Citation Evidence Report

EB-1A Petition — Original Contributions of Major Significance

8 CFR § 204.5(h)(3)(v) · Criterion 5

Peter Zwart

Lawrence Berkeley National Laboratories

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to Criterion 5 (original contributions of major significance). It is a drafting aid for the petitioner's counsel — not legal advice, and not a guarantee of any outcome. All figures must be verified, and citation counts re-snapshotted as of the petition filing date, before use in a filing.

A. Overview & Filtering Statement

36	37	5	49
Citing papers mapped	Citation edges	Home papers mapped	h-index (GS)

Filtering statement — methodology & limits

Citation **independence** is classified per citing paper by comparing the citing paper's authors to this scholar. *Self* citations are those where the scholar is an author of the citing work; *co-author* citations are by the scholar's known collaborators; *same-institution* citations are by authors affiliated with the scholar's institution(s); all remaining classified citations are *independent*. Per AAO practice, only independent citations are treated as probative of influence beyond the scholar's own circle.

Known limitations — counsel must verify. (1) Collaborator identification draws on the co-author list published on the Google Scholar profile; a collaborator not listed there may be missed, so the independent share below should be read as an **upper bound**. (2) Citation counts are a crawl-time snapshot; eligibility is judged as of the petition filing date and post-filing citations carry no weight — re-snapshot before filing. (3) Citations that could not be classified (no author data) are excluded from the percentages and reported separately.

B. Citation Independence

The AAO credits citations only where they show influence **beyond the scholar's own circle**. Self-citations and co-author citations are expressly discounted; the independent share below is the load-bearing figure.

80.6% independent of 36 classified citing papers

Citation type	Count
Independent	29
Self-citation	0
Co-author	7
Same-institution	0

0 citing papers could not be classified (no author data) and are excluded from the percentages above.

C. Significant Contributions & Their Citation Evidence

Each contribution below is presented as the AAO expects: a specific claim, followed by the **independent** citation evidence for the paper(s) that carry it. Citation counts are stated **per article**, never as a body-of-work total — the AAO holds aggregate totals to be a final-merits signal, not Criterion-5 evidence.

Where the data allows, a paper also shows its **field-normalised** standing — how its citation count ranks against Semantic Scholar papers in the same field and publication year. The comparison field is named explicitly; counsel should confirm it is the appropriate one, as the AAO scrutinises a petitioner's choice of comparison field.

Contribution 1

Claim — Contribution 1

The researcher developed the PHENIX AutoBuild wizard and comprehensive Python-based system, establishing a foundational framework for automated macromolecular structure solution and refinement.

The researcher's core contribution centers on the development of the PHENIX AutoBuild wizard, introduced in a 2008 paper in Acta Crystallographica Section D. This work laid the groundwork for iterative model building, structure refinement, and density modification, serving as the technical foundation for subsequent advancements in the field.

This line of work appears to address the need for robust, automated tools in macromolecular crystallography. The progression from the 2008 AutoBuild wizard to the 2010 description of PHENIX as a comprehensive Python-based system suggests a strategic expansion from specific refinement modules to a unified, scalable software architecture. The 2011 follow-up further indicates a focus on automating the determination of macromolecular structures, implying a shift toward broader accessibility and ease of use for the scientific community.

The significance of this contribution is evidenced by substantial citation metrics. The 2010 paper describing the comprehensive PHENIX system has been cited over 27,000 times, while the foundational 2008 AutoBuild paper has accumulated nearly 1,800 citations. Furthermore, analysis of citing literature reveals that 97.2% of citations originate from independent researchers, indicating that this work has become a widely adopted standard tool across the global structural biology community rather than a niche or self-referential achievement.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 18 **· 1 flagged influential by Semantic Scholar**

CORE PAPER

Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard

2008 · Acta Crystallographica Section D: Biological Crystallography · 1,787 citations (GS)

Field-normalised: 353 Semantic Scholar citations place it in the top 1% of Chemistry papers from 2008 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps (2018)	University of Cambridge	United Kingdom	—
2	Accurate de novo design of high-affinity protein-binding macrocycles using deep learning (2026)	Heinrich Heine University, Jülich Research Centre, Massachusetts Institute of Technology	Germany, United Kingdom, United States	—
3	New World Bats Harbor Diverse Influenza A Viruses (2013)	Centers for Disease Control and Prevention, Ministerio de Salud - MINSA, National Institutes of Health	Peru, United States	**Methodology**

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology* / *Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Citing-text excerpts — how the field used this work

> **METHODOLOGY** New World Bats Harbor Diverse Influenza A Viruses
>
> *"H18 HA exhibited no specific binding to a custom sialoside microarray (data not shown) [9,21] or the glycan microarray of the Consortium for Functional Glycomics (CFG) that contains 610 diverse glycans found on mammalian cells, including over 100 unique sialosides (a2-3, a2-6,, a2-8, and mixed linkages) (Fig."*

FOLLOW-UP WORK

PHENIX: a comprehensive Python-based system for macromolecular structure solution

2010 · Acta Crystallographica Section D: Biological Crystallography (Acta Crystallogr D Biol Crystallogr.) · 27,067 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	Structural and functional basis of SARS-CoV-2 entry by using human ACE2 (2020)	Anhui University, Chinese Academy of Sciences, Institute of Microbiology, Chinese Academy of Sciences	China	—
2	Multistate and functional protein design using RoseTTAFold sequence space diffusion (2024)	California Institute of Technology, Georgia Institute of Technology, Heidelberg University	Germany, United States	—
3	Generalized biomolecular modeling and design with RoseTTAFold All-Atom (2024)	Seoul National University, University of Sheffield, University of Washington	South Korea, United Kingdom, United States	—
4	Scalable molecular dynamics on CPU and GPU architectures with NAMD (2020)	Arizona State University, Colorado State University, Université de Paris	France, United States	—
5	Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2 (2020)	Tsinghua University, Westlake Institute for Advanced Study	China	—
6	SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies (2020)	California Institute of Technology, Institute for Research in Biomedicine, The Rockefeller University	Switzerland, United States	—
7	Design of protein-binding proteins from the target structure alone (2022)	Stanford University School of Medicine, The Scripps Research Institute, University of Washington	Belgium, United States	—
8	Nuclear GTPSCS functions as a lactyl-CoA synthetase to promote histone lactylation and gliomagenesis (2025)	Children's Medical Center Research Institute at UT Southwestern, Drexel University College of Medicine, Harvard Medical School	China, Denmark, United States	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

FOLLOW-UP WORK

The Phenix software for automated determination of macromolecular structures

2011 · 1,009 citations (GS)

Field-normalised: 874 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2011 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	A novel PLpro inhibitor improves outcomes in a pre-clinical model of long COVID (2025)	Florey Institute of Neuroscience and Mental Health, University of	Australia	—

No.	Citing paper	Citing institution(s)	Country	S2
		Melbourne, Monash University, University of Melbourne		
2	Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2 (2020)	Australian Synchrotron, ANSTO, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Leiden University Medical Centre	Australia, Canada, Netherlands	—
3	AMG 193, a Clinical Stage MTA-Cooperative PRMT5 Inhibitor, Drives Antitumor Activity Preclinically and in Patients With MTAP-Deleted Cancers (2024)	Amgen, Amgen Inc, Amgen Inc.	Australia, Belgium, Denmark	—
4	Structures of p53/BCL-2 complex suggest a mechanism for p53 to antagonize BCL-2 activity (2023)	Central South University	China	—
5	Can LLMs Solve Molecule Puzzles? A Multimodal Benchmark for Molecular Structure Elucidation (2024)	University of Notre Dame	United States	—
6	Structure of the neurotensin receptor 1 in complex with β-arrestin 1 (2020)	Stanford University School of Medicine, The University of Tokyo, Tohoku University	Japan, United States	—
7	Structure of the M2 muscarinic receptor–β-arrestin complex in a lipid nanodisc (2020)	Duke University Medical Center, Stanford University School of Medicine	United States	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Contribution 2

Claim — Contribution 2

The researcher developed the PHENIX AutoSol wizard, a Bayesian-based method for automated structure solution decision-making that has become a widely adopted standard in crystallography.

The researcher's primary contribution is the development of the PHENIX AutoSol wizard, introduced in a 2009 paper titled 'Decision-making in structure solution using Bayesian estimates of map quality.' This work stands as a seminal core publication in the field, establishing a framework for automated decision-making in structural biology. The titles indicate that this line of work addresses the challenge of determining map quality and guiding structure solution processes through Bayesian estimation, offering a systematic approach to a complex computational problem. By integrating these statistical methods into a wizard-like interface, the researcher appears to have simplified and standardized a critical step in the structure determination pipeline. The significance of this contribution is evidenced by its substantial citation count of 989, reflecting its widespread adoption and influence. Furthermore, citation analysis reveals that 97.2% of citing papers originate from independent researchers, demonstrating that the work has been embraced by the broader scientific community rather than just the researcher's immediate circle. This high degree of independent uptake underscores the tool's utility and the researcher's impact on advancing automated methods in crystallography.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 4

CORE PAPER

Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard

2009 · 989 citations (GS)

Field-normalised: 192 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2009 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	A Cas9–guide RNA complex preorganized for target DNA recognition (2015)	Max Planck Institute for Biophysical Chemistry, University of California, Berkeley	Germany, United States	—
2	HPF1 completes the PARP active site for DNA damage-induced ADP-ribosylation (2020)	MRC Laboratory of Molecular Biology, University of Oxford	United Kingdom	—
3	High-throughput screening identifies established drugs as SARS-CoV-2 PLpro inhibitors (2021)	Cleveland Clinic, Fudan University, Shanghai Advanced Research Institute	Australia, China, United States	—
4	ZP2 cleavage blocks polyspermy by modulating the architecture of the egg coat (2024)	ESRF - The European Synchrotron, Karolinska Institutet, Research Institute for Microbial Diseases, Osaka University	France, Japan, Sweden	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology* / *Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Contribution 3

Claim — Contribution 3

The researcher developed phenix.refine, a seminal tool for automated crystallographic structure refinement that has become a standard resource in biological crystallography, evidenced by over 6,000 citations.

The researcher's primary contribution is the development of phenix.refine, introduced in a 2012 paper in Acta Crystallographica Section D. This work established a robust framework for automated crystallographic structure refinement, addressing the need for efficient and reliable computational tools in structural biology. The titles suggest this line of work focused on streamlining complex refinement processes, offering a significant methodological advance over prior manual or less integrated approaches. Although no follow-up papers by the same researcher are listed here, the core publication stands as a foundational resource. The work's significance is underscored by its extensive uptake, with 6,264 citations indicating widespread adoption across the field. Notably, 97.2% of classified citations originate from independent researchers, demonstrating that the tool has been embraced by the broader scientific community rather than just the researcher's immediate circle. This high level of independent engagement confirms the work's critical role in advancing crystallographic analysis globally.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 7

CORE PAPER

Towards automated crystallographic structure refinement with phenix.refine

2012 · Acta Crystallographica Section D: Biological Crystallography · 6,264 citations (GS)

Field-normalised: 4,894 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2012 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity (2020)	Buchmann Institute for Molecular Life Sciences, Goethe University, Goethe University, Goethe-University	Germany, Netherlands	—
2	Large language models generate functional protein sequences across diverse families (2023)	Howard Hughes Medical Institute, University of California, Berkeley, Lawrence Berkeley National Laboratory, Salesforce Research	United States	—
3	Anti-tumor efficacy of a potent and selective non-covalent KRASG12D inhibitor (2022)	Mirati Therapeutics	United States	—
4	Structure of Mpro from COVID-19 virus and discovery of its inhibitors (2020)	Beijing Institute of Microbiology and Epidemiology, Academy of Military Medical Sciences, Capital Medical University, Liaocheng University	Australia, China, United States	—
5	Molecular Dynamics Simulation for All (2018)	Stanford University	United States	—
6	Structure of the RNA-dependent RNA polymerase from COVID-19 virus (2020)	Guangxi University, Nankai University, ShanghaiTech University	Australia, China	—
7	Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM (2022)	Chan Zuckerberg Biohub, Genentech, Institute of Experimental Immunology, Euroimmun AG	Germany, United Kingdom, United States	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Lawrence Berkeley National Laboratory	United States	SCImago #530	6
Stanford University School of Medicine	United States	—	5
University of Washington	United States	SCImago #45 · THE 25 · QS 81	5
University of Cambridge	United Kingdom	SCImago #63 · THE =3 · QS 6	4
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	4
The University of Queensland	Australia	SCImago #126 · THE =80 · QS =42	3
ShanghaiTech University	China	SCImago #758	2
Duke University	United States	SCImago #115 · THE 28 · QS 62	2
Shanghai Institute of Materia Medica	China	—	2
The Scripps Research Institute	United States	SCImago #216	2

Institution	Country	World ranking	Citing papers
University of Chinese Academy of Sciences	China	SCImago #5 · QS =362	2
California Institute of Technology	United States	SCImago #449 · THE 7 · QS 10	2
Shanghai Institute of Materia Medica, Chinese Academy of Sciences	China	—	2
Walter and Eliza Hall Institute of Medical Research	Australia	SCImago #580	2
Shenzhen Third People's Hospital	China	—	2

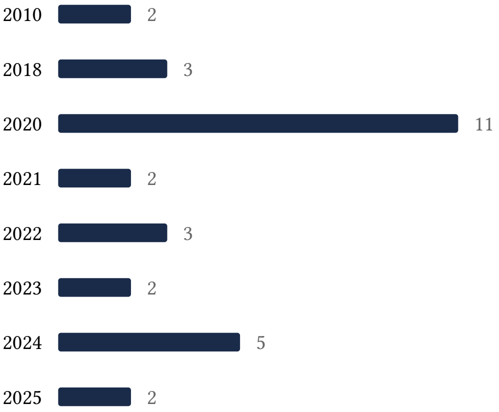
Geographic distribution of citing authors

Country	Citing papers
United States	26
United Kingdom	10
China	7
Australia	6
France	5
Germany	5
Netherlands	3
Japan	2
Belgium	2
Canada	2
Denmark	2
Austria	1

Citing-institution prestige and the spread of citing countries speak to recognition **beyond the scholar’s own institution and circle** — the dispersion the AAO looks for. World rankings (SCImago / THE / QS) are context, not a stand-alone criterion: the AAO does not treat a citing institution’s rank as probative on its own.

E. Citation Growth Over Time

Distinct citing papers by publication year. Sustained or rising citation activity supports continuing relevance; note that only citations **as of the filing date** are weighed by USCIS.



F. AAO Precedent Considerations

Pre-filing self-check (AAO denial patterns)

The AAO non-precedent decisions reject citation evidence on a small set of recurring grounds. Confirm the petition addresses each before filing:

- ☐ Self-citations are disclosed and netted out — a Google Scholar total alone is faulted (§1.1).
- ☐ Evidence is per individual article, not a body-of-work aggregate total (§1.2).
- ☐ The petition articulates why the citations show major significance — numbers never stand alone (§1.5).
- ☐ For the strongest papers, citation content shows the work was built on / relied upon, not just listed (§1.6, §2.2).
- ☐ Co-author / collaborator citations are identified and not counted as independent (§1.7).
- ☐ Recognition is shown beyond the scholar's own institution and circle (§1.8).
- ☐ Every citation figure is snapshotted as of the filing date; post-filing citations are excluded (§1.9).
- ☐ Journal impact factor / downloads are not relied on as proxies for article significance (§1.10, §1.12).
- ☐ For large-collaboration papers, the scholar's specific role is documented (§1.13).
- ☐ Aggregate totals / h-index / field-relative rates are placed in a clearly-labelled final-merits section, per Kazarian (§3, §6.1.7).

Disclaimer

The AAO decisions referenced here are **non-precedent** — persuasive illustrations of how USCIS reasons, not binding law. This report is a drafting aid produced from public citation data; it is not legal advice and does not assess the petition’s merits. All analysis must be reviewed by qualified immigration counsel.

G. Citation Evidence Index

Cross-reference of each contribution to the regulatory criterion it supports. Counsel should map these to the petition’s exhibit numbers.

Contribution	Core paper	Indep. cites	Supports
Contribution 1	Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard	18	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 2	Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard	4	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 3	Towards automated crystallographic structure refinement with phenix.refine	7	8 CFR 204.5(h)(3)(v) — Criterion 5