Publishing biomacromolecule structures: requirements and opportunities

Rita Berisio
Istituto di Biostrutture e Bioimmagini
Consiglio Nazionale delle Ricerche
Via Mezzocannone 16, Napoli
E-mail: luciana.esposito@cnr.it

Naples, Italy
29 Aug – 3 Sep 2019
Outline

Part I. I have great results. Where to publish? How?

Part II. IUCR journals. Requirements and tools

Part III. Let’s write a paper together
Part I.
You need to publish publish publish publish: Why? Where? How?

- Why publish?
- Submission
- Editorial
- Production
- Publication
Why should I publish? Can’t I keep results for myself??

Fame
Recognition by your peers

Fortune
Promotions, grant applications, research funding

Responsibility - duty
To society, taxpayer-funded research, contribution to progress
submitting an article

- Submission
- Editorial decision
- Production
- Publication
Where to publish my data?? What type of article??

…it depends on what I am trying to communicate

• Research article
• Short communication
• Letter to the editor
• Perspective
• Review/Mini review article

Do my target journals publish the type of article that I want to write?
Evaluating the target journal

Criteria you may use...

- Prestige
- Speed
- Audience
- Aesthetics
- Author service / experience
- Cost
- Likelihood of acceptance
First step: your front end is the Editor….what is he looking for??

- Title
- Abstract
  - The only pieces of your manuscript reviewers will look at before they decide to review a paper
- Cover letter
- Figures
- Presentation of the manuscript
Prior to the importance of your data... are they fair??
The first role of an Editor...

- Duplicate/Redundant publication
- Disputed Authorship
- Fabricated data
- Falsification
- Plagiarism
- Figure Manipulation
- Conflict of Interest
Once you have identified your journal...

STEP I: Preparing and submitting your manuscript

1) Read the author instructions and format

2) Take particular care of sections.

Some are more important than others...

Typical structure of a research article

- Introduction
- Method
- Results and Discussion
- Conclusion

...BUT...
The abstract – one of the most important elements of your article

- Referee and editor assessment
- Abstract and indexing / search-ability
What makes a good abstract?

In early 2007 it was determined that the compound melamine, suspected of having been involved in the deaths of numerous pets, had been fed to hogs intended for human consumption. This report describes a method for the analysis of melamine in porcine muscle tissue using solid-phase extraction (SPE) and high-performance liquid chromatography/tandem mass spectrometry (HPLC/MS/MS). Melamine was extracted in 50% acetonitrile in water. Homogenates were centrifuged and supernatants were acidified and washed with methylene chloride. The aqueous extracts were cleaned up using mixed-mode C8/strong cation exchange SPE and then concentrated, fortified with a stable isotope-labeled analog of melamine, and analyzed by HPLC/MS/MS. Gradient HPLC separation was performed using an ether-linked phenyl column with ammonium acetate/acetic acid and acetonitrile as the mobile phase. Multiple reaction monitoring (MRM) mode of two precursor-product ion transitions for melamine and one for the internal standard was used. A five point calibration curve ranging from 50 to 2000 ng/mL of melamine in solvent was used to establish instrument response. The method was validated by analysis of seven replicate porcine muscle tissue samples fortified with 10 ng/g of melamine. The mean recovery for the seven replicates was 83% with 6.5% relative standard deviation and the calculated method detection limit was 1.7 ng/g. Copyright © 2007 John Wiley & Sons, Ltd.
What makes a bad abstract?

The chlorogenic acids of Gardeniae Fructus used traditionally as a Chinese herbal medicine (zhizi) have been investigated qualitatively by liquid chromatography/multi-stage mass spectrometry (LC/MS*). Twenty-nine chlorogenic acids were detected and twenty-five characterised to regioisomer level on the basis of their fragmentation, twenty-four for the first time from this source. Assignment to the level of individual regioisomers was possible for three caffeoylquinic acids, three dicaffeoylquinic acids, three sinapoylquinic acids, four caffeoyl-sinapoylquinic acids, two feruloylsinapoylquinic acids, one p-coumaroyl-sinapoylquinic acid, three (3-hydroxy, 3-methyl)glutaroylquinic acids, two (3-hydroxy, 3-methyl)glutaroyl-feruloylquinic acids, one (3-hydroxy, 3-methyl)glutaroyl-dicaffeoylquinic acid, and one (3-hydroxy, 3-methyl)glutaroyl-cafeoyl-feruloylquinic acid. Six (3-hydroxy, 3-methyl)glutaroyl-cafeoylquinic acids were detected and two were tentatively assigned as 3-cafeoyl-4-(3-hydroxy, 3-methyl)glutaroylquinic acid and 3-cafeoyl-5-(3-hydroxy, 3-methyl)glutaroylquinic acid. The (3-hydroxy, 3-methyl)glutaroyl residue modifies the mass spectral fragmentation behavior and elution sequence compared with the chlorogenic acids that contain only a cinnamic acid residue(s). Fourteen of these twenty-nine chlorogenic acids have not previously been reported from any source. Copyright © 2010 John Wiley & Sons, Ltd.
Tips to make your article discoverable

- Make sure you have an SEO-friendly title for your article - The title needs to be descriptive and must incorporate a key phrase related to your topic. Put your keywords within the first 65 characters of the title.

- Carefully craft your abstract using keywords, keywords, keywords
  - Choose the appropriate keywords and phrases for your article. Think of a phrase of 2-4 words that a researcher might search on to find your article.
  - Consider looking up specific keywords on Google Trends or the Google Adwords keywords tool to find out which search terms are popular
  - Repeat your keywords and phrases 3-4 times throughout the abstract in a natural, contextual way.
  - BUT don’t be repetitious as search engines may un-index your article as a result.

- Provide at least five keywords or phrases in the keywords field - Include the keywords and phrases you repeated in your abstract. Provide additional relevant keywords and synonyms for those keywords as they relate to your article. Keywords are not only important for SEO, they are also used by abstracting and indexing services as a mechanism to tag research content.
Artwork

- Use one standard/common font (preferably Arial)
- Use one font size
- Avoid use of shadows/glows/reflections

- Check the author instructions with regard to reproducing colour
- For ChemDraw images use the object settings set by the journal
...so your article is written, format is good, time to submit your article with your covering letter

Together with the abstract of your paper, the cover letter is one of the first things the editor will see, so make it count!

Why is this topic important?
Why are these results significant?
What is the key result? (breakthrough!)
Why are you submitting to this journal?
Why will this journal’s readers read it?

Keep the letter as direct and short as possible. The longer it is, the easier it is to overlook something important.
The editorial workflow

Manuscript submitted

Editors examine and make Initial editorial decision
largely based on:
- Language
- Formatting / completeness
- Scope
- Is the article type correct
- Significance
- Readership
- Impact

Manuscript rejected
Manuscript sent back to author for alteration before resubmission

Ensures that the article is in a suitable state for peer review

Send to peer review
Submitted! And now??

Accept (no revisions)
...forget it. It almost never happens.

Revise – major/minor: ...frequent

Carefully consider reviewer comments
• Approach a revision decision as an opportunity to develop your paper into the best it can be
• Referee’s comments should not be seen as negative criticisms but development points
• Not all changes have to be made but require convincing arguments for changes not made

Remember! Your response may go back to reviewers. You may need to convince them and the editor!

Reject – tough luck!
Why???
(i) unclear/unimportant
(ii) Not novel/original
(iii) Conclusions do not support the data
(iv) Results not interesting
(v) Ethical questions
(vi) Unclear presentation
Most scientists regarded the new streamlined peer-review process as ‘quite an improvement.’
The Peer Review Process

Paper submitted to Editor

Editor sends paper to referees

Referees return comments

Editor makes decision

Reject
Accept
Send back for revision

Rejection before review

Rejection after review

Rejection after revision
What peer reviewers are asked to do – the referee report

- Is the motivation clear?
- Is the motivation important?
- Is the work novel and original?
- Are the conclusions supported by the data?
- Are the results important? (Are they interesting?)
- Is the presentation clear?
- Are there any ethical questions?
- Were any flaws or mistakes found?
- Should anything be added or removed?
- Are there any literature citations missing?
Editor responsibilities

- Ensure efficient, fair, and timely manuscript processing
- Ensure confidentiality of submitted manuscripts
- Make the final decision for accepting or rejecting
- Not use work reported in a submitted manuscript for their own research
- Ensure a fair selection of referees
- Act upon allegations of scientific misconduct
- Deal fairly with author appeals
Author responsibilities

- To gather and interpret data in an honest way
- To give due recognition to published work relating to their manuscript
- To give due acknowledgement to all contributors
- Notify the publisher of any errors
- To avoid undue fragmentation of work into multiple manuscripts (salami publishing)
- To ensure that a manuscript is submitted to only one journal at a time
Reviewer responsibilities

- Ensure confidentiality of manuscripts and respect privileged information
- Not to withhold a referee report for personal advantage
- Return to editor without review if there is a conflict of interest
- Inform editor quickly if not qualified or unable to review
- Judge manuscript objectively and in timely fashion
- Explain and support recommendations with arguments and references where appropriate
- Inform editor if plagiarized or falsified data is suspected
Ethical misconduct

Examples of ethical misconduct that are not tolerated:

- Falsifying data
- Fabricating data
- Plagiarism
- Multiple concurrent submissions
- Image manipulation
- Authorship misrepresentation
- Duplicate publication
Paper accepted! Next step: production
The life of an accepted article – the production process

Copyediting → Typesetting → Correction → Print product → Online product
Popular paper: High number of citations, published in a high Impact Factor Journal
What is an Impact Factor?? How is it computed??

\[ IF_y = \frac{Citations_{y-1} + Citations_{y-2}}{Publications_{y-1} + Publications_{y-2}} \]

For example, Nature had an impact factor of 41.577 in 2017

\[ IF_{2017} = \frac{Citations_{2016} + Citations_{2015}}{Publications_{2016} + Publications_{2015}} = \frac{32389 + 41701}{880 + 902} = 41.577 \]

*Article types counted in the denominator [citeable item]*
Primary research articles, Review articles, Case reports, Proceedings papers.

*Articles not counted in the denominator [not citeable item]*
Editorials, Corrections, Bibliographies, Letters, Abstracts.
Research published in a high impact factor journal is more valuable than research published in a low impact factor journal.
Pros and cons of the Impact Factor

Pros

• It is fundamentally a sound premise
• It is transparent
• It is easy to explain
• It is efficient
• After 50 years of use it is established

Cons

• Target period (window) is not appropriate for all subject areas
• Free citations to ‘non citable’ items
• A citation is not necessarily a validation
• Differences in referencing behaviour between subjects
• Misused to judge author performance
The H-index

- Equation: An individual has an index of $h$, when they have published at least $h$ papers, each of which has been cited at least $h$ times.

- The index was created in an attempt to move away from journal based measures such as the Impact Factor. An index which was applicable at the author level, accounting for the fact that the distribution of citations between articles can be tremendously skewed, even for the same author.

**Example 1**

**Bibliographic record for author A**

<table>
<thead>
<tr>
<th>Article</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Article 1</td>
<td>3</td>
</tr>
<tr>
<td>Article 2</td>
<td>3</td>
</tr>
<tr>
<td>Article 3</td>
<td>4</td>
</tr>
</tbody>
</table>

Author A has published 3 articles
3 of these articles have at least 3 citations = H index of 3
Example 2

Bibliographic record for author B

<table>
<thead>
<tr>
<th>Article</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Article 1</td>
<td>3</td>
</tr>
<tr>
<td>Article 2</td>
<td>4</td>
</tr>
<tr>
<td>Article 3</td>
<td>2</td>
</tr>
<tr>
<td>Article 4</td>
<td>9</td>
</tr>
<tr>
<td>Article 5</td>
<td>2</td>
</tr>
<tr>
<td>Article 6</td>
<td>1</td>
</tr>
</tbody>
</table>

Author B has published 6 articles. 3 of these articles have at least 3 citations = H index of 3. The author will gain a H index of 4 when a total of 4 articles achieve 4 or more citations each.
Pros and Cons of the H-index

Pros
- Elegant
- Efficient
- Transparent
- Removes bias from individual highly cited articles
- Can be mobilized to evaluate a number of criteria, not just the author (e.g. institute, country, region)

Cons
- Longitudinal bias
- No baseline
- It cannot decrease
- It may increase without an author publishing any new work
- Seminal thinkers may not necessarily publish lots of articles
- Citation index dependent – which database citation count is correct?
H-index = 8!
Part II. IUCR journals. Requirements and tools
...nanocrystals, metacrystals, amorphous materials, quasicrystals, synchrotron and XFEL studies, coherent scattering, diffraction imaging, time-resolved studies and the structure of strain and defects in materials.

...structural science of compounds and materials in the widest sense

Section C is the journal of choice for the rapid publication of articles that highlight interesting research facilitated by the determination, calculation or analysis of structures of any type, other than macromolecular structures.
...any aspect of structural biology, with a particular emphasis on the structures of biological macromolecules or the methods used to determine them. These structures may have been determined using any structural biology technique including crystallography, NMR, cryoEM and/or other techniques.

crystal structure determinations of inorganic, metal-organic and organic compounds.

...any aspect of structural biology, including structures determined using high-throughput methods or from iterative studies such as those used in the pharmaceutical industry.
identifying structural and diffusion-controlled phase transformations, structure-property relationships, structural changes of defects, interfaces and surfaces, etc.

...entire field of synchrotron radiation and free-electron laser research including instrumentation, theory, computing and scientific applications in areas such as biology, nanoscience and materials science.

...all aspects of the sciences and technologies supported by the IUCr via its commissions, including emerging fields where structural results underpin the science reported in the article. Our aim is to make IUCrJ the natural home for high-quality structural science results.
An overview of the standards required for publications. Policies of the journals that generally publish protein structures. Focus on the IUCr Journals.

**What to submit??**

...If your article reports a macromolecular crystal structure determination, you will also be asked to upload your

- PDB validation report,
- PDB code and name of your structure
- CIF file

**What is a PDB validation report??**

When coordinates and structure factors are deposited with one of the Worldwide Protein Data Bank (wwPDB) sites a PDB validation report is created and made available to the depositor.

**What is a CIF (or mmCIF)??**

mmCIF stands for macromolecular Crystallographic Information File which is a plain text file format which describes a macromolecular crystallographic experiment and its results.

The crystallographic information file uses the STAR file structure, with basic and quite simple syntax
And if you publish your biomacromolecule structure in less specialistic journals??

Manuscripts reporting new crystallographic structures of small molecules must be accompanied by a standard .cif file. A structural figure with probability ellipsoids should be included in the main supplementary information file. The structure factors for each structure should also be submitted, preferably embedded in the main .cif file, although they may be provided as a separate .hkl and/or .fcf file. Use of the 2014 version of the program SHELXL, which embeds the structure factors information in the main .cif file, is encouraged. The structure factors and structural output must be checked using IUCr's CheckCIF routine and a PDF copy of the output supplied, explaining any A- or B-level alerts.

.... At initial submission, if a paper reports any structure(s) determined by X-ray crystallography we require inclusion of a PDB Validation Report(s) before we will proceed with editorial consideration. We strongly prefer to receive the Full Validation Report, but the Preliminary Full Validation Report obtained after data deposition is also acceptable. The initial validation server report is not sufficient. We will use these reports for peer review purposes, and they will not be part of the final published paper.

...Still. You will have to prove that your model is correct.
For structures of biological macromolecules, the atomic coordinates and the related experimental data (structure factor amplitudes/intensities and/or NMR restraints and chemical shifts) must be deposited in the PDB at a member site of the Worldwide Protein Data Bank, RCSB PDB, PDBe, PDBj or BMRB. Authors must also submit the annotated PDB Summary Validation Report (designated with a Confidential watermark) to JBC for review at the time of submission. Click here for more information regarding validation reports.

...Still. You will have to prove that your model is correct.
Why all of these restrictions?? Why do we need validation reports???
The famous 2010 Fakery

2010 - Acta Crystallographica Section E issued a bombshell editorial. The journal was pulling 70 papers from two groups of researchers at the same Chinese university after discovering that the structures they reported had been fakes.

The falsified structures have many features in common: in each case, a bona fide set of intensity data, usually on a compound whose structure had been correctly determined and reported in the literature, was used to produce a number of papers, with the authors changing one or more atoms in the structure to produce what appeared to be a genuine structure determination of a new compound. The worst example generated no fewer than 18 supposedly original structures from a single common set of data.

The two main authors, Mr X — with 41 pulled papers — and Mr Y — with 39 — from Jinggangshan University, were fired in the wake of the scandal, Nature reported in January 2010. According to that article, Chinese researchers submitted “half of the 200,000-odd crystal structures” published by Acta E in the five years leading up to the retractions.
A touch of history – Letter of the IUCR director to publishers and editors of scientific works reporting crystal structure determinations – June 2011

Publication standards for crystal structures

• On behalf of the International Union of Crystallography (IUCr), we are writing to publishers and editors of scientific works reporting crystal structure determinations about the importance of retaining and validating supporting experimental data sets.

• ...Crystallography has recently seen a number of high-profile retractions of papers from the peer-reviewed literature, both in the small molecule and in the macromolecular areas. These cases frequently involved actions of scientific misconduct such as manual alterations to cell constants and atom types to produce what appeared to be genuine structure determinations of new compounds. Such cases are not always easy to detect, but once published, they compromise the quality and dependability of the scientific literature.

• The response of the IUCr has been to develop and implement measures that will help to ensure that such cases are identified before publication. It has become clear in the process that many of these problems could not have been readily discovered without the availability of the structure-factor files. Deposition of such files is currently mandatory for publications in the macromolecular field, but only required by a small number of journals publishing inorganic and small molecule structures. We therefore strongly recommend that all journals publishing crystal structures require authors to provide structure factors for checking and archiving.

• In order to help editors and publishers follow this recommendation, the IUCr checkCIF service at http://checkcif.iucr.org has been modified to allow checking of crystallographic information files (CIFs) and structure factors. In addition, the Cambridge Crystallographic Data Centre will also now allow authors to deposit structure-factor files.

Yours sincerely

Sine Larsen, President of the IUCr
Gernot Kostorz, Editor-in-chief, IUCr Journals
After the myth of all possible crystallographic frauds....

The response of the IUCr has been to develop and implement measures that will help to ensure that such cases are identified before publication.

E.g. Strictly require structure factors for checking and archiving.

The IUCr checkCIF service at http://checkcif.iucr.org to allow checking of crystallographic information files (CIFs) and structure factors.
In practical terms, the important points (small molecules):

(1) **Encourage authors to provide a CIF and structure factors** in machine-readable format for each submitted structure. These might be collected as supplementary material or deposited with the Cambridge Structural Database.

(2) Ask the authors to **check the CIF and structure factors before submission** using the checkCIF service.

(3) Use the checkCIF service to create a **validation report** for each submitted structure and make this report available to the referees.

(4) **Ensure that referees and editors know how to interpret checkCIF reports** correctly.
What IUCR tools are available to help prepare my article?

i) Preparing an enhanced figure. A toolkit is available for preparing an enhanced figure for your article using an mmCIF.

Enhanced figures in Crystallography Journals Online are interactive visualizations of crystal and molecular structures that appear in many journal articles. They use the open-source Java program Jmol and can contain author-created scripts to show different aspects of the structure.

Viewing enhanced figures

Requirements

You will require a web browser supporting JavaScript and Java, and both must be activated. Additionally, LiveConnect technology is required to operate buttons and check boxes correctly. Browsers known to work include:
You need a CIF!!!!
For macromolecules???

Macromolecular Crystallographic Information File (mmCIF)

The mmCIF file format, which has been developed under the auspices of the International Union of Crystallography (IUCr), is to extend the Crystallographic Information File (CIF) data representation used for describing small molecule structures and associated diffraction experiments.

The format of the small molecule CIF dictionary and the data files based upon that dictionary conform to a restricted version of the Self Defining Text Archive and Retrieval (STAR) representation developed by Hall. STAR permits a data organization that may be understood by analogy with a spoken language.
IUCR tools to prepare your article

ii) The IUCr WORD template is available at [http://journals.iucr.org/services/wordstyle.html](http://journals.iucr.org/services/wordstyle.html) for preparing your article.

The IUCr Word templates utilize the content management features and document styles of Word to format your manuscript and to store essential details for submission of your manuscript.

The template is also packaged with tools to access external resources, either online (e.g. the World Directory of Crystallographers) or locally (e.g. a personal database of author details collected during use of the template).

In addition, tools are provided to:

- prepare tables of experimental and geometric data from a CIF
- prepare tables of experimental data from an mmCIF
- prepare tables of experimental data for a structure in the Protein Data Bank
Alternatively, **publBio publisher software** is available at http://publbio.iucr.org for preparing your article.
How to do it:

Register on https://publbio.iucr.org/

...you will get here....

You may submit in the PDB and then...enter a PDB code OR download an mmCIF from the PDB and upload it.

!!

Enter a PDB code and/or enter a SASBDB code

Or upload a local PDB or mmCIF file

Or fetch a PDB or mmCIF file from an internet location

Or write an article without a data source
Structure of the Resuscitation Promoting Factor Interaction

R. Berisio

Synopsis
Type text here

Abstract
Type text here

Paper options

Acta Cryst. F
research communication

- Include table of crystallization data
- Include structure refinement data
- Include table of SAS experimental data
You may compile your Tables in PublBio, but ..... 

By loading a data source (PDB code or mmCIF), publBio will populate the template with as much information as possible.
When you are done with Tables Figures... save pdf for reading...then ...Submit
Part III. Let’s write a paper together using PublBio

https://publbio.iucr.org/

*From theory to Practice*