Cole-Puli International Conference Publishing Group

Science, Engineering and Technology Proceedings

Review

2025 International Conference on Science Technology, Architecture, Power and Intelligent Information Technology (APIIT 2025)

Evolution and Trends in the Electron Microscopy Data Bank

Duoyan Hu [1,*]

[1] University of Wisconsin-Madison, Madison, WI, USA
* Correspondence: Duoyan Hu, University of Wisconsin-Madison, Madison, WI, USA

Abstract: Cryo-electron microscopy (cryo-EM) and cryo-electron tomography (cryo-ET) have transformed structural biology by enabling the visualization of biological macromolecules at near-atomic resolution. Since its establishment in 2002, the Electron Microscopy Data Bank (EMDB) has played a pivotal role in this "Resolution Revolution," serving as a central repository for cryo-EM data. Over the past decade, major advances in detector technology, image processing, and analytical methodologies have led to a surge in cryo-EM structure depositions. This study provides a comprehensive analysis of temporal trends in the EMDB, highlighting the growing number of entries, improvements in resolution, and the evolution of detector usage. It emphasizes the rise of single-particle analysis (SPA) as the dominant cryo-EM technique and examines how high-resolution detectors have enhanced data quality. Additionally, it explores resolution standards across leading scientific journals, underscoring the role of EMDB not only as a data archive but also as a catalyst for innovation in structural biology.

Keywords: Electron Microscopy Data Bank (EMDB); cryogenic electron microscopy (cryo-EM); cryo-electron tomography (cryo-ET); single-particle analysis (SPA)

1. Introduction

Recent advances in cryogenic electron microscopy (cryo-EM) and cryo-electron tomography (cryo-ET) have been driven primarily by improvements in infrastructure, particularly the development of centralized databases such as the Electron Microscopy Data Bank (EMDB). These technologies have redefined the study of biological macromolecules, complexes, and even entire cells by achieving resolutions that were once unattainable using traditional techniques like X-ray crystallography. This progress, often referred to as the "Resolution Revolution," has made it possible to routinely determine atomic and near-atomic structures, with resolutions approaching 1.2 Å [1]. These capabilities have positioned cryo-EM as a leading method for imaging complex biological systems.

The EMDB, established in 2002 by the Macromolecular Structure Database group at EMBL–EBI, serves as a public archive for electron microscopy density maps of macromolecules. Over the past decade, the number of deposited cryo-EM structures in the EMDB has grown rapidly. Continued innovations in microscopy, detection technologies, and computational tools are expected to sustain and further accelerate this growth. These technological developments have not only improved structural resolution but also enabled researchers to capture the dynamics and heterogeneity inherent in biological systems [2,3].

Single-particle analysis (SPA) has emerged as the predominant method in cryo-EM, playing a central role in both conventional and in situ structural determination. The latter allows scientists to observe biological macromolecules within their native cellular environments, further expanding the scope and impact of cryo-EM. Iterative refinements to

Received: 28 April 2025
Revised: 02 May 2025
Accepted: 15 May 2025
Published: 02 June 2025

SPA and complementary approaches such as cryo-ET have led to increasingly accurate models and deeper biological insights.

As the EMDB matured, the need for robust validation tools became increasingly apparent. To address this, the cryo-EM community collaborated to develop a set of standardized validation protocols. One major outcome of these efforts was the creation of the Validation Analysis (VA) resource, a three-tiered system designed to assess and communicate the reliability of deposited structures [3]. The system includes a development tier for expert users exploring new validation methods, a production tier for applying standardized tools to new submissions, and an embargoed tier that is released shortly after publication. This framework is now integrated into the wwPDB OneDep system, enabling automated analysis during the deposition process and enhancing the overall reliability and transparency of EMDB entries [4].

The integration between the EMDB and other structural databases—such as the Protein Data Bank (PDB), the AlphaFold Protein Structure Database, and the Electron Microscopy Public Image Archive (EMPIAR)—ensures that every aspect of a cryo-EM experiment is accessible. This includes the full data pipeline, from raw images to atomic models. These cross-archive relationships allow for seamless inference of atomic coordinates (via the PDB), access to raw image data (via EMPIAR), and volume data (via the EMDB), fostering a collaborative framework that enhances structural understanding of biological systems. Furthermore, the EMDB has introduced deposition standards for raw and half-maps in certain cases, contributing to greater transparency and reproducibility in the field.

Efforts are ongoing to expand the EMDB's validation strategies, aiming to increase the availability of reliable structures and improve harmonization with other international repositories. The evolving nature of the EMDB reflects the continual refinement of the resource—new validation methods and criteria are being developed and incrementally incorporated into version 3 of the database. As cryo-EM continues to push the boundaries of biological inquiry—particularly within native cellular environments and in the study of dynamic supramolecular complexes—the EMDB has become an indispensable tool for both experimental and computational biologists. It serves not only as a record of what has been achieved but also as a driver of innovation and discovery.

The technological advancements associated with cryo-EM have had a profound effect on research productivity and scholarship, extending beyond methodological improvements alone. As the number of structures in the EMDB has grown, cryo-EM data and techniques have been adopted across an increasingly wide range of scientific disciplines. Thousands of review articles reference EMDB entries annually, illustrating its widespread influence. In 2023 alone, the number of new EMDB entries exceeded 7,000, underscoring the rapid expansion of cryo-EM publications in both volume and usage. Cryo-EM has become a foundational resource in fields such as structural biology, molecular biology, and pharmacology, where high-resolution structural data drive scientific advancement and innovation.

Today, the EMDB serves as a global hub for the structural biology community, hosting over 40,000 entries that encompass a vast array of biological structures—from individual proteins to large molecular complexes and cellular assemblies. The database is constantly updated to reflect technological progress, including the incorporation of artificial intelligence and machine learning tools to support real-time validation and data integration. As a result, the EMDB has evolved beyond a static archive; it now functions as a dynamic research platform equipped with advanced data analysis and visualization tools, enabling new directions in scientific exploration.

This study focuses on temporal trends and recent innovations in the EMDB, with particular attention to three key areas: (1) the growth of EMDB entries and the establishment of single-particle analysis (SPA) as the predominant cryo-EM technique; (2) technological advancements over time, especially in detector development and their contributions to structural data deposition; and (3) the steady improvement in resolution over the

past two decades. By examining these developments, this paper offers a historical overview of the EMDB's evolution, highlights current best practices, and reinforces the ongoing significance of this resource for the cryo-EM community.

In summary, this paper serves as a valuable reference for researchers seeking to understand the current state of cryo-EM data housed in the EMDB and how these resources can support and accelerate scientific discovery. By mapping the field's growth, technological impacts, and methodological transitions, this analysis provides a roadmap for future directions in cryo-EM. It also supports informed decisions in methodology selection, data utilization, and innovation strategies, ultimately enhancing the broader impact of cryo-EM in science and medicine.

2. Methods

In this study, we examine trends in the Electron Microscopy Data Bank (EMDB) by analyzing deposition rates, resolution improvements, and the usage of various electron microscopy (EM) methods [5]. The data were gathered from the EMDB using both the advanced search functionality on the EMDB website and a custom Python script designed to access the EMDB API programmatically.

Data collection focused on specific attributes such as deposition date, EM method (e.g., single-particle analysis [SPA], subtomogram averaging [STA], electron tomography, and helical reconstruction), and resolution ranges. These attributes were queried for each year from the inception of the EMDB in 2002 through 2024 [6]. Method-specific entries were retrieved and analyzed individually to ensure precise trend identification.

To evaluate resolution trends, entries were grouped into three categories based on reported resolution: <4 Å, 4–6 Å, and >6 Å. This classification allowed for a clearer view of improvements in structure resolution over time.

All retrieved data were exported in CSV format to facilitate processing and visualization. The datasets were then parsed using Python, with the "matplotlib" library employed to generate graphical representations of the trends [7,8]. Each EM method's annual contribution was quantified and visualized as both absolute counts and percentages, providing a comprehensive perspective on methodological evolution within the EMDB.

Processed data were organized chronologically and filtered according to deposition method [9]. Final plots were refined and formatted for clarity, ensuring effective communication of the observed trends.

3. Results

The results of the analysis provide valuable insights into the evolving landscape of cryo-electron microscopy (cryo-EM), highlighting the expansion of the Electron Microscopy Data Bank (EMDB), the increasing adoption of specific EM methodologies, and notable advancements in structural resolution, as illustrated in Figure 1, which shows the annual growth in the number of EMDB and EMPIAR entries released.

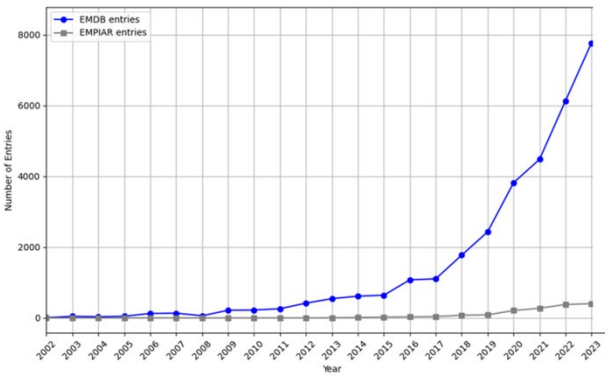

Figure 1. Number of EMDB and EMPIAR Entries Released by Year.

3.1. Growth of EMDB Entries Over Time

The growth of the EMDB has followed an exponential trajectory since its inception, with particularly rapid expansion observed from 2014 onward [10]. As shown in Figure 1, the number of entries deposited annually surged dramatically, increasing from fewer than 100 entries per year in the early 2000s to over 7,000 entries in 2023. This trend reflects the widespread adoption of cryo-EM, largely driven by key technological advancements. The period between 2014 and 2023 marks a phase of accelerated growth that aligns with the introduction of direct electron detectors and significant improvements in image processing techniques.

While the Electron Microscopy Public Image Archive (EMPIAR) has experienced more modest growth, the EMDB remains the principal repository for cryo-EM structures. The steady increase in EMPIAR deposits highlights a growing recognition of the importance of preserving raw data for subsequent analyses. Nevertheless, the volume of entries in EMPIAR remains substantially lower compared to the finalized maps archived in the EMDB. The dominance of single-particle analysis (SPA) in recent years, as well as the overall methodological distribution, is clearly reflected in Figure 2, which illustrates the number of EMDB entries by method over time.

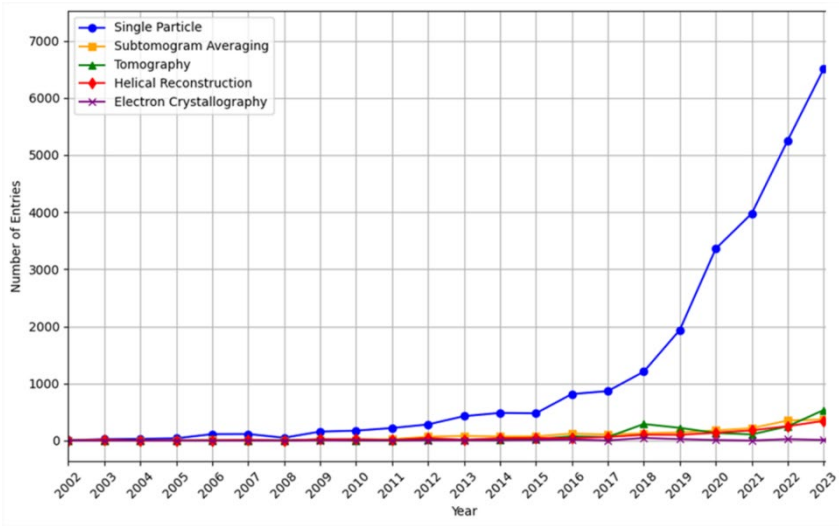

Figure 2. Number of EMDB Entries by Method Over Time.

3.2. Evolution of Electron Microscopy Methods

The rise of single-particle analysis (SPA) as the dominant method for structure determination is evident from the data presented in Figure 2. Since 2014, SPA has become the preferred approach for high-resolution structure determination, accounting for over 90% of deposited entries in recent years. The number of SPA entries surged dramatically, increasing from fewer than 500 in 2015 to more than 6,500 in 2023. This method's capability to resolve complex biological macromolecules at near-atomic resolution has established it as the gold standard for cryo-EM studies.

Subtomogram averaging and tomography have also experienced steady growth, especially since 2015. These techniques are increasingly utilized to investigate cellular and subcellular structures in their native contexts. Subtomogram averaging has demonstrated a consistent upward trend, reaching approximately 370 entries in 2023. Although tomography remains a minority method, it saw notable growth starting in 2018, with 535 entries recorded in 2023. This reflects a growing interest in approaches that capture structural details within intact cells and tissues.

Helical reconstruction, while less commonly employed compared to other methods, has maintained relatively stable usage over the years. Primarily applied to filamentous

assemblies such as fibrils and helical viruses, this method accounted for roughly 350 entries in 2024. The consistent resolution improvements observed across various methods—including helical reconstruction—are evident in Figure 3, which presents resolution trends of released EMDB entries from 2002 to 2024.

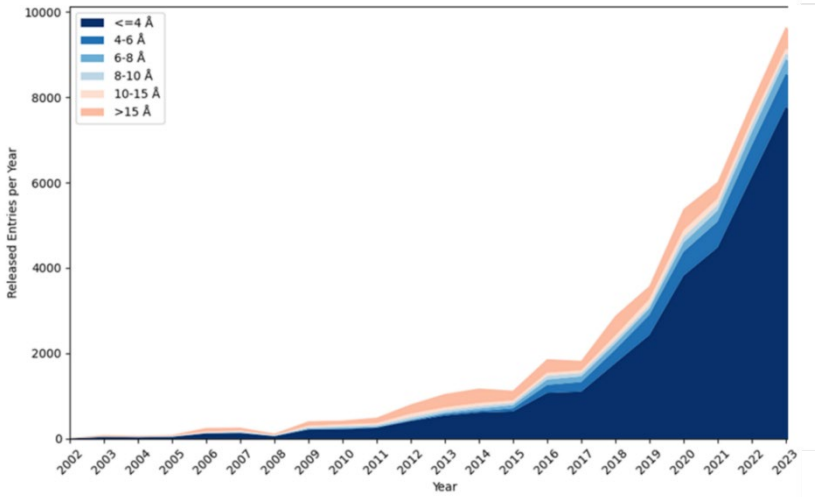

Figure 3. Resolution trends of released EMDB entries (2002-2024).

3.3. Improvement in Resolution Over Time

The resolution of structures deposited in the EMDB has improved markedly over the past decade, with the most significant advancements occurring after 2014. As illustrated in Figure 3, there has been a clear shift toward higher-resolution structures over time. In 2010, the majority of structures were resolved at resolutions greater than 6 Å, whereas by 2020, more than half of the deposited entries achieved resolutions better than 4 Å. This substantial improvement can be attributed to the adoption of direct electron detectors, advancements in image processing algorithms, and optimized cryo-EM sample preparation techniques.

Most recent depositions, from 2020 to 2023, have reached resolutions that rival or even exceed those typically attained by X-ray crystallography. The data reveal a pronounced trend toward atomic-level resolution for macromolecular structures, with a growing fraction of structures resolved within the 3–4 Å range. This trend correlates strongly with the adoption of advanced direct electron detectors, as shown in Figure 4, which illustrates the evolving usage of different detector technologies from 2002 to 2024. The availability of higher-resolution cryo-EM data in the EMDB has not only enhanced the repository's value but also empowered researchers to tackle increasingly complex biological questions with greater accuracy and detail.

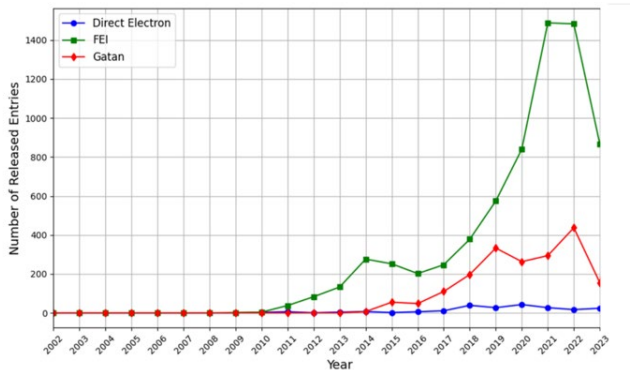

Figure 4. Trends of Released EMDB Entries for Different Detectors (2002-2024).

3.4. Trends in Detected Usage Over Time

Figure 4 illustrates the trends in EMDB entries from 2002 to 2024 based on detector technologies, including Direct Electron, FEI, and Gatan detectors. Beginning in 2014, following the introduction of direct electron detectors and improved processing methods, Gatan-based devices experienced a significant surge in popularity. The number of datasets recorded using Gatan detectors reached a peak last year, exceeding 7,000 EMDB entries within a single twelve-month period.

The growth trend for FEI detectors was more modest. Although the number of entries increased after 2016, it never matched the volume associated with Gatan-based datasets. This reflects the growing diversification of detector technologies available for electron microscopy.

While Direct Electron detectors have been pivotal in advancing resolution in cryo-EM, they have not yet become the predominant detector type. Although their usage has increased steadily since 2014, these devices have been employed more selectively, reaching their peak around 2022.

Overall, these trends highlight the technological evolution of detector systems in cryo-EM and their influence on data deposition rates. Gatan currently leads the field, largely due to its widespread use in published studies, while FEI and Direct Electron detectors tend to be employed more in specialized or high-resolution applications, as illustrated in Figure 5, which shows the cumulative growth of EMDB entries by detector type from 2002 to 2024.

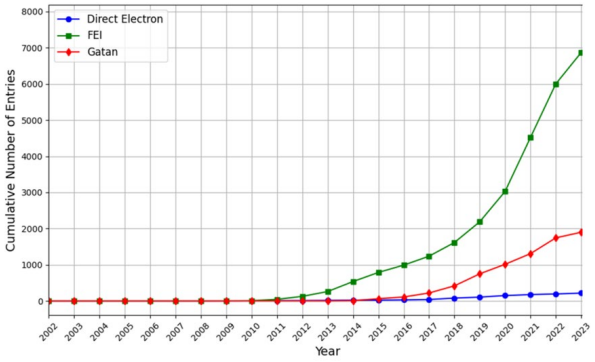

Figure 5. Cumulative Growth of EMDB Entries by Detector (2002-2024).

3.5. Increase in Cumulative EMDB Entries by Detector

The cumulative distribution of detector types is shown in Figure 5. Gatan detectors dominate the field, with nearly 70,000 cumulative entries projected by the end of 2024. This trend reflects a generational increase in the use of Gatan detectors in cryo-EM studies, playing a crucial role in driving the so-called "Resolution Revolution."

FEI detectors have also demonstrated steady cumulative growth, though their impact remains significantly less than that of Gatan. Their gradual adoption is expected to result in around 10,000 total entries by 2024.

Although Direct Electron detectors have marked a watershed moment in improving cryo-EM map resolution, their overall representation in the EMDB remains relatively small. By 2024, approximately 2,000 entries are expected to have been generated using Direct Electron detectors, indicating their more specialized yet essential role in advancing specific scientific investigations (see Figure 6).

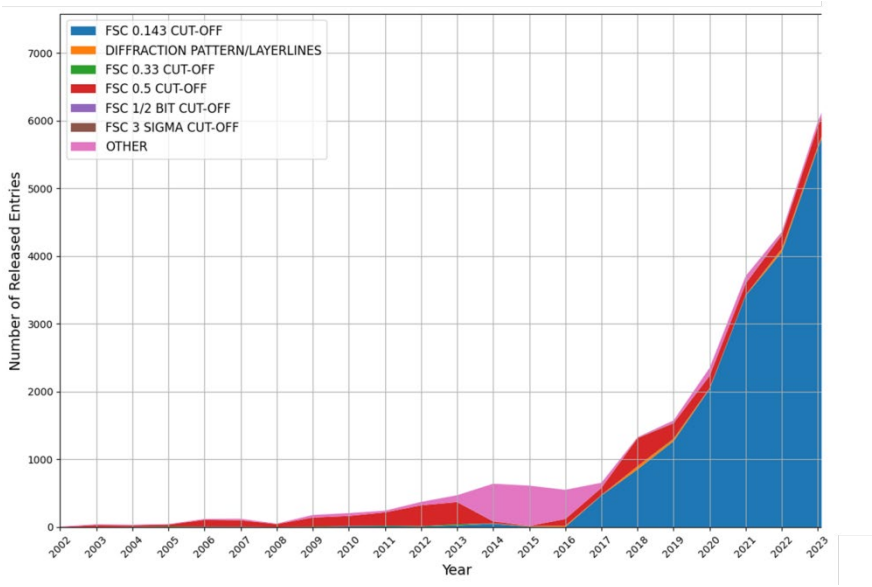

Figure 6. Trends of Released EMDB Entries by Resolution Methods (2002-2024).

Overall, these trends underscore Gatan's dominance in cryo-EM detector technology, while other detector types continue to gradually increase their presence as the field evolves.

3.6. Resolutions by Method

Figure 7 presents the trends in resolution assignment methods used for EMDB entries from 2002 to 2024. The determination of resolution near the noise threshold is predominantly based on the "FSC 0.143 cutoff" criterion, which has seen a dramatic increase in usage since early 2016 and continuing through mid-2024. This method has been employed in over 7,000 entries as of 2024, reflecting its critical role in achieving the highest resolutions in cryo-EM studies. The surge in its adoption directly correlates with advancements in detector technologies and image processing techniques, enabling unprecedented molecular detail [11].

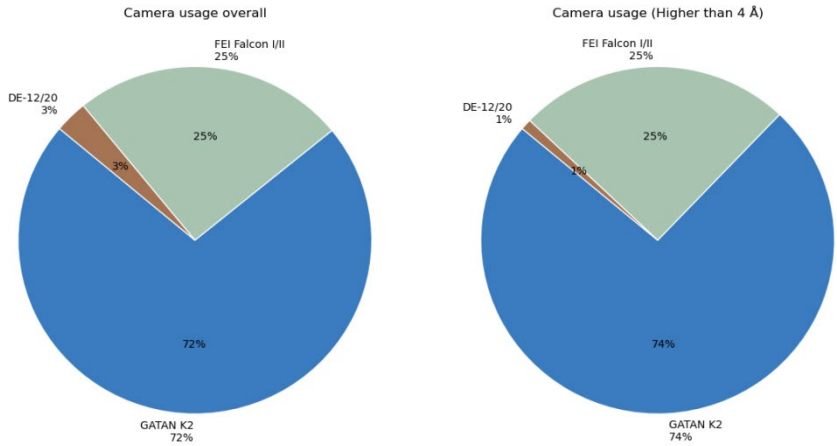

Figure 7. Camera Usage in Cryo-Em.

Other resolution assignment methods, such as "FSC 0.33 cutoff" and "FSC 0.5 cutoff," have maintained relatively stable contributions over time, appearing intermittently across several periods. The category labeled "Other" shows a more complex pattern, likely representing a variety of alternative or less standardized approaches.

The "Diffraction Pattern/Layerlines" method, which was more popular in earlier years, has steadily declined since 2016.

Together, these trends illustrate the ongoing evolution towards higher-resolution standards in cryo-EM. The dominance of the FSC 0.143 cutoff method underscores the maturation of cryo-EM into a high-resolution technique, routinely enabling visualization of molecular structures with exceptional clarity and accuracy.

3.7. Camera Usage in Cryo-EM

Camera technologies used in cryo-electron microscopy (cryo-EM) have a fundamental impact on data quality and achievable resolution. Figure 8 illustrates the usage distribution of the three most popular camera models in the EMDB: Gatan K2, FEI Falcon I/II, and Direct Electron DE-12/20, considering both all entries and those with resolutions better than 4 Å.

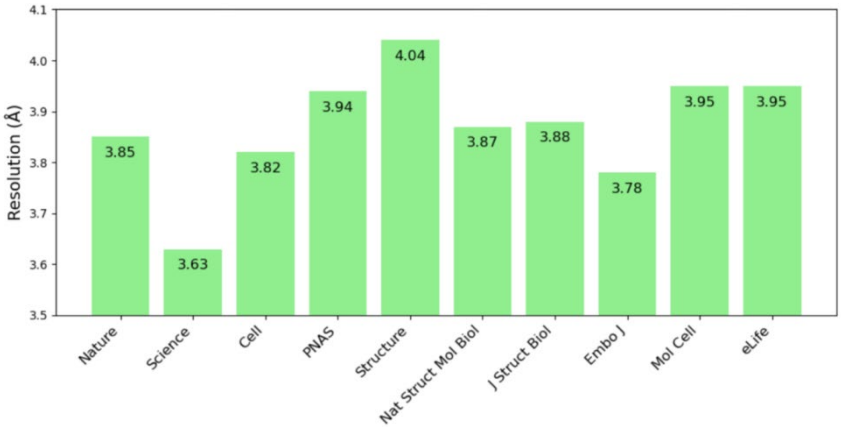

Figure 8. Journals associated resolution (higher than 5A).

The left pie chart represents the overall camera usage, revealing that Gatan K2 dominates with 72% of all entries. This significant share reflects its widespread adoption, likely due to its superior resolution capabilities and stability, particularly suited for single-particle analysis (SPA) [12]. FEI Falcon I/II follows as the second most common camera, accounting for 25% of entries, suggesting it serves as a reliable alternative in many cryo-EM applications. In contrast, the DE-12/20 camera accounts for only 3% of all entries, indicating more limited use, possibly for specific or niche applications.

Focusing on entries with resolutions better than 4 Å, which represent higher precision imaging, Gatan K2's dominance increases slightly to 74%, reinforcing its role in enabling high-resolution data collection. FEI Falcon I/II maintains a steady 25% share, while DE-12/20's presence drops to 1%, implying that it may be less optimized for the demands of high-resolution structural determination compared to the other two models.

Within the cryo-EM community, there is a clear preference for Gatan K2 and similar advanced camera technologies, as these have proven essential for producing high-quality images [13]. This data suggests that as researchers strive for ever-finer structural details, reliance on high-resolution optimized camera models is unlikely to wane. This trend reflects the broader progression in cryo-EM toward increasingly higher resolution imaging and underscores the critical role of imaging hardware in advancing the field.

3.8. Journals Associated with Resolution (Higher than 5Å)

The bar chart compares the average resolution of cryo-EM structures published in the top 10 scientific journals—namely Nature, Science, PNAS, Journal of Molecular Biology (JMB), PLOS Biology, EMBO Journal, Structure, Cell, and Journal of Cell Biology—

monitored between 2012 and 2015 [14]. Structures are categorized by resolution thresholds, particularly distinguishing those with resolutions better than 5 Å, which is considered the minimal meaningful resolution attainable in cryo-EM. This classification provides insight into the relative data quality and level of structural detail published across these journals.

As shown in the chart, *Science* exhibits the highest average resolution at 3.63 Å, followed closely by *EMBO Journal* at 3.78 Å, indicating that these journals uphold rigorous standards for cryo-EM resolution. In contrast, *Structure* has the lowest average resolution of 4.04 Å, which may reflect its publication of a broader range of structures, including more complex or less well-resolved assemblies.

Other journals maintain relatively high-resolution standards as well, with *Cell* and *Nature* averaging around 3.80 Å and 3.85 Å, respectively, positioning them between the extremes. Meanwhile, *PNAS*, *Molecular Cell*, and *eLife* show slightly higher average resolutions, ranging from approximately 3.94 Å to 3.95 Å, suggesting a somewhat broader focus on cryo-EM data resolution within their published studies.

This analysis illustrates how individual journals prioritize cryo-EM resolution quality during peer review, reflecting both community standards and editorial policies. Journals with higher average resolutions tend to publish studies emphasizing detailed molecular mechanisms, while those with wider resolution distributions often feature structural data pertaining to larger complexes or cellular components. Such diversity in publication focus ultimately supports addressing a broad spectrum of biological questions and research needs within structural biology.

4. Conclusion

The EMDB is not only a continuously updated database that keeps pace with emerging technologies but also a dynamic, living repository reflecting feedback and contributions from the scientific community. As cryo-EM methodologies continue to advance and become more widely accessible, the EMDB is poised to play an increasingly central role in driving discoveries of biological structures—thereby fostering new insights and technological innovations in molecular biology and related fields. The trends revealed here underscore the ongoing evolution of cryo-EM technologies and highlight the establishment of shared standards for publishing scientific results, which will collectively shape the future trajectory of structural biology.

References

1. S. Carvalho, J. Leite, S. Galdo-Álvarez, and O. F. Gonçalves, "The emotional movie database (EMDB): A self-report and psychophysiological study," *Appl. Psychophysiol. Biofeedback*, vol. 37, pp. 279–294, 2012.
2. W. Chiu, M. F. Schmid, G. D. Pintilie, and C. L. Lawson, "Evolution of standardization and dissemination of cryo-EM structures and data jointly by the community, PDB, and EMDB," *J. Biol. Chem.*, vol. 296, p. 100560, 2021, doi: 10.1016/j.jbc.2021.100560.
3. J. Turner, S. Abbott, N. Fonseca, R. Pye, L. Carrijo, A. K. Duraisamy *et al.*, "Correction to 'EMDB—the Electron Microscopy Data Bank'," *Nucleic Acids Res.*, vol. 52, p. D1701, 2024, doi: 10.1093/nar/gkad1199.
4. A. Ignatiou, K. Macé, A. Redzej, T. R. Costa, G. Waksman, and E. V. Orlova, "Structural analysis of protein complexes by Cryo-Electron microscopy," in *Bacterial Secretion Syst.: Methods Protocols*, 2023, pp. 431–470, doi: 10.1007/978-1-0716-3445-5_27.
5. G. J. Kleywegt, P. D. Adams, S. J. Butcher, C. L. Lawson, A. Rohou, P. B. Rosenthal *et al.*, "Community recommendations on cryoEM data archiving and validation," *IUCrJ*, vol. 11, no. 2, pp. 140–151, 2024, doi: 10.1107/S2052252524001246.
6. I. Lagerstedt, W. J. Moore, A. Patwardhan, E. Sanz-García, C. Best, J. R. Swedlow *et al.*, "Web-based visualisation and analysis of 3D electron-microscopy data from EMDB and PDB," *J. Struct. Biol.*, vol. 184, no. 2, pp. 173–181, 2013, doi: 10.1016/j.jsb.2013.09.021.
7. A. Patwardhan, "Trends in the electron microscopy data bank (EMDB)," *Biol. Crystallogr.*, vol. 73, no. 6, pp. 503–508, 2017, doi: 10.1107/S2059798317004181.
8. A. Patwardhan and C. L. Lawson, "Databases and archiving for CryoEM," *Methods Enzymol.*, vol. 579, pp. 393–412, 2016, doi: 10.1016/bs.mie.2016.04.015.

9. M. Salaga, A. Mokrowiecka, M. Zielinska, E. Malecka-Panas, R. Kordek, E. Kamysz, and J. Fichna, "New peptide inhibitor of dipeptidyl peptidase IV, EMDB-1 extends the half-life of GLP-2 and attenuates colitis in mice after topical administration," *J. Pharmacol. Exp. Ther.*, vol. 363, no. 1, pp. 92–103, 2017, doi: 10.1124/jpet.117.242586.
10. J. Salavert-Torres, A. Iudin, I. Lagerstedt, E. Sanz-García, G. J. Kleywegt, and A. Patwardhan, "Web-based volume slicer for 3D electron-microscopy data from EMDB," *J. Struct. Biol.*, vol. 194, no. 2, pp. 164–170, 2016, doi: 10.1016/j.jsb.2016.02.012.
11. H. Suzuki, T. Kawabata, and H. Nakamura, "Omokage search: shape similarity search service for biomolecular structures in both the PDB and EMDB," *Bioinformatics*, vol. 32, no. 4, pp. 619–620, 2016, doi: 10.1093/bioinformatics/btv614.
12. S. Velankar, G. Van Ginkel, Y. Alhroub, G. M. Battle, J. M. Berrisford, M. J. Conroy *et al.*, "PDBe: improved accessibility of macromolecular structure data from PDB and EMDB," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D385–D395, 2016, doi: 10.1093/nar/gkv1047.
13. Z. Wang, A. Patwardhan, and G. J. Kleywegt, "Validation analysis of EMDB entries," *Biol. Crystallogr.*, vol. 78, no. 5, pp. 542–552, 2022, doi: 10.1107/S205979832200328X.
14. The wwPDB Consortium, "EMDB—the Electron Microscopy Data Bank," *Nucleic Acids Res.*, vol. 52, no. D1, pp. D456–D465, Jan. 2024, doi: 10.1093/nar/gkad1019.