On the evolution of the quality of macromolecular models in the PDB
Supplementary Materials

Dariusz Brzezinski\textsuperscript{1,2,3,4}, Zbigniew Dauter\textsuperscript{5}, Wladek Minor\textsuperscript{4} and Mariusz Jaskolski\textsuperscript{1,6,*}

\textsuperscript{1}Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, 61-704, Poland; \textsuperscript{2}Institute of Computing Science, Poznan University of Technology, Poznan, 60-965, Poland; \textsuperscript{3}Center for Artificial Intelligence and Machine Learning, Poznan University of Technology, Poznan, 60-965, Poland; \textsuperscript{4}Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, USA; \textsuperscript{5}Synchrotron Radiation Research Section, Macromolecular Crystallography Laboratory, National Cancer Institute, Argonne National Laboratory, Argonne, USA; \textsuperscript{6}Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, 61-614, Poland;

*Correspondence e-mail: mariuszj@amu.edu.pl

Table S1. All-time journal ranking according to $P_{Q1}(t)$. The ranking includes all the journals that had at least 100 primary citations of structures in the PDB. $P_{Q1}(t)$ higher than 50\% means that the structures published in a given journal were, on average, better than 50\% of structures present in the PDB at the time of deposition. Journals with more than 1000 structures are highlighted in gray.

* Denotes journals that have average $P_{Q1}(t)$ significantly different than the expected value in the population, according to Welch’s t-test with Bonferroni correction at significance level $\alpha=0.001$.

| Rank | Journal | Avg. $P_{Q1}(t)$ \([\%]\) | Avg. resolution \([\text{Å}]\) | Structure count |
|------|---------|--------------------------|----------------------------|-----------------|
| 1    | TUBERCULOSIS (EDINB)* | 84.81 | 2.02 | 132 |
| 2    | CHEMISTRY* | 72.34 | 1.82 | 242 |
| 3    | CHEM COMMUN* | 71.88 | 1.81 | 280 |
| 4    | CHEMBIOCHEM* | 71.84 | 1.92 | 527 |
| 5    | ACS INFECT DIS* | 71.69 | 1.95 | 153 |
| 6    | EUR J MED CHEM* | 71.39 | 2.03 | 418 |
| 7    | ACS CATAL* | 71.33 | 1.94 | 241 |
| 8    | ORG BIOMOL CHEM* | 71.22 | 1.88 | 167 |
| 9    | J INORG BIOCHEM* | 71.11 | 1.78 | 171 |
| 10   | IUCRJ* | 70.62 | 1.95 | 281 |
| 11   | J SYNCHROTRON RADIAT* | 70.24 | 1.79 | 147 |
| 12   | ACS OMEGA* | 69.77 | 1.81 | 102 |
| 13   | J BIOL INORG CHEM* | 68.05 | 1.88 | 265 |
| 14   | CHEM SCI* | 67.78 | 1.87 | 268 |
| 15   | J COMPUT AIDED MOL DES | 67.57 | 1.88 | 115 |
| 16   | ANGEW CHEM* | 67.45 | 1.92 | 1065 |
| 17   | NAT CHEM* | 67.45 | 1.95 | 173 |
| 18   | FEBS J* | 66.24 | 2.02 | 1359 |
| 19   | APPL ENVIRON MICROBIOL | 66.00 | 2.01 | 112 |
| 20   | ANTIMICROB AGENTS CHEMOTHER* | 65.64 | 1.96 | 309 |
| 21   | ARCH BIOCHEM BIOPHYS* | 65.33 | 2.08 | 276 |
| 22   | GLYCOBIOLOGY | 65.13 | 1.96 | 188 |
| 23   | ACS CHEM BIOL* | 64.70 | 2.04 | 1104 |
| 24   | CEMMEDCHEM* | 64.65 | 1.94 | 556 |
| 25   | BIOCHEM J* | 64.32 | 2.11 | 1033 |
| 26   | INT J MOL. SCI | 63.83 | 2.12 | 103 |
|   | Title                        | Impact Factor | Citations |
|---|-----------------------------|---------------|-----------|
| 27 | ACTA CRYST F*               | 63.71         | 1466      |
| 28 | ACTA CRYST D*               | 63.61         | 4952      |
| 29 | J AM CHEM SOC*              | 63.37         | 2369      |
| 30 | PROTEIN ENG DES SEL         | 63.05         | 294       |
| 31 | TO BE PUBLISHED*            | 62.75         | 22421     |
| 32 | BIOPHYS J                   | 62.49         | 199       |
| 33 | MBIO                        | 62.04         | 169       |
| 34 | SCI REP*                    | 61.27         | 1847      |
| 35 | NAT CHEM BIOL               | 61.12         | 1013      |
| 36 | FEBs LETT                   | 61.10         | 814       |
| 37 | BIOCHEMISTRY*               | 60.68         | 8896      |
| 38 | PLOS PATHOG                 | 60.66         | 656       |
| 39 | CELL CHEM BIOL              | 60.51         | 902       |
| 40 | J BIOCHEM                   | 60.30         | 279       |
| 41 | PLOS ONE                    | 60.05         | 2057      |
| 42 | NAT COMMUN*                 | 60.04         | 3538      |
| 43 | FASEB J                     | 60.01         | 161       |
| 44 | BMC STRUCT BIOL             | 59.69         | 228       |
| 45 | BIOCHIMIE                   | 59.58         | 128       |
| 46 | PROTEIN SCI                 | 59.57         | 2235      |
| 47 | ACS MED CHEM LETT           | 59.41         | 1062      |
| 48 | J MED CHEM                  | 58.68         | 5525      |
| 49 | BIOCHEM BIOPHYS ACTA        | 58.15         | 600       |
| 50 | PROTEINS                    | 57.97         | 1999      |
| 51 | BIOORG MED CHEM             | 57.87         | 659       |
| 52 | MOL MICROBIOL               | 57.66         | 412       |
| 53 | J BIOL CHEM                 | 57.41         | 11055     |
| 54 | NAT MICROBIOL               | 56.88         | 111       |
| 55 | J STRUCT BIOL               | 56.58         | 1038      |
| 56 | INT J BIOL MACROMOL         | 56.48         | 168       |
| 57 | J STRUCT FUNCT GENOM        | 55.69         | 168       |
| 58 | STRUCTURE*                  | 55.54         | 5348      |
| 59 | J MOL BIOL*                 | 55.23         | 9507      |
| 60 | VIROLOGY                    | 54.62         | 126       |
| 61 | BIOCHEM BIOPHYS RES COMMUN  | 54.29         | 976       |
| 62 | MABS                        | 54.14         | 115       |
| 63 | PLANT CELL                  | 53.65         | 138       |
| 64 | J BACTERIOL                 | 53.63         | 371       |
| 65 | NUCLEIC ACIDS RES*          | 53.31         | 2127      |
| 66 | CELL HOST MICROBE           | 53.10         | 107       |
| 67 | J VIROL*                    | 52.22         | 957       |
| 68 | PNAS*                       | 52.17         | 7376      |
| 69 | SCI ADV                     | 51.69         | 182       |
| 70 | MOL PHARMACOL               | 51.48         | 129       |
| 71 | RNA                         | 51.48         | 245       |
| 72 | PLOS BIOL                   | 51.12         | 336       |
| 73 | J EXP MED                   | 50.54         | 121       |
| 74 | ELIFE*                      | 49.25         | 869       |
| 75 | J IMMUNOL*                  | 49.01         | 296       |
| 76 | PROTEIN CELL                | 48.77         | 178       |
| 77 | BIOORG MED CHEM LETT*       | 47.44         | 1590      |
| 78 | EMBG REP*                   | 47.25         | 211       |
| 79 | CELL REP*                   | 46.89         | 399       |
| 80 | GENES DEV*                  | 46.62         | 279       |
| 81 | CELL. RES*                  | 45.36         | 189       |
| 82 | SCIENCE*                    | 44.98         | 1949      |
| 83 | NATURE*                     | 44.77         | 3060      |
| 84 | NEURON*                     | 44.65         | 149       |
| 85 | NAT STRUCT MOL BIOL*        | 44.10         | 2915      |
| 86 | COMMUN BIOL*                | 43.81         | 104       |
| 87 | EMBJ*                       | 43.63         | 1910      |
| 88 | IMMUNITY*                   | 43.18         | 265       |
| 89 | MOL CELL*                   | 42.69         | 1599      |
| 90 | NAT IMMUNOL*                | 42.15         | 119       |
| 91 | CELL*                       | 40.07         | 1563      |
Figure S1. Histograms of quality metric values of structures found in the PDB.
Figure S2. Average $P_{Q_1}(t,d)$ of popular journals for each year.
Figure S3. Journal ranking over time according to $P_{Q_1}(t)$. 
Figure S4. Journal quality over time according to $P_{Qi}(t)$. 
Figure S5. Scatterplots of the values of Clashscore, Ramachandran outliers, and Rotamer outliers found in the PDB.