# Citation Evidence Report

EB-1A Petition — Original Contributions of Major Significance

8 CFR § 204.5(h)(3)(v) · Criterion 5

## Chunfu Xu

National Institute of Biological Sciences; Tsinghua Institute of Multidisciplinary Biomedical Res.

Google Scholar profile

**Generated 2026-06-08 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

| 634 | 658 | 21 | 14 |
|---|---|---|---|
| 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.

**81.2% independent** of 579 classified citing papers

| Citation type | Count |
|---|---|
| Independent | 470 |
| Self-citation | 5 |
| Co-author | 104 |
| Same-institution | 0 |

55 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 pioneered computational methods for designing protein homo-oligomers with modular specificity, establishing a foundational framework for de novo protein engineering.*

The researcher's core contribution rests on the 2016 paper 'De novo design of protein homo-oligomers with modular hydrogen-bond network–mediated specificity.' This work appears to have established a novel approach for engineering protein complexes with precise structural control, serving as the foundation for subsequent advancements in the field.

This line of work addresses the challenge of creating proteins with defined quaternary structures. The titles suggest a progression from establishing modular specificity in homo-oligomers to applying these principles for broader applications, such as designing near-infrared fluorescent proteins and transmembrane pores. This indicates a systematic expansion of the initial design framework into diverse functional protein classes.

The significance of this contribution is evidenced by the core paper's 419 citations and the follow-up work on transmembrane pores, which has garnered 192 citations. With 81.2% of citing papers originating from independent researchers, the data suggests that this methodology has been widely adopted and validated by the broader scientific community as a standard tool for protein design.

**INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 148** · **7 flagged influential by Semantic Scholar**

**CORE PAPER**

**De novo design of protein homo-oligomers with modular hydrogen-bond network–mediated specificity**

2016 · 419 citations (GS)

Field-normalised: 293 Semantic Scholar citations place it in the top 1% of Chemistry papers from 2016 indexed by Semantic Scholar, by citation count.

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 1 | Large language models generate functional protein sequences across diverse families | Salesforce Research | United States | — |
| 2 | Advances in protein structure prediction and design | University of North Carolina at Chapel Hill | United States | — |
| 3 | Computational protein design | — | — | — |
| 4 | Hallucinating symmetric protein assemblies | University of Washington | United States | — |
| 5 | Artificial intelligence-aided protein engineering: from topological data analysis to deep protein language models | Michigan State University | United States | — |
| 6 | Protein design: from the aspect of water solubility and stability | Avalon GloboCare Corp., Massachusetts Institute of Technology, Shanghai Jiao Tong University | China, Norway, Russia | — |
| 7 | Recent advances in de novo protein design: Principles, methods, and applications | Chan Zuckerberg Initiative (United States), Howard Hughes Medical Institute | United States | — |
| 8 | Design of complicated all-α protein structures | The Exploratory Research Center on Life and Living Systems, The Graduate University for Advanced Studies, SOKENDAI | Japan | — |

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 9 | Breakthroughs in computational design methods open up new frontiers for de novo protein engineering | Institute for Protein Innovation | United States | — |
| 10 | De novo protein design, a retrospective | Syracuse University, University of California, San Francisco | United States | — |
| 11 | Expanding the versatility of natural and de novo designed coiled coils and helical bundles | Friedrich Miescher Laboratory, Max Planck Institute for Developmental Biology | Germany | — |
| 12 | Protein sequence design with a learned potential | Stanford University | United States | — |
| 13 | Computational design of novel protein–protein interactions–An overview on methodological approaches and applications | École Polytechnique Fédérale de Lausanne | Switzerland | — |
| 14 | What has de novo protein design taught us about protein folding and biophysics? | University of Manchester | United Kingdom | — |
| 15 | Third generation antibody discovery methods: in silico rational design | University of Cambridge | United Kingdom | — |
| 16 | Principles of protein stability and their application in computational design | Weizmann Institute of Science | Israel | — |
| 17 | Principles for computational design of binding antibodies | Google, Inc., Weizmann Institute of Science | Israel, United States | — |
| 18 | Spatial multiplexing of fluorescent reporters for imaging signaling network dynamics | Brigham and Women's Hospital, IIT@MIT, Kansas State University | United States | — |
| 19 | Deep neural language modeling enables functional protein generation across families | Hesco (United States), Salesforce Research, Salesforce (United States) | United States | — |
| 20 | Towards functional de novo designed proteins | University of Bristol | United Kingdom | — |
| 21 | Self-assembly and regulation of protein cages from pre-organised coiled-coil modules | National Institute of Chemistry, National Research Council of Italy | Italy, Slovenia | — |
| 22 | Getting momentum: from biocatalysis to advanced synthetic biology | Universität Greifswald | Germany | — |
| 23 | Understanding a protein fold: The physics, chemistry, and biology of α-helical coiled coils | University of Bristol | United Kingdom | — |
| 24 | Accessing semiaddressable self-assembly with efficient structure enumeration | Institute of Science and Technology Austria | Austria | — |
| 25 | Time-resolved spectroscopic mapping of vibrational energy flow in proteins: Understanding thermal diffusion at the nanoscale | Osaka University | Japan | — |
| 26 | The coming of age of de novo protein design | Stanford University, University of Manchester, University of Washington | United Kingdom, United States | **Influential** |

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 27 | The “cancer immunogram” | David Geffen School of Medicine at UCLA, Netherlands Cancer Institute | Netherlands, United States | — |
| 28 | Deep learning–guided design of dynamic proteins | University of California, Irvine Medical Center, Wesleyan University | United States | — |
| 29 | Bioinspired supramolecular hydrogel from design to applications | Jilin University | China | — |
| 30 | Rational protein design | Universität Greifswald | Germany | — |

**Showing the 30 most-cited of 148 independent citing papers.**

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**

### De novo Design of Near Infrared Fluorescent Proteins

2024 · 1 citations (GS)

Field-normalised: 24 Semantic Scholar citations place it in the top 5% of Chemistry papers from 2024 indexed by Semantic Scholar, by citation count.

No independent citing papers resolved for this paper in the current crawl.

**FOLLOW-UP WORK**

### Computational design of transmembrane pores

2020 · 192 citations (GS)

Field-normalised: 135 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2020 indexed by Semantic Scholar, by citation count.

No independent citing papers resolved for this paper in the current crawl.

## Contribution 2

**Claim — Contribution 2**

*The researcher designed a hyperstable 60-subunit protein icosahedron, establishing a foundational framework for engineering complex, symmetric protein assemblies with enhanced structural integrity.*

The researcher’s primary contribution is the design of a hyperstable 60-subunit protein icosahedron, as detailed in their 2016 publication. This work represents a concrete achievement in protein engineering, focusing on the construction of large, symmetric macromolecular structures. The titles indicate a focus on stability and precise subunit organization, suggesting a methodological advance in creating robust protein cages.

This line of work appears to address the challenge of assembling large protein complexes with high fidelity and stability. By targeting a 60-subunit icosahedral geometry, the researcher likely tackled the difficulty of maintaining structural coherence in multi-component systems. The absence of follow-up papers by the same researcher in this specific dataset suggests the 2016 paper stands as a seminal, self-contained contribution that defined a new standard or capability in the field.

The significance of this work is evidenced by its substantial citation record, with 574 citations indicating broad recognition. Notably, 81.2% of the citing papers originate from independent researchers, demonstrating that the scientific community has

widely adopted or built upon these findings. This high degree of independent uptake underscores the work's utility and impact beyond the researcher's immediate circle.

**INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 340** · **4 flagged influential by Semantic Scholar**

**CORE PAPER**

### Design of a hyperstable 60-subunit protein icosahedron

2016 · 574 citations (GS)

Field-normalised: 395 Semantic Scholar citations place it in the top 1% of Materials Science papers from 2016 indexed by Semantic Scholar, by citation count.

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 1 | Computational protein design | — | — | — |
| 2 | Directed Chemical Evolution of Self-Assembling Artificial Proteins Utilizing a Supramolecular System | Indian Institute of Science Education and Research | India | — |
| 3 | Protein design: from the aspect of water solubility and stability | Avalon GloboCare Corp., Massachusetts Institute of Technology, Shanghai Jiao Tong University | China, Norway, Russia | — |
| 4 | Recent advances in de novo protein design: Principles, methods, and applications | Chan Zuckerberg Initiative (United States), Howard Hughes Medical Institute | United States | — |
| 5 | GSpect: Spectral filtering for cross-scale graph classification | North University of China, South China University of Technology | China | — |
| 6 | AlphaDesign: A de novo protein design framework based on AlphaFold | European Molecular Biology Laboratory | Germany | — |
| 7 | Protein A-like peptide generation based on generalized diffusion model | Beijing Institute of Petrochemical Technology | China | — |
| 8 | De novo protein design, a retrospective | Syracuse University, University of California, San Francisco | United States | — |
| 9 | 3D nanofabrication via directed material assembly: mechanism, method, and future | Chinese University of Hong Kong, The Chinese University of Hong Kong | China, Hong Kong | — |
| 10 | The rise of metal–organic polyhedra | Government of the Republic of Korea | South Korea | — |
| 11 | Computational protein design with deep learning neural networks | East China Normal University | China | — |
| 12 | Protein nanoparticles as vaccine platforms for human and zoonotic viruses | University of Nebraska–Lincoln | United States | — |
| 13 | Self-assembling nanocarriers from engineered proteins: Design, functionalization, and application for drug delivery | Georgia Institute of Technology | United States | — |
| 14 | Coiled coil protein origami: from modular design principles towards biotechnological applications | National Institute of Chemistry, National Research Council of Italy | Italy, Slovenia | — |

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 15 | DenseCPD: improving the accuracy of neural-network-based computational protein sequence design with DenseNet | East China Normal University | China | — |
| 16 | Rational design of metalloenzymes: From single to multiple active sites | University of South China | China | — |
| 17 | Computational design of peptide assemblies | Tianjin University | China | — |
| 18 | Bionanomaterials based on protein self-assembly: Design and applications in biotechnology | Beijing Technology and Business University, China Agricultural University | China | — |
| 19 | Spatial multiplexing of fluorescent reporters for imaging signaling network dynamics | Brigham and Women's Hospital, IIT@MIT, Kansas State University | United States | — |
| 20 | The de novo design of α-helical peptides for supramolecular self-assembly | University of Bristol | United Kingdom | — |
| 21 | Self-assembly of superstructures at all scales | Chinese Academy of Sciences, University of California, Irvine Medical Center, University of California, Riverside | China, United States | — |
| 22 | Self-assembly and regulation of protein cages from pre-organised coiled-coil modules | National Institute of Chemistry, National Research Council of Italy | Italy, Slovenia | — |
| 23 | Coiled-coil-mediated assembly of an icosahedral protein cage with extremely high thermal and chemical stability | Molecular Vista (United States), University of Michigan | United States | — |
| 24 | Virus-inspired function in engineered protein cages | ETH Zurich | Switzerland | — |
| 25 | Geometric lessons and design strategies for nanoscale protein cages | University of California, Los Angeles | United States | — |
| 26 | The coming of age of de novo protein design | Stanford University, University of Manchester, University of Washington | United Kingdom, United States | — |
| 27 | Nanoreactor design based on self-assembling protein nanocages | Beijing University of Chemical Technology, National Medical Products Administration | China | — |
| 28 | A supramolecular system mimicking the infection process of an enveloped virus through membrane fusion | Tottori University | Japan | — |
| 29 | Protein-Based Controllable Nanoarchitectonics for Desired Applications | Shaanxi Normal University, Xi'an University of Posts & Telecommunications | China | — |
| 30 | Hierarchical materials from high information content macromolecular building blocks: construction, dynamic interventions, and prediction | Pacific Northwest National Laboratory, University of Chicago, University of Tennessee | United States | — |

**Showing the 30 most-cited of 340 independent citing papers.**

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 established a framework for designing helical bundles with high thermodynamic stability, a foundational contribution evidenced by substantial independent scholarly uptake.*

The researcher's core contribution rests on the 2014 publication regarding the high thermodynamic stability of parametrically designed helical bundles. This work appears to address the challenge of engineering stable protein structures through computational design, offering a robust method for creating helical assemblies that maintain structural integrity under thermodynamic stress.

By focusing on parametric design, this line of work suggests a novel approach to predicting and optimizing stability in synthetic biology and protein engineering. The absence of follow-up papers by the same researcher in this specific dataset indicates that the 2014 paper stands as a seminal, self-contained contribution that defined a key benchmark or methodology in the field.

The significance of this work is underscored by its citation record, with 370 citations indicating strong recognition within the scientific community. Notably, 81.2% of the citing papers originate from independent researchers, suggesting that the methodology or findings have been widely adopted and validated by the broader field rather than remaining confined to the researcher's immediate circle.

**INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 0**

**CORE PAPER**

**High thermodynamic stability of parametrically designed helical bundles**

2014 · 370 citations (GS)

Field-normalised: 272 Semantic Scholar citations place it in the top 1% of Chemistry papers from 2014 indexed by Semantic Scholar, by citation count.

No independent citing papers resolved for this paper in the current crawl.

## 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 | 108 |
| University of California, Irvine Medical Center | United States | — | 28 |
| Howard Hughes Medical Institute | United States | SCImago #84 | 28 |
| University of Bristol | United Kingdom | SCImago #478 · THE =80 · QS 51 | 17 |
| Fred Hutchinson/University of Washington/Seattle Children's Cancer Consortium | United States | — | 14 |
| University of California, Los Angeles | United States | SCImago #70 · THE =18 · QS 46 | 12 |
| National Institute of Chemistry | Slovenia | — | 12 |
| ETH Zurich | Switzerland | THE 11 · QS 7 | 11 |
| University of Oxford | United Kingdom | SCImago #26 · THE 1 · QS 4 | 11 |
| University of Michigan | United States | SCImago #43 · THE 23 · QS 45 | 10 |

| Institution | Country | World ranking | Citing papers |
|---|---|---|---|
| University of North Carolina at Chapel Hill | United States | THE 78 · QS =140 | 10 |
| Osaka University | Japan | SCImago #546 · QS 91 | 9 |
| Stanford University | United States | SCImago #18 · THE =5 · QS 3 | 9 |
| Tottori University | Japan | SCImago #4167 · THE 1501+ | 9 |
| The Ohio State University | United States | THE =108 · QS 190 | 8 |

### Geographic distribution of citing authors

| Country | Citing papers |
|---|---|
| United States | 291 |
| China | 99 |
| United Kingdom | 59 |
| Japan | 44 |
| Switzerland | 31 |
| Germany | 28 |
| France | 17 |
| Poland | 14 |
| Netherlands | 14 |
| Slovenia | 13 |
| South Korea | 11 |
| India | 9 |

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.

## 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 | De novo design of protein homo-oligomers with modular hydrogen-bond network–mediated specificity | 148 | 8 CFR 204.5(h)(3)(v) — Criterion 5 |
| Contribution 2 | Design of a hyperstable 60-subunit protein icosahedron | 340 | 8 CFR 204.5(h)(3)(v) — Criterion 5 |
| Contribution 3 | High thermodynamic stability of parametrically designed helical bundles | 0 | 8 CFR 204.5(h)(3)(v) — Criterion 5 |