International Journal of Genomics and Proteomics
ISSN: 0976-4887 & E-ISSN: 0976-4895, Volume 4, Issue 1, 2013, pp.-76-83.
Available online at http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000227

DISORDERED REGIONS IN CANCER PROTEINS EXHIBIT POTENTIAL DRUG BINDING PROPERTIES

RAJU K.D.[2], RATHANKAR N.[1], JAGADEESH K.D.[3] AND NAGENDRA H.G.[3]*

[1]Apsara Innovations Pvt. Ltd., Bangalore- 560 043, Karnataka, India.
[2]Quintiles, Marathahalli, Bangalore- 560 037, Karnataka, India.
[3]Department of Biotechnology, Sir M. Visvesvaraya Institute of Technology, Bangalore- 562 157, Karnataka, India.
*Corresponding Author: Email- nagshaila@sirmvit.edu

Received: July 30, 2013; Accepted: September 03, 2013

Abstract- Intrinsically disordered proteins are those that fail to self-fold into fixed 3D geometries, and are found to be associated with innumerable abnormalities like cancer, diabetes, cardiovascular, neurodegenerative disorders etc. Based on the predictions that, about 79% of cancer proteins are disordered, an investigation was carried out to elucidate the importance of these sequences and their roles in drug binding and disease control. A search through Drug bank database revealed 210 ligand molecules associated with 253 cancer proteins. These amino acid sequences were examined with tools such as DisEMBL and SEG, to delineate the composition of their disordered regions. 127 of these 253 sequences were having 30 or more consecutive residues predicted to be "Disordered Regions". Structural homologues for 52 of these 127 protein sequences revealed quality global alignments, and 25 of these had their active site / ligand binding site in the "disordered regions", as deciphered by Accelrys Discovery Studio 2.5. Further, PDB analysis revealed that these sequences were co-crystallized with drug molecules, and these ligands were bound to the respective disordered regions in the active site. Using this as basis, docking studies were performed for 11 cancer proteins involved in key pathways, with respective drug molecules to ascertain the nature of drug-receptor interactions. The results indicated that molecules mentioned in drug bank towards treating cancer, bound to their respective receptors through the disordered regions in them, highlighting that these regions could play potentially important roles in ligand Pharmacophore interactions.

Keywords- Cancer, Disordered Regions, Docking Studies, Pharmacophore, Drug binding, Neurodegenerative disorders

Citation: Raju K.D., et al. (2013) Disordered Regions in Cancer Proteins Exhibit Potential Drug Binding Properties. International Journal of Genomics and Proteomics, ISSN: 0976-4887 & E-ISSN: 0976-4895, Volume 4, Issue 1, pp.-76-83.

Introduction

Cancer is a term used for diseases in which abnormal cells divide without control and are able to invade other tissues [1]. Cancers are caused due to various reasons, and these include - carcinogens, age, genetic mutations, immune system, diet, tobacco smoke, radiations, viruses etc [2]. National cancer Institute has grouped this disease into five major categories which include Carcinoma, Sarcoma, Leukemia & Lymphoma, Myeloma and Central nervous system cancers. Oncogenes and Tumor-suppressor genes encode proteins that play key roles in cancer development [3], cell proliferation and apoptosis [4]. Cell cycle regulatory proteins such as cyclins and cyclin-dependant kinases get over-expressed in cancer condition [5]. Similarly, other molecules like Cellular Retinoic Acid-Binding Protein 1 & 2, Cytochrome P450, Serum Albumin, GMP reductase 1 & 2, and GMP synthase are also expressed during cancer [6-10].

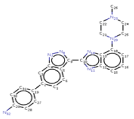
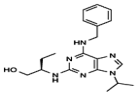
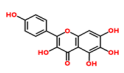
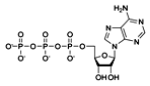
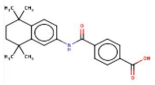
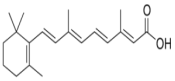
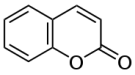
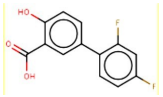
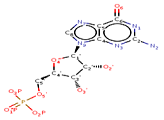
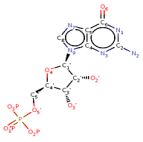
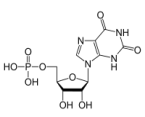
Proteins, or large segments of proteins, that lack a well-structured three-dimensional fold or to those that fail to self-fold into fixed 3D structure are termed as disordered proteins [11, 12]. Several disordered proteins are shown to be associated with human diseases, such as cancer, cardiovascular disease, amyloidoses, diabetes, neurodegenerative diseases etc [13]. Interestingly, studies carried on Human Cancer Associated Proteins in Swissprot dataset reveals that 79(+/-5) % of them contain 30 or more consecutive residues that are predicted to be disordered [14]. For example P53, a tumor suppressor protein which is implicated in more than 50% of cancers, is 37% disordered[15]. There is ample evidence to suggest that such unstructured molecules are essential for basic cellular functions. The intrinsic lack of structure appears to confer functional advantages to the protein such as, ability to bind to diverse targets, control over the thermodynamics of the binding process etc [16]. It has been indicated that a detailed investigation on disordered proteins could enable identification of potential targets for Structure Based Drug Design (SBDD) [17], which stress on the transition from disordered to ordered conformation through drug stimulation [18]. Studies on *Mycobacterium tuberculosis* have reported that intrinsically disordered proteins such as FtsW (Rv2154c), GlmU (Rv1018c) and Obg (Rv2440c) could be potential drug targets [19]. Hence, an attempt has been made in this work to analyze the disordered regions in disease protein sequences related to cancer as drug binding targets, in order to elucidate their roles in disease inhibition and suggest them as newer targets for therapy.

Materials and Methods

A formidable search was done through Drug bank [20,21] database, to retrieve all the drug molecules used to treat various types of cancer along with their respective target protein sequences. 210 drug molecules and 253 protein sequences were obtained in the process, which served as primary database for further investigations. The protocol followed is illustrated in [Fig-1]. These 253 sequences were then subjected to analysis using tools such as DisEMBL [22] and SEG [23], for delineation of disordered regions. This resulted in the 127 disordered sequences that contained stretches greater than 30 or more consecutive residues. Structural homologues were searched for these 127 sequences across PDB [24] database using the standard criteria (>/= 30% identity, e-value </= 0.001). 52 among these 127 sequences had structural counterparts along their lengths with convincing (>50 %) identities. Detailed analysis of crystal structures of these structural proteins (complexed with their respective ligands) revealed that, only for 25 of these 52 cancer related sequences, their active sites and / or ligand binding sites were part of the disordered regions, as predicted by Accelrys Discovery studio 2.5 [25]

Thus, using this as the basis cancer proteins involved in key pathways, namely Cell Division Protein Kinases 2, 5, 6 & 7, Cellular Retinoic Acid binding protein 1 & 2, Cytochrome P450 2A6, Serum Albumin Precursor, GMP reductase 1 & 2, and GMP Synthase [26-42] were investigated via docking studies (with their respective drug molecules) using FlexX (3.7) [43], to appreciate the roles of disordered regions and their prospective binding efficacy to commercial drug molecules. The details about these 11 cancer molecules and their corresponding drugs are provided in [Table-1] and [Table-2] respectively.

Table 1- Details of the 11 cancer proteins and their Ligands under study

Sl no	Protein Name	Uniprot ID and KEGG Pathway number	PDB ID (No. of residues)	Ligand in PDB (no. of atoms)	Structure of the Ligand	Disordered regions in the protein (No. of residues contained)
1	Cell division protein kinase 2	P24941 hsa:05200	3EZV (298)	Indazole Inhibitor 9 (57) C25 H25 N7 IC50: 1040nM		1-47 & 283-298 (63)
2	Cell division protein kinase 5	Q00535 hsa:05030	1UNL (292)	R-Roscovitine (52) C19 H26 N6 O IC50: 160-180nM Kd: 1900-1900000nM		1-45, 56-74, 219-231 & 277-292 (90)
3	Cell division protein kinase 6	Q00534 hsa:05212	1XO2 (326)	Tetrahydroxy-flavone (31) C15 H10 O6 IC50: 850nM		28-60, 69-82, 97-116, 145-156, 192-200, 218-228, 253-262 & 311-326 (114)
4	Cell division protein kinase 7	P50613 has:04110	1UA2 (346)	ATP (47) C10 H16 N5 O13 P3		1-35, 40-55, 159-173, 230-239 & 319-346 (100)
5	Cellular Retinoic Acid-Binding Protein 1	P62964 bta:282201	2CBR (136)	Retinobenzoic Acid (49) C22 H25 N O3 IC50: 140nM		1-12, 36-78, 87-105 & 117-128 (83)
6	Cellular Retinoic Acid-Binding Protein 2	P29373 hsa:1382	3CBS (137)	Retinoic Acid (46) C20 H24 O3 IC50: 58nM Kd: 58nM		1-11, 35-81 & 96-131 (92)
7	Cytochrome P450 2A6 beta	P11509 hsa:1548	1Z10 (494)	Coumarin (17) C9 H6 O2 270nM		29-38, 97-108, 118-144, 184-192, 253-265, 327-348, 359-390 & 464-494 (148)
8	Serum Albumin	P02768 hsa:213	2BXE (609)	5-(2,4-Di Fluoro Phenyl) -2 -Hydroxy -Benzoic Acid (26) C13 H8 F2 O3		127-139, 438-468 & 488-519 (70)
9	GMP Reductase 1	P36959 hsa:2766	2BWG (345)	Guanosine Monophosphate (38) C10 H14 N5 O8 P		1-48, 287-300 & 325-345 (81)
10	GMP Reductase 2	Q9P2T1 has:00230	2A7R (348)	Guanosine Monophosphate (38) C10 H14 N5 O8 P		1-49, 179-190, 248-259, 289-301 & 324 -348 (107)
11	GMP synthase [glutamine-hydrolyzing]	P49915 has:00983	2VXO (693)	Xanthosine Monophosphate (38) C_{10} H_{14} N_4 O_9 P		230-237, 302-341, 386-397, 402-418, 426-440, 531-538, 565-574, 625-633, 644-658 & 670-693 (141)

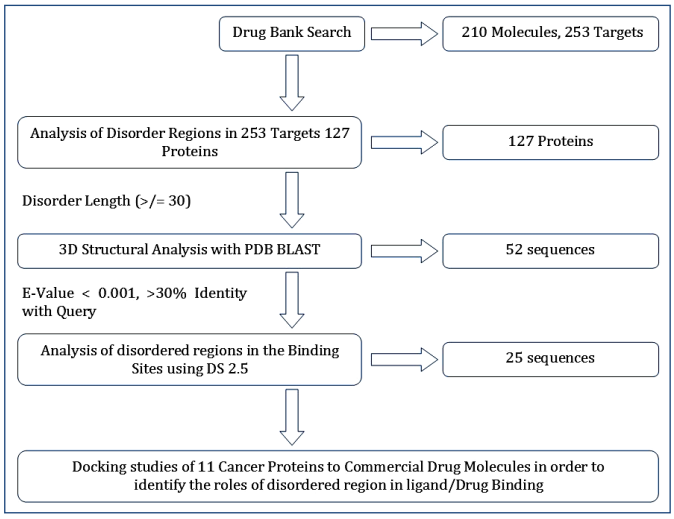


Fig. 1- Protocol followed to analyze the roles of disordered regions in drug binding

The docking studies were relevant because, an appreciable structural homology between the ligands (in their crystal structures) and the corresponding drug molecules were deciphered using Ligscore tool [44]. The RMSD values between the ligands and the respective commercial drug molecules are illustrated in [Table-3].

Table 2- Details of the commercial drugs used in the study

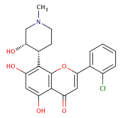
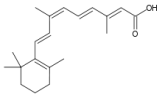
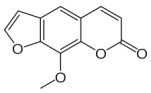
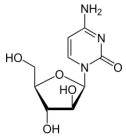
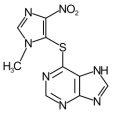
S no	Name of the Drug (no. of atoms contained)	Structure of the Drug	Ki / IC50 Values	Drug Bank ID
1	Flavopiridol (48) C21 H20 Cl N O5		IC50: 20nM Kd: 6.4 - 6400nM	DB03496
2	Alitretinoin (50) $C_{20}H_{28}O_2$		IC50: na	DB00523
3	Methoxsalen (24) C12 H8 O4		Kd: 1900nM	DB00553
4	Cytarabine (30) C9 H13 N3 O5		IC50: 0.001 to 1.0 nM	DB00987
5	Azathioprine (26) $C_9H_7N_7O_2S$		Ki:120 +/- 10 nM	DB00993

Table 3- RMSD between the ligand in PDB and the drug mentioned in drug bank

S No.	Protein	Ligand Molecule (no. of atoms contained)	Drug Molecule (no. of atoms contained)	RMSD (no. of atoms involved)
1	Cell Division Protein Kinase - 2	Indazole Inhibitor 9 (57)	Flavopiridol (48)	0.3157 (13)
2	Cell Division Protein Kinase - 5	R-Roscovitine (52)	Flavopiridol (48)	0.7127 (13)
3	Cell Division Protein Kinase - 6	Tetrahydroxyflavone (31)	Flavopiridol (48)	0.3966 (18)
4	Cell Division Protein Kinase - 7	ATP (47)	Flavopiridol (48)	0.6514 (11)
5	Cellular Retinoic Acid Binding Protein- 1	Retinobenzoic Acid (49)	Alitretinoin (50)	0.6006 (14)
6	Cellular Retinoic Acid Binding Protein- 2	Retinoic Acid (46)	Alitretinoin (50)	0.7077 (14)
7	Cytochrome P 450 2A6	Coumarin (17)	Methoxsalen (24)	0.0455 (11)
8	Serum Albumin	5-(2,4-Di Fluoro Phenyl) -2- Hydroxy-Benzoic Acid (26)	Cytarabine (30)	0.8102 (08)
9	GMP reductase 1	Guanosine Monophosphate (38)	Azathioprine (26)	0.9077 (10)
10	GMP reductase 2	Guanosine Monophosphate (38)	Azathioprine (26)	0.9077 (10)
11	GMP synthase [glutamine-hydrolyzing]	Xanthosine Monophosphate (38)	Azathioprine (26)	0.9550 (10)

To ascertain the nature of ligand-interactions with the receptors, PDB structures of these 11 cancer proteins were analyzed in detail for hydrogen bonds between the residues surrounding the ligand. These structures were then used for docking with their respective drug molecules using FlexX. Prior to this exercise, the ligand already present in crystal structure was re-docked to the receptor to account for variations by the software, if any. The docking was performed with default parameters except for "Maximum allowed overlap volume" value, which was changed to 3.6 Å from 2.9 Å, to appreciate all the possible interactions. The best docked pose was retrieved and minimized with the receptor using CHARMm [45] in Accelrys Discovery Studio 2.5. The hydrogen bonds formed by the residues surrounding the ligand or drug in docked pose are tabulated in the [Table-4a], [Table-4b], [Table-4c], [Table-4d], [Table-4e], [Table-4f], [Table-4g], [Table-4h], [Table-4i], [Table-4j], [Table-4k]. The summary of interactions with various receptors and respective ligands or drug molecules is illustrated in [Table-5].

Results and Discussions

Cell Division Protein Kinase 2 (CDK2) is a 298 amino acid long protein, targeted by the drug Flavopiridol. The disordered stretches of this protein include residues in the range 1-47 and 283-298. The docking studies with the drug molecule revealed that 14 residues react with the drug Flavopiridol, while 17 amino acids network with the Indazole Inhibitor. This clearly suggests that the drug molecule prefers to sit in the binding pocket, facilitating relevant interactions. Similarly, in the case of Cell Division Protein Kinase (CDK5), which is a 292 amino acid long protein, the disordered stretches are in between 1-45, 56-74, 219-231 and 277-292. The docking exercises with Flavopiridol indicated that 18 residues interact with the drug, while there are 21 possible contacts with the ligand R-Roscovitine, in the pocket of the receptor. Interestingly, for the 326 amino acid long Cell Division Protein Kinase 6 (CDK 6), which is having a wide disordered stretch of 114 residues belonging to 28-60, 69-82, 97-

116, 145-156, 192-200, 218-228, 253-262 and 311-326, the drug molecule interacts with 22 residues of the receptor, while the ligand Tetrahydroxyflavone makes 18 contacts only. It appears that the drug exhibits higher affinity to CDK6 molecule, than the ligand. Cell Division Protein Kinase 7 (CDK 7), which has 346 amino acids, is characterized by about 100 disordered residues belonging to the stretches 1-35, 40-55, 159-173, 230-239 and 319-346. The docking of CDK 7 with its drug molecule Flavopiridol indicated that 20 interactions are possible with the receptor, while the binding site residues coordinates with the ligand ATP via 23 contacts.

The Cellular Retinoic Acid Binding Protein 1 (CRABP 1) is a 136 amino acid long protein, targeted by the drug Alitretinoin. The disordered residues contained in the molecule belong to stretches 1-12, 36-78, 87-105 and 117-128. The docking of CRABP 1 with its drug molecule Alitretinoin was performed, which resulted in 19 interactions in the binding site. However, the receptor establishes 21 contacts with the ligand Retinobenzoic Acid. Similarly, the receptor Cellular Retinoic Acid Binding Protein 2 (CRABP 2) is a 137 residue protein, targeted by the same drug Alitretinoin; its disordered stretches include residues 1-11, 35-81 and 96-131. The proteins interact with the drug via 20 residues, while the ligand retinoic acid makes 19 interactions. This is in spite of the fact that, the structure from PDB for CRABP1 is the only non-human sequence in this set. Interestingly, the identities between the CRABP PDB sequences 2CBR (from *Bos taurus*) and 3CBS (from *Homo sapiens*) is more than 76%, and hence, the docking results could be correlated and even mapped on to human models as well.

Cytochrome P 450 2A6 is a protein containing 494 amino acids, targeted by the drug Methoxsalen; the disordered stretches of this proteins include residues 29-38, 97-108, 118-144, 184-192, 253-265, 327-348, 359-390 and 464-494, making it the longest in the set containing about 148 disordered residues. Yet, the docking results indicated that the drug molecule makes 14 interactions, while the ligand coumarin makes 13 contacts.

Serum Albumin is a 609 amino acid long protein, targeted by the drug Cytarabine. The disordered stretches of this protein include residues 127-139, 438-468 & 488-519. The docking of serum albumin with its drug molecule Cytarabine was performed and the result shows that, 19 interactions are possible with the binding site residues. On the other hand, the ligand makes 17 contacts.

GMP reductase 1, which is a 345 amino acid receptor, is targeted by the drug Azathioprine. Its disordered regions are 1-48, 287-300 & 325-345. The docking study infers that the drug molecule makes only 22 contacts in contrast to 31 interactions of the ligand GMP. Similarly, GMP reductase 2, which is also a 348 amino acid protein is targeted by the same drug Azathioprine. The number of disordered residues is 107 belonging to the stretches 1-49, 179-190, 248-259, 289-301 & 324-348. The docking results likewise indicate 31 interactions with the drug and 21 contacts with the GMP. In both of these cases, the ligand appears to be interacting strongly than the drug Azathioprine. Finally, the cancer protein GMP synthase [glutamine-hydrolyzing], which is indeed a 693 amino acid long protein, though has 141 residues belonging to the disordered stretches (230-237, 302-341, 386-397, 402-418, 426-440, 531-538, 565-574, 625-633, 644-658 and 670-693), docking exercises with the drug molecule Azathioprine highlights only 15 contacts, and the same with ligand Xanthosine Monophosphate offers 19 possible interactions.

*Table 4a-The set of interacting atoms in CDK2 (around the distance of 5Å) from the ligand (Indazole Inhibitor) and the drug (Flavopiridol) † - Hydrogen bonding residues, * - Disordered residues*

Sno	Interacting Residues of CDK2 within a distance of 5 Å	Interacting Distance with the atoms of	
		Ligand molecule (Indazole Inhibitor)	Drug molecule (Flavopiridol)
1	Ile-10†*	3.54	4.39
2	Gly-11*	---	4.41
3	Glu-12†*	---	4.3
4	Val-18*	4.38	4.24
5	Ala-31*	3.46	3.87
6	Lys-33*	3.85	---
7	Leu-55	---	---
8	Val-64	3.85	---
9	Leu-78	---	---
10	Phe-80	2.98	4.69
11	Glu-81	2.76	---
12	Phe-82	3.61	---
13	Leu-83	2.53	---
14	His-84	3.34	4.29
15	Glu-85	3.93	4.21
16	Asp-86†	3.55	2.58
17	Lys-89	---	4.79
18	Gln-131†	4.02	3.93
19	Leu-134	3.3	3.81
20	Leu-143	---	---
21	Ala-144	3.66	4.38
22	Asp-145†	4.81	2.65
23	Phe-146†	4.14	---
# of interactions within a distance of 5 Å		17	14
# of interactions of the disordered region residues		4	5

*Table 4b- The set of interacting atoms in CDK5 (around the distance of 5Å) from the ligand (Roscovitine) and the drug (Flavopiridol) † - Hydrogen bonding residues, * -Disordered residues*

S no	Interacting Residues of CDK5 within a distance of 5 Å	Interacting Distance with the atoms of	
		Ligand molecule (Roscovitine)	Drug molecule (Flavopiridol)
1	Ile-10†*	3.73	4.48
2	Gly-11*	3.8	---
3	Glu-12†*	3.41	3.71
4	Gly-13*	4.08	3.97
5	Val-18*	3.59	4.09
6	Ala-31*	3.47	3.66
7	Leu-32*	4.91	---
8	Lys-33*	4.18	3.53
9	Val-64	3.61	4.42
10	Phe-80	3.41	2.63
11	Glu-81†	3.3	4.86
12	Phe-82	4.15	3.32
13	Cys-83†	2.81	2.39
14	Asp-84†	4.56	3.51
15	Gln-85	3.57	4.17
16	Asp-86†	3.36	2.9
17	Lys-89	4.07	---
18	Gln-130†	2.83	3.8
19	Asp-131	---	---
20	Leu-133	3.53	3.6
21	Ala-143	4.24	3.43
22	Asp-144	4.26	3.21
# of interactions within a distance of 5 Å		21	18
# of interactions of the disordered region residues		8	6

*Table 4c- The set of interacting atoms in CDK6 (around the distance of 5Å) from the ligand (Tetra Hydroxy Flavone) and the drug (Flavopiridol) † -Hydrogen bonding residues, * -Disordered residues*

S no	Interacting Residues of CDK6 within a distance of 5 Å	Interacting Distance with the atoms of Ligand molecule (Tetra Hydroxy Flavone)	Drug molecule (Flavopiridol)
1	Ile-19 †	3.40	3.05
2	Gly-20 †	---	3.93
3	Glu-21	---	4.55
4	Gly-22	---	4.68
5	Val-27	4.54	3.74
6	Ala-41*	2.94	3.42
7	Lys-43 †*	3.40	4.72
8	Glu-61 †	2.39	---
9	Val-77	4.45	3.53
10	Phe-98*	3.20	3.46
11	Glu-99†*	3.02	4.72
12	His-100*	3.35	2.36
13	Val-101 †*	3.61	2.66
14	Asp-102 †*	---	3.75
15	Gln-103 †*	4.12	4.81
16	Asp-104 †*	3.12	2.60
17	Thr-107*	---	4.69
18	Gln-149 †*	2.82	3.98
19	Asn-150 †*	4.55	3.73
20	Leu-152*	3.78	3.47
21	Ala-162	4.25	4.12
22	Asp -163 †	3.27	3.47
# of interactions within a distance of 5 Å		18	22
# of interactions of the disordered region residues		11	13

*Table 4d- The set of interacting atoms in CDK7 (around the distance of 5Å) from the ligand (ATP) and the drug (Flavopiridol) † - Hydrogen bonding residues, * - Disordered residues*

S no	Interacting Residues of CDK7 within a distance of 5 Å	Interacting Distance with the atoms of Ligand molecule (ATP)	Drug molecule (Flavopiridol)
1	Leu-18†*	3.03	2.43
2	Gly-19†*	4.61	3.35
3	Glu-20*	3.39	5.00
4	Gly-21†*	4.62	4.88
5	Gln-22†*	3.12	---
6	Phe-23†*	3.62	---
7	Ala-24†*	3.11	---
8	Val-26*	3.28	3.95
9	Ala-39	3.61	3.56
10	Lys-41†*	2.94	3.20
11	Ile-75	4.63	4.79
12	Phe -91†	4.69	4.59
13	Asp-92†	.23	2.98
14	Phe-93	3.66	4.63
15	Met-94†	2.99	2.93
16	Glu-95	---	4.91
17	Thr-96	---	4.81
18	Asp-97†	4.00	4.97
19	Asp-137†	---	---
20	Lys -139†	3.60	---
21	Asn-141†	4.28	3.66
22	Asn-142†	4.32	4.17
23	Leu-144	4.10	3.14
24	Ala-154	---	4.86
25	Asp-155†	4.51	3.94
26	Gly-157	---	---
27	Lys -160†*	4.15	---
28	Ser-161†*	3.22	---
# of interactions within a distance of 5 Å		23	20
# of interactions of the disordered region residues		11	6

*Table 4e - The set of interacting atoms in CRABP1 (around the distance of 5Å) from the ligand (Retinobenzoic Acid) and the drug (Alitretinoin) † - Hydrogen bonding residues, * - Disordered residues*

S no	Interacting Residues of CRABP1 within a distance of 5 Å	Interacting Distance with the atoms of Ligand molecule (Retinobenzoic Acid)	Drug molecule (Alitretinoin)
1	Phe-15	3.92	4.19
2	Leu-19	4.52	---
3	Val-24	3.94	4.79
4	Leu-28	4.32	3.37
5	Val-31	3.76	3.14
6	Ala-32	4.13	3.77
7	Ala-35	4.19	3.71
8	Ala-36*	4.10	4.73
9	Pro-39*	3.84	---
10	Val-41*	---	---
11	Thr-54†*	3.59	4.03
12	Thr-56†*	3.27	4.07
13	Val-58*	3.55	3.77
14	Arg-59†*	3.75	4.50
15	Thr-61*	---	---
16	Thr-75*	---	---
17	Val-76†*	4.73	3.58
18	Asp-77*	3.68	4.02
19	Gly-78*	3.37	4.74
20	Arg-111†	4.24	3.37
21	Leu-120*	3.39	3.62
22	Phe-122*	4.14	3.20
23	Arg -131†	2.72	4.61
24	Tyr-133†	2.23	4.29
# of interactions within a distance of 5 Å		21	19
# of interactions of the disordered region residues		11	10

*Table 4f- The set of interacting atoms in CRABP2 (around the distance of 5Å) from the ligand (Retinoic acid) and the drug (Alitretinoin) † - Hydrogen bonding residues, * - Disordered residues*

S no	Interacting Residues of CRABP2 within a distance of 5 Å	Interacting Distance with the atoms of Ligand molecule (Retinoic acid)	Drug molecule (Alitretinoin)
1	Phe-15	4.08	3.99
2	Leu-19	2.90	4.10
3	Val-24	4.10	4.68
4	Leu-28	3.28	3.65
5	Ile-31	---	4.71
6	Ala-32	3.90	4.16
7	Ala-35*	4.04	4.43
8	Ala-36*	4.03	3.97
9	Pro-39*	3.72	3.70
10	Val-41*	---	4.29
11	Thr-54†*	4.28	3.85
12	Ser-55*	---	---
13	Thr-56*	3.70	3.80
14	Val-58*	3.66	3.22
15	Arg-59*	3.48	2.91
16	Val-76*	3.92	3.13
17	Asp-77†*	4.12	4.20
18	Arg-111†*	4.10	---
19	Leu-121*	2.99	2.55
20	Met-123*	4.59	4.30
21	Arg-132†	2.70	2.43
22	Tyr-134†	2.46	2.37
# of interactions within a distance of 5 Å		19	20
# of interactions of the disordered region residues		12	12

*Table 4g- The set of interacting atoms in Cytochrome P450 2A6 (around the distance of 5Å) from the ligand (Coumarin) and the drug (Methoxsalen) † - Hydrogen bonding residues, * - Disordered residues*

S no	Interacting Residues of Cytochrome P 450 2A6 within a distance of 5 Å	Interacting Distance with the atoms of Ligand molecule (Coumarin)	Drug molecule (Methoxsalen)
1	Phe-107*	3.78	3.49
2	Phe-111	3.74	3.64
3	Val-117	3.98	3.47
4	Phe-118*	3.81	3.58
5	Phe-209	3.97	3.54
6	Leu-296†	4.64	4.47
7	Asn-297†	3.12	3.93
8	Leu-298†	---	4.47
9	Ile-300	3.89	3.57
10	Gly-301†	3.01	3.01
11	Thr-305	3.84	3.39
12	Ile-366*	4.21	4.24
13	Leu-370*	4.00	4.73
14	Phe-480*	3.44	3.87
# of interactions within a distance of 5 Å		13	14
# of interactions of the disordered region residues		5	5

*Table 4h- The set of interacting atoms in Serum Albumin (around the distance of 5Å) from the ligand (5-(2,4-Di Fluoro Phenyl) -2 -Hydroxy -Benzoic Acid) and the drug (Cytrabine) † - Hydrogen bonding residues, * - Disordered residues*

S no	Interacting Residues of Serum Albumin within a distance of 5 Å	Interacting Distance with the atoms of Ligand molecule (5-(2,4-Di Fluoro Phenyl) -2-Hydroxy -Benzoic Acid)	Drug molecule (Cytarabine)
1	Leu-411	3.79	2.92
2	Ile-412	3.83	3.35
3	Asn-415	3.49	3.99
4	Cys-416	3.94	4.82
5	Phe-427	3.67	3.57
6	Leu-431	4.45	3.05
7	Arg-434†*	2.88	2.95
8	Try-435†*	4.88	2.90
9	Lys-438†*	3.30	2.77
10	Leu-454†*	3.20	4.27
11	Gly-455†*	---	4.61
12	Val-457†*	3.96	3.48
13	Gly-458†*	3.25	4.30
14	Ser-459*	---	4.73
15	Cys-462*	3.81	3.94
16	Leu-477	3.45	3.37
17	Arg-509†*	4.93	4.50
18	Phe-512†*	4.68	4.71
19	Ser-513†*	3.25	4.07
# of interactions within a distance of 5 Å		17	19
# of interactions of the disordered region residues		10	12

*Table 4i- The set of interacting atoms in GMP Reductase 1(around the distance of 5Å) from the ligand (GMP) and the drug (Azathioprine) † -Hydrogen bonding residues, * -Disordered residues*

S no	Interacting Residues of GMP Reductase 1 within a distance of 5 Å	Interacting Distance with the atoms of Ligand molecule (GMP)	Drug molecule (Azathioprine)
1	Ala-53	4.50	3.89
2	Asn-54†	4.71	---
3	Met-55	3.68	3.85
4	Asn-158†	4.86	---
5	Lys-177	4.49	---
6	Pro-182	4.24	4.88
7	Gly-183†	4.44	4.15
8	Ser-184†	2.43	2.60
9	Val-185†	4.76	3.60
10	Cys-186	4.71	4.39
11	Thr-187	4.86	4.32
12	Thr-188	4.19	---
13	Asp-219†	2.59	2.68
14	Gly-220†	3.95	4.86
15	Gly-221†	2.87	---
16	Cys-222†	4.33	---
17	Met-240	3.56	---
18	Leu-241†	4.67	4.86
19	Gly-242†	2.77	4.17
20	Gly-243†	2.85	3.98
21	Met-244†	4.95	---
22	Phe-266	4.55	4.84
23	Gly-268	3.19	4.15
24	Met-269†	3.15	3.19
25	Ser-270†	2.93	4.43
26	Ser-271†	3.67	---
27	Arg-286†	4.97	4.03
28	Ala-287*	3.63	3.46
29	Ser-288†*	2.77	2.40
30	Glu-289*	4.69	4.62
31	Gly-290†*	3.07	4.27
# of interactions within a distance of 5 Å		31	22
# of interactions of the disordered region residues		4	4

*Table 4j- The set of interacting atoms in GMP Reductase2 (around the distance of 5Å) from the ligand (GMP) and the drug (Azathioprine) † - Hydrogen bonding residues, * - Disordered residues*

S no	Interacting Residues of GMP Reductase 2 within a distance of 5 Å	Interacting Distance with the atoms of Ligand molecule (GMP)	Drug molecule (Azathioprine)
1	Ala-53	3.86	3.85
2	Asn-54†	4.76	---
3	Met-55	3.55	3.70
4	Asn-158†	4.58	---
5	Lys-177	4.40	---
6	Pro-182*	3.98	4.26
7	Gly-183†*	4.56	4.12
8	Ser-184†*	3.19	2.84
9	Val-185†*	4.31	4.80
10	Cys-186†*	4.71	4.14
11	Thr-187*	4.81	---
12	Thr-188*	3.32	---
13	Asp-219†	2.81	3.69
14	Gly-220†	4.02	4.82
15	Gly-221†	3.05	3.16

Table 4j- Continue…

S no	Interacting Residues of GMP Reductase 2 within a distance of 5 Å	Interacting Distance with the atoms of Ligand molecule (GMP)	Drug molecule (Azathioprine)
16	Cys-222†	4.26	4.84
17	Met-240	3.81	3.85
18	Leu-241	4.98	3.91
19	Gly-242†	2.78	3.96
20	Gly-243†	2.84	2.79
21	Met-244†	4.35	---
22	Phe-266	4.21	4.24
23	Tyr-267†	---	4.96
24	Gly-268†	4.89	4.41
25	Met-269†	3.24	2.90
26	Ser-270†	2.97	3.68
27	Ser-271†	3.77	---
28	Arg-286†	2.97	---
29	Ala-287†	4.75	---
30	Ser-288†	2.54	---
31	Glu-289*	4.67	4.90
32	Gly-290†*	3.12	4.56
# of interactions within a distance of 5 Å		31	21
# of interactions of the disordered region residues		9	6

The results from the [Table-4a], [Table-4b], [Table-4c], [Table-4d], [Table-4e], [Table-4f], [Table-4g], [Table-4h], [Table-4i], [Table-4j], [Table-4k] and the numbers of interactions as per [Table- 5], thus emphasize the involvement of disordered region residues in the binding pockets, and the participation of these residues towards hydrogen bonding interactions with the ligand and drug molecules respectively. As could be appreciated, the total numbers of interactions made by the proteins with the ligand and drug molecules respectively, in most of the cases were in the close agreement, except for GMP reductases 1 & 2.

*Table 4k- The set of interacting atoms in GMP Synthase (around the distance of 5Å) from the ligand (Xanthosine Monophosphate) and the drug (Azathioprine) † - Hydrogen bonding residues, * - Disordered residues*

S no	Interacting Residues of GMP Synthase within a distance of 5 Å	Interacting Distance with the atoms of Ligand molecule (Xanthosine Monophosphate)	Drug molecule (Azathioprine)
1	Arg-337†*	2.88	3.66
2	Ser-382	4.36	3.16
3	Gly-383†	4.22	3.29
4	Lys-384†	3.55	3.84
5	Pro-437*	4.41	---
6	Phe-438*	4.98	---
7	Pro-439†*	4.53	3.31
8	Gly-440†*	3.02	4.26
9	Pro-441	3.33	4.82
10	Ile-445	4.09	3.97
11	Arg-446	4.41	---
12	Arg-524†	4.01	2.84
13	Gln-610†	2.60	---
14	Phe-645†*	3.55	4.01
15	Lys-685†*	2.80	4.21
16	Thr-689†*	3.44	3.49
17	Thr-690†*	2.63	4.14
18	Glu-691†*	2.67	4.38
19	Glu-693*	4.97	3.12
# of interactions within a distance of 5 Å		19	12
# of interactions of the disordered region residues		11	9

Table 5- Details of Residues interacting with 11 Cancer Proteins with their Ligands and the Drug molecules respectively.

S no	Cancer Protein	Ligand Molecule	No. of interactions with Ligand	Drug Molecule	No. of interactions with Drug
1	Cell Division Protein Kinase - 2	Indazole Inhibitor 9	17	Flavopiridol	14
2	Cell Division Protein Kinase - 5	R-Roscovitine	21	Flavopiridol	18
3	Cell Division Protein Kinase - 6	Tetrahydroxyflavone	18	Flavopiridol	22
4	Cell Division Protein Kinase - 7	ATP	23	Flavopiridol	20
5	Cellular Retinoic Acid Binding Protein- 1	Retinobenzoic Acid	21	Alitretinoin	19
6	Cellular Retinoic Acid Binding Protein- 2	Retinoic Acid	19	Alitretinoin	20
7	Cytochrome P 450 2A6	Coumarin	13	Methoxsalen	14
8	Serum Albumin	5-(2,4-Di Fluoro Phenyl) -2 - Hydroxy -Benzoic Acid	17	Cytarabine	19
9	GMP reductase 1	Guanosine Monophosphate	31	Azathioprine	22
10	GMP reductase 2	Guanosine Monophosphate	31	Azathioprine	21
11	GMP synthase [glutamine-hydrolyzing]	Xanthosine Monophosphate	19	Azathioprine	15

Conclusions

The investigations thus have highlighted that residues belonging to disordered regions in these cancer targets do play vital roles in binding to the drug molecules. Based on the fact that there is a glaring homology between ligands and drug molecules, the nature of binding could be well correlated. The work establishes the fact that the commercially available drug molecules could preferentially bind to disordered regions and exercise cell proliferation and control. Accordingly, the study proposes a novel pattern of exploiting the binding site of receptors, that could be useful in pharmacophore based drug design. Our future investigation includes the detail examination of interactions with the remaining disordered cancer proteins with respective drug molecules in the cancer pathways.

Acknowledgements

The authors acknowledge the support from Visvesvaraya Technological University (VTU), Belgaum, for the financial grant extended towards this project. The infrastructural resources and encouragement provided by Sir M Visvesvaraya Institute of Technology (Sir MVIT), Bangalore, is also acknowledged.

Conflict of Interest

The authors of this manuscript declare that they do not have any conflict of interest and direct relation with any commercial entity mentioned in the paper. FlexX 2.1.2. and Accelrys Software Inc., USA, version 3.5 software, which is a licensed version.

References

[1] Horsfall F.L.Jr. (1963) *Can. Med. Assoc. J.* 89(24), 1224-1229.

[2] Doll R., Peto R. (1981) *Journal of the National Cancer Institute*, 66(6), 1191-308.

[3] Demetrios A.S. (2007) *Journal of Balkan Union of Oncology*, 12 (1), 9-12.

[4] Croce C.M. (2008) *The New England Journal of Medicine*, 358, 502-511.

[5] Georgieva J., Sinha P., Schadendorf D. (2001) *J. Clin. Pathol.*, 54, 229-23.

[6] Won J.Y., Nam E.C., Yoo S.J., Kwon H.J., Um S.J., Han H.S., Kim S.H., Byun Y., Kim S.Y. (2004) *Metabolism*, 53(8), 1007-1012.

[7] Rodriguez-Antona C., Ingelman-Sundberg M. (2006) *Oncogene*, 25(11), 1679-91.

[8] Frei E. (2011) *Diabetology & Metabolic Syndrome*, 3(1), 11.

[9] Hirst M., Haliday E., Nakamura J., Lou L. (1994) *Journal of Biological Chemistry*, 269(38), 23830-23837.

[10] Zhang J., Zhang W., Zou D., Chen G., Wan T., Zhang M., Cao X. (2003) *J. Cancer Res. Clin. Oncol.*, 129(2), 76-83.

[11] Dyson H.J., Wright P.E. (2005) *Nature Reviews Molecular Cell Biology*, 6(3), 197-208.

[12] Romero P., Obradovic Z., Li X., Garner E.C., Brown C.J., Dunker A.K. (2001) *Proteins: Structure, Function, and Bioinformatics*, 42(1), 38-48.

[13] Vladimir N., Uversky V.N., Oldfield C.J., Dunker A.K. (2008) *Annual Review of Biophysics*, 37, 215-246.

[14] Iakoucheva L.M., Brown C.J., Lawson D.J., Obradovic Z., Dunker A.K. (2002) *Journal of Molecular Biology*, 323(3), 573-584.

[15] Wells M., Tidow H., Rutherford T.J., Markwick P., Jensen M.R., Mylonas E., Svergun D.I., Blackledge M., Fersht A.R. (2008) *Proceedings of the National Academy of Sciences*, 105(15): 5762-5767.

[16] Wright P.E., Dyson H.J. (1999) *Journal of Molecular Biology*, 293(2), 321-331.

[17] Salma P., Chhatbar C., Seshadri S. (2009) *American Journal of Infectious Diseases,* 5(2), 133-141.

[18] Salma P., Chhatbar C. and Seshadri S. (2009) *American Journal of Infectious Diseases*, 5(2), 126-132.

[19] Anurag M., Dash D. (2009) *Molecular Biosystems*, 5, 1752-1757.

[20] Wishart D.S., Knox C., Guo A.C., Cheng D., Shrivastava S., Tzur D., Gautam B., Hassanali M. (2008) *Nucleic Acids Res.*, 36, 901-906.

[21] Wishart D.S., Knox C., Guo A.C., Shrivastava S., Hassanali M., Stothard P., Chang Z., Woolsey J. (2006) *Nucleic Acids Res.*, 34, 668-672.

[22] Linding R., Jensen L.J., Diella F., Bork P., Gibson T.J., Russell R.B. (2003) *Structure,* 11(11), 1316-1317.

[23] Li X., Kahveci T. (2006) *Bioinformatics,* 22(26), 2980-2987.

[24] Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. (2000) *Nucleic Acids Research*, 28, 235-242.

[25] Brooks B.R., Brooks C.L., Mackerell A.D.Jr., Nilsson L., Petrella R.J., Roux B., Won Y., Archontis G., Bartels C., Boresch S., Caflisch A., Caves L., Cui Q., Dinner A.R., Feig M., Fischer S., Gao J., Hodoscek M., Im W., Kuczera K., Lazaridis T., Ma J., Ovchinnikov V., Paci E., Pastor R.W., Post C.B., Pu J.Z., Schaefer M., Tidor B., Venable R.M., Woodcock H.L., Wu X., Yang W., York D.M., Karplus M. (2009) *J. Comput. Chem*., 30, 1545-1614.

[26] Iancu C.V., Borza T., Fromm H.J., Honzatko R.B. (2002) *J. Biol. Chem*., 277, 26779-26787.

[27] van den Elsen J.M., Kuntz D.A., Rose D.R. (2001) *The EMBO Journal*, 20(12), 3008-3017.

[28] Hirst J., Goodin D.B. (2000) *J. Biol. Chem.*, 275, 8582-8591.

[29] Stamos J., Sliwkowski M.X., Eigenbrot C. (2002) *J. Biol. Chem.,* 277, 46265-46272.

[30] Thunnissen M.M., Nordlund P., Haeggstrom J.Z. (2001) *Nat. Struct. Biol.*, 8, 131-135.

[31] Soldano K.L., Jivan A., Nicchitta C.V., Gewirth D.T. (2003) *J. Biol. Chem.* 278, 48330-48338.

[32] Liang J., Hung D.T., Schreiber S.L., Clardy J. (1996) *J. Am. Chem. Soc.*, 118, 1231-1232.

[33] Nettles J.H., Li H., Cornett B., Krahn J.M., Snyder J.P., Downing K.H. (2004) *Science,* 305, 866-869.

[34] Trujillo J.I., Kiefer J.R., Huang W., Thorarensen A., Xing L., Caspers N.L., Day J.E., Mathis K.J., Kretzmer K.K., Reitz B.A., Weinberg R.A., Stegeman R.A., Wrightstone A., Christine L., Compton R., Li X. (2009) *Bioorg. Med. Chem. Lett.,* 19, 908-911.

[35] Mapelli M., Massimilinao L., Crovace C., Seeliger M.A., Tsai L.H., Meijer L., Musacchio A. (2005) *J. Med. Chem.,* 48, 671.

[36] Lu H.S., Chang D.J., Baratte B., Meijer L., Schulze-Gahmen U. (2005) *J. Med. Chem.,* 48, 737-747.

[37] Lolli G., Lowe E.D., Brown N.R., Johnson L.N. (2004) *Structure,* 12, 2067-2079.

[38] Chaudhuri B.N., Kleywegt G.J., Broutin-L`Hermite I., Bergfors T., Senn H., Le Motte P., Partouche O., Jones T.A. (1999) *Acta Crystallography*, D55, 1850-1857.

[39] Yano J.K., Hsu M.H., Griffin K.J., Stout C.D., Johnson E.F. (2005) *Nat. Struct. Mol. Biol.*, 12, 822-823.

[40] Ghuman J., Zunszain P.A., Petitpas I., Bhattacharya A.A., Otagiri M., Curry S. (2005) *J. Mol. Biol.*, 353(1), 38-52.

[41] Li J., Wei Z., Zheng M., Gu X., Deng Y., Qiu R., Chen F., Ji C., Gong W., Xie Y., Mao Y. (2006) *J. Mol. Biol.*, 355, 980-988.

[42] Forino M., Jung D., Easton J.B., Houghton P.J. and Pellecchia M. (2005) *Journal of Medicinal Chemistry,* 48(7), 2278-2281

[43] Rarey M., Kramer B., Lengauer T. and Klebe G. (1996) *J. Mol. Biol.*, 261(3), 470-89

[44] Krammer A., Kirchhoff P.D., Jiang X., Venkatachalam C.M., Waldman M. (2005) *Journal of Molecular Graphics and Modelling*, 23(5), 395-407.

[45] Brooks B.R., Brooks C.L., Mackerell A.D., Nilsson L., Petrella R.J., Roux B., ... & Won Y., Archontis G., Bartels C., Boresch S., Caflisch A., Caves L., Cui Q., Dinner A.R., Feig M., Fischer S., Gao J., Hodoscek M., Im W., Kuczera K., Lazaridis T., Ma J., Ovchinnikov V., Paci E., Pastor R.W., Post C.B., Pu J.Z., Schaefer M., Tidorv B., Venable R.M., Woodcock H.L., Wu X., Yang W., York D.M., Karplus M. (2009) *Journal of Computational Chemistry*, 30(10), 1545-1614.