



CAS STNext[®] E-SEMINAR

COMPREHENSIVE PRIOR ART IN CHEMISTRY: WHAT YOU MIGHT MISS

Jim Brown, FIZ Karlsruhe

© 2023 American Chemical Society. All rights reserved.

 **FIZ Karlsruhe**
Leibniz Institute for Information Infrastructure

CAS 
A division of the
American Chemical Society

Agenda

- Know your topic
- Know your tools
- Know your databases
- Develop the search strategy
- Dates, dates, dates

Know your topic

- ‘*Tell me what you want and I’ll give you what you need*’ - The Doobie Brothers
- A good search strategy begins with truly understanding the topic, both specifically and generally
 - Learn about the background of the topic
 - In-depth discussions with your requester
 - Before AND as the search is ongoing

Importance of knowing your topic

- Ability to use tools to their fullest potential
- Keywords – relevant synonyms, related terms
- Classifications – relevant vs. irrelevant
- Value-add tools
 - Index terms/structures/sequences/polymers

Know your topic

- Do not be afraid to have ongoing conversations with requester!
- Use requester's expertise to your advantage
- When possible, use colleagues' expertise as well
- Inform requester of ongoing results/pitfalls
- Empowers requester as well

Know your tools

- Possible tools (which will vary based on search request)
- Keyword search
- Classification search
 - International Patent Classifications (IPCs)
 - Cooperative Patent Classifications (CPCs)
 - Japanese Patent Classifications (FCLs)
 - Japanese F-Terms (FTERMs)
 - European Patent Classifications (ECLA)*
 - In computer only Patent Classifications (ICOs)*
 - US Patent Classifications (NCLs)*

* These classes have been retired but are still available in some databases for historical searching

Know your tools (cont.)

- Value-add abstracting
 - Knowledgeable experts
- Value-add indexing
 - Index terms
 - Control Terms, Supplementary Terms (CAplus)
 - Derwent Classes, Derwent Manual Codes (DWPI)
 - Specific substance indexing (CAS REGISTRY®, Derwent Chemistry Resource (DCR), **Derwent Markush Resource (DWPIIM)**, DWPI fragmentation, ReaxysFileSub)
 - Markush indexing (MARPAT, DWPIIM)
 - Sequence indexing (CAS REGISTRY, **MARPAT**, GENESEQ, **DCR**, **DWPIIM**, USGENE, PATGENE)
 - Polymer indexing (CAS REGISTRY, **MARPAT**, Derwent World Patents Index (WPIDS, WPIX), **DCR**, **DWPIIM**)

Know your tools

- Things to consider
 - When did this term first come into being?
 - How would this concept have been covered before this term existed?
 - Concordances, reclassifications, back postings?
 - Are there other relevant terms in that system?
 - Broader, narrower and related terms

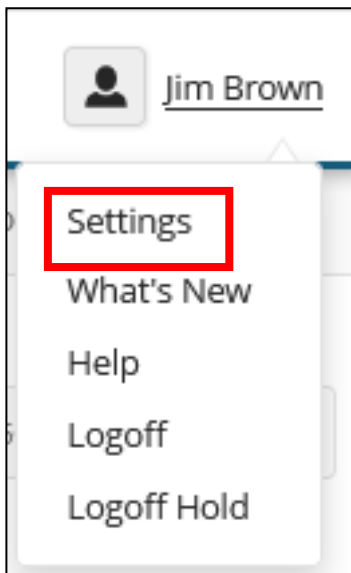
Know your databases

- Coverage
 - Country coverage
 - Date coverage
 - How far back, how current
- Document coverage
 - Areas of technology, published applications, granted patents, design, plant ...
- First-level vs. value-add

Know your tools – Keyword searching on CAS STNNext using CAplus and DWPI

- Consider where to search keyword(s)
 - Title, enhanced abstract(s), first level data
 - May be in multiple languages
 - Basic Index vs. Basic Index, Extended
- Settings associated with your STN ID
 - Plurals, spellings, abbreviations
 - Great tools, but beware of unforeseen consequences
 - Abbreviations that have more than one meaning
 - Words in another language than the one considered

Settings



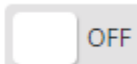
Abbreviation	ON <input type="checkbox"/>
Audit	ON <input type="checkbox"/>
Logoff History (LHIST)	ON <input type="checkbox"/>
Multiple-step (Mstep)	ON <input type="checkbox"/>
Plurals	ON <input type="checkbox"/>
Spellings	ON <input type="checkbox"/>
Structure Editor Toolbars	Floating ▾

Cost Notification (USD)

Display Cost Notifications	ON <input type="checkbox"/>
Display Cost Trigger	<input type="text" value="1"/>
Search Cost Notifications	ON <input type="checkbox"/>
Search Cost Trigger	<input type="text" value="1"/>

Plurals

Plurals



=> S BICYCLE

L1 169343 BICYCLE

=> S MOUSE

L3 131416 MOUSE

Plurals



=> S BICYCLE

169343 BICYCLE

15153 BICYCLES

L2 174493 BICYCLE

(BICYCLE OR BICYCLES)

=> S MOUSE

131416 MOUSE

134 MOUSES

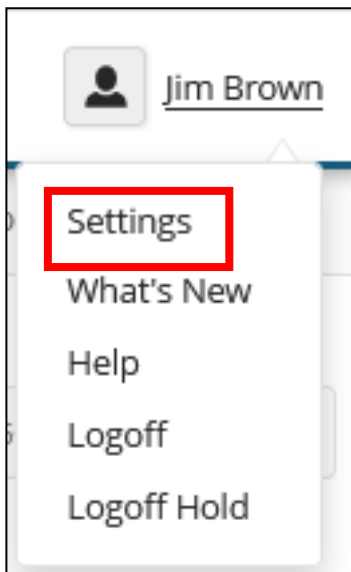
46834 MICE

13 MICES

L4 163509 MOUSE

(MOUSE OR MOUSES OR MICE OR MICES)

Settings (cont.)



Abbreviation	ON <input type="checkbox"/>
Audit	ON <input type="checkbox"/>
Logoff History (LHIST)	ON <input type="checkbox"/>
Multiple-step (Mstep)	ON <input type="checkbox"/>
Plurals	ON <input type="checkbox"/>
Spellings	ON <input type="checkbox"/>
Structure Editor Toolbars	Floating ▾

Cost Notification (USD)

Display Cost Notifications	ON <input type="checkbox"/>
Display Cost Trigger	<input type="text" value="1"/>
Search Cost Notifications	ON <input type="checkbox"/>
Search Cost Trigger	<input type="text" value="1"/>

Abbreviation

Plurals ON

Abbreviation OFF

Plurals ON

Abbreviation ON

```
=> S ALUMINUM  
  
1195676 ALUMINUM  
101 ALUMINUMS  
L1 1195709 ALUMINUM  
      (ALUMINUM OR ALUMINUMS)
```

What is the problem with this search term?

```
=> S ALUMINUM  
  
1195676 ALUMINUM  
101 ALUMINUMS  
1195709 ALUMINUM  
      (ALUMINUM OR ALUMINUMS)  
238027 AL  
5440 ALS  
243304 AL  
      (AL OR ALS)  
L2 1401791 ALUMINUM  
      (ALUMINUM OR AL)
```

Keyword searching potential issues

```
=> S ALS NOT (ALUMINUM OR ALUMINIUM OR AL)
```

```
6052 ALS/BI
```

```
1080786 ALS/BIEX
```

```
1301247 ALUMINUM/BI
```

```
583549 ALUMINUM/BIEX
```

```
848511 ALUMINIUM/BI
```

```
1206932 ALUMINIUM/BIEX
```

```
242703 AL/BI
```

```
497871 AL/BIEX
```

```
L1 1009758 ALS/BI,BIEX NOT (ALUMINUM/BI,BIEX OR ALUMINIUM/BI,BIEX OR AL/BI,  
BIEX)
```

This particular search query was conducted with the plurals and abbreviations turned off.

```
L1 ANSWER 1 OF 1009758 WPINDEX COPYRIGHT 2023 CLARIVATE on STN
```

```
Member. . .
```

```
einer Aussenscheibe (2), einer Innenscheibe (3) und einer  
thermoplastischen Zwischenschicht (4), (b) Ein- oder Aufbringen  
mindestens eines opaken Hintergrunds, insbesondere als opake Schicht,  
bevorzugt als Maskierungsstreifen (5,5'), in zumindest einem zweiten  
Teilbereich der innenraumseitigen Oberflaeche der Aussenscheibe (II), der
```

Keyword searching potential issues

L1 ANSWER 4 OF 1009758 WPINDEX COPYRIGHT 2023 CLARIVATE on STN

Member. . .

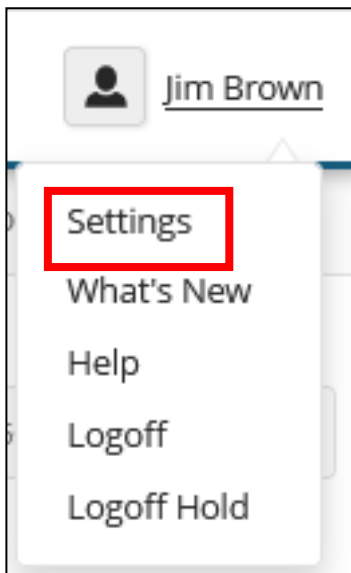
algorithm The rank of tensor RS , the value of the learning rate rn , and the tensor f by using the CP-ALS algorithm. performing decomposition to obtain initial factor matrixes $S(1)$, $S(2)$ and $S(3)$; (4) constructing regularized target optimization. . .

L2 ANSWER 3 OF 6252 WPINDEX COPYRIGHT 2023 CLARIVATE on STN

USE . . . for paralyzed patients having neurological diseases. Uses include but are not limited to brainstem infarcts, brain injury, stroke and advanced amyotrophic lateral sclerosis (ALS).

These types of issues are usually not a problem. When the aluminum concept is used in conjunction with other concepts, the false drops would decrease/disappear, but the experienced searcher would keep an eye on this type of issue.

Settings (cont.)



Abbreviation	ON <input type="checkbox"/>
Audit	ON <input type="checkbox"/>
Logoff History (LHIST)	ON <input type="checkbox"/>
Multiple-step (Mstep)	ON <input type="checkbox"/>
Plurals	ON <input type="checkbox"/>
Spellings	ON <input type="checkbox"/>
Structure Editor Toolbars	Floating ▾
Cost Notification (USD)	
Display Cost Notifications	ON <input type="checkbox"/>
Display Cost Trigger	<input type="text" value="1"/>
Search Cost Notifications	ON <input type="checkbox"/>
Search Cost Trigger	<input type="text" value="1"/>

Spellings

Plurals

ON

Abbreviation

ON

Spellings

OFF

=> S ALUMINUM

```
1195676 ALUMINUM
  101 ALUMINUMS
1195709 ALUMINUM
      (ALUMINUM OR ALUMINUMS)
238027 AL
  5440 ALS
243304 AL
      (AL OR ALS)
L2  1401791 ALUMINUM
      (ALUMINUM OR AL)
```

Plurals

ON

Abbreviation

ON

Spellings

ON

=> S ALUMINUM

```
1195676 ALUMINUM
  101 ALUMINUMS
1195709 ALUMINUM
      (ALUMINUM OR ALUMINUMS)
779399 ALUMINIUM
  56 ALUMINIUMS
779424 ALUMINIUM
      (ALUMINIUM OR ALUMINIUMS)
1476662 ALUMINUM
      (ALUMINUM OR ALUMINIUM)
238027 AL
  5440 ALS
243304 AL
      (AL OR ALS)
L3  1607088 ALUMINUM
      (ALUMINUM OR AL)
```

Spellings

Plurals ON

Abbreviation ON

Spellings OFF

```
=> S DIAPER  
  
41907 DIAPER  
12407 DIAPERS  
L1 48219 DIAPER  
      (DIAPER OR DIAPERS)
```

Turning the Spellings ON also finds non-ambiguous synonyms.

Plurals ON

Abbreviation ON

Spellings ON

```
=> S DIAPER  
  
41907 DIAPER  
12407 DIAPERS  
48219 DIAPER  
      (DIAPER OR DIAPERS)  
3515 NAPPY  
2251 NAPPIES  
5286 NAPPY  
      (NAPPY OR NAPPIES)  
L2 52518 DIAPER  
      (DIAPER OR NAPPY)
```

Basic Index

- Search a keyword without a search qualifier, the system will search the 'basic index' of that database
- The basic index changes from database to database
- Most searches require the searcher to tell the system where to look, by using a search qualifier
 - Inventors - /IN
 - Assignees - /PA

Basic index for DWPI

Search Field Name	Search Code	Search Examples	Display Codes
Basic Index* (contains single words from AB, ABDT, ABEQ, ABEX, ACTN, ACTV, ADV, ALE, DETD, DRWD, NOV, TECH, TI, TT, UADV, and USE) (18)	None or /BI (or /BIX)	S DRILLING FLUID AND EMULS? S ?PHENYLETHER? S #####DIPHENYLETHER S ULCER TREATMENT(L)ORAL	AB, ABDT, ABEQ, ABEX, ACTN, ACTV, ADV, ALE, DETD, DRWD, NOV, TECH, TI, TT, UADV, USE

- Value-add data
- Derwent enhanced titles and abstracts
- Search with no search qualifier or with **/BI** search qualifier

Basic index Extended for DWPI

Search Field Name	Search Code	Search Examples	Display Codes
Basic Index Extended* (contains single terms from author abstracts, claims, and original titles) (2)	/BIEX	S NANOCLUSTERS/BIEX	ABDE, ABEN, ABES, ABFR, ABOL, CLMDE, CLMEN, CLMES, CLMFR, CLMOL, TIDE, TIEN, TIES, TIFR, TIOL

- Available first level data
- ~ 20 patent authorities
- Original titles, original abstracts, claims
- Search with **/BIEX** search qualifier
- To have the system search both fields permanently, use the command **SET SFIELDS BI BIEX PERM**

BI and BLEX keyword search in DWPI

=> S ALUMINUM

```
1195676 ALUMINUM
  101 ALUMINUMS
1195709 ALUMINUM
      (ALUMINUM OR ALUMINUMS)
779399 ALUMINIUM
  56 ALUMINIUMS
779424 ALUMINIUM
      (ALUMINIUM OR ALUMINIUMS)
1476662 ALUMINUM
      (ALUMINUM OR ALUMINIUM)
238027 AL
  5440 ALS
243304 AL
      (AL OR ALS)
L3 1607088 ALUMINUM
      (ALUMINUM OR AL)
```

=> SET SFIELDS BI BLEX PERM

SET COMMAND COMPLETED

=> S ALUMINUM

```
1195676 ALUMINUM/BI
  101 ALUMINUMS/BI
1195709 ALUMINUM/BI
      ((ALUMINUM OR ALUMINUMS)/BI)
779399 ALUMINIUM/BI
  56 ALUMINIUMS/BI
779424 ALUMINIUM/BI
      ((ALUMINIUM OR ALUMINIUMS)/BI)
1476662 ALUMINUM/BI
      ((ALUMINUM OR ALUMINIUM)/BI)
238027 AL/BI
  5440 ALS/BI
243304 AL/BI
      ((AL OR ALS)/BI)
1607088 ALUMINUM/BI
      ((ALUMINUM OR AL)/BI)
```

```
555337 ALUMINUM/BIEX
  140 ALUMINUMS/BIEX
555375 ALUMINUM/BIEX
      ((ALUMINUM OR ALUMINUMS)/BIEX)
1072697 ALUMINIUM/BIEX
  815 ALUMINIUMS/BIEX
1072952 ALUMINIUM/BIEX
      ((ALUMINIUM OR ALUMINIUMS)/BIEX)
1403137 ALUMINUM/BIEX
      ((ALUMINUM OR ALUMINIUM)/BIEX)
456363 AL/BIEX
1058977 ALS/BIEX
1487629 AL/BIEX
      ((AL OR ALS)/BIEX)
2779659 ALUMINUM/BIEX
      ((ALUMINUM OR AL)/BIEX)
L4 3319907 ALUMINUM/BI,BIEX
```

Basic index for CPlus

Search Field Name	Search Code	Search Examples	Display Codes
Basic Index * (contains single words from title (TI), supplementary term (ST), index term (IT), and abstract (AB) fields, as well as CAS Registry Numbers)	None (or /BI or /IA)	S 50-21-5 S TRANSGENIC COTTON S ?FLUOROCARBON? S (WATER(S)OIL)/BI	AB, IT, ST, TI

- Value-add data
- CAS enhanced titles, abstracts and index terms
- Search with no search qualifier or with **/BI** search qualifier

Basic index Extended for CPlus

Search Field Name	Search Code	Search Examples	Display Codes
Basic Index * (contains single words from title (TI), supplementary term (ST), index term (IT), and abstract (AB) fields, as well as CAS Registry Numbers)	None (or /BI or /IA)	S 50-21-5 S TRANSGENIC COTTON S ?FLUOROCARBON? S (WATER(S)OIL)/BI	AB, IT, ST, TI
Basic Index plus Claims *	/BI,BIEX or /BI,CLM	S ALLOPURINOL/BI,BIEX S TRANSGENIC/BI,CLM(W)COTTON/BI,CLM	BIB CLM ALL CLM

- Original claims
- Search with **/BIEX** search qualifier
- To have the system search both fields permanently, use the command **SET SFIELDS BI BIEX PERM**

Classifications

- All of the current classification schemes get updated
- Older records may get reclassified
- Classification concordances
 - i.e., [CPC Concordances | Cooperative Patent Classification](#) covers ECLA to CPC and CPC to IPC
- Retired classification schemes do not get updated but may still be useful
- JP classifications – unique to JP documents, may provide additional search results

Classifications (cont.)

- To find relevant classifications
- Create preliminary search set
 - Keyword search, other tools
- ANALYZE set by classification scheme(s) of interest
- Use CAS STNext classification thesauri to determine quality of term
 - Consider Broader, Narrower, Related Terms

Classification example

- This example was run in Derwent World Patents Index database
- Plurals, abbreviations and spellings on
- Search fields are the basic index and the basic index extended

```
=> S ANTIBODY DRUG CONJUGATE  
  
217011 ANTIBODY/BI  
91560 ANTIBODIES/BI  
246242 ANTIBODY/BI  
      ((ANTIBODY OR ANTIBODIES)/BI)  
339002 DRUG/BI  
109531 DRUGS/BI  
388124 DRUG/BI  
      ((DRUG OR DRUGS)/BI)  
77158 CONJUGATE/BI  
10706 CONJUGATES/BI  
80742 CONJUGATE/BI  
      ((CONJUGATE OR CONJUGATES)/BI)
```

```
2588 ANTIBODY DRUG CONJUGATE/BI  
      ((ANTIBODY(W)DRUG(W)CONJUGATE)/BI)  
171692 ANTIBODY/BIEX  
73873 ANTIBODIES/BIEX  
194741 ANTIBODY/BIEX  
      ((ANTIBODY OR ANTIBODIES)/BIEX)  
366856 DRUG/BIEX  
77730 DRUGS/BIEX  
401188 DRUG/BIEX  
      ((DRUG OR DRUGS)/BIEX)  
80065 CONJUGATE/BIEX  
16867 CONJUGATES/BIEX  
86445 CONJUGATE/BIEX  
      ((CONJUGATE OR CONJUGATES)/BIEX)  
2985 ANTIBODY DRUG CONJUGATE/BIEX  
      ((ANTIBODY(W)DRUG(W)CONJUGATE)/BIEX)  
L1 3292 ANTIBODY DRUG CONJUGATE/BI,BIEX
```

The abbreviation ADC was not created by the system, nor was it included in the search strategy as it gave a lot of false drops. For example, ADC also stands for analog to digital converter.

ANALYZE the CPCs and IPCs

=> ANALYZE L1 IPC CPC 1-3292

L2 ANALYZE L1 1-3292 IPC CPC : 5728 TERMS

=> D CPC 1-

[View all](#)

TERM #	# OCC	# DOC	% DOC	IPC CPC
5	2018	1592	48.36	A61P0035-00
8	1482	1076	32.69	A61K0047-6803
10	1057	281	8.54	A61K2300-00
11	1053	861	26.15	A61K2039-505
13	1045	773	23.48	A61K0047-6849
18	742	623	18.92	C07K2317-24
20	723	589	17.89	A61K0047-6851
22	678	563	17.10	C07K2317-92
23	655	500	15.19	A61K0045-06
24	646	489	14.85	A61K0047-6889
30	572	486	14.76	C07K2317-565
33	453	353	10.72	C07K0016-30
34	442	382	11.60	C07K2317-73
36	417	342	10.39	C07K2317-76
37	412	324	9.84	C07K0016-32
40	407	303	9.20	C07K0016-28
41	405	315	9.57	A61K0047-6855
45	357	259	7.87	C07K0016-2803
48	345	207	6.29	A61K0039-39558

=> D IPC 1-

[View all](#)

TERM #	# OCC	# DOC	% DOC	IPC CPC
1	4528	2305	70.02	A61P0035-00
2	4165	2024	61.48	A61K0047-68
3	3674	1656	50.30	C07K0016-28
4	3515	1839	55.86	A61K0039-395
6	1815	981	29.80	A61K0039-00
7	1503	833	25.30	C07K0016-30
9	1092	692	21.02	C12N0015-13
12	1051	482	14.64	A61K0047-48
14	935	604	18.35	A61K0045-06
15	821	599	18.20	A61P0035-02
16	813	725	22.02	A61K0045-00
17	750	524	15.92	C07K0016-46
19	740	450	13.67	C07K0016-00
21	701	607	18.44	A61P0043-00
25	646	460	13.97	C07K0019-00
26	614	392	11.91	C07K0016-18
27	601	360	10.94	C07K0016-32
28	582	361	10.97	G01N0033-574
29	573	465	14.13	C12N0005-10

Use the CAS STNext Thesauri

- The command to access the various thesauri
- EXPAND term + ALL/IPC or
- EXPAND term +ALL/CPC

```
=> E A61P0035-00+ALL/CPC  
  
E1      0   BT4  A/CPC  
        HUMAN NECESSITIES(2013-01-01)  
E2      0   BT3  A6/CPC  
        HEALTH;AMUSEMENT(2013-01-01)  
E3      0   BT2  A61/CPC  
        MEDICAL OR VETERINARY SCIENCE;HYGIENE(2013-01-01)  
E4      420755 BT1  A61P/CPC  
        SPECIFIC THERAPEUTIC ACTIVITY OF CHEMICAL COMPOUNDS OR  
        MEDICINAL PREPARATIONS(2018-01-01)  
E5      116694 --> A61P0035-00/CPC  
        Antineoplastic agents(2018-01-01)  
E6      18371  NT1  A61P0035-02/CPC  
        specific for leukemia(2018-01-01)  
E7      10378  NT1  A61P0035-04/CPC  
        specific for metastasis(2018-01-01)  
  
***** END *****
```

This was the highest posted class in both the IPCs and CPCs, but may not be the best class code for this concept.

Use the CAS STNext Thesauri

=> E A61K0047-68+ALL/CPC			
E1	0	BT7	A/CPC HUMAN NECESSITIES(2013-01-01)
E2	0	BT6	A6/CPC HEALTH;AMUSEMENT(2013-01-01)
E3	0	BT5	A61/CPC MEDICAL OR VETERINARY SCIENCE;HYGIENE(2013-01-01)
E4	809041	BT4	A61K/CPC PREPARATIONS FOR MEDICAL, DENTAL, OR TOILET PURPOSES reference: devices or methods specially adapted for bringing pharmaceutical products into particular physical or administering forms A61J0003-00;
E5	671	BT3	A61K0047-00/CPC Medicinal preparations characterised by the non-active ingredients used, e.g. carriers or inert additives;Targeting or modifying agents chemically bound to the active ingredient(2017-08-01)
E6	1704	BT2	A61K0047-50/CPC the non-active ingredient being chemically bound to the active ingredient, e.g. polymer-drug conjugates(2017-08-01)
E7	147	BT1	A61K0047-51/CPC the non-active ingredient being a modifying agent(2017-08-01)
E8	928	-->	A61K0047-68/CPC the modifying agent being an antibody, an immunoglobulin or a fragment thereof, e.g. an Fc-fragment(2017-08-01)

E9	300	NT1	A61K0047-6801/CPC CPC-specific-text: Drug-antibody or immunoglobulin conjugates defined by the pharmacologically or therapeutically active agent(2017-08-01)
E10	2520	NT2	A61K0047-6803/CPC CPC-specific-text: Drugs conjugated to an antibody or immunoglobulin, e.g. cisplatin-antibody conjugates(2017-08-01)
E11	55	NT3	A61K0047-6805/CPC CPC-specific-text: the drug being a vinca alkaloid(2017-08-01)
E12	318	NT3	A61K0047-6807/CPC CPC-specific-text: the drug or compound being a sugar, nucleoside, nucleotide, nucleic acid, e.g. RNA antisense(2017-08-01)
E13	438	NT4	A61K0047-6809/CPC CPC-specific-text: Antibiotics, e.g. antitumor antibiotics anthracyclins, adriamycin, doxorubicin or daunomycin(2017-08-01)
E14	891	NT3	A61K0047-6811/CPC CPC-specific-text: the drug being a protein or peptide, e.g. transferrin or bleomycin(2017-08-01)
E15	361	NT4	A61K0047-6813/CPC CPC-specific-text: the drug being a peptidic cytokine, e.g. an interleukin or interferon(2017-08-01)
E16	188	NT4	A61K0047-6815/CPC CPC-specific-text: Enzymes(2017-08-01)

Classification discussion

- These classes look much better, but what is the right level to use?
- Broader/narrower than the term expanded around?
- To include all narrower classes plus the class in question in both the IPC and CPC systems, use
 - S *class*+NT/IPC,CPC
- For this example, the class for E5 is used as the broadest level
 - S A61K0047-50+NT/CPC,IPC

```
=> S A61K0047-50+NT/CPC,IPC
```

```
35398 A61K0047-50+NT/CPC (127 TERMS)
```

```
34382 A61K0047-50+NT/IPC (16 TERMS)
```

```
L3 51957 A61K0047-50+NT/CPC,IPC
```


Classifications

- Now the ADC abbreviation can be used in conjunction with the class set

```
=> S ADC AND L3

      26117 ADC/BI
      1576 ADCS/BI
      26726 ADC/BI
          ((ADC OR ADCS)/BI)
      46712 ADC/BIEX
      2409 ADCS/BIEX
      47281 ADC/BIEX
          ((ADC OR ADCS)/BIEX)
L4      1022 ADC/BI,BIEX AND L3

=> s l4 not l1

L5      97 L4 NOT L1
```

Unique ADC plus class results

L5 ANSWER 1 OF 97 WPINDEX COPYRIGHT 2023 CLARIVATE on STN
IPCI A61K0038-08 [I,A]; A61K0047-65 [I,A]; A61K0047-68 [I,A]; A61P0035-00 [I,A]; C07K0007-06 [I,A]
NOV . . . reacting pre-synthesized linkers with small molecule cytotoxins to generate linker-toxin payloads, coupling the linker-toxin payload with antibody to obtain final ADC product.

L5 ANSWER 2 OF 97 WPINDEX COPYRIGHT 2023 CLARIVATE on STN
IPCI A61K0031-4745 [I,A]; A61K0031-4745 [I,A]; A61K0039-395 [I,A]; A61K0047-65 [I,A]; A61K0047-68 [I,A]; A61K0047-68 [I,A]; A61P0035-00 [I,A]; A61P0035-00 [I,A]; A61P0035-02 [I,A]; C07K0001-107 [I,A]; C07K0016-46 [I,A]; C12N0015-13 [I,A]
ACTN . . . cell BxPC3. Results showed that in the BxPC3 single tumor model expressed in EGFR, 35 days after the first administration, ADC-6 drug had a significant tumor growth inhibitory effect on the human epidermal cancer cell BxPC3 BALB/c-Nude mouse subcutaneous xenograft tumor. . .

L5 ANSWER 7 OF 97 WPINDEX COPYRIGHT 2023 CLARIVATE on STN
IPCI A61K0031-357 [I,A]; A61K0031-4745 [I,A]; A61K0031-537 [I,A]; A61K0038-07 [I,A]; A61K0045-00 [I,A]; A61K0047-55 [I,A]; A61K0047-55 [I,A]; A61K0047-64 [I,A]; A61K0047-65 [I,A]; A61K0047-68 [I,A]; A61K0047-68 [I,A]; A61P0035-00 [I,A]; A61P0035-00 [I,A]; A61P0035-02 [I,A]

Member(0001)

TIEN PREPARATION METHOD FOR DUAL-DRUG-LINKER OF ADC AND USE THEREOF

Include Japanese classification codes

=> S A61K0047-50/FCL

L6 1872 A61K0047-50/FCL

=> S L6 NOT L3

L7 841 L6 NOT L3

=> D TI PI 1-2

L7 ANSWER 1 OF 841 WPINDEX COPYRIGHT 2023 CLARIVATE on STN

TI New tamoxifen compound used in pharmaceutical composition for treating cancer e.g. stomach cancer, colon cancer, liver cancer, pancreatic cancer, lung cancer and breast cancer, multiple myeloma, amyloidosis and autoimmune disease

PI **JP 2016222659** A 20161228 (201708) * JA 82[9]

JP 6792256 B2 20201125 (2020096) JA 82

L7 ANSWER 2 OF 841 WPINDEX COPYRIGHT 2023 CLARIVATE on STN

TI New fusion protein comprises fusion of bone morphogenetic protein 7 (BMP7) variant and human serum albumin, useful for preparing therapeutic agent for renal disease

PI **JP 2016204292** A 20161208 (201701) * JA 12[9]

Agenda

- Value-add searching
 - Understand basic indexing principles
 - Understand indexing tools available to indexer

Derwent World Patents Index

- The Derwent World Patents Index database (DWPI) is the main/ultimate Derwent database
- Patent family database
- Value add abstracting, manual codes, Derwent classes, polymer indexing
- DCR and DWPIIM indexing with roles
- Fragmentation
- Citation info

Derwent companion databases

- Derwent Chemistry Resource (DCR)
 - Specific compounds +
 - 1999, plus some earlier
- Derwent Markush Resource (DWPIM)
 - Markush structures
 - Specific compounds
 - ~ 1978, plus some earlier

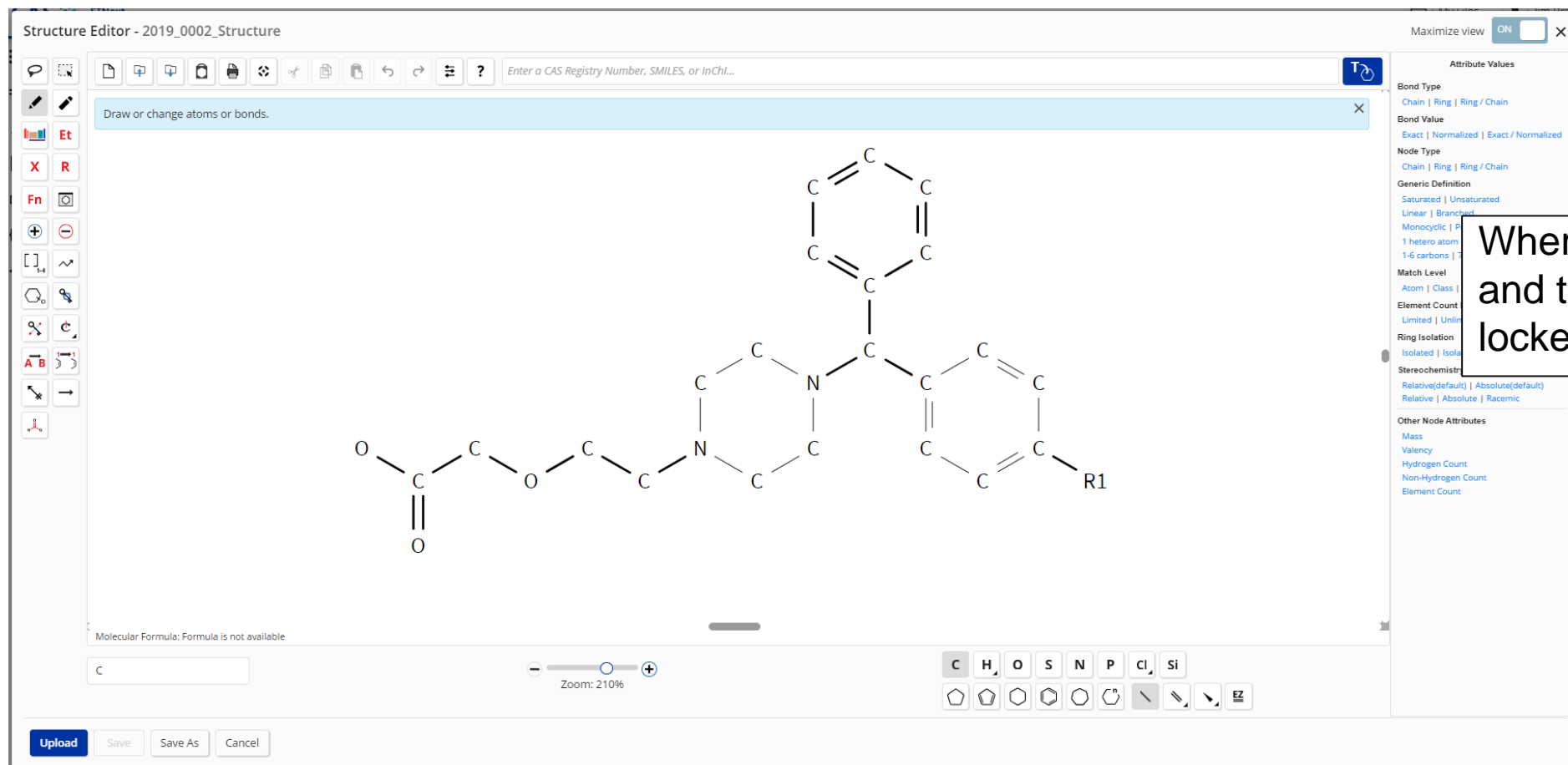
Derwent indexing of small molecules

- Fragmentation – 1963 -
- Derwent Markush Resource (DWPMIM) – ~1978 -
- Derwent Chemistry Resource (DCR) – ~1999 –
 - Some earlier indexing – DCNs and DRNs
- To cover all time periods, fragmentation must be included!

Derwent small molecule search

- Draw structure
- Structure search in DCR
- Structure search in DWPIIM
- S DCR set and DWPIIM set in DWPI to get patent family records
- Run frag search, find unique records

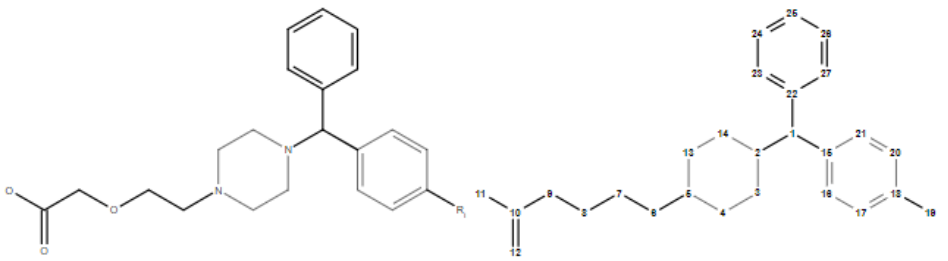
Derwent small molecule search (cont.)



Where R1 = Cl, OH, Me
and the Benzene ring is
locked.

Derwent small molecule search (cont.)

File DCR
=>
Uploading structure file: 2019_0002_Structure



R-Group Definitions
R1: Cl,OH,Me

Node Attributes
Ring Nodes : 2 3 4 5 13 14 15 16 17 18 20 21 22 23 24 25 26 27
Isolated Ring Nodes : 22 23 24 25 26 27
Chain Nodes : 1 6 7 8 9 10 11 12

Bond Attributes
Ring Bonds : 2-3 2-14 3-4 4-5 13-5 14-13 15-16 15-21 16-17 17-18 20-18 21-20 22-23 22-27 23-24 24-25 26-25 27-26
Chain Bonds : 1-2 1-15 1-22 5-6 6-7 7-8 8-9 9-10 10-11 10-12 18-19
Normalized Bonds : 15-16 15-21 16-17 17-18 20-18 21-20 22-23 22-27 23-24 24-25 26-25 27-26
Exact/Normalized Bonds : 1-2 1-15 1-22 2-3 2-14 3-4 4-5 5-6 6-7 7-8 8-9 9-10 10-11 10-12 13-5 14-13 18-19

Markush Attributes
Match Level (ATOM) : 2 3 4 5 13 14 15 16 17 18 20 21 22 23 24 25 26 27
Match Level (CLASS) : 1 6 7 8 9 10 11 12
Element Count Level (LIMITED) : 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 20 21 22 23 24 25 26 27

L1 STRUCTURE UPLOADED

=> FILE DCR

=> S L1 SSS FULL

FULL SEARCH INITIATED 11:14:46

FULL SCREEN SEARCH COMPLETED - 5204481 TO ITERATE

100.0% PROCESSED 5204481 ITERATIONS

87 ANSWERS

SEARCH TIME: 00.00.02

L2 87 SEA SSS FUL L1

=> FILE DWPIM

=> S L1 SSS FULL

100.0% PROCESSED 2638066 ITERATIONS (1 INCOMPLETE)

59 ANSWERS

SEARCH TIME: 00.03.03

L3 59 SEA SSS FUL L1

Derwent small molecule search (cont.)

=> FILE WPIX; S L2; S L3

L4 1313 L2

L5 38 L3

=> S L4 OR L5

L6 1325 L4 OR L5

=> ANA PY.B L6 1-

ANALYZE IS APPROXIMATELY 64% COMPLETE

L7 ANALYZE L6 1- PY.B : 36 TERMS

=> D 1-36

L7 ANALYZE L6 1- PY.B : 36 TERMS

TERM #	# OCC	# DOC	% DOC	PY.B
--------	-------	-------	-------	------

1	99	99	7.47	2009
2	84	84	6.34	2022
3	83	83	6.26	2010
4	72	72	5.43	2007
5	65	65	4.91	2008
6	65	65	4.91	2021
7	60	60	4.53	2014
8	59	59	4.45	2019
9	58	58	4.38	2011
10	55	55	4.15	2013
11	55	55	4.15	2020
12	54	54	4.08	2012
13	52	52	3.92	2005
14	52	52	3.92	2017
15	52	52	3.92	2018
16	44	44	3.32	2003
17	43	43	3.25	2006

18	43	43	3.25	2016
19	42	42	3.17	2023
20	41	41	3.09	2015
21	38	38	2.87	2004
22	25	25	1.89	2002
23	17	17	1.28	2001
24	16	16	1.21	1999
25	10	10	0.75	2000
26	9	9	0.68	1997
27	8	8	0.60	1998
28	6	6	0.45	1995
29	4	4	0.30	1994
30	3	3	0.23	1988
31	3	3	0.23	1990
32	2	2	0.15	1992
33	2	2	0.15	1993
34	2	2	0.15	1996
35	1	1	0.08	1989
36	1	1	0.08	1991

***** END OF L7 *****

Derwent small molecule search (cont.)

← Return to Session

Structures (152) Sort: Date Modified: Newest ▾

□ [Download] [Trash] [Folder] Move to Folder Search Files by Name [Search] [Import Sequence] [Import Structure]

□ [Structure Icon] 2019_0002_Structure ✎
14 Nov 2023 11:14 AM

[Edit]

← Upload ...

- Move
- Delete
- Download
- Generate FragCode Script**

Generate Derwent frag
code script here.

Derwent small molecule search (cont.)

Edit Script

2019_0002_Structure_FragCode_01

Check your script with the validate button.

Save As [Share] [Download] [Undo] [Redo] Validate

```
1 =>s (H602)/M0,M2,M3,M4 \>_line1
2 =>s _line1(P)(M322)/M0,M2,M3,M4 \>_line2
3 =>s _line2(P)(H641)/M0,M2,M3,M4 \>_line3
4 =>s (_line1(P)M900/M0) OR (_line1(P)M901/M2,M3,M4) OR (_line2(P)M902/M2,M3,M4) OR
_line3 \>_line4
5 =>s _line4(NOTP)(H4 OR H7)/M2,M3,M4 \>_line5
6 =>s (H401(P)H441)/M0,M2,M3,M4 \>_line6
7 =>s _line6(P)(M322)/M0,M2,M3,M4 \>_line7
8 =>s (_line6(P)M900/M0) OR (_line6(P)M901/M2,M3,M4) OR (_line7(P)M902/M2,M3,M4) OR
_line7 \>_line8
9 =>s _line8(NOTP)(H6 OR H7)/M2,M3,M4 \>_line9
10 =>s (H721(P)M341)/M0,M2,M3,M4 \>_line10
11 =>s _line10(P)(M323)/M0,M2,M3,M4 \>_line11
12 =>s (_line10(P)M900/M0) OR (_line10(P)M901/M2,M3,M4) OR (_line11(P)M902/M2,M3,M4)
OR _line11 \>_line12
13 =>s _line12(NOTP)(H4 OR H6)/M2,M3,M4 \>_line13
14 =>s _line5 OR _line9 OR _line13 \>_line14
```

Validation

Success

No errors detected

Run Save Cancel

Run the search as is, unedited, which means this is NOT a true substructure search.

Derwent small molecule search (cont.)

```
=> s (H602)/M0,M2,M3,M4
```

```
16791 (H602)/M0
```

```
335425 (H602)/M2
```

```
167468 (H602)/M3
```

```
48104 (H602)/M4
```

```
L8 533746 (H602)/M0,M2,M3,M4
```

```
=> s L8(P)(M322)/M0,M2,M3,M4
```

```
32433 (M322)/M0
```

```
418011 (M322)/M2
```

```
337872 (M322)/M3
```

```
59774 (M322)/M4
```

```
L9 218285 L8(P)(M322)/M0,M2,M3,M4
```

```
=> s L9(P)(H641)/M0,M2,M3,M4
```

```
473934 K0/M3
```

```
65708 K0/M4
```

```
L28 1408 L27(NOTP)(H3 OR H9 OR J2 OR J3 OR J4 OR J5 OR J9 OR K0)/M2,M3,M4
```

Derwent small molecule search (cont.)

=> S L28 NOT L6

L29 348 L28 NOT L6

=> ANA PY.B L29 1-348

L30 ANALYZE L29 1-348 PY.B : 47 TERMS

=> D 1-47

[View all](#)

TERM #	# OCC	# DOC	% DOC	PY.B
1	19	19	5.46	1989
2	18	18	5.17	1988
3	18	18	5.17	2005
4	18	18	5.17	2007
5	14	14	4.02	2004
6	13	13	3.74	2002
7	13	13	3.74	2006
8	12	12	3.45	2008
9	12	12	3.45	2009
10	11	11	3.16	1991
11	11	11	3.16	2010
12	9	9	2.59	1992
13	9	9	2.59	2003
14	8	8	2.30	1976
15	7	7	2.01	2011
16	6	6	1.72	1990
17	6	6	1.72	1999
18	5	5	1.44	1975
19	5	5	1.44	1983
20	5	5	1.44	1985
21	5	5	1.44	1986
22	5	5	1.44	1993
23	5	5	1.44	1995

24	5	5	1.44	2001
25	5	5	1.44	2013
26	5	5	1.44	2018
27	5	5	1.44	2021
28	4	4	1.15	1974
29	4	4	1.15	1979
30	4	4	1.15	1981
31	4	4	1.15	1987
32	3	3	0.86	1977
33	3	3	0.86	1994
34	3	3	0.86	1996
35	3	3	0.86	1997
36	3	3	0.86	2012
37	3	3	0.86	2022
38	2	2	0.57	1982
39	2	2	0.57	2014
40	2	2	0.57	2019
41	2	2	0.57	2020
42	1	1	0.29	1965
43	1	1	0.29	1969
44	1	1	0.29	1973
45	1	1	0.29	1984
46	1	1	0.29	2000
47	1	1	0.29	2023

***** END OF L30 *****

Polymers indexed in the Derwent databases

- Derwent has two polymer indexing systems in DWPI
 - Derwent enhanced polymer indexing
 - Derwent Plasdoc indexing (retires system but required for backfile searching)
- However, some polymer indexing can be found in DCR and DWPIM
- The polymer itself, or monomers, or modifiers

Polymers indexed in the CAS databases

- CAS has one polymer indexing system
 - Monomer-based indexing
 - Structural repeating units (SRUs)
- However, some polymer indexing can be found in MARPAT
 - The polymer itself, or monomers, or modifiers

Sample polymeric structure

Structure Editor - polysiloxane backbone

Maximize view X

Enter a CAS Registry Number, SMILES, or InChI...

Draw or change atoms or bonds.

Si — O — Si — O — Si — O — Si — O — Si — O

Molecular Formula: H₁₂O₅Si₅ (232.52)

Zoom: 200%

Upload Save Save As Cancel

Attribute Values

- Bond Type: Chain | Ring | Ring / Chain
- Bond Value: Exact | Normalized | Exact / Normalized
- Node Type: Chain | Ring | Ring / Chain
- Generic Definition: Saturated | Unsaturated; Linear | Branched; Monocyclic | Polycyclic; 1 hetero atom | 2+ hetero atoms; 1-6 carbons | 7+ carbons
- Match Level: Atom | Class | Any
- Element Count Level: Limited | Unlimited
- Ring Isolation: Isolated | Isolated / Embedded
- Stereochemistry: Relative(default) | Absolute(default); Relative | Absolute | Racemic
- Other Node Attributes: Mass; Valency; Hydrogen Count; Non-Hydrogen Count; Element Count

Substructure Searching polymeric structure in CAS REGISTRY

```
=> s l1 sss full
```

```
FULL SEARCH INITIATED 11:38:10
```

```
FULL SCREEN SEARCH COMPLETED - 48285 TO ITERATE
```

```
100.0% PROCESSED 48285 ITERATIONS
```

```
3343 ANSWERS
```

```
SEARCH TIME: 00.00.01
```

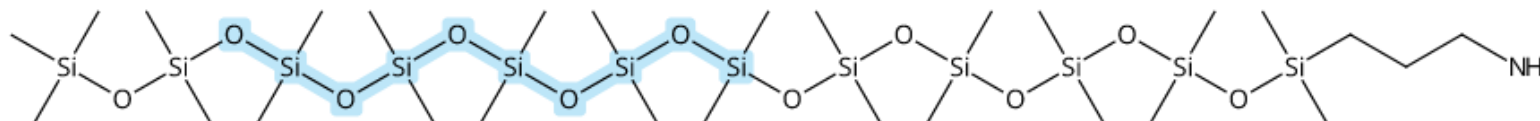
```
L4 3343 SEA SSS FUL L1
```

Substructure Searching polymeric structure in CAS REGISTRY

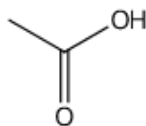
=> D STR

L4 ANSWER 1 OF 3343 REGISTRY COPYRIGHT 2023 ACS on STN

CM 1



CM 2



Substructure Searching polymeric structure in DWPIIM

L1 STRUCTURE UPLOADED

=> **s l1 sss full**

FULL SEARCH INITIATED 11:30:59

FULL SCREEN SEARCH COMPLETED - 2634767 TO ITERATE

100.0% PROCESSED 2634767 ITERATIONS

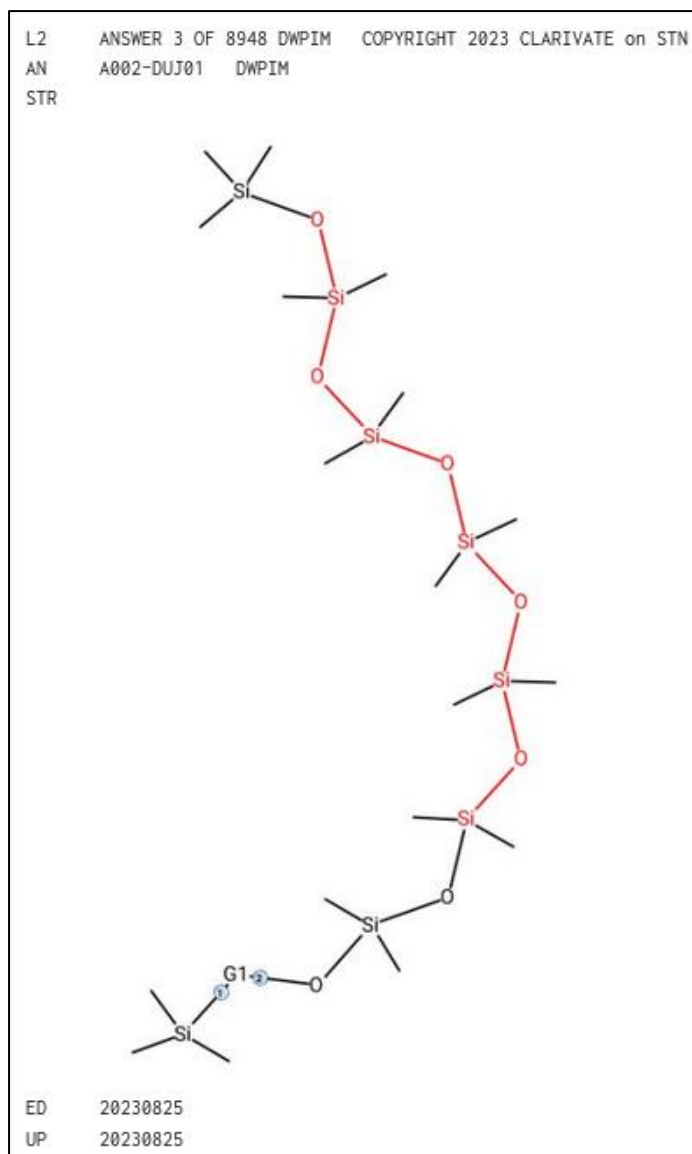
8948 ANSWERS

SEARCH TIME: 00.00.02

L2 8948 SEA SSS FUL L1

=> **d 3**

Substructure Searching polymeric structure in DWPIIM



Substructure Searching polymeric structure in DCR

```
=> s l1 sss full
```

```
FULL SEARCH INITIATED 11:34:16
```

```
FULL SCREEN SEARCH COMPLETED - 5187375 TO ITERATE
```

```
100.0% PROCESSED    5187375 ITERATIONS
```

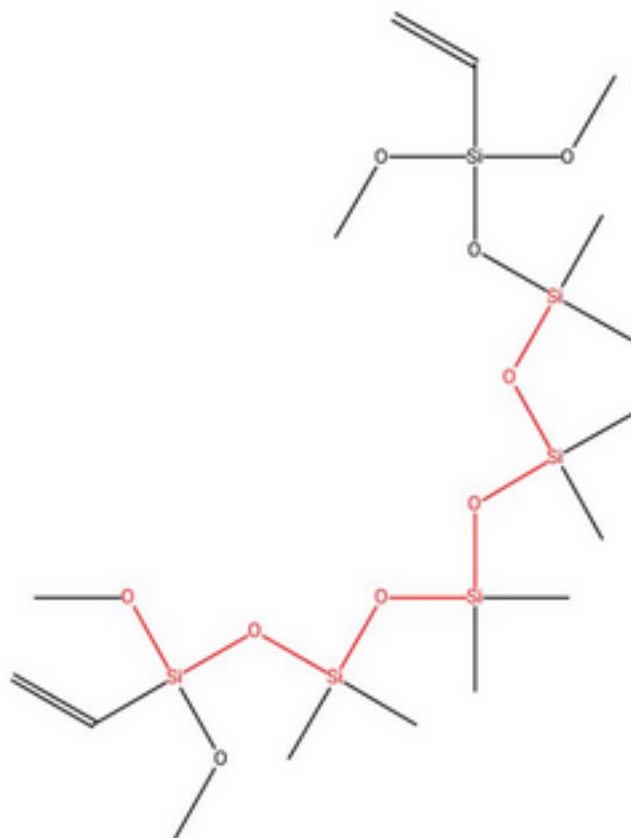
```
282 ANSWERS
```

```
SEARCH TIME: 00.00.03
```

```
L3          282 SEA SSS FUL L1
```

Substructure Searching polymeric structure in DCR (cont.)

L3 ANSWER 1 OF 282 DCR COPYRIGHT 2023 CLARIVATE on STN.
AN DCR-5993231 DCR
DCSE 5993231-0-0-0
CN.P 3,13-dietheryl-3,13-dimethoxy-5,5,7,7,9,9,11,11-octamethyl-
2,4,6,8,10,12,14-heptaoxa-3,5,7,9,11,13-hexasilapentadecane
STR



— t

PET polymer backbone

Structure Editor - PET backbone

Maximize view X

Enter a CAS Registry Number, SMILES, or InChI...

Draw or change atoms or bonds.

Molecular Formula: C₄₀H₃₄O₁₆ (770.70)

c

Zoom: 100%

C H O S N P Cl Si

Upload Save Save As Cancel

Attribute Values

- Bond Type**
Chain | Ring | Ring / Chain
- Bond Value**
Exact | Normalized | Exact / Normalized
- Node Type**
Chain | Ring | Ring / Chain
- Generic Definition**
Saturated | Unsaturated
Linear | Branched
Monocyclic | Polycyclic
1 hetero atom | 2+ hetero atoms
1-6 carbons | 7+ carbons
- Match Level**
Atom | Class | Any
- Element Count Level**
Limited | Unlimited
- Ring Isolation**
Isolated | Isolated / Embedded
- Stereochemistry**
Relative(default) | Absolute(default)
Relative | Absolute | Racemic
- Other Node Attributes**
Mass
Valency
Hydrogen Count
Non-Hydrogen Count
Element Count

Substructure searching PET polymeric structure in CAS REGISTRY

```
=> S L6 SSS FULL
```

```
FULL SEARCH INITIATED 12:24:52
```

```
FULL SCREEN SEARCH COMPLETED - 3112 TO ITERATE
```

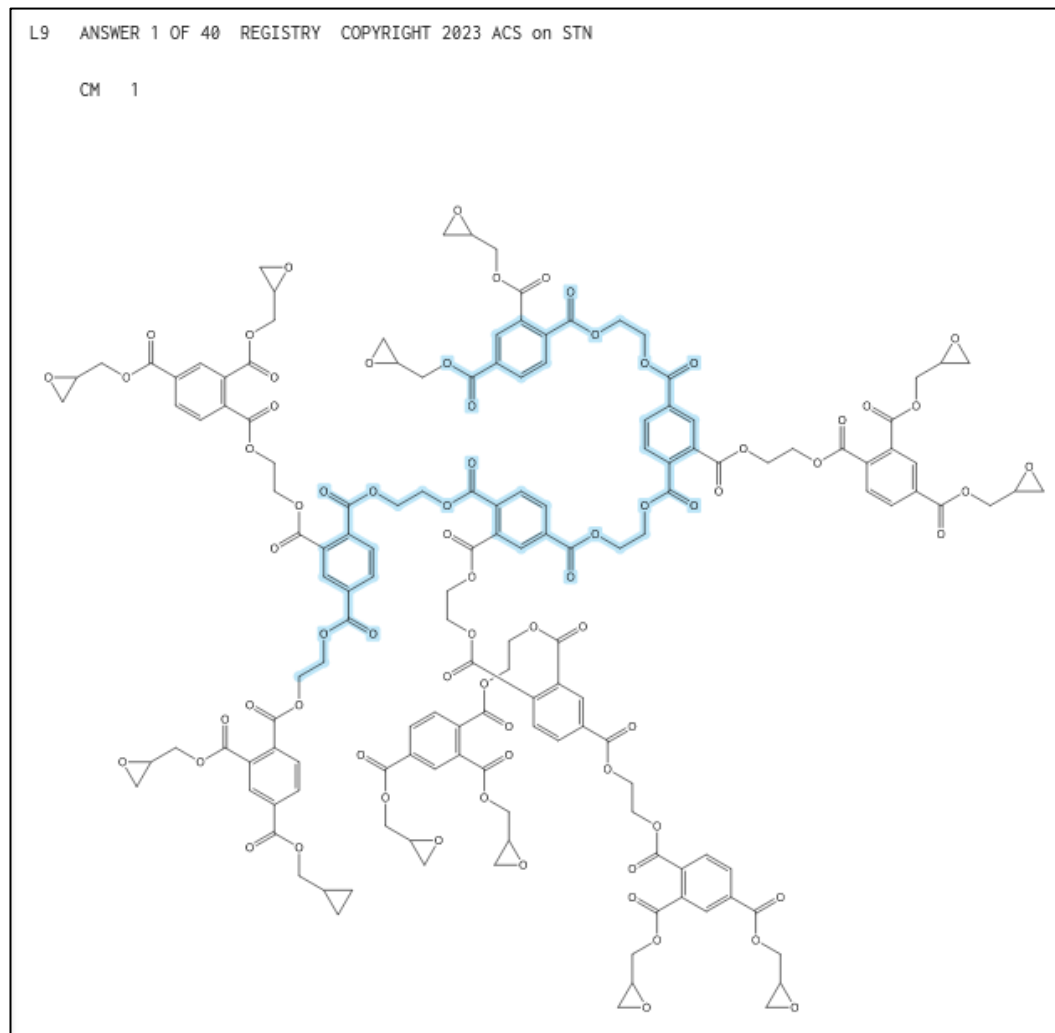
```
100.0% PROCESSED 3112 ITERATIONS
```

```
40 ANSWERS
```

```
SEARCH TIME: 00.00.01
```

```
L9 40 SEA SSS FUL L6
```

Substructure searching PET polymeric structure in CAS REGISTRY (cont.)



Substructure searching PET polymeric structure in DCR

```
L6      STRUCTURE UPLOADED

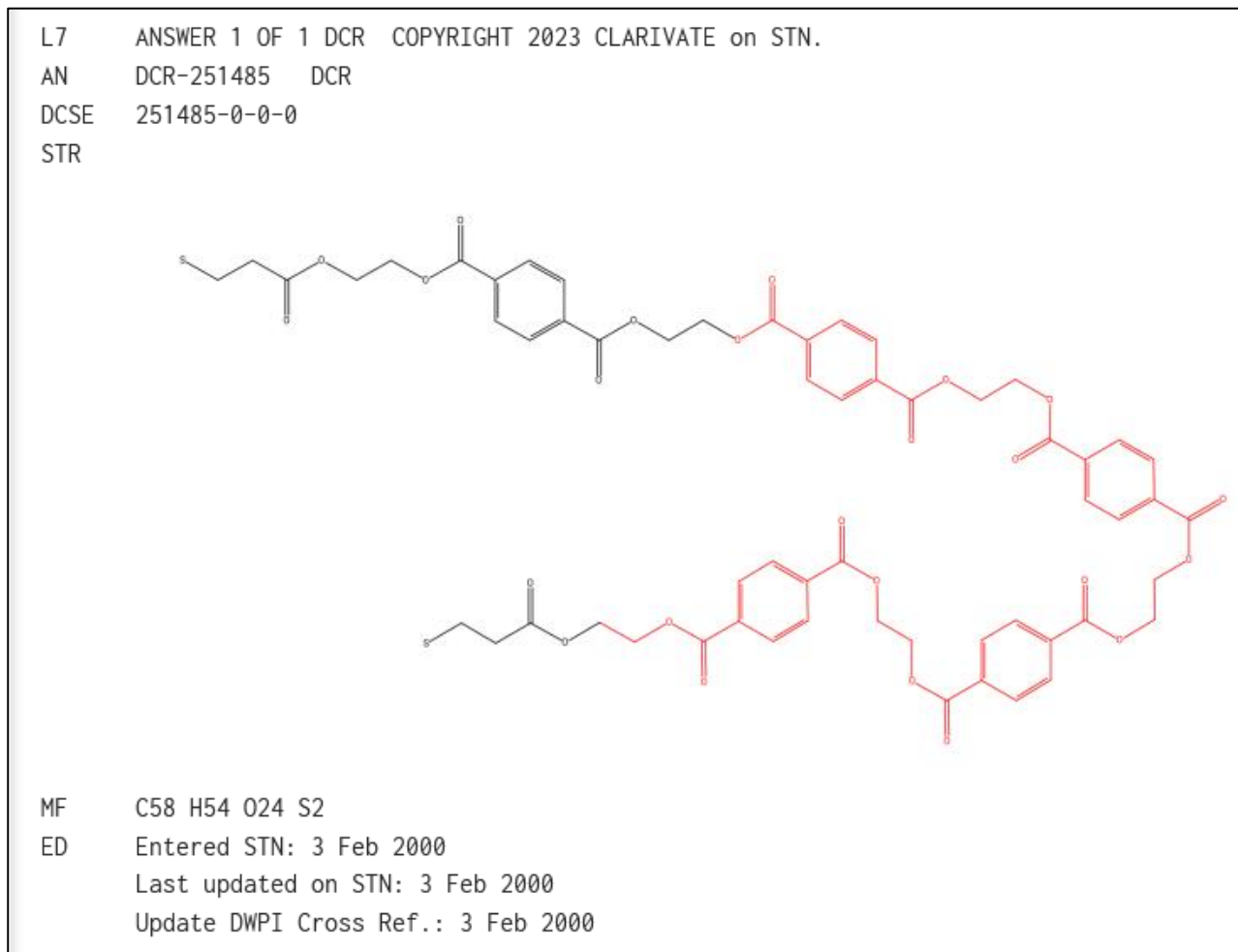
=> S L6 SSS FULL

FULL SEARCH INITIATED 12:04:45
FULL SCREEN SEARCH COMPLETED - 5187375 TO ITERATE

100.0% PROCESSED    5187375 ITERATIONS                1 ANSWERS
SEARCH TIME: 00.00.02

L7      1 SEA SSS FUL L6
```

Substructure searching PET polymeric structure in DCR (cont.)



Substructure searching PET polymeric structure in DWPIIM

=> S L6 SSS FULL

FULL SEARCH INITIATED 12:19:10

FULL SCREEN SEARCH COMPLETED - 2634767 TO ITERATE

100.0% PROCESSED 2634767 ITERATIONS

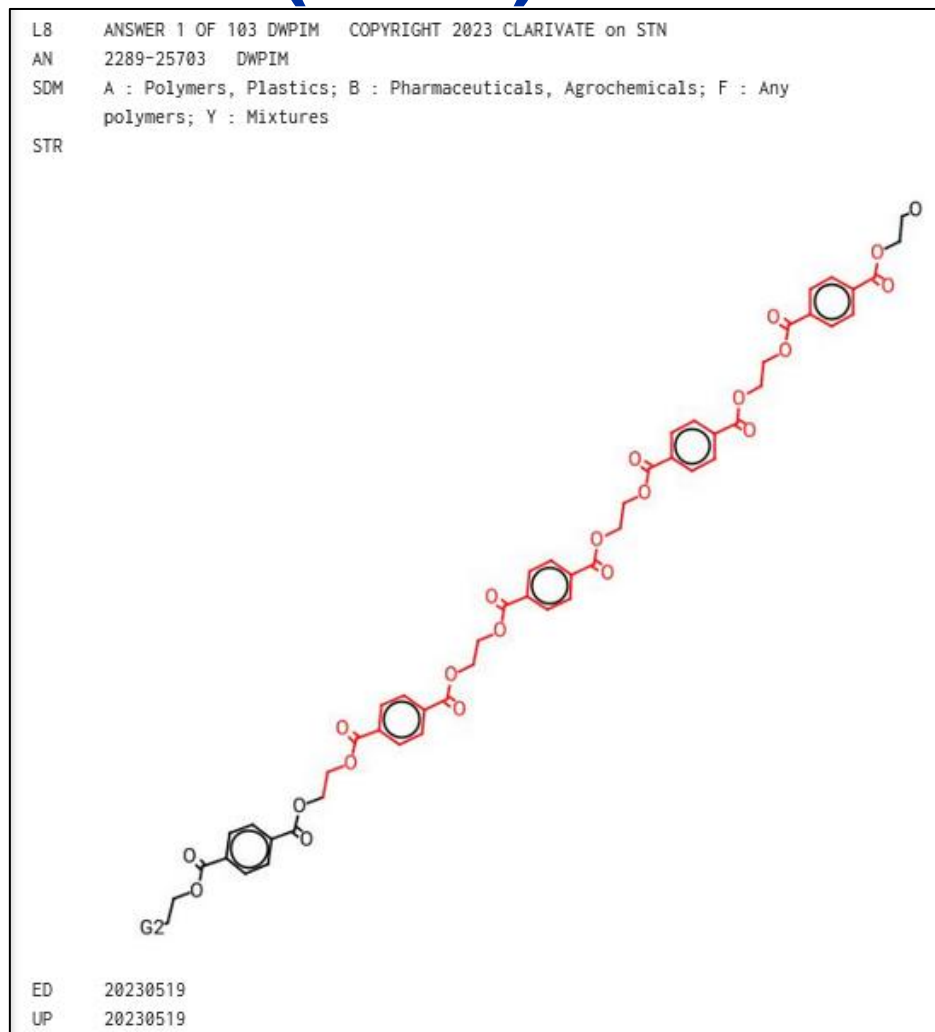
103 ANSWERS

SEARCH TIME: 00.00.34

L8

103 SEA SSS FUL L6

Substructure searching PET polymeric structure in DWPIIM (cont.)



Sequence searching

- CAS REGISTRY
 - BLAST, Motif and Structure searching
- GENESEQ, USGENE, PATGENE
 - BLAST, and Motif searching
- DCR
 - Structure searching
 - Chemical Name/Chemical Name Segment
- DWPIM/MARPAT

CAS REGISTRY record for Humira

L2 ANSWER 1 OF 1 REGISTRY COPYRIGHT 2023 ACS on STN
RN **331731-18-1** REGISTRY
ED Entered STN: 18 Apr 2001
CN Immunoglobulin G1, anti-(human tumor necrosis factor) (human monoclonal D2E7 heavy chain), disulfide with human monoclonal D2E7 light chain, dimer (CA INDEX NAME)
OTHER NAMES:
CN Adalimumab
CN Adalimumab-aqvh
CN Adalimumab-CinnaGen
CN CHS-1420
CN CinnoRA
CN D 2E7
CN D2E7
CN **Humira**
CN LU 200134
CN Yusimry
FS PROTEIN SEQUENCE
SQL 1330,451,451,214,214
NTE multichain

CAS REGISTRY record for Humira (cont.)

```
SEQ      1 EVQLVESGGG LVQPGRSLRL SCAASGFTFD DYAMHWVRQA PGKGLEWVSA
      51 ITWNSGHIDY ADSVEGRFTI SRDNAKNSLY LQMNSLRAED TAVYYCAKVS
     101 YLSTASSLDY WGQGTSLVTVS SASTKGPSVF PLAPSSKSTS GGTAALGCLV
     151 KDYFPEPVTV SWNSGALTSG VHTFPAVLQS SGLYSLSSVV TVPSSSLGTQ
     201 TYICNVNHKP SNTKVDKKVE PKSCDKTHC PPCAPELLG GPSVFLFPPK
     251 PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY
     301 NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP
     351 QVYTLPPSRD ELTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTTP
     401 VLDSGDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG
     451 K
```

```
SEQ      1 EVQLVESGGG LVQPGRSLRL SCAASGFTFD DYAMHWVRQA PGKGLEWVSA
      51 ITWNSGHIDY ADSVEGRFTI SRDNAKNSLY LQMNSLRAED TAVYYCAKVS
     101 YLSTASSLDY WGQGTSLVTVS SASTKGPSVF PLAPSSKSTS GGTAALGCLV
     151 KDYFPEPVTV SWNSGALTSG VHTFPAVLQS SGLYSLSSVV TVPSSSLGTQ
     201 TYICNVNHKP SNTKVDKKVE PKSCDKTHC PPCAPELLG GPSVFLFPPK
     251 PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY
     301 NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP
     351 QVYTLPPSRD ELTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTTTP
     401 VLDSGDGSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNHY TQKSLSLSPG
     451 K
```

```
SEQ      1 DIQMTQSPSS LSASVGDRVT ITCRASQGIR NYLAWYQQKP GKAPKLLIYA
      51 ASTLQSGVPS RFSGSGSGTD FTLTISSLQP EDVATYYCQR YNRAPYTFGQ
     101 GTKVEIKRTV AAPSVFIFPP SDEQLKSGTA SVVCLLNNFY PREAKVQWKV
     151 DNALQSGNSQ ESVTEQDSKD STYLSSTLT LSKADYEKHK VYACEVTHQG
     201 LSSPVTKSFN RGEN
```


```
SEQ      1 DIQMTQSPSS LSASVGDRVT ITCRASQGIR NYLAWYQQKP GKAPKLLIYA
      51 ASTLQSGVPS RFSGSGSGTD FTLTISSLQP EDVATYYCQR YNRAPYTFGQ
     101 GTKVEIKRTV AAPSVFIFPP SDEQLKSGTA SVVCLLNNFY PREAKVQWKV
     151 DNALQSGNSQ ESVTEQDSKD STYLSSTLT LSKADYEKHK VYACEVTHQG
     201 LSSPVTKSFN RGEN
```

Derwent indexing for Humira





- Sequences in individual GENESEQ records
- TRANSFER PNs from each GENESEQ search into DWPI, AND the two sets
- Check DCR
 - Structure?
 - Name/name segment?
 - Subject descriptor/Comment field?

Humira heavy chain

Edit Biosequence

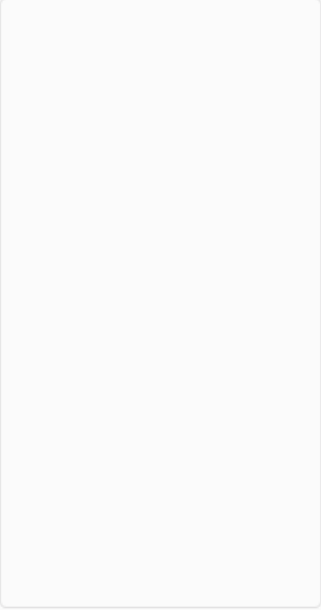
Humira heavy chain 

Check your sequence with the Validate button.

Save As   |   Validate

```
1 1 EVQLVESGGG LVQPGRSLRL SCAASGFTFD DYAMHWVRQA PGKGLEWVSA
2      51 ITWNSGHIDY ADSVEGRFTI SRDANKNSLY LQMNSLRAED TAVYYCAKVS
3      101 YLSTASSLDY WQQGLVTVS SASTKGPSVF PLAPSSKSTS GGTAALGCLV
4      151 KDYFPEPVTY SWNSGALTSG VHTFPAVLQS SGLYSLSSVY TVPSSSLGTQ
5      201 TYICNVNHPK SNTKVDKKVE PKSCDKHTC PPCAPELLG GPSVFLFPPK
6      251 PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY
7      301 NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP
8      351 QVYTLPPSRD ELTKNQVSLT CLVKGFYPSD IAVEWESNGQ PENNYKTPP
9      401 VLDSGDSFFL YSKLTVDKSR WQQGNVFSCS VMHEALHNYH TQKSLSLSPG
10     451 K
```

Validation



Upload Save Cancel

GENESEQ BLAST search for Humira heavy chain

File GENESEQ

=>

Uploading sequence file: Humira heavy chain

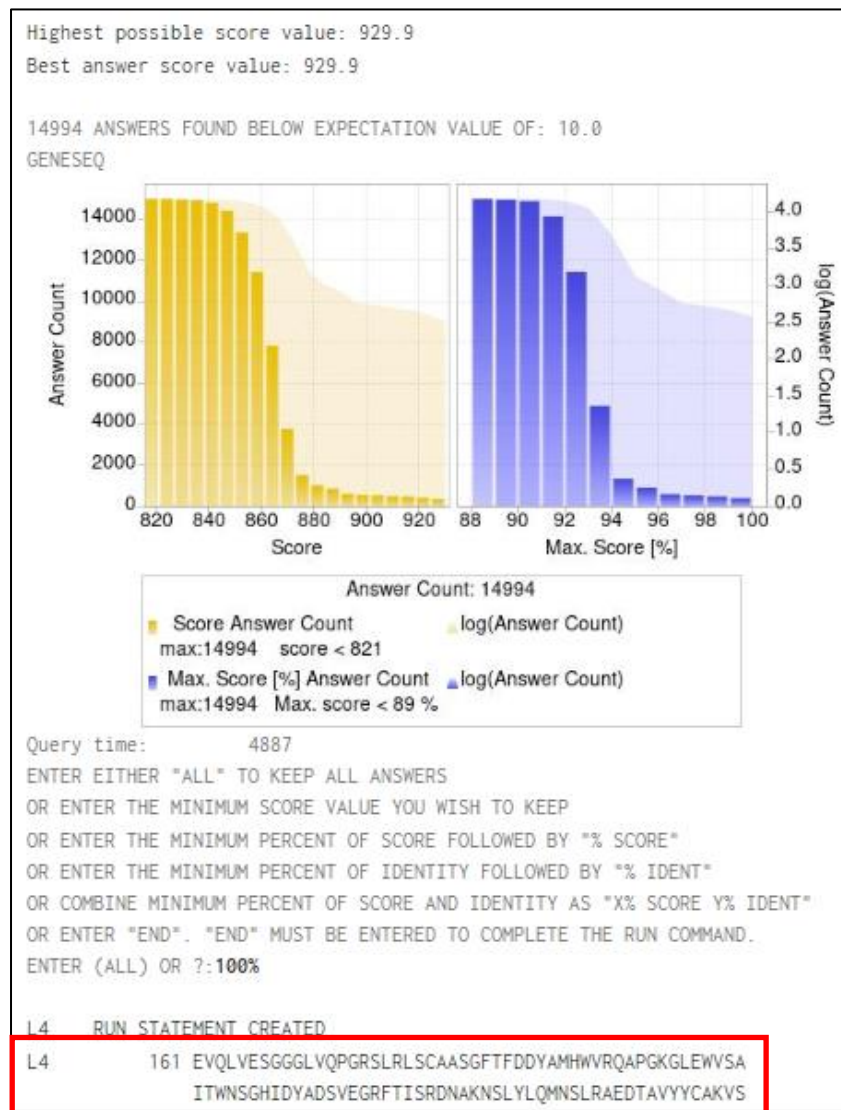
UPLOAD SUCCESSFULLY COMPLETED

L3 GENERATED

=> RUN BLAST L3/ SQP -F F

Algorithm: BLAST - BLASTP. Version: 2.12.0+

GENESEQ BLAST search for Humira heavy chain (cont.)







Humira light chain

Edit Biosequence

Humira light chain

Check your sequence with the Validate button.

Save As   |   Validate

```
1 1 DIQMTQSPSS LSASVGRVT ITCRASQGIT NYLAWYQQKPK GKAPKLLIYA
2     51 ASTLQSGVPS RFSGSGSGTD FTLTISSLQP EDVATYYCQR YNRAPYTFGQ
3     101 GTKVEIKRTV AAPSVFIFPP SDEQLKSGTA SVVCLLNIFY PREAKVQMKV
4     151 DNALQSGNSQ ESVTEQDSKD STYLSSTLT LSKADYEKHK VYACEVTHQG
5     201 LSSPVTKSFN RGECL
```

Validation

Upload Save Cancel

GENESEQ BLAST search for Humira light chain

```
File GENESEQ

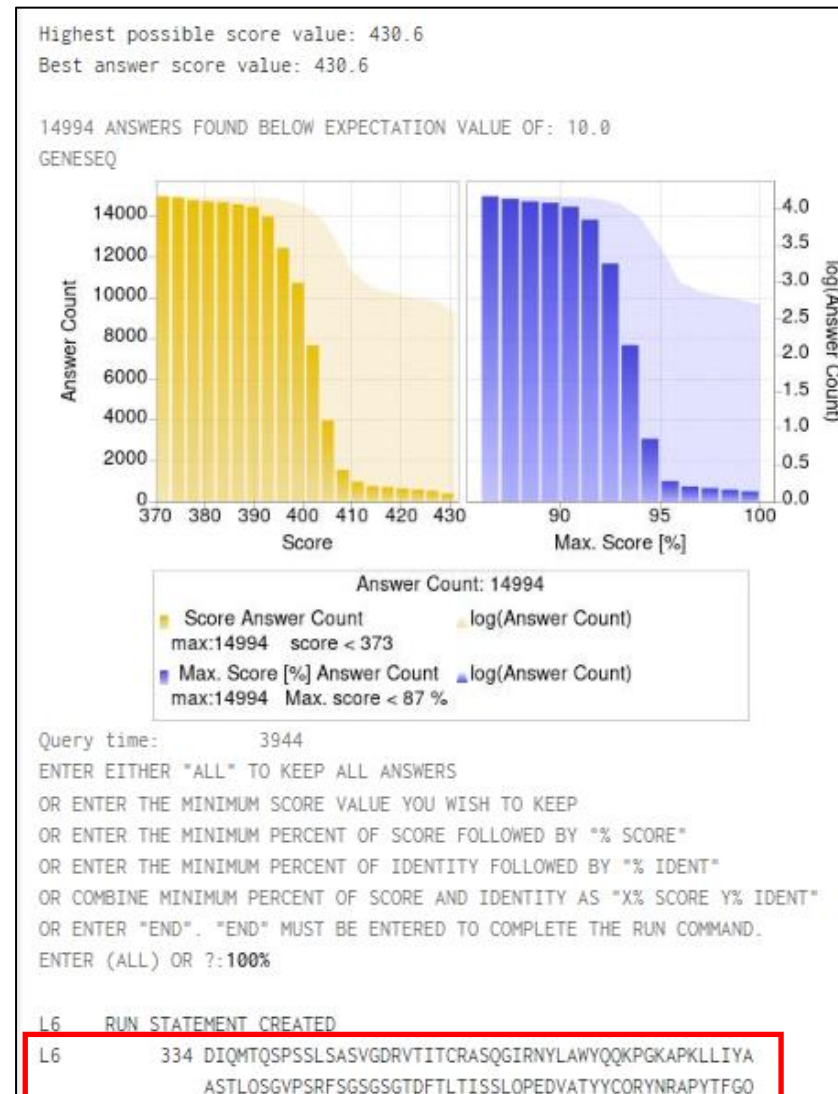
=>
Uploading sequence file: Humira light chain

UPLOAD SUCCESSFULLY COMPLETED
L5 GENERATED

=> RUN BLAST L5/SQP -F F

Algorithm: BLAST - BLASTP. Version: 2.12.0+
```


GENESEQ BLAST search for Humira light chain (cont.)



TRANSFER PNs into DWPI, AND sets together

```
=> FILE WPINDEX; TRA L4 PN 1-; TRA L6 PN 1-
```

```
L7          TRANSFER L4 1- PN :    110 TERMS  
L8          110 L7  
ALL TERMS IN L7 RETRIEVED.
```

```
L9          TRANSFER L6 1- PN :    233 TERMS  
L10         232 L9  
ALL TERMS IN L9 RETRIEVED.
```

```
=> S L8 AND L10
```

```
L11         108 L8 AND L10
```

Check DCR

=> FILE DCR; S HUMIRA/CNS

L12 1 HUMIRA/CNS

=> D

L12 ANSWER 1 OF 1 DCR COPYRIGHT 2023 CLARIVATE on STN.

AN DCR-463982 DCR

DCSE 463982-0-0-0

CN.P ADALIMUMAB

SY ADALIMUMAB; D-2-E-7; HUMIRA; TRUDEXA

STR

Substance image not available

CMT An antibody of unknown structure. Immunoglobulin G1, anti-(human tumor necrosis factor)(human monoclonal D2E7 heavy chain), disulfide with human monoclonal D2E7 light chain, dimer. Molecular weight is approximately 148,000 daltons

MF Unknown

ED Entered STN: 8 Aug 2003

Last updated on STN: 1 Aug 2014

Update DWPI Cross Ref.: 9 Nov 2023



Add that to DWPI search






=> FILE WPINDEX; S L12

L13 1605 L12

=> S L13 OR L11

L14 1668 L13 OR L11

Entered REGISTRY	14:00:41 ON 13 NOV 2023
L2 1 S HUMIRA/CN	 ...
Entered GENESEQ	14:17:01 ON 13 NOV 2023
L3 UPLOAD R BLAST	<input type="button" value="Edit"/>
L4 RUN STATEMENT CREATED	
L5 UPLOAD R BLAST	<input type="button" value="Edit"/>
L6 RUN STATEMENT CREATED	
Entered WPINDEX	14:34:39 ON 13 NOV 2023
Entered GENESEQ	14:34:39 ON 13 NOV 2023
L7 TRA L4 1- PN : 110 TERMS	
Entered WPINDEX	14:34:41 ON 13 NOV 2023
L8 110 SEA L7	 ...

Entered GENESEQ	14:34:43 ON 13 NOV 2023
L9 TRA L6 1- PN : 233 TERMS	
Entered WPINDEX	14:34:47 ON 13 NOV 2023
L10 232 SEA L9	 ...
L11 108 S L8 AND L10	 ...
Entered DCR	14:38:04 ON 13 NOV 2023
L12 1 S HUMIRA/CNS	 ...
Entered WPINDEX	14:41:30 ON 13 NOV 2023
L13 1605 S L12	 ...
L14 1668 S L13 OR L11	 ...

Dates to consider

- Date of publication is not the same as date of availability!
- Publication date vs. Entry date vs. Update date(s)
 - i.e., if record is retrieved by an index term, the update date of the indexing may become important

CAplus Publication Date vs. Entry Date

```
L2 ANSWER 7 OF 822 HCAPLUS COPYRIGHT 2023 ACS on STN
PI
PATENT NO.          KIND  DATE          APPLICATION NO.    DATE
-----
TW I710099          B    20201111      TW 2020-112808    20200416
CN 113539979        A    20211022      CN 2020-10326932  20200423
US 20210328339      A1   20211021      US 2020-16890128  20200602
US 11233324         B2   20220125
ED Entered STN: 12 Aug 2021
```

CAplus does not display Update Dates.

DWPI Publication Date vs. Entry Date

```
L3 ANSWER 1 OF 1 WPINDEX COPYRIGHT 2023 CLARIVATE on STN
PI TW I710099 B 20201111 (2021052)* ZH 38[0]
  US 20210328339 A1 20211021 (2021086) EN
  CN 113539979 A 20211022 (2021091) ZH
  US 11233324 B2 20220125 (2022009) EN
  TW 2021041740 A 20211101 (2022013) ZH
ED 20210629
UP 20230621
```