Chapter 8
The Protein Data Bank: Overview and Tools for Drug Discovery

Helen M. Berman, Peter W. Rose, Shuchismita Dutta, Christine Zardecki, and Andreas Prlić

Abstract The increasing size and complexity of the three dimensional (3D) structures of biomacromolecules in the Protein Data Bank (PDB) is a reflection of the growth in the field of structural biology. Although the PDB archive was initially used only in the field of structural biology, it has grown to become a valuable resource for understanding biology at a molecular level and is critical for designing new therapeutic options for various diseases. The many uses of the PDB archive depend upon on the tools and resources for both data management and for data access and analysis.

8.1 Introduction

The field of structural biology began in the late 1950s as scientists started to decipher the three dimensional (3D) structures of proteins. Structure determination of myoglobin [1, 2] followed closely by that of hemoglobin [3, 4] earned Perutz and Kendrew Nobel prizes in 1962. Soon members of the scientific community recognized how strong research advances could be made through a shared, public archive of data from these experiments [5, 6]. In 1971, following a meeting at Cold Spring Harbor, the Protein Data Bank (PDB) was established with seven structures [7].

H.M. Berman (✉) • S. Dutta • C. Zardecki
RCSB Protein Data Bank, Department of Chemistry and Chemical Biology and Center for Integrative Proteomics Research, Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8087, USA
e-mail: berman@rcsb.rutgers.edu

P.W. Rose • A. Prlić
RCSB Protein Data Bank, San Diego Supercomputer Center, University of California San Diego, La Jolla, CA 92093-0743, USA

Today, the PDB archive contains more than 100,000 structures and is managed by the Worldwide Protein Data Bank (wwPDB, wwpdb.org), a consortium of groups that host deposition, annotation, and distribution centers for PDB data and collaborate on a variety of projects and outreach efforts [8, 9]. While the PDB data is available as a single archive, wwPDB data centers present unique tools, resources and views of the data to facilitate scientific inquiry and analysis.

8.2 Overview

8.2.1 PDB Data

The primary data archived in the PDB are the 3D atomic coordinates of biological molecules determined using experimental methods such as X-ray crystallography, Nuclear Magnetic Resonance (NMR) and Electron Microscopy (3D EM). In addition to coordinate data PDB also archives several descriptive metadata items such as the primary citation, polymer sequence, chemical information about the ligands and macromolecules, some experimental details, and structural descriptors. Experimental data used to derive these structures (e.g. structure factors, restraints and chemical shifts) are made available, along with 3DEM map data [10].

All information regarding a particular structure is linked to an identifier (PDB ID). The original file format used to represent PDB was established 40 years ago and has very recently been replaced by PDBx/mmCIF. This newer format is computer readable and unlike the older format can accommodate large complex structural data. The PDBx/mmCIF Data Exchange Dictionary [11] consolidates content from a variety of crystallographic data dictionaries and includes extensions describing NMR, 3DEM, and protein production data. Internal data processing, annotation, and database management operations rely on the PDBx/mmCIF dictionary content and corresponding file format. As the PDBx/mmCIF file format is very extensible, it can expand and grow to support new types of information. Recently, the developers of X-ray structure determination packages have adapted PDBx/mmCIF as their standard format.

8.2.2 Data Deposition and Annotation

Once a structure has been determined, it is deposited into the PDB for processing and annotation by the wwPDB. Until recently, multiple different systems for deposition and annotation made data uniformity and exchange difficult. In the new wwPDB Common Deposition & Annotation (D&A) system, launched in 2014, data are easily transferred and shared. In addition, many aspects of the deposition and annotation practices have been improved enabling efficiency and accuracy (Fig. 8.1).
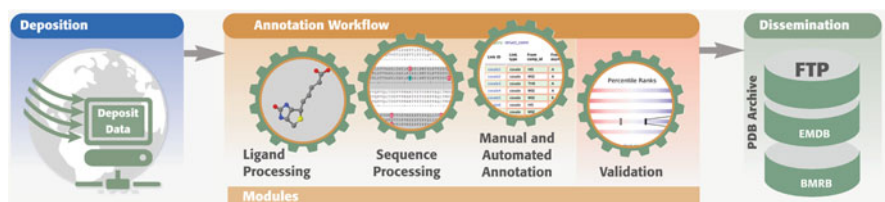
Fig. 8.1 wwPDB Common Deposition and Annotation System for PDB, EMDB and BMRB data. In this pipeline, data are submitted to the PDB using a single interface and then processed and annotated by the wwPDB using a series of focused modules [16, 32]. Data are released into the PDB FTP archive at ftp://ftp.wwpdb.org on a weekly basis

Highly qualified biocurators in the wwPDB data processing centers annotate each PDB entry to ensure accurate representation of both the structure and experiment. They review polymer sequences, small molecule chemistry, cross references to other databases, experimental details, correspondence of coordinates with primary data, protein conformation, biological assemblies, and crystal packing. During the annotation process, the wwPDB biocurators communicate with the entry authors (depositors) to make sure the data are represented in the best way possible.

To help ensure the accuracy of PDB entries, deposited data are compared with community-accepted standards during the process of validation. Method-specific Validation Task Forces (VTF) comprising of experts in X-ray Crystallography [12], NMR [13], 3DEM [14], and Small Angle Scattering [15] were convened by the wwPDB to develop consensus on validation that should be performed, and to identify software applications for validation. The VTF recommendations are now implemented in the wwPDB data processing procedures and suitable tools have been developed as part of the wwPDB Common Deposition & Annotation System.

Depositors are provided with detailed reports that include the results of data consistency, geometric and experimental data validation [16]. These reports, available as PDFs, provide an assessment of structure quality while maintaining the confidentiality of the coordinate data. Graphical depictions allow facile assessments of the overall quality as well as sequence specific features (Fig. 8.2). Currently, these wwPDB validation reports are required by several journals for manuscript review, including *eLife*, *The Journal of Biological Chemistry*, and the journals of the International Union of Crystallography. The wwPDB encourages all journal editors and referees to incorporate these reports in the manuscript submission and review process.

8.2.3   Data Distribution

The PDB archive (ftp://ftp.wwpdb.org) is updated weekly. Contents of the ftp site include experimentally determined coordinate data files, related experimental data (structure factors, constraints, and chemical shifts) and 3DEM map data. The
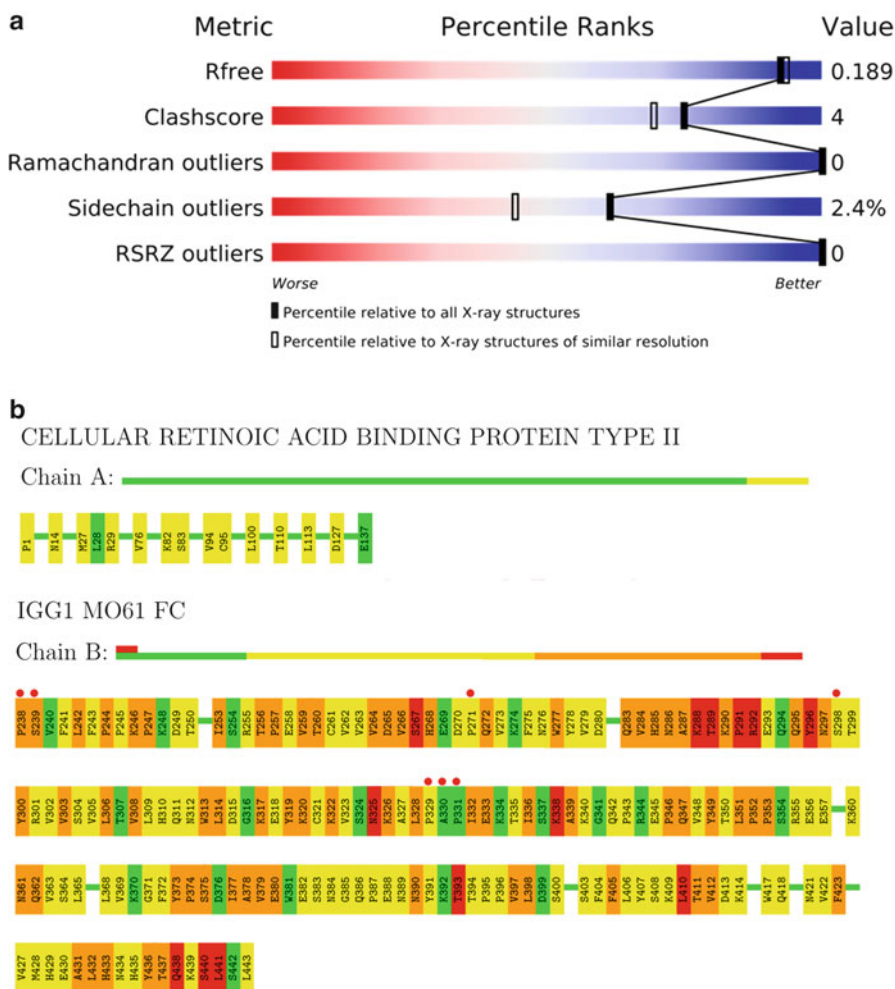
Fig. 8.2 Graphics included in the Validation Reports produced by the wwPDB. These reports, made available as PDFs, provide an assessment of structure quality while maintaining the confidentiality of the coordinate data. (**a**) The "slider" graphic gives an indication of the quality of the determined structure as compared with previously deposited PDB entries using several important global quality indicators. (**b**) Residue-property plots indicate quality information for proteins and nucleic acids on a per-residue basis. Two images are displayed for each molecule. In the *top image*, the *green*, *yellow*, *orange* and *red* segments indicate the fraction of residues with 0, 1, 2 and 3 or more types of model-only quality criteria with outliers. In the *bottom image*, the *red circle* (if present) indicates the fraction of residues that have an unusual fit to the density (RSRZ outliers) (Color figure online)

ftp site also contains the data dictionaries and external reference files (ERFs) used to describe PDB data, including the PDBx/mmCIF dictionary, the Chemical Component Dictionary (CCD) that contains detailed chemical descriptions for

standard and modified amino acids/nucleotides, small molecule ligands, and solvent molecules, and the Biologically Interesting Molecule Reference Dictionary (BIRD) that contains information about biologically interesting peptide-like antibiotic and inhibitor molecules in the PDB archive [17].

Each wwPDB member organization maintains websites with different views of the data and different services. These websites are RCSB PDB (US) at rcsb.org [18], Protein Data Bank in Europe (PDBe, United Kingdom) at pdbe.org [19], Protein Data Bank Japan (PDBj) at pdbj.org [20], and the BioMagResBank (BMRB, US) at bmrb.wisc.edu [21].

8.2.4   Growth of the PDB Archive

The number of structures contained in the archive has grown over the past ∼40 years since the creation of the PDB. In addition to structures determined by X-ray crystallography, the archive includes structures determined using NMR spectroscopy and 3D electron microscopy (3D EM) (Fig. 8.3 a–c). It is worth noting that the growth in the number of cryoEM maps is an indicator of the expected high growth rate of cryoEM-derived models that are being deposited into the PDB.

In addition, the complexity of structures deposited has increased as evidenced by growth in the number of polymers chains within each structure and the molecular weight (Fig. 8.4a, b). By reviewing the content within the PDB it is possible to see the evolution of types of methods used to determine structures. Whereas in the 1970s only relatively small structures could be studied, now we have many examples of macromolecular machines [22, 23] (Fig. 8.5). Most recently, structures have been determined using several different methods. These hybrid models are the subject of much discussion as to how to best evaluate and archive them.

8.3   RCSB PDB Resources for Drug Discovery

In addition to biological macromolecules (proteins and nucleic acids), ∼73 % of PDB entries include one or more ligands. Some of these ligands are simple, such as ions, cofactors, inhibitors, and drugs [22]. More than 1,000 PDB structures contain peptide-like inhibitors and antibiotics [17]. These ligand-bound complexes highlight the overall shapes and key functional regions of the relevant biological molecules and lay the foundations for designing molecules that can alter the function. In the 1980s when Acquired Immunodeficiency Syndrome (AIDS) was rapidly spreading through the world, structural studies of Human Immunodeficiency Virus (HIV) proteins were critical in designing specific inhibitors that have led to the development of clinically important drugs for treating HIV infection [24–26]. Similarly, there have been many studies of antibiotics that target the ribosome [27, 28].
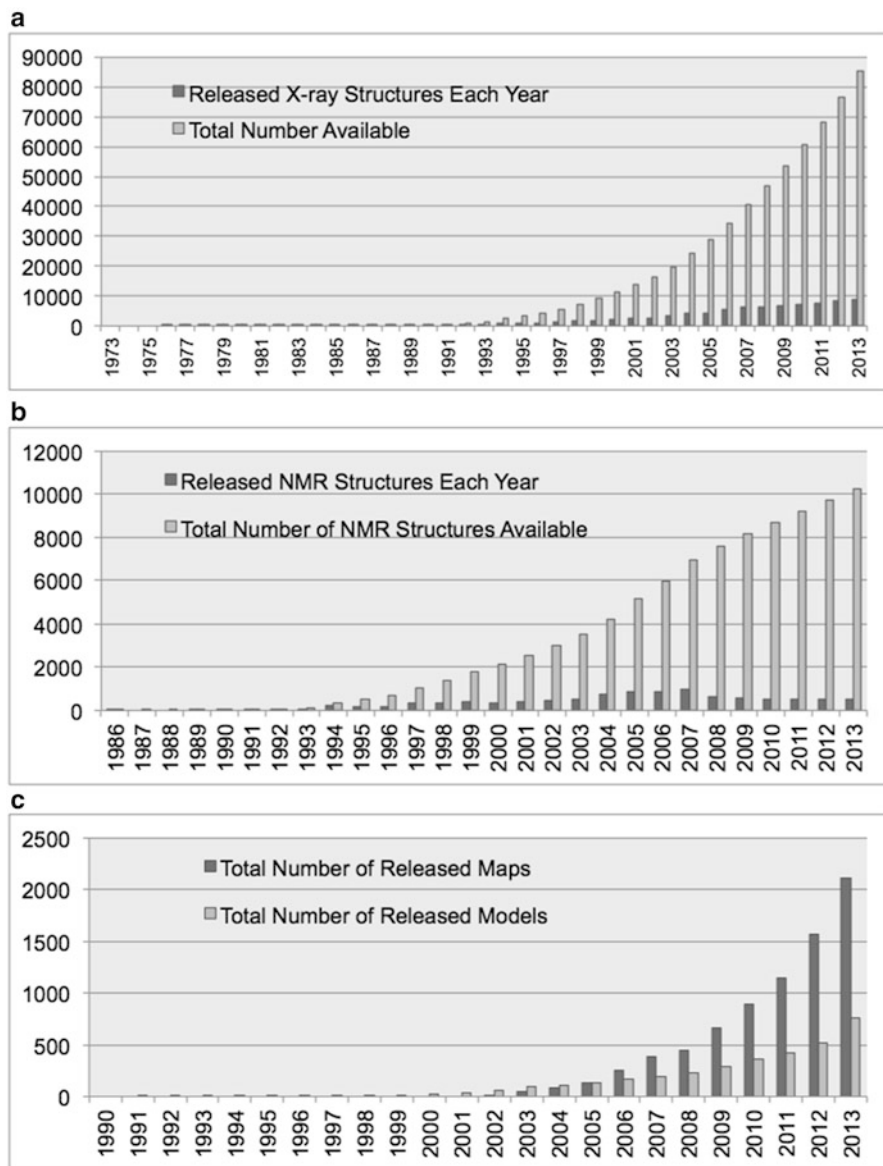
a

b

c

Fig. 8.3 Growth of the number of structures available in the PDB archive by experimental method: (**a**) X-ray crystallography, (**b**) NMR, (**c**) 3DEM

While biological polymers (proteins and nucleic acids) can be queried in the PDB by protein or gene name or its sequence, the RCSB Protein Data Bank website provides a number of resources that facilitate drug discovery-related research [29]. The following sections provide a brief description of these tools.
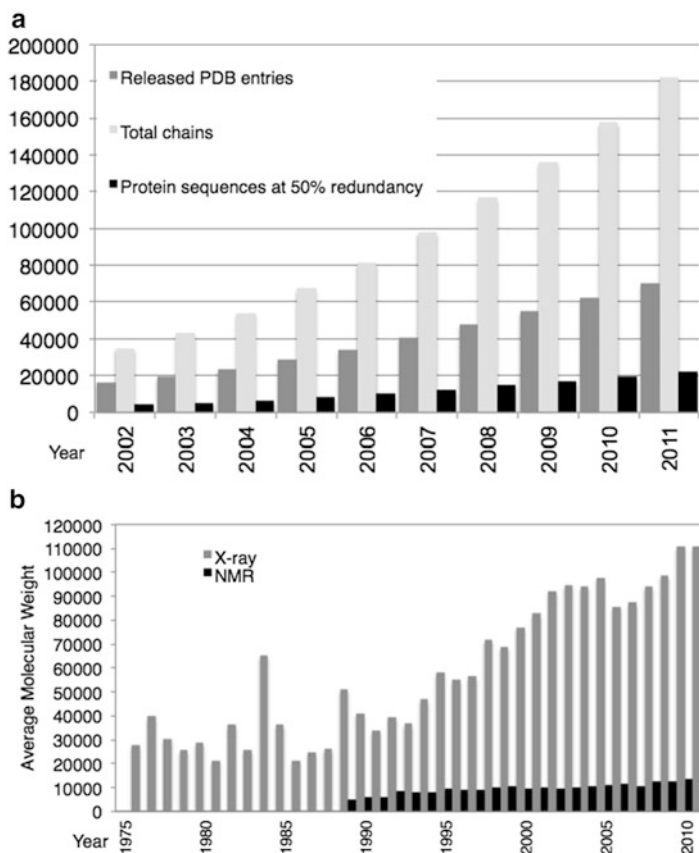
Fig. 8.4 Growth of the size and complexity of the structures available in the PDB archive. (**a**) the number of PDB entries, total related polymer chains, and protein sequences (with 50 % redundancy as calculated using blastclust [33]) available in the archive each year; (**b**) Average molecular weight of entries released each year for structures determined by X-ray crystallography (for the asymmetric unit; in *grey*) and NMR (in *black*). Calculations excluded water and counted extremely large structures as single entries. For viruses and entries that used non-crystallographic symmetry (NCS), molecular weights for the full asymmetric unit were calculated by multiplying the molecular weight of the explicit polymer chains by the number of NCS operators. The large increase shown in 1984 was due to the release of the tomato bushy stunt virus 2tbv [34] (Figures reprinted from [22])

8.3.1   Ligand Search

The most common uses of the RCSB PDB website are simple searches using the top search box on the RCSB PDB website. An autocomplete feature is available that can help guide the user to specific matches in the archive and provide relevant results. After typing a few letters in the top search bar, a suggestion box opens and organizes result sets in different categories. Each suggestion includes the number of
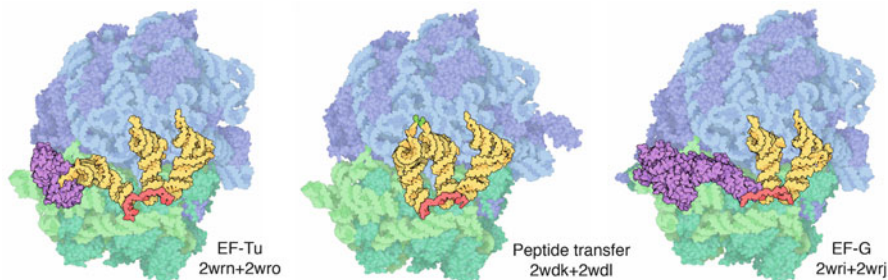
Fig. 8.5 Example of a macromolecular machine: ribosome complexes (PDB IDs 2wrn, 2wro [35], 2wdk, 2wdl [36], 2wri, 2wrj [37]. Atomic structures have been determined for ribosomes engaged in most aspects of mRNA translation. These three structures capture the ribosome in distinct phases of elongation: *left*, binding of a new tRNA assisted by elongation factor Tu; *middle*, the peptide transfer reaction; and *right*, stepping to the next reading frame by binding of elongation factor G. Image from the RCSB PDB *Molecule of the Month* feature on the Ribosome (doi: 10.2210/rcsb_pdb/mom_2010_1) and reprinted from [23]

results and links to the set of matching structures. For example, by entering the drug brand name "Glivec" or the generic name "Imatinib" the autosuggestion provides a link to the corresponding Ligand Summary Page described below.

Ligand searches by ID, name, synonym, formula, and SMILES string are possible using the top query bar. These queries are also available from the Advanced Search menu and include searching by Chemical Component identifier of the ligand, SMILES strings, chemical formula, and by chemical structure (including exact, substructure, superstructure, and similarity searches). Detailed information about ligands and drug molecules bound to macromolecules are available from the **Ligand Summary** and **Structure Summary** pages.
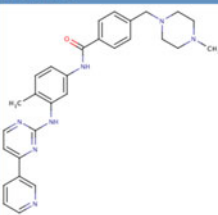
8.3.2 Ligand Summary Page

Information about the chemistry and structure of all small molecule components found in the PDB is contained in the Chemical Component Dictionary (CCD). The Ligand Summary pages present a report from the CCD are organized into widgets or boxes highlighting different types of hyperlinked information (Fig. 8.6). These widgets provide an overview of the ligand, with links to PDB entries where the component appears as a non-polymer or as a non-standard component of a polymer, links to ligand summary pages for similar ligands and stereoisomers, 2D and 3D visualization, and links to many external resources. Original data provided by the RCSB PDB are listed in blue widgets, whereas data from third parties are displayed in orange widgets.

Fig. 8.6 Ligand Summary Page (*top section*) for Imatinib (Glivec). RCSB PDB's Ligand Summary Pages provide information for all of the entries found in the wwPDB's Chemical Component Dictionary. Similar to Structure Summary pages for PDB entries, Ligand Summary Pages are organized into widgets that highlight different types of information, including a Chemical Component Summary that includes name, identifiers, synonyms, and SMILES and InChI information; links to related PDB structures where the ligand appears as a free ligand; links to other Summary Pages for similar ligands and stereoisomers, and links to information about the chemical component at external resources. These summaries can be accessed by performing a ligand search, selecting a ligand from a PDB entry's Structure Summary page, and from the Ligand Hits tab for query results. In the example shown, Glivec is present in 16 PDB entries as a co-crystal structure. Drug annotation is provided by DrugBank [31]

8.3.3 Ligand Summary Reports

For queries that return a set of ligands, the results can be saved as **Ligand Summary Reports** in form of a comma separated value (CSV) file or an Excel spreadsheet. These reports include information about the ligands, such as formula, molecular weight, name, SMILES string, and lists of PDB entries that include the ligand. The report can be expanded to show a sub-table of all PDB entries that contain the ligand as a free ligand and those that contain the ligand as part of a polymer.

8.3.4   Structure Summary Page

Structure Summary pages provide details about specific structure entries in the PDB. It describes all polymers and ligands included in the entry, some details about the experiment, links to the primary citation, and presents resources to interactively visualize the entry. Special support is also offered for the analysis of ligands associated with PDB entries. Any ligands included in a PDB entry are listed in the **Ligand Chemical Component** widget of the entry's **Structure Summary** page. This area displays a 2D chemical structure image, name and formula of each ligand, link to the **Ligand Summary** page, and provides access to 2D and 3D binding site visualization.

8.3.5   Binding Site Visualization

In order to understand the neighborhood of the ligand in the PDB entry and its interactions, 2D interaction diagrams are generated by PoseView [30] and show which atoms or areas of the ligand and the polymer interact with each other, as well as the type of interaction (Fig. 8.7). Interactions are determined by geometric criteria.

Ligand Explorer is a 3D viewer that visualizes the interactions of bound ligands in protein and nucleic acids structures (Fig. 8.8). It has options to turn on the display of interactions including hydrogen bonds, hydrophobic contacts, water mediated hydrogen bonds, and metal interactions. Several types of binding site surfaces can be generated including opaque and transparent solid surfaces, meshes, and dotted surfaces, color coded by hydrophobicity or chain identifier.

8.3.6   Drug and Drug Target Mapping

A detailed mapping of drugs by chemical structure and drug targets by protein sequence is available from the **Drug and Drug Target Mapping** page, which is accessible from the **Search** menu on the RCSB PDB website. Two tables provide access to information about drugs and drug targets from DrugBank [31] that are mapped to PDB entries with each weekly update.

- Drugs Bound to Primary Targets: Lists drugs bound to primary target(s), or a homolog of primary target(s), i.e., co-crystal structures of drugs.
- Primary Drug Targets: Lists primary drug targets in the PDB, regardless if the drug molecule is part of the PDB entry (e.g., apo forms of drug targets, drug target with different bound ligands). Biotherapeutics, such as complexes with monoclonal antibodies, are included.
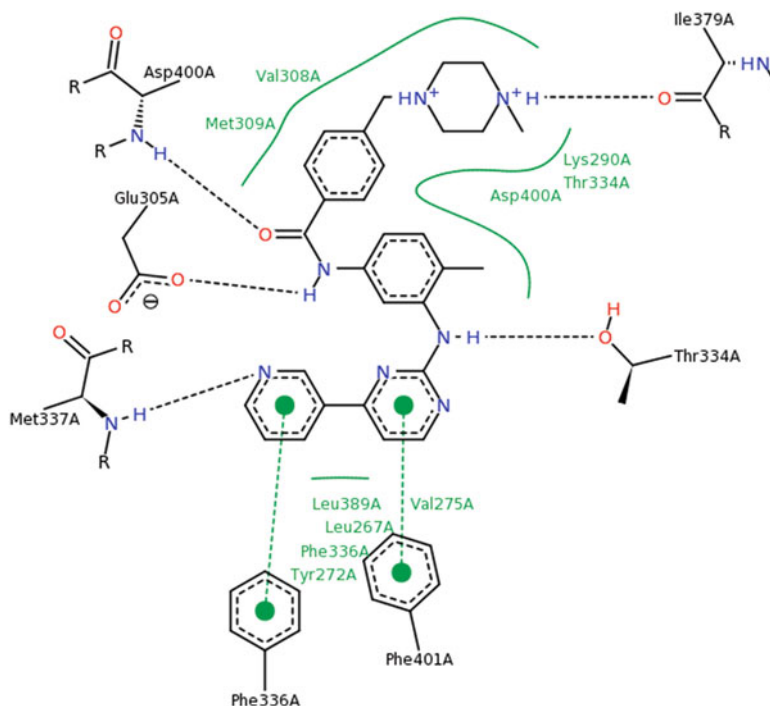
Fig. 8.7 2D macromolecule-ligand interaction diagram of Imatinib (Glivec) bound to Proto-oncogene tyrosine-protein kinase ABL1 (PDB Id: 1OPJ, [38]) generated by PoseView (*black dashed lines*: hydrogen bonds and salt bridges, *green solid lines*: hydrophobic interactions, *green dashed lines*: Pi-Pi interactions) (Color figure online)

These tables can be searched, filtered, sorted, and downloaded as Excel Spreadsheets.

8.4 Summary

The PDB was established in 1971 to archive the experimentally determined 3D structures of biological macromolecules. Today, the archive contains the atomic coordinates and experimental data for more than 100,000 proteins, nucleic acids, and large macromolecular machines. Under the management of the wwPDB collaboration, a new data deposition and annotation tool has been developed to efficiently receive and carefully annotate PDB depositions before public release in the archive.

The RCSB PDB website offers a number of different resources to search, visualize, compare and analyze PDB data. Many of these tools are focused on the study of drug complexes available in the archive.
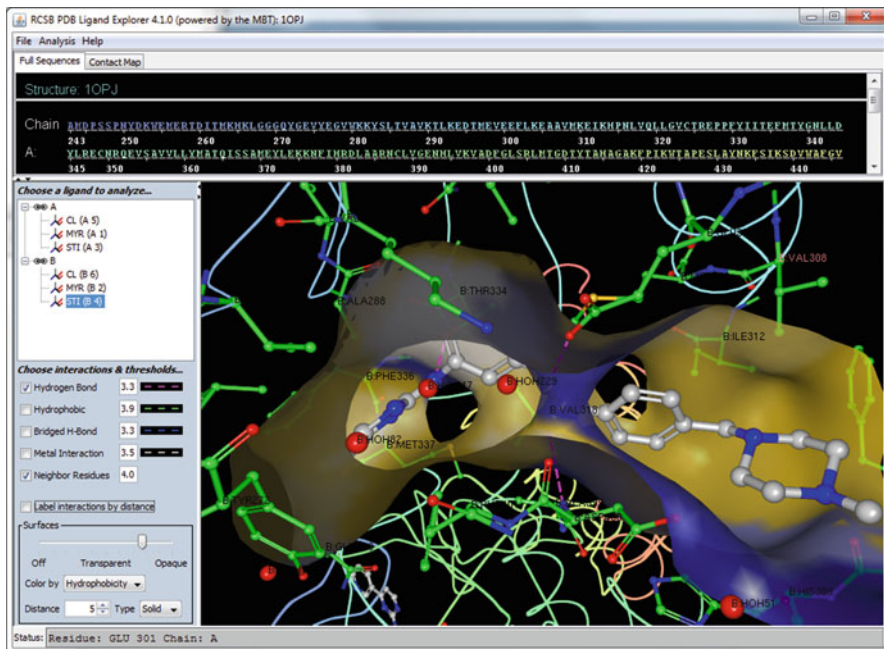
Fig. 8.8 Ligand Explorer 3D view of Imatinib (Glivec) bound to Proto-oncogene tyrosine-protein kinase ABL1 (PDB ID: 1OPJ [38]). The binding pocket is delineated by a surface color-coded by hydrophobicity of the binding site residues (*yellow*: hydrophobic, *blue*: hydrophilic). The vertical cross-section looking into the binding site is transparent and shows the residues linking the drug-binding pocket (Color figure online)

References

1. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC (1958) A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature 181: 662–666
2. Kendrew JC, Dickerson RE, Strandberg BE, Hart RG, Davies DR, Phillips DC, Shore VC (1960) Structure of myoglobin: a three-dimensional Fourier synthesis at 2 A. resolution. Nature 185(4711):422–427
3. Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G, North ACT (1960) Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5 Å resolution, obtained by X-ray analysis. Nature 185:416–422
4. Bolton W, Perutz MF (1970) Three dimensional fourier synthesis of horse deoxyhaemoglobin at 2.8 Ångstrom units resolution. Nature 228(271):551–552

5. Berman HM, Kleywegt GJ, Nakamura H, Markley JL (2013) How community has shaped the Protein Data Bank. Structure 21(9):1485–1491
6. Berman H (2008) The Protein Data Bank: a historical perspective. Acta Crystallogr A: Found Crystallogr 64:88–95
7. Protein Data Bank (1971) Protein Data Bank. Nat New Biol 233:223
8. Berman HM, Henrick K, Nakamura H (2003) Announcing the worldwide Protein Data Bank. Nat Struct Biol 10(12):980
9. Berman HM, Henrick K, Kleywegt G, Nakamura H, Markley J (2012) The worldwide Protein Data Bank. In: Arnold E, Himmel DM, Rossmann MG (eds) International tables for X-ray crystallography, vol F, Crystallography of biological macromolecules. Springer, Dordrecht, pp 827–832
10. Lawson CL, Baker ML, Best C, Bi C, Dougherty M, Feng P, van Ginkel G, Devkota B, Lagerstedt I, Ludtke SJ, Newman RH, Oldfield TJ, Rees I, Sahni G, Sala R, Velankar S, Warren J, Westbrook JD, Henrick K, Kleywegt GJ, Berman HM, Chiu W (2011) EMDataBank.org: unified data resource for CryoEM. Nucleic Acids Res 39:D456–D464
11. Westbrook J, Henrick K, Ulrich EL, Berman HM (2005) 3.6.2 The Protein Data Bank exchange data dictionary. In: Hall SR, McMahon B (eds) International tables for crystallography, vol G. Definition and exchange of crystallographic data. Springer, Dordrecht, pp 195–198
12. Read RJ, Adams PD, Arendall WB III, Brunger AT, Emsley P, Joosten RP, Kleywegt GJ, Krissinel EB, Lutteke T, Otwinowski Z, Perrakis A, Richardson JS, Sheffler WH, Smith JL, Tickle IJ, Vriend G, Zwart PH (2011) A new generation of crystallographic validation tools for the Protein Data Bank. Structure 19(10):1395–1412
13. Montelione GT, Nilges M, Bax A, Güntert P, Herrmann T, Markley JL, Richardson J, Schwieters C, Vuister GW, Vranken W, Wishart D (2013) Recommendations of the wwPDB NMR structure validation task force. Structure 21:1563–1570
14. Henderson R, Sali A, Baker ML, Carragher B, Devkota B, Downing KH, Egelman EH, Feng Z, Frank J, Grigorieff N, Jiang W, Ludtke SJ, Medalia O, Penczek PA, Rosenthal PB, Rossmann MG, Schmid MF, Schroder GF, Steven AC, Stokes DL, Westbrook JD, Wriggers W, Yang H, Young J, Berman HM, Chiu W, Kleywegt GJ, Lawson CL (2012) Outcome of the first electron microscopy validation task force meeting. Structure 20(2):205–214
15. Trewhella J, Hendrickson WA, Sato M, Schwede T, Svergun D, Tainer JA, Westbrook J, Kleywegt GJ, Berman HM (2013) Meeting report of the wwPDB small-angle scattering task force: data requirements for biomolecular modeling and the PDB. Structure 21:875–881
16. Gore S, Velankar S, Kleywegt GJ (2012) Implementing an X-ray validation pipeline for the Protein Data Bank. Acta Crystallogr D68:478–483
17. Dutta S, Dimitropoulos D, Feng Z, Periskova I, Sen S, Shao C, Westbrook J, Young J, Zhuravleva M, Kleywegt G, Berman H (2013) Improving the representation of peptide-like inhibitor and antibiotic molecules in the Protein Data Bank. Biopolymers 101(6):659–668
18. Berman HM, Westbrook JD, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242
19. Gutmanas A, Alhroub Y, Battle GM, Berrisford JM, Bochet E, Conroy MJ, Dana JM, Fernandez Montecelo MA, van Ginkel G, Gore SP, Haslam P, Hatherley R, Hendrickx PM, Hirshberg M, Lagerstedt I, Mir S, Mukhopadhyay A, Oldfield TJ, Patwardhan A, Rinaldi L, Sahni G, Sanz-Garcia E, Sen S, Slowley RA, Velankar S, Wainwright ME, Kleywegt GJ (2014) PDBe: Protein Data Bank in Europe. Nucleic Acids Res 42(1):D285–D291
20. Kinjo AR, Suzuki H, Yamashita R, Ikegawa Y, Kudou T, Igarashi R, Kengaku Y, Cho H, Standley DM, Nakagawa A, Nakamura H (2012) Protein Data Bank Japan (PDBj): maintaining a structural data archive and resource description framework format. Nucleic Acids Res 40:D453–D460
21. Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte CF, Tolmie DE, Kent Wenger R, Yao H, Markley JL (2008) BioMagResBank. Nucleic Acids Res 36:D402–D408

22. Berman HM, Coimbatore Narayanan B, Costanzo LD, Dutta S, Ghosh S, Hudson BP, Lawson CL, Peisach E, Prlic A, Rose PW, Shao C, Yang H, Young J, Zardecki C (2013) Trendspotting in the protein data bank. FEBS Lett 587(8):1036–1045

23. Goodsell DS, Burley SK, Berman HM (2013) Revealing structural views of biology. Biopolymers 99(11):817–824

24. Wlodawer A, Miller M, Jaskolski M, Sathyanarayana BK, Baldwin E, Weber IT, Selk LM, Clawson L, Schneider J, Kent SB (1989) Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. Science 245(4918):616–621

25. Wensing AM, van Maarseveen NM, Nijhuis M (2010) Fifteen years of HIV protease inhibitors: raising the barrier to resistance. Antiviral Res 85(1):59–74

26. Cihlar T, Ray AS (2010) Nucleoside and nucleotide HIV reverse transcriptase inhibitors: 25 years after zidovudine. Antiviral Res 85(1):39–58

27. Yonath A (2005) Antibiotics targeting ribosomes: resistance, selectivity, synergism and cellular regulation. Annu Rev Biochem 74:649–679

28. Bulkley D, Innis CA, Blaha G, Steitz TA (2010) Revisiting the structures of several antibiotics bound to the bacterial ribosome. Proc Natl Acad Sci U S A 107(40):17158–17163

29. Rose PW, Bi C, Bluhm WF, Christie CH, Dimitropoulos D, Dutta S, Green RK, Goodsell DS, Prlic A, Quesada M, Quinn GB, Ramos AG, Westbrook JD, Young J, Zardecki C, Berman HM, Bourne PE (2013) The RCSB Protein Data Bank: new resources for research and education. Nucleic Acids Res 41(D1):D475–D482

30. Stierand K, Rarey M (2010) Drawing the PDB: protein − ligand complexes in two dimensions. Med Chem Lett 1:540–545

31. Knox C, Law V, Jewison T, Liu P, Ly S, Frolkis A, Pon A, Banco K, Mak C, Neveu V, Djoumbou Y, Eisner R, Guo AC, Wishart DS (2011) DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Nucleic Acids Res 39(Database issue):D1035–D1041

32. Young JY, Feng Z, Dimitropoulos D, Sala R, Westbrook J, Zhuravleva M, Shao C, Quesada M, Peisach E, Berman HM (2013) Chemical annotation of small and peptide-like molecules at the Protein Data Bank. Database 2013:bat079

33. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

34. Hopper P, Harrison SC, Sauer RT (1984) Structure of tomato bushy stunt virus. V. Coat protein sequence determination and its structural implications. J Mol Biol 177(4):701–713

35. Schmeing TM, Voorhees RM, Kelley AC, Gao YG, Murphy FVT, Weir JR, Ramakrishnan V (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. Science 326(5953):688–694

36. Voorhees RM, Weixlbaumer A, Loakes D, Kelley AC, Ramakrishnan V (2009) Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. Nat Struct Mol Biol 16(5):528–533

37. Gao YG, Selmer M, Dunham CM, Weixlbaumer A, Kelley AC, Ramakrishnan V (2009) The structure of the ribosome with elongation factor G trapped in the posttranslocational state. Science 326(5953):694–699

38. Nagar B, Hantschel O, Young MA, Scheffzek K, Veach D, Bornmann W, Clarkson B, Superti-Furga G, Kuriyan J (2003) Structural basis for the autoinhibition of c-Abl tyrosine kinase. Cell 112(6):859–871