



A new insight into identification of in silico analysis of natural compounds targeting GPR120

Nagaraju Chinthakunta¹ · Srinivasulu Cheemanapalli¹ · Surekha Chinthakunta² · C. M. Anuradha³ · Suresh Kumar Chitta¹

Received: 20 October 2017 / Revised: 6 April 2018 / Accepted: 7 April 2018 / Published online: 14 May 2018
© The Author(s) 2018

Abstract

G-protein coupled receptor (GPR120) is an omega-3 fatty acid receptor that inhibits macrophage-induced tissue inflammation. Recent studies revealed GPR120 promotes colorectal carcinoma through modulation of VEGF, IL-8, PGE2, and NF-kB expression. However, three-dimensional structure of GPR120 is not yet available in Protein Data Bank (PDB). In the present study, we focused on a 3-D structural model of GPR120 has been constructed using homology modeling techniques. The structural quality of the predicted GPR120 model was verified using Procheck, Whatif, ProSA, and Verify 3D. After this chemical database of natural compounds have been constructed and screened for its druggability using molinspiration server. Molecular docking studies of natural compounds on GPR120 model revealed that silibinin (−6.87 kcal/mol), withanolide (−6.19 kcal/mol), limonene (−6.17 kcal/mol), and cervical (−6.15 kcal/mol) have shown good docking interactions with active site residues of the target. Active site residues of Arg280, Asp275, and Gly122 showed hydrogen-bonding interactions with predicted compounds. Based on these in silico findings, we proposed that virtual screening of natural compounds against of GPR120 is a novel approach to find potential anti-colorectal cancer therapeutics.

Keywords Colorectal cancer · G-protein coupled receptor 120 · Homology modeling natural compounds · Rule of five · Docking studies · Potential therapeutics

1 Introduction

Colon or colorectal cancer is a type of cancer that starts in the large intestine (colon) or the rectum (end of the colon). The origin of colon cancer is gastrointestinal tract of the epithelial cell lining present in the colon or rectum. Commonly mutations occur in the intestinal crypt cells (Ionov et al. 1993; Abdul Khalek et al. 2010). Colorectal cancer (CRC) is the third among most common malignancies worldwide (Shike et al. 1990) and the second leading cause of cancer deaths in the United States. It is

estimated that 132,700 new cases were diagnosed in the United States in 2015 and 49,700 deaths occurred due to this disease (American Cancer Society 2015). Better medication is still a significant cause of cancer-associated deaths. CRC is based on a complex of diseases arising from multistep process events in enterocytes, including genetic, epigenetic events, and abnormal signaling in basic cellular pathways. Thus, it is a hallmark of clinical value to identify potential molecules for tumor-preventive strategies (Roberta Bertorelle et al. 2014).

In G-protein-coupled receptors (GPCRs), ligands bind specifically to GPCRs to stimulate and induce a variety of cellular responses via several second messenger pathways; e.g., modulation of cyclic-AMP production, the phospholipase C pathway, ion channels, and MAPK (Ulloa-Aguirre et al. 1999; Gether 2000; Schulte and Fredholm 2003). They are important signaling molecules for many aspects of cellular functions, including vision, olfaction, behavior, and autonomic transmission nervous system (Morris and Malbon 1999). Besides, they also regulate many characteristic features of tumorigenesis, including proliferation,

✉ Suresh Kumar Chitta
suresh.chitta@skuniversity.ac.in; chitta34c@gmail.com

¹ Bioinformatics Infrastructure Facility, Department of Biochemistry, Sri Krishnadevaraya University, Anantapur, Andhra Pradesh 515003, India

² Department of Botany, Rayalaseema University, Kurnool, Andhra Pradesh 518007, India

³ Department of Biotechnology, Sri Krishnadevaraya University, Anantapur, Andhra Pradesh 515003, India

invasion, survival at the secondary site, and immune cell function, as well as several cancer-associated signaling pathways (Feigin 2013). These properties permitted the widespread development of GPCR-targeted drugs, which represent nearly 30% of all currently used therapeutics (Lappano and Maggiolini 2011; Dorsam and Gutkind 2007). In particular, G-protein-coupled receptor 120 (GPR120), the most enigmatic member of this large family, has generated attention because of its potential role in the regulation of metabolic and inflammatory diseases such as obesity and type 2 diabetes.

G-protein coupled receptor 120 is a functional omega-3 FA receptor/sensor and mediates powerful insulin sensitizing and anti-diabetic effects by repressing macrophage-induced tissue inflammation (Oh et al. 2010). It is highly expressed in adipose tissue and proinflammatory macrophages, while activation of GPR120 affected LPS- and TNF- α -induced inflammatory signaling responses (Oh et al. 2010; Ichimura et al. 2012). Activation of GPR120 signaling induced the expression and secretion of proangiogenic mediators of CRC cells which promoted the angiogenesis. The PI3K/Akt–NF- κ B pathway is activated by GPR120 signaling and required for GPR120 signaling-induced angiogenic switching in CRC cells. Furthermore, GPR120 activation enhanced motility of CRC cells and induced epithelial–mesenchymal transition (EMT) of CRC cells (Wu et al. 2013).

There is an increasing demand for natural compounds that improve human health. The World Health Organization estimated that approximately 80% of the world's inhabitants rely on the traditional medicine for their primary health care (Farnsworth et al. 1985). Plants have long been used in the treatment of cancer (Hartwell 1971), and many nutritive and non-nutritive phytochemicals with diversified pharmacological properties have shown promising responses for the prevention and/or intervention of various cancers (Surh 2003). These products, especially phytochemicals, have been extensively studied and have exhibited anti-carcinogenic activities by interfering with the initiation, development, and progression of cancer through the modulation of various mechanisms including cellular proliferation, differentiation, apoptosis, angiogenesis, and metastasis (Rajesh et al. 2015). We focused on construction of GPR120 model using *in silico* tools and refinement of structure by docking studies with natural compounds which are believed to help in understanding of structural features and the interactions of natural compounds with GPR120 which may be helpful in designing of novel inhibitors of colorectal cancer.

2 Materials and methods

In the present study, all the calculations were performed in a workplace by AMD 64 bits dual processing hi end server machines. Molecular modeling tasks were performed with Modeller9v3; docking calculations were performed with AutoDock 4.0. Unless otherwise stated, default settings were used during all calculations.

2.1 Sequence alignments

All the analysis was carried out by AMD 64 bits dual processing hi end server machines. The sequence of G-protein coupled receptor120 (gi: 82581671) was obtained from the National Centre for Biotechnology Information (NCBI. <http://www.ncbi.nlm.nih.gov/>). Local alignments were predicted using Blastp (Basic Local Alignment Search Tool) (Altschul et al. 1997) at the NCBI and the homologous entries were obtained from the protein data bank (Berman et al. 2000). The Blastp alignment was further refined using sequence alignments in the Clustal W/X 1.83 with default parameters (Thompson et al. 1994).

2.2 3D model construction

The Blastp alignment was used for homology modeling built in Modeller9v3 (<http://www.salilab.org/modeller/9v3>) which generated structures by applying spatial restraints. A bundle of 100 models from the random generation of the starting structure was calculated and subsequently the best model (with the low RMS value of superposition using Swiss-pdb viewer) (Guex and Peitsch 1997). To gain a better relaxation and much apart arrangement of the atoms, refinement was done on the built GPR120 model by energy minimization (EM). The stabilization was assessed by graphics visualization.

2.3 Evaluation of the homology model

The stereochemical parameters of the energy minimized GPR120 model were assessed by Procheck (Laskowski et al. 1993), Whatif (Vriend 1990), Errat (Colovos and Yeates 1993), ProSA (Sippl 1993) and Verify 3D (Bowie et al. 1991; Luthy et al. 1992). Verify 3D was used to assess whether a primary sequence is compatible with the current 3D structural model. The compatibility between the amino acid side chains in the model is a validation criterion. Torsion angle restraints for the side chains of each amino acid in the predicted GPR120 model were determined using a

Fig. 1 Multiple sequence alignment of GPR120 receptor and the template 4N6H. Highly conserved residues are represented by as stars



Web server Predictor (Berjanskii et al. 2006). The predictor assigns an error in the predicted chi (χ) torsion angle and including grains by combining its confidence scores with predicted or identified secondary structures and local sequence identity. Secondary structural conformations for the developed GPR120 model were predicted by Pdbsum (Laskowski et al. 2005).

2.4 Docking studies

2.4.1 Selection and screening of natural ligands

To fulfil the aim of constructing a novel ligand for GPR120, we selected a library of 100 molecules from the previous publications and browsing Internet. The selected library of ligands was tested for Lipinski’s rule of five using molinspiration server (Lipinski et al. 2001) for their ability to follow the rule of five. Auto Dock 4.0/ADT (Goodsell and Morris 1998) program was used to investigate ligand binding to structurally refined GPR120 model using a grid spacing of 0.375 Å and the grid points in X, Y, and Z axis were set to 60×60×60. The search was based on the Lamarckian genetic algorithm (Miyamoto and Kollman 1992; Oprea et al. 2001) and the results were analyzed using binding energy. For each ligand, a docking experiment consisting of 100 stimulations was performed and the analysis was based on binding-free energies and root-mean-square deviation (RMSD) values. Docking with natural Compounds was also performed onto GPR120 model with the same parameters and PMV 1.4.5 viewer was then used to observe the interactions of the docked compounds to the GPR120 model (Kitchen et al. 2004) and we submitted the developed 3D model of GPR120 to Protein Model Data Base (PMDB) (Castrignano et al. 2006), which maintains 3D models obtained by structure prediction methods.

3 Results and discussion

3.1 Sequence alignments

The coordinating 3D structure of Human Delta Opioid 7tm Receptor (PDB ID: 4N6H) (Fenalti et al. 2014). We found more than 70 crystallographic structures showing high identity score with respect to G-protein coupled receptor 120 using BLASTp results. We selected the Human Delta Opioid 7tm Receptor structure as template and the sequence identity

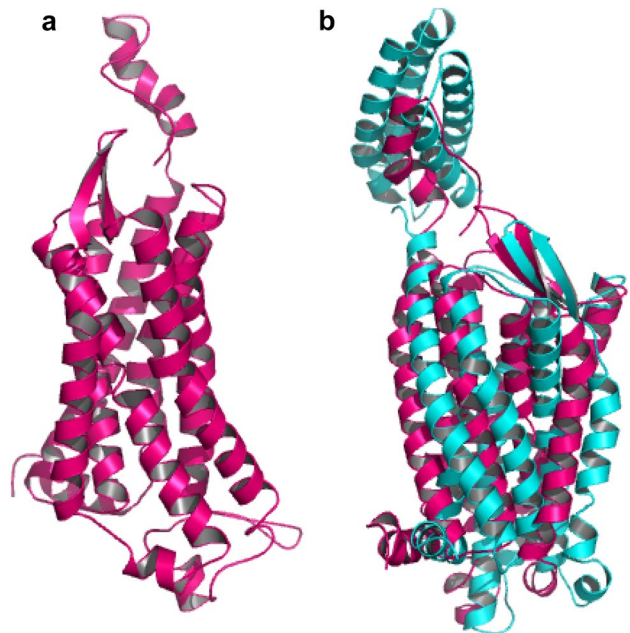


Fig. 2 a Predicted 3-D structure of GPR120 using Modeller 9v3. b Superimposed structures of GPR120 (hot pink) and 4N6H (cyan)

between G-protein coupled receptor 120 and template 4N6H has 26% similarity having a resolution of 1.80 Å making it an excellent template. The most significant step in homology modeling process is to obtain the correct sequence alignment of the target sequence with the homologues. Finally,

we performed an alignment between the selected template and the G-protein coupled receptor 120 using the ClustalX 1.8 with default parameters (Thompson et al. 1994). The sequence alignment performed homology modeling is shown in Fig. 1.

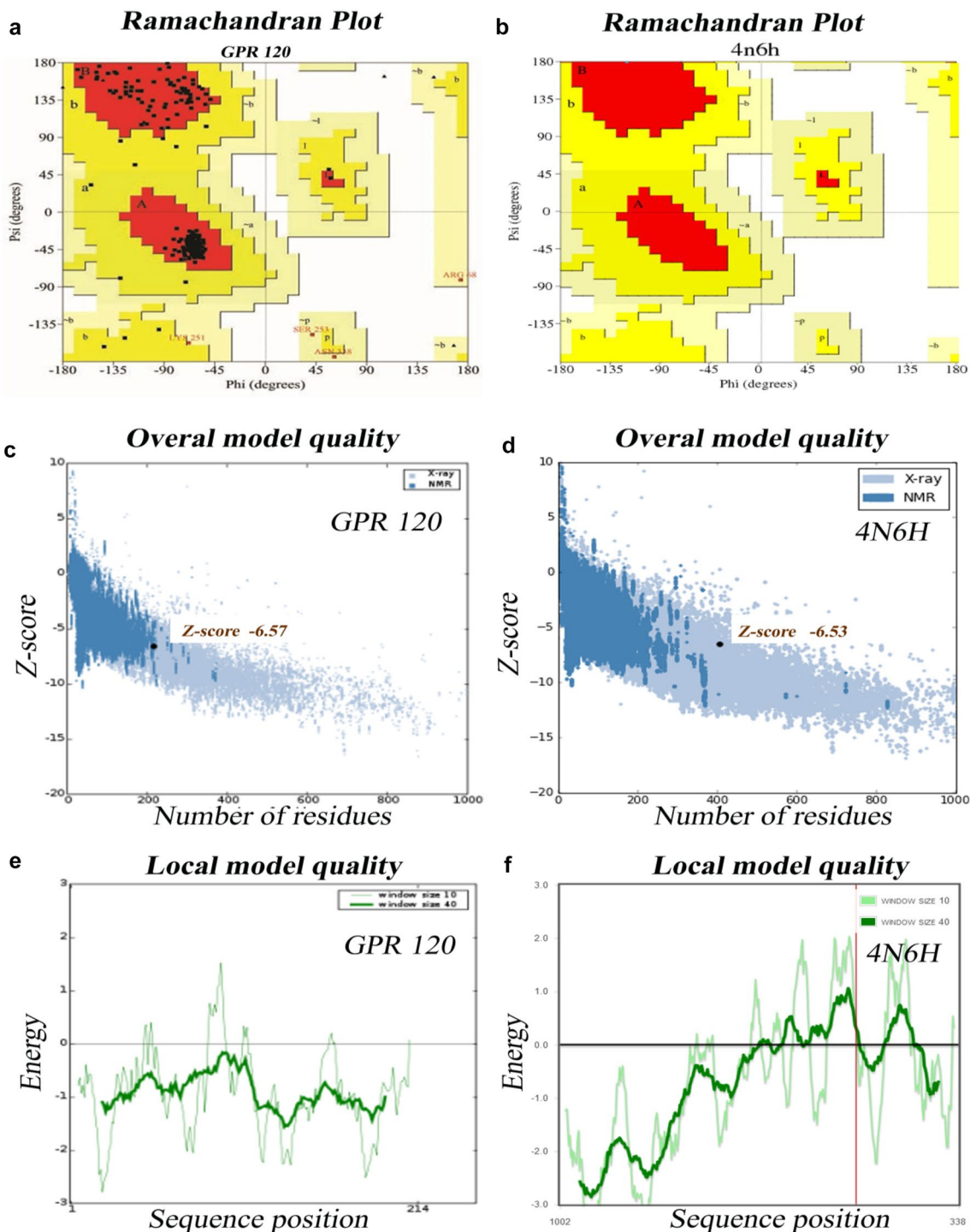


Fig. 3 **a** ProSA-web Z-scores of all protein chains in PDB determined by X-ray crystallography (light blue) and NMR spectroscopy (dark blue) with respect to their length. The Z-score of GPR120 present in that range represented in large black dot. **b** Energy plot for the predicted GPR120

3.2 Homology modeling

The search using the BLASTp alignment algorithm within the PDB database showed various potential templates for molecular modeling purposes. More than 70 crystallographic structures showed high identity score with and maximum query coverage respect to G-protein coupled receptor. The coordinates of the crystal structures of Human Delta Opioid 7m Receptor (PDB ID: 4N6H) (Castrignano et al. 2006) were used as a template to build the structure of G-protein coupled receptor 120. The 3D models of the G-protein coupled receptor 120 were built by Modeller 9v3. One hundred models were generated and the crystal structure of the template was saved for further refinement and validation (Fig. 2a, b). Furthermore, refinement was performed to obtain the best conformation of the developed model of G-protein coupled receptor 120.

3.3 Structural validation of developed model

The constructed model was subjected to validation using Ramachandran plot with Procheck program by checking the detailed residue-by-residue stereochemical quality of a protein structure (Laskowski et al. 1993). The Ramachandran plot revealed that 100% of the residues in homology model were in favored and allowed regions. The main structural elements of the optimized GPR120 homology model are shown in Fig. 3. In comparison with the templates, the homology model had a similar Ramachandran plot with 0.0% residues in disallowed regions. The total quality G-factor was -0.1 , which indicates a good quality model (acceptable values of the G-factor in Procheck are between 0 and -0.5 , with the best models displaying values close to zero) showed in (Table 1). The Errat is a so-called “overall quality factor” for non-bonded atomic interactions and higher scores mean higher quality (Sippl 1993). The normally accepted range is > 50 for a high-quality model (Colovos and Yeates 1993). In the current case, the Errat score for the GPR120 model is 54.366, which well within the range of a high-quality model.

Table 1 Ramachandran plot statistics

Residues in most favored regions	328	95.1%
Residues in additional allowed regions	13	3.8%
Residues in generously allowed regions	4	1.2%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non proline residues	345	100.00%
Number of end-residues (excl. Gly and Pro)	1	
Number of glycine residues (shown as triangle)	15	
Number of proline residues	16	
Total number of residues	377	

Analysis of the energy minimized GPR120 model with Whatif web interface (Vriend 1990) revealed that RMS Z-Scores for bond angles and bond lengths are all close to 1 and also within the limits of template. Detailed structural investigation of the predicted GPR120 model with Pdbsum, a secondary structure prediction server, revealed that 225 (59.7%) residues are in α -helices and 3 (0.8%) residues are in 3–10 helix and 137 (36.3%) residues are in other conformations (Fig. 4, Laskowski et al. 2005). The tertiary structure of GPR120 showed close similarity to crystallized 4N6H, with a backbone RMS value of GPR120–4N6H, is between 0.58 and 0.72 Å, respectively. The low RMS values for backbone superposition reflect the high structural conservation of this complex through evaluation, making it a good system for homology modeling.

3.4 Screening and docking studies of natural inhibitors of GPR120

Docking is frequently used to predict the binding orientation of small molecule drug candidates to their protein targets to predict the affinity and activity of the small molecule. Hence, docking plays an important role in the rational design of drugs (Kitchen et al. 2004). Docking studies were performed to gain insight into the binding interaction between constructed model of GPR120 and selected 100 natural compounds.

3.4.1 Selection and screening of ligand molecules

One hundred natural compounds used as ligand molecules were taken from the National Centre for Biotechnology Information (NCBI) Pub-Chem database. These molecules were downloaded in Canonical SMILES format and converted to Protein Data Bank (PDB) coordinates file using Online SMILIS translate (<http://cactus.nci.nih.gov/translate/>). The selected ligand molecules were checked through the Molinspiration online server (<http://www.molinspiration.com/cgi-bin/properties>) for identifying their drug-likeness properties, and only 89 molecules that obey the Lipinski’s rule of five were used for further docking analysis (Table 2).

3.4.2 Docking studies of natural inhibitors with GPR120 model

Docking studies were performed to gain insight into the binding conformation of lead molecules with GPR120 model. A library of 100 lead molecules was constructed and screened for satisfying the minimal criteria of ADME for

Secondary structure summary

Strand	Alpha helix	3-10 helix	Other	Total residues
12 (3.2%)	225 (59.7%)	3 (0.8%)	137 (36.3%)	377

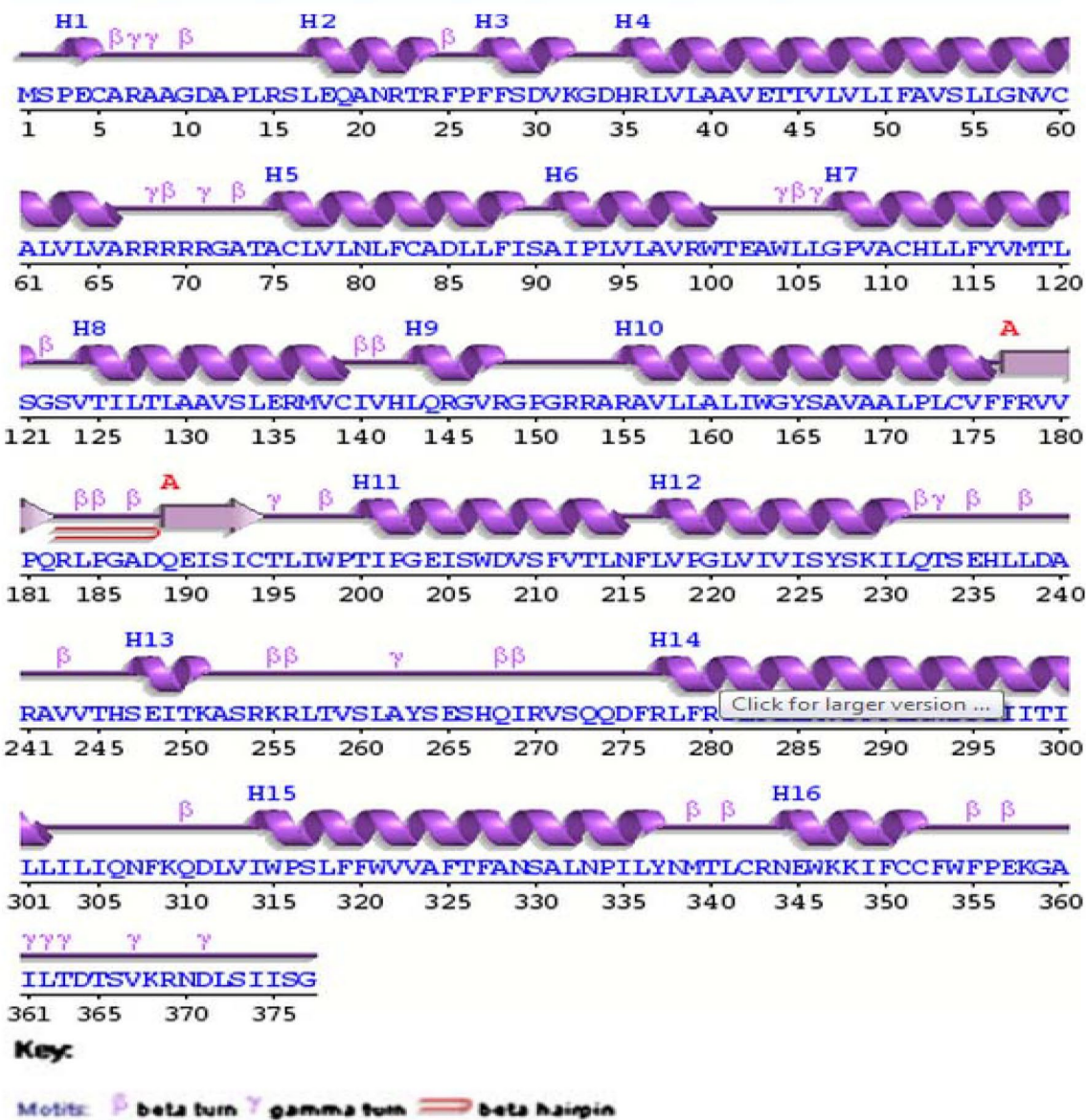


Fig. 4 Secondary structure wiring diagram for the GPR120 showing the location of secondary structure elements

further analysis, using molinspiration. Among the 100 lead molecules, 89 molecules were selected based on the criteria of satisfying Lipinski’s rule of five with zero violations. All docking calculations were carried out using AutoDock 4.0/ADT and the dlG files generated were analyzed for their binding conformations. Analysis was based on free energy of binding, lowest docked energy, and calculated RMSD values (Table 3). The total clusters of docking conformations,

with the 89 docked lead molecules, showed negative binding energies. Among all docking conformations, Silybin (Wing Ying Cheung et al. 2010), Withanolide D (Susmita Mondal et al. 2010), Limonene (Vigushin et al. 1998), and Carvacrol (Hailong et al. 2012), respectively, gave the best predicted binding-free energy of -6.19 , -6.87 , -6.17 , and -6.15 kcal mol⁻¹ to the GPR120 (Fig. 5, Table 4), and the

Table 2 Drug-like properties of natural compounds used in docking studies onto G-protein coupled receptor 120

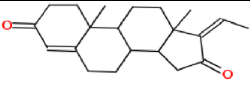
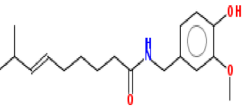
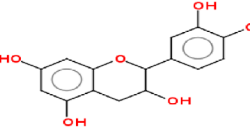
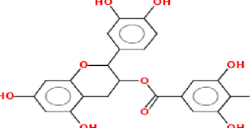
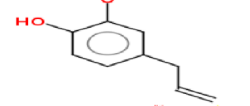
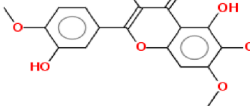
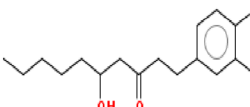
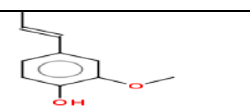
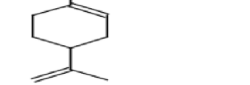
S. no.	Name of compound	Structure of compound	Mol. wt	Log <i>p</i> value	H. donor	H. acceptor	TPSA	Rot. bonds
1	Guggulsterone s E		312.453	3.62	0	2	34.14	0
2	Capsaicin		305.418	3.103	2	4	58.56	9
3	(-)-Epicatechine		290.271	1.369	5	6	110.37	1
4	Epicatechingallate		442.376	2.537	7	10	177.13	4
5	Eugenol		164.204	2.1	1	2	29.46	3
6	Eupatin		360.318	2.314	3	8	118.59	4
7	Gingerol		294.391	3.217	2	4	66.76	10
8	Isoeugenol		164.204	2.379	1	2	29.46	2
9	Limonene		136.238	3.615	0	0	0.0	1

Table 2 (continued)

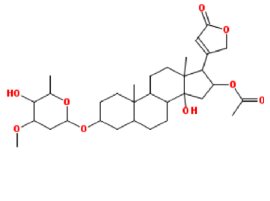
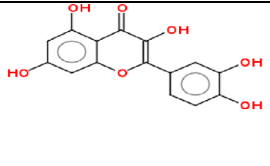
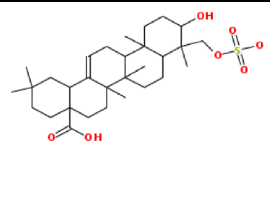
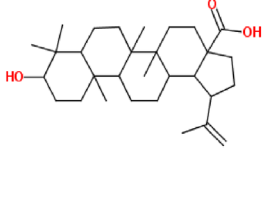
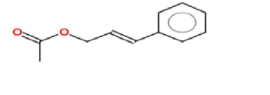
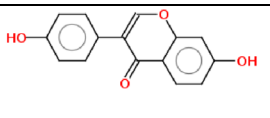
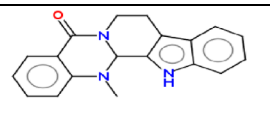
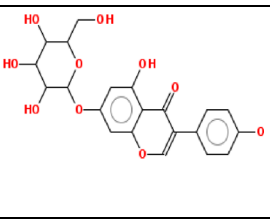
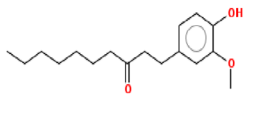
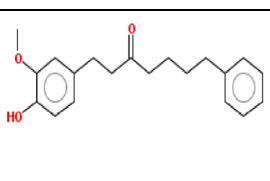
10	Oliandrin		576.727	2.877	2	9	120.76	6
11	Quercetin		302.238	1.683	5	7	131.23	1
12	Triterpenoid		552.774	3.224	3	7	121.13	4
13	Betulinic acid		456.711	7.04	2	3	57.527	2
14	Cinnamyl acetate		176.215	2.736	0	2	26.30	4
15	Daidzein		254.241	2.559	2	4	70.66	1
16	Evodimine		303.365	2.819	1	4	39.33	0
17	Genistin		432.381	0.48	6	10	170.04	4
18	Paradol		278.392	4.602	1	3	46.53	10
19	Yakuchinone-A		312.409	4.241	1	3	46.53	9

Table 2 (continued)

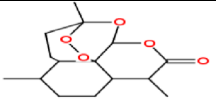
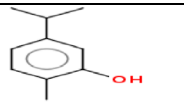
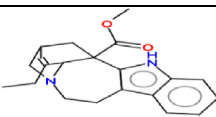
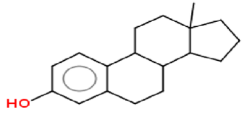
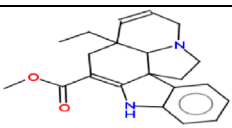
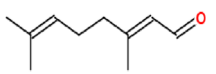
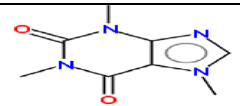
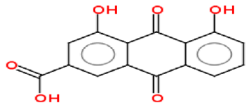
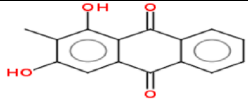
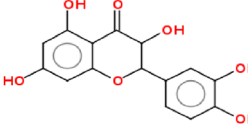
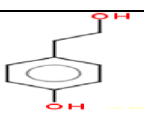
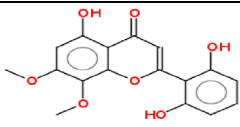
20	Artemisine		282.336	3.316	0	5	54.07	0
21	Carvacrol		150.221	3.815	1	1	20.22	1
22	Catharanthine		366.435	3.987	1	4	45.33	3
23	Phenolic Steroid		256.389	4.58	1	1	20.22	0
24	Tabersonine		336.435	3.687	1	4	41.57	3
25	Citral		152.237	3.654	0	1	17.071	4
26	Caffeine		194.194	0.063	0	6	61.83	0
27	Rhein		284.223	2.997	3	6	111.89	1
28	Rubiadin		254.241	3.719	2	4	74.59	0
29	Taxifolin		304.254	0.712	5	7	127.44	1
30	Tyrosol		138.166	1.005	2	2	40.45	2
31	Viscidulin 11		330.292	2.731	3	7	109.36	3

Table 2 (continued)

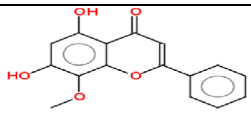
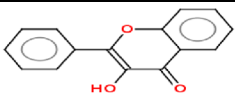
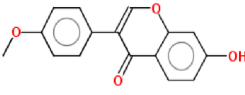
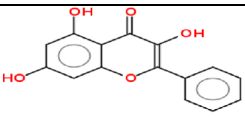
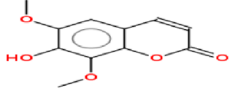
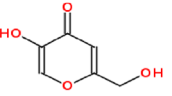
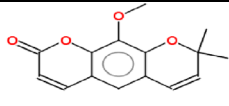
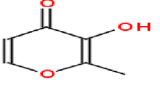
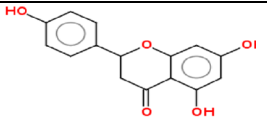
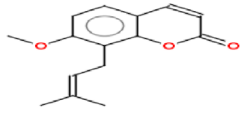
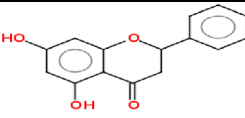
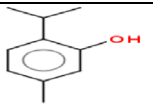
32	Wogonin		284.267	2.958	2	5	79.90	2
33	3-Hydroxyflavone		238.242	3.446	1	3	50.43	1
34	Formononetin		268.268	3.095	1	4	59.67	2
35	Galangin		270.24	2.651	3	5	90.89	1
36	Isofraxidin		222.196	1.543	1	5	68.90	2
37	Kojic acid		142.11	-0.888	2	4	70.66	1
38	Luvangetin		258.273	3.177	0	4	48.67	1
39	Maltol		126.111	-0.243	1	3	50.43	0
40	Naringenin		272.256	2.177	3	5	86.98	1
41	Osthol		244.29	3.83	0	3	39.44	3
42	Pinocembrin		256.257	2.596	2	4	66.76	1
43	Thymol		150.221	3.342	1	1	20.228	1

Table 2 (continued)

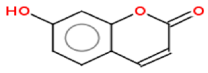
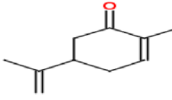
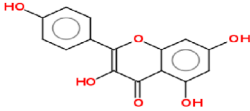
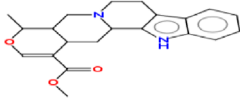
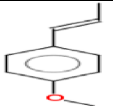
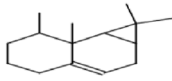
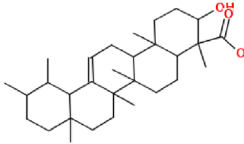
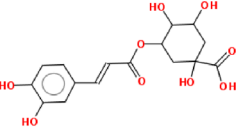
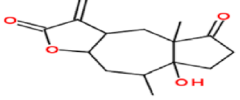
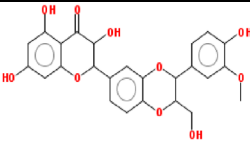
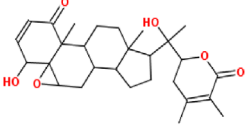
44	Ubelleferone		162.144	1.511	1	3	50.43	0
45	d-Carvone		150.221	2.513	0	1	17.071	1
46	Kaempferol		286.239	2.172	4	6	111.12	1
47	Raubasine		352.434	3.414	1	5	54.56	2
48	Anthole		148.205	3.096	0	1	9.23	2
49	Aristolene		204.357	4.837	0	0	0.0	0
50	Boswellic acid		456.711	6.789	2	3	57.52	1
51	Chlorogenic acid		354.311	-0.453	6	9	164.74	5
52	Sequiterpenela c-ctone		264.321	0.336	1	4	63.60	0
53	Silybin		482.441	1.465	5	10	135.14	4
54	Withanolide D		470.66	4.153	2	6	96.36	2

Table 2 (continued)

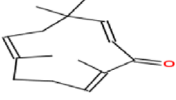
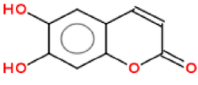
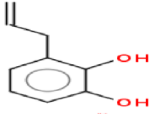
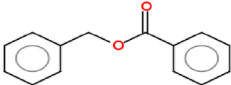
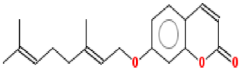
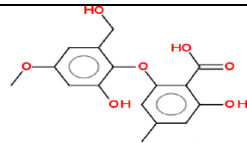
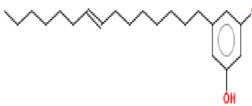
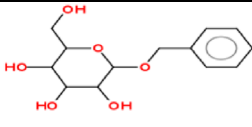
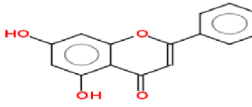
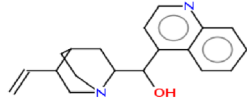
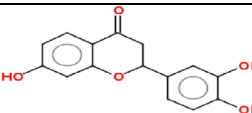
55	Zerumbone		218.34	4.204	0	1	17.07	0
56	Asculetin		178.13	1.021	2	4	70.66	0
57	Allylpyrocatechol		150.17	1.998	2	2	40.45	0
58	Ascabio		212.28	3.703	0	2	26.30	4
59	Aurapten		298.32	5.568	0	3	39.44	6
60	Barcelonic acid A		320.27	2.359	4	7	116.45	5
61	Bilobol		318.51	7.843	2	2	40.45	13
62	Benzyl glycopyranoside		270.21	-0.432	4	6	99.38	4
63	Chrysin		254.21	2.943	2	4	70.66	1
64	Cinchonine		294.38	3.028	1	3	36.35	3
65	Butin		272.256	1.711	3	5	86.98	1

Table 2 (continued)

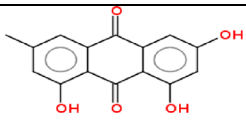
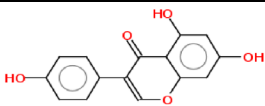
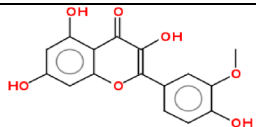
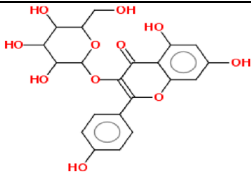
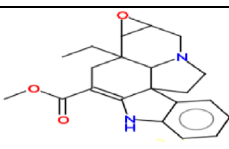
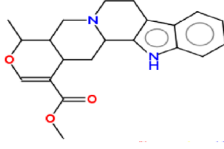
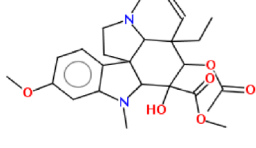
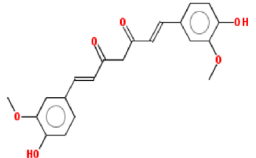
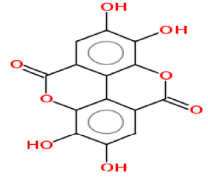
66	Emodin		270.24	3.008	3	5	94.82	0
67	Genistein		270.24	2.268	3	5	90.89	1
68	Isorhamnetin		316.25	1.99	4	7	120.35	2
69	Kaemperol-3-glucoside		448.38	0.125	7	11	190.27	4
70	Lochnericine		352.44	2.988	1	5	54.09	3
71	Raubasine		352.44	3.414	1	5	54.56	2
72	Vindiline		456.59	2.601	1	8	88.54	6
73	Curcumin		368.35	2.303	2	6	93.06	8
74	Ellagic acid		305.14	0.943	4	8	141.33	0

Table 2 (continued)

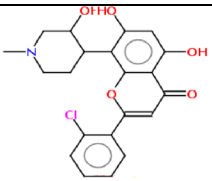
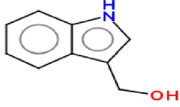
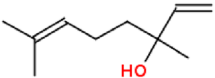
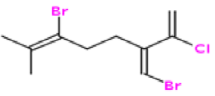
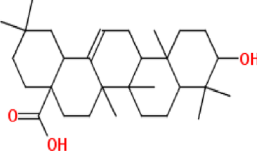
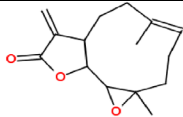
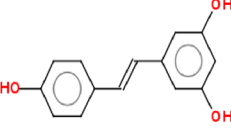
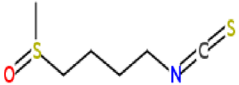
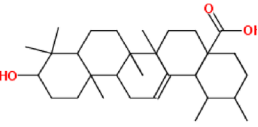
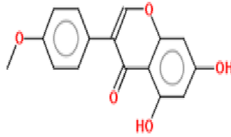
75	Flavopiridol		401.86	3.455	3	6	94.13	2
76	Indole-3-cabinol		147.17	1.426	2	2	36.01	1
77	Linalool		154.23	3.213	1	1	20.22	4
78	Halogenated mono terpene		328.45	4.826	0	0	0.0	4
79	Oleanolic acid		456.71	6.725	2	3	57.52	1
80	Parthanolide		248.32	2.089	0	3	38.88	0
81	Resveratrol		228.27	2.986	3	3	60.68	2
82	Sulforafan		177.24	1.146	0	0	29.435	5
83	Ursolic acid		456.711	6.789	2	3	33.0	1
84	Biochanin A		284.267	2.804	2	5	79.90	2

Table 2 (continued)

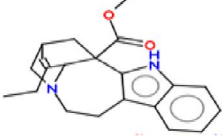
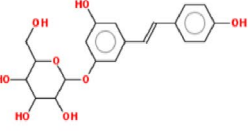
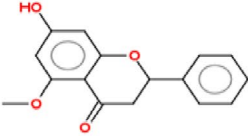
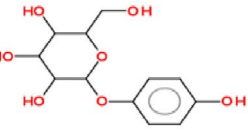
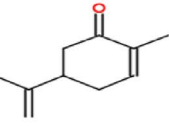
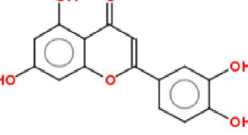
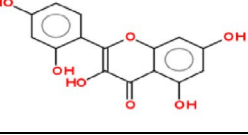
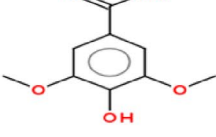
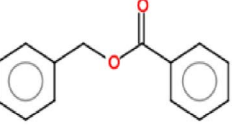
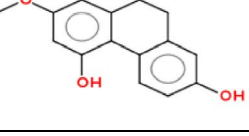
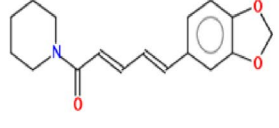
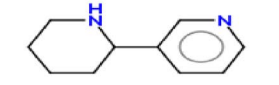
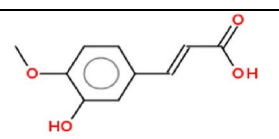
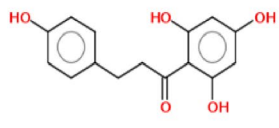
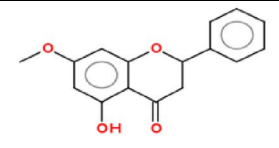
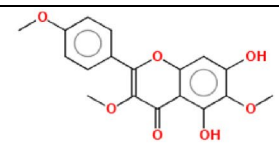
85	Catharanthine		336.435	3.987	1	4	45.33	3
86	Z-PICEID		390.388	1.199	6	8	139.83	5
87	Alpinetin		270.284	2.664	1	4	55.76	2
88	Arbutin		272.253	-0.808	5	7	119.60	3
89	Carvone		150.221	2.513	0	1	17.07	1
90	Luteoin		286.239	1.974	4	6	111.12	1
91	Morin		302	1.881	5	7	131.35	1
92	Syngic acid		198.174	1.204	2	5	75.99	3
93	Benzyl benzoate		212.248	3.703	0	2	26.305	4
94	Lusianthridine		242.274	2.04	2	3	49.69	2

Table 2 (continued)

95	Piperine		285.343	3.332	0	4	38.77	3
96	Anabesine		162.236	0.449	1	2	24.91	1
97	Isoferulic acid		194.186	1.249	2	4	66.76	3
98	Phloretin		274.272	2.656	4	5	97.98	4
99	Pinostrobin		270.284	3.132	1	4	55.76	2
100	Santin		344.319	3.0	2	7	98.36	4

corresponding references clearly uncloak which are all under clinical trails.

To confirm the binding mode of natural substrate, docking was performed on the GPR120 model; natural substrate docking revealed that the amino acids Ile50 in orange, Phe88 in yellow, Ala91 in salmon, Gly122 in green, Ile191 in red, Asn215 in magenta, Asp275 in cyan, Arg280 in orange, Phe289 in salmon, Ile295 in yellow, Ile297 in spiltpea, Ile300 in salmon, Ile301 in yellow, and Phe325 in deep-salmon color (Fig. 5) played vital role to in binding the natural substrates and except Asn215, Asp275, and Arg280 all hydrophobic amino acids.

4 Conclusion

In this study, we have developed a three-dimensional structure of GPR120 receptor through homology modeling using delta opioid 7tm receptor (PDB ID: 4N6H) as a template. The generated model was assessed by several validation tools like Procheck, Errat, whatif, ProSA 2007, and Verify 3D. All above-mentioned tools revealed that the model is reliable. This model was also submitted to PMDB server (PDB: PM0079568) for public assessment. From the available scientific literature, 100 natural compounds have been selected and 89 compounds followed

Table 3 Binding energies of docked natural compounds calculated by AutoDock

S. no.	Compound name	CID no.	Lowest binding energy (kcal/mol)	Inhibition constant ($\mu\text{M}/\text{mM}$)
1	Guggulsterones E	6439929	-4.44	73.25
2	Capsacian	1548943	-4.42	553.27
3	(-)Epicatchine	72276	-4.57	447.16
4	Euenol	3314	-3.8	308.86
5	Flavonoid	5317287	-5.05	16.30
6	Gingerol	442793	-3.86	20.13
7	Isoeugenol	853433	-5.4	110.45
8	Limonene	22311	-6.17	29.84
9	Quercetin	5280343	-4.41	583.29
10	Cinnamyl acetate	5282110	-6.05	36.67
11	Daidzein	5281708	-5.25	141.64
12	Evodimine	151289	-5.0	215.46
13	Paradol	94378	-3.99	11.90
14	Yakuchinone-A	133145	-4.35	651
15	Artemisine	68827	-5.24	144.17
16	Carvacrol	10364	-6.15	30.87
17	Catharanthine	197771	-5.48	96.85
18	Phenolic steroid	439726	4.71	354.94
19	Tabersonine	20485	-5.53	88.58
20	Citral	643779	-4.95	235.43
21	Caffeine	2519	-3.41	3190
22	Rhein	10168	-4.7	359.53
23	Rubiadin	124062	-4.53	477.98
24	Taxifolin	439533	-4.68	373.32
25	Tyrosol	10393	-4.63	400.67
26	Viscidulin 11	5322059	-5.35	120.69
27	Wogonin	5281703	-5.27	137.15
28	3-Hydroxyflavone	11349	-5.07	192.3
29	Formononetin	5280378	-4.39	608.88
30	Galangin	5281616	-4.78	315.78
31	Isofraxidin	5318565	-4.21	818.78
32	Kojic acid	3840	-4.56	457.88
33	Luvangetin	343582	-4.57	446.42
34	Maltol	8369	-4.2	829.62
35	Naringenin	932	-4.84	284.29
36	Osthol	10228	-4.87	268.85
37	Pinocembrin	68071	-5.25	141.93
38	Thymol	6989	-3.5	320
39	Ubelleferone	5281426	-5.98	41.05
40	D-Carvone	16724	-5.63	75.18
41	Kaempferol	5280863	-4.42	75.18
42	Raubasine	441975	-5.13	173.58
43	Anthole	637563	-4.1	991.37
44	Aristolene	530421	-5.41	107.8
45	Sequiterpenelactone	338659	-5.58	81.2
46	Silybin	31553	-6.87	9.12
47	Withanolide D	161671	-6.19	78.71
48	Zerumbone	5470187	-3.7	1860
49	Asculetin	5281416	-4.51	496.78

Table 3 (continued)

S. no.	Compound name	CID no.	Lowest binding energy (kcal/mol)	Inhibition constant ($\mu\text{M}/\text{mM}$)
50	Allylpyrocatechol	292101	5.09	185.5
51	Ascabiol	2345	-4.51	496.78
52	Barcelonic acid A	10358625	-3.73	18.60
53	Bilobol	5281852	-2.97	1.86
54	Benzyl glycopyranoside	188977	-4.24	782.76
55	Chrysin	5281607	-4.81	297.75
56	Cinchonine	90454	-4.89	260.15
57	Butin	92775	-5.57	82.55
58	Emodin	3220	-4.62	409.35
59	Genistein	5280961	-5.64	73.25
60	Isorhamnetin	5281654	-4.96	0.21
61	Kaemperol-3-glucoside	5282102	-4.42	571.91
62	Lochnericine	11382599	-5.19	156.7
63	Raubasine	441975	5.11	179.54
64	Vindiline	260535	-5.92	45.74
65	Curcumin	969516	-5.06	196.36
66	Flavopiridol	5287969	-3.98	1210
67	Indole-3-cabinol	3712	-3.53	2600
68	Linool	6549	3.53	2590
69	Halogenated mono terpene	11493622	-4.45	544.66
70	Parthanolide	6473881	-4.92	248.68
71	Resveratrol	445154	-5.76	59.59
72	Sulforafan	5350	-4.82	290.63
73	Biochanin A	5280373	-4.96	232.7
74	Catharanthine	197771	-5.58	0.22
75	Z-PICEID	10178463	-5.72	63.73
76	Alpinetin	4053302	-4.92	245.55
77	Arbutin	440936	-5.04	202.08
78	Carvone	7439	-5.27	137.55
79	Luteoin	5280445	-5.76	60.28
80	Morin	5281670	-4.3	705.91
81	Syrenic acid	10742	-3.27	4030
82	Benzyl benzoate	2345	-6.41	19.92
83	Lusianthridine	442702	-3.72	1.88
84	Piperine	638024	-4.59	432.06
85	Anabasine	2181	-5.69	67.14
86	Isoferulic acid	736186	-5.8	80.97
87	Phloretin	4788	-3.72	1890
88	Pinostrobin	73201	-5.29	132.4
89	Santin	5281695	-4.91	251.48

the rule of five. These 89 natural compounds were docked with GPR120 receptor and the following four compounds have exhibited the highest binding energy levels in the

order as CID: 31553 > CID: 161671 > CID: 22311 > CID: 10364, which infers the favorable rank score, docking score, and hydrogen-bonding energies. Furthermore,

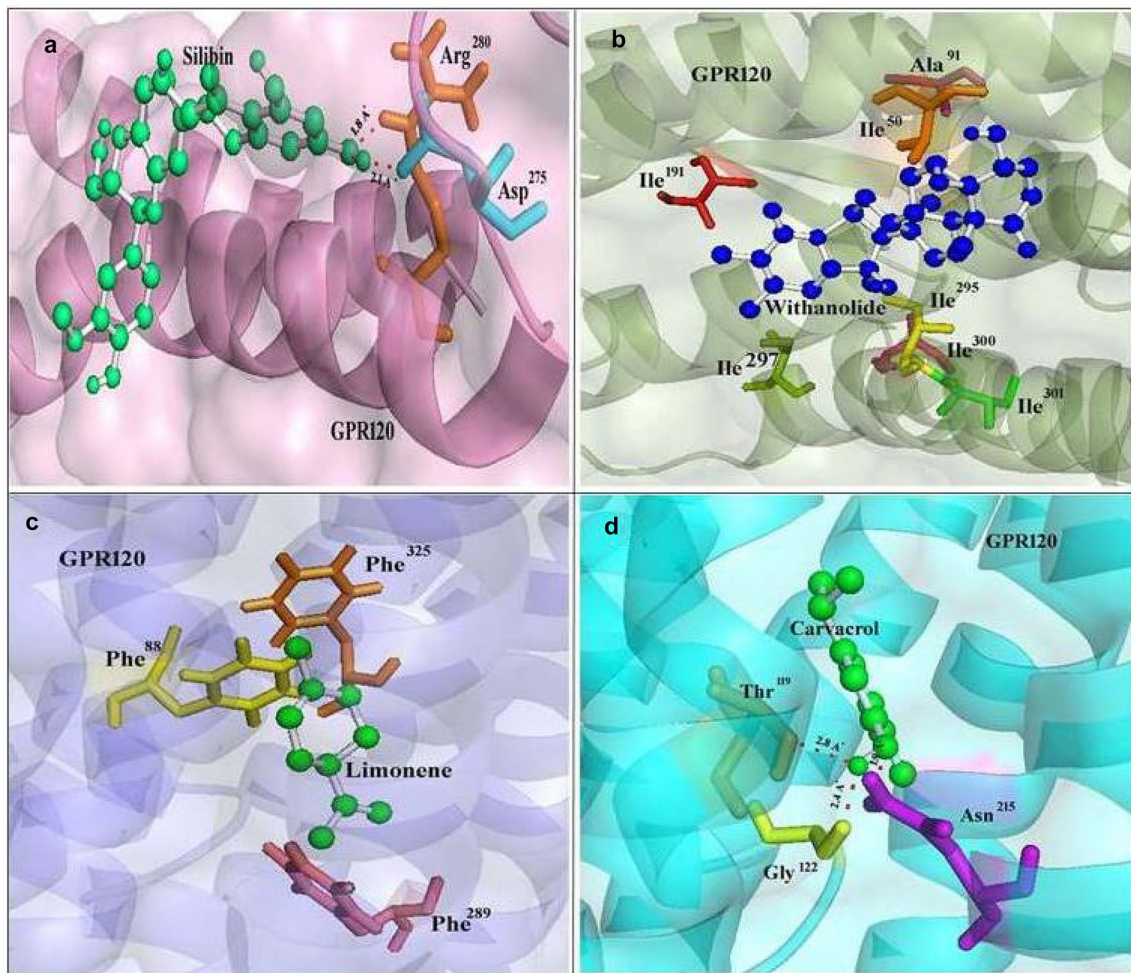


Fig. 5 **a** Docking conformation of natural compound of Silibin on G-protein coupled receptor 120 homology model. **b** Docking conformation of natural compound of Withanolide on G-protein coupled receptor 120 homology model. **c** Docking conformation natural compound of limonene on G-protein coupled receptor 120 homology

model. **d** Docking conformation of natural compound of carvacrol on G-protein coupled receptor 120 homology models. Built model of G-protein coupled receptor 120 is represented in cartoon and 60% of electrostatic surface. Ligands are represented by ball and stick and the residues interacting with are represented by stics

Table 4 High score-binding energies of docked natural compounds calculated by AutoDock

Docked molecule	Compound Cid no.	Cluster rank	Cluster number	Binding energy (kcal/mol)	RMSD (Å)
Silybin	31553	1	3	-6.87	0.49
Withanolide D	161671	1	3	-6.19	0.82
Limonene	22311	2	41	-6.17	0.13
Carvacrol	10364	1	19	-6.15	0.65

more hydrophobic interactions were observed in the binding pocket of the GPR120. These four drugs are under clinical trials and thereby help promising therapeutics for colorectal cancer.

Acknowledgements This work was supported by Department of Biotechnology (DBT) Bioinformatics Infrastructure Facility (BIF) and a Fellowship from Indian Council of Medical Research (ICMR), Govt. of India, New Delhi (to SC).

Compliance with ethical standards

Conflict of interest The authors declare that no conflicts of interest exist.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abdul Khalek FJ, Gallicano GI, Mishra L (2010) Colon cancer stem cells. *Gastrointest Cancer Res* S16–S23
- Altshul SF, Madden TL, Schaffer AA, Zhang J, Zhang W, Mille W, Lipman D (1997) Gapped BLAST and PSI BLAST—a new generation of protein data base search programs. *J Nucl Acids Res* 25:3389–3402
- American Cancer Society (2015) Cancer facts and figures. Atlanta. Available online
- Berjanskii MV, Neal S, Wishart DS (2006) PREDITOR a web server for predicting torsion angle restraints. *Nucl Acids Res* 34:W63–W69
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucl Acids Res* 28:235–242
- Bertorelle R, Rampazzo E, Pucciarelli S, Nitti D, De Rossi A (2014) Telomeres, telomerase and colorectal cancer. *World J Gastroenterol* 28(20):1940–1950
- Bowie JU, Luthy R, Eisenberg D (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253:164–170
- Castrignano T, De Meo PDO, Cozzetto D, Talamo IG, Tramontano A (2006) The PMBD Protein Model Database. *Nucl Acids Res* 34:306–309
- Colovos C, Yeates TO (1993) Verification of protein structure: patterns of nonbonded atomic interactions. *Protein Sci* 2:1511–1519
- Dorsam RT, Gutkind JS (2007) G-protein coupled receptors and cancer. *Nat Rev Cancer* 7:79–94
- Farnsworth NR, Akerele O, Bingel AS, Soejarto DD, Guo Z (1985) Medicinal plants in therapy. *Bull World Health Organ* 63:965–981
- Feigin ME (2013) Harnessing the genome for characterization of G-protein coupled receptors in cancer pathogenesis. *FEBS J* 280(19):4729–4738
- Fenalti G, Giguere PM, Katritch V, Huang XP, Thompson AA, Cherezov V, Roth BL, Stevens RC (2014) Molecular control of δ -opioid receptor signalling. *Nature* 506(7487):191–196
- Gether U (2000) Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocr Rev* 21:90–113
- Goodsell DS, Morris GM (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19(14):1639–1662
- Guex N, Peitsch MC (1997) SWISS MODEL and the Swiss PDB viewer: an environment for comparative protein modeling. *Electrophoresis* 18:2714–2723
- Hartwell JL (1971) Plants used against cancer a survey. *Lloydia* 34(4):386–425
<http://www.salilab.org/modeller/9v3>
- Ichimura A, Hirasawa A, Poulain-Godefroy O, Bonnefond A, Hara T, Yengo L (2012) Dysfunction of lipid sensor GPR120 leads to obesity in both mouse and human. *Nature* 483:350–354
- Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M (1993) Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363(6429):558–561
- Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery—methods and applications. *Nat Rev Drug Discov* 3:935–949
- Lappano R, Maggiolini M (2011) G-protein coupled receptors Novel targets for drug discovery in cancer. *Nat Rev Drug Discov* 10:47–60
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK a program to check the stereo chemical quality of protein structure. *J Appl Crystallogr* 26:283–291
- Laskowski RA, Watson JD, Thornton JM (2005) ProFunc a server for predicting function from 3D structure. *Nucl Acids Res* 33:W89–W93
- Lipinski CA, Lombardo F, Dominy W, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 46:3–26
- Luthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. *Nature* 356:83–85
- Miyamoto S, Kollman PA (1992) SETTLE an analytical version of the SHAKE and RATTLE algorithms for rigid water models. *J Comput Chem* 13:952–962
- Mondal S, Mandal C, Sangwan R, Chandra S, Mandal C (2010) Withanolide D induces apoptosis in leukemia by targeting the activation of neutral sphingomyelinase-ceramide cascade mediated by synergistic activation of c-Jun N-terminal kinase and p38 mitogen-activated protein kinase. *Mol Cancer* 5(5):239
- Morris AJ, Malbon CC (1999) Physiological regulation of G protein-linked signaling. *Physiol Rev* 79:1373–1430
- NCBI. <http://www.ncbi.nlm.nih.gov/>
- Oh DY, Talukdar S, Bae EJ, Imamura T, Hidetaka Morinaga W, Fan Q, Li P, Lu WJ, Watkins SM, Olefsky JM (2010) GPR120 is an omega-3 fatty acid receptor mediating potent anti-inflammatory and insulin sensitizing effects. *Cell* 142(5):687–698
- Oprea TI, Davis AM, Teague SJ, Leeson PD (2001) Is there a difference between leads and drugs? A historical perspective. *J Chem Inf Comput Sci* 41:1308–1315
- Rajesh E, Sankari LS, Malathi L, Krupaa JR (2015) Naturally occurring products in cancer therapy. *J Pharm Bioallied Sci* 7(Suppl 1):S181–S183
- Schulte G, Fredholm BB (2003) Signaling from adenosine receptors to mitogen-activated protein kinases. *Cell Signal* 15:813–827
- Shike M, Winawer SJ, Greenwald PH (1990) Primary prevention of colorectal cancer. The WHO Collaborating Centre for the Prevention of Colorectal Cancer. *Bull World Health Organ* 68(3):377–385
- Sippl MJ (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins* 17:355–362
- Surh YJ (2003) Cancer chemoprevention with dietary phytochemicals. *Nat Rev Cancer* 3(768–780):14
- Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W—improving the sensitivity of progressive multiple sequence alignment through sequence weighting position-specific gap penalties and weight matrix choice. *Nucl Acids Res* 22:4673–4680
- Ulloa-Aguirre A, Stanislaus D, Janovick JA, Conn PM (1999) Structure–activity relationships of G protein-coupled receptors. *Arch Med Res* 30:420–435
- Vigushin DM, Poon GK, Boddy A, English J, Halbert GW, Pagonis C, Jarman M, Coombes RC, Cancer Research Campaign Phase I/II Clinical Trials Committee (1998) Phase I and pharmacokinetic study of α -limonene in patients with advanced cancer. *Cancer Chemother Pharmacol* 42(2):111–117

- Vriend G (1990) WHATIF—a molecular modeling and drug design program. *J Mol Graph* 8:52–56
- Wing Ying Cheung C, Gibbons N, Wayne Johnson D, Lawrence Nicol D (2010) Silibinin—a promising new treatment for cancer. *Anti Cancer Agents Med Chem* 10(3):186–195
- Wu Q, Wang H, Zhao X, Shi Y, Jin M, Wan B, Xu H, Cheng Y, Ge H, Zhang Y (2013) Identification of G-protein coupled receptor 120 as a tumor promoting receptor that induces angiogenesis and migration in human colorectal carcinoma. *Oncogene* 32:5541–5550
- Yu H, Zhang Z-L, Chen J, Pei A, Hua F, Qian X, He J, Liu C-F, Xu X (2012) Carvacrol a food additive provides neuroprotection on focal cerebral ischemia/reperfusion injury in mice. *PLoS ONE* 7(3):e33584