



Structure-based approaches to the indexing and retrieval of patent chemistry

Tim Miller
Head of Research
May 2010



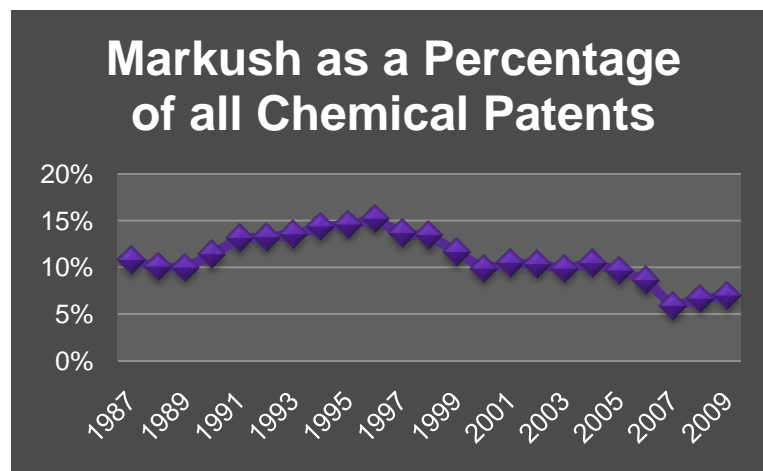
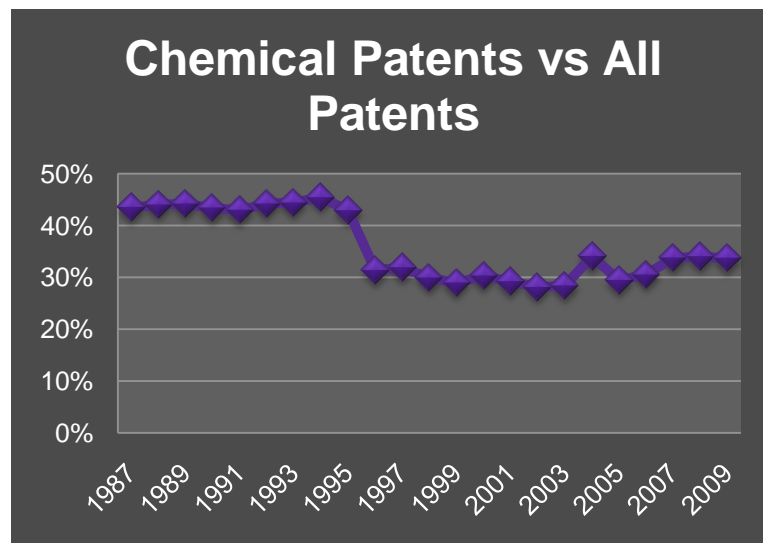
THOMSON REUTERS

TOPICS

- Chemistry in Patents
- Structure Indexing of patents
- New developments
- Challenges yet to be overcome

Chemistry in Patents

- Importance
 - Patents are essential to protect chemical inventions
 - 70% of patent information is never published elsewhere.
- Volumes
 - Chemical patenting is steady
 - Markush patenting as a percentage of all chemical patents is decreasing



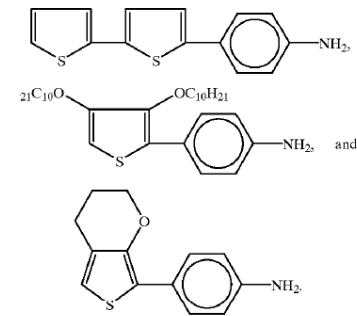
Chemistry in Patents

- Specific Compounds
- Markush
- Reactions & new Intermediates
- Polymers
- Inorganics

(57)

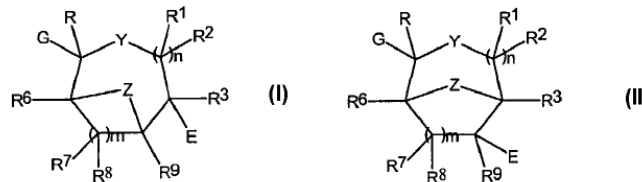
ABSTRACT

Provided are polymers prepared by the polymerization of monomers of the formulae:



1 Claim, No Drawings

(54) Title: PROCESS FOR PREPARATION OF BICYCLIC AND POLYCYCLIC MOLECULES



(57) Abstract: A method of synthesis of a bicyclic or polycyclic compound of formula (I) or formula (II) in which: E represents an electrophile; each of R, R1, R2, R3, R6, R7, R8, R9 and X independently represents the common organic substituent groups defined in claim 1; Y represents C(r12)R13, O, NR14, or S; Z represents O, NR15, S or CR16W; G represents W or X; W

represents an electron withdrawing group; X has the same definition as R, and W= W; and each of n and m represents an integer from 0 to 100. The method comprises the steps of (a) activating a compound of formula (III); 8b) subjecting a compound of formula (IV) to nucleophilic addition with the activated form of compound III; (c) subjecting the product of step (b) to ring closing metathesis; and (d) subjecting the product of step (c) to stereoselective ring closure. The methods of invention are useful in the synthesis of candidate pharmaceutical agents or intermediates in drugs synthesis.

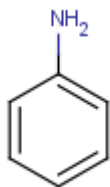
Indexing Patents, the Early Days

- Fragmentation Codes

- GREMAS
- DWPI
- CLAIMS/IFI

- Full Structures

- CAS Registry
- Beilstein



Aniline is the only specific compound given in the patent

The original Markush claim

Claim 1. The process for the manufacture of dyes which comprised coupling with a halogen-substituted pyrazolone, a diazotized unsulphonated material selected from the group consisting of aniline, homologues of aniline and halogen substitution products of aniline.

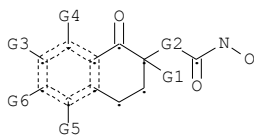
Code	Fragment
H1	Amine essentially present
M521	1 mononuclear heterocyclic ring
M520	No mononuclear heterocyclic ring
M320	No multivalent carbon chains
M210	C1-6 alkyl chain
M270	Alkyl attached to heteroatom
M273	Heteroatom is N
F011	Substitution on 1-position of heterocycle
H6411	Halogen linked to aromatic ring
H6422	Halogens linked to aromatic ring

Starting material indexed using
Derwent
Fragmentation codes

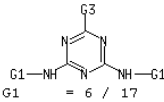
The 1980s – Topological Markush Searching Arrives

- University of Sheffield
 - GENSAAL
- Derwent, INPI and Questel
 - Markush DARC
- CAS
 - MARPAT

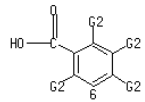
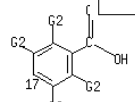
Screenshot of Markush DARC

MARKUSH/DARC	1/ 1	CN :0069-16801	MMS
-FG: 0-	-GM: 3/ 6-		AV NU CR
		H	
		O	
		N	
		CN	
		HAL	
1-O-CHK	1-N-CHK	CHK	ARY
1-O-ARY	1-N-CHK	1-F	HEA
1-O-CHK-ARY	1-N-CO1-CHK	F	HEF
		1-N	
		1-N	
		1-N	
CHK3=C1-4.			

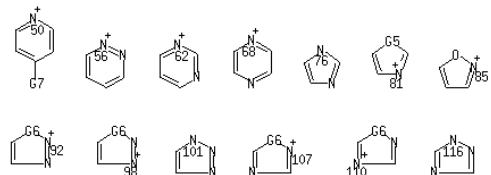
MSTR 1



G1-NH G3
G1 = 6 / 17

G2 = H / R / (Specifically claimed: alkyl (opt. substd. by 1 or more G4))
G3 = heterocycle <containing 1 or more N, attached through 1 or more N> (opt. substd.) / (Specifically claimed: 50 / 56 / 62 / 68 / 76 / 81 / 85 / 92 / 98 / 101 / 107 / 110 / 116 / 120 / 126 / 129 / 136 / 143 / 146 / 151 / 157 / 162 / 181 / 168 / 189 / 192 / 203) / (Examples: 235 / morpholino / pyrrolidino / piperidino / piperazino)



Screenshot of MARPAT

The New Millennium – Technology to Extract Structures from Patents

- Text mining
 - Temis, IBM, ReelTwo, etc.
- Name to Structure
 - ACD/Labs, CambridgeSoft, ChemAxon, etc.
- Chemical OCR
 - CLiDE, Kekule, SCAI
- New Products & Services
 - Elsevier's Patent Chemistry Database, SureChem, etc.

Today – A New Generation of Markush Tools?

- ChemAxon
 - Search, Enumerate
 - Available in current release of Marvin/JChem
- Digital Chemistry
 - Search, Enumerate
 - In prototype in TORUS
- DecrIPt
 - Enumerate, Rank overlaps
 - Available as service
- Symyx

Look at the cool things we can do!

- Quick, easy searches to establish where best to focus my efforts
- Overlap and difference analysis to find holes in an IP portfolio
- IP Screening of my combinatorial libraries

Are we done then?

Can the combination of new Markush-capable systems and smart text + image processing open the world of patent chemistry to the masses?

Challenge #1

**THERE'S A LOT OF IT, AND ITS
NOT PRETTY**



The “poor bloody indexer”

- There are a lot of Markush patents
 - They come in many languages
 - And they can be very complex
- Thomson Reuters has to employ around 90 Indexers just to keep up!

US 2002/026948 A1

Nov. 10, 2002

Chemical structures and text from a patent document, including a Markush structure (1) and various chemical formulas.

Chemical structures and text from a patent document, including a Markush structure (1) and various chemical formulas.

Size Matters

- The original Markush encompassed “only” a few thousand compounds
- Recent Markush are somewhat larger

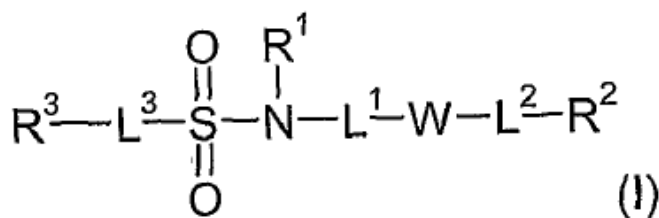
The screenshot displays the 'Markush Enumeration' software interface. At the top, a window titled 'Markush Enumeration' shows a Markush template for a pyrazolo[1,5-a]pyridine derivative with five substituent positions labeled R1 through R5. Below the template, the R-group definitions are listed: R1 = H, C, halogen; R2 = H, C, halogen; R3 = H, C, halogen; R4 = H, C, halogen; R5 = H, C, halogen. The main window, 'Enumerate a Markush structure', contains the following controls:

- Enumeration options:**
 - Full enumeration
 - Random enumeration
 - Markush reduction according to the hit
- Max structures:** 12
- Output to file:**
- Display options:**
 - Rows: 3
 - Columns: 4
 - Show R-groups
 - Colouring
- Full enumeration of this structure produces ~ 10³⁶ structures**
- Enumerate** button
- 12 structures enumerated**

Below the controls, a 3x4 grid of 12 chemical structures is displayed, representing the enumerated compounds. Each structure is a variation of the Markush template with specific R-groups substituted at the R1 through R5 positions.

Extraction Nightmares

(54) Title: NOVEL SULPHONAMIDE DERIVATIVES AS GLUCOCORTICOID RECEPTOR MODULATORS FOR THE TREATMENT OF INFLAMMATORY DISEASES



(57) Abstract: A compound of formula (I) or a pharmaceutically acceptable salt thereof; compositions comprising them, processes for preparing them and their use in medical therapy (for example modulating the glucocorticoid receptor in a warm blooded animal).

Enumerate a Markush structure

Enumeration options:

- Full enumeration
- Random enumeration
- Markush reduction according to the hit

Max structures: 12

Output to file

Display options:

Rows: 3

Columns: 4

Show R-groups

Colouring

Full enumeration of this structure produces ~ 10²⁶ structures

Enumerate

12 structures enumerated

Close



Challenge #2

CONTEXTUAL INFORMATION ASSOCIATION

Connecting Contextual Information

ID: 119

vmns

Markush structure	compound number
	0327-34307
	0327-34308
	0327-34309
	0327-34310
	0327-34311
	0327-34312

Title: New hydroxybenzoate salts formed as the reaction product of E-metanicotine compound and hydroxybenzoic acid, useful for treating e.g. CNS disorder, HIV-dementia, epilepsy, mania, and depression

Description: A hydroxybenzoate salt formed as the reaction product of an E-metanicotine compound of formula $Cy-C(E)=CE-(C(E)2)m-CEE1-N(Z1)-Z2$ (I), a hydroxybenzoic acid of formula (II) (where the hydroxy group can be present at a position ortho, meta or para to the carboxylic acid group), is new. The molar ratio of (I) to (II) is 1:2 - 2:1. Cy = 5- or 6-membered heteroaryl ring other than 5-isopropoxyppyridine (preferably rings of formula (i) or (ii)); E, E1 = H or alkyl (optionally halo substituted); Z1, Z2 = H or alkyl; m = 1 - 6; Z = non-hydrogen substituent selected from alkenyl, heterocyclyl, (cyclo)alkyl, aryl, alkylaryl, arylalkyl (all optionally substituted), F, Cl, Br

Assignees: ...

Patents: WO2006053039-A2 *
US20060122238-A1
EP1814853-A2
AU2005304575-A1
IN200701331-P2
CN101068784-A
JP2008519766-W
US20080249142-A1

Use: In the preparation of a medicament for treating a disorder that results from an alteration in normal neurotransmitter release, and CNS disorder (claimed). Also for ameliorating any of the symptoms associated with those conditions, diseases, and disorders including neurological disorders, neurodegenerative

Mechanism Of Action: CNS nicotinic receptor binder. The CNS nicotinic receptor binding efficacy of hydroxybenzoate salts with relevant receptor sites can be determined according to the methods described in US5597919. The results showed low binding constants thus the salts exhibits good high affinity binding to ...

Activity: CNS-Gen.; Neuroprotective; Antiinflammatory; Tranquilizer; Neuroleptic; Nootropic; Anti-HIV; Antiparkinsonian; Cerebroprotective; Anticonvulsant; Muscular-Gen.; Antimanic; Antidepressant; Vasotropic; Antismoking; Antiaddictive; Antialcoholic; Anorectic; Gastrointestinal-Gen.; Antidiarrheic; Antilulcer; Vulnerary.

Provisos, other information that can't be represented in the structure

Biological information

Even more context - How we do it in DWPI

Indexing
"paragraphs"
based on
IBM card
records

CMC UPB 20060724

M2 *01* C316 D022 D029 D621 D699 F011 F012 F013 F014 F015 F016 **Structural**

F021 F029 F111 F199 F431 F432 F499 F530 F541 F542 F552 F553 F554 F555 F556 F557 F558 F559 F560 F561 F562 F563 F564 F565 F566 F567 F568 F569 F570 F571 F572 F573 F574 F575 F576 F577 F578 F579 F580 F581 F582 F583 F584 F585 F586 F587 F588 F589 F590 F591 F592 F593 F594 F595 F596 F597 F598 F599 F600 F601 F602 F603 F604 F605 F606 F607 F608 F609 F610 F611 F612 F613 F614 F615 F616 F617 F618 F619 F620 F621 F622 F623 F624 F625 F626 F627 F628 F629 F630 F631 F632 F633 F634 F635 F636 F637 F638 F639 F640 F641 F642 F643 F644 F645 F646 F647 F648 F649 F650 F651 F652 F653 F654 F655 F656 F657 F658 F659 F660 F661 F662 F663 F664 F665 F666 F667 F668 F669 F670 F671 F672 F673 F674 F675 F676 F677 F678 F679 F680 F681 F682 F683 F684 F685 F686 F687 F688 F689 F690 F691 F692 F693 F694 F695 F696 F697 F698 F699 F700 F701 F702 F703 F704 F705 F706 F707 F708 F709 F710 F711 F712 F713 F714 F715 F716 F717 F718 F719 F720 F721 F722 F723 F724 F725 F726 F727 F728 F729 F730 F731 F732 F733 F734 F735 F736 F737 F738 F739 F740 F741 F742 F743 F744 F745 F746 F747 F748 F749 F750 F751 F752 F753 F754 F755 F756 F757 F758 F759 F760 F761 F762 F763 F764 F765 F766 F767 F768 F769 F770 F771 F772 F773 F774 F775 F776 F777 F778 F779 F780 F781 F782 F783 F784 F785 F786 F787 F788 F789 F790 F791 F792 F793 F794 F795 F796 F797 F798 F799 F800 F801 F802 F803 F804 F805 F806 F807 F808 F809 F810 F811 F812 F813 F814 F815 F816 F817 F818 F819 F820 F821 F822 F823 F824 F825 F826 F827 F828 F829 F830 F831 F832 F833 F834 F835 F836 F837 F838 F839 F840 F841 F842 F843 F844 F845 F846 F847 F848 F849 F850 F851 F852 F853 F854 F855 F856 F857 F858 F859 F860 F861 F862 F863 F864 F865 F866 F867 F868 F869 F870 F871 F872 F873 F874 F875 F876 F877 F878 F879 F880 F881 F882 F883 F884 F885 F886 F887 F888 F889 F890 F891 F892 F893 F894 F895 F896 F897 F898 F899 F900 F901 F902 F903 F904 F905 F906 F907 F908 F909 F910 F911 F912 F913 F914 F915 F916 F917 F918 F919 F920 F921 F922 F923 F924 F925 F926 F927 F928 F929 F930 F931 F932 F933 F934 F935 F936 F937 F938 F939 F940 F941 F942 F943 F944 F945 F946 F947 F948 F949 F950 F951 F952 F953 F954 F955 F956 F957 F958 F959 F960 F961 F962 F963 F964 F965 F966 F967 F968 F969 F970 F971 F972 F973 F974 F975 F976 F977 F978 F979 F980 F981 F982 F983 F984 F985 F986 F987 F988 F989 F990 F991 F992 F993 F994 F995 F996 F997 F998 F999

Synthesis

Descriptors

RIN: 00210 00211 00212
MCN: 0327-34301-K 0327-34301-P

M2 *02*
M720 N221 N225 N242 N311 N421 N422 N512 N513 P210 P420 P442 P444
P446 P448 P451 P510 P517 P617 P625 P641 P642 P646 P714 P731 P735
P738 P942 M905 M904

MCN: 0327-34302-K 0327-34302-P

M2 *03*
M533 M540 M541 M542 M543 M630 M650 M720 N221 N225 N242 N311 N421
N422 N512 N513 P210 P420 P442 P444 P446 P448 P451 P510 P517 P617
P625 P641 P642 P646 P714 P731 P735 P738 P942 M905 M904
MCN: 0327-34304-K 0327-34304-P

M2 *04*
N513 P210 P420 P442 P444 P446 P448 P451 P510 P517 P617 P625 P641
P642 P646 P714 P731 P735 P738 P942 M905 M904
MCN: 0327-34303-K 0327-34303-P

M2 *05*
M542 M543 M630 M650 M720 N221 N225 N242 N311 N421 N422 N512 N513
P210 P420 P442 P444 P446 P448 P451 P510 P517 P617 P625 P641 P642
P646 P714 P731 P735 P738 P942 M905 M904
MCN: 0327-34305-K 0327-34305-P

M2 *06*
M541 M542 M543 M630 M650 M720 N221 N225 N242 N311 N421 N422 N512
N513 P210 P420 P442 P444 P446 P448 P451 P510 P517 P617 P625 P641
P642 P646 P714 P731 P735 P738 P942 M905 M904
MCN: 0327-34306-K 0327-34306-P

M2 *07* M510 M511 M520 M521 M530 M531 M540 M541 M630 M650 M720 N221 N225
N242 N311 N421 N422 N512 N513 P210 P420 P442 P444 P446 P448 P451
P510 P517 P617 P625 P641 P642 P646 P714 P731 P735 P738 P942
M905 M904

MCN: 0327-34307-K 0327-34307-P

M2 *08* M531 M540 M630 M650 M720 N221 N225 N242 N311 N421 N422 N512 N513
P210 P420 P442 P444 P446 P448 P451 P510 P517 P617 P625 P641 P642
P646 P714 P731 P735 P738 P942 M905 M904

MCN: 0327-34308-K 0327-34308-P

M2 *09* M531 M540 M630 M650 M720 N221 N225 N242 N311 N421 N422 N512 N513
P210 P420 P442 P444 P446 P448 P451 P510 P517 P617 P625 P641 P642
P646 P714 P731 P735 P738 P942 M905 M904

MCN: 0327-34309-K 0327-34309-P

M2 *10* M630 M650 M720 N221 N225 N242 N311 N421 N422 N512 N513 P210 P420
P442 P444 P446 P448 P451 P510 P517 P617 P625 P641 P642 P646 P714
P731 P735 P738 P942 M905 M904

MCN: 0327-34310-K 0327-34310-P

M2 *11* M521 M530 M531 M540 M630 M650 M720 M800 N221 N225 N242 N311 N421
N422 N512 N513 P210 P420 P442 P444 P446 P448 P451 P510 P517 P617
P625 P641 P642 P646 P714 P731 P735 P738 P942 M905 M904
DCN: RAMSYO-K RAMSYO-P

DCR: 1310317-K 1310317-P

M2 *12* M391 M413 M510 M520 M521 M530 M531 M540 M630 M650 M720 M800 N221
N225 N242 N311 N421 N422 N512 N513 P210 P420 P442 P444 P446 P448
P451 P510 P517 P617 P625 P641 P642 P646 P714 P731 P735 P738 P942
M905 M904

DCN: RAMSY-P-K RAMSY-P

DCR: 1310318-K 1310318-P

M2 *13* N512 N513 P210 P420 P442 P444 P446 P448 P451 P510 P517 P617 P625
P641 P642 P646 P714 P731 P735 P738 P942 M905 M904
DCN: RAMSYQ-K RAMSYQ-P

DCR: 1310319-K 1310319-P

M2 *14* M520 M521 M522 M530 M531 M532 M540 M541 M720 N209 N221 N225 N231
N242 N309 N311 N361 N421 N422 N512 N513 P446 M905 M904
RIN: 00210 00211 00212

MCN: 0327-34311-K 0327-34311-P

M2 *15* M393 M413 M415 M510 M520 M521 M530 M540 M541 M720 N209 N221 N225
N231 N242 N309 N311 N361 N421 N422 N512 N513 P446 M905 M904
RIN: 00060 00061 00074 00076 00081 00083 00084 00087 00088 00089
00094 00096 00102 00105 00110 00115 00131 00133 00135 00137
00138 00734 009740 11555 41038 41246 42005 45813

MCN: 0327-34312-K 0327-34312-P

M2 *16* F013 F015 F620 H1 H102 H181 H7 H721 M210 M211 M240 M273 M281
M314 M321 M332 M342 M373 M391 M413 M510 M521 M530 M540 M720 N209
N221 N225 N231 N242 N309 N311 N361 N421 N422 N512 N513 P446
M905 M904

DCN: RAMSYR-K RAMSYR-P

DCR: 1310320-K 1310320-P

M2 *17* F013 F015 F620 H1 H103 H181 H7 H721 M210 M211 M240 M273 M281
M282 M314 M321 M332 M342 M373 M391 M413 M510 M521 M530 M540 M720
N209 N221 N225 N231 N242 N309 N311 N361 N421 N422 N512 N513 P446
M905 M904

DCN: RAMSYS-K RAMSYS-P

DCR: 1310321-K 1310321-P

**Specific
compounds**



Interpretation

- The indexing in Paragraph 1 describes the synthesis of a known compound

M2 *01* C316 D022 D029 D621 D699 F011 F012 F013 F014 F015 F016 F019 F020
 F021 F029 F111 F199 F431 F432 F499 F530 F541 F542 F551 F552 F580
 F599 G001 G002 G003 G010 G011 G012 G013 G014 G015 G016 G017 G018
 G019 G020 G021 G022 G029 G030 G031 G032 G039 G040 G050 G051 G100
 G111 G112 G113 G221 G299 G553 G563 H100 H101 H102 H103 H121 H122
 H141 H142 H143 H161 H181 H182 H183 H321 H341 H342 H343 H401 H402
 H403 H404 H405 H421 H441 H442 H443 H444 H481 H482 H483 H494 H521
 H522 H523 H541 H542 H543 H561 H581 H582 H583 H592 H594 H599 H601
 H602 H603 H604 H608 H609 H621 H641 H642 H643 H681 H682 H683 H685
 H689 H715 H721 H722 H723 J011 J012 J013 J014 J111 J131 J132 J133
 J171 J172 J173 J211 J221 J222 J231 J232 J241 J242 J261 J271 J272
 J273 J311 J321 J322 J331 J332 J341 J342 J361 J371 J372 J373 J411
 J431 J432 J471 J521 J581 J582 J583 J592 K442 K499 K510 K599 K742
 K799 L142 L143 L199 L462 L463 L499 L921 L922 L941 L943 L999 M111
 M112 M113 M115 M116 M119 M121 M122 M123 M124 M125 M126 M129 M131
 M132 M133 M135 M136 M137 M139 M141 M142 M143 M150 M210 M211 M212
 M213 M214 M215 M216 M220 M221 M222 M223 M224 M225 M226 M231 M232
 M233 M240 M273 M280 M281 M311 M314 M315 M316 M320 M321 M322 M323
 M331 M332 M333 M334 M340 M342 M343 M344 M349 M352 M353 M373 M391
 M392 M393 M412 M413 M414 M510 M511 M512 M520 M521 M522 M523 M530
 M531 M532 M533 M540 M541 M542 M543 M630 M650 M720 N221 N225 N242
 N311 N421 N422 N512 N513 P210 P420 P442 P444 P446 P448 P451 P510
 P517 P617 P625 P641 P642 P646 P714 P731 P735 P738 P942 M905
 M904
 RIN: 00210 00211 00212
MCN: 0327-34301-K 0327-34301-P

P210 : Antiviral
 P420 : Antiinflammatory
 P442 : Anticonvulsant
 P444 : Antiparkinson
 P446 : Neuroleptic
 P448 : Anxiolytic
 P451 : Antidepressant
 P510 : Autnomic Nervous System
 P517 : Muscle Relaxant
 P617 : Antimetabolite
 P625 : Hormone Activity
 P641 : Alcoholism
 P642 : Smoking
 P646 : Antidote
 P714 : Anabolic
 P731 : Anorectic
 P735 : Antidiarrhoeal
 P738 : Ulcers
 P942 : Wound Treatment

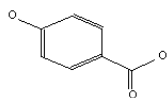
N221 : C=C-H H bond broken
 N225 : C-Hal bond broken
 N242 : C-O bond broken
 N311 : C-C bond formed
 N421 : Acid conditions
 N422 : Basic conditions
 N512 : 10-30°C
 N513 : 30-200°C

A: substance analysed/detected	Q: product defined by starting materials
C: catalyst	R: removing/purifying agent
D: detecting agent	S: starting material
E: excipient	T: therapeutic
K: known (always output for recent data if not N)	U: use of single compound
M: component of a mixture	V: reagent
N: new compound	X: substance removed
P: known compound produced	Z: miscellaneous

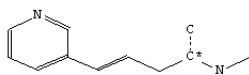
Matching Specifically Disclosed Compounds

CN.S METHYL- ((E)- (S)-1-METHYL-4-PYRIDIN-3-YL-BUT-3-ENYL)-AMINE
4-HYDROXY-BENZOATE
SDCN RAMSYO

CM 1



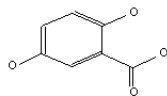
CM 2



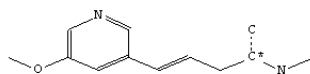
AN.S DCR-1310318

CN.S [(E)- (S)-4-(5-METHOXY-PYRIDIN-3-YL)-1-METHYL-BUT-3-ENYL]-METHYL-AMINE
2,5-DIHYDROXY-BENZOATE
SDCN RAMSYP

CM 1

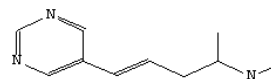


CM 2

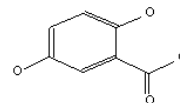


AN.S DCR-1310319

CN.S METHYL- ((E)-1-METHYL-4-PYRIDIN-3-YL-BUT-3-ENYL)-AMINE
2,5-DIHYDROXY-BENZOATE



CM 2



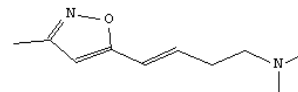
AN.S DCR-1310320

CN.S Methyl-[(E)-4-(3-methyl-isoxazol-5-yl)-but-3-enyl]-amine
SDCN RAMSYR



AN.S DCR-1310321

CN.S Dimethyl-[(E)-4-(3-methyl-isoxazol-5-yl)-but-3-enyl]-amine
SDCN RAMSYS



Challenge #3

MAKING SENSE OF MARKUSH



Where is my hit – and does it matter?

- Ranking of hits based on “nearness to the core of the invention”

Enumeration options:

- Full enumeration
- Random enumeration
- Markush reduction according to the hit

Max structures:

Output to file

Display options

Rows:

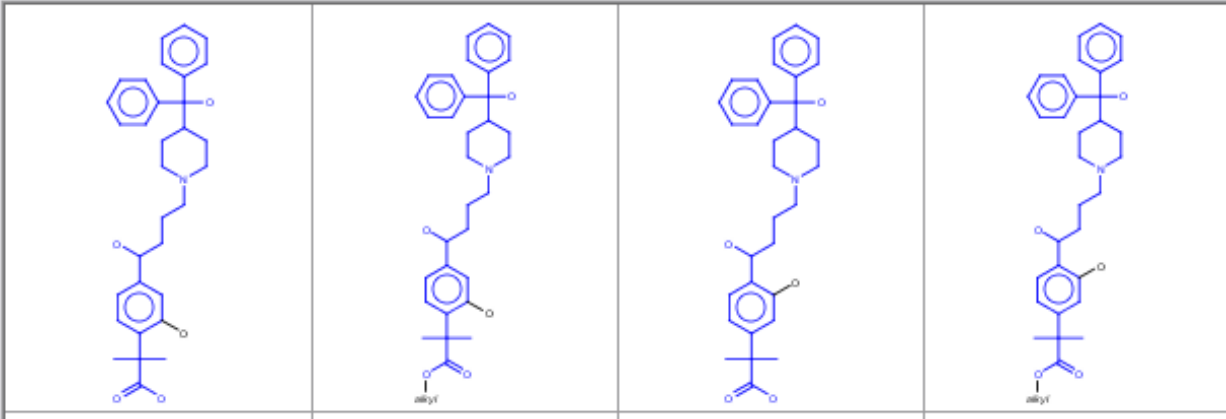
Columns:

Show R-groups

Colouring

Full enumeration of this structure produces 240 structures

4 structures enumerated

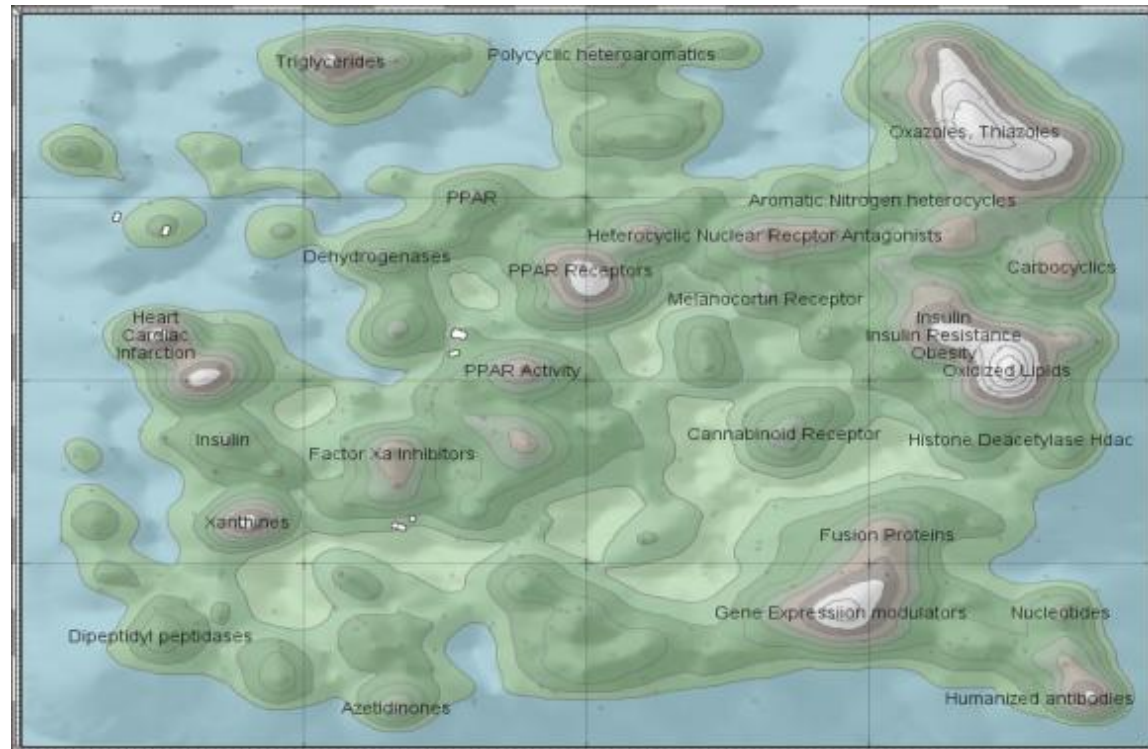


The screenshot shows a software interface for chemical enumeration. It features two main control panels: 'Enumeration options' and 'Display options'. The 'Enumeration options' panel includes radio buttons for 'Full enumeration', 'Random enumeration', and 'Markush reduction according to the hit' (which is selected), a 'Max structures' input field set to 10, and an 'Output to file' checkbox. The 'Display options' panel includes 'Rows' and 'Columns' dropdown menus set to 3 and 4 respectively, a 'Show R-groups' checkbox, and a 'Colouring' checkbox which is checked. To the right of these panels, text indicates 'Full enumeration of this structure produces 240 structures' and a button labeled 'Enumerate'. Below the controls, a status bar shows '4 structures enumerated'. The main area displays a 3x4 grid of chemical structures, all rendered in blue. Each structure is a complex molecule with a central core consisting of a benzene ring, a piperidine ring, and a piperazine ring, with various side chains and functional groups.

JChem's useful selective enumeration

Visualising Results

- Can we rank results in terms of “nearness”?
- Can we visualise the patent landscape in some way?



Thank You

Any questions?