

Searching and Subsets in ConQuest

Developed using
2022.1.0 CSD Release
CSD version 5.43 (November 2021) + 1 update

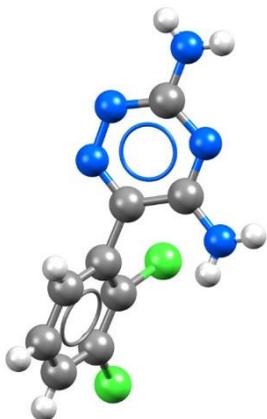


Table of Contents	
Introduction.....	2
Objectives.....	2
Pre-required skills.....	2
Materials.....	2
Example 1. Basic Searching in ConQuest.....	3
Conclusion	5
Example 2. Restricting searches to a subset	6
Conclusion	7
Further Exercises	8
Conclusions.....	8
Next steps.....	8
Bonus Exercise. Loading Subsets.....	9
Glossary	12
Review. ConQuest Interface.....	13
Review. Draw Window	14
ConQuest sketching conventions.....	14

Introduction

ConQuest is the desktop search interface to the Cambridge Structural Database (CSD). All text, numeric and structural data stored within the CSD can be searched using ConQuest. ConQuest provides an extensive range of flexible search options including searching based on compound name, formula, elemental composition, and literature search to name a few. With over a million entries in the database, individual structures can be missed. Subsets break the database down into more manageable subcategories and is an efficient way to search a certain category of compounds. As of the 2022.1 release, the latest additions to the subsets are the Election diffraction, High pressure, Hydrates, Polymorphs and three [MOF](#) dimensionality subsets. The teaching subset is also now available in ConQuest.

Before beginning this workshop, ensure that you have a registered copy of CSD-Core or CSD-Enterprise installed on your computer. Please contact your site administrator or workshop host for further information.

Objectives

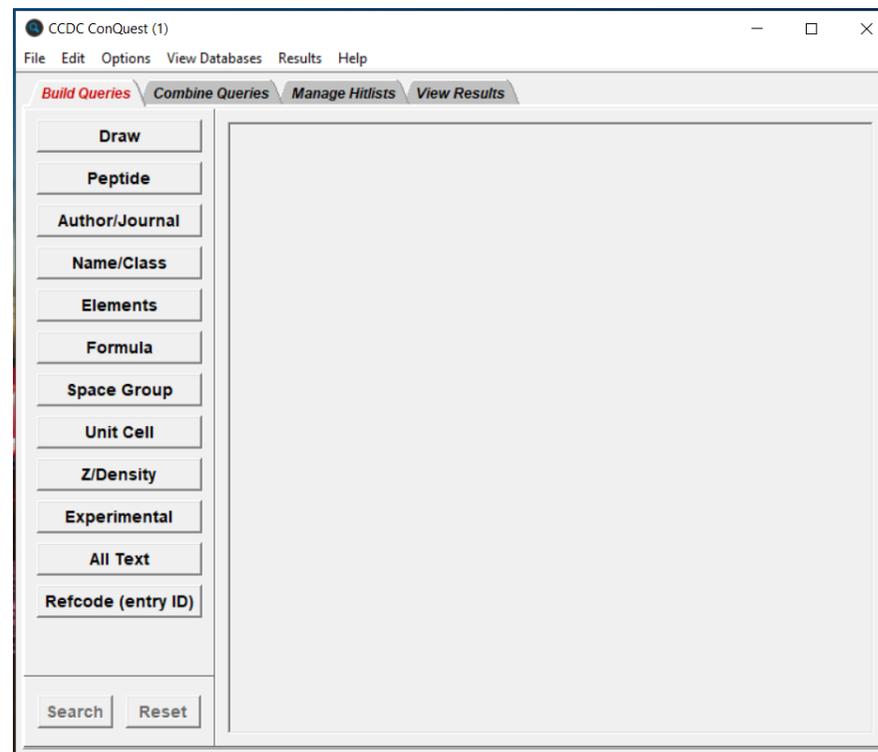
This workshop will guide you through sketching and using curated subsets in the CSD. At the end of this tutorial, you should be able to:

- Draw a 2D structure in ConQuest and conduct a basic search.
- Restrict a search to a chosen subset.
- Locate and load subsets.
- Search structures at the intersection of two subsets.

This workshop will take approximately **25** minutes to be completed.

Note: The words in *Blue Italic* in the text are reported in the [Glossary](#) at the end of this handout. A *review* of the ConQuest interface and of the ConQuest Draw window is also at the end of this handout.

Note: Your results could differ based on the CSD version that you have installed. Do not worry if you are on a different version and get a different number of structures in your [hitlists](#). A similar order of magnitude could be expected and don't forget to always approach the results critically!



Pre-required skills

This is an entry-level session; no previous experience is required.

Materials

There are no additional materials required for this workshop.

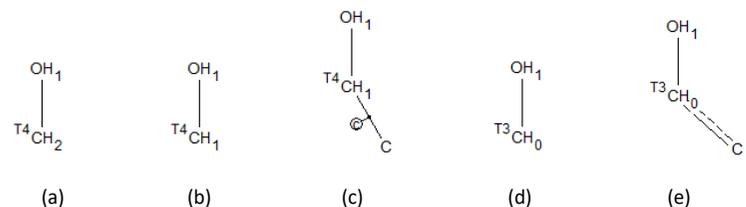
Example 1. Basic Searching in ConQuest

All drawing takes place in the central white area of the *Draw* window. In addition to creating 2D chemical structure sketches, the *Draw* window allows for the inclusion of 3D parameters for searching or for filtering.

For this example, we will sketch and search for fragments (a) and (b). If you complete the workshop early and are interested, you can draw the remaining three fragments and conduct a search.

This example will demonstrate sketching techniques in ConQuest.

1. Launch ConQuest by clicking the ConQuest Icon  on your desktop or launching it from the Start or Applications menu and open the sketcher by clicking the **Draw** button.
2. Sketch a simple alcohol fragment. First, add a carbon atom by clicking anywhere in the drawing area, making sure you have carbon selected ("C" in the white square at the bottom of the draw window). Then, add an oxygen atom to it with a single bond. To do so, select "O", oxygen, from the elements at the bottom of the draw window, make sure the bond type selected is "Single", and click and drag from the carbon atom.
3. To search for fragments with one hydrogen bonded to the oxygen, right click on the O, choose *Hydrogens* and then 1 from the drop-down menu to add the H atom.
4. To ensure that the carbon atom is bonded to at least 2 hydrogen atoms and one other atom in addition to the OH group, first add two H atoms as in **Step 3** above, but this time select 2 hydrogens. Then right-click the C again and choose *Number of Bonded Atoms*, and then 4 from the drop-down menu. This will produce fragment "a" from the list show in the introduction to this exercise.
5. To run the search, click the **Search** button at the bottom right of the Draw window.

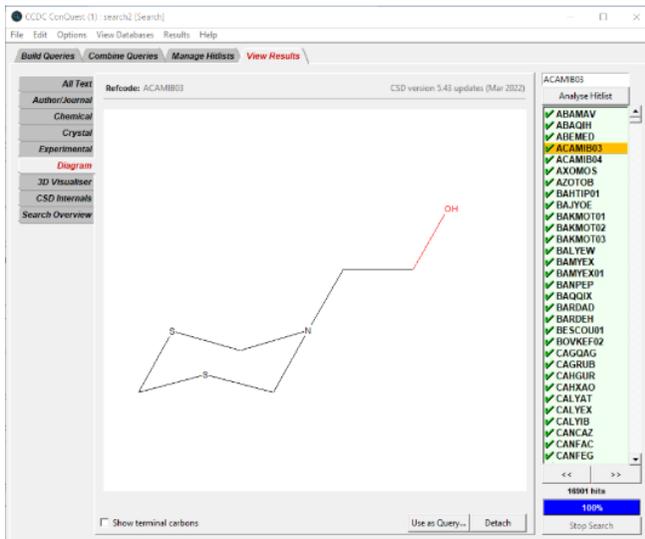


These subtly different definitions would retrieve the following types of functional group:

- primary alcohols
- cyclic and acyclic secondary alcohols
- cyclic secondary alcohols only
- unsaturated cyclic and acyclic alcohols and carboxylic acids
- unsaturated cyclic alcohols only e.g. phenols

In the **Search Setup** pop-up, tick the boxes for *3D coordinates determined* and *Only Organics* (make sure both the box next to *Only* and the one next to *Organics* are selected). Then click the **Start Search** button.

6. Based on your CSD version, you should get results in the order of or over 16,901 with this query. Use the arrow keys to scroll through the reffcode list to view the results. Explore the results using the tabs on the right-hand side, for example to see the 3D structure, or learn more about the crystal structure and the experimental conditions. Observe how the fragment we searched for is highlighted in the 2D diagram and on the 3D structure. *What do you find?*



7. When you are done looking at your results, click the *Build Queries* tab to return to the Query window. You should see your first search fragment listed as Query 1.
8. We are now going to search for a secondary alcohol, as shown as fragment b in the list at next to the introduction. We will save time, by editing the drawing that we did for our first search. Click the **Edit...** button in the query window to return to the **Draw** window.
9. Right-click on the C atom and choose *Hydrogens*, and *1* from the drop-down menu as in **Step 3** above. Leave the *Number of Bonded Atoms* set to 4 to match a secondary alcohol (fragment "b").

5

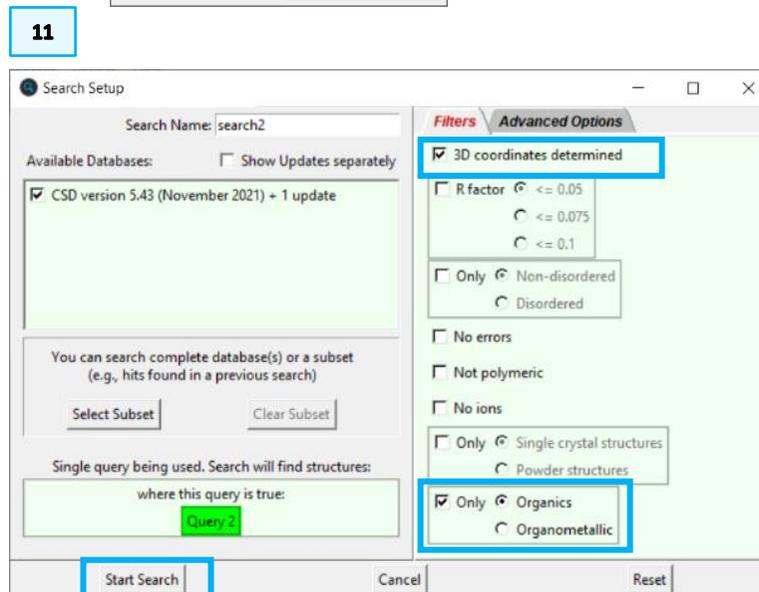
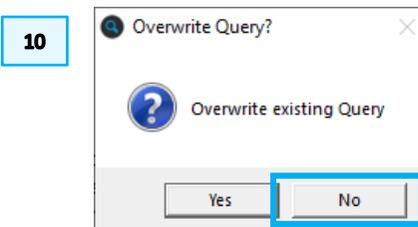
7

9

- Click **Search** as in **Step 5** above. A pop-up warning will ask if you want to overwrite the existing Query. Click **No** to save this as a new Query.
- In the **Search Setup**, make sure the “3D coordinates determined” and “Only Organics” filters are ticked. Then, start the search.
- Based on your CSD version, you should get in the order or over 30,650 hits with this new query. Use the arrow keys to scroll through the refcode list to view the results. *Observe: how are the hits different from the first search?*

Conclusion

You should now be familiar with sketching substructures using the ConQuest *Draw* window; how to edit searches; how to change bond and atom types; and how to view the data for each hit in your results list.



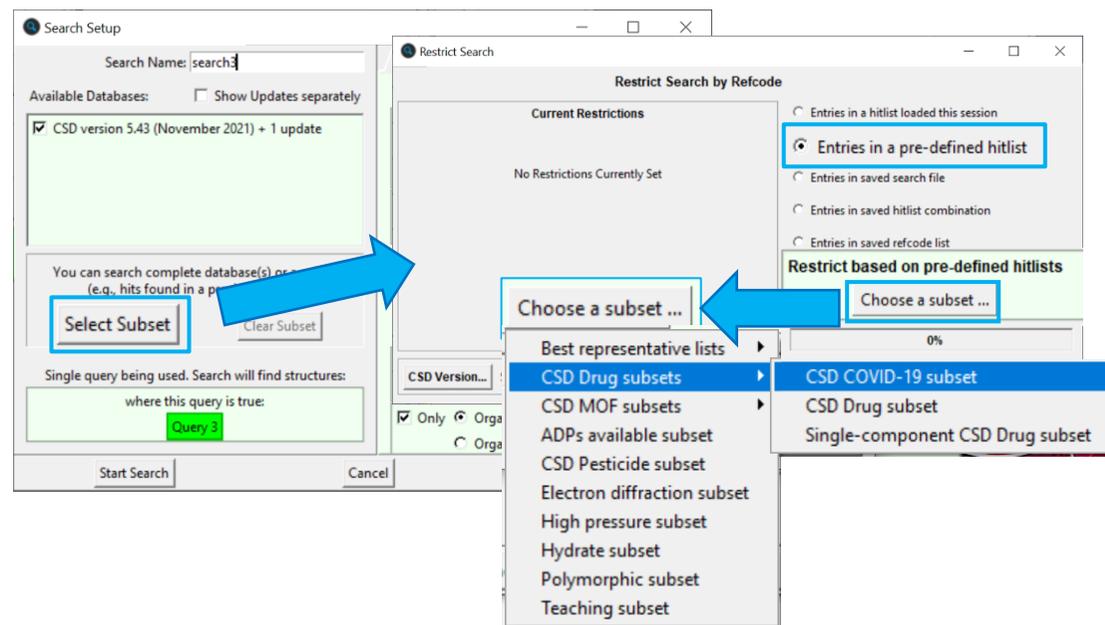
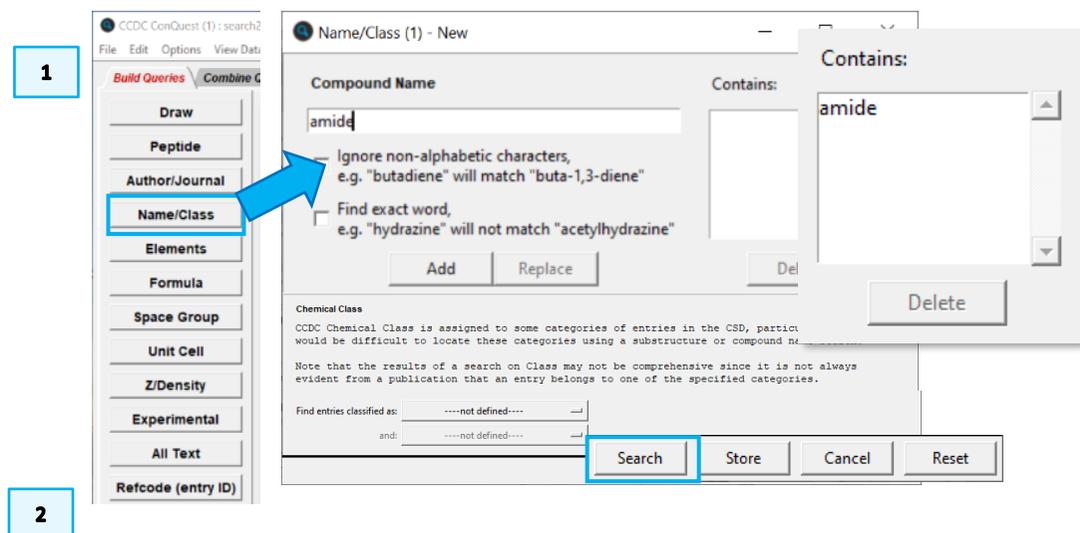
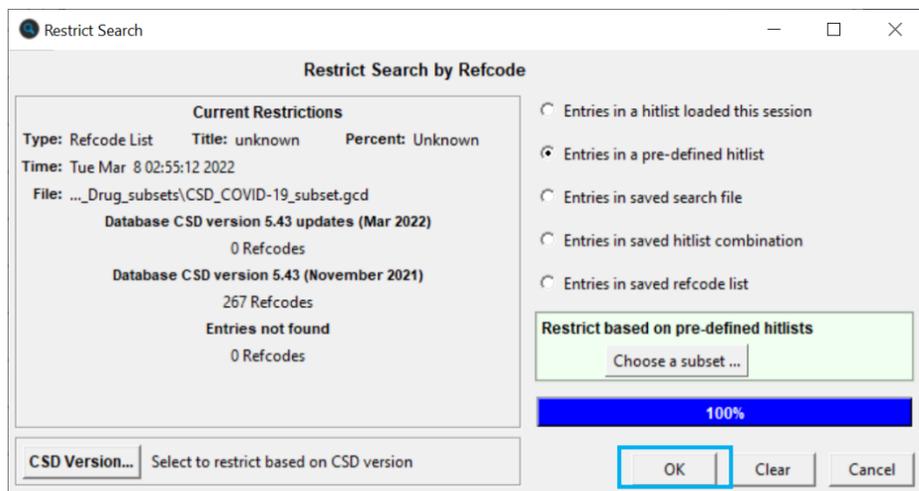
12

Example 2. Restricting searches to a subset

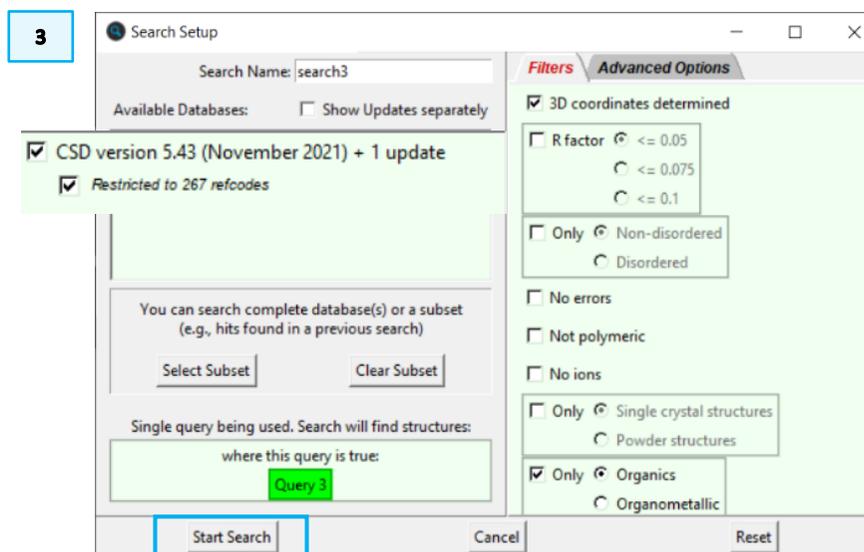
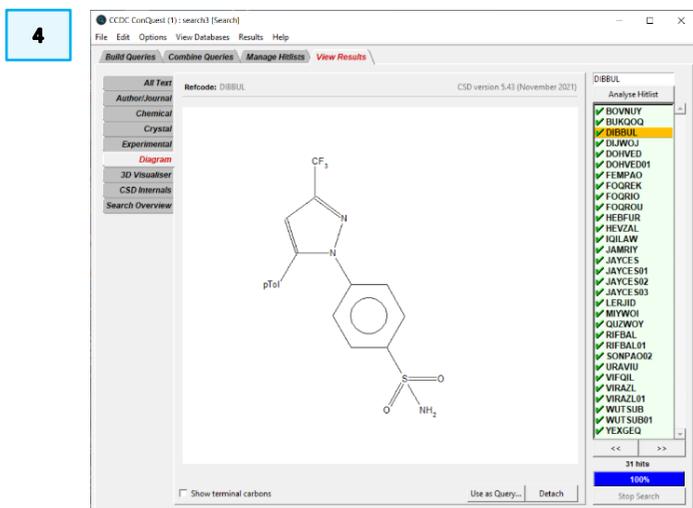
There are three ways to use subsets; you can load subsets in ConQuest to view the structures, restrict your search to a specific subset or through the CSD Python API. To begin with, we will focus on how to restrict the search to a subset. You can find how to load subsets in the bonus exercise. If instead you are interested in accessing subsets through the API, we recommend looking at the [API workshops on this page](#) and [user guide](#).

For the purposes of this workshop, we will assume that you are interested in CSD entries with “amide” in the compound name in the [CSD COVID-19 subset](#).

1. Click on the *Build Queries* tab and click on the **Name/Class** button to start a text-based search. Type the word *amide* in the “Compound Name” box. Click **Add** to include the word amide in the “Contains:” box. Click **Search** to begin the search.
2. In the *Search Setup* dialog box, we will select the subset we want to search from. Click on the **Select Subset** button to open the *Restrict Search* window. In this window, select *Entries in a pre-defined hitlist* to load the CSD defined subsets. Click on Choose a subset..., from the drop-down menu, select **CSD Drug subsets > CSD COVID-19 subset**. Once the subset is loaded, click **OK** to close the window.

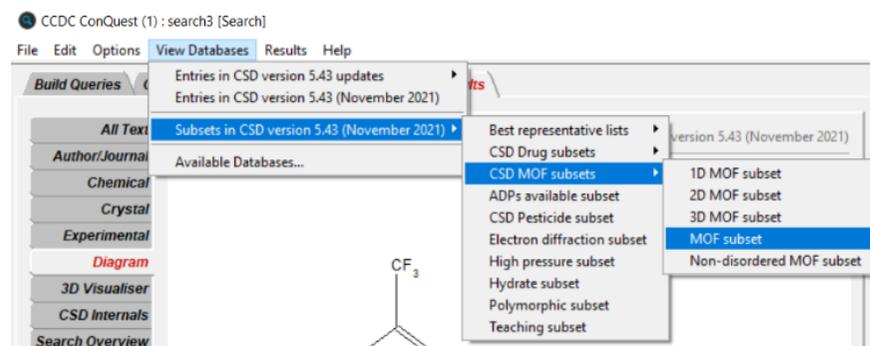


- In the *Search Setup* window, note that under *Available Databases*, there is an additional line below the CSD version that indicates the search is restricted to 267 refcodes, which is the number of structures in the CSD COVID-19 subset (note: this number might change based on your CSD version). Make sure to tick the boxes for “3D coordinates determined” and “Only Organics” to apply these filters. Then click **Start Search** to begin the search.
- There should be at least 31 structures returned. You can look through the results to assess where the amide groups are located and use such information to either refine or expand future searches.



Conclusion

You have seen one way to use subsets in your ConQuest searching by restricting a search to the CSD COVID-19 subset. You have also seen how to create a query based on the compound name. There are many subsets that you can explore. For example, the MOF subset can be extremely useful if you want to search specifically for [MOFs](#).



Further Exercises

- Can you sketch the remaining alcohol substructure from Example 1 and search the entire database?
- Can you find structures that are common to the Hydrate and MOF subsets? (*Hint*: see bonus exercise)
- How would you conduct a name search for other functional groups in the CSD COVID-19 subset?

Conclusions

This workshop introduced how to conduct basic searches and how to use subsets in ConQuest. You should now be familiar with:

- Sketching substructures using the ConQuest *Draw* window.
- Restricting searches to a specific subset.

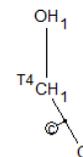
In the bonus exercise you will see in more detail:

- Loading pre-defined subsets.
- Finding structures in multiple subsets.

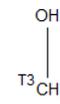
Next steps

After this workshop, you can explore more ConQuest exercises in the self-guided workshops available in the [CSD-Core](https://www.ccdc.cam.ac.uk/Community/educationalresources/workshop-materials/csd-core-workshops/) workshops area on our website:

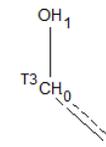
<https://www.ccdc.cam.ac.uk/Community/educationalresources/workshop-materials/csd-core-workshops/>. Please see the [ConQuest documentation and other available ConQuest documents here](#) .



(c)



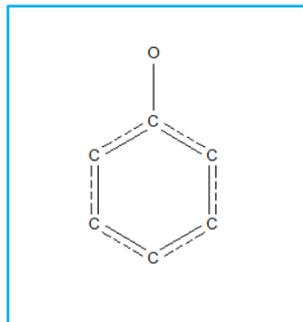
(d)



(e)

Bonus Exercise. Loading Subsets

There are three ways to use subsets; you can load subsets in ConQuest to view the structures, restrict your search to a specific subset or through the CSD Python API. For this bonus exercise, we will focus on the first method. If you are interested in accessing subsets through the API, we recommend looking at the [API workshop](#) and [user guide](#).



For the purposes of this exercise, we will assume that you are interested in CSD entries with a specific functional group, which are in the [CSD Drug subset](#) (containing the approved Drug list by the DrugBank) and in the [CSD teaching subset](#).

1. Launch ConQuest by clicking the ConQuest Icon  on your desktop or launching it from the Start or Applications menu. In the main ConQuest window, click on **View Databases** then click on *Subsets in CSD version 5.43 (November 2021)*. A list of available subsets will be displayed.
2. You will first load the drug subset by clicking on *CSD Drug subsets* and then on *CSD Drug subset*. It will take a few seconds to load and there should be at least 13,700 structures in the subset. Repeat the procedure to load the *Teaching subset*. There should be at least 837 entries. With both subsets loaded, *how can you determine how many structures are in both subsets?*
3. You can answer this question by using the combine [hitlists](#) function from the Manage Hitlists tab. Click on the *Manage Hitlists* tab. We want to find out how many structures are in both subsets. Under the *Combine Hitlists* menu, List A should be CSD_Drug_subset. To select the teaching subset as List B, click on List B and in the drop-down box, choose Teaching subset. Under *Generate a List of Entries*, the radio next to "common to List A and List B" should be selected. Click **OK**, a new hitlist named "combination1" should be generated with 107 hits.

Name	Hits	Type
CSD_Drug_subset	13791	Refcode List
combination1	107	Combination
teaching_subset	837	Refcode List

1

CCDC ConQuest (1)

File Edit Options View Databases Results Help

Build Queries

Draw

Peptide

Author/Journal

Name/Class

Elements

Entries in CSD version 5.43 updates

Entries in CSD version 5.43 (November 2021)

Subsets in CSD version 5.43 (November 2021)

Available Databases...

Best representative lists

CSD Drug subsets

CSD MOF subsets

ADPs available subset

CSD Pesticide subset

Electron diffraction subset

High pressure subset

Hydrate subset

Polymorphic subset

Teaching subset

CSD COVID-19 subset

CSD Drug subset

Single-component CSD Drug subset

2

CCDC ConQuest (1) : CSD Drug subset [Refcode List] - CSD Drug subset.gcd

File Edit Options View Databases Results Help

Build Queries Combine Queries Manage Hitlists View Results

Author/Journal

Chemical

Crystal

Experimental

Diagram

3D Visualiser

CSD Itemsets

Search Overview

Refcode: ABARQ

CSD version 5.43 (November 2021)

ABARQ

ABARQ0

ABARQ1

ABARQ2

ABARQ3

ABARQ4

ABARQ5

ABARQ6

ABARQ7

ABARQ8

ABARQ9

ABARQ10

ABARQ11

ABARQ12

ABARQ13

ABARQ14

ABARQ15

ABARQ16

ABARQ17

ABARQ18

ABARQ19

ABARQ20

ABARQ21

ABARQ22

ABARQ23

ABARQ24

ABARQ25

ABARQ26

ABARQ27

ABARQ28

ABARQ29

ABARQ30

ABARQ31

ABARQ32

ABARQ33

ABARQ34

ABARQ35

ABARQ36

ABARQ37

ABARQ38

ABARQ39

ABARQ40

ABARQ41

ABARQ42

ABARQ43

ABARQ44

ABARQ45

ABARQ46

ABARQ47

ABARQ48

ABARQ49

ABARQ50

ABARQ51

ABARQ52

ABARQ53

ABARQ54

ABARQ55

ABARQ56

ABARQ57

ABARQ58

ABARQ59

ABARQ60

ABARQ61

ABARQ62

ABARQ63

ABARQ64

ABARQ65

ABARQ66

ABARQ67

ABARQ68

ABARQ69

ABARQ70

ABARQ71

ABARQ72

ABARQ73

ABARQ74

ABARQ75

ABARQ76

ABARQ77

ABARQ78

ABARQ79

ABARQ80

ABARQ81

ABARQ82

ABARQ83

ABARQ84

ABARQ85

ABARQ86

ABARQ87

ABARQ88

ABARQ89

ABARQ90

ABARQ91

ABARQ92

ABARQ93

ABARQ94

ABARQ95

ABARQ96

ABARQ97

ABARQ98

ABARQ99

ABARQ100

ABARQ101

ABARQ102

ABARQ103

ABARQ104

ABARQ105

ABARQ106

ABARQ107

ABARQ108

ABARQ109

ABARQ110

ABARQ111

ABARQ112

ABARQ113

ABARQ114

ABARQ115

ABARQ116

ABARQ117

ABARQ118

ABARQ119

ABARQ120

ABARQ121

ABARQ122

ABARQ123

ABARQ124

ABARQ125

ABARQ126

ABARQ127

ABARQ128

ABARQ129

ABARQ130

ABARQ131

ABARQ132

ABARQ133

ABARQ134

ABARQ135

ABARQ136

ABARQ137

ABARQ138

ABARQ139

ABARQ140

ABARQ141

ABARQ142

ABARQ143

ABARQ144

ABARQ145

ABARQ146

ABARQ147

ABARQ148

ABARQ149

ABARQ150

ABARQ151

ABARQ152

ABARQ153

ABARQ154

ABARQ155

ABARQ156

ABARQ157

ABARQ158

ABARQ159

ABARQ160

ABARQ161

ABARQ162

ABARQ163

ABARQ164

ABARQ165

ABARQ166

ABARQ167

ABARQ168

ABARQ169

ABARQ170

ABARQ171

ABARQ172

ABARQ173

ABARQ174

ABARQ175

ABARQ176

ABARQ177

ABARQ178

ABARQ179

ABARQ180

ABARQ181

ABARQ182

ABARQ183

ABARQ184

ABARQ185

ABARQ186

ABARQ187

ABARQ188

ABARQ189

ABARQ190

ABARQ191

ABARQ192

ABARQ193

ABARQ194

ABARQ195

ABARQ196

ABARQ197

ABARQ198

ABARQ199

ABARQ200

ABARQ201

ABARQ202

ABARQ203

ABARQ204

ABARQ205

ABARQ206

ABARQ207

ABARQ208

ABARQ209

ABARQ210

ABARQ211

ABARQ212

ABARQ213

ABARQ214

ABARQ215

ABARQ216

ABARQ217

ABARQ218

ABARQ219

ABARQ220

ABARQ221

ABARQ222

ABARQ223

ABARQ224

ABARQ225

ABARQ226

ABARQ227

ABARQ228

ABARQ229

ABARQ230

ABARQ231

ABARQ232

ABARQ233

ABARQ234

ABARQ235

ABARQ236

ABARQ237

ABARQ238

ABARQ239

ABARQ240

ABARQ241

ABARQ242

ABARQ243

ABARQ244

ABARQ245

ABARQ246

ABARQ247

ABARQ248

ABARQ249

ABARQ250

ABARQ251

ABARQ252

ABARQ253

ABARQ254

ABARQ255

ABARQ256

ABARQ257

ABARQ258

ABARQ259

ABARQ260

ABARQ261

ABARQ262

ABARQ263

ABARQ264

ABARQ265

ABARQ266

ABARQ267

ABARQ268

ABARQ269

ABARQ270

ABARQ271

ABARQ272

ABARQ273

ABARQ274

ABARQ275

ABARQ276

ABARQ277

ABARQ278

ABARQ279

ABARQ280

ABARQ281

ABARQ282

ABARQ283

ABARQ284

ABARQ285

ABARQ286

ABARQ287

ABARQ288

ABARQ289

ABARQ290

ABARQ291

ABARQ292

ABARQ293

ABARQ294

ABARQ295

ABARQ296

ABARQ297

ABARQ298

ABARQ299

ABARQ300

ABARQ301

ABARQ302

ABARQ303

ABARQ304

ABARQ305

ABARQ306

ABARQ307

ABARQ308

ABARQ309

ABARQ310

ABARQ311

ABARQ312

ABARQ313

ABARQ314

ABARQ315

ABARQ316

ABARQ317

ABARQ318

ABARQ319

ABARQ320

ABARQ321

ABARQ322

ABARQ323

ABARQ324

ABARQ325

ABARQ326

ABARQ327

ABARQ328

ABARQ329

ABARQ330

ABARQ331

ABARQ332

ABARQ333

ABARQ334

ABARQ335

ABARQ336

ABARQ337

ABARQ338

ABARQ339

ABARQ340

ABARQ341

ABARQ342

ABARQ343

ABARQ344

ABARQ345

ABARQ346

ABARQ347

ABARQ348

ABARQ349

ABARQ350

ABARQ351

ABARQ352

ABARQ353

ABARQ354

ABARQ355

ABARQ356

ABARQ357

ABARQ358

ABARQ359

ABARQ360

ABARQ361

ABARQ362

ABARQ363

ABARQ364

ABARQ365

ABARQ366

ABARQ367

ABARQ368

ABARQ369

ABARQ370

ABARQ371

ABARQ372

ABARQ373

ABARQ374

ABARQ375

ABARQ376

ABARQ377

ABARQ378

ABARQ379

ABARQ380

ABARQ381

ABARQ382

ABARQ383

ABARQ384

ABARQ385

ABARQ386

ABARQ387

ABARQ388

ABARQ389

ABARQ390

ABARQ391

ABARQ392

ABARQ393

ABARQ394

ABARQ395

ABARQ396

ABARQ397

ABARQ398

ABARQ399

ABARQ400

ABARQ401

ABARQ402

ABARQ403

ABARQ404

ABARQ405

ABARQ406

ABARQ407

ABARQ408

ABARQ409

ABARQ410

ABARQ411

ABARQ412

ABARQ413

ABARQ414

ABARQ415

ABARQ416

ABARQ417

ABARQ418

ABARQ419

ABARQ420

ABARQ421

ABARQ422

ABARQ423

ABARQ424

ABARQ425

ABARQ426

ABARQ427

ABARQ428

ABARQ429

ABARQ430

ABARQ431

ABARQ432

ABARQ433

ABARQ434

ABARQ435

ABARQ436

ABARQ437

ABARQ438

ABARQ439

ABARQ440

ABARQ441

ABARQ442

ABARQ443

ABARQ444

ABARQ445

ABARQ446

ABARQ447

ABARQ448

ABARQ449

ABARQ450

ABARQ451

ABARQ452

ABARQ453

ABARQ454

ABARQ455

ABARQ456

ABARQ457

ABARQ458

ABARQ459

ABARQ460

ABARQ461

ABARQ462

ABARQ463

ABARQ464

ABARQ465

ABARQ466

ABARQ467

ABARQ468

ABARQ469

ABARQ470

ABARQ471

ABARQ472

ABARQ473

ABARQ474

ABARQ475

ABARQ476

ABARQ477

ABARQ478

ABARQ479

ABARQ480

ABARQ481

ABARQ482

ABARQ483

ABARQ484

ABARQ485

ABARQ486

ABARQ487

ABARQ488

ABARQ489

ABARQ490

ABARQ491

ABARQ492

ABARQ493

ABARQ494

ABARQ495

ABARQ496

ABARQ497

ABARQ498

ABARQ499

ABARQ500

ABARQ501

ABARQ502

ABARQ503

ABARQ504

ABARQ505

ABARQ506

ABARQ507

ABARQ508

ABARQ509

ABARQ510

ABARQ511

ABARQ512

ABARQ513

ABARQ514

ABARQ515

ABARQ516

ABARQ517

ABARQ518

ABARQ519

ABARQ520

ABARQ521

ABARQ522

ABARQ523

ABARQ524

ABARQ525

ABARQ526

ABARQ527

ABARQ528

ABARQ529

ABARQ530

ABARQ531

ABARQ532

ABARQ533

ABARQ534

ABARQ535

ABARQ536

ABARQ537

ABARQ538

ABARQ539

ABARQ540

ABARQ541

ABARQ542

ABARQ543

ABARQ544

ABARQ545

ABARQ546

ABARQ547

ABARQ548

ABARQ549

ABARQ550

ABARQ551

ABARQ552

ABARQ553

ABARQ554

ABARQ555

ABARQ556

ABARQ557

ABARQ558

ABARQ559

ABARQ560

ABARQ561

ABARQ562

ABARQ563

ABARQ564

ABARQ565

ABARQ566

ABARQ567

ABARQ568

ABARQ569

ABARQ570

ABARQ571

ABARQ572

ABARQ573

ABARQ574

ABARQ575

ABARQ576

ABARQ577

ABARQ578

ABARQ579

ABARQ580

ABARQ581

ABARQ582

ABARQ583

ABARQ584

ABARQ585

ABARQ586

ABARQ587

ABARQ588

ABARQ589

ABARQ590

ABARQ591

ABARQ592

ABARQ593

ABARQ594

ABARQ595

ABARQ596

ABARQ597

ABARQ598

ABARQ599

ABARQ600

ABARQ601

ABARQ602

ABARQ603

ABARQ604

ABARQ605

ABARQ606

ABARQ607

ABARQ608

ABARQ609

ABARQ610

ABARQ611

ABARQ612

ABARQ613

ABARQ614

ABARQ615

ABARQ616

ABARQ617

ABARQ618

ABARQ619

ABARQ620

ABARQ621

ABARQ622

ABARQ623

ABARQ624

ABARQ625

ABARQ626

ABARQ627

ABARQ628

ABARQ629

ABARQ630

ABARQ631

ABARQ632

ABARQ633

ABARQ634

ABARQ635

ABARQ636

ABARQ637

ABARQ638

ABARQ639

ABARQ640

ABARQ641

ABARQ642

ABARQ643

ABARQ644

ABARQ645

ABARQ646

ABARQ647

ABARQ648

ABARQ649

ABARQ650

ABARQ651

ABARQ652

ABARQ653

ABARQ654

ABARQ655

ABARQ656

ABARQ657

ABARQ658

ABARQ659

ABARQ660

ABARQ661

ABARQ662

ABARQ663

ABARQ664

ABARQ665

ABARQ666

ABARQ667

ABARQ668

ABARQ669

ABARQ670

ABARQ671

ABARQ672

ABARQ673

ABARQ674

ABARQ675

ABARQ676

ABARQ677

ABARQ678

ABARQ679

ABARQ680

ABARQ681

ABARQ682

ABARQ683

ABARQ684

ABARQ685

ABARQ686

ABARQ687

ABARQ688

ABARQ689

ABARQ690

ABARQ691

ABARQ692

ABARQ693

ABARQ694

ABARQ695

ABARQ696

ABARQ697

ABARQ698

ABARQ699

ABARQ700

ABARQ701

ABARQ702

ABARQ703

ABARQ704

ABARQ705

ABARQ706

ABARQ707

ABARQ708

ABARQ709

ABARQ710

ABARQ711

ABARQ712

ABARQ713

ABARQ714

ABARQ715

ABARQ716

ABARQ717

ABARQ718

ABARQ719

ABARQ720

ABARQ721

ABARQ722

ABARQ723

ABARQ724

ABARQ725

ABARQ726

ABARQ727

ABARQ728

ABARQ729

ABARQ730

ABARQ731

ABARQ732

ABARQ733

ABARQ734

ABARQ735

ABARQ736

ABARQ737

ABARQ738

ABARQ739

ABARQ740

ABARQ741

ABARQ742

ABARQ743

ABARQ744

ABARQ745

ABARQ746

ABARQ747

ABARQ748

ABARQ749

ABARQ750

ABARQ751

ABARQ752

ABARQ753

ABARQ754

ABARQ755

ABARQ756

ABARQ757

ABARQ758

ABARQ759

ABARQ760

ABARQ761

ABARQ762

ABARQ763

ABARQ764

ABARQ765

ABARQ766

ABARQ767

ABARQ768

ABARQ769

ABARQ770

ABARQ771

ABARQ772

ABARQ773

ABARQ774

ABARQ775

ABARQ776

ABARQ777

ABARQ778

ABARQ779

ABARQ780

ABARQ781

ABARQ782

ABARQ783

ABARQ784

ABARQ785

ABARQ786

ABARQ787

ABARQ788

ABARQ789

ABARQ790

ABARQ791

ABARQ792

ABARQ793

ABARQ794

ABARQ795

ABARQ796

ABARQ797

ABARQ798

ABARQ799

ABARQ800

ABARQ801

ABARQ802

ABARQ803

ABARQ804

ABARQ805

ABARQ806

ABARQ807

ABARQ808

ABARQ809

ABARQ810

ABARQ811

ABARQ812

ABARQ813

ABARQ814

ABARQ815

ABARQ816

ABARQ817

ABARQ818

ABARQ819

ABARQ820

ABARQ821

ABARQ822

ABARQ823

ABARQ824

ABARQ825

ABARQ826

ABARQ827

ABARQ828

ABARQ829

ABARQ830

ABARQ831

ABARQ832

ABARQ833

ABARQ834

ABARQ835

ABARQ836

ABARQ837

ABARQ838

ABARQ839

ABARQ840

ABARQ841

ABARQ842

ABARQ843

ABARQ844

ABARQ845

ABARQ846

ABARQ847

ABARQ848

ABARQ849

ABARQ850

ABARQ851

ABARQ852

ABARQ853

ABARQ854

ABARQ855

ABARQ856

ABARQ857

ABARQ858

ABARQ859

ABARQ860

ABARQ861

ABARQ862

ABARQ863

ABARQ864

ABARQ865

ABARQ866

ABARQ867

ABARQ868

ABARQ869

ABARQ870

ABARQ871

ABARQ872

ABARQ873

ABARQ874

ABARQ875

ABARQ876

ABARQ877

ABARQ878

ABARQ879

ABARQ880

ABARQ881

ABARQ882

ABARQ883

ABARQ884

ABARQ885

ABARQ886

ABARQ887

ABARQ888

ABARQ889

ABARQ890

ABARQ891

ABARQ892

ABARQ893

ABARQ894

ABARQ895

ABARQ896

ABARQ897

ABARQ898

ABARQ899

ABARQ900

ABARQ901

ABARQ902

ABARQ903

ABARQ904

ABARQ905

ABARQ

4. To view the structures in “combination1” in the main ConQuest window, make sure combination1 is selected (it will be highlighted in red), then click the **View** button.

We are interested in looking for a structure with a phenol functional group that belongs in both the drug and teaching subset. *How can we conduct such search, limiting the search to the results from the combination in the previous step?*

5. In the main ConQuest window, click on the *Build Queries* tab, click on the **Draw** button, and sketch the motif in the figure to the right, which is an aromatic ring with an oxygen attached with a single bond to one of the carbons. To sketch the aromatic ring, you can use the pre-designed one on the left-hand side of the Draw window. Once you are done sketching, click **Search** to proceed to the *Search Setup* window.

5

4

Name	Hits	Type
CSD_Drug_subset	13791	Refcode List
combination1	107	Combination
teaching_subset	837	Refcode List

6. The *Search Setup* window is where we can select the subset to be used for the search. Click the **Select Subset** button to open the *Restrict Search* window. The subset we want to load is the “combination1” hitlist that was previously generated. Since the hitlist was generated this session, select *Entries in a hitlist loaded this session*. There should be a drop-down menu to select which of the hitlists generated this session you want to use. Click on the menu and select **combination1**. Once the hitlist is loaded, click **OK** to close the window.

6

Restrict Search by Refcode

Current Restrictions

No Restrictions Currently Set

Entries in a hitlist loaded this session

Entries in a pre-defined hitlist

Entries in saved search file

Entries in saved hitlist combination

Entries in saved refcode list

Restrict based on search this session

CSD_Drug_subset

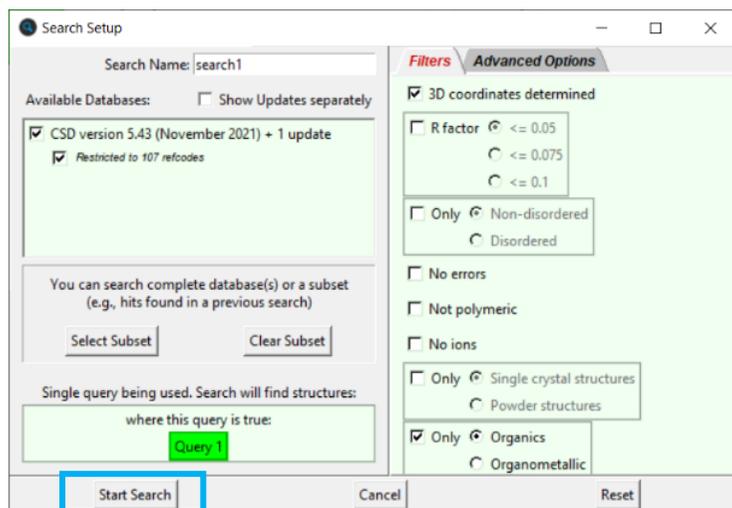
CSD_Drug_subset

teaching_subset

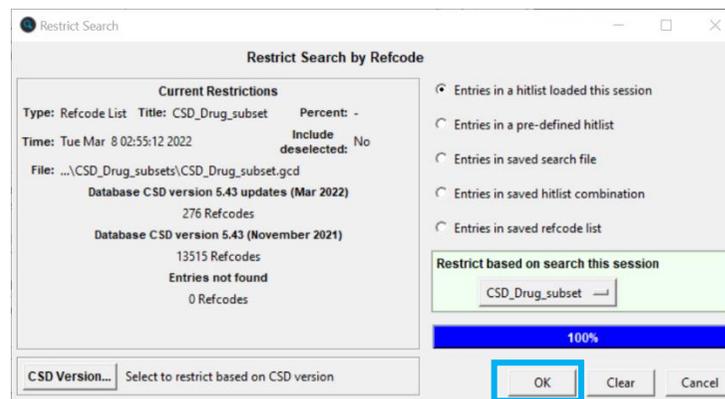
combination1

7. Now in the *Search Setup* dialogue box, note that under *Available Databases*, there is an additional line below the CSD version that indicates the search is restricted to 107 refcodes. Make sure both boxes are ticked. Be sure to also tick the boxes for “3D coordinates determined” and “Only Organics” to apply these filters. Then click **Start Search** to begin the search.
8. At least 23 structures should be returned. Shown in the diagram is a structure of Aspirin (ACSALA01).

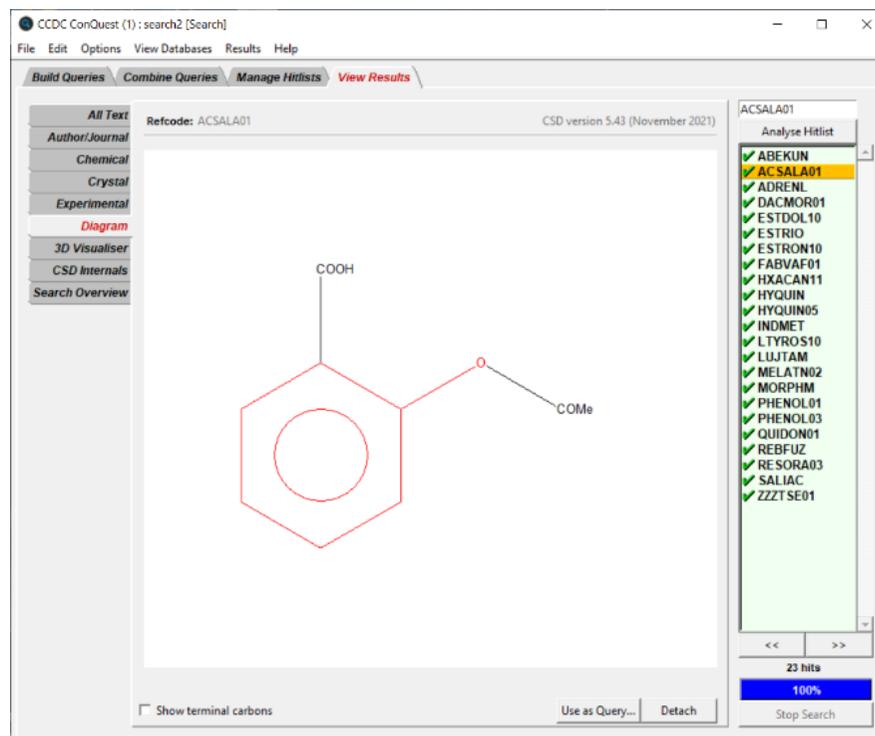
7



6



8



Glossary

Amide

An amide is a compound with the group RCONR'R''.

CSD COVID-19 subset

The subset contains CSD entries that have been reported as small molecule drug candidates for the treatment of COVID-19.

CSD Drug subset

Subsets that contain approved drug compounds as defined by the DrugBank database (<https://www.drugbank.ca/>). Further information can be found here: The CSD Drug Subset: The changing chemistry and crystallography of small molecule pharmaceuticals. Mathew J. Bryant, Simon N. Black, Helen Blade, Robert Docherty, Andrew G.P. Maloney, Stefan C. Taylor, J. Pharm. Sci., 108 1655-1662, 2019 [DOI: 10.1016/j.xphs.2018.12.011] CSD COVID-19 subset

Hitlist

A hitlist is a subset of CSD entries which can include search results, refcode lists, or the results of combining these.

Metal-Organic Complex

A chemical compound that contains a metal and organic ligand.

Metal-Organic Frameworks (MOFs)

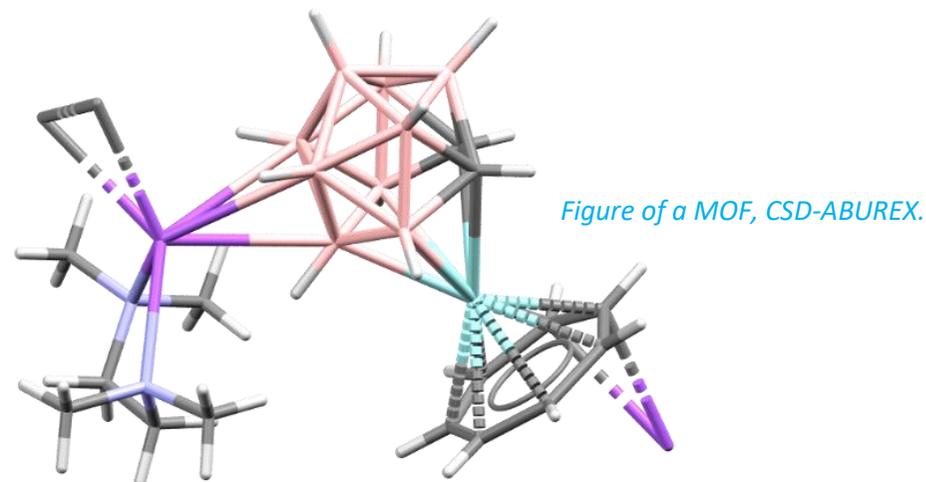
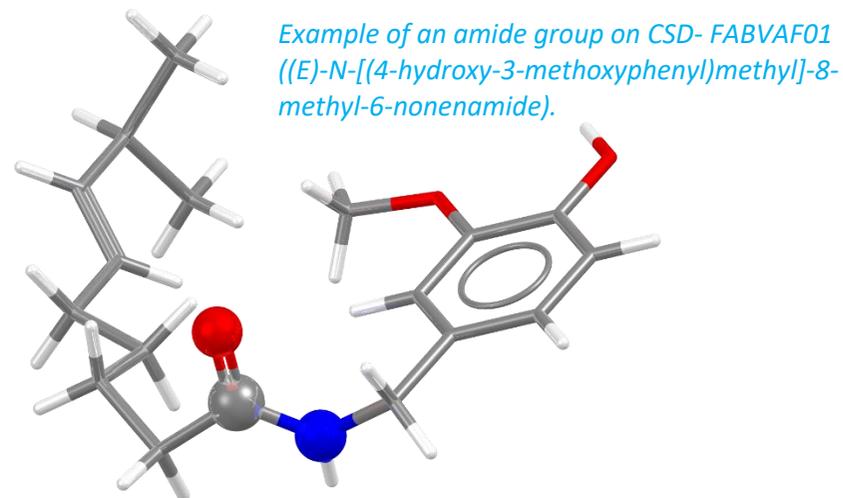
A MOF, or metal organic framework, is a material formed of metal clusters or nodes, linked by organic molecules.

Only Organics

Selecting this filter eliminates from the search any structure that contains a transition metal, lanthanide, actinide or any of Al, Ga, In, Tl, Ge, Sn, Pb, Sb, Bi, Po.

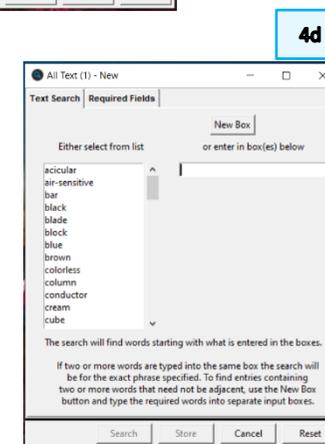
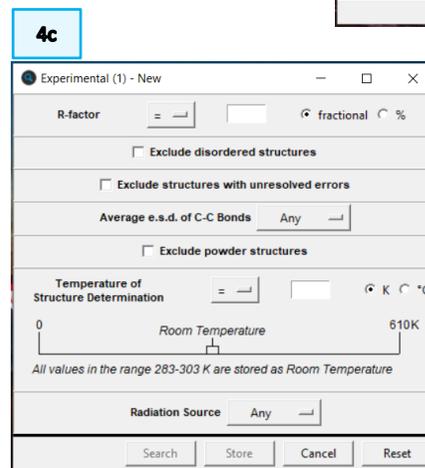
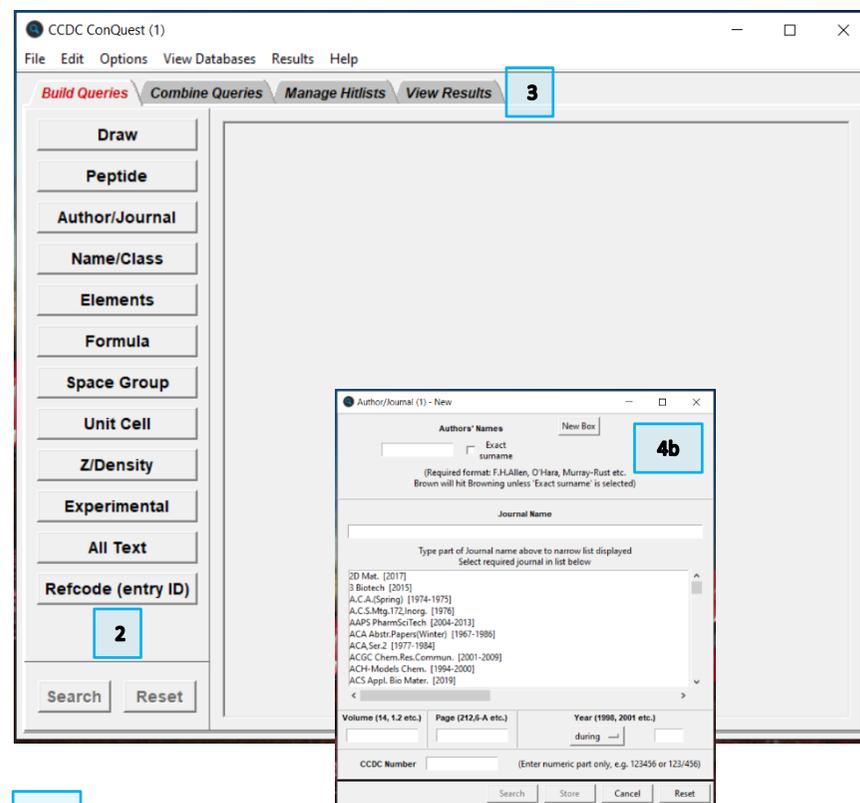
Teaching subset

The CSD Educational Collection of structures containing 700+ structures carefully selected to enhance chemistry learning is now available to view in ConQuest (having previously been available in Mercury)



Review. ConQuest Interface

1. Launch ConQuest by clicking the ConQuest Icon  on your desktop or launching it from the Start or Applications menu.
2. The ConQuest main window shows all the search routines you can perform on the left-hand side of the window.
3. The row of tabs across the top of the window will guide you through the steps of the search process.
4. Some example searches are
 - a. Draw – substructure and 3D information searching
 - b. Author/Journal – bibliographic searching
 - c. Experimental – experimental set up searching
 - d. All Text – generic text-based searching
5. The majority of the searching we will do in these tutorials will be substructure searching, so we will focus on the Draw tab here.



Review. Draw Window

All drawing takes place in the central white area of the *Draw* window. In addition to creating 2D chemical structure sketches, the *Draw* window allows for the inclusion of 3D parameters for searching or for filtering.

ConQuest sketching conventions

- Left click in the sketcher to insert the selected atom type
- Left click and drag to sketch two bonded atoms
- Use the **Edit** button to modify properties of or delete atoms, bonds or entire substructures
- Right-click on atoms or bonds to modify their properties
- Use the **Templates...** button to pick from a list of CSD editor devised and drawn substructures
- Use the **More...** button to find less frequently used element types, or generic atom type groups (e.g. halogens), or define custom element combinations (e.g. C or N or O).

