How structural biologists and the Protein Data Bank contributed to recent US FDA new drug approvals

John D. Westbrook[1] and Stephen K. Burley[1,2,3]

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

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

Author E-mail Addresses: John.Westbrook@rcsb.org, Stephen.Burley@RCSB.org

Discovery and development of 210 new molecular entities [NMEs, new drugs] approved by the US Food and Drug Administration 2010-2016 was facilitated by 3D structural information generated by structural biologists worldwide and distributed on an open access basis by the Protein Data Bank [PDB]. The molecular targets for 94% of these NMEs are known. The PDB archive contains 5,914 structures containing one of the known targets and/or a new drug, providing structural coverage for 88% of the recently approved NMEs across all therapeutic areas. More than half of the 5,914 structures were published and made available by the PDB at no charge with no restrictions on usage >10 years before drug approval. Citation analyses revealed that these 5,914 PDB structures significantly impacted the very large body of publicly-funded research reported in publications on the NME targets that motivated biopharmaceutical company investment in discovery and development programs that produced the NMEs.