JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

Cite This: J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

pubs.acs.org/jcim

Discovering Putative Protein Targets of Small Molecules: A Study of the p53 Activator Nutlin

Minh N. Nguyen,[†,$] Neeladri Sen,[‖,$] Meiyin Lin,[‡,$] Thomas Leonard Joseph,[†] Candida Vaz,[†,▽] Vivek Tanavde,[†,○] Luke Way,[#] Ted Hupp,[*,#] Chandra S. Verma,[*,†,§,⊥] and M. S. Madhusudhan[*,‖]

[†]Bioinformatics Institute, 30 Biopolis Street, #07-01, Matrix, Singapore 138671

[‡]Hwa Chong Institution, 661 Bukit Timah Road, Singapore 269734

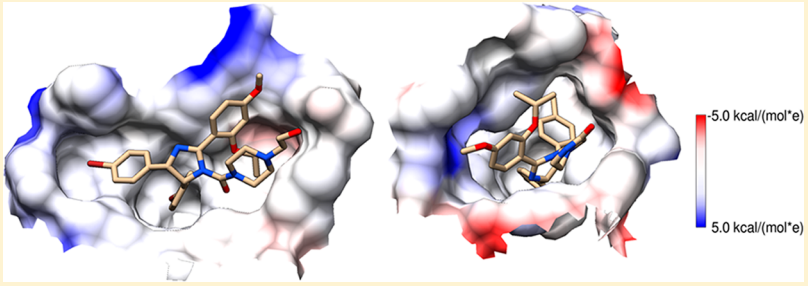
[§]Department of Biological Sciences, 16 Science Drive 4, National University of Singapore, Singapore 117558

[‖]Indian Institute of Science Education and Research Pune (IISER Pune), Pune 411008, India

[⊥]School of Biological Sciences, 60 Nanyang Drive, Nanyang Technological University, Singapore 637551

[#]University of Edinburgh, Edinburgh Cancer Research Centre, Edinburgh, U.K. EH4 2XR

Supporting Information



ABSTRACT: Small molecule drugs bind to a pocket in disease causing target proteins based on complementarity in shape and physicochemical properties. There is a likelihood that other proteins could have binding sites that are structurally similar to the target protein. Binding to these other proteins could alter their activities leading to off target effects of the drug. One such small molecule drug Nutlin binds the protein MDM2, which is upregulated in several types of cancer and is a negative regulator of the tumor suppressor protein p53. To investigate the off target effects of Nutlin, we present here a shape-based data mining effort. We extracted the binding pocket of Nutlin from the crystal structure of Nutlin bound MDM2. We next mined the protein structural database (PDB) for putative binding pockets in other human protein structures that were similar in shape to the Nutlin pocket in MDM2 using our topology-independent structural superimposition tool CLICK. We detected 49 proteins which have binding pockets that were structurally similar to the Nutlin binding site of MDM2. All of the potential complexes were evaluated using molecular mechanics and AutoDock Vina docking scores. Further, molecular dynamics simulations were carried out on four of the predicted Nutlin−protein complexes. The binding of Nutlin to one of these proteins, gamma glutamyl hydrolase, was also experimentally validated by a thermal shift assay. These findings provide a platform for identifying potential off-target effects of existing/new drugs and also opens the possibilities for repurposing drugs/ligands.

1. INTRODUCTION

Current drug discovery efforts usually identify one or a few protein target(s) and attempt to inhibit these using small molecules/peptides. Most inhibitors/drugs are identified using various computational/experimental screens followed by rounds of rational manipulation and extensive experimental validation.[1] The drugs bind selectively to a pocket on the targets because of complementarity in shape and physicochemical properties. However, the diversity of protein shapes is limited, and it is likely that similar binding pockets could be found in other proteins. The binding of the drug to these off-target proteins could either lead to adverse drug reactions[2] or indicate an alternate use of the drug.[3] A relevant observation to support this claim would be that of protein kinases, which have structurally similar ATP binding pockets.[4] A drug designed against the ATP binding pocket of one kinase often also binds to other kinases,[5] making the drug less specific. An efficient drug discovery effort would be strengthened with the identification of putative binding pockets on off-target proteins.

Several computational tools have been developed for identification of structural similarities between the 3-dimensional structures of proteins or parts of proteins.[6−13] Our tool CLICK[6] can compare the 3D structures or even substructures of molecules. The CLICK program creates small cliques of points from representative atoms of spatially proximal amino

Received: October 30, 2018
Published: February 22, 2019

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

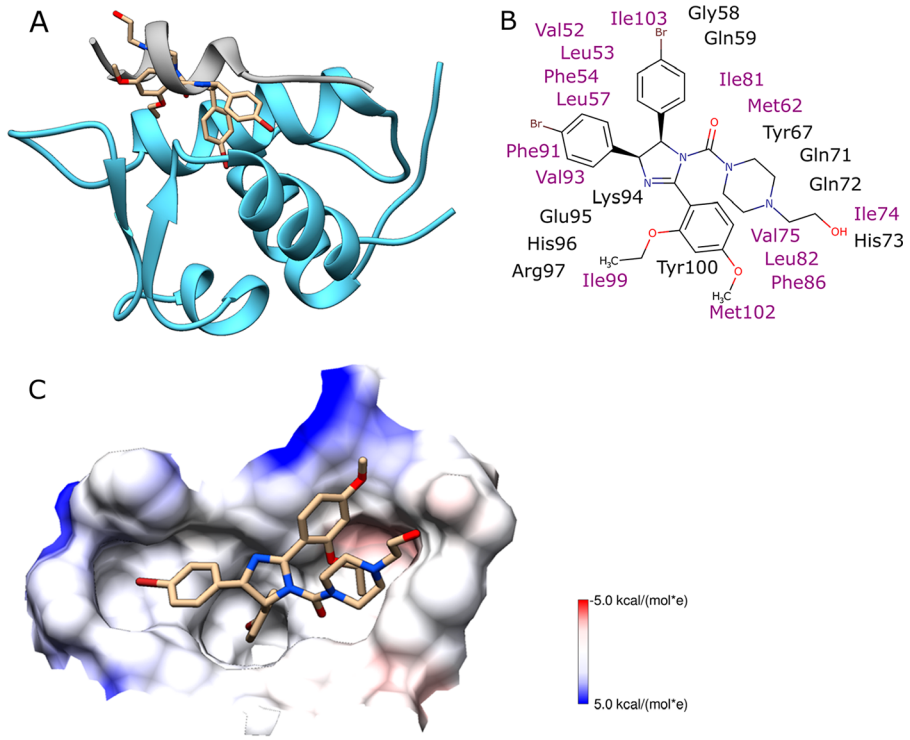


Figure 1. (A) Superimposition of the crystal structure complexes of MDM2 (blue ribbons) with Nutlin-2 (tan sticks, PDB ID: 1rv1) and p53 (gray ribbons, PDB ID: 1ycr). (B) 2D representations of the interactions, within 6 Å, of Nutlin 2 with residues of MDM2 from the crystal structures and MD snapshots. Hydrophobic residues are colored in purple. (C) Surface representation of the binding pocket residues shown in (B) along with the bound Nutlin (tan sticks). The binding pocket is colored as per the Chimera[41] rendered columbic charge representation, where shades of blue and red represent positively and negatively charged regions, respectively.

acid residues (3−7 in number). These cliques are then superimposed by a 3D least-squares fit. To guide the matching of cliques, other features such as solvent accessibility, secondary structure, and residue depth can also be used. CLICK is capable of aligning structures with dissimilar topologies, conformations, or even molecular types. These unique properties make CLICK particularly well suited for comparing protein substructures, such as ligand binding sites.[6]

In this study, we test the efficacy of CLICK in identifying similar binding pockets of the small molecule Nutlin. Nutlin is known to bind MDM2, a negative regulator of the tumor suppressor protein p53.[14,15] Upregulated activities of MDM2 in several cancers results in increased degradation of p53 and hence is being pursued as a potential therapeutic target.[16] The interactions between MDM2 and Nutlin have been explored using several experimental techniques including crystallography (PDB ID: 1RV1/4J3E),[15,17,18] which shows that Nutlin occupies a hydrophobic pocket in the N-terminal domain of MDM2 and mimics key residues of the N-terminal region of p53 which occupy the same pocket (PDB ID: 1YCR)[19](Figure 1a). In cancer cells with upregulated levels of MDM2,[20] the abolition of the interactions between MDM2 and p53 by peptides or small molecules such as Nutlin is demonstrably sufficient to induce activation of p53.[15] Thus, Nutlin-like molecules could be potential drugs for such cancers, several of which are currently in clinical trials.[21] However, there is some evidence of toxicity by this approach,[22] and hence, there is a need to develop robust methods to prescreen potential drugs for potential adverse reactions. The current study provides a tentative list of proteins that may be targets of Nutlin in addition to MDM2. If validated, this technique has the potential of adding specificity filters in drug design for detecting off target effects of small molecule compounds resulting in cost savings. We have used the program CLICK to predict potential protein targets of Nutlin other than MDM2. Further, we computed the binding free energy of Nutlin with these proteins using the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) protocol of Amber11. We also compared our binding poses with the poses predicted by docking Nutlin onto the CLICK predicted binding pocket using AutoDock Vina. Molecular dynamics simulations were carried out on 4 of the putative proteins complexed to Nutlin to probe the stability of these complexes. Finally, we experimentally validated the binding of Nutlin to one of the predicted off target proteins and showed an impact on its thermostability.

2. MATERIALS AND METHODS

2.1. Nutlin Binding Sites on MDM2. The binding site residues within 6 Å of Nutlin-2 and of Nutlin-3a were extracted from the structures of their complexes with MDM2 (PDB IDs: 1RV1:B for Nutlin-2 and PDB ID: 4J3E:A for Nutlin-3a). Henceforth, Nutlin will be used to refer to both Nutlin-2 and Nutlin-3a.

In order to account for binding site flexibility, Nutlin binding site residues were also extracted from 5 snapshots from a molecular dynamics (MD) simulation trajectory of MDM2, at 3, 6, 9, 12, and 15 ns.[17]

2.2. Data Set of Representative Protein Structures. To search for putative non-MDM2 proteins that could bind Nutlin, 4239 crystal structures of proteins were selected from the PDB using the program PISCES,[23] such that the proteins (a) were all human proteins, (b) were resolved at resolutions

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Table 1. Details of the Actual and Predicted Human Proteins That Bind Nutlin

Name of actual/ predicted Nutlin binding proteins	PDB code	Binding energy with Nutlin-2 (kcal/mol)	Binding energy with Nutlin-3a (kcal/mol)	AutoDock Vina binding energy with Nutlin-3a (kcal/mol)	RMSD (Å)[a]	Template[b]	Gene name	CLICK Z-score	CLICK SO (%)	CLICK RMSD (Å)	Number of residues within 6 Å of Nutlin	Seq Identity (%)	Number of aligned residues in Active/ functional sites	Pathway protein involved	Nutlin binding sites with superimposed residues in their active/ functional sites
AP-2 Complex Subunit Beta	2g30:A	−75.57	−57.08	−11.4	2.3	3 ns	AP2B1	6.85	100.00	1.95	48	7.41	0 res	Endocrine and other factor-regulated calcium reabsorption; Endocytosis; Huntington's disease; Synaptic vesicle cycle	No
Glutataryl-CoA Dehydro-genase, mitochon-drial	1siq:A	−70.65	−55.99	−11.0	0.54	15 ns	GCDH	6.85	97.06	1.94	53	7.41	0 res	Fatty acid metabolism; Lysine degradation; Metabolic pathways; Tryptophan metabolism	No
Gamma-glutamyl hydrolase	1l9x:A	−67.93	−57.58	−10.9	0.56	3 ns	GGH	6.85	100.00	1.97	46	11.11	1 res	Folate biosynthesis	Yes
Steryl-sulfatase	1p49:A	−64.35	−56.49	−10.0	0.33	6 ns	STS	7.58	100.00	1.67	37	3.70	1 res	Steroid hormone biosynthesis	Yes
Peptidylprolyl Isomerase domain and WD Repeat Containing Protein 1	2a2n:A	−61.82	−53.03	−4.0	1.35	Crystal	PPWD1	2.24	85.12	2.71	48	0.00	0 res	Protein folding	No
Stromelysin-1	1hy7:A	−61.00	−63.05	−10.6	0.51	Crystal	MMP3	2.24	84.30	2.77	47	3.70	8 res	Rheumatoid arthritis	Yes
MDM2 (3 ns)	-	−55.31	−50.83			3 ns					25				Yes
Interferon-gamma	1fyh:A	−48.92	−39.89	−8.6	5.16	Crystal	IFNG	2.24	84.30	2.70	27	7.41	0 res	African trypanosomiasis; Amoebiasis; Antigen processing and presentation; Chagas disease (American trypanosomiasis); Cytokine-cytokine receptor interaction; Influenza A; Jak-STAT signaling pathway; Leishmaniasis; Malaria; Measles; Natural killer cell mediated cytotoxicity; Osteoclast differentiation;	No

J. Chem. Inf. Model. XXXX, XXX, XXX−XXX DOI: 10.1021/acs.jcim.8b00762

Table 1. continued

Name of actual/ predicted Nutlin binding proteins	PDB code	Binding energy with Nutlin-2 (kcal/mol)	Binding energy with Nutlin-3a (kcal/mol)	AutoDock Vina binding energy with Nutlin-3a (kcal/mol)	RMSD (Å)[a]	Template[b]	Gene name	CLICK Z-score	CLICK SO (%)	CLICK RMSD (Å)	Number of residues within 6 Å of Nutlin	Seq Identity (%)	Number of aligned residues in Active/ functional sites	Pathway protein involved	Nutlin binding sites with superimposed residues in their active/ functional sites
														Proteasome; Regulation of autophagy; Salmonella infection; T cell receptor signaling pathway; TGF-beta signaling pathway; Toxoplasmosis; Tuberculosis; Type I diabetes mellitus	
MDM2 (6 ns)	-	−48.44	−44.93			6 ns					24				Yes
Human Dead box RNA helicase DDX52	3dkp:A	−47.92	−39.60	−8.8	0.52	6 ns	DDX52	6.61	96.67	1.97	32	11.11	0 res	RNA helicase	No
MDM2 (crystal)	1rv1:B/ 4j3e:A	−47.21	−45.02	−8.4	0.48	Crystal	MDM2	13.41	100.00	0.00	24/25	100.0			Yes
MDM2 (12 ns)		−45.18	−43.72			12 ns					24				Yes
MDM2 (15 ns)		−44.66	−43.11			15 ns					24				Yes
HDM2	2axi:A	−42.74	−41.60			Crystal	MDM2	11.79	100.00	0.61	29	95.8			Yes
Hexokinase-2	2nzt:A	−39.83	−34.20	−6.7	4.71	6 ns	HK2	6.61	96.67	1.91	22	7.41	0 res	Amino sugar and nucleotide sugar metabolism; Butirosin and neomycin biosynthesis; Carbohydrate digestion and absorption; Fructose and mannose metabolism; Galactose metabolism; Glycolysis/ Gluconeogenesis; Insulin signaling pathway; Metabolic pathways; Starch and sucrose metabolism; Type II diabetes mellitus	No
Multiple PDZ domain protein	2o2t:A	−39.46	−38.01	−7.4	6.58	Crystal	MPDZ	3.69	79.17	2.19	21	7.41	0 res	Tight junction	No

J. Chem. Inf. Model. XXXX, XXX, XXX−XXX
DOI: 10.1021/acs.jcim.8b00762

Table 1. continued

Name of actual/ predicted Nutlin binding proteins	PDB code	Binding energy with Nutlin-2 (kcal/mol)	Binding energy with Nutlin-3a (kcal/mol)	AutoDock Vina binding energy with Nutlin-3a (kcal/mol)	RMSD (Å)[a]	Template[b]	Gene name	CLICK Z-score	CLICK SO (%)	CLICK RMSD (Å)	Number of residues within 6 Å of Nutlin	Seq Identity (%)	Number of aligned residues in Active/ functional sites	Pathway protein involved	Nutlin binding sites with superimposed residues in their active/ functional sites
MDM4	3fea:A	−38.62	−39.32	−6.1	0.73	Crystal	MDM4	10.64	100.00	1.00	24	51.85			Yes
Enhancer of MRNA-Decapping protein 3	2vc8:A	−35.55	−33.67	−5.8	0.79	Crystal	EDC3	3.94	87.50	2.09	22	14.81	0 res	Gene Expression; Metabolism of RNA	No
Ketohexoki-nase	2hlz:A	−34.48	−28.51	−7.7	6.23	Crystal	KHK	3.94	83.33	2.00	22	3.70	1 res	Fructose and mannose metabolism; Metabolic pathways	Yes
Programmed cell death protein 10	3ajm:A	−31.54	−30.85	−6.8	2.41	Crystal	PDCD10	6.12	87.50	1.76	23	7.41	5 res	Apoptotic pathways	Yes
Endothelial Nitric-oxide synthase	1m9m:A	−30.89	−29.16	−6.8	2.32	Crystal	NOS3	5.64	83.33	1.84	23	7.41	0 res	Arginine and proline metabolism; Calcium signaling pathway; Metabolic pathways; VEGF signaling pathway	No
Glucocorti-coid receptor 2	3gn8:A	−30.13	−28.72	−7.5	3.60	Crystal	NCOA2	5.15	87.50	2.11	26	3.70	0 res	Developmental biology; Metabolism	No
INTERLEU-KIN-5	1hul:A	−30.01	−29.51	−7.0	4.67	Crystal	IL5	4.91	75.00	1.87	27	3.70	0 res	Allograft rejection; Asthma; Autoimmune thyroid disease; Cytokine-cytokine receptor interaction; Fc epsilon RI signaling pathway; Hematopoietic cell lineage; Intestinal immune network for IgA production; Jak-STAT signaling pathway; T cell receptor signaling pathway	No
CHITI-NASE-3 like protein 1	1hjx:A	−29.92	−26.64	−6.9	4.20	Crystal	CHI3L1	3.69	75.00	1.94	25	7.41	0 res	Amino sugar and nucleotide sugar metabolism	No
RecQ-mediated genome instability protein 1	3mxn:A	−29.63	−26.77	−6.2	3.43	Crystal	RMI1	3.94	87.50	2.08	22	3.70	0 res	Fanconi anemia pathway	No

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Table 1. continued

Name of actual/ predicted Nutlin binding proteins	PDB code	Binding energy with Nutlin-2 (kcal/mol)	Binding energy with Nutlin-3a (kcal/mol)	AutoDock Vina binding energy with Nutlin-3a (kcal/mol)	RMSD (Å)[a]	Template[b]	Gene name	CLICK Z-score	CLICK SO (%)	CLICK RMSD (Å)	Number of residues within 6 Å of Nutlin	Seq Identity (%)	Number of aligned residues in Active/ functional sites	Pathway protein involved	Nutlin binding sites with superimposed residues in their active/ functional sites
GTP-binding protein Di-Ras1	2gf0:A	−28.64	−28.58	−6.2	2.66	Crystal	DIRAS1	4.18	79.17	1.99	19	11.11	0 res	GTPase activity	No
Dual-specificity Phospha-tase DUPD1	2y96:A	−28.09	−29.93	−6.3	2.31	Crystal	DUPD1	3.94	79.17	2.19	23	0.00	0 res	Dephosphorylation activity	No
Cytochrome P450 1B1	3pm0:A	−28.07	−26.67	−7.7	3.72	Crystal	CYP1B1	5.64	83.33	1.89	23	3.70	2 res	Metabolism of xenobiotics by cytochrome P450; Steroid hormone biosynthesis; Tryptophan metabolism	Yes
Ubiquitin thioesterase ZRANB1	3zrh:A	−27.61	−26.70	−6.6	3.68	Crystal	ZRANB1	3.94	83.33	2.11	22	3.70	0 res	Cell migration, hydrolysis of ester, thioester, amide, peptide, isopeptide	No
Leukotriene C4 synthase	3pcv:A	−27.47	−26.89	−6.8	3.29	15 ns	LTC4S	6.85	91.18	1.59	21	0.00	1 res	Arachidonic acid metabolism; Metabolic pathways	Yes
Golgi reassembly stacking protein 2	3rle:A	−26.99	−23.67	−6.8	3.17	Crystal	GORASP2	3.69	79.17	2.12	21	7.41	2 res	Assembly and golgi stacking of golgi cisternae	Yes
Phosphopan-tothenoyl-cysteine synthetase	1p9o:A	−26.49	−22.20	−7.0	3.26	Crystal	PPCS	3.69	79.17	2.13	24	0.00	0 res	Metabolic pathways; Pantothenate and CoA biosynthesis	No
Serine/ threonine-protein kinase/ endoribo-nuclease IRE1	3p23:A	−25.84	−25.49	−7.0	3.97	Crystal	ERN1	4.66	79.17	1.77	21	7.41	0 res	Alzheimer's disease; Protein processing in endoplasmic reticulum	No
Cytochrome P450 2C9	1r9o:A	−25.67	−22.20	−8.0	3.70	Crystal	CYP2C9	4.66	79.17	1.78	24	0.00	0 res	Arachidonic acid metabolism; Drug metabolism - cytochrome P450; Linoleic acid metabolism; Metabolic pathways; Metabolism of xenobiotics by cytochrome P450; Retinol metabolism	No

J. Chem. Inf. Model. XXXX, XXX, XXX−XXX
DOI: 10.1021/acs.jcim.8b00762

Table 1. continued

Name of actual/ predicted Nutlin binding proteins	PDB code	Binding energy with Nutlin-2 (kcal/mol)	Binding energy with Nutlin-3a (kcal/mol)	AutoDock Vina binding energy with Nutlin-3a (kcal/mol)	RMSD (Å)[a]	Template[b]	Gene name	CLICK Z-score	CLICK SO (%)	CLICK RMSD (Å)	Number of residues within 6 Å of Nutlin	Seq Identity (%)	Number of aligned residues in Active/ functional sites	Pathway protein involved	Nutlin binding sites with superimposed residues in their active/ functional sites
C-type lectin domain family 4 member K	3p7g:A	−25.57	−24.70	−6.8	3.60	Crystal	CD207	4.42	75.00	1.69	21	0.00	1 res	Immune system	Yes
Chloride intracellular channel protein 2	2r4v:A	−24.84	−20.66	−6.4	3.78	Crystal	CLIC2	4.42	79.17	1.89	20	3.70	0 res	Chloride transmembrane transport, glutathione peroxidase activity	No
Lanosterol 14-alpha demethy-lase	3ld6:A	−24.76	−23.48	−7.1	3.25	Crystal	CYP51A1	4.18	83.33	2.16	28	3.70	0 res	Metabolic pathways; Steroid biosynthesis	No
Heparin-binding growth factor 1	3o3q:A	−24.54	−24.07	−5.4	5.64	Crystal	FGF1	3.94	79.17	2.10	24	7.41	0 res	MAPK signaling pathway; Melanoma; Pathways in cancer; Regulation of actin cytoskeleton	No
Hypothetical ubiquitin-conjugating enzyme LOC55284	2a7l:A	−24.44	−23.86	−6.8	3.22	Crystal	UBE2W	5.39	87.50	1.86	24	11.11	0 res	Ubiquitin mediated proteolysis	No
Protein-Glutamine Gamma-Glutamyl-transferase KNo	2xzz:A	−24.12	−25.41	−6.6	2.45	Crystal	TGM1	2.97	79.17	2.30	21	0.00	0 res		No
General vesicular transport factor p115	2w3c:A	−23.75	−29.05	−6.0	3.58	Crystal	USO1	4.42	75.00	1.88	20	0.00	0 res		No
Protein SET	2e50:A	−23.59	−22.18	−8.1	5.43	Crystal	SET	2.97	70.83	1.95	22	3.70	0 res	Gene expression; Metabolism of RNA	No
Receptor tyrosine-protein kinase erbB-4	2r4b:A	−23.48	−18.18	−5.8	11.45	Crystal	ERBB4	4.42	75.00	1.77	15	0.00	2 res	Tyrosine kinase activity,	Yes
Glutathione Transferase Zeta	1fw1:A	−23.12	−17.80	−6.2	6.88	Crystal	GSTZ1	3.45	79.17	2.09	21	0.00	0 res	Drug metabolism - cytochrome P450; Glutathione metabolism; Metabolic pathways; Metabolism of xenobiotics by	No

J. Chem. Inf. Model. XXXX, XXX, XXX−XXX
DOI: 10.1021/acs.jcim.8b00762

Table 1. continued

Name of actual/ predicted Nutlin binding proteins	PDB code	Binding energy with Nutlin-2 (kcal/mol)	Binding energy with Nutlin-3a (kcal/mol)	AutoDock Vina binding energy with Nutlin-3a (kcal/mol)	RMSD (Å)[a]	Template[b]	Gene name	CLICK Z-score	CLICK SO (%)	CLICK RMSD (Å)	Number of residues within 6 Å of Nutlin	Seq Identity (%)	Number of aligned residues in Active/ functional sites	Pathway protein involved	Nutlin binding sites with superimposed residues in their active/ functional sites
														cytochrome P450; Tyrosine metabolism	
INTERLEUKIN-17A	2vxs:A	−22.33	−21.15	−7.0	3.34	Crystal	IL17A	3.45	75.00	2.09	20	3.70	0 res	Cytokine-cytokine receptor interaction; Rheumatoid arthritis	No
Serine-pyruvate aminotransferase	3r9a:A	−21.59	−19.08	−6.7	4.63	Crystal	AGXT	3.69	83.33	2.06	23	0.00	0 res	Alanine, aspartate and glutamate metabolism; Glycine, serine and threonine metabolism; Metabolic pathways; Peroxisome	No
Receptor tyrosine-protein kinase erbB-2	3pp0:A	−21.55	−21.60	−6.3	5.58	Crystal	ERBB2	4.42	75.00	1.75	17	0.00	1 res	Calcium signaling pathway; Endocytosis; ErbB signaling pathway	Yes
Proto-oncogene Tyrosine protein Kinase Receptor RET	2ivs:A	−21.54	−21.77	−6.6	12.51	Crystal	RET	5.64	79.17	1.38	17	7.41	1 res	Endocytosis; Pathways in cancer; Thyroid cancer	Yes
Sulfotransferase 1C2	2gwh:A	−21.39	−20.20	−5.9	4.68	Crystal	SULT1C4	5.15	79.17	1.83	22	7.41	1 res	Metabolism	Yes
Fibroblast growth factor homologous factor 1	1q1u:A	−21.18	−19.79	−6.0	2.98	Crystal	FGF12	2.97	70.83	1.95	22	0.00	0 res	MAPK signaling pathway; Melanoma; Pathways in cancer; Regulation of actin cytoskeleton	No
Pigment Epithelium-Derived factor	1imv:A	−20.41	−16.13	−6.0	2.56	Crystal	SERPINF1	3.69	75.00	1.94	20	3.70	1 res	Inhibition of angiogenesis	Yes
Activin receptor type IIB	2qlu:A	−19.77	−16.61	−6.1	4.60	Crystal	ACVR2B	3.21	70.83	1.71	20	7.41	0 res	Cytokine-cytokine receptor interaction; TGF-beta signaling pathway	No
5′-Deoxy 5′-Methylthioadeno-	1cb0:A	−18.41	−24.24	−7.5	9.03	Crystal	MTAP	4.42	75.00	1.82	20	0.00	2 res	Cysteine and methionine	Yes

J. Chem. Inf. Model. XXXX, XXX, XXX−XXX
DOI: 10.1021/acs.jcim.8b00762

Table 1. continued

Name of actual/ predicted Nutlin binding proteins	PDB code	Binding energy with Nutlin-2 (kcal/mol)	Binding energy with Nutlin-3a (kcal/mol)	AutoDock Vina binding energy with Nutlin-3a (kcal/mol)	RMSD (Å)[a]	Template[b]	Gene name	CLICK Z-score	CLICK SO (%)	CLICK RMSD (Å)	Number of residues within 6 Å of Nutlin	Seq Identity (%)	Number of aligned residues in Active/ functional sites	Pathway protein involved	Nutlin binding sites with superimposed residues in their active/ functional sites
sine phosphorylase														metabolism; Metabolic pathways	
Alkaline phosphatase, placental type	3mk1:A	−17.82	−14.82	−5.7	6.93	Crystal	ALPP	3.69	79.17	1.97	27	3.70	0 res	Folate biosynthesis; Metabolic pathways	No
Tumor necrosis factor receptor superfamily member 1B	3alq:R	−17.17	N/A	−4.6	4.56	Crystal	TNFRSF1B	3.21	70.83	1.83	18	0.00	0 res	Adipocytokine signaling pathway; Amyotrophic lateral sclerosis (ALS); Cytokine-cytokine receptor interaction	No
Dual specificity mitogen-activated protein kinase kinase 4	3aln:A	−16.81	−19.05	−6.4	6.17	Crystal	MAP2K4	3.69	75.00	1.93	19	14.81	1 res	Chagas disease (American trypanosomiasis); Epithelial cell signaling in *Helicobacter pylori* infection; ErbB signaling pathway; Fc epsilon RI signaling pathway; GnRH signaling pathway; HTLV-I infection; Influenza A; MAPK signaling pathway; Toll-like receptor signaling pathway	Yes

[a]RMSD between central imidazole ring of Nutlin-3a as predicted by CLICK (after Amber11 energy minimization) and as predicted by AutoDock Vina. [b]Column informs if the putative binding site has been identified using the crystal structure or the MD snapshot of Nutlin (3 ns, 6 ns, 9 ns, 12 ns, 15 ns).

J. Chem. Inf. Model. XXXX, XXX, XXX−XXX
DOI: 10.1021/acs.jcim.8b00762

higher than 3 Å, (c) had R-factor < 0.3, (d) were not more than 95% sequentially identical to one another, and (e) had a length greater than 40 residues. The complete list of human proteins used for the study can be found in Supporting Information Section 2.

2.3. CLICK Searching. The Nutlin binding site(s) were structurally superimposed on proteins in our data set using CLICK (http://cospi.iiserpune.ac.in/click). Our CLICK program superimposes two molecular structures, even if they are topologically dissimilar, by a 3D least-square fit of their representative atoms. In this case, the C^{α} and C^{β} atoms of the residues were chosen as representative atoms for structural superimposition. A clique of points is made with the representative atoms such that no pair of atoms within a clique is separated by more than a distance threshold of 10 Å. The clique size was earlier optimized to contain 7 residues.[11] To ensure that equivalent residues occupy similar environments in their respective proteins, a match was only made if the residue depth difference was less than 2.25 Å. Residue depth is defined as the closest distance of the residue from bulk solvent.[24] The CLICK program produces a Z-score for reliability of match, and we had previously established that a score of 2 and above was indicative of a significant comparison. The objective of these comparisons was to match regions on proteins that structurally resembled the Nutlin binding pocket on MDM2.

2.4. Eliminating Hits Clashing with Nutlin. In our protocol, we superimposed the proteins from the database (section 2.2) onto the Nutlin binding sites. Proteins that had regions that matched the Nutlin binding sites with significant Z-scores were termed as "hits". Both Nutlin-2 and Nutlin-3a were then independently transferred as rigid bodies onto the hit protein to form a complex. The complex was energy minimized using Amber11.[25] Steric hindrances in the complexes (with either of the Nutlins) were quantified by a clash score. A clash results when the intermolecular atomic distance between two non-hydrogen atoms of the hit and Nutlin is less than 2.0 Å. Ideally, we would want no short contacts between the atoms of the protein and the ligand. However, we tolerated a few short contacts, empirically set to 5 short contacts involving the protein side chains and 1 short contact involving the protein main chain. Our tolerance levels were decided upon following the logic that short contacts with the side chain could be more easily resolved (moving individual side chains) as opposed to making conformational changes to the main chain.

2.5. Validation by Scoring the Poses of Nutlin. *2.5.1. Single Point Binding Energy Calculations and Hydrophobicity of the Binding Pocket.* A single point binding energy of the observed/predicted complex was computed using the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method with the GB module of Amber11. The binding free energy of Nutlin was also calculated on the human analogs of MDM2, Hdm2, and MDM4, which are structurally similar to MDM2. The binding energy, as computed here, is essentially the enthalpy change (ΔH) as a result of binding. The more negative ΔH is, the tighter Nutlin binds to the protein target. The shortlisted hits were rank ordered by binding energies (Table 1) calculated using Amber11.

2.5.2. Docking of Nutlin-3a onto the Target Protein Using AutoDock Vina. To validate the binding site and binding pose, AutoDock Vina[26] was used to dock Nutlin-3a onto the energy minimized structures of the target proteins obtained from Section 2.4. The AutoDock exercise was not carried out with the crystal structures of these protein as in many of the cases the binding pocket could not accommodate the ligand before conformational changes (Supporting Information section 1). Polar hydrogens were added and charges were assigned to atoms of both the target protein and Nutlin-3a. A 30*30*30 $Å^3$ box (dimensions chosen considering the size of Nutlin) for docking of Nutlin-3a was centered on its N1 atom from the CLICK predicted structure. The binding free energy and the corresponding RMSD (calculated between the central 5 membered ring atoms − N, N1, C10, C11, C18) to the CLICK computed binding pose were calculated for the best AutoDock pose (Table 1).

2.6. Identifying Functions/Pathways of Putative Nutlin Targets. The UniProt (Universal Protein Resource, http://www.uniprot.org) database was mined for information relating to the biological role and function of the hit protein, its interactions with other proteins, binding sites, and post-translational modifications.[27] The Ingenuity Pathway Analysis (IPA) software,[28] KEGG (http://www.kegg.jp), and Reactome (http://www.reactome.org) were used to identify the pathways that the hit proteins were involved in.

2.7. Molecular Dynamics (MD) Simulations of Nutlin Targets. MD simulations were carried out on a subset of the proteins found to bind Nutlin, namely gamma-glutamyl hydrolase (GGH) (PDB ID: 1L9X:A), steryl sulfatase (PDB ID: 1P49:A), and interferon-gamma (PDBID: 1FYH:A) and Human dead box RNA helicase (PDB ID: 3DKP:A). The rationale for choosing only these systems for MD simulations studies is mentioned in the Results section 3.3. Simulations (for all systems) were carried out in triplicate on the unliganded native structures of the proteins as well as on their modeled complexes with Nutlin-3a. The simulations were carried out using Gromacs[29,30] with the Amber99SB-ILDN force field.[31] Parameters for Nutlin-3a were obtained using antechamber.[32,33] Each system was solvated in a cubic water box whose sides were at a minimum distance of 25 Å from any protein atom. Charge neutrality was achieved by adding sodium or chloride counterions. The particle mesh Ewald sum method was used for treating electrostatic interactions; LINCS[34] was used to constrain the hydrogen bond lengths, enabling a time step of 2 fs. Initially, the whole system was minimized for 5000 steps or until the maximum force was <1000 kJ/mol/nm. The system was then heated to 300 K in an NVT ensemble simulation for 100 ps using a Berendsen thermostat. The system was subsequently equilibrated in an NPT ensemble simulation for 100 ps to stabilize the pressure using a Parrinello−Rahman barostat. Finally, each system was simulated for a maximum of 100 ns and structural snapshots were captured every 10 ps. Simulations were stopped when the distance between Nutlin-3a and the geometric center of the protein increased by 10 Å compared to the starting structure. The 10 Å distance was empirically chosen as indicative of the ligand irreversibly leaving the binding pocket. The temperature, potential energy, and kinetic energies were monitored during the simulations to check for anomalies.

3. RESULTS

3.1. Nutlin Binding Pocket. We have studied two variants of Nutlin: Nutlin-2 and Nutlin-3. Both variants are similarly structured with a central 5-membered imidazole ring, three of whose atoms are connected to 6-membered aromatic rings. Of these three aromatic rings, two are halogenphenyls

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(bromophenyl in Nutlin-2 and chlorophenyl in Nutlin-3). The other aromatic ring of Nutlin-2 being ethoxy methoxy phenyl whereas Nutlin-3 being methoxy-2-(propan-2-yloxy) phenyl group. Another atom of the central imidazole ring is connected to a pipirazine ring. The pipirazine group in Nutlin-2 is hydroxyl ethyl piprizine whereas in Nutlin-3 it is piprazin-2-one. Nutlin-3 exists in 2 enantiomeric forms Nutin-3a and Nutlin-3b where Nutlin-3b is a 150-fold less potent inhibitor of MDM2 than Nutlin-3a.[35]

The binding sites of Nutlin-2 and Nutlin-3a on MDM2 are almost identical and predominantly hydrophobic. In the crystal structures of Nutlin-2 and Nutlin-3a bound to MDM2 (PDB codes 1rv1 and 4j3e respectively), there are 24 and 25 residues within a distance of 6 Å from Nutlin-2 and Nutlin-3a, respectively (Figure 1B). Though the side chains are important in receptor ligand interactions, we only considered C^{α} and C^{β} atoms of the residues to constitute our binding site descriptor, to get a description of the binding pocket and an approximate orientation of the side chains. Atoms in the side chains, especially in the solvent exposed regions, are flexible and in their apo-structures may not be positioned appropriately for ligand binding (Supporting Information − section 1). In order to account for the dynamics of the protein, snapshots from MD trajectories of Nutlin bound MDM2 were also considered.[17] The number of binding site residues from the MD snapshots varies between 23 to 25 but retains their predominantly hydrophobic characteristic (Table 1) with 18 of these residues being hydrophobic (Figure 1C). While there are a few polar and charged amino acids in the pocket, their side chains are often pointed away from Nutlin. Sometimes, such as in the MD snapshot after 3 ns, the Nutlin binds deeper inside the cavity and hydrogen bonds with the side-chain of Gln72 of MDM2.

3.2. Identification of Putative Nutlin Binding Proteins. We used our CLICK program to identify proteins from among a set of 4239 human proteins that had regions structurally similar to the Nutlin binding site of MDM2. A structural overlap of 70% or above using C^{α} and C^{β} atoms and a Z-score of 2 or greater were empirically chosen thresholds (previously optimized, unpublished data) to determine meaningful matches (Table 1). For each of these hits, a model was constructed with the Nutlin in its new putative binding site. To begin with, the model is simply the coordinates of the Nutlin transferred onto the new hit after superimposing with the MDM2 binding site. This complex is energy minimized, and the resultant structure is examined for steric clashes. Models with severe clashes (as described in the Materials and Methods) are discarded. This search protocol for alternate binding partners yielded 49 human proteins (Table 1). Only 2 of the 49 hits, MDM4 and Hdm2 (52 and 96% sequence identity, respectively), are homologues of MDM2. The other putative predicted targets of Nutlin are unrelated to MDM2.

In 16 of these 49 predicted target sites, a putative binding site residue (within 6 Å of Nutlin) is involved in protein function either as an active site residue or one that undergoes post translational modification such as glycosylation (Table 1). The functions of these proteins are likely to be affected on binding Nutlin. The other 33 predicted target sites, while viable for ligand binding by our predictions, are not close to any known functional site of the protein. While the binding of Nutlin to these sites may have indirect functional consequences, we focused only on some of the hits where the functional consequences could be directly affected. The hit proteins play a role in various biosynthetic pathways including endocytosis, protein folding, metabolism, apoptosis, signaling, cell migration, immune system, transport of ions, proteolysis, etc. (Table 1).

Only one of the predicted target sites, in 5′-deoxy-5′-methylthioadenosine phosphorylase (PDBID: 1CB0:A/3OZC:A), had a bound ligand pCl-phenylthioDADMelmmA (PDBID: 3OZC:A). Interestingly, a part of this ligand, an aromatic ring, superimposes on one of the aromatic rings of Nutlin-3a (Supporting Information Figure S1).

One of the 16 proteins whose function is likely to be affected on Nutlin binding is gamma glutamyl hydrolase. Overexpression of this protein has been found to be associated with several cancers, including breast and bladder, and rheumatoid arthritis.[36] In principle, Nutlin could be repurposed to serve as a drug to combat the above diseased conditions. A more detailed analysis of the Nutlin binding to gamma glutamyl hydrolase is presented in the sections below.

3.3. Computational Measurement of Nutlin−Protein Complex Stability. We measured the strengths of association between Nutlin and the putative hit through direct and indirect computations. We directly measured the binding free energies using single point molecular mechanics computations and using the AutoDock Vina energy function. Indirectly, we assessed the strength of the complex by subjecting it to MD simulations and determined if the association was stable.

The single point binding energies of Nutlin to its original targets, MDM2, including its MD snapshots, MDM4, and Hdm2, all lie in the range of around −55 to −39 kcal/mol. The binding energy among these targets is the lowest for the 3 ns snapshot of MDM2 to Nutlin-2/Nutlin-3a (−55 and −51 kcal/mol, respectively) where the Nutlin binds deep inside the cavity and hydrogen bonds to Gln72. The binding free energy for the 49 hits for the Nutlin binding site ranges from −76 to −15 kcal/mol (Table 1). Eight (AP-2 Complex subunit beta, mitochondrial Glutaryl-CoA dehydrogenase, gamma glutamyl hydrolase, Streyl sulfatase, Stromelysin-1, Interferon-gamma, human dead box RNA helicase DDX52, peptidylprolyl isomerase domain, and WD Repeat Containing Protein 1) of the 49 proteins appear to bind Nutlin better than the original targets (marked in Table 1).

Nutlin-3a was docked onto the CLICK discovered binding pockets of the 49 putative alternate target proteins using AutoDock Vina. The RMSD to the CLICK-predicted pose and the binding energies were computed for the best bound complexes (Table 1). The best pose was the one that had the least AutoDock energy. All the AutoDock binding energies were in the range of −11.4 to −4.0 kcal/mol, indicative of favorable binding events. The binding energy of Nutlin-3a onto MDM2 was −8.4 kcal/mol with a binding pose RMSD of 0.48 Å. Seven proteins (AP-2 complex subunit beta, mitochondrial glutaryl-CoA dehydrogenase, gamma glutamyl hydrolase, streyl sulfatase, stromelysin-1, interferon-gamma, human dead box RNA helicase DDX52) had better AutoDock binding energies than MDM2. All of these also had better single point energy scores than MDM2 as described above. Interestingly, both methods compute the binding energy between Nutlin and the AP-2 Complex Subunit Beta (PDB ID: 2g30:A) as the best scoring interaction.

Overall, the single point energy scores show a similar trend as the AutoDock scores. Protein−Nutlin complexes that score well with one measure also do so with the other. The correlation between the single point scores and the AutoDock

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

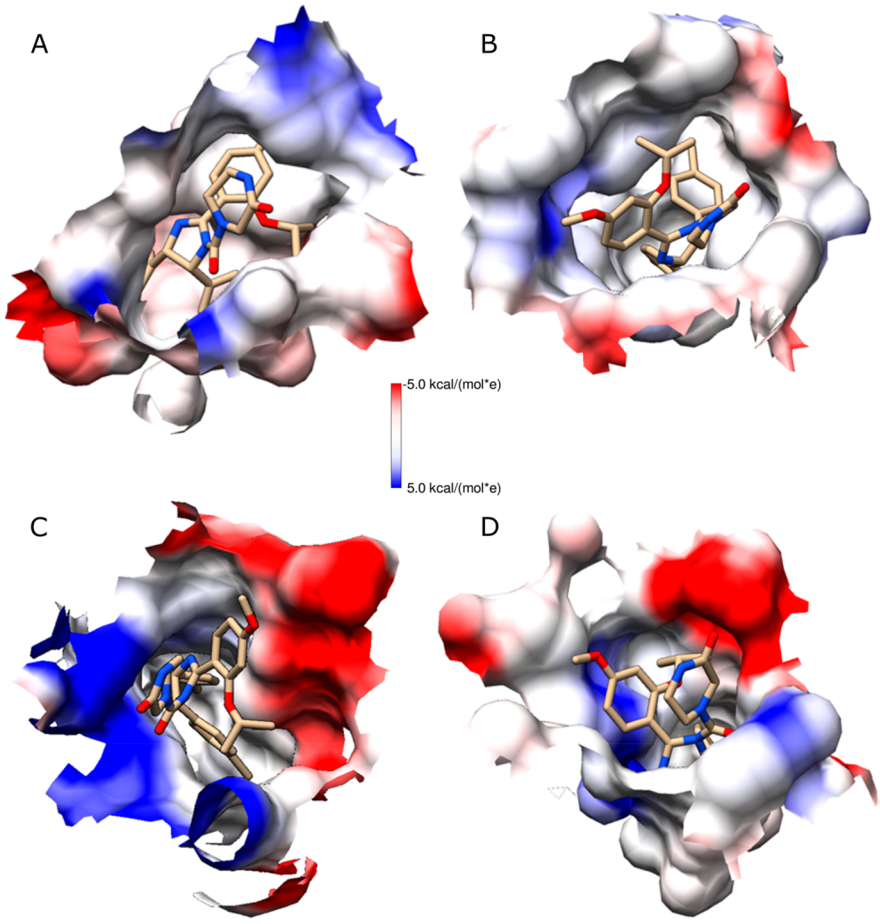


Figure 2. Surface representation of the predicted binding pocket of Nutlin (within 6 Å) in (A) gamma-glutamyl hydrolase (1L9X:A), (B) human dead box RNA helicase (3DKP:A), (C) interferon-gamma (1FYH:A) (D), steryl sulfatase (1P49:A). The binding pocket is colored as per the Chimera rendered columbic charge representation.

scores was 0.69 (Supporting Information Figure S3a). We also compared the AutoDock poses to the CLICK-predicted poses. Here again, the trends show that the larger the deviation (higher RMSD) from the CLICK pose, the less favorable is the energy (Table 1, Supporting Information Figure 3b, 3c).

Though we have computed binding energies in two different ways, these may not necessarily be indicative of favorable (or unfavorable) binding. These energy/scoring functions are inexact and do not always capture the surface chemistry accurately. In this case, the ligand is hydrophobic and we believe that the binding surface of its receptor should similarly be hydrophobic, as seen in the Nutlin-MDM2 complex crystal structure. The hydrophobicities of the eight predicted binding pockets from proteins that had better single point binding energies than MDM2 were examined manually (using Columbic coloring in Chimera). Two of the eight proteins, including gamma glutamyl hydrolase and human deadbox RNA helicase, showed predominantly hydrophobic pockets and were expected to bind stably to Nutlin. The other six proteins had polar patches in their binding pockets or had polar pocket peripheries. Either one of these characteristics were deemed as destabilizing toward binding Nutlin (Figure 2).

To test the validity of the hydrophobicity conjecture proposed above, four of these eight proteins and their Nutlin bound complexes were subjected to triplicate 100 ns MD simulations. Of the four hits, two had hydrophobic pockets (gamma glutamyl hydrolase and human deadbox RNA helicase) while the other two had some polar residues lining the pockets (inteferon gamma) and/or the pocket periphery (steryl sulphatase) (Figure 2).

We measured the stability of the Nultin-bound protein complexes by analyzing the trajectories of the distances of the center of Nutlin-3a from the center of the protein during the MD simulations (Figure 3). Nutlin-3a remained in the predicted binding site for the two proteins (Gamma glutamyl hydrolase and Human dead box RNA helicase) (in all the triplicate simulations) with a hydrophobic pocket and rim throughout the course of the simulation. The average distances between Nutlin-3a and gamma glutamyl hydrolase and human dead box RNA helicase were 13.0 Å (±0.8 Å) and 18.1 Å (±1.3 Å), respectively. The distances between the centers of the Nutlin to that of the protein in complexes with a hydrophilic binding site/periphery (Interferon Gamma and Steryl Sulfatase) increased to greater than 10 Å of the initial value (in all the triplicate simulations). At this stage the simulation was stopped. Such large deviations from the initial position are indicative of an irreversible dissociation event. In order to check if the Nutlin bound complex showed unusual fluctuations of their residues, average root-mean-square fluctuations (RMSF) of the residues with respect to the average position of the residues during the simulations were calculated. Nutin bound and Nutlin unbound complexes of Gamma glutamyl

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

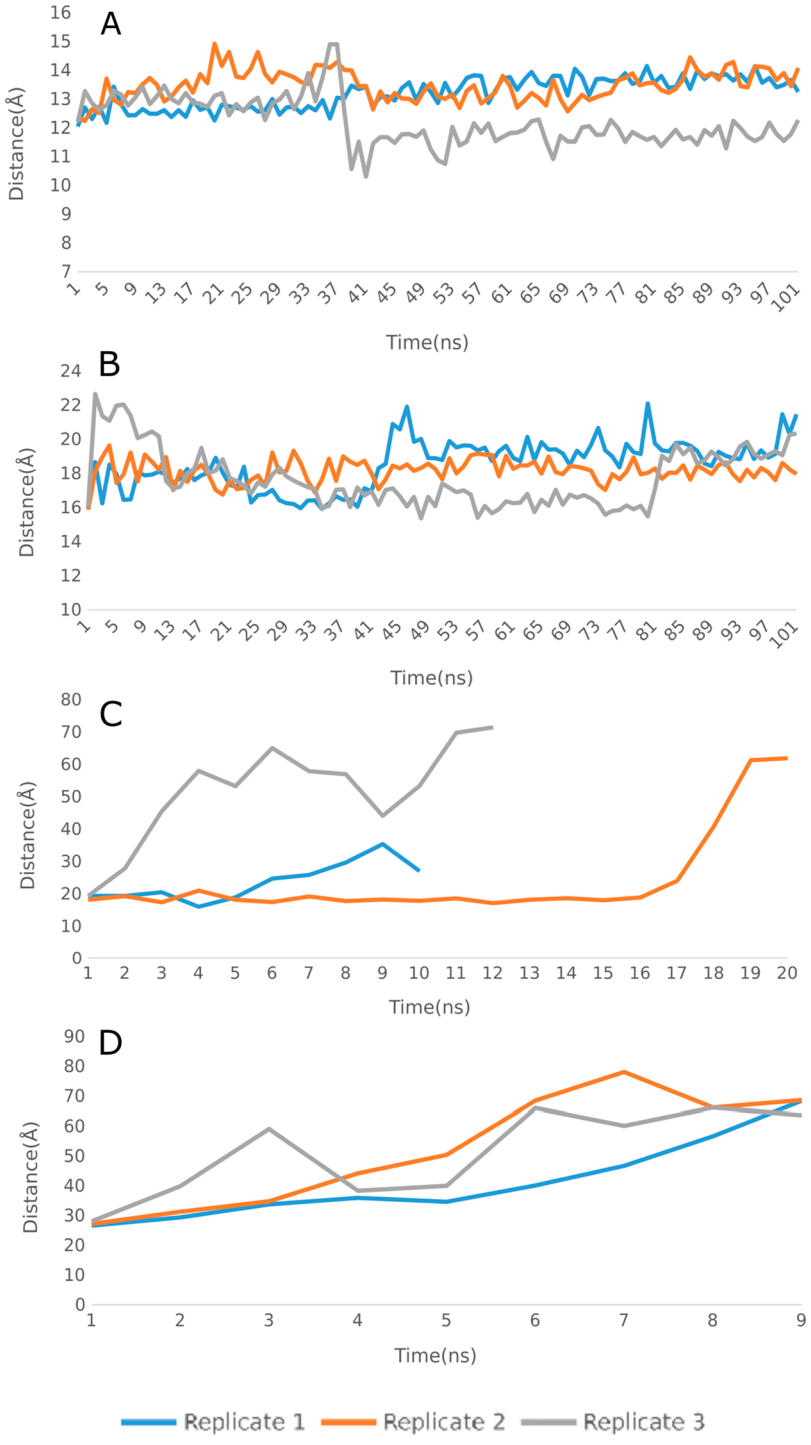


Figure 3. Distance trajectories of the center of the protein from the center of the Nutlin-3a for (A) gamma-glutamyl hydrolase (1L9X:A), (B) human dead box RNA helicase (3DKP:A), (C) interferon-gamma (1FYH:A), (D) steryl sulfatase (1P49:A). The three different trajectories from the triplicate simulations are depicted in different colors.

hydrolase and Human dead box RNA helicase show a similar RMSF, indicating the stability of the complex (Figure 4). The stability of two of the simulations with a predominantly hydrophobic rim and pocket indicate that along with matching 3D structural environment, the physicochemical properties should match for efficient binding.

3.4. Thermal Shift Assay. Our computational analysis predicts a stable association between Nutlin-3a and gamma glutamyl hydrolase. Not only does it have a favorable binding energy (by both scoring schemes), its predicted binding pocket is hydrophobic and the bound Nutlin-3a does not dissociate during the triplicate 100 ns MD runs. Gamma glutamyl hydrolase has an AutoDock score of −10.9 kcal/mol and an RMSD of 0.54 Å between the AutoDock relaxed posed and the CLICK binding pose. The Nutlin-3a binding to this protein was hence chosen for experimental validation. Gamma glutamyl hydrolase (100 ng in PBS buffer) was obtained from Abcam. It was diluted in PBS buffer containing DMSO at 0.1% final concentration

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

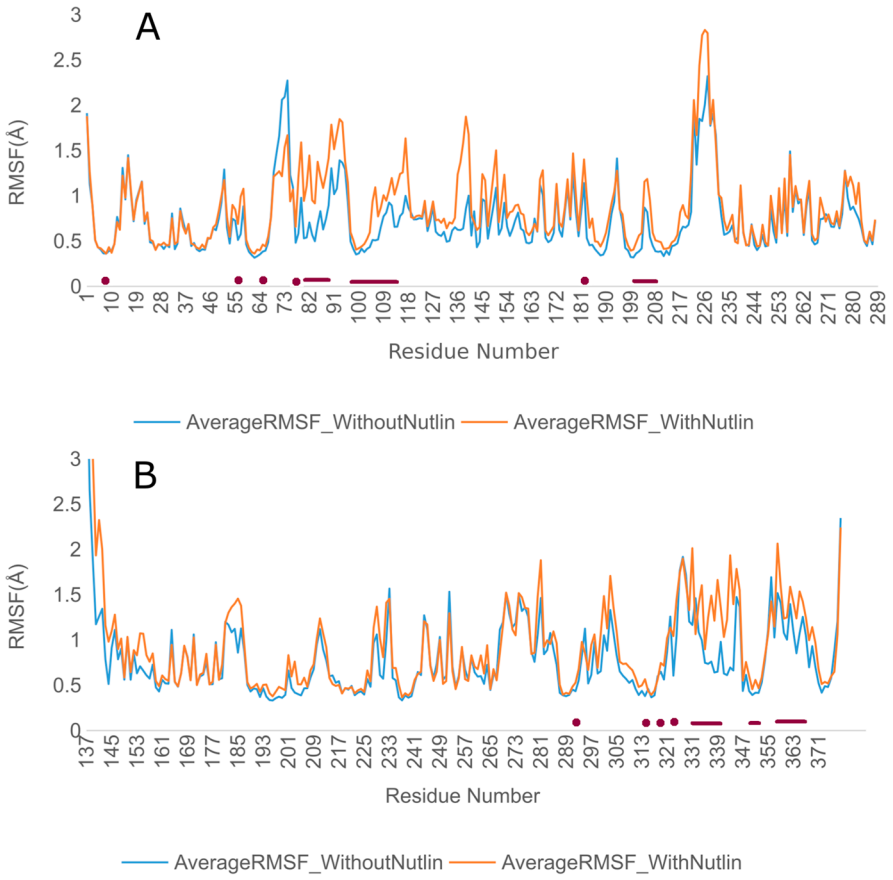


Figure 4. RMSF of individual residues with respect to the average structure generated from the MD simulation for the Nutlin-3a bound (orange) and unbound (blue) structures of (A) gamma-glutamyl hydrolase (1L9X:A) and (B) human dead box RNA helicase (3DKP:A). The regions in the protein that are within 6 Å of Nutlin-3a are depicted by purple dots or lines on the *x* axis.

(25 μL) with or without Nutlin-3a at a concentration of 1 μM, 2.5 μM, 5 μM, 10 μM, 20 μM, and 40 μM. The dose titration with the above-mentioned concentrations at 42 °C revealed that at least 40 μM of Nutlin-3a was required to protect the enzyme gamma-glytamyl hydrolase from thermal denaturation (Supporting Information Figure S4). Following an incubation of the enzyme with (+) and without (−) Nutlin-3a at a concentration of 40 μM (containing DMSO at 0.1% final concentration) at 37 °C (Figure 5, lanes 1, 2, 5, 6) or 42 °C (Figure 5, lanes 3, 4, 7, 8) for 30 min, the samples were either centrifuged (Figure 5B, lanes 5−8) or not centrifuged as controls (Figure 5A, lanes 1−4). Following this procedure, the soluble supernatant was recovered. The data (Figure 5B lanes 5 and 7) show that Nutlin-3a can promote enhanced thermal protection of gamma glutamyl hydrolase from heat aggregation/denaturation. These data suggest that, in principle, our screens are able to identify novel functional modes of binding of the small molecule Nutlin-3a.

3.5. Comparison of CLICK with Other Methods for Identifying Putative Nutlin Binding Site. In order to check the efficacy of CLICK in identifying putative binding sites for Nutlin, it was compared to other methods for investigating similarities in Nutlin binding sites. To the best of our knowledge, there are currently two methods/servers for doing this task, SMAP[37] and idTarget.[38] We tested the predictive ability of these servers for prospective Nutlin binders. The SMAP server (http://smap.nbcr.net; this site is currently unreachable) lists five best hits with significant p-values (<1.0 × 10^{-4}) (Table 2). It searches for regions in other

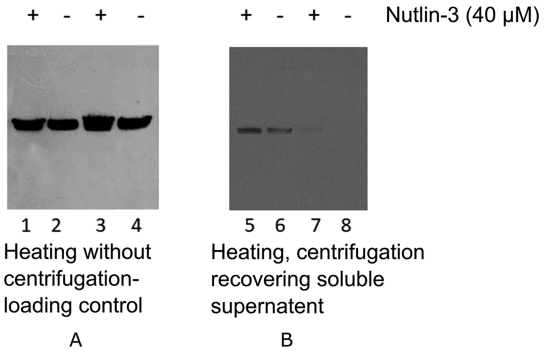


Figure 5. Data showing that incubation of gamma glutamyl hydrolase (100 ng in 25 μL PBS containing DMSO at 0.1% final concentration) with Nutlin-3a (40 $\mu\mu$M containing DMSO at 0.1% final concentration) can stimulate the protection of the protein from heat induced aggregation/denaturation, which is suggestive of binding as modeled. The right four lanes show the effects on Gamma glutamyl hydrolase stability as a function of temperature and Nutlin-3a. (+) indicates presence of Nutlin-3a whereas (−) indicates its absence. (A) The left four lanes are loading controls that were processed without centrifugation and (B) after centrifugation in the soluble fraction. Samples in lanes 1, 2, 5, and 6 are incubated at 37 °C while samples in lanes 3, 4, 7, and 8 are incubated at 42 °C.

proteins that are structurally similar to the Nutlin binding site on MDM2. SMAP is used for the comparison and the similarity search of protein 3D motifs independent of sequence order and has been applied for predicting drug side effects and to repurpose existing drugs for new indications. All the

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Table 2. Results from SMAP Server to Find Nutlin Binding Pockets in Human Protein Structures

PDB ID	Number of steric clashes
1z1m:A	62
2qag:C	4
3dzu:A	20
2qag:B	10
1lv2:A	43

5 best hits of SMAP have 4 to 62 clashes with Nutlin (Table 2), and their MM/GBSA binding energies could not be calculated.

The second method, idTarget, predicts possible binding targets of a small chemical molecule via a divide-and-conquer docking approach, using the scoring scheme from AutoDock4. idTarget has been shown to be able to reproduce known off-targets of drugs or drug-like compounds. The idTarget server (http://idtarget.rcas.sinica.edu.tw) detected 34 hits with Z score < −0.5 and using the same clash tolerance as described in section 2.4. We have made a simplistic assumption in this study that binding to Nutlin should also mean similarity in binding site. In all the putative targets identified by CLICK, the predicted binding sites overlap with the MDM2 site by more than 70%. Only three of the idTarget hits have an overlap of 70.21% with the MDM2 site. The overlaps of the predicted hits by idTarget range between 23−70% (Table 3). Though, idTarget used AutoDock scoring schemes to score the predictions and are all shown to have favorable interactions energy with Nutlin, the binding free energies calculated using MM/GBSA of Amber 11 show positive values. 2r4v (Chloride intracellular channel protein 2) is identified as a target both by CLICK and idTarget, but they identify different regions on the protein as binding sites. The 7 alternative targets identified by CLICK that have lower single point energy and AutoDock binding affinity scores than MDM2 were predicted neither by SMAP nor by idTarget.

4. DISCUSSION

Broadly speaking, for the productive binding/interaction of biomolecules, there needs to be complementarity in geometry and chemistry. In this study, we have showcased the utility of our CLICK software in detecting protein substructures (binding sites) with similar geometry. We had previously shown that CLICK could detect ATP binding sites by structural similarity.[6] Here, we have used CLICK to detect putative binding sites on proteins that are structurally similar to the Nutlin binding pocket on MDM2. This was effected by mining a nonredundant structural database of human proteins for regions of proteins that are structurally similar to the Nutlin binding sites obtained from the MDM2−Nutlin complex crystal structure and snapshots from its MD simulations. We found 49 human proteins that have regions that are structurally similar to the Nutlin binding site on MDM2. To ensure that the geometric similarities were significant, we ensured that at least 70% of the residues in the putative hits overlapped. Additionally, when placing the Nutlin in these predicted pockets there were less than 1 main chain and 5 side chain clashes. We believe that such stringent criteria would exclude some of the known Nutlin binders, such as Bcl-X_L,[39] but help in minimizing false positives. In future applications, we are exploring the use of sub optimal matches, which would have predicted Bcl-X_L as one of the putative binders.

Table 3. Results from idTarget Server to Find Nutlin Binding Pockets in Human Protein Structures

PDB ID	Number of steric clashes	Binding energy (kcal/mol)	Number of binding site residues in the predicted binding site within 6 Å	Structure Overlap with respect to the number of residues in Nutlin-2 binding site
3kjd	2	965.05	32	61.70
3l0l	0	590.71	38	70.21
3i28	1	6927.15	34	68.09
2zb4	1	395.69	39	68.09
3inm	0	18.91	23	51.06
3g2f	1	127.59	32	70.21
2x7g	0	64.62	32	65.96
1wb0	0	5.54	7	23.40
2vuw	1	95.31	35	63.83
2h7c	1	596.55	20	48.93
3e7e	0	127.37	29	61.70
2ipx	2	88.84	28	53.19
2wax	0	250.74	13	42.55
3bgv	0	1003.23	28	51.06
3c0i	0	472.41	35	70.21
1z70	0	1788.47	20	44.68
2jc9	1	66.97	25	61.70
2wef	1	167.99	34	65.96
2r4v	0	99.25	19	59.57
3bpt	0	257.11	28	55.32
3epy	0	126.83	15	51.06
2pez	0	59.56	26	51.06
3ebb	1	3945.18	18	48.94
2ql9	1	232.35	7	29.79
1s1d	0	160.35	23	59.57
2vfk	1	32.01	21	48.94
2wm1	0	448.78	15	42.55
3iai	0	46.17	30	68.09
1wl4	0	524.28	22	51.06
3ijj	2	602.19	15	42.55
1d4a	1	84.94	20	53.19
3e9k	1	155.08	21	57.45
2zg1	2	203.13	16	46.81
1elv	1	153.43	16	48.94

Having satisfactorily obtained similarly structured pockets, we next evaluated the chemistry of interaction or simply the binding energies of the predicted Nutlin−protein complex. These computations were done using an MM/GBSA scoring scheme as well as the AutoDock energy scores. These putative alternate targets of Nutlin also had favorable energies of binding as computed/predicted by single point energy calculations and by the AutoDock energy function. The AutoDock computed energies for 7 Nutlin−protein complexes had lower values than the native Nutlin−MDM2 complex. Consistently, all 7 of these complexes also had better MM/GBSA scores than Nutlin−MDM2, in addition to the peptidylprolyl isomerase domain and WD Repeat Containing Protein 1. The rank ordering of the two scoring scheme of the putative alternate targets was also similar (correlation coefficient of 0.69). On comparing the RMSD of the two poses and the single point energies, we again found good correlation (coefficient 0.57). Interestingly, the smaller the RMSD between the two predicted poses the more favourable the binding energy (by both methods).

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Despite the good agreement between our single point energy scores and AutoDock evaluations, we are aware that these molecular mechanics and empirical scoring schemes are often not very accurate. We evaluated the physicochemical nature of the ligand and its receptor site. Nutlin is predominantly hydrophobic, and in its complex with MDM2, it is bound in a hydrophobic site. We chose two predicted binders each with hydrophobic and nonhydrophobic pockets and subjected them to triplicate MD simulations. The Nutlin bound to Interferon gamma and Steryl sulfatase did not remain bound to the predicted nonhydrophobic pocket. Whereas, Nutlin remained bound to the hydrophobic pockets of gamma glutamyl hydrolase and human dead box RNA helicase in triplicate 100 ns MD trajectories.

In order to assess whether the binding of the ligand/drug would alter/affect the functioning of the putative hits, we attempted to correlate functional information to the amino acids that constitute the binding site. Function is associated with all amino acids in the active site of enzymes and at sites of post translational modifications. Simplistically, we have assumed that any binding in the close proximity to these functional sites would impair protein activity. We found 16 predicted targets whose functions are likely to be affected on Nutlin binding. It is possible that (some of the) other predicted target sites may affect protein function through allostery.

Given our somewhat modest resources and the inhibitive cost of Nutlin, we experimentally validated the binding of Nutlin-3a to Gamma glutamyl hydrolase. This enzyme has a hydrophobic binding site and does not dissociate with Nutlin in MD simulations, and we predict it would have its function (glycosylation) affected on binding Nutlin. We showed that Nutlin-3a can protect gamma glutamyl hydrolase from thermal denaturation.

An important implication of our study is that this procedure can be used not just to discover alternate binding sites for known ligands/inhibitors/drugs, but it could serve as a platform to repurpose known drugs. For instance, the levels of gamma glutamyl hydrolase have been implicated in several disease conditions including several cancers and arthritis, and perhaps the binding of Nutlin could influence favorable therapeutic outcomes. In this study, in conjunction with the energy scores computed by two different methods, we felt it was necessary to manually look into ligand−receptor specific properties, in this case hydrophobicity. For a larger more general application of this method, an automated classification of the ligand and receptor/binding sites would be required. While we are working toward that end, it is beyond the scope of this study.

Another positive aspect of using the CLICK software was that we identified hits that are not readily identified by other docking procedures. For instance, when we relaxed/scored the ligand−protein complexes with AutoDock Vina, we had to make use of our complex structures as a starting point and could not begin with the crystal structures. This is because our predicted models, after energy minimization, had opened to slot in the ligand while the binding sites on the crystal structures were seldom in a conformation conducive to ligand binding.

We also compared the performance of our method to two other methods, SMAP and idTarget. Except for one target, none of the proteins identified by our method was predicted by the other methods. All the targets identified by the SMAP server had several clashes with the Nutlin, and hence the single point energies could not be calculated. The MM/GBSA energies for the hits identified by idTarget were consistently unfavorable. Even if this assessment may not be completely accurate, we also noticed that the structural overlap between the idTarget hits to the MDM2 binding site were seldom, if at all, as high as our predictions. We used a threshold of 70% similarity to filter our CLICK identified hits. idTarget predictions had structural overlaps in the range of 23−70%.

In conclusion, the program CLICK has been used to identify the possible proteins that Nutlin can bind to. The participation of these proteins in different biological pathways hints at likely off-target effects such as toxicity. Experimental techniques such as CETSA[40] have the potential to identify the drug target but in general are time-consuming and/or expensive. In contrast, CLICK can exhaustively and quickly search large sets of protein structures to identify the best target candidates and can hence reduce the experimental tests to a limited number of proteins, resulting in reduction in time and cost efficiency. Hence CLICK can be used as an initial screening tool for cost-effective toxicology studies of drugs. In this paper we present a list of proteins that could potentially bind to Nutlin, which can be used to validate their binding and off target effects. The best hits presented in this paper are only a partial list of targets for Nutlin binding, as not all proteins have known 3D structures. A limitation of our method is that it is dependent on the availability of experimentally determined 3D structures of proteins. We believe that a larger study could be envisioned utilizing homologous structures or models. However, that is also beyond the scope of this current study.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.8b00762.

Superimposition of Nutlin binding pocket and 4CT ligand in PDB ID: 3OZC:A (Figure S1); docking of Nutlin-3a on gamma glutamyl hydrolase (Section 1, Figure S2); correlations between binding free energies calculated by MM/GBSA and AutoDock Vina (Figure S3a); correlations between RMSD between the predicted binding poses and binding free energy calculated by MM/GBSA and AutoDock Vina (Figure S3b, S3c); dose titration of Nutlin-3a (Figure S4); PDB IDs of human proteins used in this study (Section 2) (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: madhusudhan@iiserpune.ac.in.
*E-mail: ted.hupp@ed.ac.uk.
*E-mail: chandra@bii.a-star.edu.sg.

ORCID

Minh N. Nguyen: 0000-0003-4949-2571
Chandra S. Verma: 0000-0003-0733-9798
M. S. Madhusudhan: 0000-0002-2889-5884

Present Addresses

○(V.T.) Biological and Life Sciences, School of Arts and Sciences, Ahmedabad University, Central Campus, Navrangpura, Ahmedabad − 380009, Gujarat, India.
▽(C.V.) Singapore Institute for Clinical Sciences (SICS), Brenner Centre for Molecular Medicine, 30 Medical Drive, Singapore 117609.

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Author Contributions

[$]M.N.N., N.S., and L.M. contributed equally to the work.

Funding

Funding was provided by Biomedical Research Council (A*STAR), Singapore. M.S.M. would also like to acknowledge Wellcome Trust-DBT India alliance for a senior fellowship. N.S. would like to acknowledge a CSIR-SPMF fellowship for funding. M.N.N. would like to acknowledge the Biomedical Research Council (BMRC)-Economic Development Board (EDB) Industry Alignment Fund (IAF311017G and H18/01/a0/B14), A*STAR, Singapore for funding.

Notes

The authors declare no competing financial interest.

C.S.V. is the founder/scientific consultant of Sinopsee Therapeutics, a biotechnology company developing molecules for therapeutic purposes; the current work has no conflict with the company.

■ ACKNOWLEDGMENTS

The authors would like to thank Sanjana Nair for discussions.

■ REFERENCES

(1) Perry, M. E. The Regulation of the P53-Mediated Stress Response by MDM2 and MDM4. *Cold Spring Harbor Perspect. Biol.* **2010**, *2*, a000968.

(2) International Drug Monitoring: The Role of National Centres. Report of a WHO Meeting. *World Heal. Organ. - Tech. Rep. Ser.* **1972**.

(3) Campillos, M.; Kuhn, M.; Gavin, A.-C.; Jensen, L. J.; Bork, P. Drug Target Identification Using Side-Effect Similarity. *Science* **2008**, *321*, 263−266.

(4) Cohen, P.; Tcherpakov, M. Will the Ubiquitin System Furnish as Many Drug Targets as Protein Kinases? *Cell* **2010**, *143*, 686−693.

(5) Uitdehaag, J. C. M.; Verkaar, F.; Alwan, H.; De Man, J.; Buijsman, R. C.; Zaman, G. J. R. A Guide to Picking the Most Selective Kinase Inhibitor Tool Compounds for Pharmacological Validation of Drug Targets. *Br. J. Pharmacol.* **2012**, *166*, 858−876.

(6) Nguyen, M. N.; Tan, K. P.; Madhusudhan, M. S. CLICK–Topology-Independent Comparison of Biomolecular 3D Structures. *Nucleic Acids Res.* **2011**, *39* (Web Serverissue), W24−8.

(7) Konagurthu, A. S.; Whisstock, J. C.; Stuckey, P. J.; Lesk, A. M. MUSTANG: A Multiple Structural Alignment Algorithm. *Proteins: Struct., Funct., Genet.* **2006**, *64*, 559

(8) Holm, L.; Sander, C. Protein Structure Comparison by Alignment of Distance Matrices. *J. Mol. Biol.* **1993**, *233*, 123−138.

(9) Nussinov, R.; Wolfson, H. J. Efficient Detection of Three-Dimensional Structural Motifs in Biological Macromolecules by Computer Vision Techniques. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 10495−10499.

(10) Braberg, H.; Webb, B. M.; Tjioe, E.; Pieper, U.; Sali, A.; Madhusudhan, M. S. Salign: A Web Server for Alignment of Multiple Protein Sequences and Structures. *Bioinformatics* **2012**, *28* (15), 2072−2073.

(11) Nguyen, M. N.; Madhusudhan, M. S. Biological Insights from Topology Independent Comparison of Protein 3D Structures. *Nucleic Acids Res.* **2011**, *39* (14), e94.

(12) Konc, J.; Janežič, D. ProBiS Algorithm for Detection of Structurally Similar Protein Binding Sites by Local Structural Alignment. *Bioinformatics* **2010**, *26*, 1160−1168.

(13) Štular, T.; Lešnik, S.; Rožman, K.; Schink, J.; Zdouc, M.; Ghysels, A.; Liu, F.; Aldrich, C. C.; Haupt, V. J.; Salentin, S.; Simone, Daminelli; Schroeder, M.; Langer, T.; Gobec, S.; Janežič, D.; Konc, J. Discovery of Mycobacterium Tuberculosis InhA Inhibitors by Binding Sites Comparison and Ligands Prediction. *J. Med. Chem.* **2016**, *59*, 11069−11078.

(14) Momand, J.; Zambetti, G. P.; Olson, D. C.; George, D.; Levine, A. J. The MDM-2 Oncogene Product Forms a Complex with the P53 Protein and Inhibits P53-Mediated Transactivation. *Cell* **1992**, *69*, 1237−1245.

(15) Brown, C. J.; Cheok, C. F.; Verma, C. S.; Lane, D. P. Reactivation of P53: From Peptides to Small Molecules. *Trends Pharmacol. Sci.* **2011**, *32*, 53−62.

(16) Khoo, K. H.; Verma, C. S.; Lane, D. P. Drugging the P53 Pathway: Understanding the Route to Clinical Efficacy. *Nat. Rev. Drug Discovery* **2014**, *13*, 217−236.

(17) Joseph, T. L.; Madhumalar, A.; Brown, C. J.; Lane, D. P.; Verma, C. Differential Binding of P53 and Nutlin to MDM2 and MDMX: Computational Studies. *Cell Cycle* **2010**, *9*, 1167−1181.

(18) Vu, B.; Wovkulich, P.; Pizzolato, G.; Lovey, A.; Ding, Q.; Jiang, N.; Liu, J. J.; Zhao, C.; Glenn, K.; Wen, Y.; Tovar, C.; Packman, K.; Vassilev, L.; Graves, B. Discovery of RG7112: A Small-Molecule MDM2 Inhibitor in Clinical Development. *ACS Med. Chem. Lett.* **2013**, *4*, 466−469.

(19) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. Structure of the MDM2 Oncoprotein Bound to the P53 Tumor Suppressor Transactivation Domain. *Science (Washington, DC, U. S.)* **1996**, *274*, 948−953.

(20) Berberich, S. J. RNAi Knockdown of HdmX or Hdm2 Leads to New Insights into P53 Signaling. *Cell Cycle* **2010**, *9*, 3640−3641.

(21) Secchiero, P.; Bosco, R.; Celeghini, C.; Zauli, G. Recent Advances in the Therapeutic Perspectives of Nutlin-3. *Curr. Pharm. Des.* **2011**, *17*, 569−577.

(22) Burgess, A.; Chia, K. M.; Haupt, S.; Thomas, D.; Haupt, Y.; Lim, E. Clinical Overview of MDM2/X-Targeted Therapies. *Front. Oncol.* **2016**, *6* (January), 1−7.

(23) Wang, G.; Dunbrack, R. L. PISCES: A Protein Sequence Culling Server. *Bioinformatics* **2003**, *19* (12), 1589−1591.

(24) Chakravarty, S.; Varadarajan, R. Residue Depth: A Novel Parameter for the Analysis of Protein Structure and Stability. *Structure* **1999**, *7*, 723−732.

(25) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 11*; 2010.

(26) Trott, O.; Schroer, A. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2009**, *31*, 455−461.

(27) Bairoch, A.; Apweiler, R.; Wu, C. H.; Barker, W. C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M.; Martin, M. J.; Natale, D. A.; O'Donovn, C.; Redaschi, N.; Yeh, L. L. The Universal Protein Resource (UniProt). *Nucleic Acids Res.* **2005**, *33*, 154−159.

(28) Jiménez-Marín, Á.; Collado-Romero, M.; Ramirez-Boo, M.; Arce, C.; Garrido, J. J. Biological Pathway Analysis by ArrayUnlock and Ingenuity Pathway Analysis. *BMC Proc.* **2009**, *3*, S6.

(29) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.

(30) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845−854.

(31) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber Ff99SB Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.

(32) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(33) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.
(34) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(35) ElSawy, K. M.; Verma, C. S.; Lane, D. P.; Caves, L. S. D. On the Origin of the Stereoselective Affinity of Nutlin-3 Geometrical Isomers for the MDM2 Protein. *Cell Cycle* **2013**, *12*, 3727−3735.
(36) Świerkot, J.; Ślęzak, R.; Paweł Karpiński, J. P.; Leszek Noga, J. S.; Wiland, P. Associations between Single-Nucleotide Polymorphisms of RFC-1, GGH, MTHFR, TYMS, and TCII Genes and the Efficacy and Toxicity of Methotrexate Treatment in Patients with Rheumatoid Arthritis. *Polym. Arch. Med. Wewn.* **2015**, *38*, 8−15.
(37) Ren, J.; Xie, L.; Li, W. W.; Bourne, P. E. SMAP-WS: A Parallel Web Service for Structural Proteome-Wide Ligand-Binding Site Comparison. *Nucleic Acids Res.* **2010**, *38*, W441−444.
(38) Wang, J. C.; Chu, P. Y.; Chen, C. M.; Lin, J. H. IdTarget: A Web Server for Identifying Protein Targets of Small Chemical Molecules with Robust Scoring Functions and a Divide-and-Conquer Docking Approach. *Nucleic Acids Res.* **2012**, *40*, W393−399.
(39) Shin, J. S.; Ha, J. H.; He, F.; Muto, Y.; Ryu, K. S.; Yoon, H. S.; Kang, S.; Park, S. G.; Park, B. C.; Choi, S. U.; et al. Structural Insights into the Dual-Targeting Mechanism of Nutlin-3. *Biochem. Biophys. Res. Commun.* **2012**, *420*, 48−53.
(40) Jafari, R.; Almqvist, H.; Axelsson, H.; Ignatushchenko, M.; Lundbäck, T.; Nordlund, P.; Molina, D. M. The Cellular Thermal Shift Assay for Evaluating Drug Target Interactions in Cells. *Nat. Protoc.* **2014**, *9*, 2100−2122.
(41) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.

DOI: 10.1021/acs.jcim.8b00762
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX