Figure 5.12 (c, f). Furthermore, Radius of Gyration (Rg) measures the 'compactness' of the protein, which is calculated for the protein backbone and is equivalent to its principal moment of inertia. The Rg of Alpelisib was found to be (i) in the range of 19.1Å - 20.5 Å for BRAF and 15.2 Å – 15.9 Å for NRAS (Figure 5.13 (a, d)), (ii) for Obatoclax, Rg was 19Å – 19.9 Å for BRAF and 15.2 Å – 16.1 Å for NRAS (Figure 5.13 (b, e)), and (iii) for Cetuximab, Rg was 18.6 Å – 19.4 Å for BRAF and 15.2Å - 16 Å for NRAS (Figure 5.13 (c, f)) respectively, indeed the Rg determines the compactness of the protein during the simulation process. The overall radius of gyration analysis of BRAF and NRAS shows that these two proteins were not able to maintain their proper folding, and shows large changes during the simulation, except in case of the BRAF-Alpelisib complex.

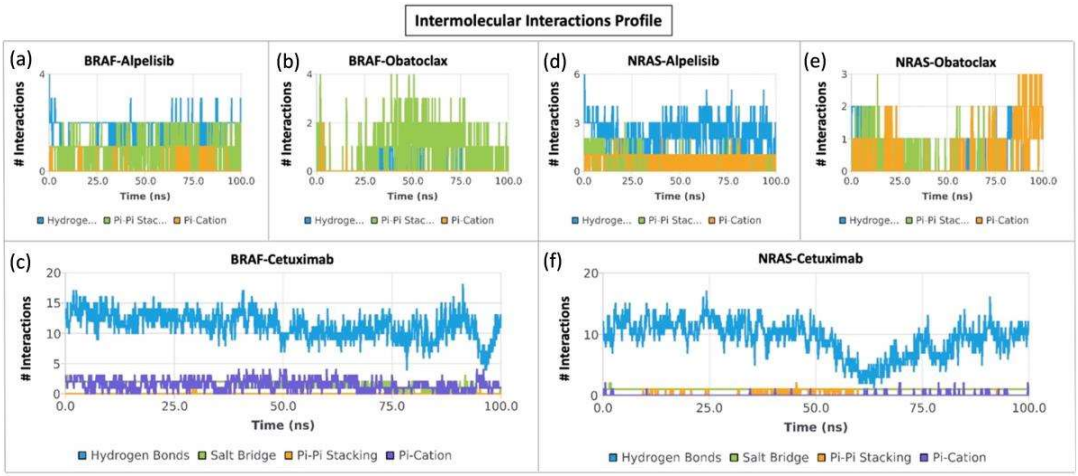

Figure 5.12 Intermolecular interactions between the protein and ligand were calculated during the simulation to measure the stability of the complexes. The interaction profiles of the following complexes are depicted in the figure: (a) BRAF-Alpelisib, (b) BRAF-Obatoclax, (c) BRAF-Cetuximab, (d) NRAS-Alpelisib, (e) NRAS-Obatoclax, and (f) NRAS-Cetuximab.

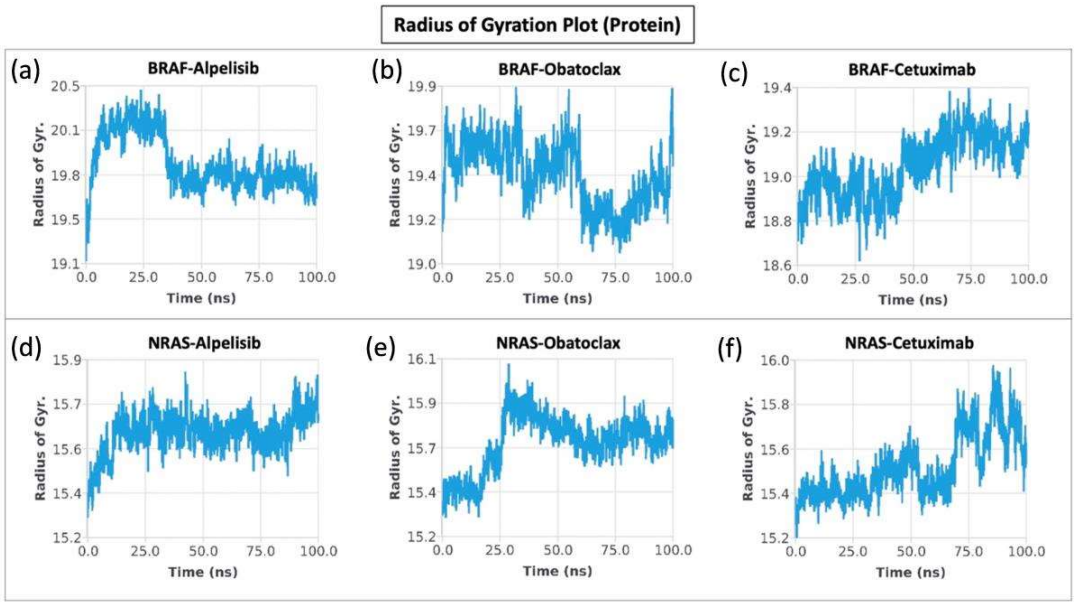

Figure 5.13 Radius of gyration plot of the protein-ligand complexes illustrating changes in protein compactness throughout the simulation. In this figure, the radius of gyration of the protein is depicted for each complex: (a) BRAF-Alpelisib complex, (b) BRAF-Obatoclax complex, (c) BRAF-Cetuximab complex, (d) NRAS-Alpelisib complex, (e) NRAS-Obatoclax complex, and (f) NRAS-Cetuximab complex.

5.4 Discussion

The mechanism of the melanoma process has been the subject of numerous studies over the past few years, nevertheless the incidence and death rates of melanoma are continuously rising globally. In contrast to earlier studies that only focused on several genes or a single database or cohort, our work chose 3 high-quality gene profile datasets from several universal databases in order to comprehensively investigate the driven-genes and biological pathways involved in melanoma. Numerous types of cancer exhibit the episodic natural phenomena of spontaneous regression of malignant tumours including melanoma, in which the tumour is totally eliminated permanently without toxicity effects. In-depth research into this unusual natural occurrence may shed light on the feasibility of reproducing such a

regression process on a clinical melanoma tumour in the patient context, without significant toxicity nor negative side effects.

In our present study, we have investigated the spontaneously regressing melanoma microarray data and endeavoured to explore the driver genes and biological pathways involved in melanoma regression. Finally, we have identified 344 *upregulated* and 160 *downregulated* DEGs. The enrichment analysis shows that the significant *downregulated* genes were associated with the mitotic cell cycle process, DNA replication, G2/M checkpoints, G2/M DNA damage checkpoint signaling pathways, which may play crucial role in melanoma regression (Figure 5.4(b)). We have also identified that 27 melanoma driving genes are either involved in (i) cell cycle regulation and in cell cycle division or (ii) signaling pathway activation, mainly as MAPK and PI3K/AKT signaling pathways. The enrichment analysis shows that these 27 melanoma regression driving genes were enriched in Chk1/Chk2 (cds1) mediated inactivation of cyclin B: Cdk1complex, cell cycle, Mitotic, Cyclin a/B1 associated event during G2/M transition, cell cycle checkpoints and phosphorylation of protein involved in the G2/M transition by cyclin A: Cdc2 complexes (Figure 6). Moreover, Figure 5.14 illustrates the cell cycle progression showing some of the driving genes or active sites, for example: SOX2 genes interact with cyclin D to control cell cycle [137], the downregulation of FGFR2 gene arrests the cell cycle at G1 phase [138], STAT1 gene directly interacts with cyclin D1/Cdk4 to arrest the cell cycle at G1 phase [139], PSMB9 gene actuates the M/G1 transition phase [140], TFAP2A act on the resting phase of cell cycle [141], and so on.

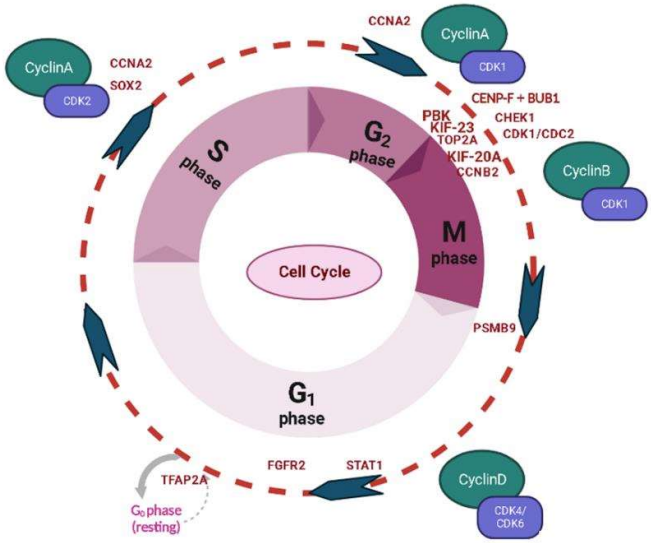

Figure 5.14 The cell cycle process under the genes that actuate the spontaneous regression of malignant melanoma tumour.

Our investigation indicates that most of the DEGs were involved in either MAPK or PI3K/AKT signaling pathways or both. Both these proteins were activated by RAS family proteins BRAF and NRAS, which are oncogenes, very frequent in malignant melanoma [142–144]. In this study we have used modelling techniques like docking and molecular dynamics approach to examine the inhibitory role of Alpelisib and Cetuximab on NRAS and BRAF genes, which are the most crucial genes in melanoma regression via MAPK and PI3 kinase signaling pathways. Utilizing MD Simulation studies, we have investigated the atomic/molecular level changes in the structural and physical aspects of the protein and even the individual amino acids.

Thereafter, we analyzed the changes in the behavioural pattern of the protein and stability of interaction of the ligand with each of the proteins using by means of computational analysis using RMSD, RMSF, Rg, SSE, H-Bonds, Salt-interactions, and Protein-Ligand interactions. The conformational changes between the different drugs interacting with our proteins have been visualized using a time-series analysis of the RMSD

duration over the simulation. The plot thus generated for the protein-ligand complexes annotates the functional variations of the individual binding site residues, thereby depicting the changes in ligand binding behavior. Indeed, the RMSF calculations for each amino-acid residue of the protein annotates the measurements of local fluctuations throughout the period of simulation (100ns). The changes in peak denotes fluctuations while the green extended lines depict stable interactions with the respective amino-acid residues of the protein [Figure A4 (a, b, c, d)]. The radius of gyration (Rg) is an essential component in determining protein stability and protein folding patterns. A higher Rg value indicates that protein unfolding has occurred, and on the contrary a lower Rg value means that there has been proper folding or compactness of the protein.

By studying several interactions, such as H-bond and salt-bridges, the thermodynamic stability and folding pattern of the protein can be analysed, since these bonds play an important role in the formation of secondary structure of proteins. Any changes in the bonding pattern could result in misfolding, thus indicating structural and thereafter functional impairments in the protein. This would further lead to alterations in the molecular pathways.

Through the studies conducted, we have found that individually the drugs Alpelisib and Cetuximab, but not Obatoclax show stable binding with both NRAS and BRAF. Thus, the two drug (Alpelisib and Cetuximab) can be re-purposed readily for inhibiting progression or inducing regression of melanoma, since these drugs are already in use for treatment of several cancers (but not melanoma) and that these drugs are also FDA approved.

We have also found that the binding sites of Cetuximab and Alpelisib are not the same since Alpelisib is a small molecule (semi synthetic chemical drug) binding to the allosteric

site of protein whereas Cetuximab is a monoclonal antibody binding to the variable region of the protein (antigen). Therefore, these two drugs can be used as a combination, which would allow for better efficiency and so a lower dose quantity. Furthermore, the active pharmacophore moieties of Alpelisib and Cetuximab could be combined in a macromolecule which may be administered in a signal dosing.

5.5 Conclusion

Our current study suggests that the Ras-Raf-MEK-ERK and PI3K-PKT signaling pathways is activated in the majority of melanoma tumours. This activation occurs through either NRAS or BRAF mutations, both of which arise early during melanoma pathogenesis and are preserved throughout tumour progression. Our present investigation points out that the pharmaceuticals Alpelisib and Cetuximab may be potential repurposed drugs which can target both the pathways together instead of targeting one gene at a time, and this conjoint activity may enable therapeutic duplication of the process of spontaneous regression phenomenon on malignant melanoma patients.