Searching and analysing molecular geometries using ConQuest and Mercury (CQ-006)

Developed using 2023.1 CSD Release CSD version 5.44 (April 2023)

Table of Contents	
Introduction	2
Learning Outcomes	2
Pre-required Skills	2
Materials	2
Urea conformation in the CSD	3
Data Analysis in Mercury	5
Conclusion	5
Exercises	6
Summary	6
Next Steps	6
Feedback	6
Glossary	7
Review. ConQuest Interface	8
Review. Draw Window	9
ConQuest sketching conventions	9





2

Introduction

ConQuest is the desktop search interface to the Cambridge Structural Database (CSD). All textual, 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. The data from structures in the CSD can be used to show the most likely values a particular bond, angle, *torsion* or ring would adopt. This information can be included in the ConQuest search parameters with further analysis carried out in Mercury.

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.

Learning Outcomes

After completing this workshop, you will be able to:

- Add 3D parameters in your ConQuest search.
- Use the data analysis function in Mercury.
- Explore conformation using ConQuest search results.

This workshop will take approximately **30** minutes to be completed. The words in <u>Blue Italic</u> in the text are reported in the <u>Glossary</u> at the end of this handout. A <u>review</u> of the ConQuest interface is also at the end of this handout.

Pre-required Skills

The following exercises assume you have a working knowledge of the programs ConQuest and Mercury, namely, how to conduct basic structural search in ConQuest and how to display and manipulate structures in Mercury. You can find the basics of visualization in Mercury as a <u>CSDU module</u> or as a self-guided workshop in the <u>CSD-Core workshop</u> area on our website. You can also find ConQuest workshops in the CSD-Core workshop area.

Materials

There are no additional materials required for this workshop.





Urea conformation in the CSD

There are over 1.2 million structures in the CSD providing a wealth of structural data. Structural chemists can use tools in the CSD portfolio to validate the threedimensional conformation of a particular molecule or to verify the most observed conformer for a molecule. There are a couple of ways to explore conformation in the CSD, including Mogul Geometry Check. However, in this workshop, we will explore how to define parameters of interest in a ConQuest search and analyse the results using the data analysis tool in Mercury. We will investigate the preferred *conformation* of a urea in the CSD.

- 1. Launch ConQuest by clicking the ConQuest Icon ${}^{igodoldsymbol{ extsf{O}}}$ on your desktop or launching it from the Start or Applications menu Open the sketcher by clicking the Draw button.
- 2. Sketch the urea molecule, if you need a refresher on how to sketch, a review of the ConQuest interface is at the end of this handout.
- We want to omit rings in the dihedral angle of interest to limit our results to ureas. Right-click on the C-N bond and click on **Cyclicity** and select Acyclic.
- Now we define the dihedral angle of interest. Click on ADD 3D to launch the 4. Geometric Parameters window. Select the H-N-N-H angle in order and click **Define** next to Torsion. Since signs of individual torsion angles can be ignored, we can use the Apply Function option to take the absolute value of the angles.





N,N"-propane-1,3-diylbis(3-(3-pyridinyl)urea) molecule, refcode AGOLOY with the urea substructure highlighted





4



Help

2

- 5. In the Geometric Parameters window, click on **Options** under TOR1 to launch the 3D Limits and Options window. In the *APPLY FUNCTION* drop down menu, select **Absolute** and click **OK** and click **Done** in the Geometric Parameters windows. You should be returned to the Draw window with the abs(TOR1) parameter defined. We are now ready to conduct our search. Click Search to launch the Search Setup window.
- 6. We want to search the entire database for structures that match our sketched substructure. In the *Search Setup* window, make sure to tick the box for "3D coordinates determined". Click **Start Search** to begin the search.
- 7. There should be at least 2886.¹ structures returned. Note that the value for the torsion angle parameter you defined is shown in the upper right corner of the *View Results* tab. You can use the arrow keys to scroll through the refcodes to investigate the hits returned by this query.
- 8. We can use Mercury's **Data Analysis** module to further analyse the data. To launch this, click the **Analyse Hitlist** button and choose *Analyse Data* from the drop-down menu.
- 9. If you are interested in analysing other parameters, you can tick them off in the dialogue box that appears. We will skip this step for now. Simply click **Analyse in Mercury.**
- 10. This will launch the Mercury app with the refcode list loaded and the **Data Analysis** window. The **Data Analysis** window is often hidden behind the main Mercury app, so you may have to minimise Mercury to see it.

AGOLOY		
Analyse Hitlist		
ABIBEU	Visualise Structures	
✓ ACURUN	Analyse Data	
ACUSAU		
✓ ADEZEP	Select All	
ADIFAV	Deselect All	
✓ AFIPIQ	Invert Selection	
AFIPIQ01		

¹ As of version 2023.1 of the CSD.

	Colored also iteration with the i	- shade for a the shades hade			
	Select the items you wish to i	nclude from the choices belo	N		-
File type: Mercury data fi	ile (.c2m) 🔟			9	
✓ Include Defined Para	meters				
Crystal data					
R-factor	🔲 Space Gp. Symbol	Space Gp. Number	No. of Coordinates		ates
Z Value	Z Prime	Study Temp.	Calc. Density		
Cell data					
🗆 a	□ b	□ c			
🗌 Alpha	🗖 Beta	🗖 Gamma	Cell Volume		
Reduced Cell a	Reduced Cell b	Reduced Cell c	Reduced Cell Vo		/olume
🗖 Reduced Cell Alpha	Reduced Cell Beta	Reduced Cell Gamma			
Other					
Publication Year	Unique Chemical Units	Multiplier Sum	Compound I		me



GOLOY (P21/c) - Mercury	😯 Data Analysis					- 🗆 ×
File Edit Selection Display Calculate CSD-Community CSD-Core CSD-Materials CSD-Theory CSD-F	article File Options					
Project Made: Pick Atoms V Clear Measurements (A) O C Show Labels for All atoms	search1 Spreadsheet 1		8 >			
Style: Ball and Stick v Colour: by Element v Manage Styles Publication v Atom selections:	File Tools Descriptors	File Tools Descriptors Display Selection Plots Statistics				8 0
□ Animate Default view: b ∨ a b c a* b* c* x x+ y- y+ z- z+ x+90 x+90 y-90	y+90 Find identifier	Find identifier				
	Identifier	NAME	Owery	Fragment	abs(TORI)	
	search1(ABIBEUI0	ABIBEU	1	1	2,6390	
	search1IACURUNI1	ACURUN	1	· · · · ·	0.6990	
	search1(ACUSAU(2	ACUSAU	1	1	3.7130	
	search1[ADEZEP[3	ADEZEP	1	1	173,2070	
	search1(ADEZEP)4	ADEZEP	1	2	176.1110	
	search1[ADIFAV[5	ADIFAV	1	1	169.0340	
	search1(AFIPIQ(6	AFIPIQ	1	1	177.2730	
	search1 AFIPIQ01 7	AFIPIQ01	1	1	166.0760	
	search1 AFOKIS 8	AFORIS	1	1	0.0880	
	search1 AFOKIS 9	AFOKIS	1	2	7.3640	
	search1[AFOKOY]10	AFOROY	1	1	4.0360	
	search1 AGOKIR 11	AGOKIR	1	1	158.4010	
	search1 AGOKIR01 12	AGOKIR01	1	1	177.9410	
	search1 AGOKOX 13	AGOKOX	1	1	9.6190	
	search1[AGOKUD]14	AGOKUD	1	1	0.7420	
	search1 AGOLAK 15	AGOLAK	1	1	7.0970	
	search1[AGOLI5]16	AGOUS	1	1	0.0170	
10	search1 AGOLIS01 17	AGOLIS01	1	1	27.2880	
	search1[AGOLOY]18	AGOLOY	1	1	6.9570	
	search1 AGOLOY 19	AGOLOY	1	2	4.8260	
	search1 AGOLUE 20	AGOLUE	1	1	4.7950	
	search1[AGOEUE]21	AGOLUE	1	2	16.3390	
	search1[AGOMAL]22	AGOMAL	1	1	11.0120	
	search1[AGOMAL]23	AGOMAL	1	2	16.3320	
	search (JAGOMEP)24	AGOMEP			3.9220	
	search (pAGOMEP)25	AGOMEP		2	2.0190	
	scorch1/AGOMIT[20	AUABAE			3.9000	
Display Options	search1ph/ARAE129	AMADAE			23.6/60	
- Display	scorch1/AHARAE/28	AMARAE		2	0.2760	
Pacing Floor Content - Compatibility	search1AHARAE120	AHARAF		4	12,8630	
Asymmetric linit I i i Read I have defended	search1/A (APEII31	ATARE			9 2840	
Ante cantra	search1(A)U(OEB2	AJUIOF			0.4590	
	search1IAJUIOFI33	AJUIOF	- i	2	6.3020	
Reset	search1(A)UUU134	AJUIUL	1	ĩ	1.4590	
	search1[AKELIM[35	AKELIM	1		172.0430	~
Press the left mouse button and move the mouse to rotate the structure						

Data Analysis in Mercury.

Now that you have performed the search and obtained the torsion angle parameter, we can use the data analysis module in Mercury to visualize the values.

- 1. Look at the **Data Analysis** window. This shows all the data from your defined geometric parameter, sorted by refcode.
- 2. To view the distribution of torsion angles in this hitlist, click the **abs(TOR1)** header in the column to highlight the column, then click *Plots* from the top menu and then select *Histogram*.
- 3. Most of the urea containing structures in the CSD adopt a cis conformer. Click on the bar of the histogram around 0 degrees to highlight the corresponding entries in the spreadsheet.
- 4. It will be easier to scroll through the refcodes in that region by sorting the spreadsheet by abs(TOR1). Right-click on the **abs(TOR1)** heading and click *Sort*.
- 5. You should now see the highlighted refcodes in the spreadsheet that correspond to the angles clustered around 0 degrees. Clicking on each entry in the spreadsheet will place the corresponding structure in the Mercury display.
- As you scroll through the highlighted refcodes, can you observe a trend in the structures clustered here? Review some of the structures in the 150 – 180degree cluster (trans conformers).

Conclusion

From the search of the CSD and plotting the results, we can see that the most observed conformer for urea in the CSD is the cis conformer.



Exercises

In the example above, we conducted a search of the entire database. How could the distribution of our histogram change if we defined a narrower search?

- Return to the ConQuest, using Query 1, include all the filters in the Search Setup window.
- Use the data analysis module in Mercury to create a histogram on your torsion angles. How does this distribution differ from our initial search?

Summary

This workshop introduced how to use ConQuest searches and data analysis in Mercury to review conformation of a specific molecule.

In the example you have seen how to add 3D parameters to your query to mine the CSD for even more structural information. The data analysis window allows you to plot your data in a variety of ways, determine statistics on your datasets, and to perform calculations on the values of the parameters returned by the search.

You should now be familiar with:

- Adding 3D parameters in your search.
- Conducting data analysis in Mercury.

Next Steps

After this workshop, you can explore more exercises in the self-guided workshops available in the <u>CSD-Core</u> (ConQuest) and <u>CSD-Materials</u> (Mercury) workshops area on our website.

https://www.ccdc.cam.ac.uk/community/training-and-learning/workshopmaterials/ . Please see the <u>ConQuest documentation</u> for more details.

Feedback

We hope this workshop improved your understanding of ConQuest and you found it useful for your work. As we aim to continuously improve our training materials, we would love to get your feedback. Click on <u>this link</u> to a survey, it will take less than **5** minutes to complete. The feedback is anonymous. You will be asked to insert the workshop code, which for this self-guided workshop is CQ-006. Thank you!





Glossary

Conformation

The shape, or more precisely, the spatial arrangement of a molecule is called conformation. The same molecule can present different conformations (different shapes).

Torsion Angles

Torsion angles are used to describe conformations around rotatable bonds. The torsion angle between 4 atoms A-B-C-D is the angle by which the vector A-B must be rotated in order to eclipse the vector C-D when views along the vector B-C.

N-t-Butylaminocarbonyl-N-2chlorophenoxyacetylthiourea structure (refcode LIDCEH) with the urea substructure in a trans conformation highlighted

7

(acetyl)-(1-(2-(1-(pyridin-2yl)ethylideneamino)ethyl)-3-ethylurea)palladium(ii) tetrafluoroborate structure (refcode, DEDDUM) with the urea substructure in a cis conformation highlighted



8

Review. ConQuest Interface

- 1. Launch ConQuest by clicking the ConQuest Icon (S) 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.

CCDC ConQuest (1)				-		×
File Edit Options View Databases Results Build Queries Combine Queries Manag	Help ge Hitlists View	Results 3				
Draw						-1
Peptide						
Author/Journal						
Name/Class						
Elements						
Formula						
Space Group	Author/Journal (1)	lenn .		7		
Unit Cell	• Hunoyoumur (r)	Authors' Names	New Box			
Z/Density	(Re Brown	surname quired format: F.H.Allen, O'Ha will hit Browning unless 'Exac	ra, Murray-Rust etc. t surname' is selected)			
Experimental	Journal Kame					
All Text	Type part of Journal name above to narrow list displayed Select required journal in list below					
Refcode (entry ID) 2 Search Reset	Binete (2015) A.C.A.(Sping) [1974-1] A.C.S.Mg.Zinerg, [1974-1] A.C.S.Mg.Zinerg, [1 A.C.A.Spitz Zinerg, [1 A.C.A.Spitz Zinerg, [1 A.C.A.Spitz Jipor, Jipol A.C.G.C.Den-Res.Comm A.C.H.Models Chem. [] A.C.S.Appl. Bio Mater. [C Volume (14, 1.2 etc.)	75] 976] 064-2013] 11(967-1986] 14(2001-2009] 954-2000] 919] 14ge (212,6-A etc.)	Vear (1998, 2001 etc.) during			
	CCDC Number	(Enter nu Search	meric part only, e.g. 123456 or 123/456) Store Cancel Reset	1		
4c				_		4d
Experimental (1) - New -			All Text (1) - New		_	
R-factor = 🕫 fi	actional C %		Text Search Required Fields			
Exclude disordered structures			Either select from list	ore	nter in box(es)	below
Exclude structures with unresolved en	rors		acicular air-sensitive bar	I		
Average e.s.d. of C-C Bonds Any	-		black blade block			
Temperature of =	⊙ко∘с		blue brown colorless			
0 Room Temperature	610K		conductor cream cube			
All values in the range 283-303 K are stored as Room	Temperature		The search will find words sta	rting with what	it is entered in	the boxes.
Radiation Source Any			two or more words that n button and type the requ	specified. To fi eed not be adj ired words int	nd entries con acent, use the o separate inp	taining New Box out boxes.
Search Store Can	el Reset		Search	Store	Cancel	Reset

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).

