

Foundation Review revised and submitted to *Drug Discovery Today*, January 8th 2020.

Impact of Protein Data Bank on Anti-neoplastic Approvals

John D. Westbrook¹, Rose Soskind²⁺, Brian P. Hudson¹,
and Stephen K. Burley^{1,3,4,5*}

¹ Research Collaboratory for Structural Bioinformatics Protein Data Bank, Institute for Quantitative Biomedicine, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

² Ernest Mario School of Pharmacy, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

³ Rutgers Cancer Institute of New Jersey, Robert Wood Johnson Medical School, New Brunswick, NJ 08903, USA

⁴ Research Collaboratory for Structural Bioinformatics Protein Data Bank, San Diego Supercomputer Center, University of California, San Diego, La Jolla, CA 92093, USA

⁵ Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92093, USA

+ Present Address: NYU Langone Health, 550 1st Avenue, New York, NY 10016, USA

* *Corresponding Author:* Burley, S.K. (Stephen.Burley@RCSB.org)

Corresponding Author Mobile Telephone: +1-858-361-6961

Other Author E-mail addresses:

JDW: John.Westbrook@rcsb.org

RS: Rose.Soskind@nyulangone.org

BPH: Brian.Hudson@rcsb.org

Other Elements

Key Words:

Protein Data Bank, Structure-guided Drug Discovery, US FDA Oncology Approvals, Small-molecule Drug, Biologic Drug, Open Access

Teaser:

Open access to three-dimensional macromolecular structure information managed by the Protein Data Bank facilitated discovery/development of more than 90% of new anti-neoplastic agents approved by the US FDA 2010-2018.

Highlights

Protein Data Bank (PDB) provides open access to >160K 3D structures of biomolecules.

79 anticancer NMEs with known molecular targets were approved by US FDA 2010-2018.

PDB provides target structures for >90% of anticancer NMEs approved 2010-2018.

PDB provides target-NME co-structures for >50% of anticancer NMEs approved 2010-2018.

PDB facilitated discovery/development of >90% of anticancer NMEs approved 2010-2018.

Three Author Biographies:

John D. Westbrook, Ph.D. is a computational chemist and the lead data and software architect for the RCSB Protein Data Bank at Rutgers, The State University of New Jersey. He serves on data standards committees for the International Union of Crystallography, the American Crystallographic Association, and Research Data Alliance. Awards and prizes include Biocuration Career Award from the International Biocuration Society, Rutgers University Supercomputer Fellowship, Rutgers University Johnson Fellowship, Raymond Davis Memorial Fellowship from the Society of Photographic Science and Engineering, and Minolta Corporation Fellowship in Imaging Science.

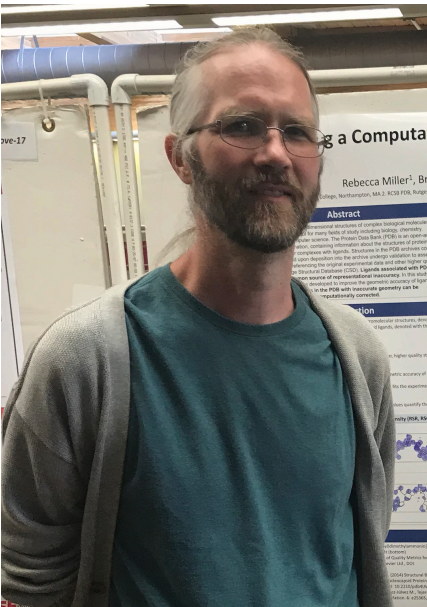
Brian P. Hudson, Ph.D. is an expert in structural biology and biological database curation. He currently serves as a Senior Biocurator at the RCSB Protein Data Bank at Rutgers, the State University of New Jersey.

Stephen K. Burley, M.D., D.Phil. is an expert in the areas of molecular biophysics, structural biology, bioinformatics, data science, structure/fragment-based drug discovery, and clinical medicine/oncology. Burley currently serves as University Professor and Henry Rutgers Chair, Founding Director of the Institute for Quantitative Biomedicine, and Director of the RCSB Protein Data Bank at Rutgers, The State University of New Jersey. He is a Member of the Rutgers Cancer Institute of New Jersey, where he Co-Leads Cancer Pharmacology. Previous work experience includes senior leadership roles at Lilly Research Laboratories, SGX Pharmaceuticals, Inc., The Rockefeller University, and the Howard Hughes Medical Institute.

Three Author Photographs:



John D. Westbrook



Brian P. Hudson



Stephen K. Burley

Open access to three-dimensional (3D) structure information from the Protein Data Bank (PDB) facilitated discovery/development of more than 90% of the 79 new anti-neoplastic agents (54 small-molecules, 25 biologics) with known molecular targets approved by the US FDA 2010-2018. Analyses of PDB holdings, the scientific literature, and related documents for each drug-target combination revealed that the impact of public-domain 3D structure data was broad and substantial, ranging from understanding target biology (~95% of all targets), to identifying a given target as likely druggable (~95% of all targets), to structure-guided lead optimization (>70% of all small-molecule drugs). In addition to aggregate impact assessments, illustrative case studies are presented for three protein kinase inhibitors, an allosteric enzyme inhibitor, and seven advanced-stage melanoma therapeutics.

Introduction

Over the past two decades, protein crystallography and structure-guided drug discovery have become established tools used throughout the biopharmaceutical industry [1,2]. 3D structures of biological macromolecules can inform our understanding of target biology (reviewed in [3]). They can confirm that a given protein target is likely to be druggable with small-molecule and/or biologic agents (reviewed in [4]). And, in the most favorable cases, protein crystallography can enable structure-guided optimization of affinity of small-molecule leads [1]. 3D structural data have also proven useful in overcoming some of the other challenges (*e.g.*, avoiding unwanted off-target binding) inherent in turning biochemically active compounds into potent drug-like molecules suitable for safety and efficacy testing in animals and humans [5]. In the realm of biologics (~20% of approved drugs in the current era [6]), 3D structural information is also used routinely to inform engineering of monoclonal antibodies and other protein-based therapeutics [7,8].

Public-domain 3D structure information is distributed on an open-access basis by a single, global data resource known as the Protein Data Bank (PDB [9]). Since 2008, publication of new macromolecular structures in most scientific journals has been contingent on mandatory deposition to the PDB of the 3D atomic coordinates constituting the structure together with experimental data and related metadata. Many governmental/non-governmental research funders also require PDB deposition of macromolecular structure data by their grantees. When the PDB was established in 1971 as the first open-access digital data resource in biology, it housed only seven protein structures [9].

Today, the PDB is regarded as a global public good vital to research and education/training across the biological and biomedical sciences. At the time of this publication, the PDB housed >160,000 experimentally determined, atomic-level 3D structures of biological macromolecules (*i.e.*, proteins, DNA, and RNA), many of which have been visualized in the act of binding one or more small-molecule ligands including United States Food and Drug Administration (US FDA) approved drugs. Since 2003, the PDB has been managed jointly according to the *FAIR* Principles of *Findability-Accessibility-Interoperability-Reusability* [10] by the Worldwide Protein Data Bank (wwPDB) partnership [11,12] (including the US Research Collaboratory for Structural Bioinformatics Protein Data Bank or RCSB PDB [13,14], Protein Data Bank in Europe [15], Protein Data Bank Japan [16], and BioMagResBank [17]).

The RCSB PDB (rcsb.org) recently published a quantitative overview of the impact of PDB structures on 210 New Medical Entities (NMEs or new drugs) approved by US FDA 2010-2016 [18]. This work built on previously published analyses of lessons learned from 20

years of anti-cancer drugs [19] and went beyond individual case studies (*e.g.*, [20]) and presentations at scientific meetings that described the impact of structure-guided drug discovery and of protein crystallographers working in industry. In all, we documented that nearly 6,000 atomic-level 3D structures of molecular targets stored in the PDB archive facilitated discovery and development of ~88% of the 210 new drugs approved 2010-2016 across all therapeutic areas.

Given the large number of recent US FDA drug approvals for oncology indications, we now review the ways that open access to PDB data facilitated discovery and development of 79 anti-neoplastic agents with known molecular targets (54 small molecules, 25 biologics) approved 2010-2018. In addition to an aggregate review of PDB impact on new drug approvals, we review three case studies illustrating the impact of PDB data, including three hinge-binding CDK4/CDK6 inhibitors (palbociclib, ribociclib, and abemaciclib), an isocitrate dehydrogenase 2 (IDH2) allosteric inhibitor (enasidenib), and seven therapeutic agents that have transformed clinical management of advanced-stage melanoma (Protein Kinase Inhibitors: vemurafenib, dabrafenib, trametinib, and encorafenib; Antibodies: ipilimumab, pertuzumab, and nivolumab).

Anti-neoplastic Drugs Approved by US FDA 2010-2018

A total of 81 anti-neoplastic NMEs were approved by US FDA 2010-2018. Two of these newly approved drugs, trabectedin and ingenol mebutate, were not considered in this review because their molecular targets are unknown. All but three of the 79 remaining anti-neoplastic agents target human proteins. Nearly three-quarters (54/79, ~68%) of the NMEs are of low molecular weight (<1000 Da, denoted LMW-NMEs), all targeting human proteins. We classified the remaining NMEs (25/79, ~32%; ≥1000Da) as Biologic-NMEs. Three of these Biologic-NMEs have small-molecule targets: dinutuximab, which is a monoclonal antibody that recognizes glycolipid GD2; and two L-asparaginases that hydrolyze L-asparagine yielding L-aspartate and ammonia. The remaining 22 Biologic-NMEs (22/25, 88%) target extracellular human proteins.

Impact of PDB Structures on Anti-neoplastic Drug Approvals

We searched the entire PDB archive using corresponding reference amino acid sequences from UniProt (uniprot.org [21]) to identify 3D structures that include all or part of the known macromolecular target for each of the 79 anti-neoplastic NMEs (Table 1). As of September 2019, the archive contained one or more structures for 74 of the 79 NME targets (~95%). Every LMW-NME has at least one target structure present in the PDB (54/54, 100%). The LMW-NMEs themselves are also well represented in the PDB. For more than three-quarters of the LMW-NMEs (41/54, ~76%), one or more public-domain

PDB structures reveal at the atomic level precisely how the drug binds to the corresponding NME target protein, and, in some cases, to other so-called off-target proteins. Eighty percent of the Biologic-NMEs (20/25) have one or more target structures in the PDB. For approximately one-half of the Biologic-NMEs (13/25, 52%), the PDB archive houses one or more 3D structures of the drug itself and/or the drug-target complex. For both the LMW-NMEs and the Biologic NMEs, >95% of the target structures were deposited to the PDB at least a decade before the drug was approved for clinical use by US FDA.

Small-molecule NMEs. The 54 LMW-NMEs target 13 distinct classes of proteins (Table 2). Without exception, known protein targets of every one of the 54 LMW-NMEs were represented in the PDB archive. In all, we identified 2,115 “Relevant Structures,” which include PDB structures containing the following: (a) a reference or a mutant/variant form of the target protein; (b) a LMW-NME bound to a reference or mutant/variant form of its target protein; (c) a LMW-NME bound to a potential alternative target protein; or (d) a LMW-NME bound to a possible off-target protein. The number of Relevant Structures identified for each target or target class ranges from <10 for IDH2 to 1,136 for the protein kinases.

More than 98% of the 2,115 Relevant Structures were deposited to the PDB well before the LMW-NME was approved for clinical use by US FDA. The median time between PDB deposition and approval exceeded 14 years (data not shown). The vast majority of the 2,115 Relevant Structures (1,807/2,115; ~85%) were reported in a PubMed-indexed publication around the time of PDB deposition. These 1,807 papers had garnered a total of 172,653 literature citations as of September 2019, giving an average of >95 citations/primary publication. For reference, the average number of literature citations/primary publication across the entire PDB archive is ~50 [22].

Review of PDB archival holdings and the scientific literature pertaining to each NME target/LMW-NME combination summarized in Tables 2 and 3 revealed that public domain 3D structure data facilitated discovery and development of all 54 LMW-NMEs in the following ways:

(i) Target Biology: Atomic-level 3D structures provide functional insights that are not always readily apparent from amino acid sequence (reviewed in [22]).

In every case, the PDB provides one or more experimentally determined structure of each unique NME target.

(ii) Target Druggability: 3D structures enable visualization of surface features deemed likely to bind small organic compounds for inhibition of enzymatic action or other biochemical or biological function (reviewed in [4]).

In every case, a PDB structure(s) revealed one or more potential small-molecule binding sites, either on the surface of the NME target or within a protein-protein interface (e.g., the homodimeric enzymes IDH1 and IDH2).

(iii) Small-molecule Binding: Co-crystal structure studies permit 3D assessment of binding of tool compounds or small-molecule hits coming from biochemical, biophysical [23], or fragment [24] screening campaigns, thereby aiding medicinal chemistry decision making (e.g., [25]).

For more than 41 of 54 of cases (>76%), one or more structures of a small molecule bound to the NME target were freely available from the PDB.

(iv) Structure-Guided Lead Optimization: Co-crystal structures are widely used across the biopharmaceutical industry to guide optimization of potency (reviewed in [1]). In the most favorable cases, knowledge of co-crystal structures with potential off targets (e.g., GSK-3 β , inhibition of which causes hyperglycemia) can also be employed to help ensure the desired selectivity profile and reduce the likelihood of off-target toxicity. In the absence of experimental co-crystal structures of the target protein, *in silico* docking tools are typically used to guide lead optimization (reviewed in [26]). Where an experimental 3D structure of the target protein is not available, homology models are routinely combined with these same *in silico* docking tools. Machine learning approaches are also being used with increasing frequency to drive medicinal chemistry campaigns (reviewed in [27]).

In 39/54 (~72%) of cases, there is direct or indirect evidence from the PDB archive and the scientific literature that structure-guided lead optimization with the target protein or computational tools with public domain PDB structures have been employed by one or more biopharmaceutical companies prosecuting the NME target (Table 3).

Not surprisingly, the vast majority of the 39 LMW-NMEs identified as confirmed or probable products of structure-guided discovery correspond to 28 of the LMW-NMEs targeting one or more protein kinases (Table 3).

24 of the kinase inhibitors were confirmed as products of structure-guided drug discovery (“Yes” in Table 3) on the basis of direct evidence from the scientific

literature (or private communications) that the sponsor company or its predecessor (for acquired programs) or a competitor company used crystallography and/or computational modeling to study how each LMW-NME bound to its target protein [28-46].

- 4 kinase inhibitors were identified as probable products of structure-guided drug discovery (“Prob” in Table 3) on the basis of indirect evidence. In these cases, a PDB structure of the target protein was publicly available 10 or more years prior to approval and an academic research group used crystallography to study each LMW-NME bound to its target protein. We classified these four less clear-cut cases as probable, because we think it highly likely that the sponsor company was in possession of the same or similar data given the ubiquity of expert protein crystallography and computational chemistry teams across the industry [47-49].
- 4 kinase inhibitors for which a PDB structure of the target protein was publicly available less than 10 years prior to approval were identified as possible products of structure-guided drug discovery (“Poss” in Table 3). In these cases we were unable to find direct or indirect evidence that they were products of structure-guided drug discovery.
- 1 kinase inhibitor, midostaurin, was identified as a natural product derivative (“Nat Prod” in Table 3).

The remaining 11 LMW-NMEs identified as confirmed or probable products of structure-guided discovery target other classes of proteins (Table 3).

- 3 LMW-NMEs, including two isocitrate dehydrogenase inhibitors (ivosidentib, enasidenib) and venetoclax (targeting BCL-2) were confirmed as products of structure-guided drug discovery (“Yes” in Table 3) on the basis of direct evidence from the scientific literature (or private communications with industry experts) that the sponsor company used crystallography to study how each LMW-NME bound to its target protein [50,51].
- 8 LMW-NMEs, including four poly ADP-ribose polymerase (PARP) inhibitors, two non-steroidal antiandrogens, and two histone deacetylase (HDAC) inhibitors, were identified as probable products of structure-guided drug discovery (“Prob” in Table 3) on the basis of indirect evidence. In these cases, a PDB structure of the target protein was publicly available at least 10 prior to approval. With the exception of the two antiandrogens an academic research group had used crystallography to study each LMW-NME bound to its target protein. We classified

these eight less clear-cut cases as probable, because we think it highly likely that the sponsor company or its predecessor (for acquired programs) was in possession of the same or similar data given the ubiquity of expert protein crystallography and computational chemistry teams across the industry [52,53].

- 6 LMW-NMEs were identified as unlikely to be products of structure-guided drug discovery (“Unl” in Table 3). In these cases, a PDB structure of the target protein was either not publicly available at the time of approval or only became available shortly before approval.
- 4 LMW-NMEs were identified as a natural product derivatives (“Nat Prod” in Table 3).

The breadth and depth of PDB structures and publications coming from industry revealed by our analyses confirm that 3D structures are impacting discovery of LMW-NMEs in real time. Conservative estimates suggest that X-ray crystal structures of proteins held as trade secrets inside company firewalls across the biopharmaceutical industry are comparable in aggregate to PDB archival holdings (*i.e.*, ~160,000 structures). Willingness on the part of industry to share a subset of these data with academic researchers is essential for the long-term health of the experimental and computational eco-systems that support structure-guided drug discovery. It is encouraging that approximately two-thirds (27/41, ~66%) of the PDB structures of the anti-neoplastic LMW-NMEs bound to their targets enumerated in Table 3 were deposited by industrial protein crystallography teams. (N.B.: Given the highly competitive nature of biopharmaceutical industry, PDB deposition of structures from biopharmaceutical companies often lags the actual research.) Equally encouraging is the fact that a number of biopharmaceutical companies generously contributed “post-competitive” co-crystal structures and affinity data that enabled blinded computational docking/scoring challenges organized over the past five years by the Drug Design Data Resource (D3R, <https://drugdesigndata.org>; [54-56]).

(v) Optimization of ADME Properties: Finally, 3D structures of proteins are also used to overcome **ADME** (**A**bsorption-**D**istribution-**M**etabolism-**E**xcretion) issues (reviewed in [5]).

Of particular relevance are PDB structures of cytochrome P450 enzymes (earliest PDB ID: 1og2 [57]), the P-glycoprotein multi-drug transporter (earliest PDB ID: 3g5u [58]), and the human ether-a-go-go related potassium channel (earliest PDB ID: 5va1 [59]). Notwithstanding availability of 3D structures for these and other

ADME related proteins in the PDB, we were unable to find evidence in the scientific literature that 3D structure was per se used to overcome ADME issues for any of the 54 LMW-NMEs enumerated in Table 3.

Biologic NMEs: The 25 Biologic-NMEs approved by US FDA 2010-2018 were divided into three types (Table 4). Most (20/25, 80%) are either monoclonal antibodies (16/25, 64%; 12 unique targets) or antibody-drug conjugates (ADCs: 4/25, 16%; 3 unique targets), which consist of a monoclonal antibody linked to either a protein toxin (Moxetumomab pasudotox-tdfk) or a small-molecule drug (Brentuximab vedotin, Ado-trastuzumab emtansine, and Inotuzumab ozogamicin). Five “Other” Biological-NMEs (Table 4) include two L-asparaginases (Calaspargase pegol-mknl; Asparaginase *Erwinia chrysanthemi*), extracellular portions of Vascular Endothelial Growth Factor Receptor (VEGFR) 1 and VEGFR 2 fused to human IgG1 immunoglobulin Fc domains (Ziv-aflibercept, targeting VEGF1 and VEGF2), Human Interleukin 3 (IL3) fused to Diphtheria Toxin (Tagraxofusp-erzs, targeting the IL3 receptor), and a radiolabeled oligopeptide that targets Somatostatin Receptors (Lutetium Lu 177 dotatate).

For 24 of the 25 Biologic-NMEs (96%), we identified more than 479 “Relevant Structures” in the PDB, which include the following: (1) the reference or mutant/variant form of the protein targeted by the Biologic-NME; (2) all or part of the Biologic-NME itself; or (3) all or part of the Biologic-NME bound to a reference or mutant/variant form of its target protein.

All 479 Relevant Structures were deposited to the PDB well before the NME was approved by US FDA for clinical use (data not shown). The median time between PDB deposition and FDA approval exceeded 9.5 years. The vast majority of the 479 Relevant Structures (415/479, 87%) were reported in a PubMed-indexed publication around the time of PDB deposition. These 415 papers had garnered 42,115 literature citations as of September 2019, giving an average of >101 citations/primary publication *versus* the average across the entire PDB archive of ~50 [22].

Review of PDB archival holdings and the scientific literature pertaining to each NME target/Biologic-NME combination summarized in Table 4 revealed that public domain 3D structure data facilitated discovery and development of more than 90% of the 25 Biologic-NMEs as follows:

(1) Target Biology: Atomic-level 3D structures provide functional insights that are not always readily apparent from amino acid sequence (reviewed in [22]).

For the 22 Biologic-NMEs with known protein targets, the PDB archive contains one

or more experimentally determined structure for all but two of the targets (20/22, 90%).

(2) Protein Engineering: Across the biopharmaceutical industry, antibody engineering depends critically on our knowledge of 3D structures (reviewed in [7,8,60]). X-ray crystallographic studies of human and mouse antibodies began to bear fruit as early as the 1970s and continue to do so. The first human protein structure in the PDB was that of a Bence-Jones immunoglobulin light-chain dimer (PDB ID: 1rei [61]). The first PDB structure of an Fab fragment was that of McPC603, a phosphocholine-binding mouse myeloma protein (PDB ID: 1mcp [62], deposited in 1985). The first PDB structure of single chain Fv was that of Se155-4 bound to a trisaccharide ligand (PDB ID: 1mfa [63], deposited in 1994). Today, thousands of antibody structures are represented in the PDB, ranging from entire immunoglobulins to Fab fragments and single chain Fvs.

Design of the limited repertoire of molecular scaffolds used across the biopharmaceutical industry utilized knowledge of PDB structures, making all 20 of the recently approved antibodies or ADCs indirect products of 3D structure. It was not possible from public domain information to determine whether or not project-specific structure data directly drove engineering of a particular antibody or ADC. Consultation with industry experts revealed that proprietary 3D structures held inside company firewalls are used in a substantial number of cases but by no means the majority (private communications). In a limited number of cases, structures of these antibody frameworks have been publicly disclosed and deposited to the PDB (e.g., PDB ID: 4kmt [64], 5i15, 5i16, 5i17, 5i18, 5i19, 5i1a, 5i1c, 5i1d, 5i1e, 5i1g, 5i1h, 5i1i, 5i1j, 5i1k and 5i1l [65]).

Going beyond conventional antibody scaffolds, PDB structure data are also in routine use across the biopharmaceutical industry to guide design of various bispecific and trispecific agents. Three such molecules were approved by US FDA 2010-2018.

Blinatumomab approximates T-cells to the surfaces of malignant B-cells by simultaneously targeting CD3 on the T-cell and CD19 on the B-cell using two antibody variable region heterodimers (V_L - V_H) fused together by a linker between the two V_H segments [66].

Tagraxofusp-erzs consists of interleukin 3 (IL3) fused to a truncated form of the diphtheria toxin (DT) protein. First PDB structures of IL3 (PDB ID: 1jli [67]) and DT (PDB ID: 1ddt [68]) were both made publicly available in the mid 1990s, well

before approval in 2018. Frankel et al. (2000) [69] described use of PDB ID: 1ddt [68] to guide design of various IL3-DT fusions.

Ziv-aflibercept is a triple-fusion protein consisting of extracellular portions of VEGFR1 fused to corresponding extracellular portions of VEGFR2 fused to an IgG1-Fc domain. The dimeric assembly targets free VEGF1 and VEGF2 growth factors for internalization and degradation by cells bearing Fc receptors. The first PDB structures of IgG Fc (PDB ID: 1fc1 [70]) and VEGFR1 (PDB ID: 1flt [71]) were made publicly available in 1981 and 1995, respectively, well before US FDA approval in 2012.

(3) **Molecular Recognition:** Use of 3D structures to understand how antibodies bind their target proteins has contributed to biologic drug discovery in various ways. For example, structural studies of anti-HER2 antibodies showed that they bind to distinct antigenic epitopes, revealing the molecular underpinnings of effective combination antibody therapy for breast cancer (reviewed in [72]) with pertuzumab (PDB ID: 1s78 [73]) and trastuzumab (PDB ID: 1n8z [74]). Co-crystal structures of monoclonal antibodies recognizing their targets and other biophysical findings also provide detailed maps of target binding sites. This information is frequently used in patent applications to strengthen intellectual property protection claims for Biologic-NMEs [75].

PDB structures provide insights into how 8 of 16 (50%) antibodies and 1 of 4 (25%) antibody-drug conjugates bind to their protein targets.

Case Studies

Going beyond these aggregate analyses, we now review three case studies illustrating the impact of PDB data: (i) three CDK4/CDK6 inhibitors; (ii) an IDH2 inhibitor; and (iii) seven therapeutic agents for treatment of advanced-stage melanoma. The following considerations influenced selection of NMEs for inclusion in these three case studies.

The CDK4/CDK6 inhibitors were selected because they exemplify parallel use of structure-guided drug discovery by three large biopharmaceutical companies that competed head-to-head on targeting precisely the same binding site in two closely related protein kinases (*i.e.*, the hinge regions), building on PDB structure data for many protein kinases including one of the targets (human CDK6 [76]) that entered the public domain ~17 years prior to US FDA approval.

The IDH2 allosteric inhibitor was selected because it exemplified use of structure-guided drug discovery by a small biotechnology company, building on a lone PDB structure of a highly similar mammalian homolog of human IDH2 (*i.e.*, porcine IDH2 [77]) that entered the public domain ~15 years prior to US FDA approval. The PDB structure of the LMW-NME bound to IDH2 was deposited by the sponsor company ~1 year in advance of approval [50]).

Finally, seven NMEs that have transformed clinical management of advanced-stage melanoma were selected for detailed review, including four protein kinase inhibitors and three monoclonal antibodies targeting five distinct human proteins in aggregate. They were discovered and developed by six companies competing intensively in the same clinical arena to address very considerable unmet medical needs. These seven discovery/development efforts built on understanding of target biology and target druggability in 3D, which was facilitated by open access to thousands of PDB structures of cellular signaling proteins and the four target proteins and their complexes with proteins and small-molecule ligands. Three of the four LMW-NMEs were the product of structure-guided drug discovery campaigns targeting BRAF. PDB structure data preceded US FDA approval by only seven years in the earliest instance (*i.e.*, vemurafenib, approved in 2011). The relatively short timeline reflects the combined impact of understanding target biology and target druggability on target selection (*i.e.*, mutant BRAF), structure-guided lead compound discovery/optimization, and highly focused clinical trial design that together with a companion diagnostic supported accelerated approval by US FDA. Protein engineering of the three Biologic-NMEs was, at a minimum, indirectly facilitated by open access to the extensive collection of antibody structures housed within the PDB archive. Co-crystal structures of all three Biologic-NMEs with their target proteins are also freely available from the PDB.

CDK4/CDK6 Case Study: Two closely related cyclin-dependent kinases, CDK4 and CDK6, are responsible for controlling progression through the G1 phase of the cell cycle, playing central roles in cell proliferation and tumorigenesis. The first PDB structures of human CDK4 and human CDK6 both entered the public domain more than a decade ago (CDK4-PDB ID: 2w96 [78]; CDK6-PDB ID: 1bi8 [76]). Efforts to discover and develop CDK4 and CDK6 inhibitors as targeted cancer therapies began in the early 1990s (reviewed in [79]), culminating in US FDA approval of three dual CDK4/CDK6 inhibitors for treatment of breast cancer (palbociclib, ribociclib, and abemaciclib). All three of these LMW-NMEs came from structure-guided drug discovery efforts carried out independently by different sponsor companies. Each discovery team could rely on open access to tens of CDK structures and thousands of other protein kinase structures previously archived in the PDB. Co-crystal structures of each new drug bound to CDK6 were generously deposited to the PDB by Pfizer. Figure 1 compares the earliest structures of CDK4 (PDB ID: 2w96)

and CDK6 (PDB ID: 1bi7) with co-crystal structures for palbociclib (PDB ID: 5l2i [31]), ribociclib (PDB ID: 5l2t [31]), and abemaciclib (PDB ID: 5l2s [31]) bound to CDK6. Close inspection of the modes of inhibitor binding reveals both common (*e.g.*, hydrogen bonding engagement of the hinge region) and disparate features of CDK6-ligand interactions for the three inhibitors (Figure 1D).

IDH2 Case Study. The first 3D structure of a mammalian IDH2 (porcine, 96% identical in amino acid sequence to human) was deposited to the PDB in 2002 by academic researchers (PDB ID: 1lwd [77]). IDH2 is a homodimeric, NADP(+)-dependent, mitochondrial enzyme responsible for catalyzing oxidative decarboxylation of isocitrate to 2-oxoglutarate. Certain IDH2 gene mutations confer a gain-of-function on malignant cells, resulting in accumulation and secretion of the oncometabolite (*R*)-2-hydroxyglutarate (reviewed in [80]). As of September 2019, the PDB archive housed six X-ray structures of human IDH2, all of which were contributed by biopharmaceutical companies (Novartis or Agios). Agios deposited the earliest human IDH2 structure (PDB ID: 4ja8 [81]), which revealed the allosteric mechanism by which one of their proprietary compounds (AGI-6780) inhibited the R140Q form of IDH2 by binding within the dimer interface (data not shown). A structure-guided drug discovery campaign at Agios subsequently yielded enasidenib (Figure 2, PDB ID: 5i96 [50]). This LMW-NME was approved by US FDA in 2017 for relapsed or refractory acute myeloid leukemia in individuals with specific mutations of the IDH2 gene confirmed by an FDA-approved diagnostic test.

Advanced-stage Melanoma Case Study: Changing clinical management paradigms for advanced-stage (Stages 3 and 4) melanoma provide compelling evidence for the transformative impact of 3D structure information and structure-guided drug discovery on US FDA drug approvals. Ten years ago, treatment options for this disease were decidedly limited and of marginal benefit (*i.e.*, median overall survival ~9 months). Today, median overall survival exceeds 2 years [82], and is expected to increase further with optimization of standard-of-care using these and other recently approved agents.

Seven NMEs were approved for treatment of unresectable or metastatic melanoma by US FDA 2010-2018. Four of these new drugs are LMW-NMEs that inhibit protein kinases (vemurafenib, approved 2011; dabrafenib, 2013; trametinib, 2013; encorafenib, 2018). The remainder are Biologic-NMEs that target cytotoxic T lymphocyte protein-4 (CTLA-4) (ipilimumab, approved 2011) or programmed death receptor-1 (PD-1) (pembrolizumab, 2014; nivolumab, 2014) and block downregulation of T-cell function by tumor cells.

The first PDB structure of the catalytic domain of wild-type human BRAF (PDB ID: 1uwh [83]) was deposited by academic researchers in 2004. At that time, >30 BRAF gene

mutations had been associated with human cancers. Most of these mutations mapped to the activation segment or the P-loop within the catalytic domain, where they were thought to destabilize the inactive conformation of the enzyme. A second PDB structure of mutant BRAF (PDB ID: 1uwj [83]) contributed by the same group revealed a tool compound (BAY43-9006) binding to the inactive conformation of the enzyme. It was subsequently documented that the V600E mutant form of BRAF is present in ~50% of late-stage, metastatic melanomas (reviewed in [84]), making this mutant enzyme a highly attractive discovery target.

Vemurafenib was discovered by Plexxikon in the course of a well-publicized fragment-based, structure-guided lead-optimization campaign targeting V600E BRAF (PDB ID: 3c4c, 3c4d, 3c4e, 3c4f, and 4fk3 [85]). The Plexxikon structure of vemurafenib bound to V600E BRAF (PDB ID: 3og7 [43]) is illustrated in Figure 3A. Dabrafenib was discovered by GSK with the aid of computational docking into one of the PDB structures of V600E BRAF determined by Plexxikon [86]. Two PDB structures of dabrafenib bound to other mutant forms of BRAF were subsequently contributed to the PDB by Boehringer Ingelheim (PDB ID: 5cs2 [87]) and Genentech (PDB ID: 5hie [88]). Vemurafenib binding to V600E BRAF is compared in Figure 3B to that of dabrafenib (PDB ID: 4xv2 [34]). Close inspection of the modes of inhibitor binding reveals both common (*e.g.*, hydrogen bonding engagement of the hinge region) and disparate features of BRAF-ligand interactions for the two inhibitors (Figure 3B).

BRAF V600E mutations result in constitutive activation of the signalling pathway that includes the mitogen-activated protein kinases MEK1 and MEK2 (reviewed in [89]), making these enzymes attractive drug discovery targets for advanced-stage melanoma. The earliest public-domain human MEK1 (PDB ID: 1s9j [90]) and MEK2 (PDB ID: 1s9i [90]) structures were deposited to PDB in 2004 by Pfizer. Trametinib was discovered by Japan Tobacco using medicinal chemistry optimization of a high-throughput screening hit [91]. Trametinib is not represented in the PDB. Unlike the three ATP-competitive inhibitors of BRAF (vemurafenib, dabrafenib, and encorafenib), trametinib is an allosteric inhibitor of MEK1/MEK2 [92]. Trametinib inhibits BRAF V600 mutation positive melanoma cell growth *in vitro* and *in vivo*. (N.B.: Trametinib was approved for the treatment of patients who have not received prior BRAF inhibitor therapy.)

Vemurafenib and dabrafenib were each initially approved for single-agent treatment of patients with unresectable or metastatic melanoma with BRAF V600E mutation confirmed by an FDA-approved test. Initial results were promising, with objective tumor responses in approximately half of patients with advanced-stage melanoma. However, the duration of responses proved limited in most patients, with progression-free survival of ~6 months due to emergence of acquired resistance following activation of MEK1

and/or MEK2 [93]. In 2015, cobimetinib, an additional MEK inhibitor, was approved in combination with vemurafenib for treatment of unresectable or metastatic melanoma with a BRAF V600E or V600K mutation. Cobimetinib was discovered by Exelixis during the course of a structure-guided drug discovery campaign (PDB ID: 4an2 [40]).

The remaining LMW-NME targeting BRAF (encorafenib) was approved in 2018 for use in combination with the MEK inhibitor binimetinib for treatment of patients with unresectable or metastatic melanoma with BRAF V600E mutation confirmed by an FDA-approved diagnostic test. Encorafenib was discovered by the Novartis Institutes for Biomedical Research during the course of a structure-guided drug discovery campaign [94]. Encorafenib is not represented in the PDB. Binimetinib is a mitogen-activated protein kinase 1/2 (MEK 1/2) inhibitor discovered by Array Biopharma, Inc. during the course of a structure-guided drug discovery campaign (private communication). The encorafenib/binimetinib combination showed significant clinical benefit *versus* encorafenib or vemurafenib used as single agents [95]. (See [96] for a comprehensive review of the structural biology of small-molecule BRAF inhibitors.)

Modulation of T-cell mediated immunity is a medically important phenomenon that has been significantly impacted by structural biologists and the PDB. The archive currently houses >750 related PDB structures, which together reveal the molecular mechanisms underpinning antigen presentation to T-cell receptors and also explain much of T-cell regulation in 3D. The earliest such contribution was the landmark crystal structure of the major histocompatibility complex (MHC) (PDB ID: 1hla [97]). Subsequently deposited PDB structures revealed how MHC presents linear peptide antigens to T-cells (*e.g.*, PDB ID: 1hsa [98]) and in turn how MHC-peptide antigen complexes are recognized by T-cell receptors (*e.g.*, PDB ID: 1ao7 [99]). Thereafter, structural biologists revealed at the atomic level many of the protein-protein interactions responsible for regulating T-cells. Various biopharmaceutical companies acted on these insights by successfully targeting immune checkpoints leading to US FDA approval of seven antibody therapeutics 2010-2018.

The first of these Biologic-NMEs (ipilimumab, Bristol-Myers Squibb, approved 2011) targets CTLA-4, thereby blocking negative regulation of T-cells by B7-1 or B7-2 proteins found on the surface of tumor cells. The PDB houses multiple structures of CTLA-4 (earliest PDB ID: 1ah1 [100]), including those of CTLA-4 binding to B7-1 (PDB ID: 1il8 [101]) and CTLA-4 binding to B7-2 (PDB ID: 1ah1 [102]). Publication of the co-crystal structure of the Fab fragment of ipilimumab recognizing CTLA-4 followed some years after drug approval (PDB ID: 5tru [103]).

Nivolumab (Bristol-Myers Squibb) and pembrolizumab (Merck) were both approved in 2014. These antibodies target PD-1, thereby blocking downregulation of T-cells due to PD-1 binding to programmed death receptor-ligand 1 (PD-L1) or PD-L2 found on the surface of tumor cells (Figure 4). The PDB houses multiple structures of PD-1 (earliest PDB ID: 1npu [104]), PD-L1 (earliest PDB ID: 3bis [105]), and PD-L2 (earliest PDB ID: 3bov [106]). In addition, the PDB contains structures of PD-1/PD-L1 complexes (earliest PDB ID: 3bik [105]; Figure 4A) and PD-1/PD-L2 complexes (earliest PDB ID: 3bp5 [106]; Figure 4B). Structures of both nivolumab (earliest PDB ID: 5ggq [107]) and pembrolizumab (earliest PDB ID: 5dk3 [108]) are similarly available from the PDB. The PDB also contains multiple structures of nivolumab/PD-1 complexes (earliest PDB ID: 5ggr [107]) and pembrolizumab/PD-1 complexes (earliest PDB ID: 5jxe [109]; Figure 4C).

Current standard-of-care for advanced-stage melanoma [110,111] begins with either pembrolizumab or nivolumab, particularly in individuals whose tumor cells do not possess mutant BRAF. Both of these Biologic-NMEs can shrink tumors for long periods of time in favorable cases (*e.g.*, President Jimmy Carter, who benefited from pembrolizumab). Ipilimumab is not typically used as first line treatment, although it can be combined with nivolumab or pembrolizumab to improve the likelihood of a tumor response. If a BRAF gene mutation is detected in the affected individual's tumor, combination therapy with a small-molecule BRAF inhibitor plus a MEK inhibitor (*e.g.*, vemurafenib/cobimetinib, dabrafenib/trametinib, or encorafenib/binimetinib) can be used as an alternative first line treatment strategy. At present, optimal choices as to first line treatment, combinations of antibodies, and combinations of antibodies with targeted agents are being evaluated in clinical trials. Prognoses for individuals with advanced-stage melanoma appear likely to improve further as clinical oncologists and dermatologists gain more experience using these new agents.

Concluding Remarks

This review documents that PDB structure data contribute broadly to oncology drug discovery/development in the biopharmaceutical industry (and to a lesser extent in academe). For the 54 LMW-NMEs analyzed, all of which have known protein targets, there is evidence from the PDB and/or the scientific literature that discovery and development of every one of these new drugs was facilitated by the availability of public-domain 3D structure information. In >70% of cases, the LMW-NMEs were the product of biopharmaceutical company structure-guided drug discovery efforts, involving co-crystal structure studies and/or computational docking into crystal structures, *etc.* For the 25 Biologic-NMEs analyzed there is again evidence from the PDB and/or the scientific literature that discovery and development of more than 90% these new drugs were

facilitated directly or indirectly by the availability of public-domain 3D structure information.

With year-on-year growth in the number of structures in the PDB approaching 10%, the impact of the resource and structure-guided approaches on drug discovery/development is destined to remain significant. Moreover, the growing number of PDB structures coming from cryo-electron microscopy since the advent of the “Resolution Revolution” [112], promises even broader 3D structural coverage of the human proteome. We can expect deposition of new PDB structures of many of the integral membrane proteins and other macromolecular machines that are currently being sub-optimally targeted with relatively non-specific agents or are considered to be undruggable [113].

The long-standing requirement for PDB deposition of 3D atomic coordinates and experimental data and metadata upon publication ensures that this valuable information is made immediately available to basic and applied researchers around the world without limitations on usage. Moreover, expert biocuration and standardized validation of the experimental data and the atomic coordinates across the PDB help to ensure that the archive as a whole can be mined for new knowledge using statistical tools [114,115] or machine learning approaches [27].

As custodian of the PDB Core Archive, the wwPDB partnership is committed to the ***FAIR*** Principles [10], which help ensure the broadest possible use of public domain biomedical research data. The PDB has been recognized as a Core Certified Repository by CoreTrustSeal (coretrustseal.org). This international, community-based, non-governmental, non-profit organization promotes sustainable and trustworthy data infrastructures of which the PDB is widely regarded as a gold-standard exemplar.

Supplementary Information is provided describing the assembly and analysis of the data set described in this review.

Acknowledgements We thank the more than 30,000 structural biologists who have deposited structures to the PDB Core Archive since 2000. We also thank Drs. James (Jay) Bradner, Kirk L. Clark, Marc C. Deller, Dashyant Dhanak, Jose Duarte, Gary L. Gilliland, Fred D. Ledley, Catherine E. Peishoff, Alexander Rose, Joan Segura, Paul A. Sprengeler, Jonathan H. Watanabe, Christine Zardecki, and X.F. Steven Zheng for their insightful comments, and we gratefully acknowledge contributions from all members of the Research Collaboratory for Structural Bioinformatics PDB and our Worldwide Protein Data Bank partners. RCSB PDB is jointly funded by the National Science Foundation (DBI-1832184), the US Department of Energy (DE-SC0019749), and the National Cancer Institute, National Institute of Allergy and Infectious Diseases, and National Institute of General Medical Sciences of the National Institutes of Health under grant R01GM133198.

Author Contributions J.D.W., R.S, B.P.H., and S.K.B. assembled and analyzed the data and wrote the manuscript.

Author Information The authors declare no competing financial interests.

References

- 1 Blundell, T.L. (2017) Protein crystallography and drug discovery: recollections of knowledge exchange between academia and industry. *IUCrJ* 4 (Pt 4), 308-321
- 2 Klebe, G. (2013) *Drug design: methodology, concepts, and mode-of-action*, Berlin: Springer
- 3 Burley, S.K. et al. (2018) RCSB Protein Data Bank: Sustaining a living digital data resource that enables breakthroughs in scientific research and biomedical education. *Protein Sci* 27 (1), 316-330
- 4 Brown, K.K. et al. (2018) Approaches to target tractability assessment - a practical perspective. *Medchemcomm* 9 (4), 606-613
- 5 Stoll, F. et al. (2011) Utility of protein structures in overcoming ADMET-related issues of drug-like compounds. *Drug Discov Today* 16 (11-12), 530-538
- 6 Mullard, A. (2016) 2015 FDA drug approvals. *Nat Rev Drug Discov* 15 (2), 73-76
- 7 Gilliland, G.L. et al. (2012) Leveraging SBDD in protein therapeutic development: antibody engineering. *Methods Mol Biol* 841, 321-349
- 8 Chiu, M.L. and Gilliland, G.L. (2016) Engineering antibody therapeutics. *Curr Opin Struct Biol* 38, 163-173
- 9 Protein Data Bank. (1971) Crystallography: Protein Data Bank. *Nature (London), New Biol.* 233 (42), 223-223
- 10 Wilkinson, M.D. et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3 (160018), 1-9
- 11 Berman, H.M. et al. (2003) Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* 10 (12), 980
- 12 wwPDB consortium. (2019) Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res* 47 (D1), D520-D528
- 13 Berman, H.M. et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28 (1), 235-242

- 14 Burley, S.K. et al. (2019) RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res* 47 (D1), D464-D474
- 15 Mir, S. et al. (2018) PDBe: towards reusable data delivery infrastructure at protein data bank in Europe. *Nucleic Acids Res* 46 (D1), D486-D492
- 16 Kinjo, A.R. et al. (2017) Protein Data Bank Japan (PDBj): updated user interfaces, resource description framework, analysis tools for large structures. *Nucleic Acids Res* 45 (D1), D282-D288
- 17 Ulrich, E.L. et al. (2008) BioMagResBank. *Nucleic Acids Res* 36 (Database issue), D402-408
- 18 Westbrook, J.D. and Burley, S.K. (2019) How Structural Biologists and the Protein Data Bank Contributed to Recent FDA New Drug Approvals. *Structure* 27, 211-217
- 19 Liu, Z. et al. (2017) Lessons Learned from Two Decades of Anticancer Drugs. *Trends Pharmacol Sci* 38 (10), 852-872
- 20 Hu, T. et al. (2018) The impact of structural biology in medicine illustrated with four case studies. *J Mol Med (Berl)* 96 (1), 9-19
- 21 The UniProt Consortium. (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 45 (D1), D158-D169
- 22 Burley, S.K. et al. (2018) RCSB Protein Data Bank: Sustaining a living digital data resource that enables breakthroughs in scientific research and biomedical education. *Protein Sci* 27, 316-330
- 23 Renaud, J.P. et al. (2016) Biophysics in drug discovery: impact, challenges and opportunities. *Nat Rev Drug Discov* 15 (10), 679-698
- 24 Erlanson, D.A. et al. (2016) Twenty years on: the impact of fragments on drug discovery. *Nat Rev Drug Discov* 15 (9), 605-619
- 25 Burns, M.C. et al. (2018) High-throughput screening identifies small molecules that bind to the RAS:SOS:RAS complex and perturb RAS signaling. *Anal Biochem* 548, 44-52
- 26 Lin, J.H. (2016) Review structure- and dynamics-based computational design of anticancer drugs. *Biopolymers* 105 (1), 2-9

- 27 Lo, Y.C. et al. (2018) Machine learning in chemoinformatics and drug discovery. *Drug Discov Today* 23 (8), 1538-1546
- 28 Kawase, T. et al. (2019) Effect of Fms-like tyrosine kinase 3 (FLT3) ligand (FL) on antitumor activity of gilteritinib, a FLT3 inhibitor, in mice xenografted with FL-overexpressing cells. *Oncotarget* 10 (58), 6111-6123
- 29 Matsuki, M. et al. (2018) Lenvatinib inhibits angiogenesis and tumor fibroblast growth factor signaling pathways in human hepatocellular carcinoma models. *Cancer Med* 7 (6), 2641-2653
- 30 Bender, A.T. et al. (2017) Ability of Bruton's Tyrosine Kinase Inhibitors to Sequester Y551 and Prevent Phosphorylation Determines Potency for Inhibition of Fc Receptor but not B-Cell Receptor Signaling. *Mol Pharmacol* 91 (3), 208-219
- 31 Chen, P. et al. (2016) Spectrum and Degree of CDK Drug Interactions Predicts Clinical Performance. *Mol Cancer Ther* 15 (10), 2273-2281
- 32 Scott, W.J. et al. (2016) Discovery and SAR of Novel 2,3-Dihydroimidazo[1,2-c]quinazoline PI3K Inhibitors: Identification of Copanlisib (BAY 80-6946). *ChemMedChem* 11 (14), 1517-1530
- 33 Huang, W.S. et al. (2016) Discovery of Brigatinib (AP26113), a Phosphine Oxide-Containing, Potent, Orally Active Inhibitor of Anaplastic Lymphoma Kinase. *J Med Chem* 59 (10), 4948-4964
- 34 Zhang, C. et al. (2015) RAF inhibitors that evade paradoxical MAPK pathway activation. *Nature* 526 (7574), 583-586
- 35 Somoza, J.R. et al. (2015) Structural, biochemical, and biophysical characterization of idelalisib binding to phosphoinositide 3-kinase delta. *J Biol Chem* 290 (13), 8439-8446
- 36 Friboulet, L. et al. (2014) The ALK inhibitor ceritinib overcomes crizotinib resistance in non-small cell lung cancer. *Cancer Discov* 4 (6), 662-673
- 37 Gajiwala, K.S. et al. (2013) Insights into the aberrant activity of mutant EGFR kinase domain and drug recognition. *Structure* 21 (2), 209-219
- 38 McTigue, M. et al. (2012) Molecular conformations, interactions, and properties associated with drug efficiency and clinical performance among VEGFR TK inhibitors. *Proc Natl Acad Sci U S A* 109 (45), 18281-18289

- 39 Solca, F. et al. (2012) Target binding properties and cellular activity of afatinib (BIBW 2992), an irreversible ErbB family blocker. *J Pharmacol Exp Ther* 343 (2), 342-350
- 40 Rice, K.D. et al. (2012) Novel Carboxamide-Based Allosteric MEK Inhibitors: Discovery and Optimization Efforts toward XL518 (GDC-0973). *ACS Med Chem Lett* 3 (5), 416-421
- 41 Cui, J.J. et al. (2011) Structure based drug design of crizotinib (PF-02341066), a potent and selective dual inhibitor of mesenchymal-epithelial transition factor (c-MET) kinase and anaplastic lymphoma kinase (ALK). *J Med Chem* 54 (18), 6342-6363
- 42 Sakamoto, H. et al. (2011) CH5424802, a selective ALK inhibitor capable of blocking the resistant gatekeeper mutant. *Cancer Cell* 19 (5), 679-690
- 43 Bollag, G. et al. (2010) Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* 467 (7315), 596-599
- 44 O'Hare, T. et al. (2009) AP24534, a pan-BCR-ABL inhibitor for chronic myeloid leukemia, potently inhibits the T315I mutant and overcomes mutation-based resistance. *Cancer Cell* 16 (5), 401-412
- 45 Qian, F. et al. (2009) Inhibition of tumor cell growth, invasion, and metastasis by EXEL-2880 (XL880, GSK1363089), a novel inhibitor of HGF and VEGF receptor tyrosine kinases. *Cancer Res* 69 (20), 8009-8016
- 46 Johnson, T.W. et al. (2014) Discovery of (10R)-7-amino-12-fluoro-2,10,16-trimethyl-15-oxo-10,15,16,17-tetrahydro-2H-8,4-(metheno)pyrazolo[4,3-h][2,5,11]-benzoxadiazacyclotetradecine-3-carbonitrile (PF-06463922), a macrocyclic inhibitor of anaplastic lymphoma kinase (ALK) and c-ros oncogene 1 (ROS1) with preclinical brain exposure and broad-spectrum potency against ALK-resistant mutations. *J Med Chem* 57 (11), 4720-4744
- 47 Yosaatmadja, Y. et al. (2015) Binding mode of the breakthrough inhibitor AZD9291 to epidermal growth factor receptor revealed. *J Struct Biol* 192 (3), 539-544
- 48 Levinson, N.M. and Boxer, S.G. (2014) A conserved water-mediated hydrogen bond network defines bosutinib's kinase selectivity. *Nat Chem Biol* 10 (2), 127-132
- 49 Knowles, P.P. et al. (2006) Structure and chemical inhibition of the RET tyrosine kinase domain. *J Biol Chem* 281 (44), 33577-33587

- 50 Yen, K. et al. (2017) AG-221, a First-in-Class Therapy Targeting Acute Myeloid Leukemia Harboring Oncogenic IDH2 Mutations. *Cancer Discov* 7 (5), 478-493
- 51 Birkinshaw, R.W. et al. (2019) Structures of BCL-2 in complex with venetoclax reveal the molecular basis of resistance mutations. *Nat Commun* 10 (1), 2385
- 52 Thorsell, A.G. et al. (2017) Structural Basis for Potency and Promiscuity in Poly(ADP-ribose) Polymerase (PARP) and Tankyrase Inhibitors. *J Med Chem* 60 (4), 1262-1271
- 53 Hai, Y. and Christianson, D.W. (2016) Histone deacetylase 6 structure and molecular basis of catalysis and inhibition. *Nat Chem Biol* 12 (9), 741-747
- 54 Gathiaka, S. et al. (2016) D3R grand challenge 2015: Evaluation of protein-ligand pose and affinity predictions. *J Comput Aided Mol Des* 30 (9), 651-668
- 55 Gaieb, Z. et al. (2018) D3R Grand Challenge 2: blind prediction of protein-ligand poses, affinity rankings, and relative binding free energies. *J Comput Aided Mol Des* 32 (1), 1-20
- 56 Gaieb, Z. et al. (2019) D3R Grand Challenge 3: blind prediction of protein-ligand poses and affinity rankings. *J Comput Aided Mol Des* 33 (1), 1-18
- 57 Williams, P.A. et al. (2003) Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* 424 (6947), 464-468
- 58 Martinez, L. et al. (2014) Understanding polyspecificity within the substrate-binding cavity of the human multidrug resistance P-glycoprotein. *FEBS J* 281 (3), 673-682
- 59 Wang, W. and MacKinnon, R. (2017) Cryo-EM Structure of the Open Human Ether-a-go-go-Related K(+) Channel hERG. *Cell* 169 (3), 422-430 e410
- 60 Chiu, M.L. et al. (2019) Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies (Basel)* 8 (4)
- 61 Epp, O. et al. (1975) The molecular structure of a dimer composed of the variable portions of the Bence-Jones protein REI refined at 2.0-A resolution. *Biochemistry* 14 (22), 4943-4952
- 62 Satow, Y. et al. (1986) Phosphocholine binding immunoglobulin Fab McPC603. An X-ray diffraction study at 2.7 Å. *J Mol Biol* 190 (4), 593-604

- 63 Zdanov, A. et al. (1994) Structure of a single-chain antibody variable domain (Fv) fragment complexed with a carbohydrate antigen at 1.7-A resolution. *Proc Natl Acad Sci U S A* 91 (14), 6423-6427
- 64 Teplyakov, A. et al. (2014) Antibody modeling assessment II. Structures and models. *Proteins* 82 (8), 1563-1582
- 65 Teplyakov, A. et al. (2016) Structural diversity in a human antibody germline library. *MAbs* 8 (6), 1045-1063
- 66 Burt, R. et al. (2019) Blinatumomab, a bispecific B-cell and T-cell engaging antibody, in the treatment of B-cell malignancies. *Hum Vaccin Immunother* 15 (3), 594-602
- 67 Maier, W. et al. (1996) [Lermoyez syndrome--electrocochleographic studies]. *Laryngorhinootologie* 75 (6), 372-376
- 68 Bennett, M.J. et al. (1994) Refined structure of dimeric diphtheria toxin at 2.0 A resolution. *Protein Sci* 3 (9), 1444-1463
- 69 Frankel, A.E. et al. (2000) Characterization of diphtheria fusion proteins targeted to the human interleukin-3 receptor. *Protein Eng* 13 (8), 575-581
- 70 Deisenhofer, J. (1981) Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9- and 2.8-A resolution. *Biochemistry* 20 (9), 2361-2370
- 71 Wiesmann, C. et al. (1997) Crystal structure at 1.7 A resolution of VEGF in complex with domain 2 of the Flt-1 receptor. *Cell* 91 (5), 695-704
- 72 Olson, E.M. (2012) Maximizing human epidermal growth factor receptor 2 inhibition: a new oncologic paradigm in the era of targeted therapy. *J Clin Oncol* 30 (14), 1712-1714
- 73 Franklin, M.C. et al. (2004) Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. *Cancer Cell* 5 (4), 317-328
- 74 Cho, H.S. et al. (2003) Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* 421 (6924), 756-760
- 75 Deng, X. et al. (2018) Enhancing antibody patent protection using epitope mapping information. *MAbs* 10 (2), 204-209

- 76 Russo, A.A. et al. (1998) Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature* 395 (6699), 237-243
- 77 Ceccarelli, C. et al. (2002) Crystal structure of porcine mitochondrial NADP+-dependent isocitrate dehydrogenase complexed with Mn²⁺ and isocitrate. Insights into the enzyme mechanism. *J Biol Chem* 277 (45), 43454-43462
- 78 Day, P.J. et al. (2009) Crystal structure of human CDK4 in complex with a D-type cyclin. *Proc Natl Acad Sci U S A* 106 (11), 4166-4170
- 79 Sherr, C.J. et al. (2016) Targeting CDK4 and CDK6: From Discovery to Therapy. *Cancer Discov* 6 (4), 353-367
- 80 Molenaar, R.J. et al. (2018) Wild-type and mutated IDH1/2 enzymes and therapy responses. *Oncogene* 37 (15), 1949-1960
- 81 Wang, F. et al. (2013) Targeted inhibition of mutant IDH2 in leukemia cells induces cellular differentiation. *Science* 340 (6132), 622-626
- 82 Luke, J.J. et al. (2017) Targeted agents and immunotherapies: optimizing outcomes in melanoma. *Nat Rev Clin Oncol* 14 (8), 463-482
- 83 Wan, P.T. et al. (2004) Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* 116 (6), 855-867
- 84 Gray-Schopfer, V. et al. (2007) Melanoma biology and new targeted therapy. *Nature* 445 (7130), 851-857
- 85 Tsai, J. et al. (2008) Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc Natl Acad Sci U S A* 105 (8), 3041-3046
- 86 Rheault, T.R. et al. (2013) Discovery of Dabrafenib: A Selective Inhibitor of Raf Kinases with Antitumor Activity against B-Raf-Driven Tumors. *ACS Med Chem Lett* 4 (3), 358-362
- 87 Waizenegger, I.C. et al. (2016) A Novel RAF Kinase Inhibitor with DFG-Out-Binding Mode: High Efficacy in BRAF-Mutant Tumor Xenograft Models in the Absence of Normal Tissue Hyperproliferation. *Mol Cancer Ther* 15 (3), 354-365
- 88 Foster, S.A. et al. (2016) Activation Mechanism of Oncogenic Deletion Mutations in BRAF, EGFR, and HER2. *Cancer Cell* 29 (4), 477-493

- 89 Acosta, A.M. and Kadkol, S.S. (2016) Mitogen-Activated Protein Kinase Signaling Pathway in Cutaneous Melanoma: An Updated Review. *Arch Pathol Lab Med* 140 (11), 1290-1296
- 90 Ohren, J.F. et al. (2004) Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. *Nat Struct Mol Biol* 11 (12), 1192-1197
- 91 Abe, H. et al. (2011) Discovery of a Highly Potent and Selective MEK Inhibitor: GSK1120212 (JTP-74057 DMSO Solvate). *ACS Med Chem Lett* 2 (4), 320-324
- 92 Gilmartin, A.G. et al. (2011) GSK1120212 (JTP-74057) is an inhibitor of MEK activity and activation with favorable pharmacokinetic properties for sustained in vivo pathway inhibition. *Clin Cancer Res* 17 (5), 989-1000
- 93 Eroglu, Z. and Ribas, A. (2016) Combination therapy with BRAF and MEK inhibitors for melanoma: latest evidence and place in therapy. *Ther Adv Med Oncol* 8 (1), 48-56
- 94 Drahl, C. (2013) LGX818, Made To Fight Melanoma. *Chemical & Engineering News* 91, 14
- 95 Dummer, R. et al. (2018) Encorafenib plus binimetinib versus vemurafenib or encorafenib in patients with BRAF-mutant melanoma (COLUMBUS): a multicentre, open-label, randomised phase 3 trial. *Lancet Oncol* 19 (5), 603-615
- 96 Agianian, B. and Gavathiotis, E. (2018) Current Insights of BRAF Inhibitors in Cancer. *J Med Chem* 61 (14), 5775-5793
- 97 Bjorkman, P.J. et al. (1987) Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329, 506
- 98 Madden, D.R. et al. (1992) The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. *Cell* 70 (6), 1035-1048
- 99 Garboczi, D.N. et al. (1996) Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* 384 (6605), 134-141
- 100 Metzler, W.J. et al. (1997) Solution structure of human CTLA-4 and delineation of a CD80/CD86 binding site conserved in CD28. *Nat Struct Biol* 4 (7), 527-531

- 101 Stamper, C.C. et al. (2001) Crystal structure of the B7-1/CTLA-4 complex that inhibits human immune responses. *Nature* 410 (6828), 608-611
- 102 Schwartz, J.C. et al. (2001) Structural basis for co-stimulation by the human CTLA-4/B7-2 complex. *Nature* 410 (6828), 604-608
- 103 Ramagopal, U.A. et al. (2017) Structural basis for cancer immunotherapy by the first-in-class checkpoint inhibitor ipilimumab. *Proc Natl Acad Sci U S A* 114 (21), E4223-E4232
- 104 Zhang, X. et al. (2004) Structural and functional analysis of the costimulatory receptor programmed death-1. *Immunity* 20 (3), 337-347
- 105 Lin, D.Y. et al. (2008) The PD-1/PD-L1 complex resembles the antigen-binding Fv domains of antibodies and T cell receptors. *Proc Natl Acad Sci U S A* 105 (8), 3011-3016
- 106 Lazar-Molnar, E. et al. (2008) Crystal structure of the complex between programmed death-1 (PD-1) and its ligand PD-L2. *Proc Natl Acad Sci U S A* 105 (30), 10483-10488
- 107 Lee, J.Y. et al. (2016) Structural basis of checkpoint blockade by monoclonal antibodies in cancer immunotherapy. *Nat Commun* 7, 13354
- 108 Scapin, G. et al. (2015) Structure of full-length human anti-PD1 therapeutic IgG4 antibody pembrolizumab. *Nat Struct Mol Biol* 22 (12), 953-958
- 109 Na, Z. et al. (2017) Structural basis for blocking PD-1-mediated immune suppression by therapeutic antibody pembrolizumab. *Cell Res* 27 (1), 147-150
- 110 Rozeman, E.A. et al. (2018) Advanced Melanoma: Current Treatment Options, Biomarkers, and Future Perspectives. *Am J Clin Dermatol* 19 (3), 303-317
- 111 Domingues, B. et al. (2018) Melanoma treatment in review. *Immunotargets Ther* 7, 35-49
- 112 Kuhlbrandt, W. (2014) Biochemistry. The resolution revolution. *Science* 343 (6178), 1443-1444
- 113 Goodsell, D.S. et al. (2020) RCSB Protein Data Bank: Enabling biomedical research and drug discovery. *Protein Sci* 29, 52-65

- 114 Shao, C. et al. (2018) Outlier analyses of the Protein Data Bank archive using a probability-density-ranking approach. *Sci Data* 5, 180293
- 115 Shao, C. et al. (2017) Multivariate Analyses of Quality Metrics for Crystal Structures in the Protein Data Bank Archive. *Structure* 25, 458-468
- 116 McTigue, M.A. et al. (1999) Crystal structure of the kinase domain of human vascular endothelial growth factor receptor 2: a key enzyme in angiogenesis. *Structure* 7 (3), 319-330
- 117 Xu, W. et al. (1997) Three-dimensional structure of the tyrosine kinase c-Src. *Nature* 385 (6617), 595-602
- 118 Mol, C.D. et al. (2003) Structure of a c-kit product complex reveals the basis for kinase transactivation. *J Biol Chem* 278 (34), 31461-31464
- 119 Nagar, B. et al. (2002) Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). *Cancer Res* 62 (15), 4236-4243
- 120 Stamos, J. et al. (2002) Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *J Biol Chem* 277 (48), 46265-46272
- 121 Yun, C.H. et al. (2008) The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc Natl Acad Sci U S A* 105 (6), 2070-2075
- 122 Koshiba, S. et al. (2010) Structural basis for the recognition of nucleophosmin-anaplastic lymphoma kinase oncoprotein by the phosphotyrosine binding domain of Suc1-associated neurotrophic factor-induced tyrosine-phosphorylated target-2. *J Struct Funct Genomics* 11 (2), 125-141
- 123 Schiering, N. et al. (2003) Crystal structure of the tyrosine kinase domain of the hepatocyte growth factor receptor c-Met and its complex with the microbial alkaloid K-252a. *Proc Natl Acad Sci U S A* 100 (22), 12654-12659
- 124 Huang, C.H. et al. (2007) The structure of a human p110alpha/p85alpha complex elucidates the effects of oncogenic PI3Kalpha mutations. *Science* 318 (5857), 1744-1748
- 125 Hyvonen, M. and Saraste, M. (1997) Structure of the PH domain and Btk motif from Bruton's tyrosine kinase: molecular explanations for X-linked agammaglobulinaemia. *EMBO J* 16 (12), 3396-3404

- 126 Skerratt, S.E. et al. (2016) The Discovery of a Potent, Selective, and Peripherally Restricted Pan-Trk Inhibitor (PF-06273340) for the Treatment of Pain. *J Med Chem* 59 (22), 10084-10099
- 127 Griffith, J. et al. (2004) The structural basis for autoinhibition of FLT3 by the juxtamembrane domain. *Mol Cell* 13 (2), 169-178
- 128 Xu, X. et al. (2004) Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity. *J Biol Chem* 279 (32), 33946-33957
- 129 Petros, A.M. et al. (2001) Solution structure of the antiapoptotic protein bcl-2. *Proc Natl Acad Sci U S A* 98 (6), 3012-3017
- 130 Kinoshita, T. et al. (2004) Inhibitor-induced structural change of the active site of human poly(ADP-ribose) polymerase. *FEBS Lett* 556 (1-3), 43-46
- 131 Matias, P.M. et al. (2000) Structural evidence for ligand specificity in the binding domain of the human androgen receptor. Implications for pathogenic gene mutations. *J Biol Chem* 275 (34), 26164-26171
- 132 Somoza, J.R. et al. (2004) Structural snapshots of human HDAC8 provide insights into the class I histone deacetylases. *Structure* 12 (7), 1325-1334
- 133 Byrne, E.F.X. et al. (2016) Structural basis of Smoothed regulation by its extracellular domains. *Nature* 535 (7613), 517-522
- 134 DeVore, N.M. and Scott, E.E. (2012) Structures of cytochrome P450 17A1 with prostate cancer drugs abiraterone and TOK-001. *Nature* 482 (7383), 116-119
- 135 Angers, S. et al. (2006) Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* 443 (7111), 590-593
- 136 Sievers, Q.L. et al. (2018) Defining the human C2H2 zinc finger degrome targeted by thalidomide analogs through CRBN. *Science* 362 (6414)
- 137 Harshbarger, W. et al. (2015) Crystal structure of the human 20S proteasome in complex with carfilzomib. *Structure* 23 (2), 418-424
- 138 Schrader, J. et al. (2016) The inhibition mechanism of human 20S proteasomes enables next-generation inhibitor design. *Science* 353 (6299), 594-598

- 139 Aldaz, H. et al. (2005) Insights into microtubule nucleation from the crystal structure of human gamma-tubulin. *Nature* 435 (7041), 523-527
- 140 Brown, A. et al. (2014) Structure of the large ribosomal subunit from human mitochondria. *Science* 346 (6210), 718-722
- 141 Gurel, G. et al. (2009) U2504 determines the species specificity of the A-site cleft antibiotics: the structures of tiamulin, homoharringtonine, and bruceantin bound to the ribosome. *J Mol Biol* 389 (1), 146-156
- 142 Alexeeva, M. et al. (2015) The structure of a dual-specificity tyrosine phosphorylation-regulated kinase 1A-PKC412 complex reveals disulfide-bridge formation with the anomalous catalytic loop HRD(HCD) cysteine. *Acta Crystallogr D Biol Crystallogr* 71 (Pt 5), 1207-1215

Figure Legends

Figure 1. Hinge-binding inhibitors targeting two cyclin-dependent kinases. (A) CDK4 (purple) bound to cyclin-D1 (green) (PDB ID: 2w96). (B) CDK6 (blue) bound to multiple tumor suppressor (red) (PDB ID: 1bi7). (C) CDK6 (blue) bound to abemaciclib (yellow) (PDB ID: 5l2s). (D) Active site of CDK6 (PDB ID: 5l2s) showing bound abemaciclib (yellow; PDB ID: 5l2s), overlaid with palbociclib (red; PDB ID: 5l2i) and ribociclib (green; PDB ID: 5l2t).

Figure 2. Allosteric inhibitor enasidenib (yellow) targeting homodimeric IDH2 (green and blue) (PDB ID: 5i96).

Figure 3. Hinge-binding inhibitors targeting mutant BRAF. (A) V600E BRAF Kinase (blue) bound to vemurafenib (yellow) (PDB ID: 3og7). (B) Active site of V600E BRAF kinase (PDB ID: 3og7) showing vemurafenib (yellow; PDB ID: 3og7) overlaid with dabrafenib (red; PDB ID: 4xv2).

Figure 4. Immune checkpoint blockade. (A) PD-1 (red) bound to PD-L1 (blue) (PDB ID: 3bik). (B) PD-1 (red) bound to PD-L2 (light blue) (PDB ID: 3bp5). (C) PD-1 (red) recognition by pembrolizumab Fab (green) (PDB ID: 5ggs). (D) PD-1 (red) recognition by nivolumab Fab (green) (PDB ID: 5ggr).

Table Legends

Table 1. Overview of PDB holdings for anti-neoplastic NMEs and their known molecular targets approved 2010-2018.

Table 2. PDB holdings for anti-neoplastic LMW-NMEs approved 2010-2018.

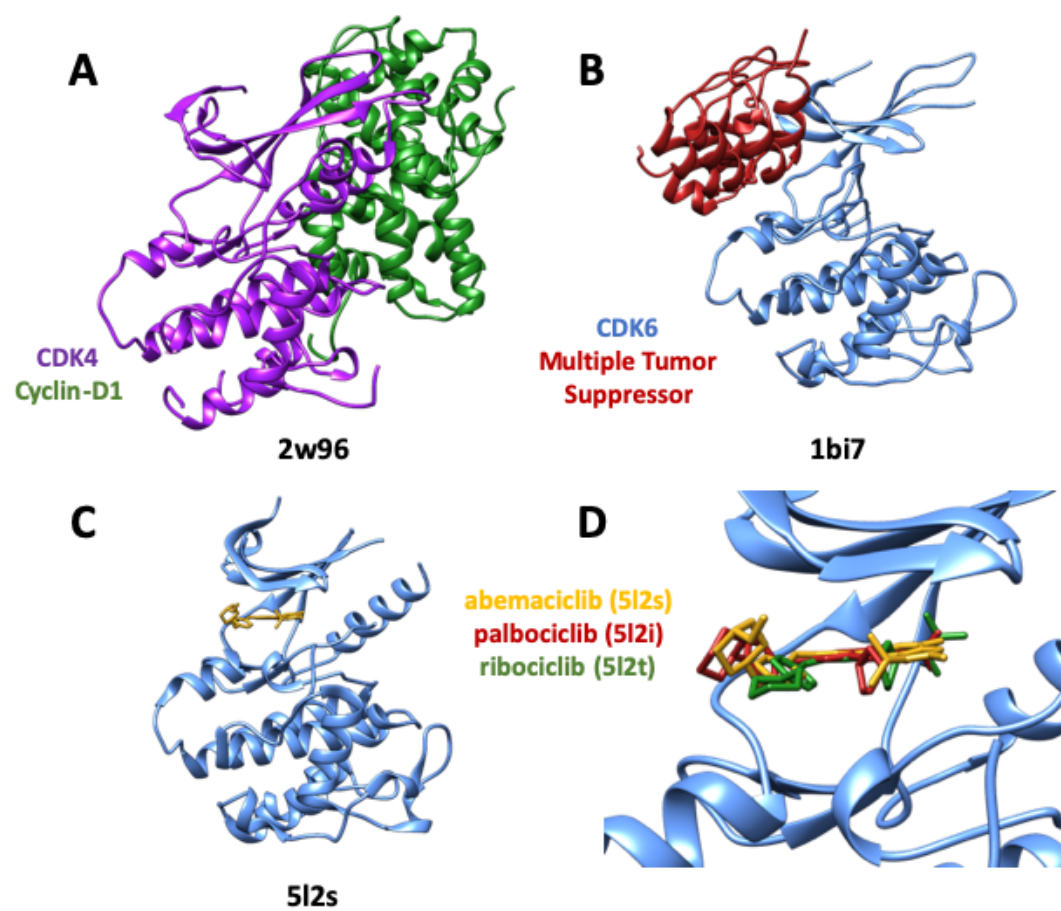
* **Bold** indicates the targets or target classes for which 3D structure information and structure-guided drug discovery facilitated approval of 39/54 (~72%) newly approved LMW-NMEs. Target Name Abbreviations - IDH1: isocitrate dehydrogenase 1; IDH2: isocitrate dehydrogenase 2; BCL-2: B-cell lymphoma 2; PARP: poly ADP-ribose polymerase; HDAC: histone deacetylase; and CYP17A1: cytochrome p450 17A1.

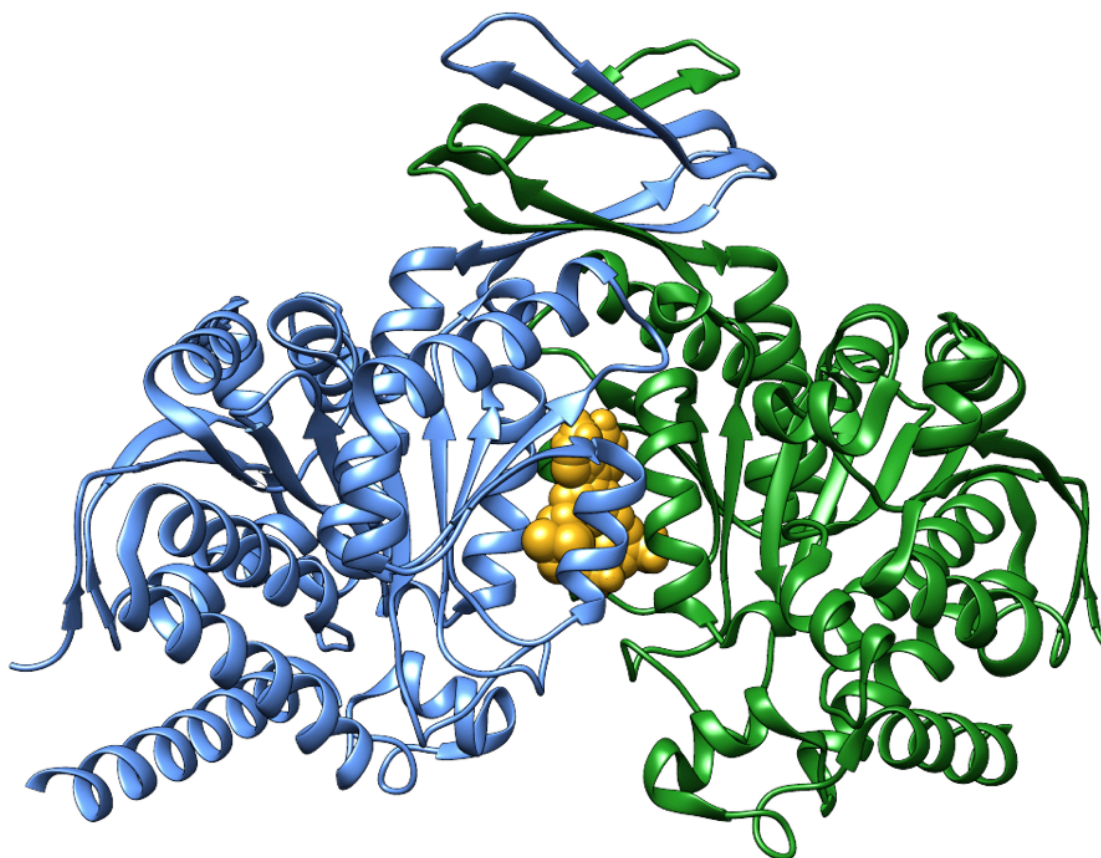
Table 3. Evidence summary for structure-guided drug discovery (SGDD) of LMW-NMEs approved 2010-2018.

** indicates LMW-NMEs featured in the three case studies described at the end of this review. Target Name Abbreviations – VEGFR: vascular endothelial growth factor receptor; EGFR: epidermal growth factor receptor; ALK: anaplastic lymphoma kinase; CDK: cyclin-dependent kinase; PI3K: phosphoinositide 3-kinase; BTK: Bruton's tyrosine kinase; TRK: and Tropomyosin receptor kinase.

Table 4. PDB holdings for anti-neoplastic Biologic-NMEs approved 2010-2018.

*** Other: calaspargase pegol-mknl, asparaginase *Erwinia chrysanthemi*, ziv-aflibercept, tagraxofusp-erzs, lutetium Lu 177 dotatate.





IDH2 (5i96)
enasidenib

Figure 3. Hinge-binding inhibitors targeting mutant BRAF. (A) V600E BRAF Kinase (blue) bound to vemurafenib (yellow) (PDB ID: 3og7). (B) Active site of V600E BRAF kinase (PDB ID: 3og7) showing vemurafenib (yellow; PDB ID: 3og7) overlaid with dabrafenib (red; PDB ID: 4xv2).

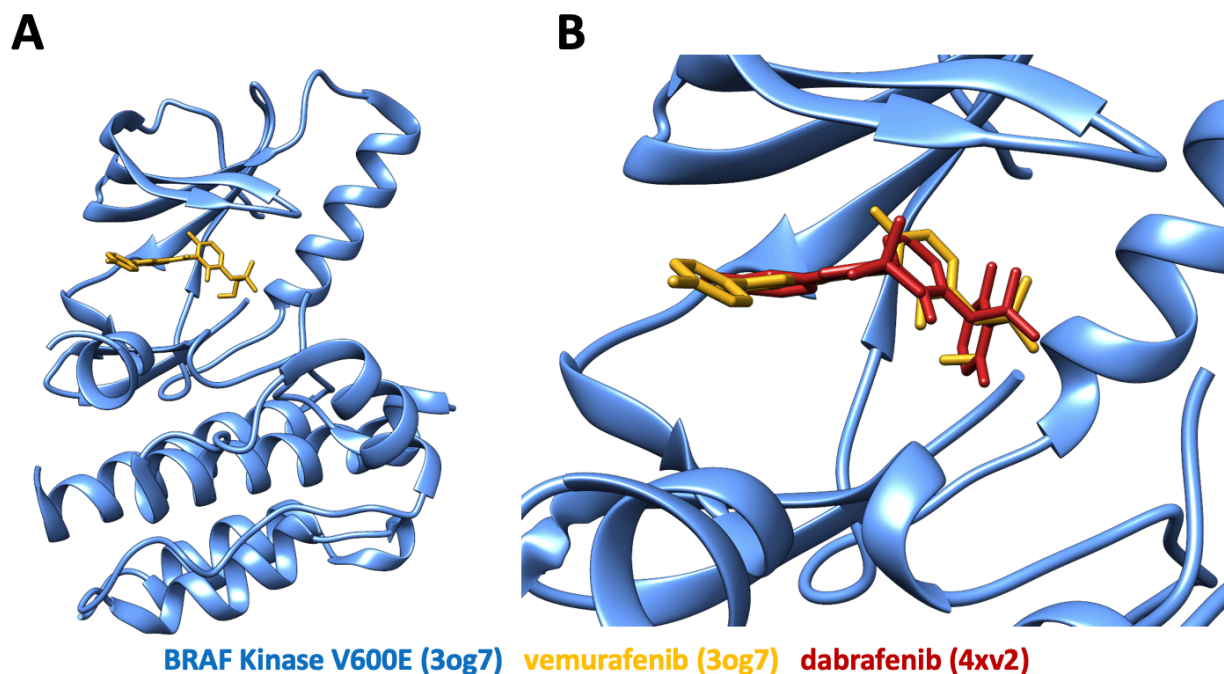
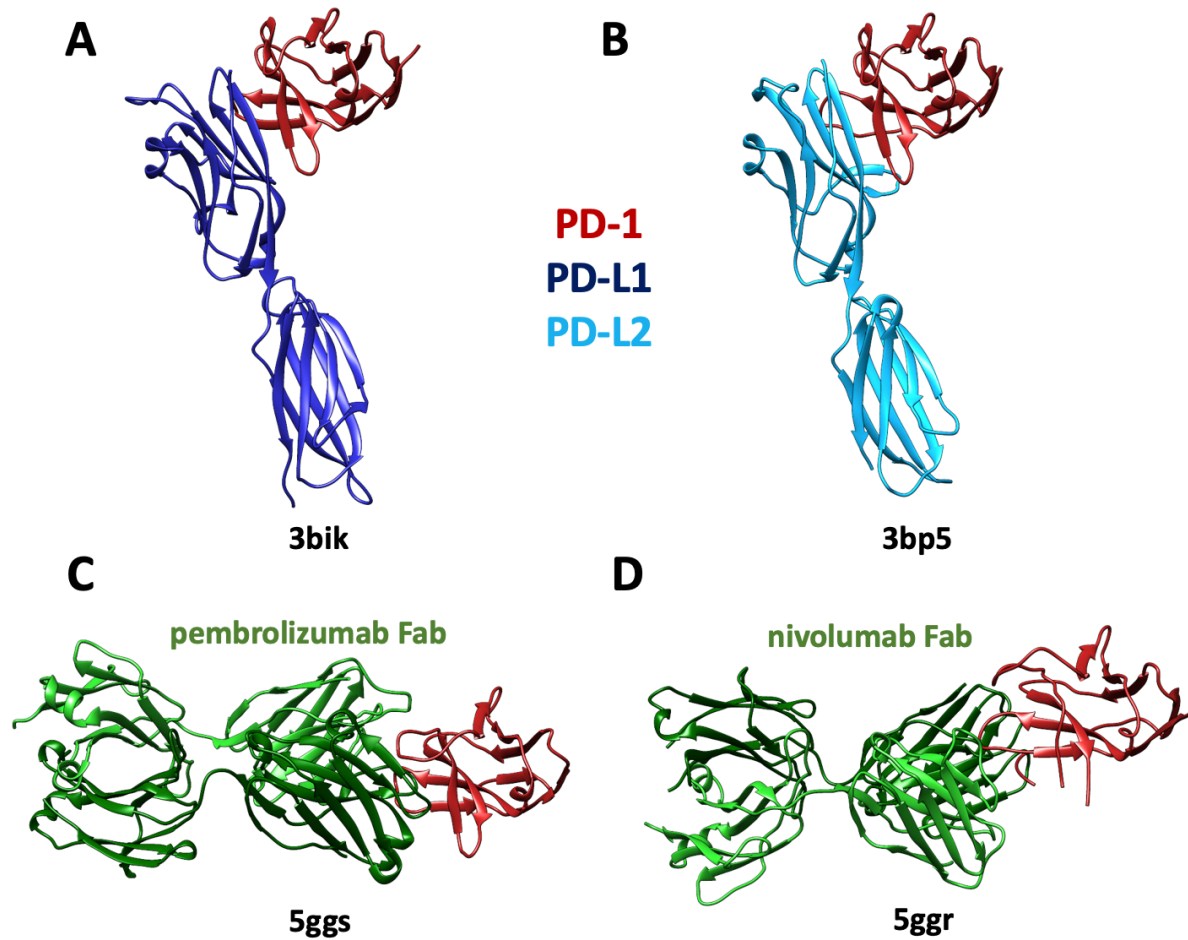
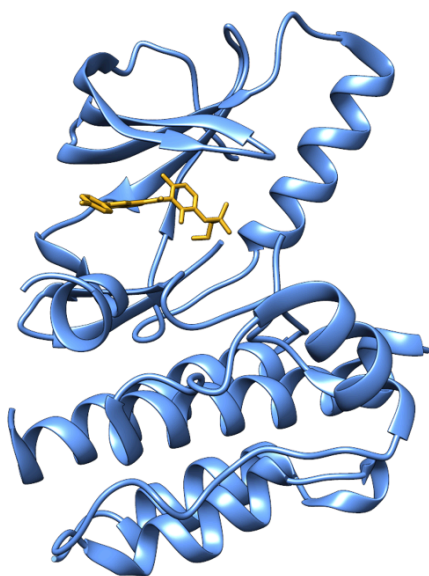
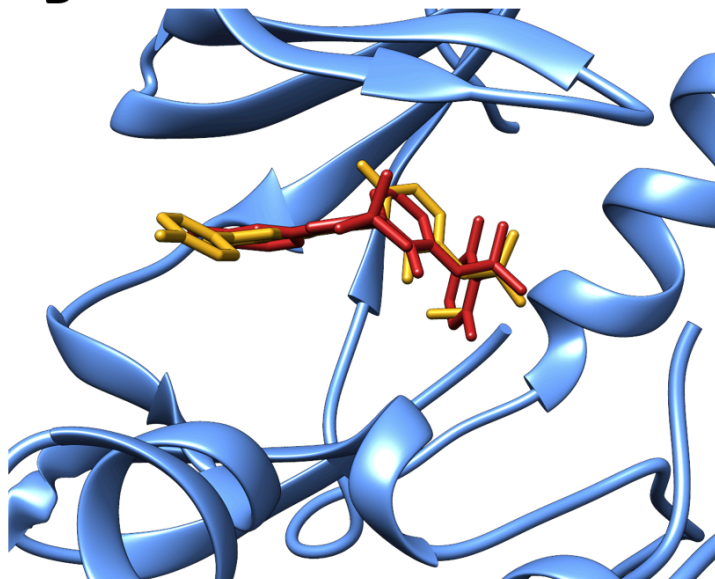


Figure 4. Immune checkpoint blockade. (A) PD-1(red) bound to PD-L1 (blue) (PDB ID: 3bik). (B) PD-1 (red) bound to PD-L2 (light blue) (PDB ID: 3bp5). (C) PD-1 (red) recognition by pembrolizumab Fab (green) (PDB ID: 5ggs). (D) PD-1 (red) recognition by nivolumab Fab (green) (PDB ID: 5ggr).



A**B**

BRAF Kinase V600E (3og7) **vemurafenib (3og7)** **dabrafenib (4xv2)**

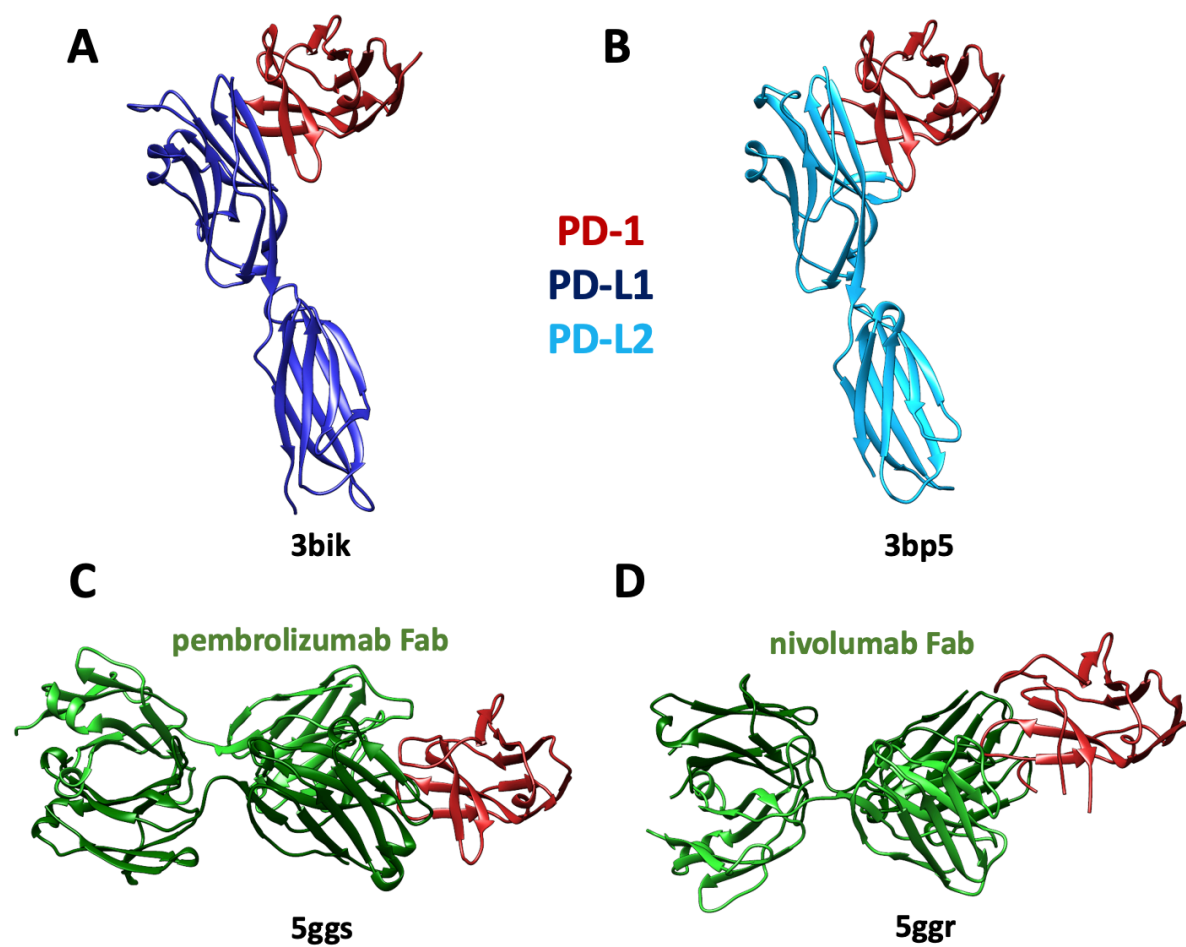


Table 1

NME Class	Number	Number with Target Structure(s) in PDB (95% identity)	Number with NME Structure(s) in PDB	Number with NME-Target Complex Structure(s) in PDB
Anti-neoplastic agents with known molecular targets	79	74 (~94%)	53 (~67%)	47 (~59%)
LMW-NMEs	54	54 (100%)	40 (~74%)	38 (~70%)
Biologic-NMEs	25	20 (80%)	13 (52%)	9 (36%)

Table 2

LMW-NME Target Class	Number in Target Class	Number with target structure(s) in PDB	Total unique PDB IDs for NME target Structures (95% Identity)	Number with target/NME complex structure(s) in PDB
Protein Kinases*	33	33 (100%)	1,136	26 (~76%)
IDH1	1	1 (100%)	40	0 (0%)
IDH2	1	1 (100%)	9	1 (100%)
BCL-2	1	1 (100%)	26	1 (100%)
PARPs	4	4 (100%)	89	4 (100%)
Androgen Receptor	2	2 (100%)	96	0 (0%)
HDACs	2	2 (100%)	92	2 (100%)
Smoothened	3	3 (100%)	11	1 (~33%)
CYP17A1	1	1 (100%)	13	1 (100%)
E3 Ubiquitin Ligase	1	1 (100%)	50	1 (100%)
Proteasome	2	2 (100%)	62	2 (100%)
Tubulin	2	2 (100%)	249	1 (50%)
Ribosome A Site	1	1 (100%)	134	1 (100%)
All	54	54 (100%)	2007	41 (~74%)

Table 3

Generic Drug Name	Sponsor Company	Target Protein Class	Target Protein	Earliest >95% Identical Target (Domain) PDB ID/Year [Literature Citation]	US FDA Approval Year	Target/ LMW-NME Complex PDB ID [Literature Citation]	Source of Target-Drug Complex PDB ID (Academe, Industry)	SGDD (Yes, Prob, Poss, Unl, Nat Prod)
vemurafenib**	Roche	Kinase	BRAF	1uwh/2004 [83]	2011	3og7 [43]	Industry	Yes
dabrafenib**	GSK	Kinase	BRAF	“	2013	4xv2 [34]	Industry	Yes
encorafenib**	Novartis	Kinase	BRAF	“	2018			Yes
vandetanib	AstraZeneca	Kinase	VEGFRs	1vr2/1999 [116]	2011	2ivu [49]	Academe	Prob
axitinib	Pfizer	Kinase	VEGFRs	“	2012	4ag8 [38]	Industry	Yes
lenvatinib	Esai	Kinase	VEGFRs	“	2015	5zv2 [29]	Industry	Yes
bosutinib	Pfizer	Kinase	SRC	1fmk/1997 [117]	2012	4mxo [48]	Academe	Prob
regorafenib	Bayer	Kinase	KIT	1pkg/2003 [118]	2012			Poss
ponatinib	Ariad	Kinase	T315I ABL	1iep/2001 [119]	2012	3ik3 [44]	Academe/ Industry	Yes
neratinib	Puma	Kinase	EGFRs	1m14/2002 [120]	2017	2jiv [121]	Academe	Prob
dacomitinib	Pfizer	Kinase	EGFRs	“	2018	4i23 [37]	Industry	Yes
osimertinib	AstraZeneca	Kinase	EGFRs	“	2015	4zau [47]	Academe	Prob
afatinib	Boehringer Ingelheim	Kinase	EGFRs	“	2013	4g5j [39]	Industry	Yes
crizotinib	Pfizer	Kinase	ALK	2yt2/2007 [122]	2011	2xp2 [41]	Industry	Yes
ceritinib	Novartis	Kinase	ALK	“	2014	4mkc [36]	Academe/ Industry	Yes
alectinib	Roche	Kinase	ALK	“	2015	3aox [42]	Industry	Yes
brigatinib	Ariad	Kinase	ALK	“	2107	6mx8 [33]	Industry	Yes
lorlatinib	Pfizer	Kinase	ALK	“	2018	4cli [46]	Industry	Yes
palbociclib	Pfizer	Kinase	CDK4/6	1bi8/1998 [76]	2015	5l2i [31]	Industry	Yes
ribociclib	Novartis	Kinase	CDK4/6	“	2017	5lt2 [31]	Industry	Yes
abemaciclib	Lilly	Kinase	CDK4/6	“	2017	5l2s [31]	Industry	Yes
cobimetinib	Exelixis	Kinase	MEK	1s9j/2004 [90]	2015	4an2 [40]	Industry	Yes
binimetinib	Array Biopharma	Kinase	MEK	“	2018			Yes
trametinib**	JapanTobacco	Kinase	MEK	“	2013			Poss
cabozantinib	Bristol Myers Squibb	Kinase	MET	1r0p/2003 [123]	2016	3lq8 [45]	Industry	Yes
idelalisib	Gilead	Kinase	PI3Ks	2rd0/2007 [124]	2014	4xe0 [35]	Industry	Yes
copanlisib	Bayer	Kinase	PI3Ks	“	2017	5g2n [32]	Industry	Yes
duvelisib	Intellikine	Kinase	PI3Ks	“	2018			Poss
ibrutinib	Celera	Kinase	BTK	1btk/1997 [125]	2013	5p9i [30]	Industry	Yes
acalabrutinib	Acerta	Kinase	BTK	“	2017			Yes

larotrectinib	Array Biopharma	Kinase	TRKs	5jfw/2016 [126]	2018			Poss
gilteritinib	Astellas	Kinase	FLT3	1rjb/2004 [127]	2018	6jqr [28]	Industry	Yes
ivosidenib	Agios	Enzyme	IDH1	1t09/2004 [128]	2018			Yes
enasidenib**	Agios	Enzyme	IDH2	1lwd/2002 [77]	2017	5i96 [50]	Industry	Yes
venetoclax	Abbott	Programmed Cell Death	BCL-2	1g5m/2000 [129]	2016	6o0k [51]	Industry	Yes
olaparib	AstraZeneca	Enzyme	PARPs	1uk0/2004 [130]	2014	4tvj [52]	Academe	Prob
rucaparib	Clovis Oncology	Enzyme	PARPs	“	2016	4rv6 [52]	Academe	Prob
niraparib	GSK	Enzyme	PARPs	“	2017	4r6e [52]	Academe	Prob
talazoparib	Pfizer	Enzyme	PARPs	“	2018	4und [52]	Academe	Prob
enzalutamide	Medivation	Nuclear Hormone Receptor	Androgen Receptor	1e3g/2000 [131]	2012			Prob
apalutamide	Janssen	Nuclear Hormone Receptor	Androgen Receptor	“	2018			Prob
belinostat	Spectrum Pharmaceuticals	Epigenetic	HDACs	1t64/2004 [132]	2014	5een [53]	Academe	Prob
panobinostat	Novartis	Epigenetic	HDACs	“	2015	5ef8 [53]	Academe	Prob
vismodegib	Roche	GPCR	Smoothened	4jkv/2013 [132]	2012	5l7i [133]	Academe	Unl
sonidegib	Sun Pharma	GPCR	Smoothened	“	2015			Unl
glasdegib	Pfizer	GPCR	Smoothened	“	2018			Unl
abiraterone acetate	Apotex	Cytochrome P450	CYP17A1	3ruk/2012 [134]	2011	3ruk [134]	Academe	Unl
pomalidomide	Celgene	Protein Degradation	E3 Ubiquitin Ligase	2hye/2006 [135]	2013	6h0f [136]	Academe	Unl
ixazomib citrate	Millennium	Protein Degradation	Proteasome	4r3o/2015 [137]	2015	5lf7 [138]	Academe	Unl
carfilzomib	Amgen	Protein Degradation	Proteasome	“	2012	4r67 [137]	Academe	Nat Prod
cabazitaxel	Sanofi-Aventis	Cell Division	Tubulin	1z5v/2005 [139]	2010			Nat Prod
eribulin	Eisai	Cell Division	Tubulin	“	2010			Nat Prod
omacetaxine mepesuccinate	Teva	Ribosome	A site	3j7y/2014 [140]	2012	3g6e [141]	Academe	Nat Prod
midostaurin	Millennium	Kinase	Multiple Kinases	N/A	2017	4nct [142]	Academe	Nat Prod

Table 4

Types of Biologic-NMEs	Number of Biologic-NMEs of each type	Number of Biologic-NMEs with PDB target structure(s)	Unique PDB IDs for Biologic-NME target structures (95% Identity)	Biologic-NME structure(s) in PDB
Antibodies	16	14 (~88%)	395	8 (50%)
ADCs	4	4 (100%)	30	1 (25%)
Other**	5	2 (40%)	48	4 (80%)
All	25	20 (80%)	405	13 (52%)