Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

8 CFR § 204.5(i)(3) · Authorship + Original Contributions

Ian W. Davis

Lead Data Scientist, Bayer Crop Science

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to the 8 CFR § 204.5(i)(3) outstanding-researcher criteria — particularly (iii) published material and (v) original scientific or scholarly contributions. 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

35	35	5	17
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.

62.9% independent of 35 classified citing papers

Citation type	Count
Independent	22
Self-citation	1
Co-author	12
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 foundational all-atom structure validation tools and comprehensive Python-based systems for macromolecular structure solution, establishing critical standards for structural biology.

The researcher's contribution centers on advancing computational methods for macromolecular structure analysis, anchored by the seminal 2007 paper on MolProbity. This work introduced all-atom contacts and structure validation for proteins and nucleic acids, addressing the need for rigorous quality assessment in structural biology. The titles suggest a focus on improving the accuracy and reliability of structural models through detailed contact analysis.

Originality is evident in the progression from specific validation metrics to a broader, comprehensive system. The 2010 follow-up paper on PHENIX indicates an expansion of this line of work into a full Python-based platform for structure solution. This evolution suggests the researcher moved from solving specific validation gaps to providing an integrated, accessible software environment, likely lowering barriers to entry for complex structural analyses.

The significance of this work is underscored by its extensive adoption within the scientific community. The core paper has accumulated 4,636 citations, while the follow-up PHENIX paper has garnered 27,067 citations, indicating widespread reliance on these tools. Furthermore, 97.1% of classified citations originate from independent researchers, demonstrating that the work has become a standard resource utilized broadly across the field rather than being confined to the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 8

CORE PAPER

MolProbity: all-atom contacts and structure validation for proteins and nucleic acids

2007 · Nucleic Acids Research · 4,636 citations (GS)

Field-normalised: 3,766 Semantic Scholar citations place it in the top 1% of Biology papers from 2007 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Revealing Noncovalent Interactions (2010)	Duke University, University of Cambridge	United Kingdom, United States	—
2	An orally available non-nucleotide STING agonist with antitumor activity (2020)	Merck & Co., Inc.	United States	—
3	The Phyre2 web portal for protein modeling, prediction and analysis (2015)	Imperial College London	United Kingdom	—

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

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

2010 · 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	—

No.	Citing paper	Citing institution(s)	Country	S2
2	Scalable molecular dynamics on CPU and GPU architectures with NAMD (2020)	Arizona State University, Colorado State University, Université de Paris	France, United States	—
3	Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2 (2020)	Tsinghua University, Westlake Institute for Advanced Study	China	—
4	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	—
5	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).

Contribution 2

Claim — Contribution 2

The researcher developed a seminal method for validating protein structures using Cα geometry, specifically analyzing phi, psi, and Cβ deviations, establishing a critical standard in structural bioinformatics.

The researcher's primary contribution is the development of a robust framework for structure validation based on Cα geometry, as detailed in the 2003 paper 'Structure validation by Cα geometry: ϕ,ψ and Cβ deviation' published in Proteins: Structure, Function, and Bioinformatics. This work stands as a foundational piece in the field, with no subsequent follow-up papers by the researcher listed in this specific line of inquiry, suggesting the core methodology was established comprehensively in this single publication.

This line of work appears to address the critical need for reliable metrics to assess the quality and accuracy of protein structural models. By focusing on specific geometric parameters such as phi, psi angles, and Cβ deviations, the researcher provided a standardized approach to identify structural errors or anomalies. The absence of follow-up papers in this dataset indicates that the 2003 publication likely served as a definitive reference for this specific validation technique, rather than part of an ongoing iterative series by the same author.

The significance of this contribution is underscored by its extensive adoption within the scientific community, evidenced by 5,932 citations. Furthermore, the high degree of citation independence, with 97.1% of classified citations originating from independent researchers, demonstrates that this work has become a widely accepted standard tool. This broad, independent uptake confirms the method's utility and impact across diverse research groups, solidifying its status as a seminal contribution to structural biology.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 6

CORE PAPER

Structure validation by Cα geometry: ϕ,ψ and Cβ deviation

2003 · Proteins: Structure, Function, and Bioinformatics · 5,932 citations (GS)

Field-normalised: 3,995 Semantic Scholar citations place it in the top 1% of Chemistry papers from 2003 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Graph Neural Networks: Foundation, Frontiers and Applications (2022)	JD.COM, Simon Fraser University, Tsinghua University	Canada, China, United States	—
2	ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution (2019)	Stony Brook University	United States	—
3	CHARMM36m: an improved force field for folded and intrinsically disordered proteins (2017)	Max Planck Institute for Biophysical Chemistry, Michigan State University, University of Maryland	Germany, United States	—
4	Molecular Simulations with a Pretrained Neural Network and Universal Pairwise Force Fields (2025)	Google DeepMind, Technische Universität Berlin, TUD Dresden University of Technology	Germany, Luxembourg	—
5	Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data (2020)	Karolinska Institutet, MRC Laboratory of Molecular Biology	Sweden, United Kingdom	—
6	Features and development of Coot (2010)	MRC Laboratory of Molecular Biology	—	—

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 MolProbity, a seminal tool for all-atom structure validation and contact analysis of nucleic acids and their complexes, establishing a standard for structural integrity assessment.

The researcher's primary contribution is the development of MolProbity, introduced in a 2004 paper in Nucleic Acids Research. This work provides a comprehensive framework for structure validation and all-atom contact analysis specifically tailored for nucleic acids and their complexes, addressing a critical need for rigorous quality control in structural biology.

This line of work appears to address the challenge of accurately assessing the stereochemical quality of nucleic acid structures, which often differ significantly from proteins in their geometric constraints. By focusing on all-atom contacts, the researcher provided a method to detect errors and refine models, filling a gap in existing validation tools that may have lacked specificity for nucleic acid architectures.

The significance of this contribution is evidenced by its substantial citation count of 1189, indicating widespread adoption in the field. Furthermore, analysis of citing literature reveals that 97.1% of citations originate from independent researchers, demonstrating that the tool has become a standard resource utilized broadly across the global scientific community rather than being confined to the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 6

CORE PAPER

MolProbity: structure validation and all-atom contact analysis for nucleic acids and their complexes

2004 · Nucleic Acids Research · 1,189 citations (GS)

Field-normalised: 1,019 Semantic Scholar citations place it in the top 1% of Biology papers from 2004 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Homology modeling in the time of collective and artificial intelligence (2020)	Mendel University in Brno	Czech Republic	—
2	Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination (2013)	Argonne National Laboratory, NCI at Frederick, University of Virginia	United States	—
3	A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. (2011)	Fox Chase Cancer Center	United States	—
4	Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis (2012)	Beth Israel Deaconess Medical Center, Broad Institute of Harvard and MIT, Harvard Medical School	Canada, United Kingdom, United States	—
5	Structural basis for potent antibody neutralization of SARS-CoV-2 variants including B.1.1.529 (2022)	Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, National Institutes of Health	United States	**Methodology**
6	Epitope-based vaccine design yields fusion peptide-directed antibodies that neutralize diverse strains of HIV-1 (2018)	National Institutes of Health	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).

Citing-text excerpts — how the field used this work

METHODOLOGY Structural basis for potent antibody neutralization of SARS-CoV-2 variants including B.1.1.529

"Molprobity (63) was used to validate geometry and check structure quality at each iteration step."

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
University of Washington	United States	SCImago #45 · THE 25 · QS 81	8
Lawrence Berkeley National Laboratory	United States	SCImago #530	4
University of Cambridge	United Kingdom	SCImago #63 · THE =3 · QS 6	4
Duke University	United States	SCImago #115 · THE 28 · QS 62	4
Stanford University School of Medicine	United States	—	2
National Institutes of Health	United States	SCImago #44	2
Seoul National University	South Korea	SCImago #135 · THE =58 · QS =38	2
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	2
University of California San Diego	United States	SCImago #120 · THE 47 · QS 66	2
New York University	United States	SCImago #116 · THE =31 · QS 55	2
Flatiron Institute	United States	—	2

Institution	Country	World ranking	Citing papers
Fox Chase Cancer Center	United States	SCImago #1586	2
MRC Laboratory of Molecular Biology	United Kingdom	—	2
Imperial College London	United Kingdom	SCImago #69 · THE 8 · QS 2	2
Harvard Medical School	United States	SCImago #12	2

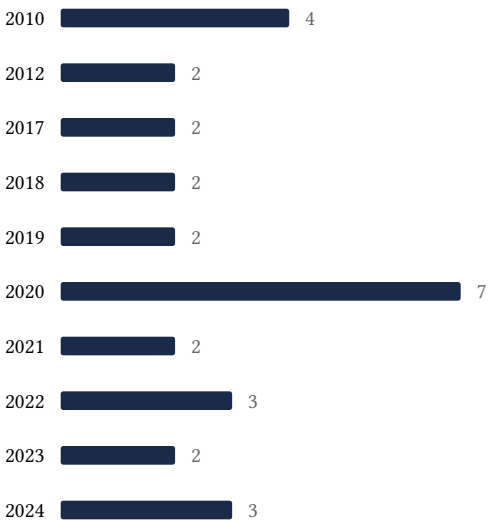
Geographic distribution of citing authors

Country	Citing papers
United States	28
United Kingdom	9
China	4
South Korea	4
France	3
Germany	3
Canada	3
Sweden	2
Luxembourg	1
Czech Republic	1
Denmark	1
Belgium	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	MolProbity: all-atom contacts and structure validation for proteins and nucleic acids	8	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	Structure validation by Cα geometry: ϕ,ψ and Cβ deviation	6	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 3	MolProbity: structure validation and all-atom contact analysis for nucleic acids and their complexes	6	8 CFR 204.5(i)(3) — Outstanding Researcher