© 2018 IJSRST | Volume 4 | Issue 5 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology

Molecular Characterization of human KRAS protein using Bioinformatics tools

Vinod P Sinoorkar, Smita S Kashetti*, Akanksha S Degaonkar, Vedika V Bhola
Walchand Centre for Biotechnology, Solapur, Maharashtra, India

ABSTRACT

Cancer is an abnormal growth of cells caused by multiple changes in gene expression leading to dysregulated balance of cell proliferation and cell death and ultimately evolving into a population of cells that can invade tissue and metastasize to distant sites , causing significant morbidity and if untreated death of host. Cancer develops due to genetic damage to DNA and epigenetic changes. These changes affect the normal functions of the cell, including cell proliferation, programmed cell death (apoptosis) and DNA repair. GTPase Kras also known as V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog and Kras is a stable protein having length of about 189 a.a and molecular weight about 21655.83, in humans it is encoded by the KRAS gene. The protein product of the normal KRAS gene performs an essential function in normal tissue signalling, and the mutation of a KRAS gene is an essential step in the development of many cancers. Hence the present investigation was carried out to understand the molecular features of Kras protein by retrieving the protein sequence information from major protein database of bioinformatics.This analysis forms the base for the detailed understanding of molecular mechanism of Kras during onset of various cancers and further work on this, can pave a new dimension for the treatment of cancer.

Keywords : Cancer, KRAS, Bioinformatics, Mutation, Molecular features

I. INTRODUCTION

Cancer is an abnormal growth of cells caused by multiple changes in gene expression leading to dysregulated balance of cell proliferation and cell death and ultimately evolving into a population of cells that can invade tissue and metastasize to distant sites, causing significant morbidity and if untreated death of host . There are number of cancers are frequently occurs in a society few of them includes the lung cancer, oral cancer, liver cancer and leukemia. This type of cancer is caused due to inherited factor [1], smoking, tobacco chewing, oncogene activation or mutation in functional genes [2] etc. There are several molecules are included in a cancer but the present study includes the kras protein which is encoded by Kras gene. The protein product of the normal KRAS gene performs an essential function in normal tissue signalling, and the mutation of a KRAS gene is an essential step in the development of many cancers. Like other members of the Ras family, the Kras protein is a GTPase and is an early player in many signal transduction pathways. Kras is usually tethered to cell membranes because of the presence of an isoprene group on its C-terminus. Cytogenetic location the Kras is present at the short arm of chromosome no.12 [3] at position 12.1. Molecular location on chromosome 12:- 25,204,788 to 25,250,930bp. The kras molecule involves in activation of the RAS/RAF/MEK/MAPK pathway occurs frequently in all types of cancer [4]. The GTPase ras molecule auto phoshorylates the raf molecule and its downstream molecules even in the absence of a ligand leads to cause a cancer. Due to technological advances in the genomics disciplines, cancer research is going through a revolution. The technological advances that lead to the post-genome era have allowed molecular biologists to make meticulous studies on the DNA (genome), the mRNA (transcriptome) and the protein sequences (proteome).

Initiatives that intend to describe cancer in a global dimension are providing an opportunity for investigators to have more useful data to analyse and integrate in novel ways. Cancer bioinformatics deals with the organization and analysis of the data so that important trends and patterns can be identified – the ultimate goal being the discovery of new therapeutic and/or diagnostic protocols for cancer [5]. There are many newly available resources in these areas that may be unfamiliar to most cancer researchers wanting to incorporate bioinformatics tools and analyses into their work, and also to bioinformaticians looking for real data to develop and test algorithms. This review reveals the inter dependence of cancer research and bioinformatics and highlight the most appropriate and useful resources available to cancer researchers. These include not only public databases, but general and specific bioinformatics tools which can be useful to the cancer researcher. The primary foci are function and structure prediction tools of protein genes. The result is a useful reference to cancer researchers and bioinformaticians studying cancer alike [6].With this detailed understanding of cancer and literature survey, the present study was carried out to understand the molecular features and role of KRAS in various types of cancers using bioinformatics approach.

II. MATERIALS AND METHODS

A. Retrieval of protein sequence information of Kras

For the study of Kras molecule, its amino acid sequence was retrieved from the major protein sequence databases like Uniprot KB and NCBI protein database. The sequence obtained was stored in Fasta format with its accession number.

B. Analysis of Physicochemical properties

The analysis of physicochemical properties of normal and mutated Kras was done by using protein prediction tool Protparam which gives a detailed information of protein like isoelectrical point, theoretical pI, extension coefficient, half-life, instability index, Aliphatic index, Grand average, hydropathicity, etc.

C. Secondary structure prediction of Kras

The secondary structure prediction of normal and mutated Kras was carried out by using online secondary structure prediction tool SOPMA, which gives the information of Alpha helix, beta sheets, extended strands and random coils.

D. Prediction of Tertiary structures of Kras

The Tertiary structure of normal and mutated Kras was obtained by using SWISS Model tool by selecting the template with maximum homology and with optimised parameters. The obtained structure was stored in pdb format for visualization.

E. Visualization of tertiary structure of Kras

The predicted Tertiary structure of normal and mutated Kras was visualized by using structure visualization tool Rasmol. Visualization was done using different models and formats to understand structural features of normal and mutated Kras.

III. RESULTS AND DISCUSSION

A. Retrieval of protein sequence information of Kras

For the study of normal and mutated Kras molecule, its amino acid sequence was retrieved from the major protein sequence databases like Uniprot KB and NCBI protein database. The sequence obtained was stored in Fasta format with its accession number. The length of the sequence was found to be 189a.a as shown in figure 2(a) (b).

Accession no.-p01116

>sp|P01116|RASK_HUMAN GTPase KRas OS=Homo sapiens GN=KRAS PE=1 SV=1MTEYKLVVVGAGGVGKSALTIQLIQNHFVDE YDPTIEDSYRKQVVIDGETCLLDILDTAGQEEYSAM RDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKR

VKDSEDVPMVLVGNKCDLPSRTVDTKQAQDLARS
YGIPFIETSAKTRQRVEDAFYTLVREIRQYRLKKISK
EEKTPGCVKIKKCIIM

Figure 2(a). Showing Protein sequence of normal human KRAS

>sp|P01116|RASK_HUMAN GTPase KRas OS=Homo sapiens GN=KRAS PE=1 SV=1MTEYKLVVVGAVGVGKSALTIQLIQNHFVDE
YDPTIEDSYRKQVVIDGETCLLDILDTAGQEEYSAM
RDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKR
VKDSEDVPMVLVGNKCDLPSRTVDTKQAQDLARS
YGIPFIETSAKTRQRVEDAFYTLVREIRQYRLKKISK
EEKTPGCVKIKKCIIM

Figure 2(b) Showing Protein sequence of mutated human KRAS

B. Analysis of Physicochemical properties

The analysis of physicochemical properties of normal and mutated Kras was done by using protein prediction tool i.e. Protparam which gives a detailed information of protein i.e. isoelectrical point, theoretical pI, extension coefficient, half-life, instability index, Aliphatic index, Grand average, hydropathicity, etc., were predicted, calculated and tabulated as shown in table 1(a),(b).

Table 1(a) Showing Physicochemical properties of normal Kras

Properties	Values
No. of amino acid	189
Molecular weight	21655.83
Theoretical pI	6.33
Total no. of negative charged residue (Asp+Glu)	29
Total no .of positively charged residue(Arg +Lys)	28
Atomic composition	
1)Carbon(C)	953
2)Hydrogen(H)	1533
3)Nitrogen(N)	261
4)Oxygen(o)	293
5)Sulphur(s)	10
Total no .of atoms	3050
Extinction coefficient No.	13660
Half life	30hrs
Instability index	Stable (38.56)
Aliphatic index	85.03
Grand average of a hydropathicity	-0.432

Table 1(b) Showing Physicochemical properties of mutated Kras

Properties	Values
No. of amino acid	189
Molecular weight	21697.91
Theoretical pI	6.33
Total no. of negatively charged residue (Asp+Glu)	29
Total no .of positively charged residue(Arg+Lys)	28
Atomic composition	
1)Carbon(C)	956
2)Hydrogen(H)	1539
3)Nitrogen(N)	261
4)Oxygen(o)	293
5)Sulphur(s)	10
Total no.of atoms	3059
Extinction coefficient No.	13660
Half life	30hrs
Instability index	Stable(37.46)
Aliphatic index	86.56
Grand average of a hydropathicity	-0.407

C. Secondary structure predictions of Kras

Secondary structure prediction of normal and mutated Kras was carried out by using SOPMA and all the secondary structural elements like alpha helix,

beta sheets, random coils and extended strands were predicted as shown in table 2 (a) (b)and figure 3(a) (b).

Table 2 (a) Showing secondary structure information of normal Kras

Structural component	Residues	Percentage
Alpha helix (Hh)	81	42.86%
3_{10} helix (Gg)	0	0.00%
Pi helix (Ii)	0	0.00%
Beta bridge (Bb)	0	0.00%
Extended strand(Ee)	39	20.63%
Beta turn (Tt)	19	10.05%
Bend region(Ss)	0	0.00%
Random coil (Cc)	50	26.46%
Ambiguous States	0	0.0%
Other States	0	0.0%

Table 2 (b) Showing secondary structure information of mutated Kras

Structural component	Residues	Percentage
Alpha helix (Hh)	79	41.80%
3_{10} helix (Gg)	0	0.00%
Pi helix (Ii)	0	0.00%
Beta bridge (Bb)	0	0.00%
Extended strand(Ee)	43	22.75%
Beta turn (Tt)	17	8.99%
Bend region(Ss)	0	0.00%
Random coil (Cc)	50	26.46%
Ambiguous States	0	0.00%
Other States	0	0.00%

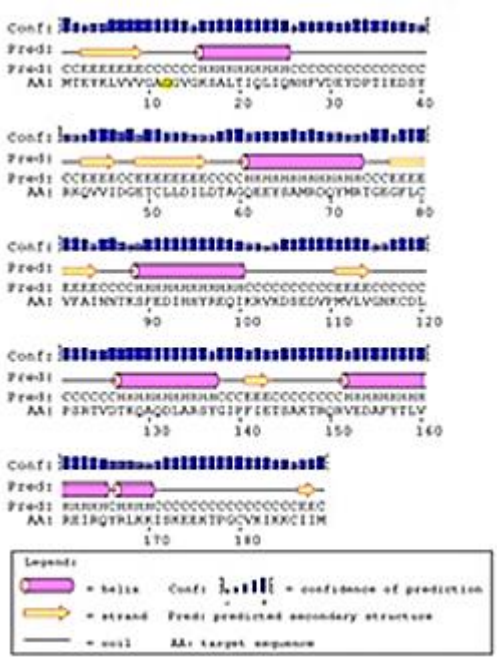

Fig 3(a) showing secondary structure information of normal Kras

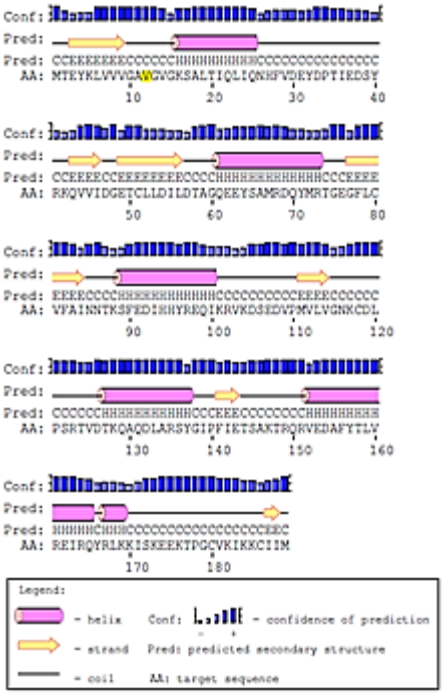

Fig 3(b) showing secondary structure information of mutated Kras

D. Prediction Tertiary structures of Kras

The Tertiary structure of normal and mutated Kras was obtained by using SWISS Model tool by selecting the template with maximum homology and with optimised parameters. The obtained structure was stored in pdb format for visualization. The details of

template selected for structure prediction was as shown on table 3(a) (b).

Table 3 (a) Showing Template details for normal Kras

Name	Title	identity	Oligostate
1)6q21.1.B	C-H-Ras p21 protein catalytic domain	92.40	Homo tetramer
2)6q21.1.C	C-H-Ras p21 protein catalytic domain	92.40	Homo tetramer
3)1aa9.1.A	C-HA-RAS	92.40	monomer
4)1ioz.1A	Transforming protienp21/H-RAS-1	92.40	monomer
5)1q21.1.A	C-H-RASp21 protein catalytic Domain	92.40	Homodimer
6)3con.1.A	GTPase N-RAS	93.49	Monomer

Table 3 (b) Showing Template details for Mutated Kras

Name	**Title**	**identity**	**Oligostate**
1)1ioz.1.A	Transforming protienp21/H-RAS-1	91.81	monomer
2)1aa9.1.A	C-HA-RAS	91.81	monomer
3)6q21.1.B	C-H-RASp21 protein catalytic Domain	91.81	Homo tetramer
4)6q21.1.C	C-H-RASp21 protein catalytic Domain	91.81	Homo tetramer
5)6q21.1.D	C-H-RAS p1 protein catalytic Domain	91.81	Mono tetramer
6)3con.1.A	GTPase N-RAS	92.90	monomer

E. Visualization of tertiary structure of Kras

The predicted Tertiary structure of normal and mutated Kras was visualized by using structure visualization tool Rasmol. Visualization was done using different models and formats to understand structural features of Kras. The various models were represented as shown in figures 4(a)(b).

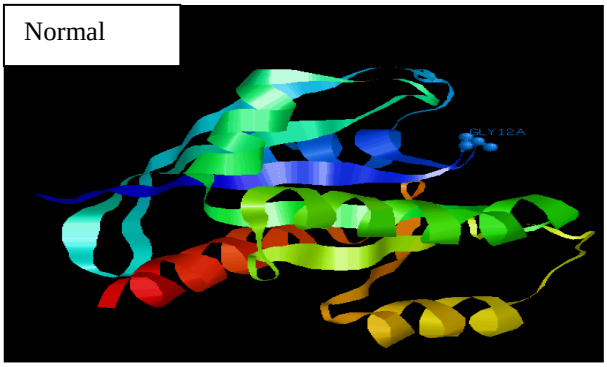

Figure 4(a). 3D structure of normal kras highlighting residue with position

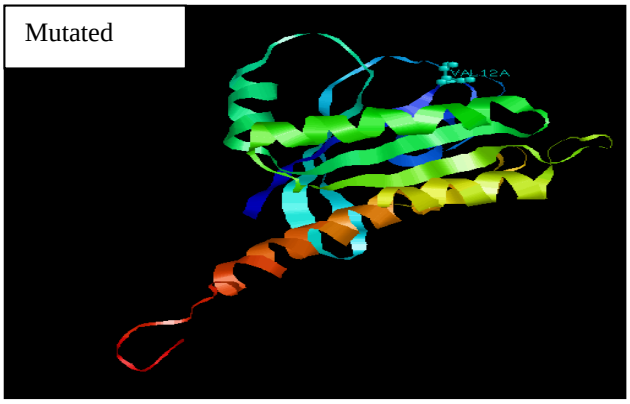

Figure 4(b). 3D structure of mutated kras highlighting residue with position

IV. Conclusion

Cancer is an abnormal growth of cells caused by multiple changes in gene expression leading to dysregulated balance of cell proliferation and cell death and ultimately evolving into a population of cells that can invade tissue and metastasize to distant sites , causing significant morbidity and if untreated death of host. Cancer develops following genetic damage to DNA and epigenetic changes. These changes affect the normal functions of the cell, including cell proliferation, programmed cell death (apoptosis) and DNA repair. As more damage accumulates, the risk of cancer increases. In order to make an impact on cancer, we must understand the molecular exhibit multiple genetic lesions including mutations activating the dominant cellular proto-oncogenes as well as the accumulation of multiple molecular abnormalities overtime. GTPase KRas also

known as V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog and KRAS, is a protein that in humans is encoded by the KRAS gene. The protein product of the normal KRAS gene performs an essential function in normal tissue signalling, and the mutation of a KRAS gene is an essential step in the development of many cancers. Hence the present investigation was carried out to understand the molecular features of Kras protein by retrieving the protein sequence information from major protein database, analysis of physicochemical parameters using Protparam tool. The analysis emphasise that Kras has 189 residues, high molecular weight protein with slightly negative charge and highly stable. The secondary structure of both normal and mutated Kras proteins predicted by SOPMA, are dominated by alpha helix followed by random coils and extended strands. The tertiary structures of normal and mutated Kras were predicted by Swiss model tool and visualized by Rasmol with the residue replacement highlighted in both. This preliminary analysis forms the base for the detailed understanding of molecular mechanism of Kras during onset of various cancers and further work on this, can have a new dimension for the treatment of cancer.

V. Acknowledgment

The authors thanks to the Walchand College of Arts and Science and department of Biotechnology (Solapur, Maharashtra, India) for providing a constant encouragement, anonymous references and suggestions to project.

VI. REFERENCES

[1]. Yang, "Genetic susceptibility to lung cancer and co-morbidities "Journal of Thoracic Disease 2013.

[2]. Herbst, “Lung cancer" New England Journal of medicine, 2008, 359.

[3]. Alex A. Adjei, “K-ras as a Target for Lung Cancer Therapy”, Journal of Thoracic Oncology, June 2008 Vol. 3, No. 6, Supplement 2, S16 - S163.

[4]. Jill E. Larsen and John D. Minna, “Molecular Biology of Lung Cancer: Clinical Implications”, Clin Chest Med, 2011, 703–740.

[5]. Beatriz Stransky and Pedro Galante, “Application of Bioinformatics in Cancer Research”, Springer, 2009.

[6]. Disuke kihara, “Bioinformatics research for cancer research with an emphasis on gene function and structure prediction tools”, Cancer informatics, 2007.

[7]. Nishtha Singh, Sonal Upadhyay, et al , “In silico Analysis of Protein” J Bioinform, Genomics, Proteomics, 07 September 2016,1-7.

[8]. Anshul Tiwari, Monika Sharma, et al, “In silico Characterization of Retinal S-antigen and Retinol Binding Protein-3: Target against Eales’ Disease”Int.J. Bioautomation, 2014, 18(4), 287-296.

[9]. Razumilava, “Classification, diagnosis and management of cholangio carcinoma”, 2013.

[10]. Kranenburg O “The KRAS oncogenes: past, present and future, 2005.

[11]. Merck Manual Professional Edition, "Lung Carcinoma: Tumors of the Lungs", Online edition, Retrieved 15 August 2007.

[12]. Jakopovic, et al; "Targeting the epigenome in lung cancer: expanding approaches to epigenetic therapy", 2013.

[13]. D Behera, “Lung Cancer in India NewDelhi”, 2008.