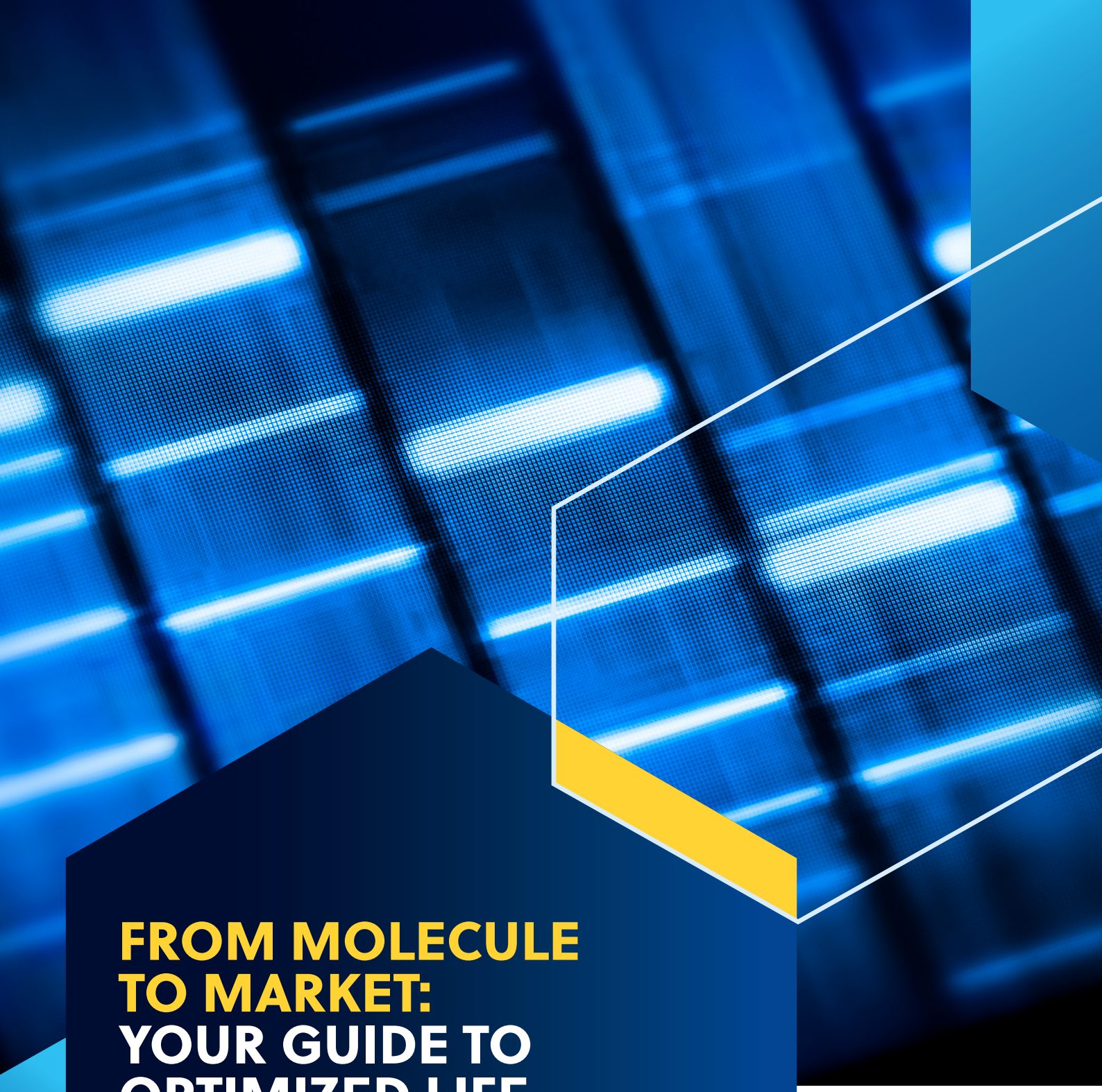


FROM MOLECULE TO MARKET: YOUR GUIDE TO OPTIMIZED LIFE SCIENCES AND BIOTECH IP SEARCH RESULTS

CAS
A division of the
American Chemical Society

Innovation hinges on comprehensive intellectual property (IP) strategies in the evolving life sciences and biotech industries. The key to confident decision-making lies in thorough, accurate patent search results, whether you are evaluating patentability, assessing freedom to operate (FTO), or tracking bioactivity data.

This guide equips life sciences and biotech IP searchers with strategies, tools, and tips to ensure precise, efficient searches that fuel innovation and mitigate risk.



Strategic IP searches: The foundation of innovation and risk management in life sciences

IP searches serve as the cornerstone of innovation in life sciences industries. From small molecules to biologics, competitive intelligence helps you identify opportunities, avoid duplicate efforts, and ensure freedom to operate. The ability to navigate complex patent landscapes and uncover critical insights can make or break an organization's strategy.

The stakes are high. Missing prior art or failing to map the competitive landscape accurately can lead to wasted R&D resources, legal disputes, or unprotected innovations. These risks are magnified as the volume of scientific and patent data grows, with patents becoming more data-dense each year.



Real-world challenges

Modern IP professionals in scientific fields face significant hurdles.

Data fragmentation: Critical data is scattered across patent and scientific databases, grant disclosures, clinical trial results, and niche industry reports. Limited integration between patent and non-patent literature creates blind spots, making it difficult to uncover all relevant prior art, emerging technologies, and potential risks.

Limited search functions: Basic filters and rigid search parameters lack the precision and flexibility needed for complex scientific IP searches. Search engines often struggle with non-standard terminologies, chemical structures, biological sequences, and data embedded in images and figures, forcing users to sift through irrelevant results or miss key findings.

Inefficient workflows: Disconnected systems and manual processes consume valuable time and resources, diverting focus from strategic analysis. Searchers must reconcile inconsistencies across siloed resources—merging data from multiple formats and platforms—before validating and sharing meaningful insights. These inefficiencies increase the risk of oversight, slow decision-making, and hindered innovation.

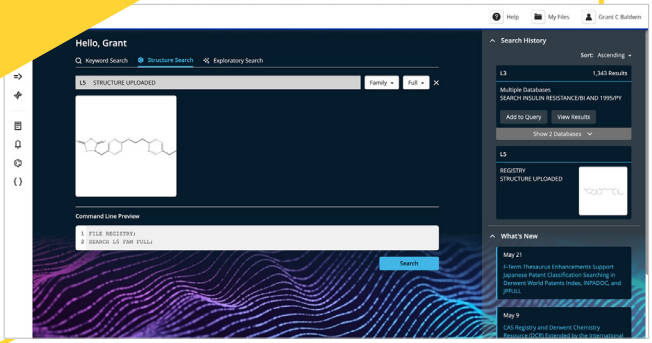
Overcome challenges with comprehensive life sciences content, search capabilities, and powerful display options

CAS IP Finder, powered by STN™, connects searchers to the world's largest collection of normalized and standardized life sciences content, including SAR, ADME, and toxicity data.

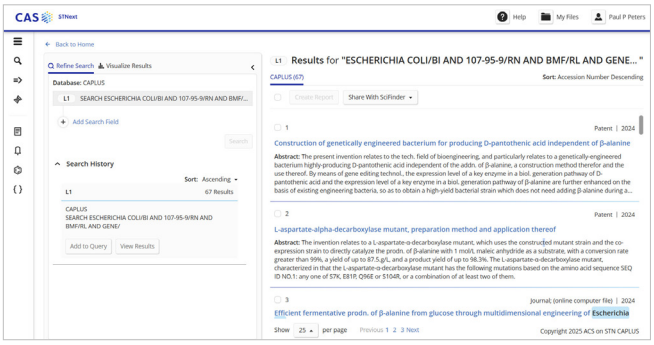
By integrating 100+ global scientific, patent, and technical data sources into a single, intuitive platform, CAS IP Finder™ delivers the tools you need to streamline workflows, uncover critical insights, and make confident decisions.

The platform offers the only unified, complete access to these databases for the most accurate and timely search results. The content is supported with new search and display capabilities, allowing users to search bioactivity content with an unprecedented level of accuracy and precision.

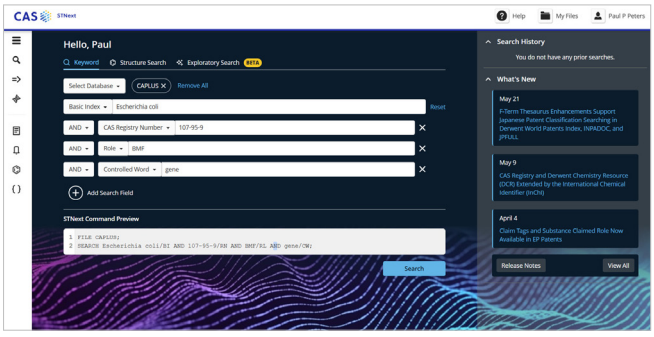
Thanks to a powerful new query builder and modernized interface, searchers can execute keyword lookups quickly, perform point-and-click searches, and use structure-drawing tools to complement or even replace traditional command-line searches in CAS IP Finder.



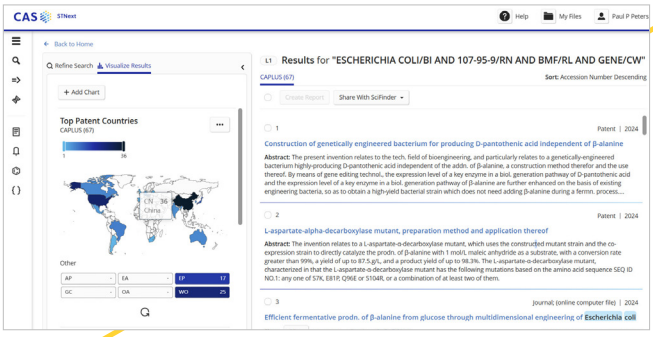
Best-in-class structure search



Intuitive results viewer



Query builder for point-and-click search



Visualization tools



Essential searches, databases, and tools for a comprehensive life sciences search strategy

Search type	Target information	Example IP applications
Nucleic acid sequence	DNA/RNA sequences, including genes, genomes, or nucleotide patterns	Validating novelty in gene therapy patents, ensuring freedom to operate for CRISPR innovations, or tracking new mRNA vaccine developments
Amino acid sequence	Protein sequences, such as enzymes, therapeutic antibodies, or fusion proteins	Assessing patentability for monoclonal antibodies, comparing protein variants, or discovering competitive overlaps in biosimilars
Bioactivity	Functional, SAR, ADME, and toxicity data for molecules or compounds of interest	Investigating structure-activity relationships, assessing compound toxicity, mapping ADME data to chemical entities, or supporting claims with bioactivity evidence
Chemical structure search	Exemplified chemical structures or generic Markush representations to identify related compound classes	Exploring unclaimed chemical spaces in drug discovery or mapping prior art for enzyme inhibitors, small-molecule biologics, etc.
Biological species	Specific biological species or organisms, including pathogens or model organisms	Identifying species-specific innovations in agricultural biotech, tracking pathogen data for vaccine development, or exploring prior art in microbiology-related patents

Nucleic acid and amino acid sequence search

Sequence searchable databases

CAS IP Finder provides access to multiple sequence databases with unsurpassed global content, covering millions of nucleotide and amino acid sequences from patents and scientific literature. This often hard-to-find data is now accessible through multiple search pathways for more efficient, accurate results.

You can complete comprehensive and precise multi-database sequence searches across leading scientific content collections, including:

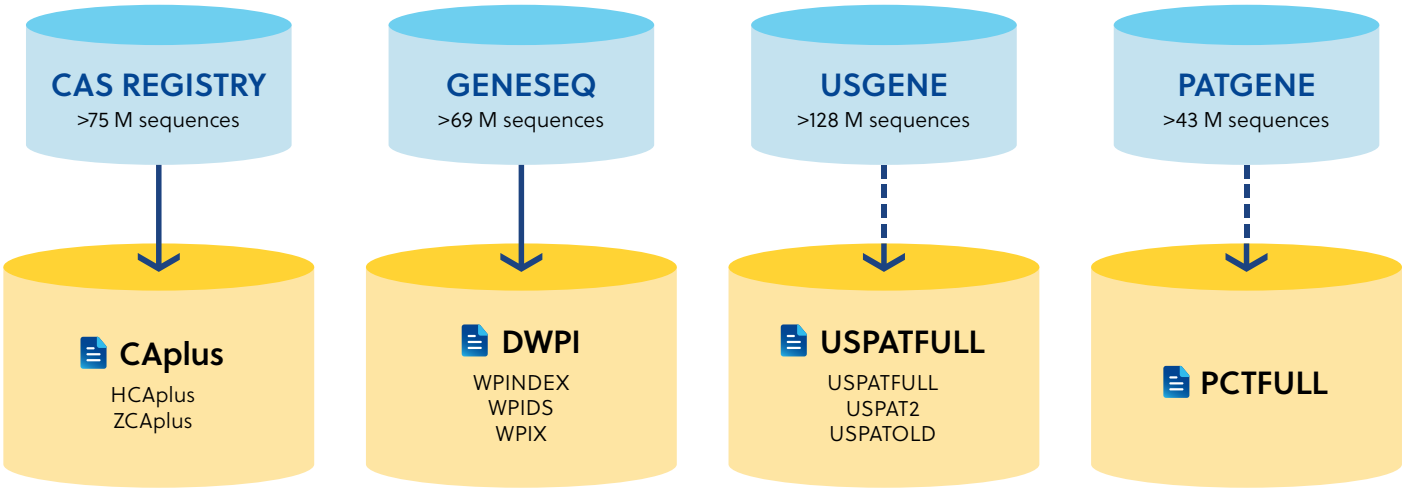
CAS REGISTRY®: The world's largest curated database of chemical substances, offering access to detailed information on over 290 million molecules, including their structures, properties, and roles.

GENESEQ: Annotated protein and nucleic acid sequences from basic patent documents across 50+ patent authorities in the Derwent World Patents Index (DWPI).

PATGENE: A specialized database providing access to all peptide and nucleic acid sequences submitted electronically to the World Intellectual Property Organization (WIPO) by patent applicants.

USGENE: A focused repository of biological sequence (DNA, RNA, and protein) data from published patent applications and issued patents of the United States Patent and Trademark Office (USPTO).

Link sequence files from sequence searchable databases to bibliographic files for more effective searches



Sequence search methods

BLAST
NCBI BLAST for advanced similarity searching

GETSIM
FASTA-based search algorithm for comparison of the entire sequence length; more computational time required. Not available in CAS REGISTRY

GETSEQ
Sequence Code Match (SCM) for simple sequence queries to find exact sequences or with controlled variations

Sequence Code Match searching (SCM)
Search the sequence query as is or with some variation using special symbols

- Exact (the same sequence and length)
- Exact Family (peptide searches only)
- Subsequence (query sequence embedded in a larger sequence)
- Subsequence Family (peptide searches only)

Bioactivity search

The CAS Bioactivity file simplifies the searching process by streamlining many required steps.

With CAS Bioactivity accessed through CAS IP Finder, searchers can:

- Stay in one tool for increased efficiency.
- Search against a specialized bioactivity file with comprehensive coverage and robust search options.
- Search against content that has already been aggregated and integrated with wider content resources.
- Pinpoint the location of relevant data in primary literature, reducing the time spent on manual reviews.



Chemical structure search

Use exact or substructure queries based on goals:

- Use exact structure queries for precise results.
- Opt for substructure searches to capture derivatives or modifications for broader insights.

Leverage Markush searches to avoid missing important results:

- Include Markush structures to search for patent claims covering families of related compounds.
- Perform searches in complementary Markush databases such as MARPAT and the Derwent Markush Resource (DWPIM).

Use advanced query refinement features:

- Apply chemical drawing tools to create or modify structural nodes directly within the search interface.
- Combine chemical structure searches with sequence or text-based queries for a multi-dimensional approach.

Check contextual relevance:

- Cross-reference results with bioactivity data to understand the functional implications of a structure.
- Review linked patent claims to see how the structure is described or protected.

Optimize for workflow efficiency:

- Save frequently used structures as templates for recurring searches.
- Leverage automated alerts to track updates to relevant structures or patents.



Biological species search

The same organism might be known under many different names. An example of this is the common laboratory research model *Saccharomyces cerevisiae*, which is also referred to as Baker's Yeast, Fermipan, and Biosaf, among other names. Take advantage of advanced search capabilities in CAS IP Finder to avoid missing results:

- Use the query builder or look up database-specific controlled term hierarchies to create targeted queries that include scientific names, abbreviated forms, and any known product or lab jargon names.
- Search controlled terms for precise matches and use relationship codes to include further taxonomic levels.
- Consider searching on a broader taxonomic level, such as the genus or family.
- Try multiple search pathways, including keyword search and point-and-click query refining, for quick yet comprehensive searches.

Key features and tips for comprehensive life sciences IP searches

Optimize search strategies

- Leverage multiple search pathways by using the query builder, keyword search, command line, and structure searches to maximize precision and coverage.
- Refine queries with CAS Thesauri on STN® and keyword indexing for domain-specific results.

Enhance workflow efficiency

- Eliminate manual data aggregation and focus on strategic analysis.
- Automate scripts, revisit searches easily, and use flexible search pathways to reduce effort and time, enabling faster decision-making.

Access comprehensive databases

- Use the CAS Content Collection™ for curated scientific content from top journals, patents, and non-patent literature.
- Search specialized databases for bioactivity, sequences, and biochemical pathways.

Build collaborative insights

- Share search results with cross-functional teams for seamless collaboration.
- Leverage on-demand support and training from CAS experts to enhance workflows.

Take the next step toward smarter life sciences IP strategies

CAS IP Finder empowers life sciences professionals with precision tools and integrated insights, enabling confident decision-making at every stage of innovation. Its intuitive workflows, comprehensive data coverage, and advanced search tools make it an indispensable asset for navigating complex IP landscapes.

Explore CAS IP Finder and discover the tools that make innovation possible.

Learn more about CAS IP Finder at link.cas.org/ip-finder



CAS
A division of the
American Chemical Society

© 2025 American Chemical Society. All rights reserved.
IPFWHP102653ENG250908