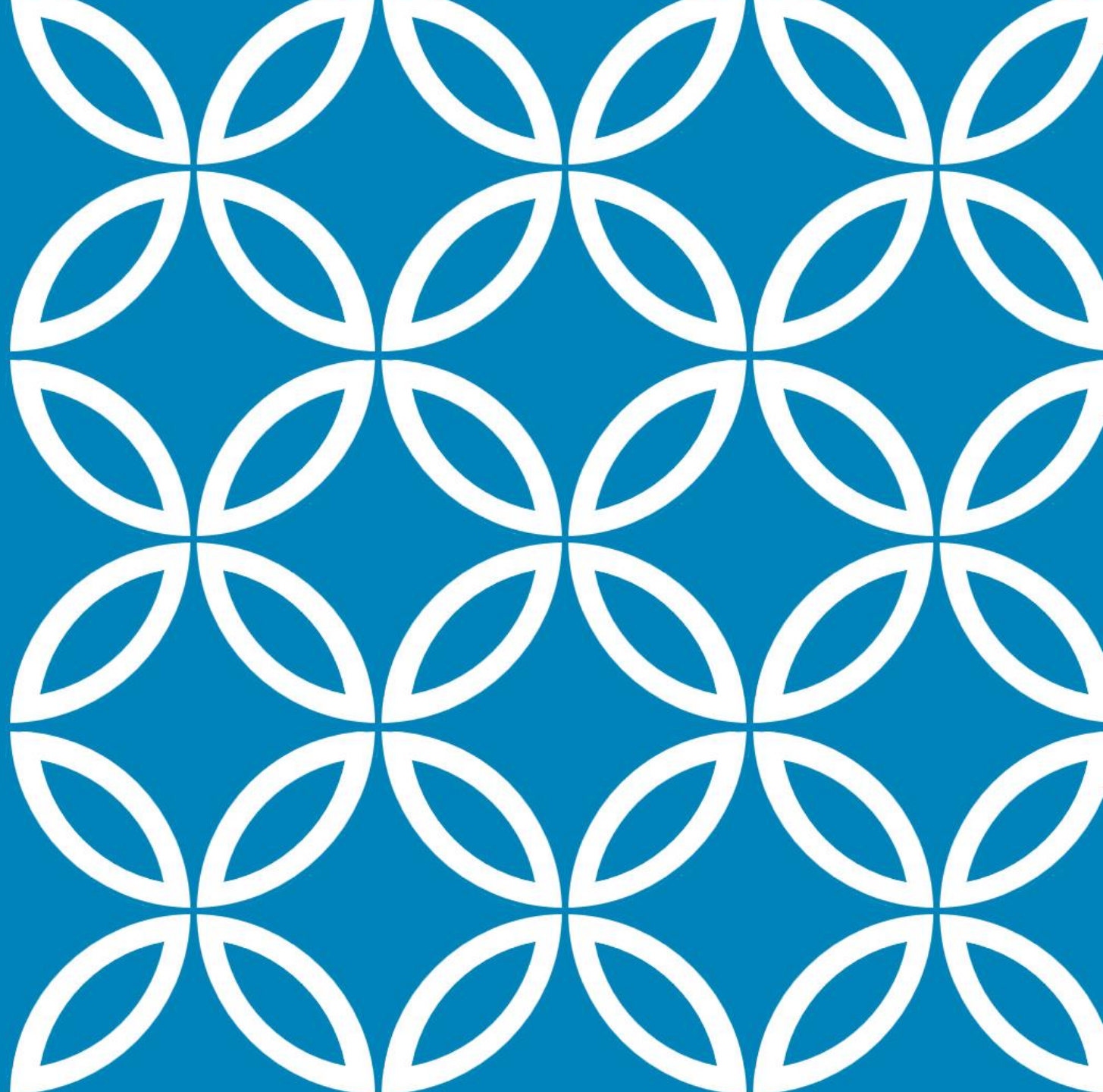


CURRENT PRACTICES IN (HOME LAB) RAW DATA ARCHIVAL

Responses to a survey published
at:

<https://www.iucr.org/news/newsletter/volume-28/number-1/raw-data-availability-the-small-molecule-crystallography-perspective>

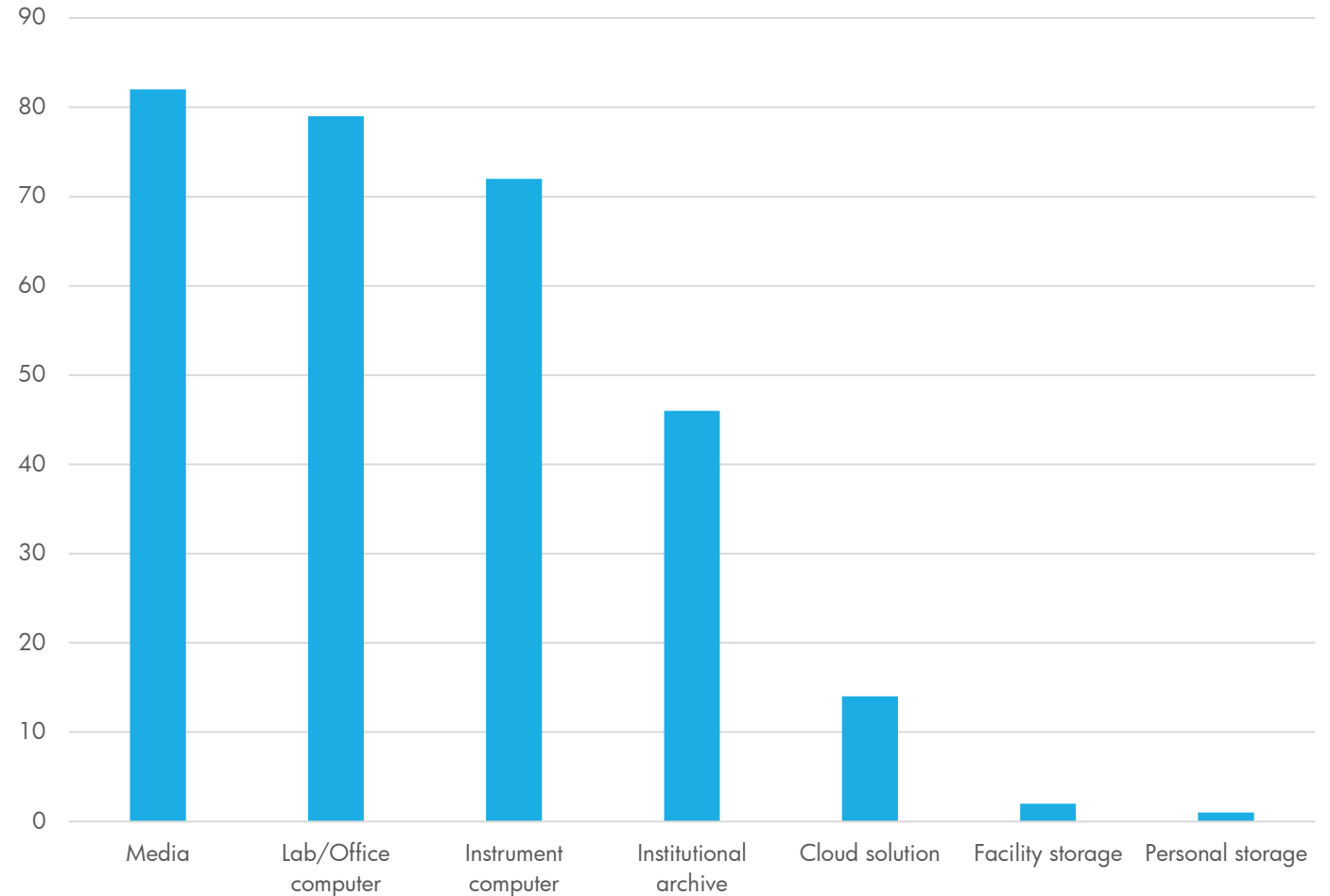


A SNAPSHOT OF CURRENT ARCHIVAL PRACTICE



Mainly 'Do It Yourself' with resources readily to hand in the lab

Around 20% use more 'modern' (ie networked) approaches

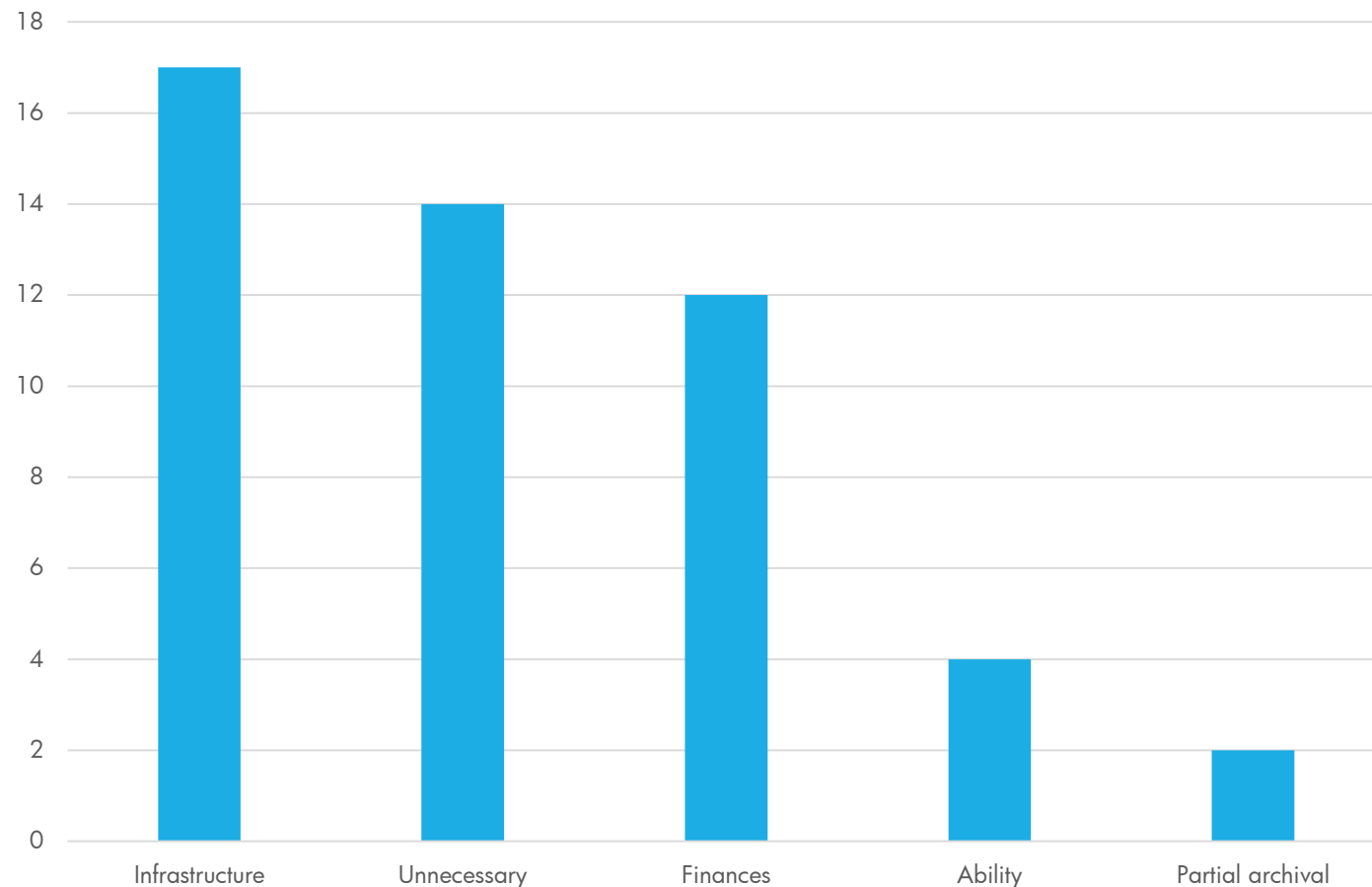




WHY NOT ARCHIVE RAW DATA?

38/193 people DON'T archive raw data

Half of these probably would if there was support available



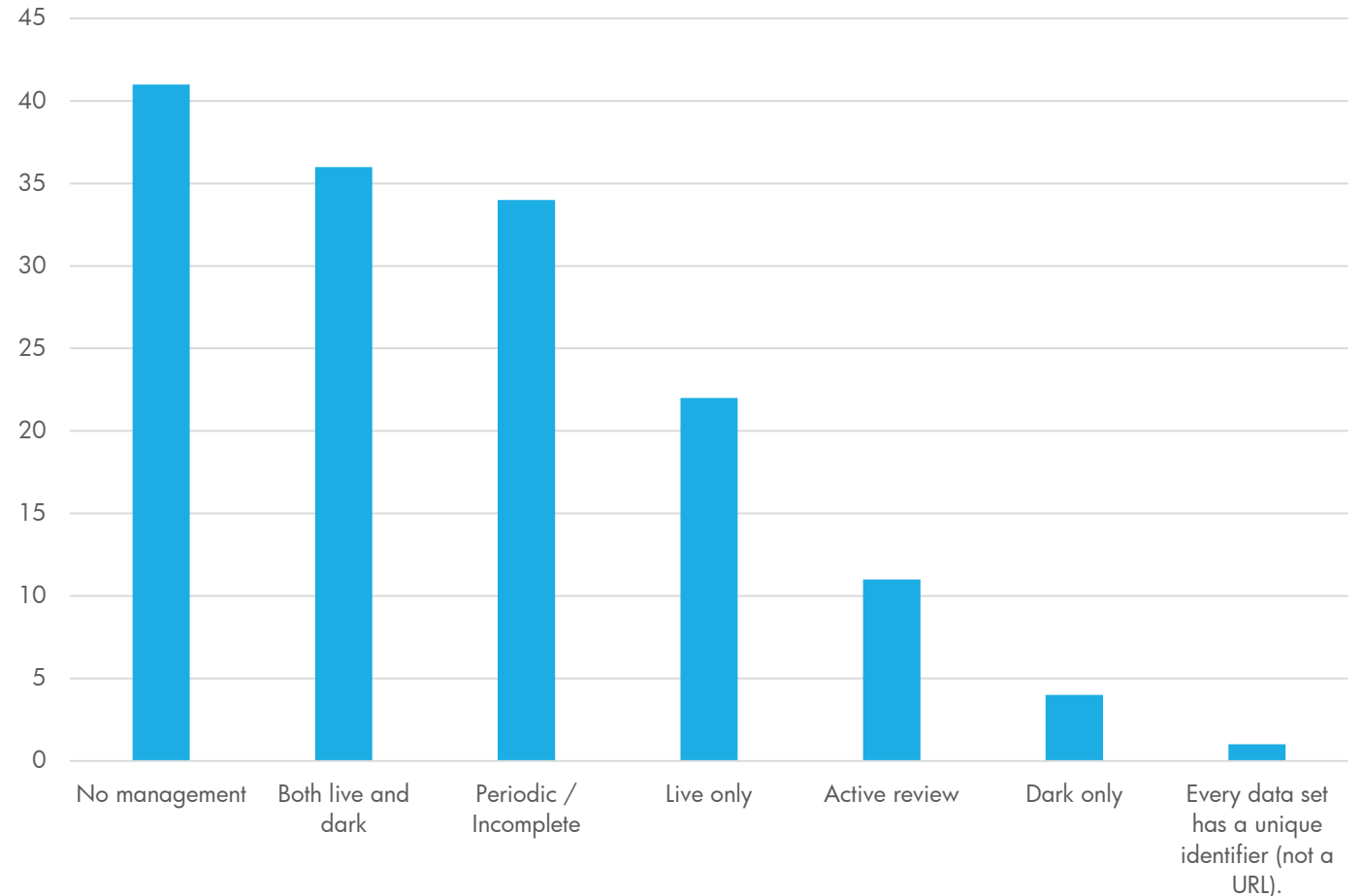
ARE ARCHIVES 'MANAGED'?



Largest response is no

Majority perform some sort of activity

Very little active review over time



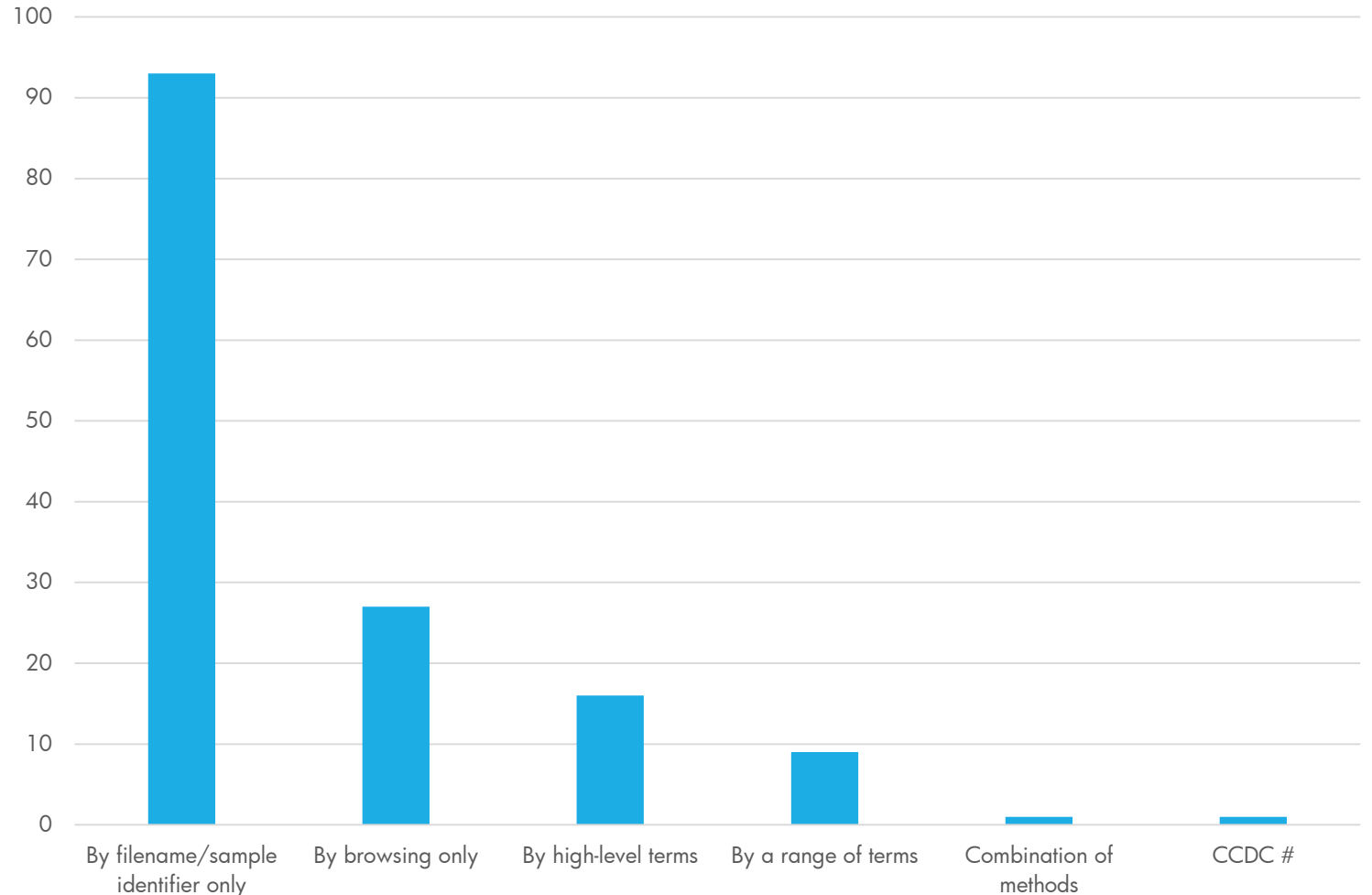
CAN YOU SEARCH YOUR ARCHIVE?



Very few 'advanced' approaches

Mostly tools provided already by operating system

Requires prior knowledge to search

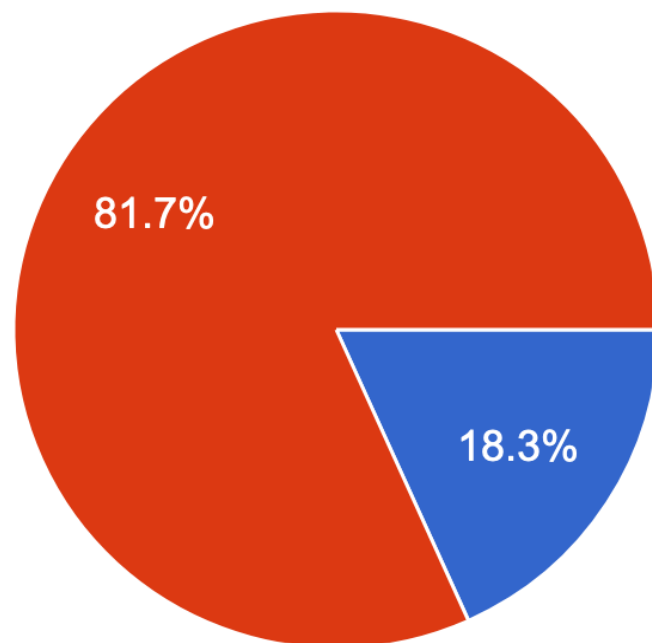




EXTENT OF 'OPENNESS'

Is your raw data externally available?

186 responses



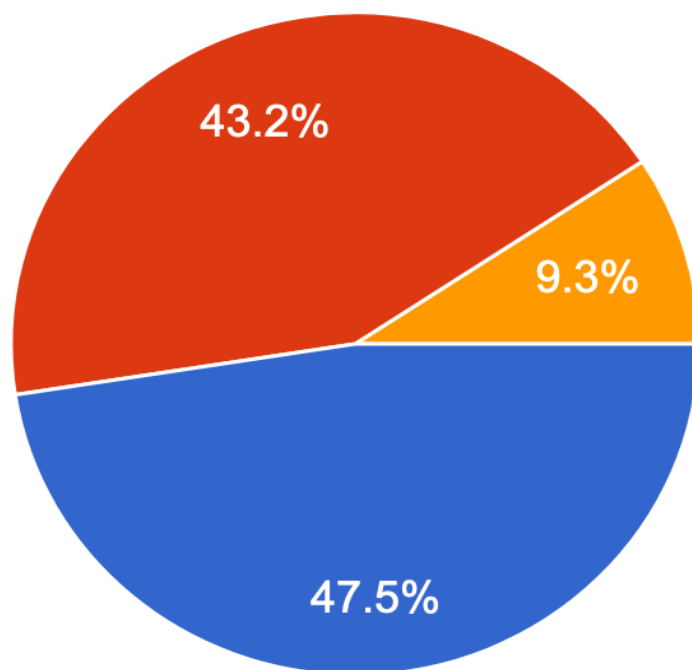
- Yes (e.g. open access; collaborators; other researchers requesting access)
- No



WHAT IF...?

If it were to become Policy to make funded research raw diffraction data accessible after 3 years since measurement, would you endeavour to comply?

183 responses



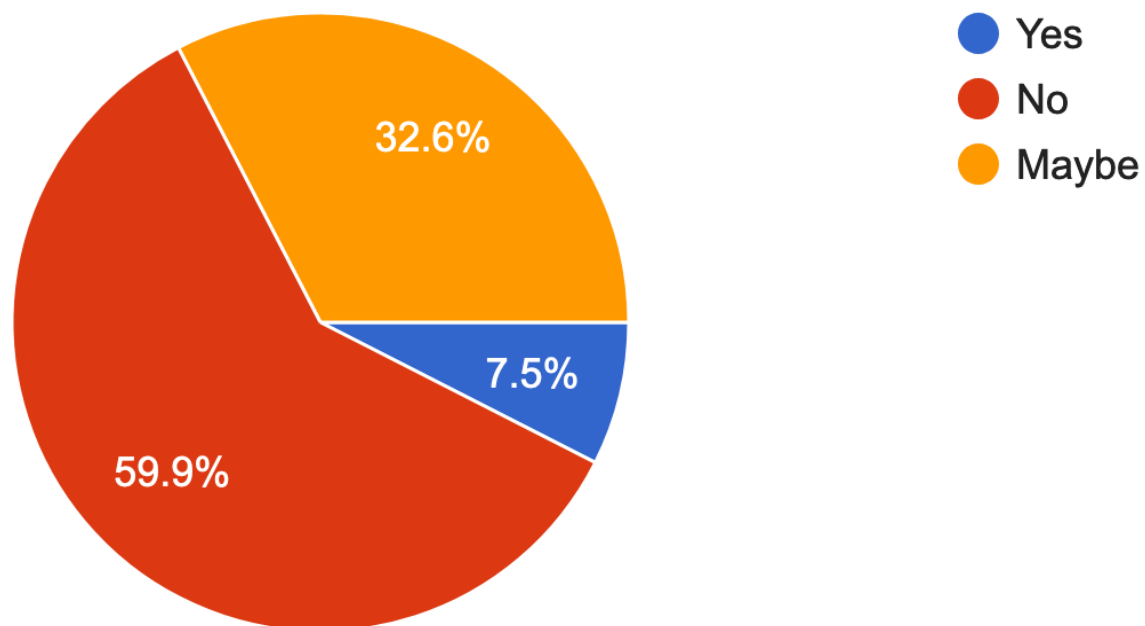
- Yes, I would do this for all my raw data
- Yes, but I would only do so in cases where I had to
- No



AFFORDABILITY

Would you anticipate being able to pay for external archive/repository services and facilities e.g. through research grants or institutional funding

187 responses

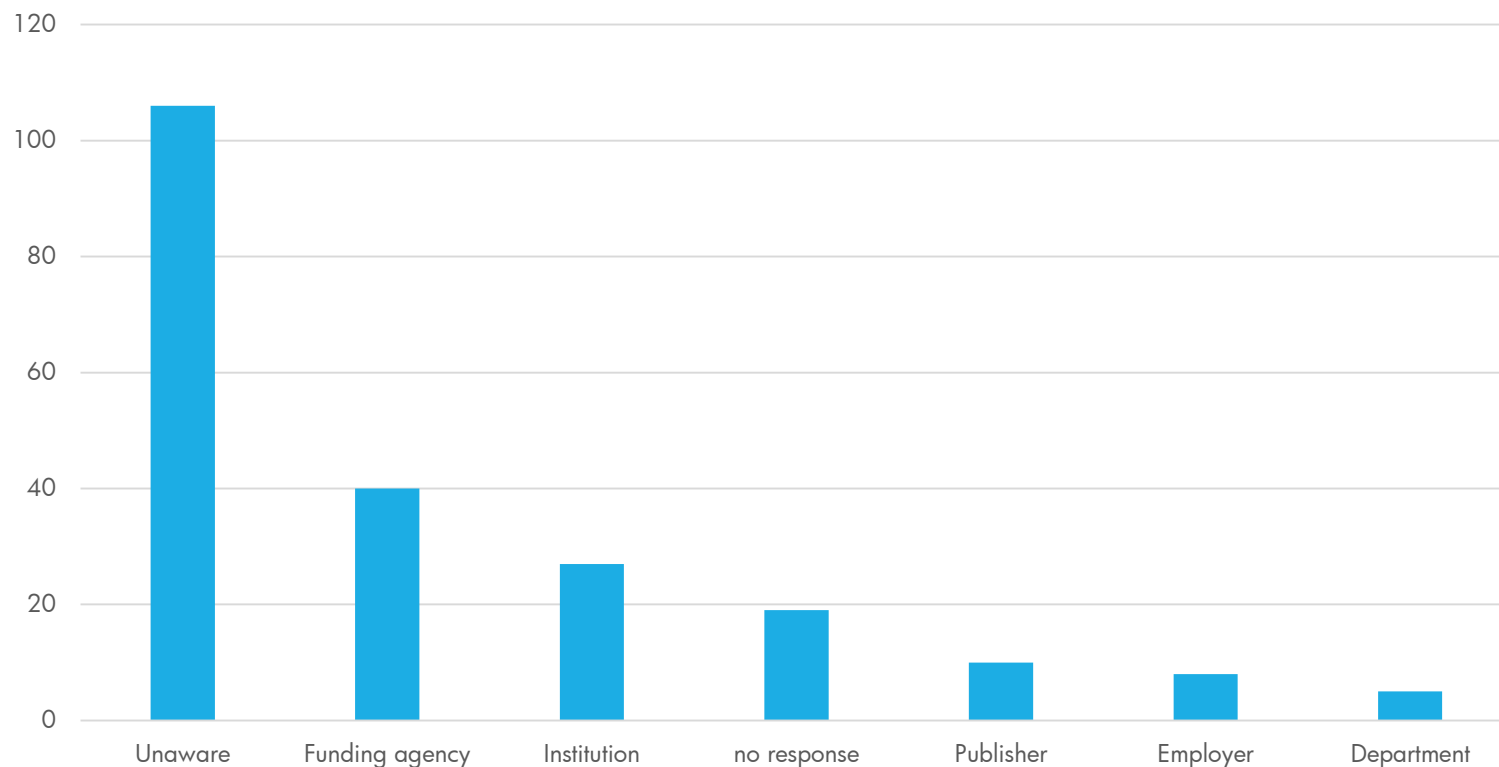




POLICY AWARENESS

Over half unaware of any governing policy

Funding agencies are the most prevalent, with institutions also featuring



TEST CASE...



nature

Explore content | Journal information | Publish with us

nature > articles > article

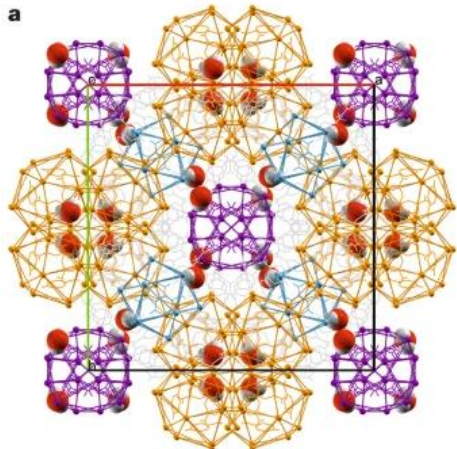
Article | Published: 10 February 2021

Complex structures arising from the self-assembly of a simple organic salt

Riccardo Montis, Luca Fusaro, Andrea Falqui, Michael B. Hursthouse, Nikolay Tumanov, Simon J. Coles, Terry L. Threlfall, Peter N. Horton, Rachid Sougrat, Anaïs Lafontaine, Gérard Coquerel & A. David Rae

Nature 590, 275–278 (2021) | Cite this article

7042 Accesses | 2 Citations | 61 Altmetric | Metrics



Data availability

Full crystallographic details in CIF format have been deposited in the Cambridge Crystallographic Data Centre database (deposition numbers: phase 1, CCDC 1540139 (unconstrained) and 1540140 (constrained); phase 2, CCDC 1897427; phase 3, CCDC 1540141; phase 4, 1897428). Copies of this information may be obtained free of charge from deposit@ccdc.cam.ac.uk or <http://www.ccdc.cam.ac.uk>. Raw single-crystal diffraction data corresponding to the structures of phases 1–4 have been deposited in the Zenodo repository at the following locations: phase 1, <https://doi.org/10.5281/zenodo.2595089>; phase 2, <https://doi.org/10.5281/zenodo.2585776>; phase 3, <https://doi.org/10.5281/zenodo.2593670>; phase 4, <https://doi.org/10.5281/zenodo.2593677>. The CCD images (in either Rigaku IMG or Bruker KCD format) have been deposited, along with instrument parameters and all files associated with image processing. This will enable the reader to fully validate these structural models and, for those who wish to investigate alternative approaches to modelling these results or develop them further, it will be possible to do so without having to synthesize the materials and collect diffraction data. NMR results are

zenodo

Search



Upload Communities

Log in Sign up

March 15, 2019

Dataset Open Access

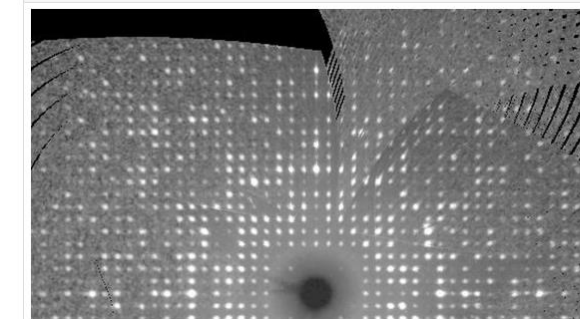
Single crystal diffraction raw data for Fampridine hydrochloride Phase 1

Coles, Simon; Montis, Riccardo; Horton, Peter; Hursthouse, Michael

A set of diffraction images (in both Denzo .x and Bruker Nonius .kcd formats) collected on a Bruker Nonius KappaCCD diffractometer on an FR591 rotating anode generator equipped with confocal mirrors.

The sample is Phase 1 of the Fampridine hydrochloride organic salt - a system that forms numerous complicated phases. The authors wish to make the raw data available so that those who wish to explore the modelling of these exceptionally complex systems may process the data themselves.

Preview



Files (384.3 MB)

Name	Size	Preview	Download
0kl.jpg	55.7 kB		
md5:afe35e5223ba22723ed25b18a1be19			
0kl.syn	609.1 kB		
md5:c50261f74f2715170a01b4156108249			
apex2-0625-00315-625.zip	1.4 MB		
md5:f2c674b08525f33361fc9f05a142b622			

151

views

302

downloads

See more details...

Indexed in

OpenAIRE

Publication date:

March 15, 2019

DOI:

[DOI: 10.5281/zenodo.2595089](https://doi.org/10.5281/zenodo.2595089)

Communities:

Chemical Crystallography

License (for files):

[Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/)

Versions

Version 1

[10.5281/zenodo.2595089](https://doi.org/10.5281/zenodo.2595089)

Mar 15, 2019

Cite all versions? You can cite all versions by using the DOI [10.5281/zenodo.2595088](https://doi.org/10.5281/zenodo.2595088). This DOI represents all versions, and will always resolve to the latest one. Read more.

Share



Cite as

Coles, Simon, Montis, Riccardo, Horton, Peter, & Hursthouse, Michael. (2019). Single crystal diffraction raw data for Fampridine hydrochloride Phase 1 [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.2595089>



CONCLUSIONS?

This behaviour reflects general data management practice

'Archival' can imply keeping data safely out of sight, but some current data management practice could easily be converted to more 'open'

Only a modest amount of culture change required

'Lab finances' are not a significantly influencing factor for local archive, but are for publishing

Some infrastructure and guidance development is necessary...

- Index/Search system (not reliant on operating system/single lab ID)
- Persistent IDs
- Standard archive approach (metadata + packaging format)
- Mechanism to clearly link raw data with 'results' data
- System to establish if an image is 'clean'
- System to establish if all Bragg diffraction in raw data accounted for in the model
- Ability to have control over when/if to publish

A start: IUCrData to publish "Raw" Data Letters...

TODAY'S AGENDA



Day 2: Thursday 12 August		
Session 2: What would a 'standard' look like? – best practice		
14:00-14:10 (08:00-08:10)	Simon Coles (Southampton)	Current practices in raw data archival (home laboratories)
14:10-14:35 (08:10-08:35)	Graeme Winter (Diamond)	Future outlook for a curated raw data archive
14:35-15:00 (08:35-09:00)	Loes Kroon-Batenburg (Utrecht)	How to make structures better in the future and the use of published raw data in crystallographic software development
15:00-15:20 (09:00-09:20)	Coffee	
15:20-15:45 (09:20-09:45)	Teodor Ivănoaica (Extreme Light Infrastructure)	Data Policy at Large Research Infrastructure. Data Retention challenge
15:45-16:10 (09:45-10:10)	Natalie Johnson (CCDC)	The value of integration of raw data with publishing procedures and results databases
16:35 (10:35)	Panel discussion	16:35 (10:35)
17:00 (11:00)	Close of Day 2	