



OPEN

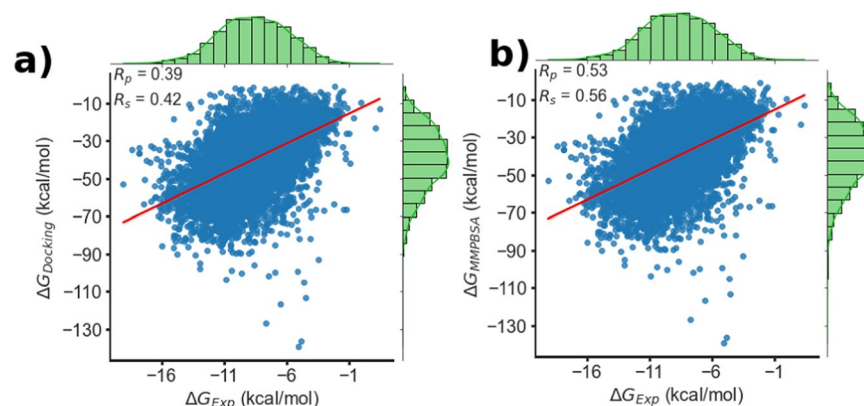
Author Correction: PLAS-20k: Extended Dataset of Protein-Ligand Affinities from MD Simulations for Machine Learning Applications

Divya B. Korlepara, C. S. Vasavi, Rakesh Srivastava , Pradeep Kumar Pal, Saalim H. Raza, Vishal Kumar, Shivam Pandit, Aathira G. Nair, Sanjana Pandey, Shubham Sharma , Shruti Jeurkar, Kavita Thakran, Reena Jaglan, Shivangi Verma, Indhu Ramachandran, Prathit Chatterjee, Divya Nayar & U. Deva Priyakumar

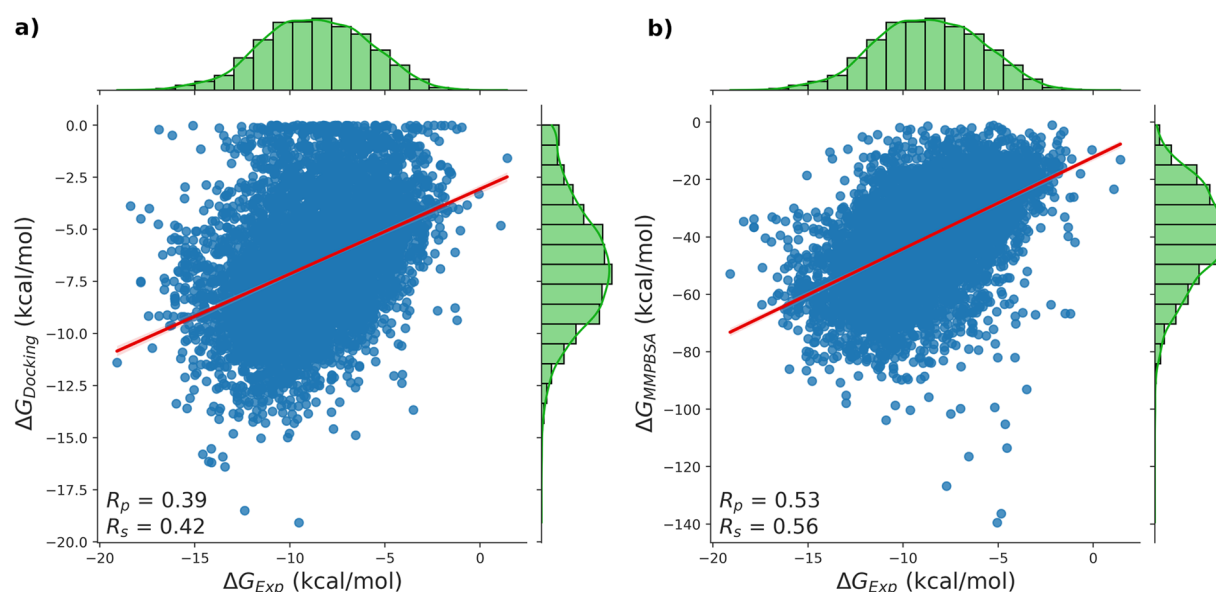
Correction to: *Scientific Data* <https://doi.org/10.1038/s41597-023-02872-y>, published online 09 February 2024

The original version of Figure 2 contained only the b) panel, incorrectly duplicated, instead of both a) and b).

Previous figure 2



Correct figure 2



There was also an error in the Technical Validation section. In the original version we stated that a subset of 6842 complexes from PLAS-20k had experimental binding affinities whereas in fact this number was 6622. This error was made due to the original size of PLAS-20k being reduced from 20,000 to 19,500 complexes during the dataset creation but this reduction in size was not also accounted for in the validation sample. The correct (lower) sample size has been included in the manuscript (6842 replaced with 6622, where mentioned in the Technical Validation section and Figure 2 caption). Furthermore, the Pearson correlation coefficient (R_p) derived from this sample have been corrected to account for the correct sample size (originally stated as 0.50, now corrected to 0.53).

The dataset file at figshare has also been amended so the correct number of PLC data points are present, and to remove a duplicated file. The DOI for the previous version is <https://doi.org/10.6084/m9.figshare.c.6742521.v1>. The DOI for the updated files is <https://doi.org/10.6084/m9.figshare.c.6742521.v2>, which has been corrected in the data citation in the manuscript (reference 39).

Finally, two individuals were missing from the Acknowledgements section - Ms. Ananya Priya and Mr. Suyash Gupta. These individuals, along with the others currently listed, helped in data post-processing and website development.

All changes – the replacement of the figure, text changes, and change to the dataset DOI, have been applied to the pdf and HTML versions of the paper.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024