Bioexcel Workshop

SHARING DATA FROM MOLECULAR SIMULATIONS

Sunday 25th November - Tuesday 27th November

Bioexcel Workshop

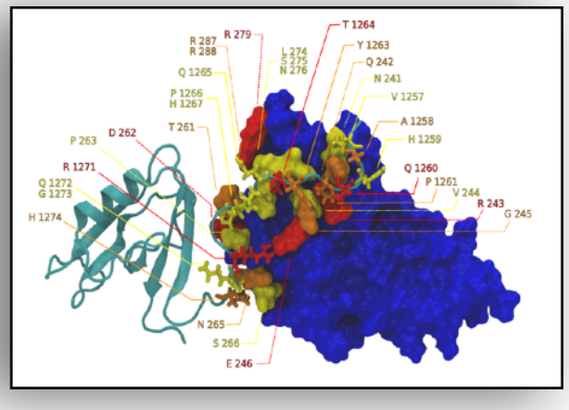
SHARING DATA FROM MOLECULAR SIMULATIONS

WELCOME!

Sunday 25th November - Tuesday 16th November

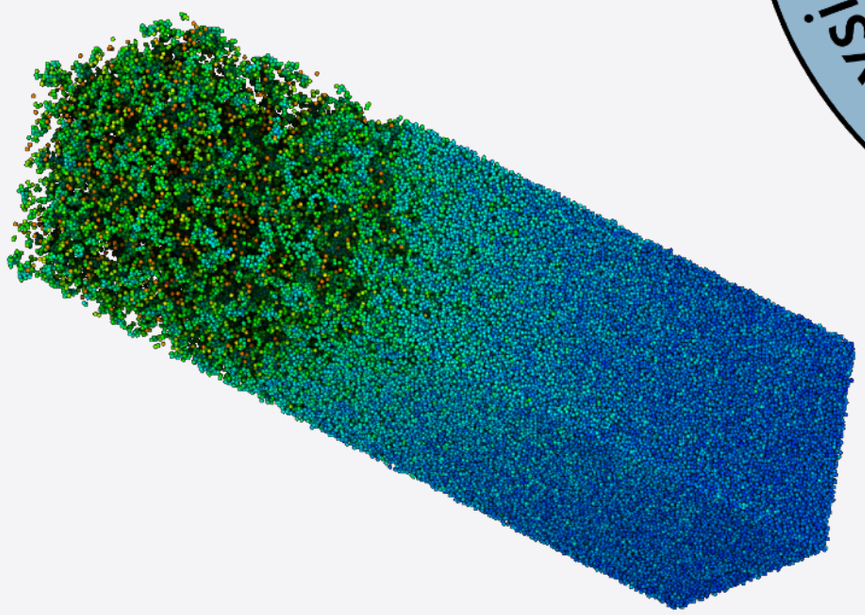
ME, MYSELF AND I

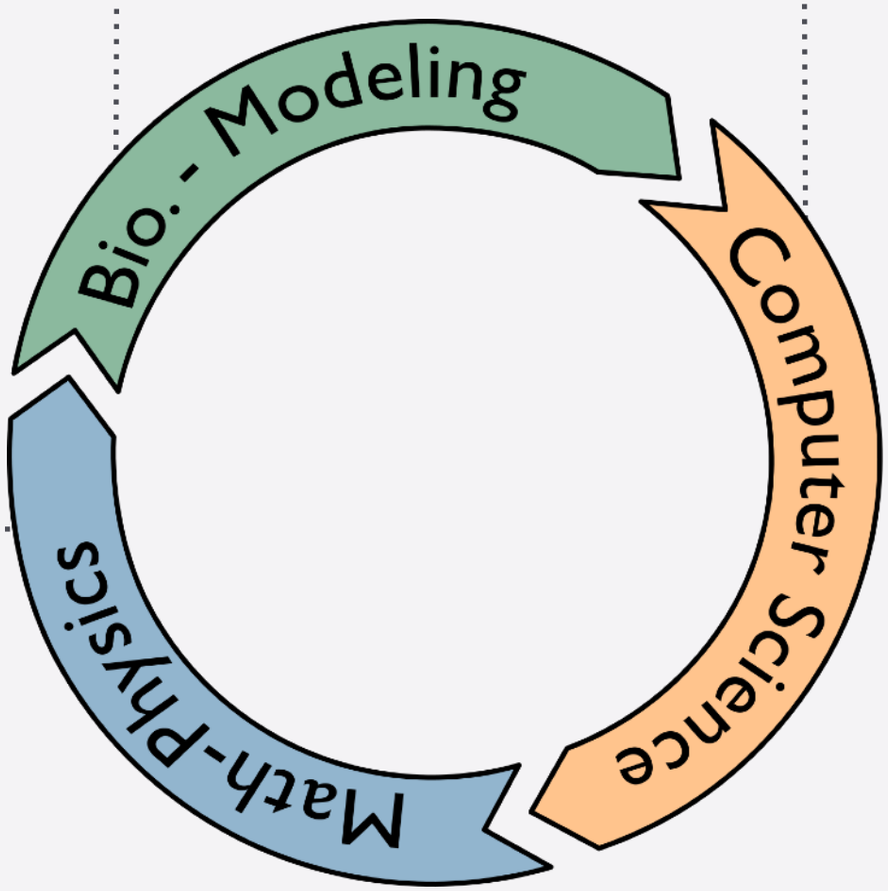
PROTEIN DOCKING



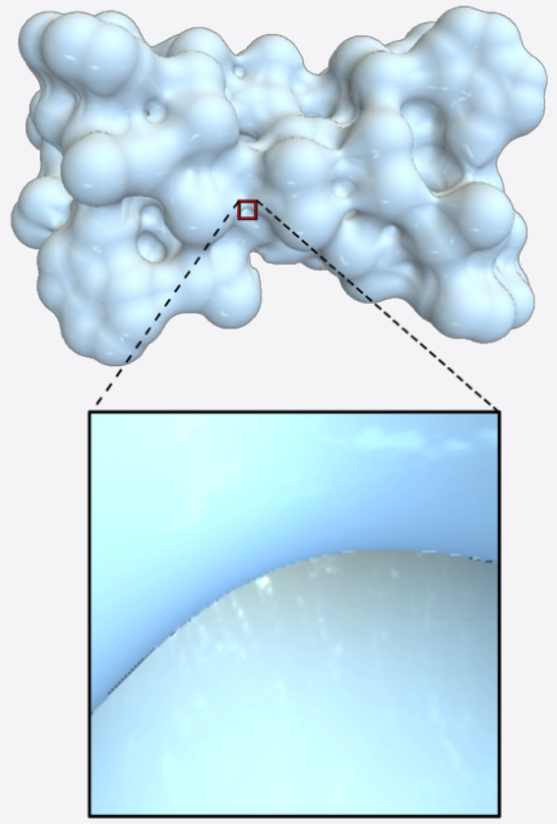
Déliot*, **Chavent*** *et al.* , *BBRC*, ***2009***

MATERIAL PHYSICS





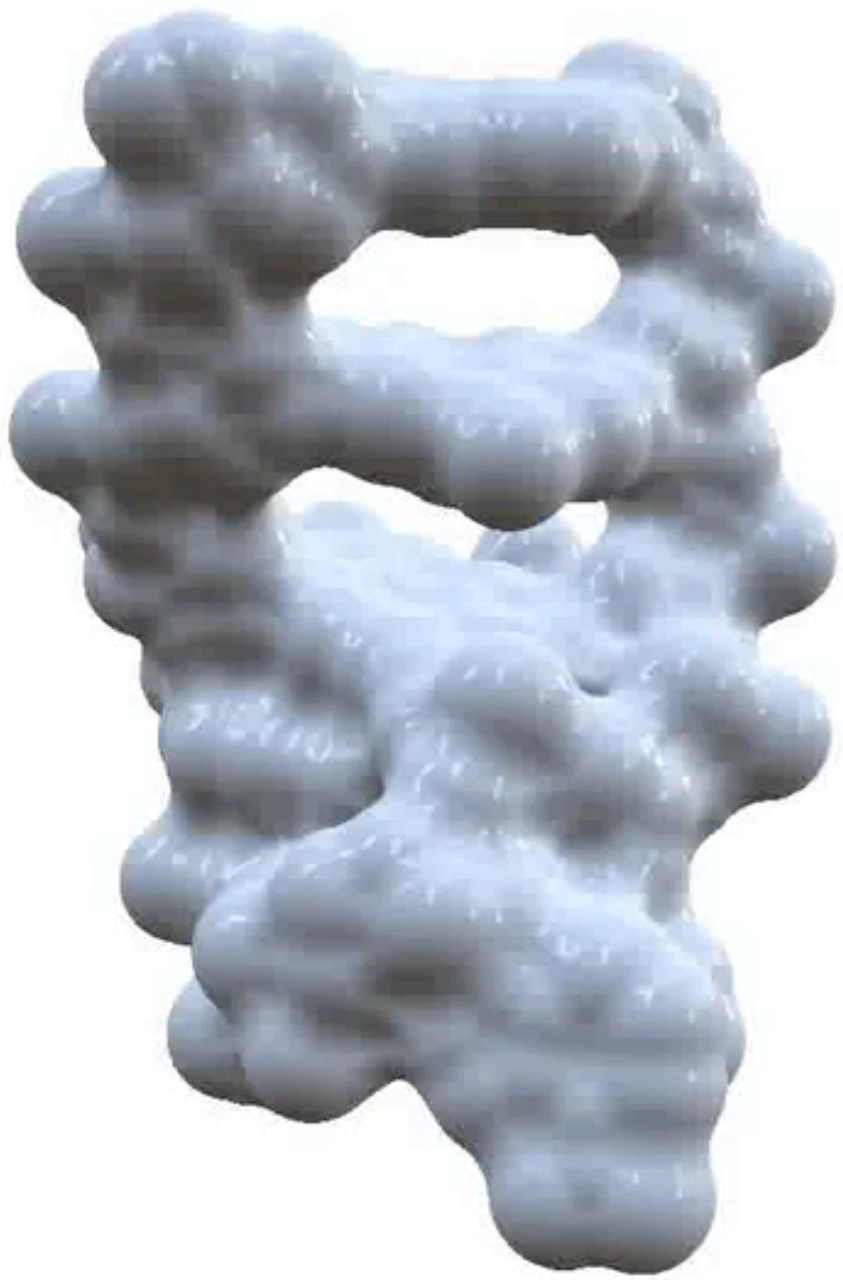
MOLECULAR SURFACE



Chavent *et al.* , *J. Mol. Graph. Mod.* , ***2008***

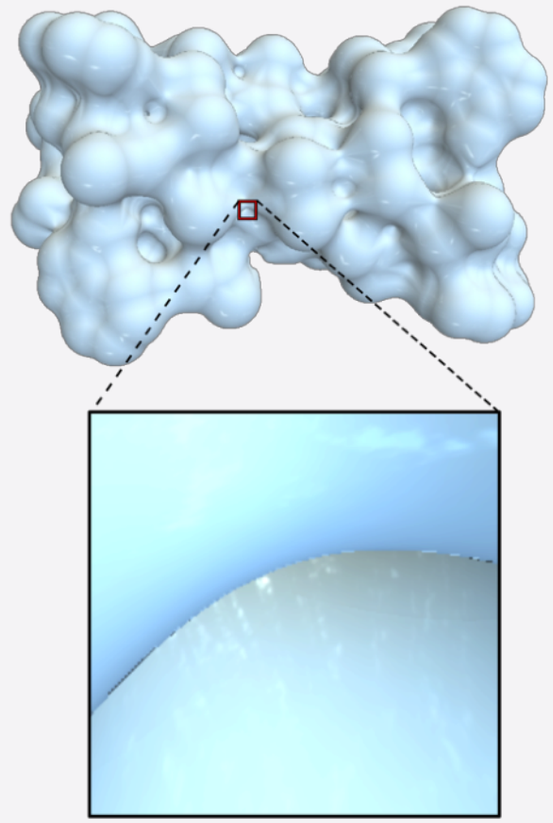
Chavent *et al.* , *J. Comp. Chem.*, ***2011***

MOLECULAR SKIN SURFACE



Edelsbrunner , *Disc. Comp. Geom.*, ***1999***

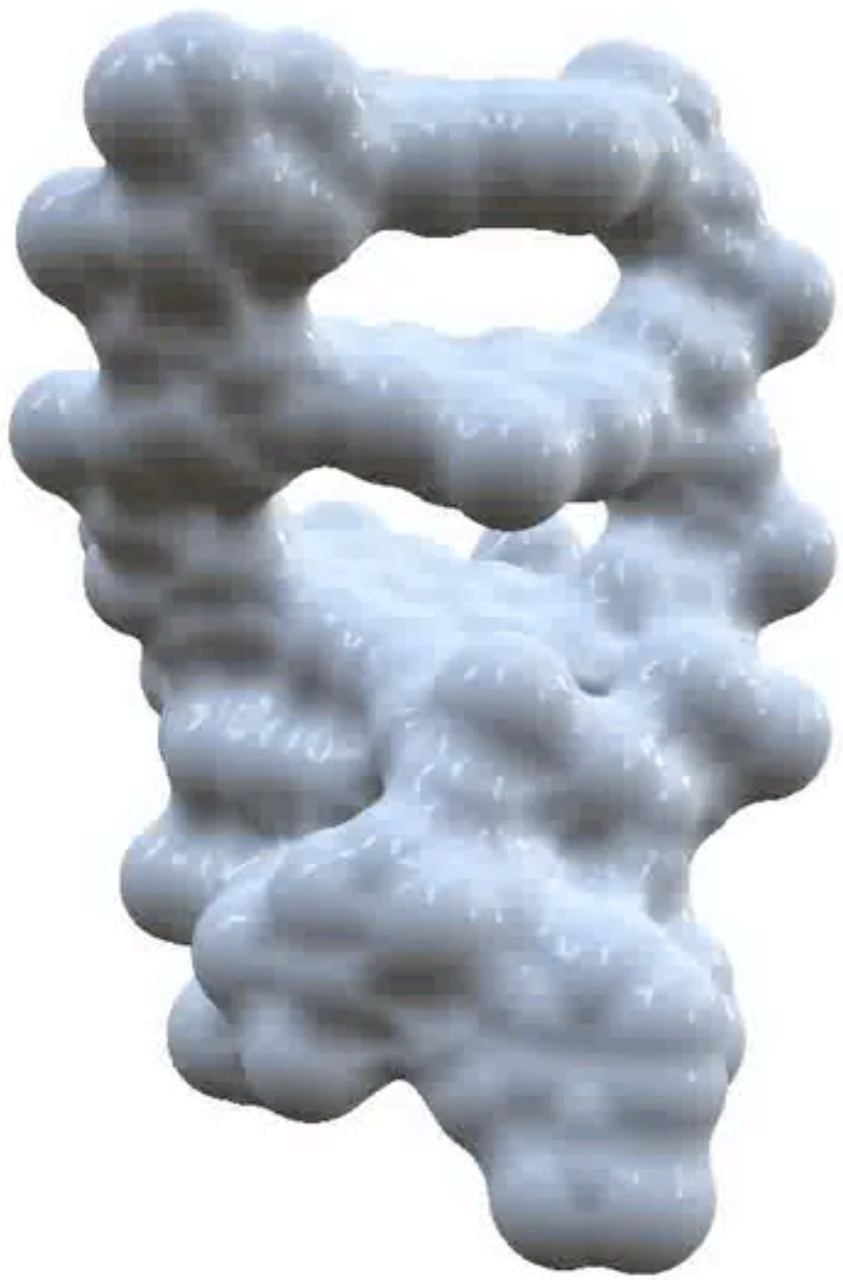
MOLECULAR SURFACE



Chavent *et al. , J. Mol. Graph. Mod. ,* ***2008***

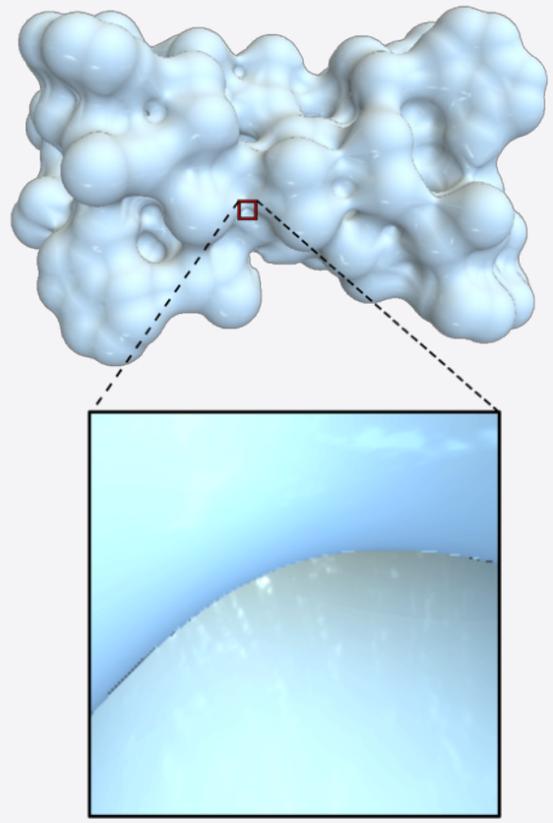
Chavent *et al. , J. Comp. Chem.,* ***2011***

MOLECULAR SKIN SURFACE



Edelsbrunner , *Disc. Comp. Geom.*, ***1999***

MOLECULAR SURFACE

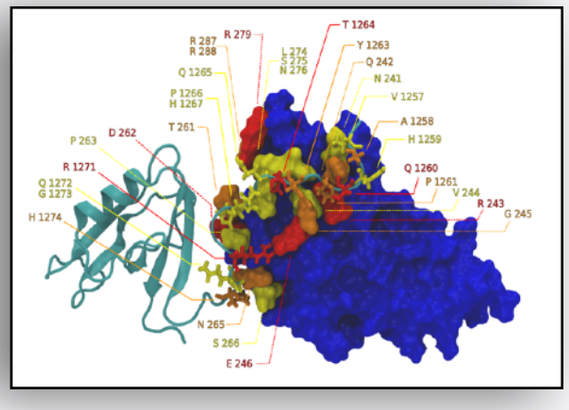


Chavent *et al. , J. Mol. Graph. Mod. ,* ***2008***

Chavent *et al. , J. Comp. Chem.,* ***2011***

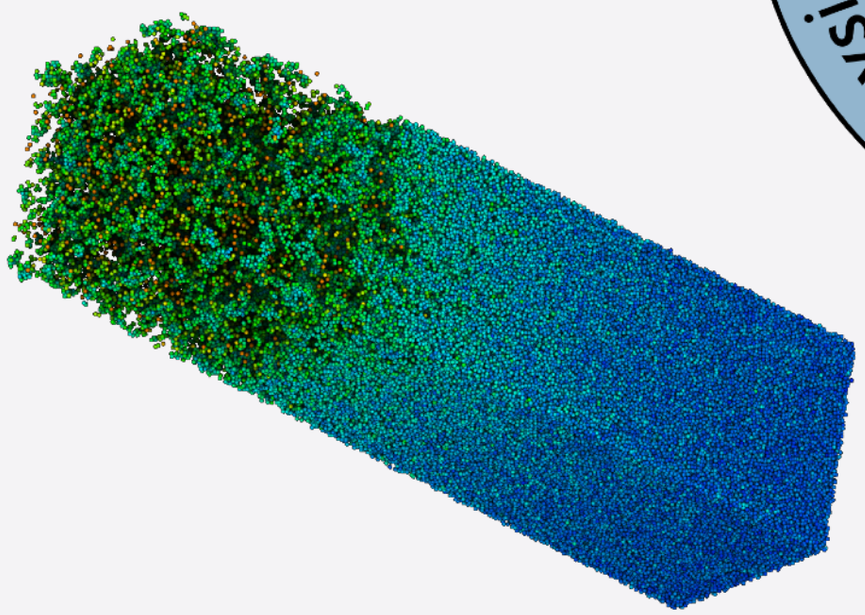
ME, MYSELF AND I

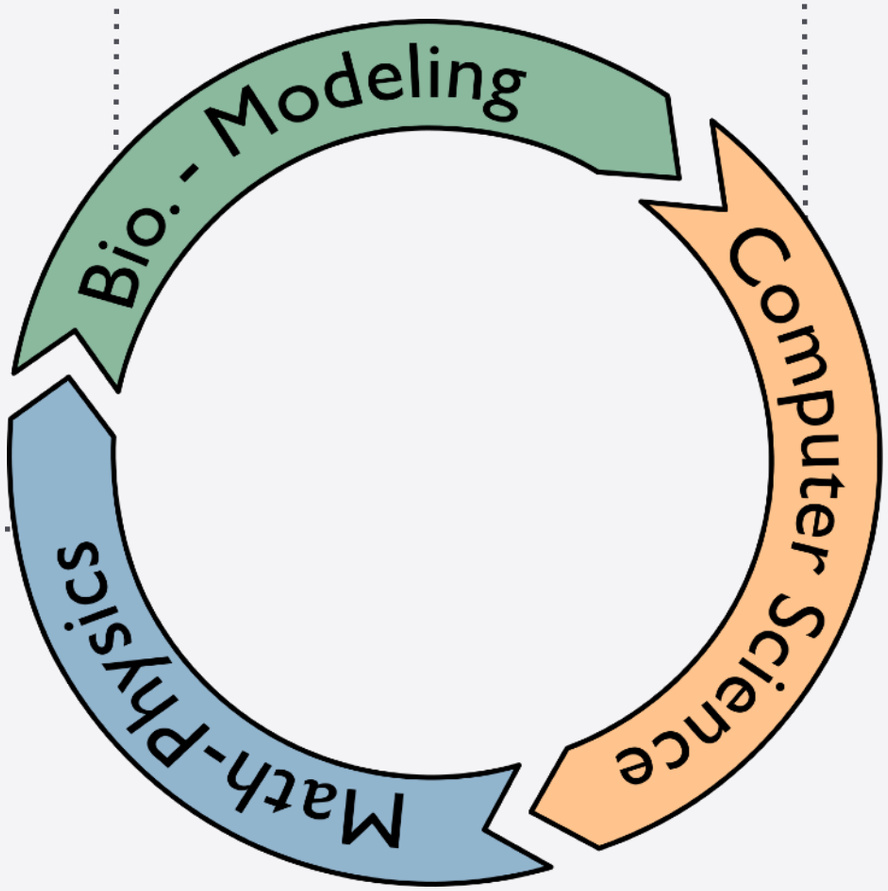
PROTEIN DOCKING



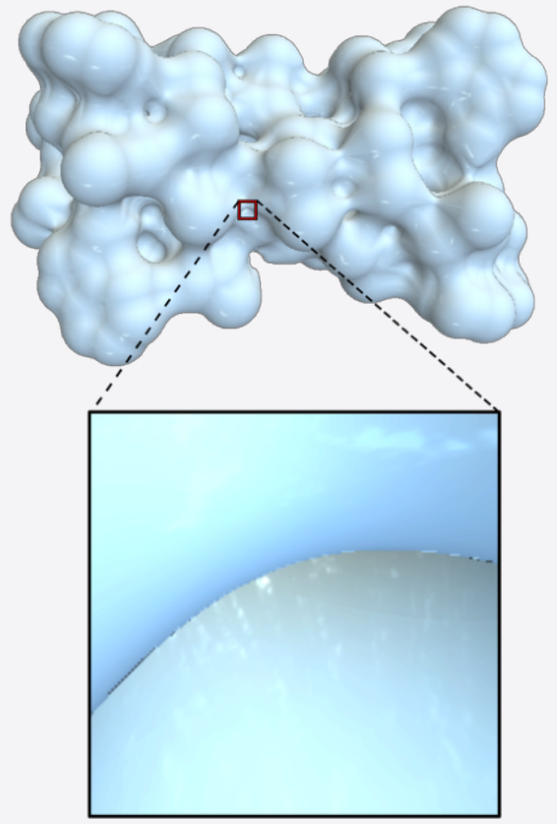
Déliot*, **Chavent*** *et al.* , *BBRC*, ***2009***

MATERIAL PHYSICS





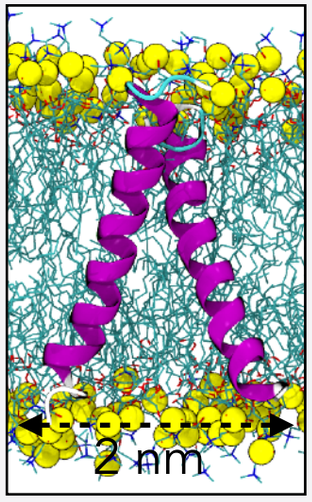
MOLECULAR SURFACE



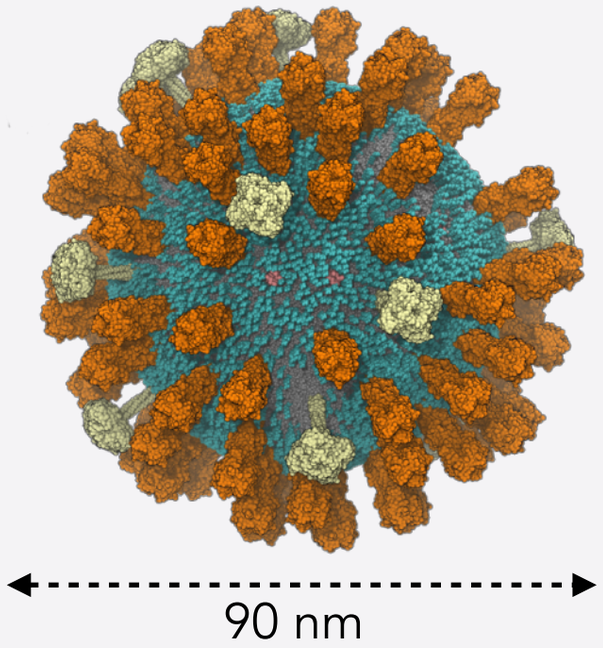
Chavent *et al.* , *J. Mol. Graph. Mod.* , ***2008***

Chavent *et al.* , *J. Comp. Chem.*, ***2011***

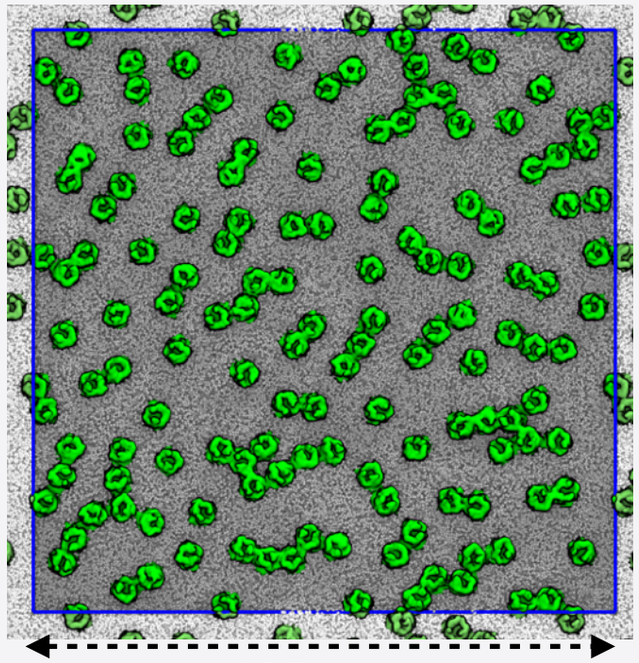
ME, MYSELF AND I



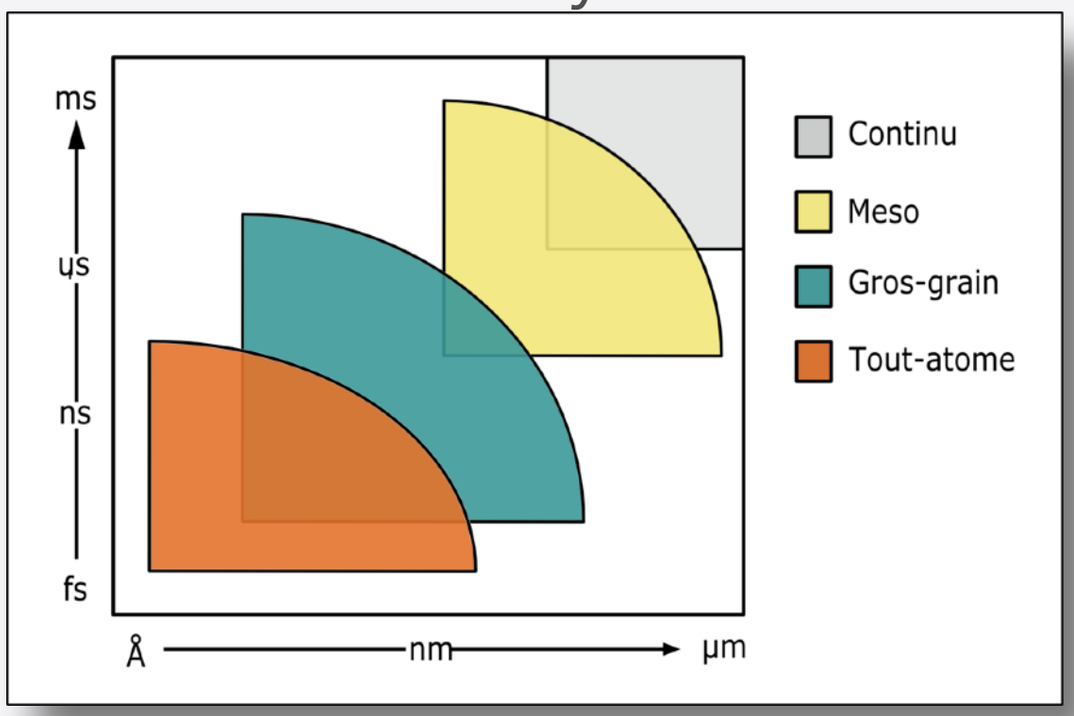
Chavent, *et al.* , *Biochemistry*, ***2014***

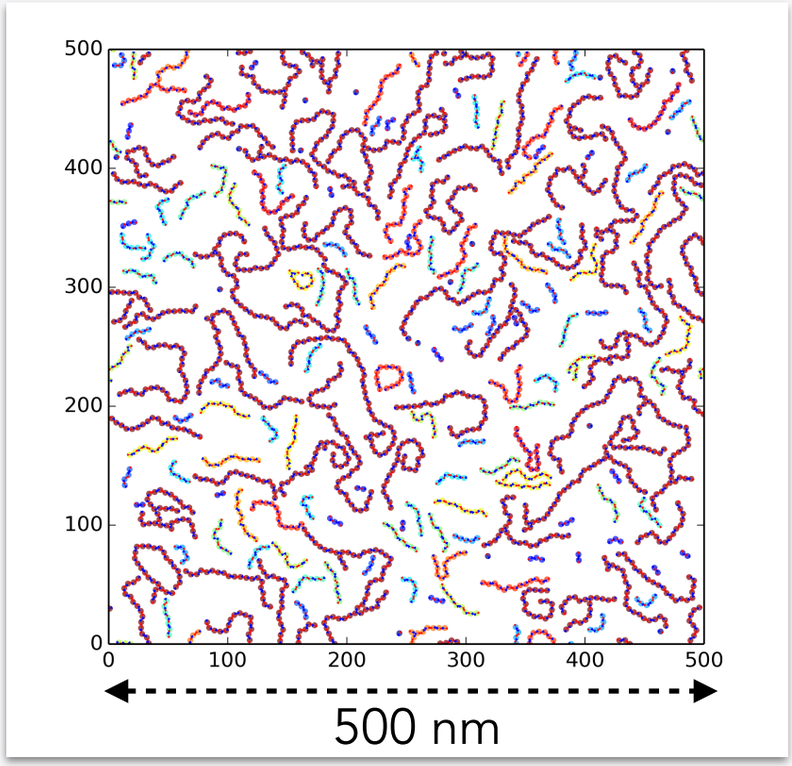


Reddy, *et al.* , *Structure*, ***2015***



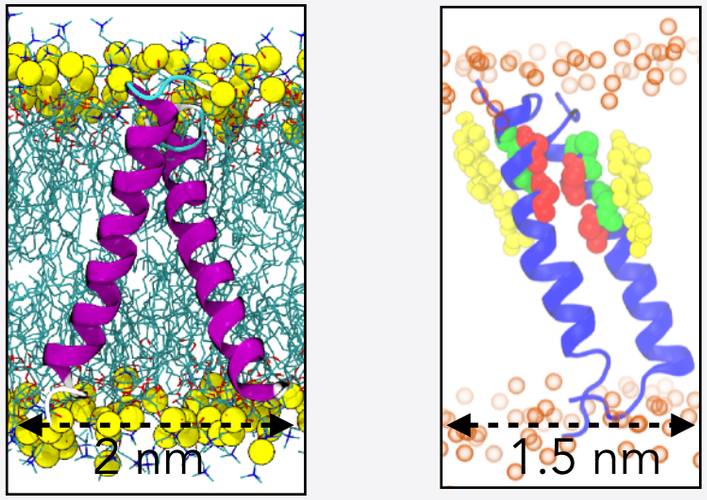
Molecular dynamics





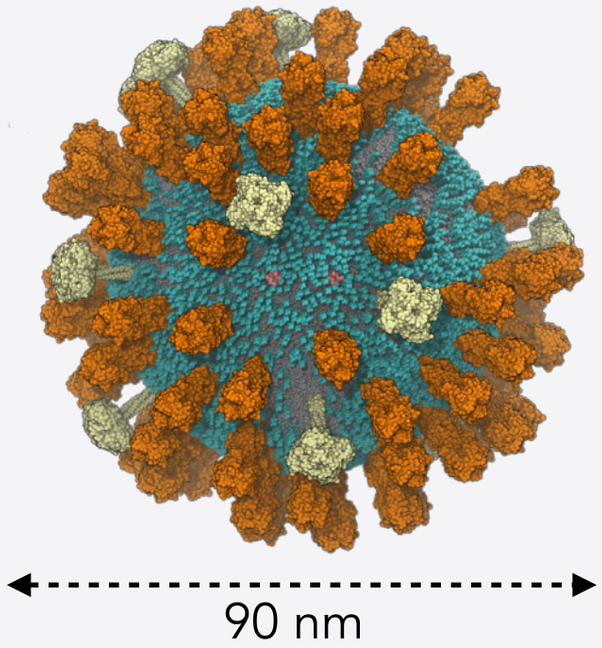
Chavent, *et al.* , *Nature Com.*, ***2018***

ME, MYSELF AND I

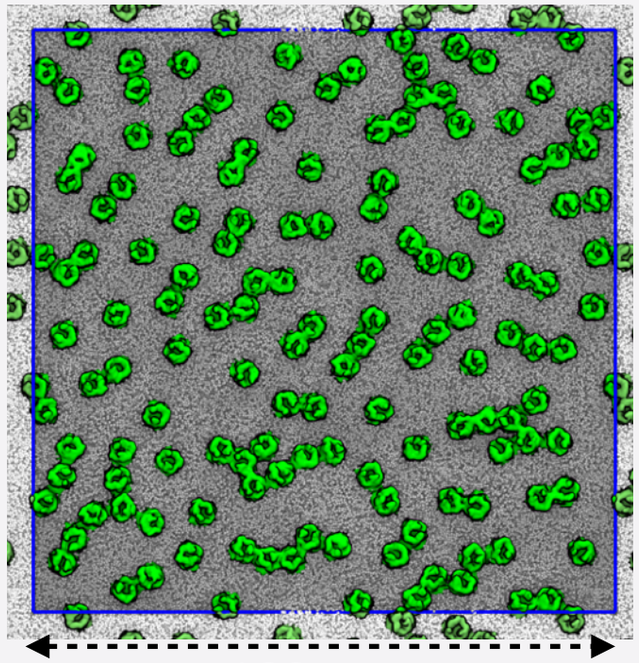


Chavent, *et al.* , *Biochemistry*, ***2014***

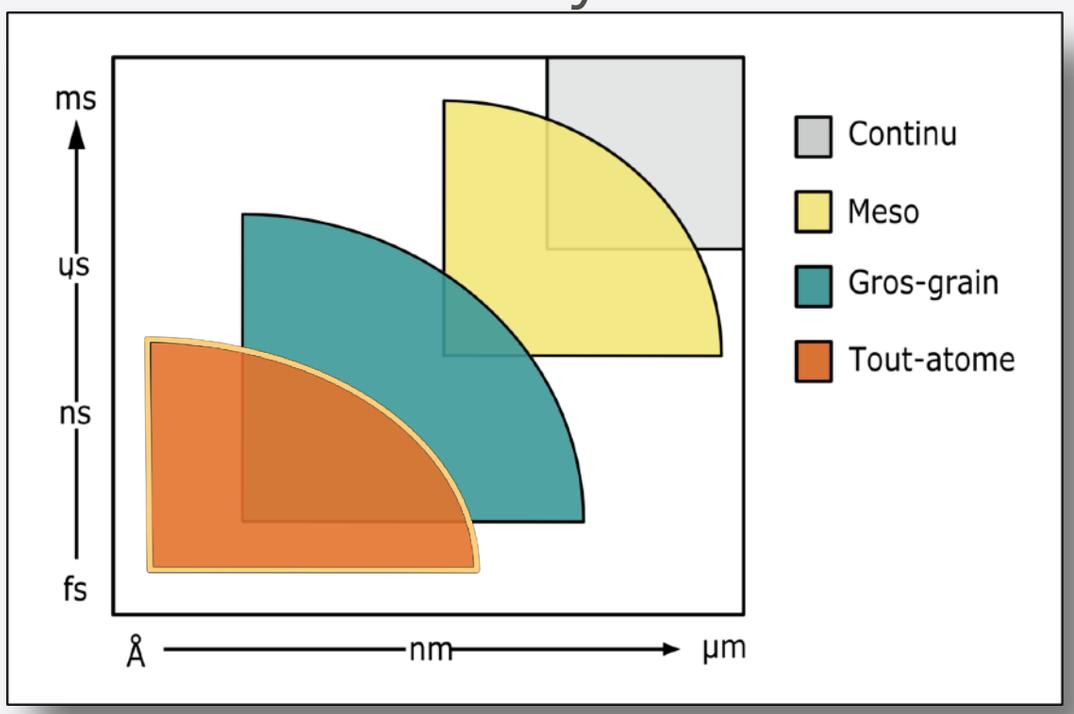
Atomistic (AT)

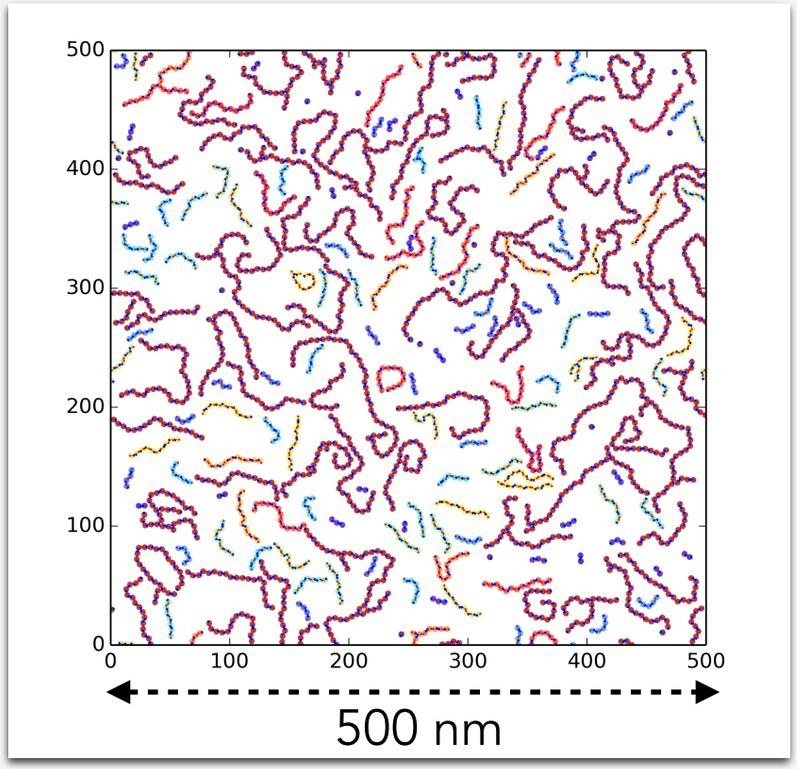


Reddy, *et al.* , *Structure*, ***2015***



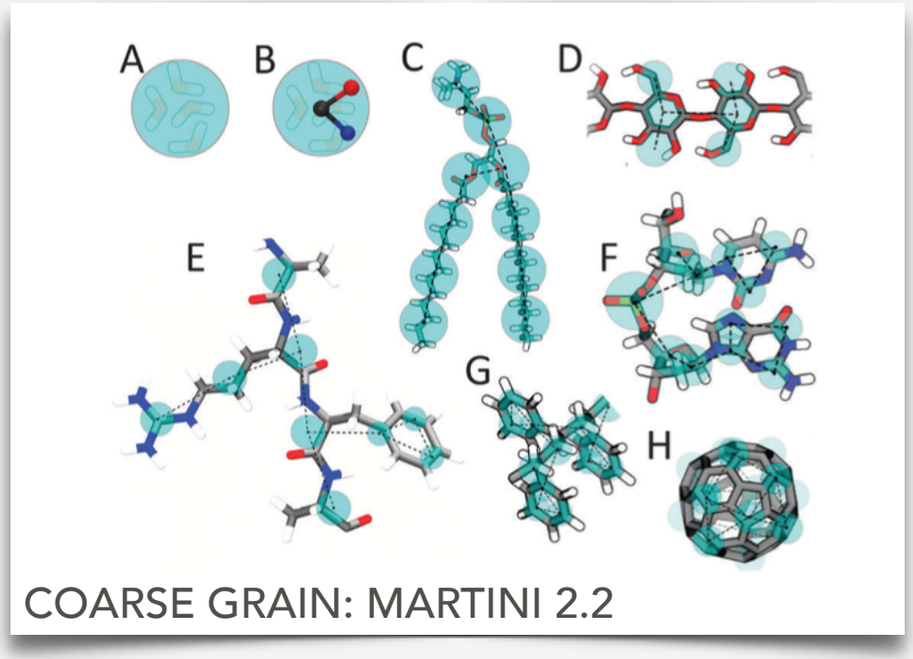
Molecular dynamics





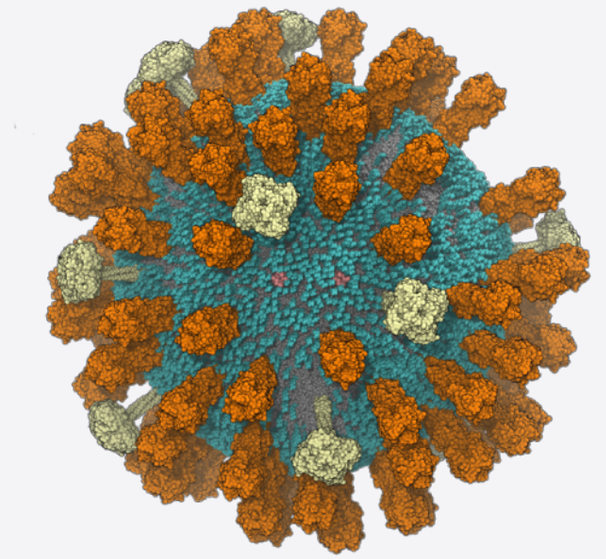
Chavent, *et al.* , *Nature Com.*, ***2018***

ME, MYSELF AND I

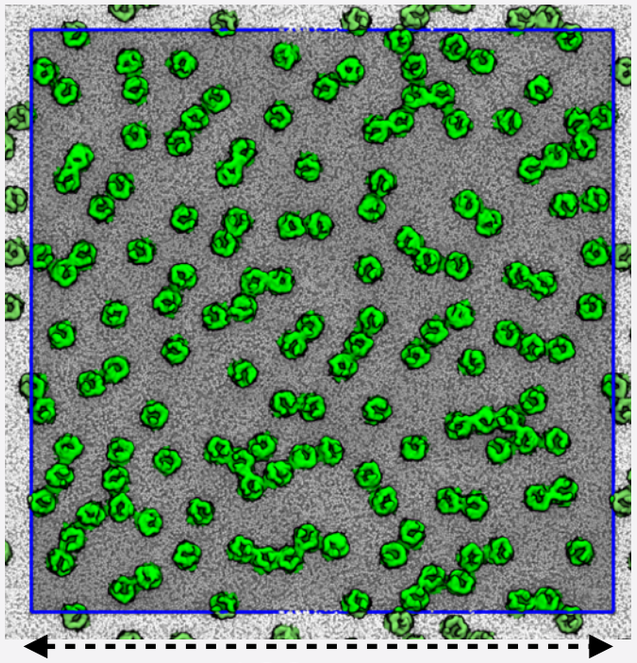


COARSE GRAIN: MARTINI 2.2

S. J. Marrink & P. T. Tielman, *Chem. Soc. Rev.* ***2013***



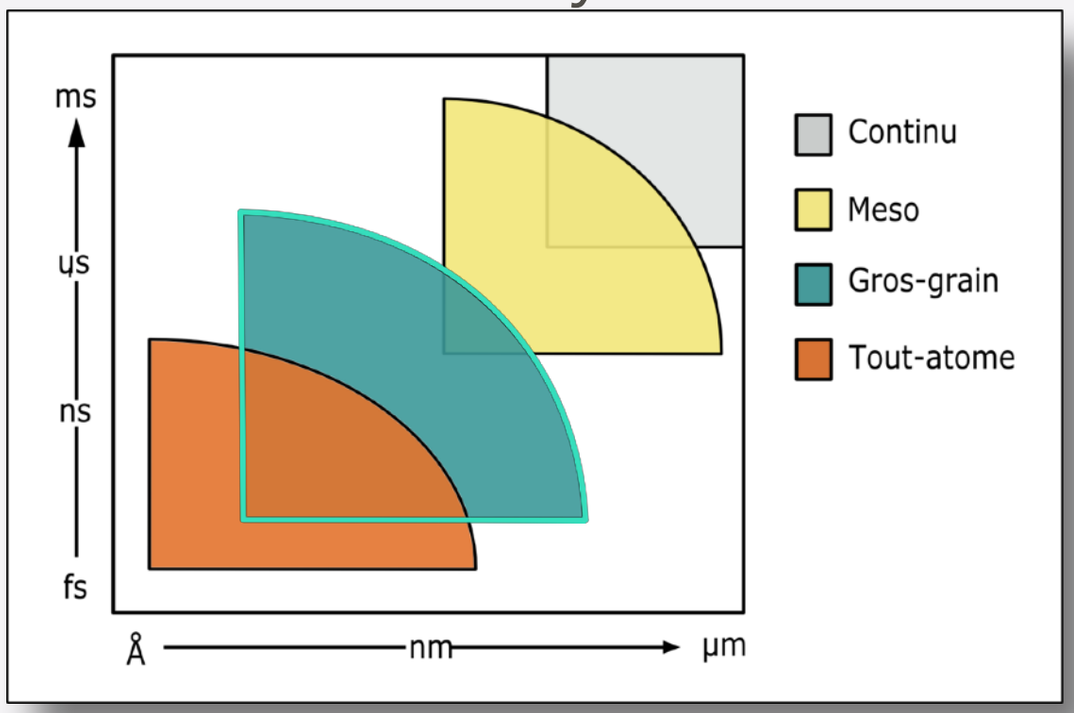
90 nm



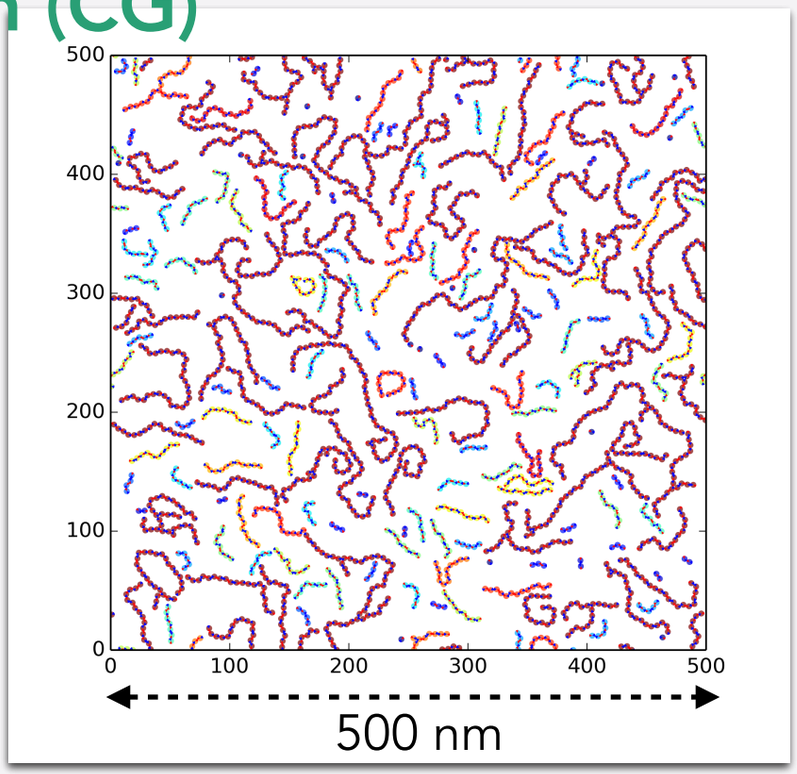
120 nm

Reddy, *et al.* , *Structure*, ***2015***

Molecular dynamics

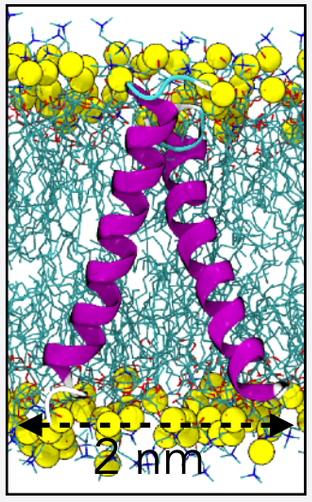


Coarse-Grain (CG)

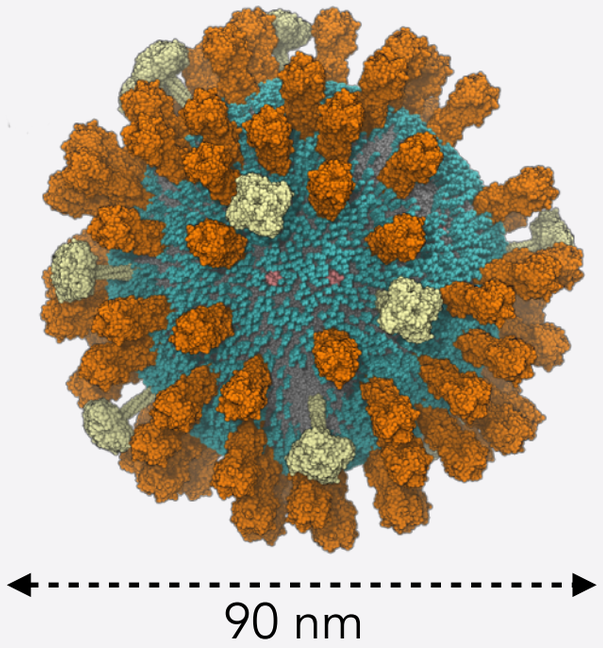


Chavent, *et al.* , *Nature Com.*, ***2018***

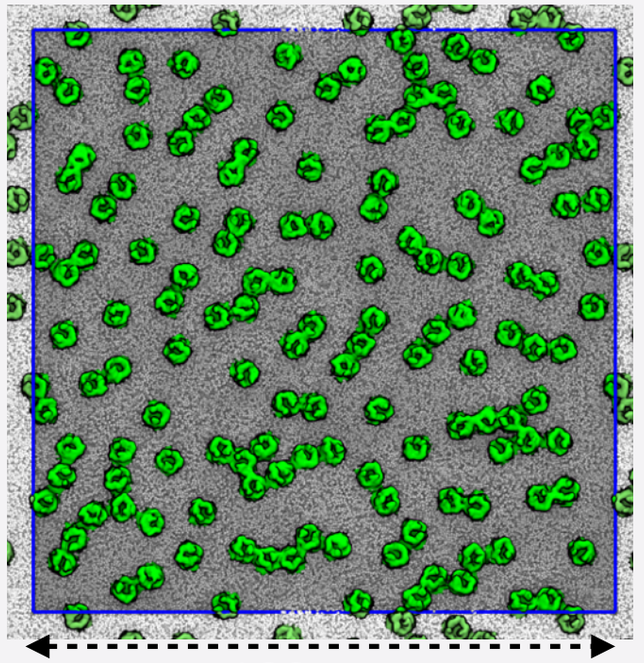
ME, MYSELF AND I



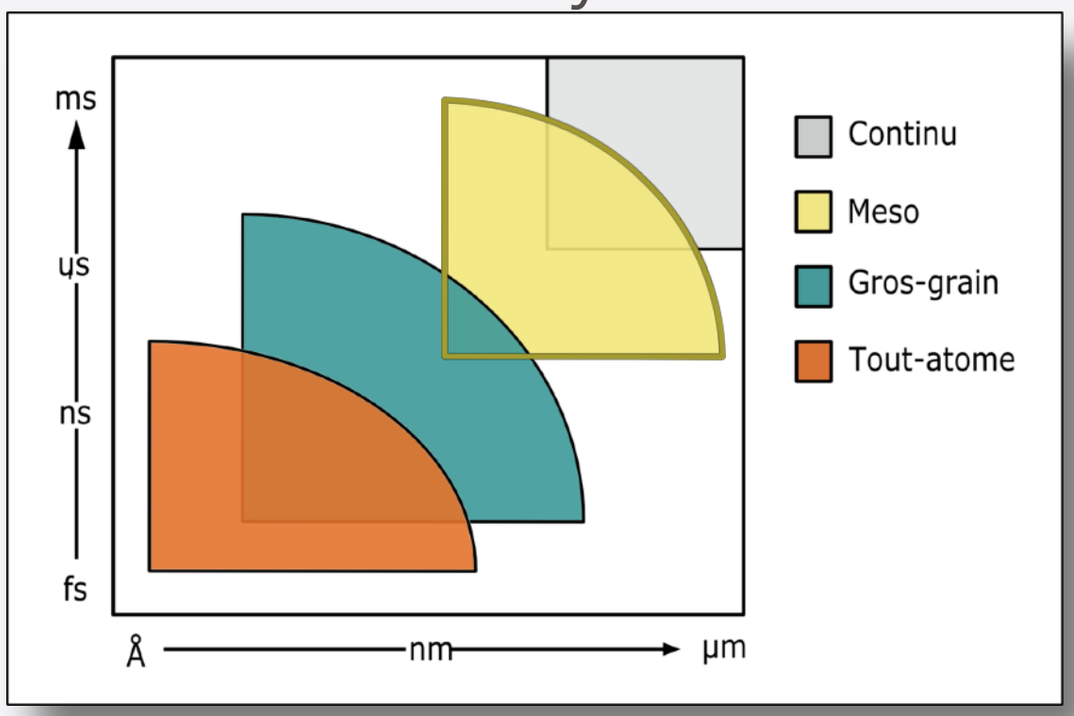
Chavent, *et al.* , *Biochemistry*, ***2014***



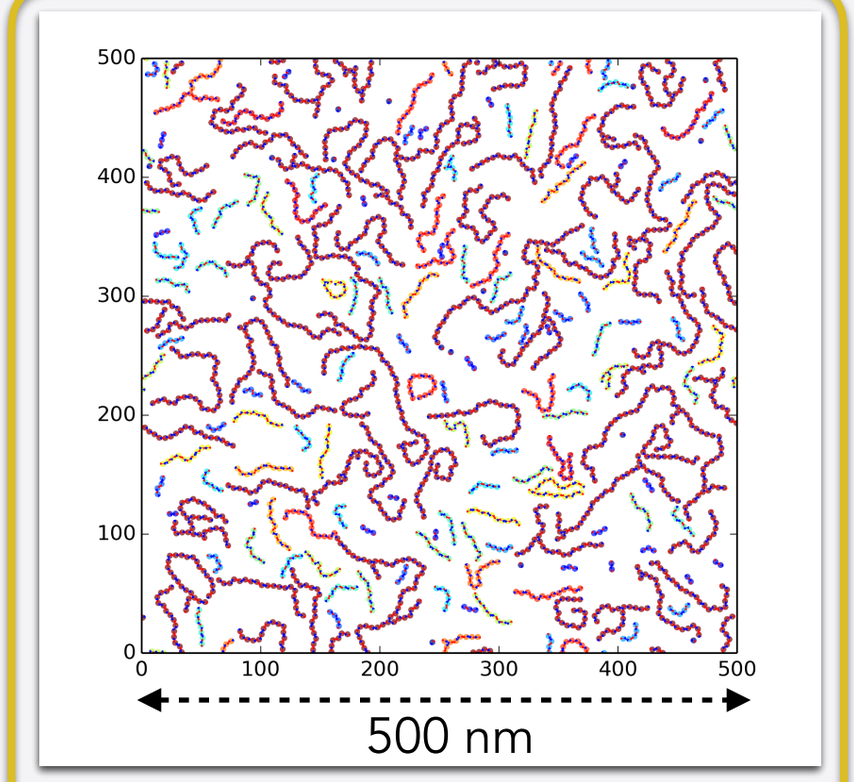
Reddy, *et al.* , *Structure*, ***2015***



Molecular dynamics



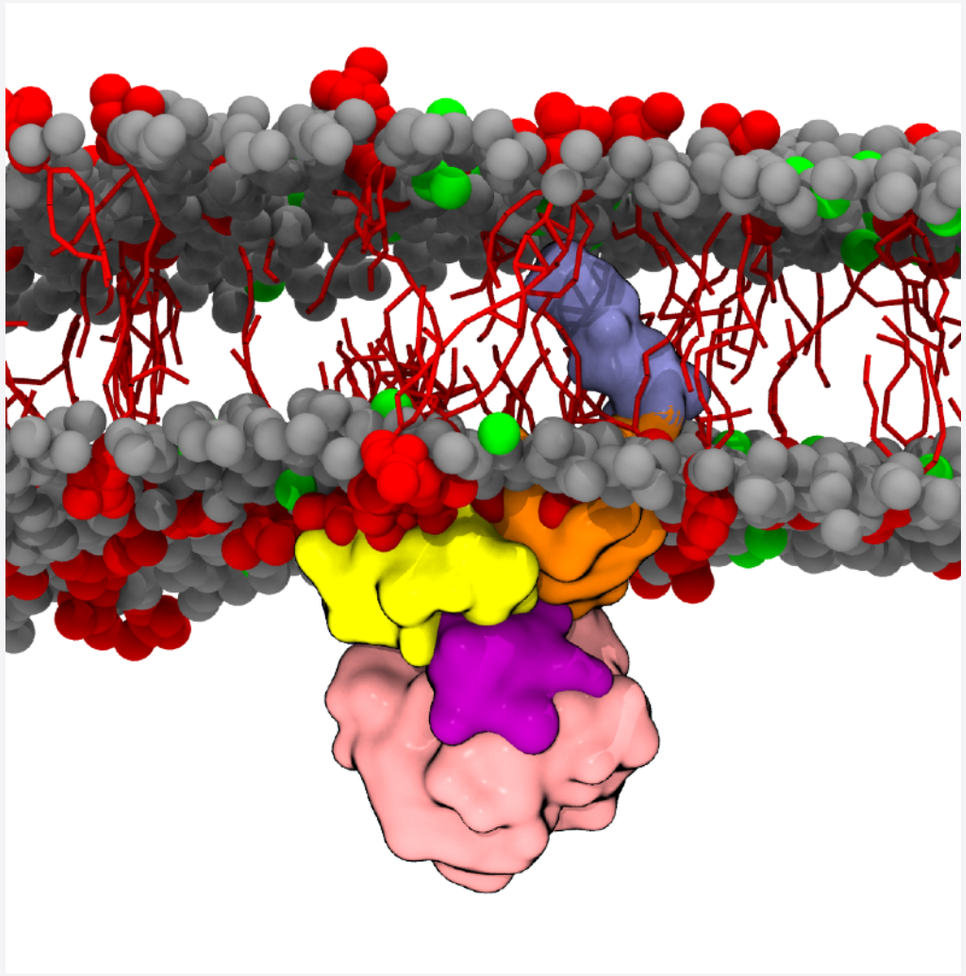
Meso



Chavent, *et al.* , *Nature Com.*, ***2018***

HOW TO SHARE MD DATA ?

EphA2 KINASE interacting with PIP2 molecules

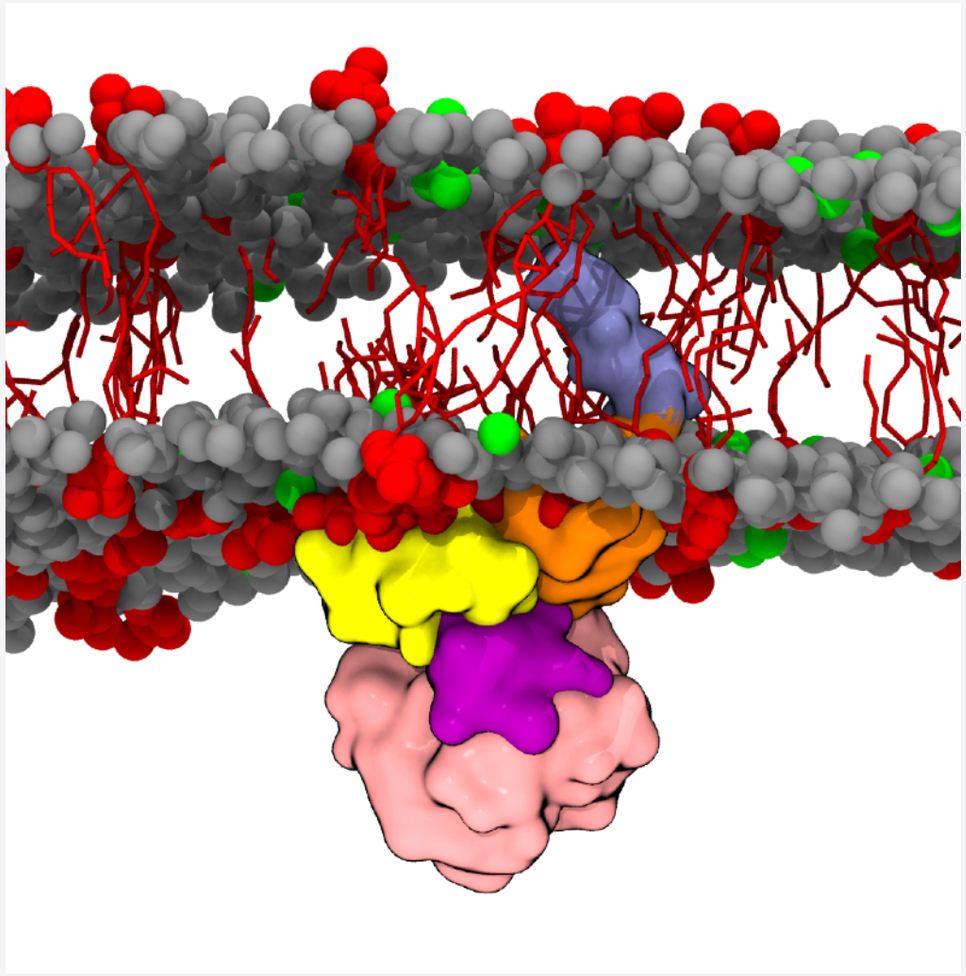


last frame -> file size: 10-50Mb

Chavent, *et al.* , *Structure*, ***2018***

HOW TO SHARE MD DATA ?

EphA2 KINASE interacting with PIP2 molecules



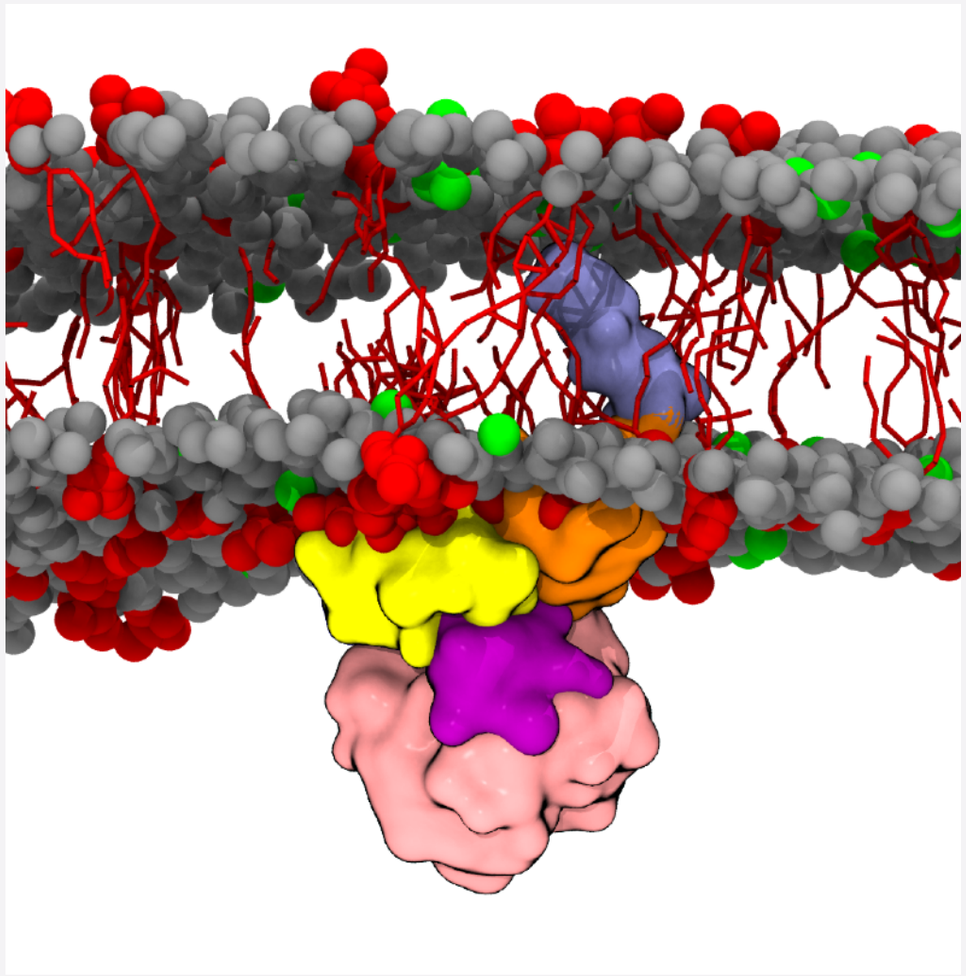
last frame -> file size: 10-50Mb

Zenodo and link in SI

Chavent, *et al.* , *Structure*, ***2018***

HOW TO SHARE MD DATA ?

EphA2 KINASE interacting with PIP2 molecules



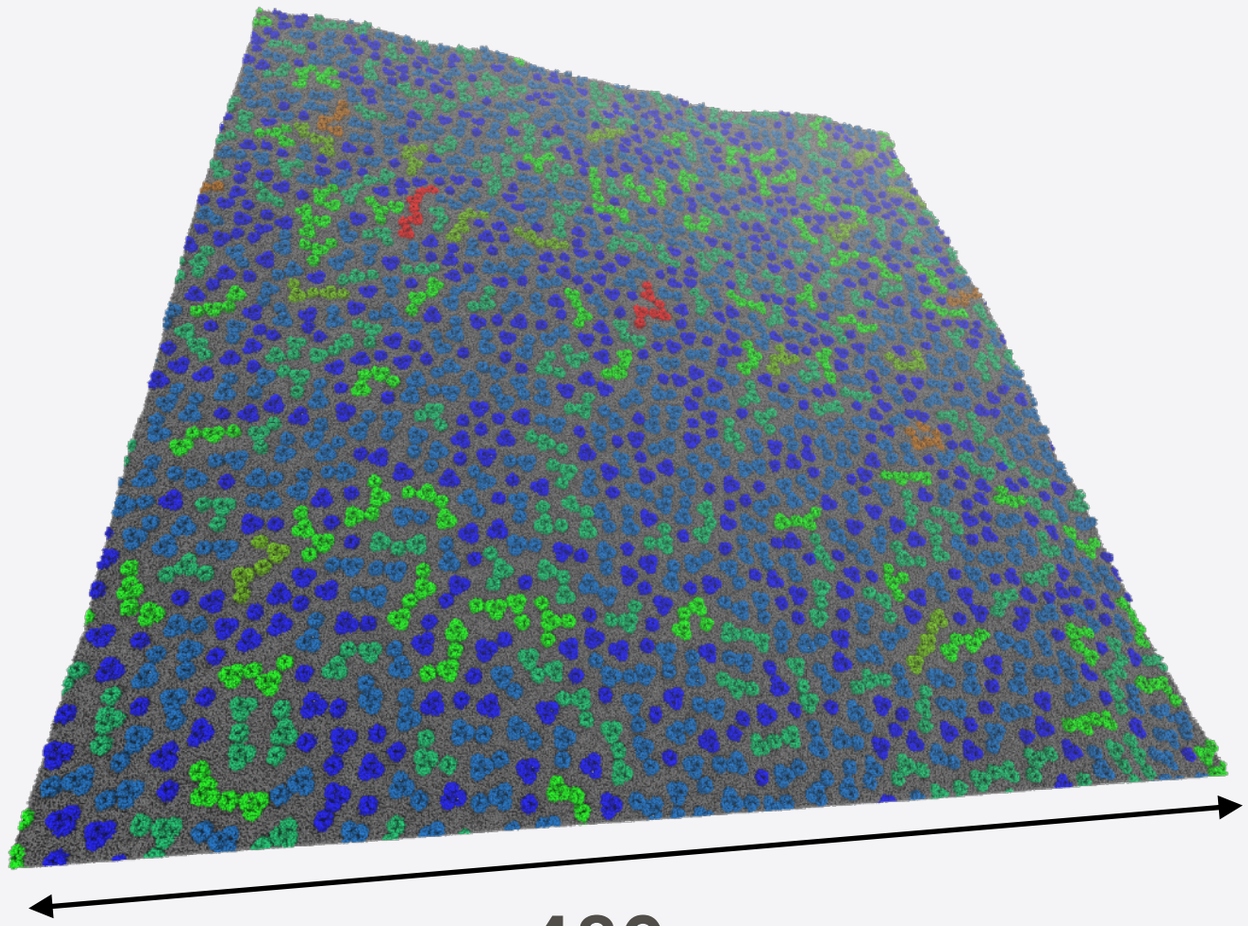
last frame -> file size: 10-50Mb

Zenodo and link in SI

Chavent, *et al.* , *Structure*, ***2018***

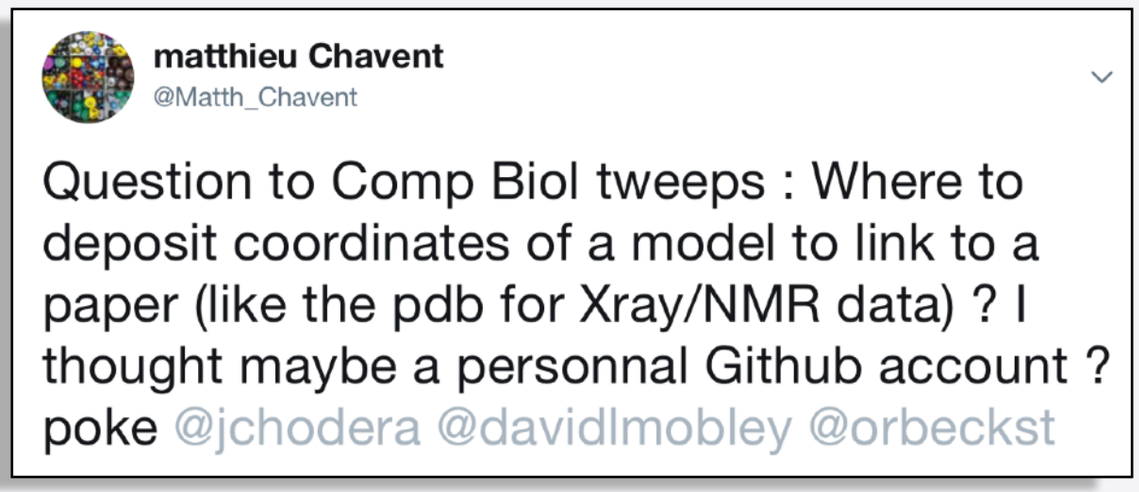
mix of BtuB and OmpF proteins

> 2000 proteins
~22M particles (with water)
1.5TB of data (2.5µs trajectory)
1 GB for 1 snapshot

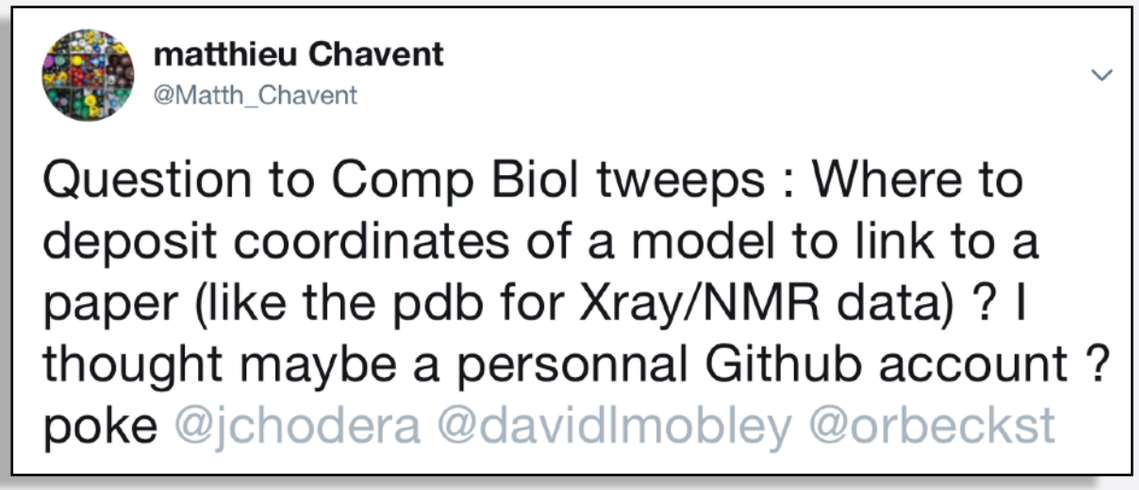


Chavent, *et al.* , *Nature Com.*, ***2018***

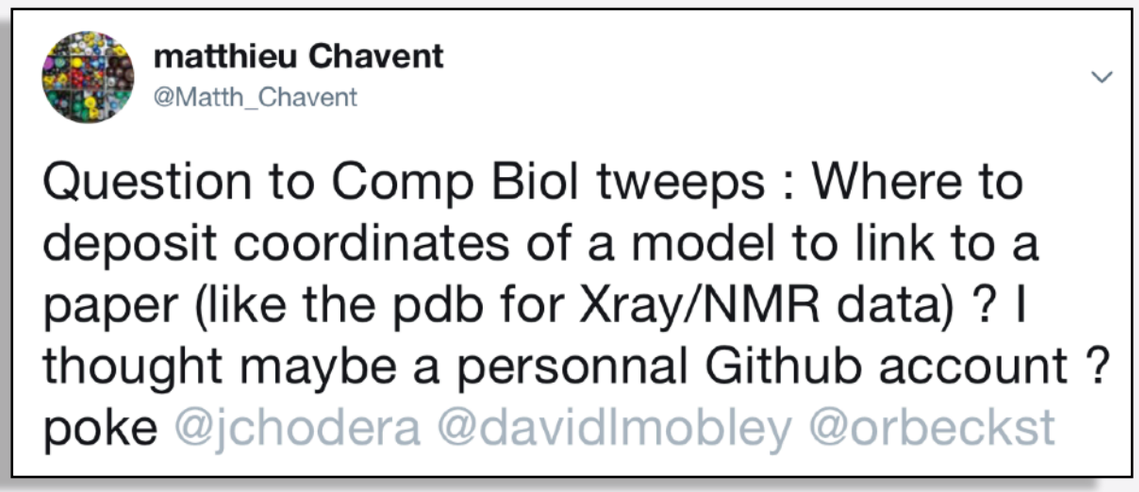
ASKING TWITTER



ASKING TWITTER



ASKING TWITTER



ASKING TWITTER

Andrew Marsh @marshgroup · 24 avr.

En réponse à @AnthonyC1Nash @Matth_Chavent et 3 autres

.@figshare @datadryad also good places. Librarians have told me the more the better to maximise exposure and @altmetric impact.

John Chodera @jchodera · 24 avr.

En réponse à @Matth_Chavent @davidlmobley @orbeckst

Your friendly journal editors have thought about this question! @Naturepub suggests @figshare and @datadryad for cases that are not clearly handled by standardized databases. nature.com/authors/polici…

Anthony Nash
@AnthonyC1Nash

Abonné

En réponse à @Matth_Chavent @jchodera et 2 autres

Following. Wouldn't be interested to know.

Traduire le Tweet

10:46 - 24 avr. 2018

matthieu Chavent
@Matth_Chavent

Question to Comp Biol tweeps : Where to deposit coordinates of a model to link to a paper (like the pdb for Xray/NMR data) ? I thought maybe a personnal Github account ? poke @jchodera @davidlmobley @orbeckst

ASKING TWITTER

Andrew Marsh @marshgroup · 24 avr.

En réponse à @AnthonyC1Nash @Matth_Chavent et 3 autres

.@figshare @datadryad also good places. Librarians have told me the more the better to maximise exposure and @altmetric impact.

John Chodera @jchodera · 24 avr.

En réponse à @Matth_Chavent @davidlmobley @orbeckst

Your friendly journal editors have thought about this question! @Naturepub suggests @figshare and @datadryad for cases that are not clearly handled by standardized databases. nature.com/authors/polici…

Anthony Nash
@AnthonyC1Nash

Abonné

En réponse à @Matth_Chavent @jchodera et 2 autres

Following. Wouldn't be interested to know.

Traduire le Tweet

10:46 - 24 avr. 2018

matthieu Chavent
@Matth_Chavent

Question to Comp Biol tweeps : Where to deposit coordinates of a model to link to a paper (like the pdb for Xray/NMR data) ? I thought maybe a personnal Github account ? poke @jchodera @davidlmobley @orbeckst

Mikael Trellet @mika_trellet · 24 avr.

En réponse à @Matth_Chavent @jchodera et 2 autres

Partial answer: if this is a model you got from integrative/hybrid methods have a look at pdb-dev.wwpdb.org

ASKING TWITTER

Andrew Marsh @marshgroup · 24 avr.

En réponse à @AnthonyC1Nash @Matth_Chavent et 3 autres

.@figshare @datadryad also good places. Librarians have told me the more the better to maximise exposure and @altmetric impact.

John Chodera @jchodera · 24 avr.

En réponse à @Matth_Chavent @davidlmobley @orbeckst

Your friendly journal editors have thought about this question! @Naturepub suggests @figshare and @datadryad for cases that are not clearly handled by standardized databases. nature.com/authors/polici…

Anthony Nash
@AnthonyC1Nash

Abonné

En réponse à @Matth_Chavent @jchodera et 2 autres

Following. Wouldn't be interested to know.

Traduire le Tweet

10:46 - 24 avr. 2018

matthieu Chavent
@Matth_Chavent

Question to Comp Biol tweeps : Where to deposit coordinates of a model to link to a paper (like the pdb for Xray/NMR data) ? I thought maybe a personnal Github account ? poke @jchodera @davidlmobley @orbeckst

David Mobley @davidlmobley · 24 avr.

I like to think of two categories: (a) broad dissemination/ease of use in the short term, and (b) long-term archival even after I have forgotten about it or am dead. For (a) former, GitHub, @figshare, @datadryad (haven't used it yet but it looks good) and others are good options.

Mikael Trellet @mika_trellet · 24 avr.

En réponse à @Matth_Chavent @jchodera et 2 autres

Partial answer: if this is a model you got from integrative/hybrid methods have a look at pdb-dev.wwpdb.org

ASKING TWITTER

Andrew Marsh @marshgroup · 24 avr.

En réponse à @AnthonyC1Nash @Matth_Chavent et 3 autres

.@figshare @datadryad also good places. Librarians have told me the more the better to maximise exposure and @altmetric impact.

John Chodera @jchodera · 24 avr.

En réponse à @Matth_Chavent @davidlmobley @orbeckst

Your friendly journal editors have thought about this question! @Naturepub suggests @figshare and @datadryad for cases that are not clearly handled by standardized databases. nature.com/authors/polici…

Anthony Nash
@AnthonyC1Nash

Abonné

En réponse à @Matth_Chavent @jchodera et 2 autres

Following. Wouldn't be interested to know.

Traduire le Tweet

10:46 - 24 avr. 2018

matthieu Chavent
@Matth_Chavent

Question to Comp Biol tweeps : Where to deposit coordinates of a model to link to a paper (like the pdb for Xray/NMR data) ? I thought maybe a personnal Github account ? poke @jchodera @davidlmobley @orbeckst

David Mobley @davidlmobley · 24 avr.

I like to think of two categories: (a) broad dissemination/ease of use in the short term, and (b) long-term archival even after I have forgotten about it or am dead. For (a) former, GitHub, @figshare, @datadryad (haven't used it yet but it looks good) and others are good options.

Pierre Poulain @pierrepo · 24 avr.

;-) no need to put the data elsewhere. @ZENODO_ORG is backed up in @CERN_FR cloud infrastructure and implements FAIR principles about.zenodo.org/principles/ Pretty safe!

Mikael Trellet @mika_trellet · 24 avr.

En réponse à @Matth_Chavent @jchodera et 2 autres

Partial answer: if this is a model you got from integrative/hybrid methods have a look at pdb-dev.wwpdb.org

ASKING TWITTER

Andrew Marsh @marshgroup · 24 avr.

En réponse à @AnthonyC1Nash @Matth_Chavent et 3 autres

.@figshare @datadryad also good places. Librarians have told me the more the better to maximise exposure and @altmetric impact.

John Chodera @jchodera · 24 avr.

En réponse à @Matth_Chavent @davidlmobley @orbeckst

Your friendly journal editors have thought about this question! @Naturepub suggests @figshare and @datadryad for cases that are not clearly handled by standardized databases. nature.com/authors/polici…

Anthony Nash
@AnthonyC1Nash

Abonné

En réponse à @Matth_Chavent @jchodera et 2 autres

Following. Wouldn't be interested to know.

Traduire le Tweet

10:46 - 24 avr. 2018

matthieu Chavent
@Matth_Chavent

Question to Comp Biol tweeps : Where to deposit coordinates of a model to link to a paper (like the pdb for Xray/NMR data) ? I thought maybe a personnal Github account ? poke @jchodera @davidlmobley @orbeckst

David Mobley @davidlmobley · 24 avr.

I like to think of two categories: (a) broad dissemination/ease of use in the short term, and (b) long-term archival even after I have forgotten about it or am dead. For (a) former, GitHub, @figshare, @datadryad (haven't used it yet but it looks good) and others are good options.

Pierre Poulain @pierrepo · 24 avr.

;-) no need to put the data elsewhere. @ZENODO_ORG is backed up in @CERN_FR cloud infrastructure and implements FAIR principles about.zenodo.org/principles/ Pretty safe!

Dr Lucy Forrest @LucyRForrest · 25 avr.

For the record, lately I also try to attach the structural model as supplementary data. As a backup.

Mikael Trellet @mika_trellet · 24 avr.

En réponse à @Matth_Chavent @jchodera et 2 autres

Partial answer: if this is a model you got from integrative/hybrid methods have a look at pdb-dev.wwpdb.org

ASKING TWITTER

Andrew Marsh @marshgroup · 24 avr.

En réponse à @AnthonyC1Nash @Matth_Chavent et 3 autres

.@figshare @datadryad also good places. Librarians have told me the more the better to maximise exposure and @altmetric impact.

John Chodera @jchodera · 24 avr.

En réponse à @Matth_Chavent @davidlmobley @orbeckst

Your friendly journal editors have thought about this question! @Naturepub suggests @figshare and @datadryad for cases that are not clearly handled by standardized databases. nature.com/authors/polici…

Anthony Nash
@AnthonyC1Nash

Abonné

En réponse à @Matth_Chavent @jchodera et 2 autres

Following. Wouldn't be interested to know.

Traduire le Tweet

10:46 - 24 avr. 2018

matthieu Chavent
@Matth_Chavent

Question to Comp Biol tweeps : Where to deposit coordinates of a model to link to a paper (like the pdb for Xray/NMR data) ? I thought maybe a personnal Github account ? poke @jchodera @davidlmobley @orbeckst

David Mobley @davidlmobley · 24 avr.

I like to think of two categories: (a) broad dissemination/ease of use in the short term, and (b) long-term archival even after I have forgotten about it or am dead. For (a) former, GitHub, @figshare, @datadryad (haven't used it yet but it looks good) and others are good options.

Pierre Poulain @pierrepo · 24 avr.

;-) no need to put the data elsewhere. @ZENODO_ORG is backed up in @CERN_FR cloud infrastructure and implements FAIR principles about.zenodo.org/principles/ Pretty safe!

Dr Lucy Forrest @LucyRForrest · 25 avr.

For the record, lately I also try to attach the structural model as supplementary data. As a backup.

Kresten Lindorff-Larsen @LindorffLarsen · 27 août

En réponse à @Matth_Chavent @jchodera et 2 autres

For single structures we have previously also just had coordinates as supporting information, e.g. here sciencedirect.com/science/articl… Obviously this isn't a good solution for trajectories

Mikael Trellet @mika_trellet · 24 avr.

En réponse à @Matth_Chavent @jchodera et 2 autres

Partial answer: if this is a model you got from integrative/hybrid methods have a look at pdb-dev.wwpdb.org

ASKING TWITTER

Andrew Marsh @marshgroup · 24 avr.

En réponse à @AnthonyC1Nash @Matth_Chavent et 3 autres

.@figshare @datadryad also good places. Librarians have told me the more the better to maximise exposure and @altmetric impact.

John Chodera @jchodera · 24 avr.

En réponse à @Matth_Chavent @davidlmobley @orbeckst

Your friendly journal editors have thought about this question! @Naturepub suggests @figshare and @datadryad for cases that are not clearly handled by standardized databases. nature.com/authors/polici…

Anthony Nash
@AnthonyC1Nash

Abonné

En réponse à @Matth_Chavent @jchodera et 2 autres

Following. Wouldn't be interested to know.

Traduire le Tweet

10:46 - 24 avr. 2018

matthieu Chavent
@Matth_Chavent

Question to Comp Biol tweeps : Where to deposit coordinates of a model to link to a paper (like the pdb for Xray/NMR data) ? I thought maybe a personnal Github account ? poke @jchodera @davidlmobley @orbeckst

Lucie Delemotte @DelemotteLab · 27 août

Same here, we use OSF. Individual files can't be more than 5GB, but there doesn't seem to be a limit on number of files!

David Mobley @davidlmobley · 24 avr.

I like to think of two categories: (a) broad dissemination/ease of use in the short term, and (b) long-term archival even after I have forgotten about it or am dead. For (a) former, GitHub, @figshare, @datadryad (haven't used it yet but it looks good) and others are good options.

Pierre Poulain @pierrepo · 24 avr.

;-) no need to put the data elsewhere. @ZENODO_ORG is backed up in @CERN_FR cloud infrastructure and implements FAIR principles about.zenodo.org/principles/ Pretty safe!

Dr Lucy Forrest @LucyRForrest · 25 avr.

For the record, lately I also try to attach the structural model as supplementary data. As a backup.

Kresten Lindorff-Larsen @LindorffLarsen · 27 août

En réponse à @Matth_Chavent @jchodera et 2 autres

For single structures we have previously also just had coordinates as supporting information, e.g. here sciencedirect.com/science/articl… Obviously this isn't a good solution for trajectories

Mikael Trellet @mika_trellet · 24 avr.

En réponse à @Matth_Chavent @jchodera et 2 autres

Partial answer: if this is a model you got from integrative/hybrid methods have a look at pdb-dev.wwpdb.org

ASKING TWITTER

matthieu Chavent
@Matth_Chavent

My brain:
Don’t say it
Don’t say it
Don’t say it
Don’t say it
Don’t say it
Don’t say it
Don’t say it
Don’t say it
Don’t say it

ASKING TWITTER

Andrew Marsh @marshgroup · 24 avr.

En réponse à @AnthonyC1Nash @Matth_Chavent et 3 autres

.@figshare @datadryad also good places. Librarians have told me the more the better to maximise exposure and @altmetric impact.

John Chodera @jchodera · 24 avr.

En réponse à @Matth_Chavent @davidlmobley @orbeckst

Your friendly journal editors have thought about this question! @Naturepub suggests @figshare and @datadryad for cases that are not clearly handled by standardized databases. nature.com/authors/polici…

Anthony Nash
@AnthonyC1Nash

Abonné

En réponse à @Matth_Chavent @jchodera et 2 autres

Following. Wouldn't be interested to know.

Traduire le Tweet

10:46 - 24 avr. 2018

matthieu Chavent
@Matth_Chavent

En réponse à @pierrepo @davidlmobley et 10 autres

By the way guys, as this dicussion is going very well, we may write a short "10 simple rules to store and share your comp chem data" or is it already done ?

Lucie Delemotte @DelemotteLab · 27 août

Same here, we use OSF. Individual files can't be more than 5GB, but there doesn't seem to be a limit on number of files!

David Mobley @davidlmobley · 24 avr.

I like to think of two categories: (a) broad dissemination/ease of use in the short term, and (b) long-term archival even after I have forgotten about it or am dead. For (a) former, GitHub, @figshare, @datadryad (haven't used it yet but it looks good) and others are good options.

Pierre Poulain @pierrepo · 24 avr.

;-) no need to put the data elsewhere. @ZENODO_ORG is backed up in @CERN_FR cloud infrastructure and implements FAIR principles about.zenodo.org/principles/ Pretty safe!

Dr Lucy Forrest @LucyRForrest · 25 avr.

For the record, lately I also try to attach the structural model as supplementary data. As a backup.

Kresten Lindorff-Larsen @LindorffLarsen · 27 août

En réponse à @Matth_Chavent @jchodera et 2 autres

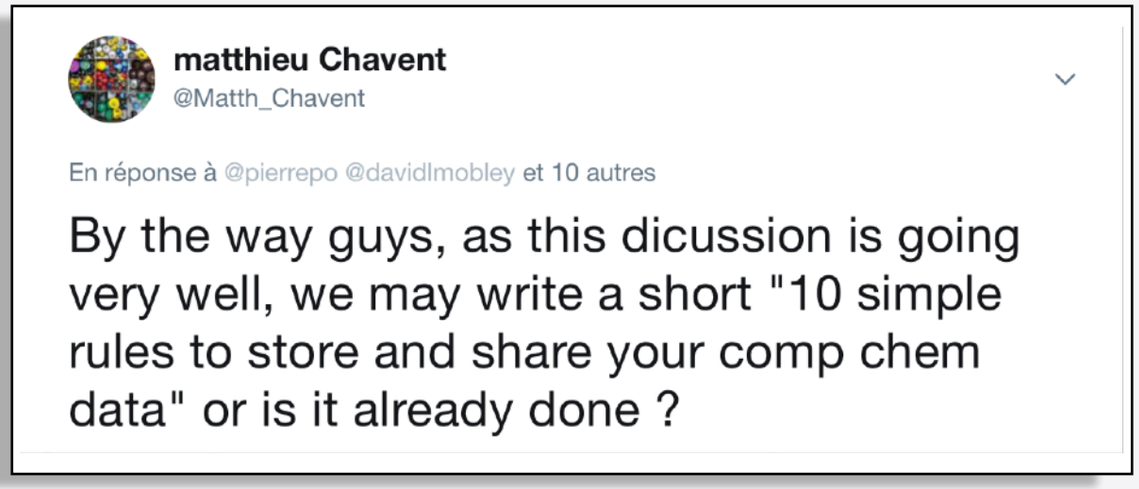
For single structures we have previously also just had coordinates as supporting information, e.g. here sciencedirect.com/science/articl… Obviously this isn't a good solution for trajectories

Mikael Trellet @mika_trellet · 24 avr.

En réponse à @Matth_Chavent @jchodera et 2 autres

Partial answer: if this is a model you got from integrative/hybrid methods have a look at pdb-dev.wwpdb.org

ASKING TWITTER



ASKING TWITTER

David Mobley @davidlmobley · 24 avr.

En réponse à @Matth_Chavent @pierrepo et 10 autres

How about a GitHub repo? Then eventually turn it into a "Best Practices" minireview for LiveCoMS, the Living Journal of Computational Molecular Science? A what to do and why to do it that way.

matthieu Chavent
@Matth_Chavent

En réponse à @pierrepo @davidlmobley et 10 autres

By the way guys, as this dicussion is going very well, we may write a short "10 simple rules to store and share your comp chem data" or is it already done ?

ASKING TWITTER

David Mobley @davidlmobley · 24 avr.

En réponse à @Matth_Chavent @pierrepo et 10 autres

How about a GitHub repo? Then eventually turn it into a "Best Practices" minireview for LiveCoMS, the Living Journal of Computational Molecular Science? A what to do and why to do it that way.

Traduire le Tweet

1 6 12

matthieu Chavent @Matth_Chavent · 24 avr.

This is a good idea and it would be very useful!

Traduire le Tweet

4 3

matthieu Chavent @Matth_Chavent · 24 avr.

I would be happy to contribute but not sure that I am the most legitimate for leading this project. If somebody want to push a bit further happy to help as much as I can

ASKING TWITTER

David Mobley @davidlmobley · 24 avr.

En réponse à @Matth_Chavent @pierrepo et 10 autres

How about a GitHub repo? Then eventually turn it into a "Best Practices" minireview for LiveCoMS, the Living Journal of Computational Molecular Science? A what to do and why to do it that way.

Traduire le Tweet

1 6 12

matthieu Chavent @Matth_Chavent · 24 avr.

This is a good idea and it would be very useful!

Traduire le Tweet

4 3

matthieu Chavent @Matth_Chavent · 24 avr.

I would be happy to contribute but not sure that I am the most legitimate for leading this project. If somebody want to push a bit further happy to help as much as I can

Traduire le Tweet

1 1

David Mobley @davidlmobley · 24 avr.

Key thing, IMO, is to find someone who's willing to be the driver (you, if you want) then get people who are "legitimate" to sign on. The largest part of getting something important done is, well, doing it. Legitimacy follows. :)

Traduire le Tweet

ASKING TWITTER

David Mobley @davidlmobley · 24 avr.

En réponse à @Matth_Chavent @pierrepo et 10 autres

How about a GitHub repo? Then eventually turn it into a "Best Practices" minireview for LiveCoMS, the Living Journal of Computational Molecular Science? A what to do and why to do it that way.

Traduire le Tweet

1 6 12

matthieu Chavent @Matth_Chavent · 24 avr.

This is a good idea and it would be very useful!

Traduire le Tweet

4 3

matthieu Chavent @Matth_Chavent · 24 avr.

I would be happy to contribute but not sure that I am the most legitimate for leading this project. If somebody want to push a bit further happy to help as much as I can

Traduire le Tweet

1 1

David Mobley @davidlmobley · 24 avr.

Key thing, IMO, is to find someone who's willing to be the driver (you, if you want) then get people who are "legitimate" to sign on. The largest part of getting something important done is, well, doing it. Legitimacy follows. :)

Traduire le Tweet

1 2

matthieu Chavent @Matth_Chavent · 24 avr.

Well, if I can find enough people to help I am happy to lead this. Please DM me if you want to participate to start this project

WORKSHOP FUNDING AND ORGANIZATION

Rossen Apostolov

bioexcel

Erik Lindahl

WORKSHOP FUNDING AND ORGANIZATION

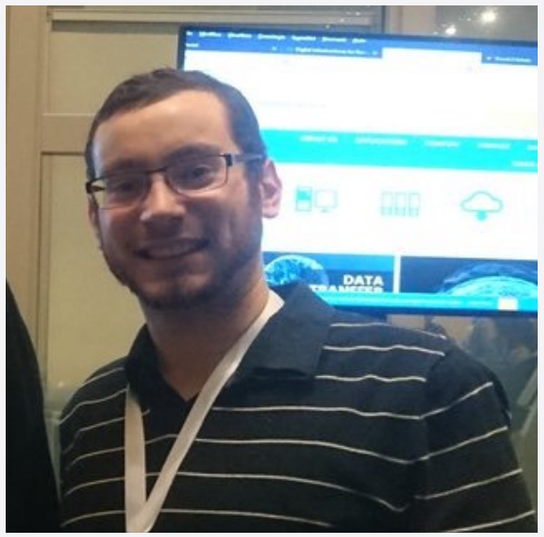
Rossen Apostolov

bioexcel

Erik Lindahl

Lucie Delemotte
(SciLifeLab)

Reba Howard
(SciLifeLab)



Mikael Trellet
(Utrecht Uni)

Daniel Smith
(MolSSI)

WHAT ARE THE GOALS OF THIS WORKSHOP ?

WHAT ARE THE GOALS OF THIS WORKSHOP ?

Discuss

(during the session and before/after)

WHAT ARE THE GOALS OF THIS WORKSHOP ?

Discuss

(during the session and before/after)

Find best pratices

(what are the solutions we already have?)

WHAT ARE THE GOALS OF THIS WORKSHOP ?

Discuss

(during the session and before/after)

Find best pratices

(what are the solutions we already have?)

Highlight issues - find new solutions

(what can we develop in the future?)

WHAT ARE THE GOALS OF THIS WORKSHOP ?

Discuss

(during the session and before/after)

Find best pratices

(what are the solutions we already have?)

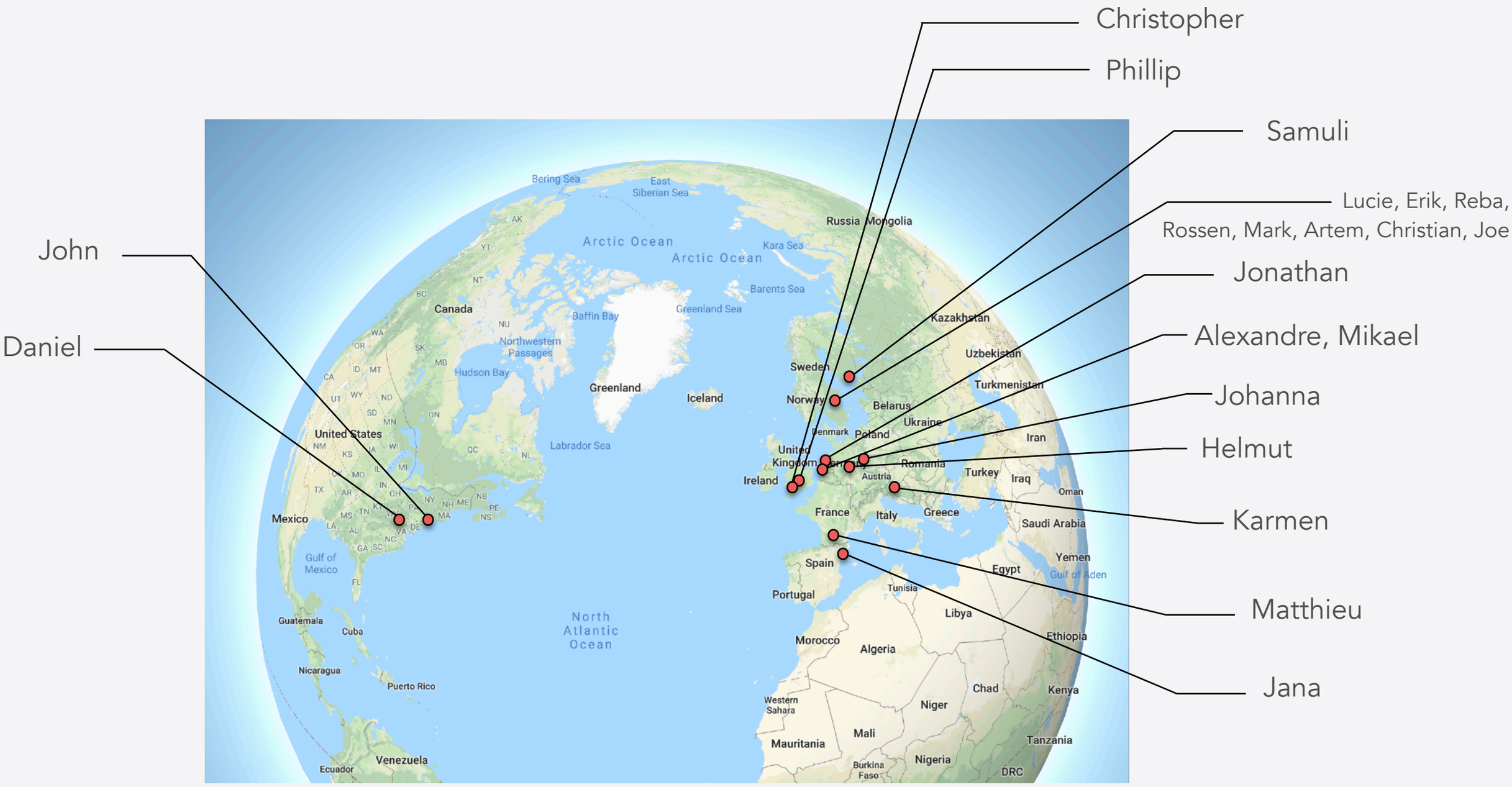
Highlight issues - find new solutions

(what can we develop in the future?)

Write - disseminate

(gather these pieces of information in a white paper)

WHO ARE WE ?


Christopher
Phillip
Samuli
Lucie, Erik, Reba,
Rossen, Mark, Artem, Christian, Joe
Jonathan
Alexandre, Mikael
Johanna
Helmut
Karmen
Matthieu
Jana
John
Daniel
Bering Sea
East Siberian Sea
Russia
Mongolia
Arctic Ocean
Kara Sea
Barents Sea
Canada
Baffin Bay
Greenland Sea
Northwestern Passages
Hudson Bay
Greenland
Iceland
Labrador Sea
United States
Mexico
Gulf of Mexico
North Atlantic Ocean
Kazakhstan
Uzbekistan
Turkmenistan
Sweden
Norway
Belarus
Ukraine
Poland
Denmark
Iran
Romania
Turkey
Iraq
Oman
Saudi Arabia
Yemen
Egypt
Gulf of Aden
Ireland
Austria
France
Italy
Greece
Spain
Portugal
Tunisia
Libya
Morocco
Algeria
Ethiopia
Chad
Kenya
Niger
Mali
Western Sahara
Mauritania
Nigeria
Burkina Faso
Tanzania
DRC
Guatemala
Cuba
Nicaragua
Puerto Rico
Venezuela
Ecuador

WHAT HAVE WE DONE ?

MemProt MD Search About

MemProtMD

GROMACS
FAST. FLEXIBLE. FREE.

GPCRMD

HOME ABOUT

EOSC-hub

MolSSI

mdbox

MDsrv

A fast & easy way of sharing MD simulations

bioexcel

BioSimSpace

Open Force Field Initiative

An open source, open science, and open data approach to better force fields

The NMRlipids project

HOW THIS WORKSHOP WILL WORK

Discuss

(no computing devices during the sessions
- except for taking notes or for tweeting)

Discuss

(no computing devices during the sessions
- except for taking notes or for tweeting)

Disseminate

(Twitter: #SDMS18 , recordings)

Discuss

(no computing devices during the sessions
- except for taking notes or for tweeting)

Disseminate

(Twitter: #SDMS18 , recordings)

Few more things …

(Mikeal and Reba)