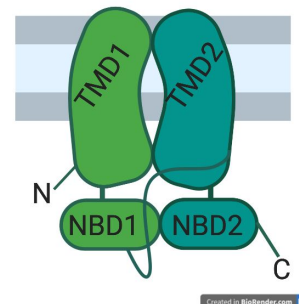


## Virtual Classroom: Model Building Tutorial. May 2020.

In this tutorial we will build a model of the **C**ystic **F**ibrosis **T**ransmembrane **C**onductance **R**egulator (CFTR). CFTR belongs to the ATP binding cassette (ABC) transporter superfamily but functions as an anion channel in epithelial cells. Mutations in CFTR that lead to channel dysfunction are causative for cystic fibrosis, making a high resolution structure of this protein highly sought for understanding the molecular basis of the disease and mechanisms of action of small molecule therapeutics.

The ABC transporter family is characterized by two nucleotide binding domains (NBDs) that bind and hydrolyze ATP. The NBDs are connected to transmembrane domains (TMDs). CFTR is a pseudo-dimer consisting of a single polypeptide. An overall topology is shown in the cartoon on the right.



Structures of isolated NBDs have been solved using X-ray crystallography and structures of the full length channel were only more recently solved using cryo-EM (the first published in 2017 by Liu et al. in Cell). We will use a published map, deposited in the EMDB from one of these structures:

Molecular structure of the ATP-bound, phosphorylated human CFTR; Zhe Zhang, Fangyu Liu, Jue Chen; *Proc Nat Acad Sci*. Dec 2018, 115(50)12757-12762; DOI: 10.1073/pnas.1815287115

We will start with a model that has most of the TMD already built, but is missing two helices and both NBDs. We will use Chimera to build the NBDs by assembling existing structures and fitting them roughly to the map. We will then use Coot to do manual real-space building and refinement of the NBDs and also build in the two missing TMD helices *de novo*.

**Please complete this pre-tutorial checklist prior to the tutorial, some of these steps may be time consuming if you do not already have software installed:**

- ☐ Download tutorial files (models and maps) here:  
[https://drive.google.com/open?id=1Wiwi\\_ugPOUQoR5uBLpTUHo5T-lw00oT9](https://drive.google.com/open?id=1Wiwi_ugPOUQoR5uBLpTUHo5T-lw00oT9)
- ☐ Install Chimera (downloads are here: <https://www.cgl.ucsf.edu/chimera/download.html>)
- ☐ Launch Chimera and open a model file (Menu: File>Open then find any model (.pdb) in the tutorial files you downloaded)
- ☐ Open a map file in Chimera (Menu: File>Open then find any map (.map) in the tutorial files)
- ☐ Install Coot (for Linux download from <https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/>. For Mac download from [http://scottlab.ucsc.edu/xtal/wiki/index.php/Installing\\_Coot\\_on\\_OS\\_X](http://scottlab.ucsc.edu/xtal/wiki/index.php/Installing_Coot_on_OS_X). For windows download from <https://bernhardcl.github.io/coot/wincoot-download.html>) We will be using Coot version 0.8 for this tutorial (version 0.9 is fine too if you already have it).
- ☐ Launch Coot and open a model file (Menu: File>Open Coordinates then find any model (.pdb) in the tutorial files)
- ☐ Open a map file in Coot (Menu: File>Open Map then find any map (.map) in the tutorial files).

If you can check all of these boxes, you're ready to start!










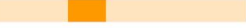




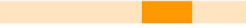





## Tutorial outline:

- Part 1 - Use Uniprot to become familiar with the sequence and make a plan
- Part 2 - Collect model files and fit to a map in Chimera
- Part 3 - Refine NBDs into map using Coot
- Part 4 - Build missing pieces (protein & ligands) using Coot
- Part 5 - Validation


## Part 1 - Use Uniprot to become familiar with the sequence and make a plan

We will start by using Uniprot to familiarize ourselves with the domain architecture and sequence of the protein: <https://www.uniprot.org/uniprot/P13569>

Importantly we can easily find domain boundaries:

Family & Domains <sup>i</sup>					
Domains and Repeats					
Feature key	Position(s)	Description	Actions	Graphical view	Length
Domain <sup>i</sup>	81 – 365	ABC transmembrane type-1 1  PROSITE-ProRule annotation 	 Add  BLAST		285
Domain <sup>i</sup>	423 – 646	ABC transporter 1  PROSITE-ProRule annotation 	 Add  BLAST		224
Domain <sup>i</sup>	859 – 1155	ABC transmembrane type-1 2  PROSITE-ProRule annotation 	 Add  BLAST		297
Domain <sup>i</sup>	1210 – 1443	ABC transporter 2  PROSITE-ProRule annotation 	 Add  BLAST		234

And existing structures that are available in public repositories:

Structure <sup>i</sup>					
					
PDB Entry	Method	Resolution	Chain	Positions	Links
5D3F	X-ray	2.74 Å	C	747-774	PDBsum PDBe RCSB PDB PDBj PDBsum
5TF7	X-ray	1.93 Å	A	387-646	PDBe RCSB PDB PDBj PDBsum
5TF8	X-ray	1.86 Å	A	387-646	PDBe RCSB PDB PDBj PDBsum
5TFA	X-ray	1.87 Å	A	387-646	PDBe RCSB PDB PDBj PDBsum
5TFB	X-ray	1.87 Å	A	387-646	PDBe RCSB PDB PDBj PDBsum

Let's identify starting models for the NBDs:

For NBD1 we'll use 5TF7 (residues 387-646) and for NBD2 we'll use 3GD7 (residues 1193-1427).

We can download these from the PDB or fetch them into Chimera directly. We'll use the PDB search to find an NBD dimer as a useful template (I found 1L2T) and download our map from the EMDataResource (<https://www.emdataresource.org/> search ID 9230)

What we've collected so far: A map (emd\_9230.map), a model of the TMDs (CFTR\_TMDs.pdb), a model of NBD1 (5TF7.pdb), a model of NBD2 (3GD7.pdb) and an ATP bound NBD dimer as a template (1L2T.pdb). Now we assemble:

## Part 2 - Collect model files and fit to a map in Chimera

For all of Part 2, we will use Chimera. Chimera has a user guide here:

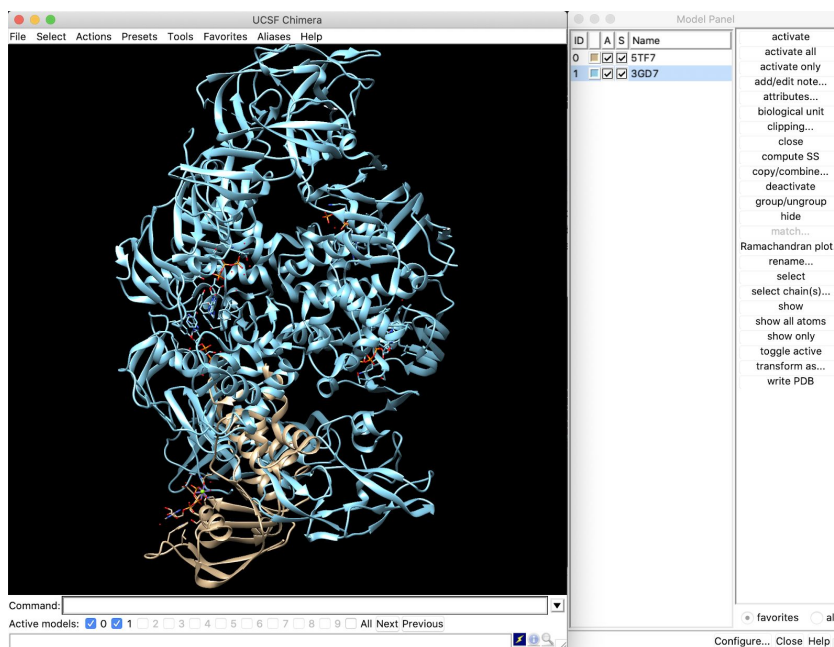
<https://www.cgl.ucsf.edu/chimera/docs/UsersGuide/index.html>

In this tutorial we will only touch on a few useful functions - there are many we will not cover! We will make use of GUI buttons and the Chimera command line.

**2.1** Open the two NBD models, either using the File>Open menu or fetching by ID from the PDB:

- 5TF7 - This model is an NBD1 monomer with Mg-ATP and solvent
- 3GD7 - This model is an NBD2 fusion protein tetramer.

For NBD2 (3GD7) we don't care about the fusion protein or the crystallographic copies that make the tetrameric model, we only need an NBD monomer (residues 1193-1427). We can generate the monomer by opening the PDB file in a text editor and manually deleting the unneeded parts of the model, or we can do this using selections in Chimera. We will use command line selections. Open the command line from: Tools>General Controls>Command Line.



Also, let's open the Model Panel, this will help us keep track of the various objects we have open. Open it from Tools>General Controls>Model Panel. In the model panel you can easily see the ID for each model (here, 0 & 1) and options to **Show** and **Activate** each model individually.

**2.2** For making selections in Chimera, the basic syntax is here:

[https://www.cgl.ucsf.edu/chimera/docs/UsersGuide/midas/frameatom\\_spec.html](https://www.cgl.ucsf.edu/chimera/docs/UsersGuide/midas/frameatom_spec.html)

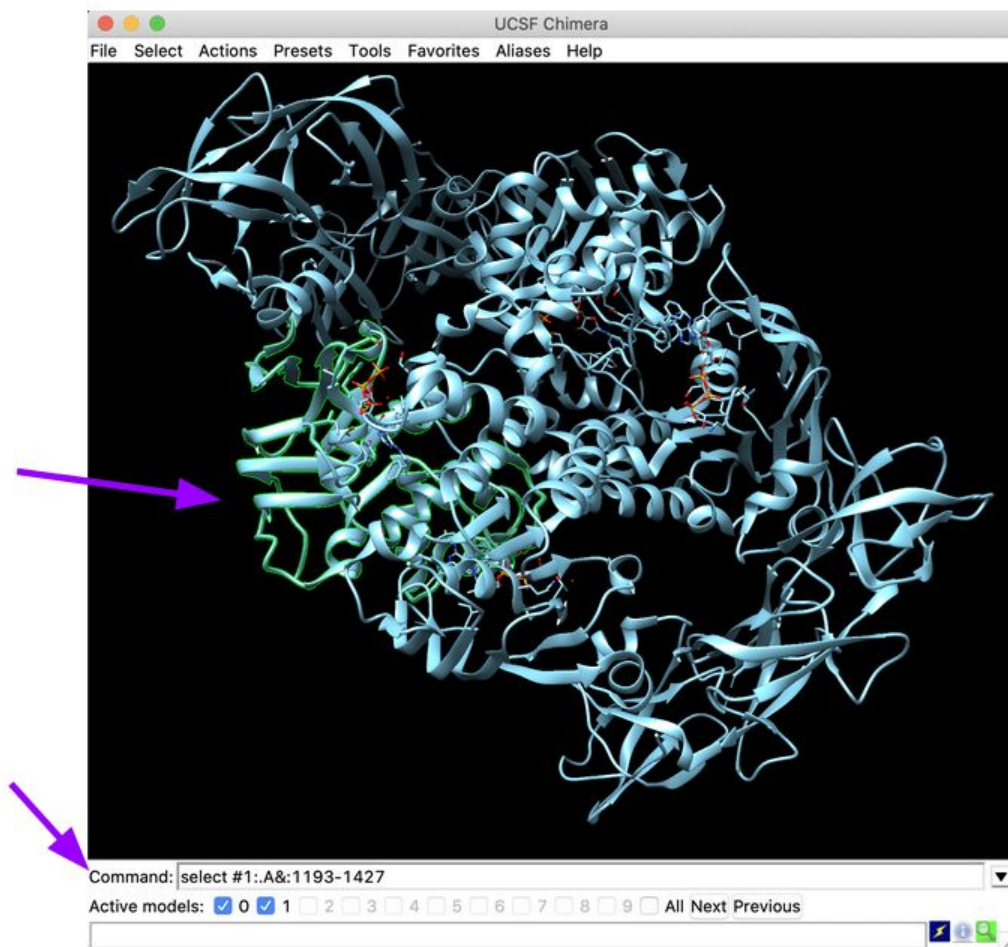
Most important for now, we can select a #(model), :.(ChainID), or :(residue number).

In this case to choose residues 1193-1427 of chain A in 3GD7 we type:

`select #1:.A&:1193-1427`

(make sure 3GD7 is model #1 in your session using the Model Panel. If it's not, choose the correct model number for your session).

Selected residues are highlighted with a green outline. We can then save just those residues to a new file.





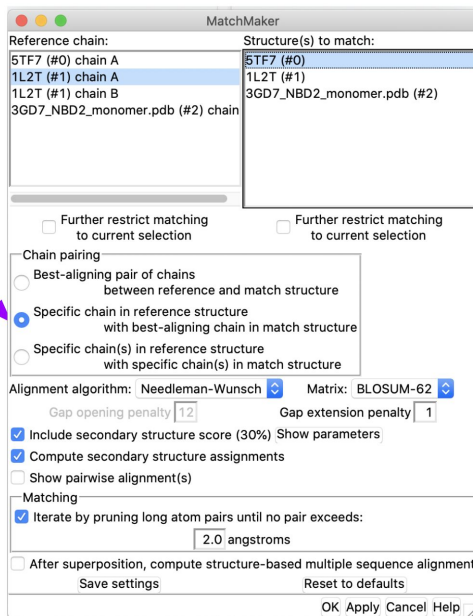
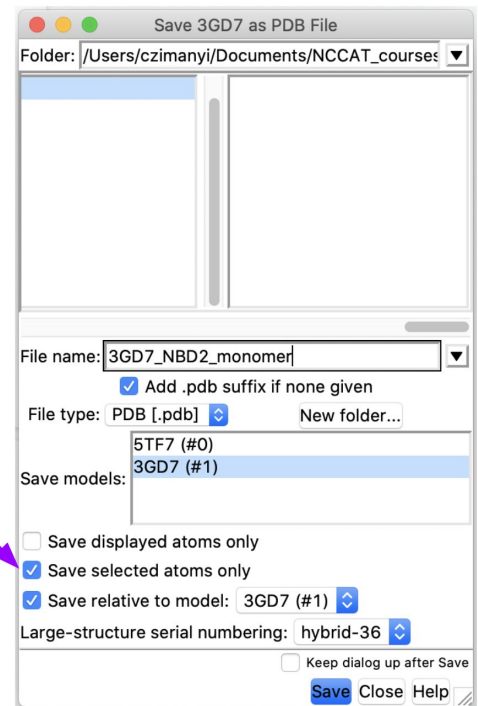
Go to Actions>Write PDB. Make sure to save selected atoms only. Give it a useful name and save it in a useful place. Now we can close the original model and open our newly saved monomer. You'll notice we have lost the Mg-ATP that was in the original model because we didn't select it. No worries, we'll add it back later by hand in Coot.

*Chimera tip: For interactive selections you can also use: Control + left mouse to click on what you want to select or you can use the Select menu options.*

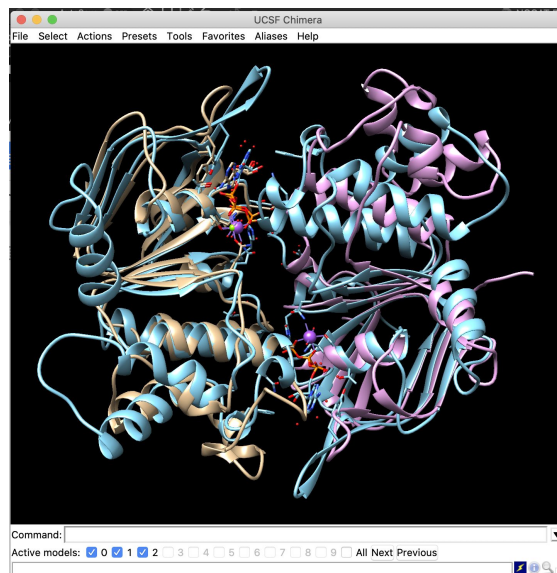
**2.3** Our two NBDs are in random positions based on our starting models. To make fitting them to the map easier, let's assemble them as a dimer, by aligning to a known ATP bound NBD dimer structure. Open 1L2T. This is a heterodimer with two chains, A & B. We will align NBD1 (5TF7) to chain A, and NBD2 (3GD7) to chain B.

We will use the MatchMaker in Chimera to do the alignment. Choose Tools>Structure Comparison>MatchMaker

Choose to pair a "specific chain in reference structure" so that you can choose chain A as the reference and choose 5TF7 as the structure to match. All other defaults should be fine. Then in a second run of MatchMaker, match 3GD7 monomer model to chain B of 1L2T.



Now I have an approximate dimer overlaid on my template that looks like this:



We can close the template from the Model Panel, (we're done with it) and next we can fit this approximate dimer into a map. First let's save our progress. You can save a Chimera session from the File menu.

**2.4** Open the map, either using the menu: File>Open or File>Fetch by ID and choose 9230 from the EMDB. Let's open the TMD model now as well (CFTR\_TMDs.pdb). This should already be well aligned in the map.

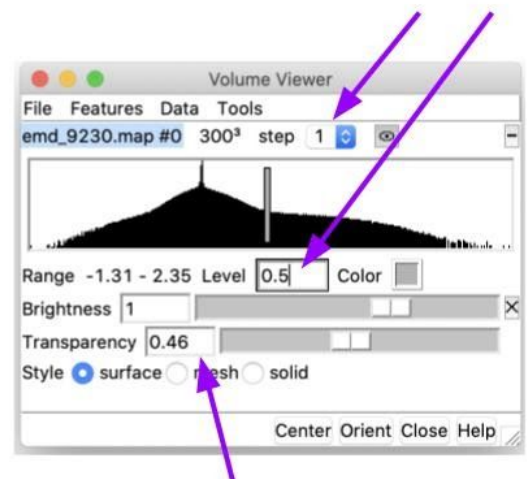
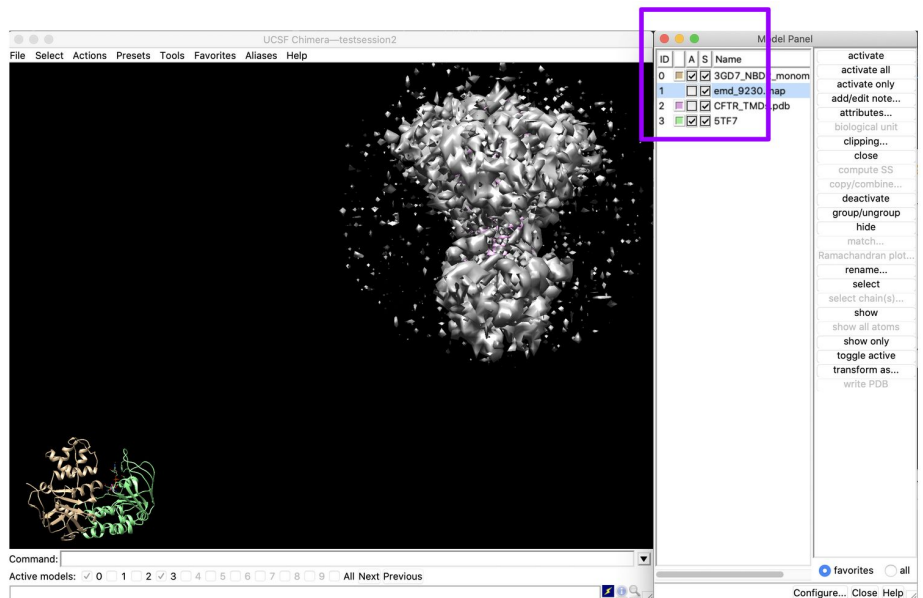
You'll notice the map and TMD model is far away in real space from the NBD models, so let's move the NBDs towards the map with the help of the Model Panel. (It's best practice to move models relative to a map).

Uncheck the A column for the map and we can now move the models using middle mouse (or option + left mouse) and drag them towards the map. You may need a few rounds of (de)activating the map to move in all three dimensions. Once you have the NBDs in the right neighborhood, let's make the map easier to fit into.

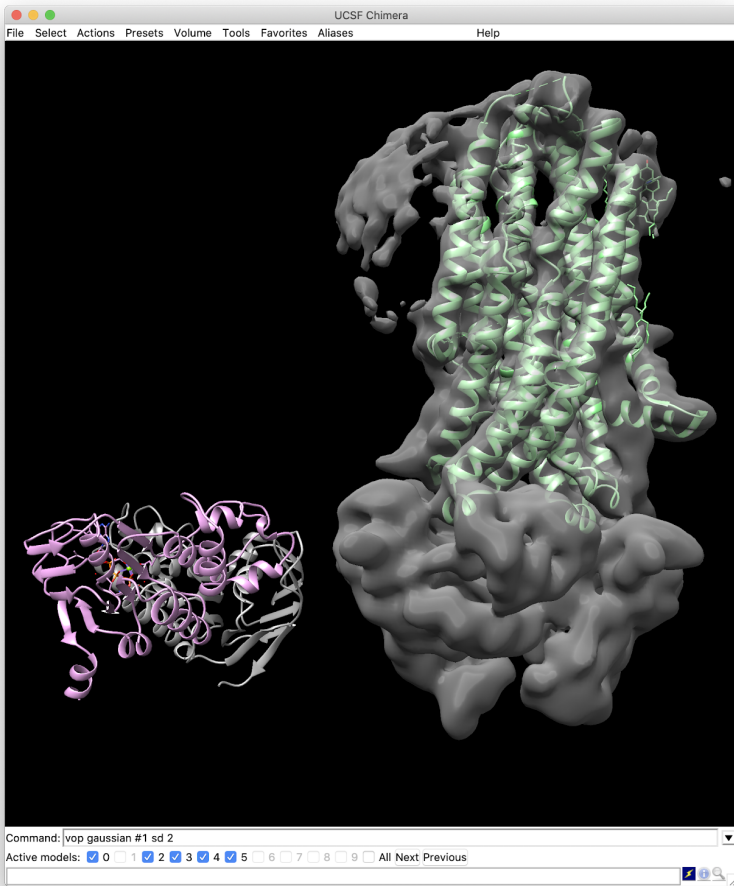
First, we can change the representation: Use the Volume viewer (Tools>Volume Data>Volume viewer) to change the contour to the **level** recommended by the depositors (here, 0.5). We can also use the Features>Brightness and **Transparency**: to adjust the transparency. Also, set the **Step** to 1.

This is a good place to start, but there are many other useful manipulations of volume representation that may be useful for you. See some here:

<https://www.cgl.ucsf.edu/chimera/data/tutorials/maps08/volume-basics.html>

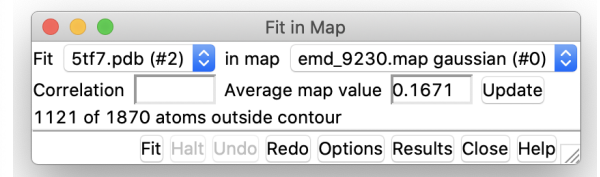


We want to start with a basic fit and in high resolution maps, it can be hard to see what's what. To see larger features more prominently we can smooth our map using a gaussian blurring feature in Chimera with the command **>vop gaussian #1 sd 2**



This will create a second map as a new object and this is the map we'll fit into.

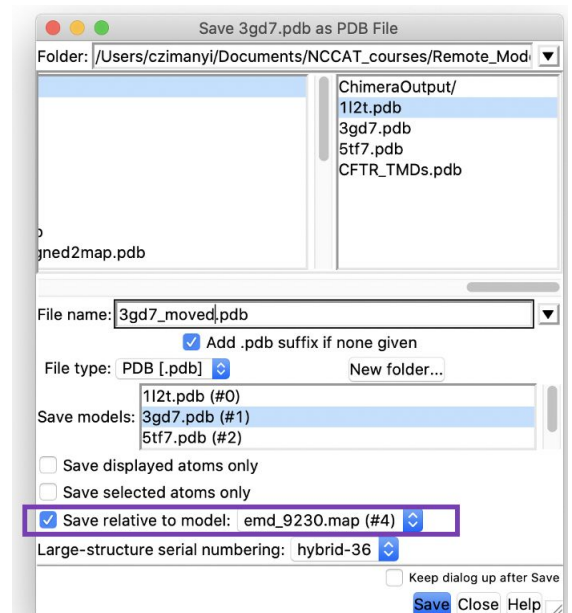
Now we will fit the NBDs into the map. Chimera does a good job of calculating a fit of a model into a map, but the search is local, so the model needs to be well positioned to start. We will put the NBDs into the map as best as we can by hand, by translating and rotating our approximate dimer. Once we're inside the density we can use the Fit in Map tool in Volume>Fit in Map



We can fit each chain independently or both together by making a selection with both chains and choosing Fit: selection. Here we just want a global fit. We will need to make some adjustments that aren't rigid body movements of the full chain, but we'll do these in Coot.

Once you have both NBD's positioned, we want to save them so that they are in the right place relative to the map. In the Save PDB... dialog choose to save the model relative to the map object. Although we can merge the coordinates into one file in Chimera, in this example we will do it in Coot. Save both NBDs and we're done with Chimera.

(A quick note - to cheat and/or save time, you can always check your progress against the full published model: 6MSM.pdb).



## Part 3 - Refine NBDs into map using Coot

For parts 3 & 4 we will use Coot. Coot has a user guide here:

<https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/web/docs/coot.html>

If you're opening Coot from the command line (in Linux or MacOS) open it from the directory containing the files you want to work with (this isn't strictly necessary but makes life easier).

**3.1** Open the three model files that were aligned to the map in Chimera (File>Open Coordinates).

Here are some useful navigation tips for Coot:

Manipulate:

- Ctrl-click-drag (or middle mouse) to translate view
- Click-drag to rotate view
- Right-click drag (or right mouse) to zoom in/out
- Middle-click on atom to center
- Shift-click (or double click) to label atom/residue

View:

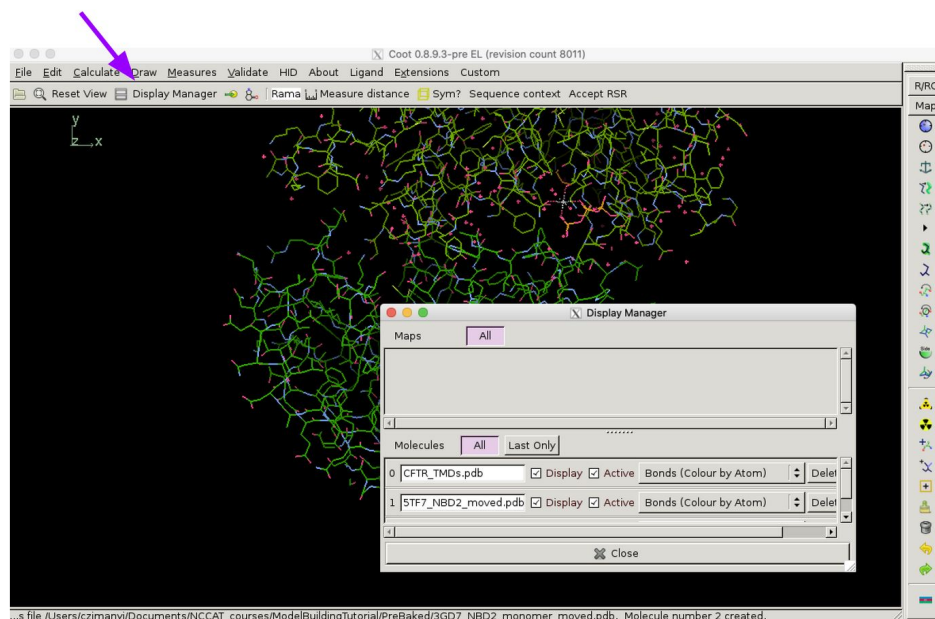
- Ctrl-right-click-drag up/down to translate slab in/out of plane
- Ctrl-right-click-drag left/right to change thickness of slab
- +/- keys will adjust the map level
- [ and ] keys will toggle the model display type

Similar to the model panel in Chimera, Coot has a Display Manager. You can find the Display Manager in Draw>Display Manager or the button on the top of the display.

**3.2** Merge the three models into a single file using Edit>Merge Molecules.

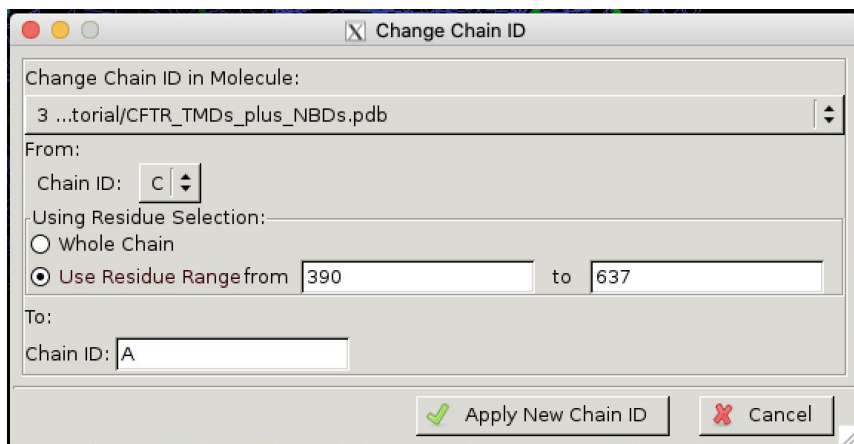
Choose all three models and merge into any of one the existing models. Save the file you merged into with a new and meaningful name. Then we can delete the unmerged models.

You'll notice, if you label on the atoms in the different domains, that the merged chains were given unique ChainIDs. You can also visualize this using the sequence viewer in Draw>Sequence View.





**3.3** We want everything to be part of Chain A so we will change the Chain IDs using Edit>Change Chain ID. Because we want to change Chains C & D to an ID that already exists, we have to ensure there are no residue number overlaps. For NBD2 this is not a problem, but the NBD1 model has ligands and waters with overlapping residue numbers so we have to specify the exact range of residues to change (here we'll only reassign protein atoms). [NBD1 is 390-637 and if you've forgotten that range, the sequence viewer can help you find it quickly.]



Close and reopen the sequence viewer to make sure this worked as intended. You'll notice chain C still exists with a number of heteroatoms. 701 & 702 are Mg-ATP. We want to keep this for now. Atoms 801-968 are waters. We don't want to include these in our model, so let's remove them. This time, let's do it manually in a text editor. Save your current file with a useful name (I called mine CFTR\_TMDs\_plus\_NBDs\_chainA.pdb). Then open the .pdb file with your favorite text editor.

Waters have the three letter code HOH and are labeled as HETATMs. Find them, delete them, and save the file with a new name and the extension.pdb (I called mine CFTR\_TMDs\_plus\_NBDs\_noHOH.pdb).

CFTR_TMDs_plus_NBDs_chainA.pdb											
ATOM	8784	O	UNK	B	17	132.963	159.236	133.243	1.00	208.79	O
ATOM	8785	CB	UNK	B	17	136.237	158.788	131.808	1.00	206.29	C
TER	8786		UNK	B	17						
HETATM	8787	Mg	Mg	C	701	162.161	160.510	103.475	1.00	16.77	Mg
HETATM	8788	PG	ATP	C	702	162.686	157.843	101.706	1.00	21.31	P
HETATM	8789	O1G	ATP	C	702	162.305	158.087	100.266	1.00	22.79	O
HETATM	8790	O2G	ATP	C	702	162.541	156.363	102.037	1.00	20.59	O
HETATM	8791	O3G	ATP	C	702	161.867	158.732	102.623	1.00	20.01	O
HETATM	8792	PB	ATP	C	702	164.988	159.619	101.992	1.00	18.06	P
HETATM	8793	O1B	ATP	C	702	165.430	160.070	100.624	1.00	18.62	O
HETATM	8794	O2B	ATP	C	702	164.107	160.646	102.664	1.00	14.31	O
HETATM	8795	O3B	ATP	C	702	164.279	158.125	101.902	1.00	19.12	O
HETATM	8796	PA	ATP	C	702	166.733	159.834	104.337	1.00	20.32	P
HETATM	8797	O1A	ATP	C	702	167.219	161.239	104.223	1.00	17.23	O
HETATM	8798	O2A	ATP	C	702	165.676	159.656	105.361	1.00	19.78	O
HETATM	8799	O3A	ATP	C	702	166.350	159.325	102.853	1.00	19.12	O
HETATM	8800	O5'	ATP	C	702	167.943	158.822	104.739	1.00	23.07	O
HETATM	8801	C5'	ATP	C	702	169.254	159.046	104.246	1.00	20.78	C
HETATM	8802	C4'	ATP	C	702	170.175	158.416	105.280	1.00	25.93	C
HETATM	8803	O4'	ATP	C	702	170.137	159.145	106.372	1.00	29.90	O
HETATM	8804	C3'	ATP	C	702	169.651	157.024	105.688	1.00	29.96	C
HETATM	8805	O3'	ATP	C	702	170.713	156.151	105.799	1.00	36.69	O
HETATM	8806	C2'	ATP	C	702	169.015	157.228	107.056	1.00	29.88	C

**3.4** Open this file in Coot. If it looks good, we can close the model with waters. Then open the map (I find this to be somewhat slow on my MacBook). You can change the map display using Edit>Map parameters and also the properties button in the Display Manager. Set the level to the suggested level of 0.5 in the properties window.

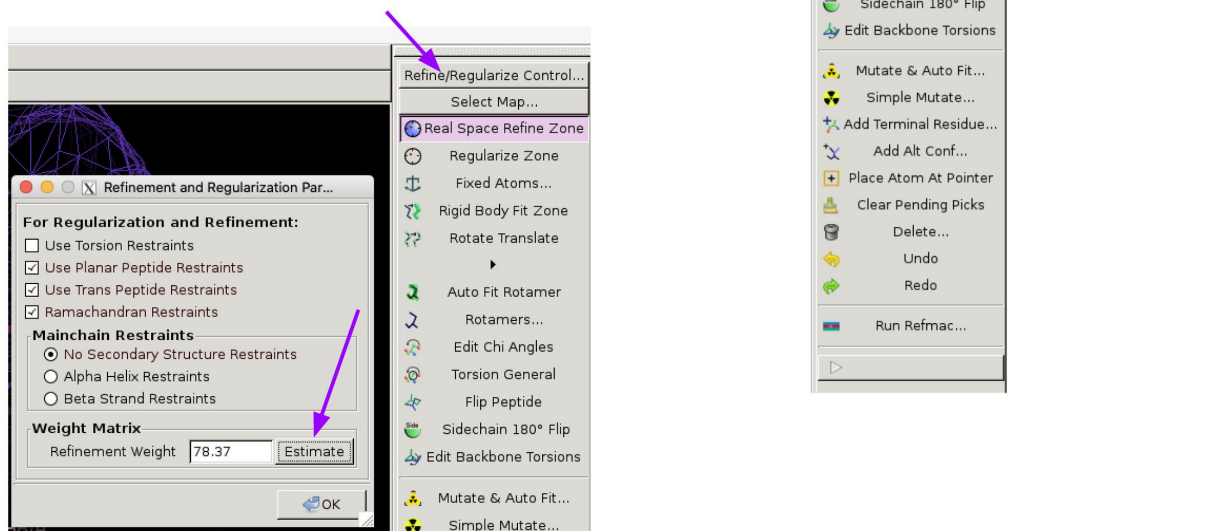
**3.5** Now we can use the modeling tools. It can help to toggle between representation types in the Display Manager (or with the [ ] keys) to get a sense of what will need work. For example I can see that the beta-strand containing Asn1262 is placed in the density fairly well, but Ser1276 needs to be repositioned substantially.



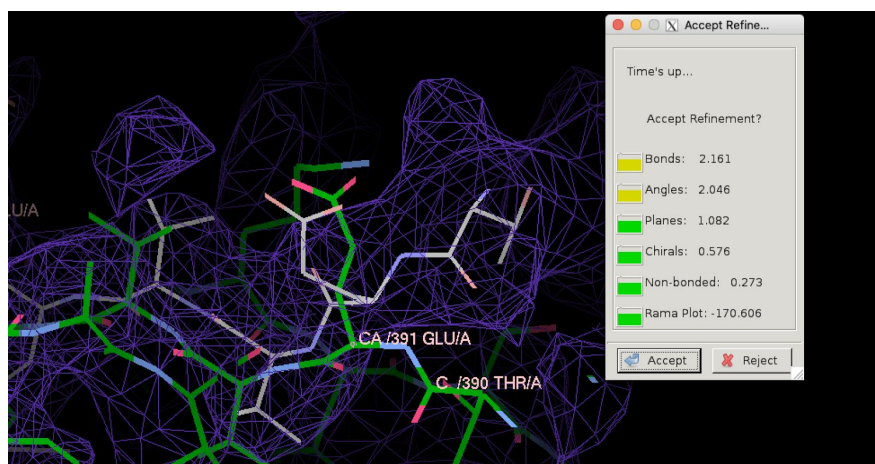
Your model may differ slightly based on your fit from Chimera, so I won't outline detailed steps, we will just work through our own models using the different refinement modes as needed.

Different refinement modes are useful for different tasks. The two major modes are rigid body and dynamic refinement.

For dynamic refinement guided by map restraints (Real Space Refine Zone), we want to set the refinement weight using the Refine/Regularize control parameters. Click “Estimate” to give a good starting refinement weight. This usually works well. If you need to adjust that weight, increasing it puts more weight on the map and decreasing it puts more weight on geometry.



I suggest using Real Space Refine Zone with larger selections of atoms (10's of residues) initially for things that are close but not exactly in the density. This will be a little slow but works. Then you can use shorter (1-4 residue) selections to spot clean. Dynamic refinement will be guided by color coded validation parameters. It isn't vital for these to be perfect at every step, especially at early stages of building but they are a helpful guide. If things go awry you can reject the refinement.





For bigger issues and pieces far from fitting into the density, the rigid body options will be more useful. You can translate and rotate any selected range of atoms. In some cases where there is a large rearrangement that can't be adjusted with rigid body motion, it can be useful to just delete residues and build them back with the Add Terminal Residue button.

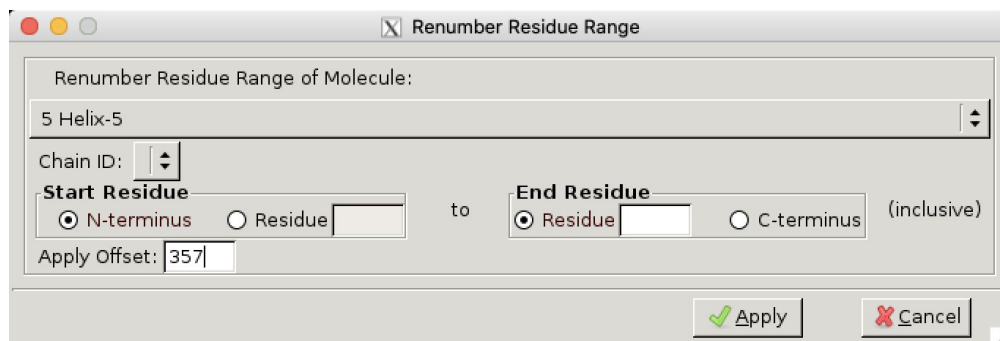
Spend some time playing with the refinement modes to get used to them.

**Note - Coot does not autosave work, so save your progress often!**

## Part 4 - Build missing pieces (protein & ligands) using Coot

Not all starting models will be complete. Here we are missing two long helices in the TMD.

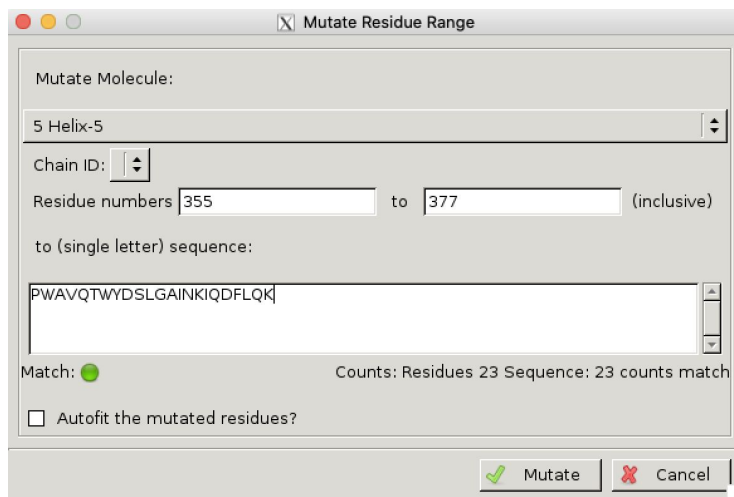
**4.1** Find the density with the missing helices by navigating around the map (and if you can't find them, they are adjacent to residues 354 and 1144. Draw>Go To Atom... is another way to navigate quickly). Find the unmodelled density following residue 354. Navigate so that the center pointer is near the middle of the missing helix. Choose Calculate>Other Modelling Tools from the menu and click on "Place Helix Here". Coot will auto generate a helix into the density. This is a new molecule (which you can see if you look in the Display Manager). We can attach it properly by deleting any overlapping residues, and changing the residue numbering (Edit>Renummer residues). We know the first new residue should be 355 so apply the appropriate offset.



You can then merge the helix into the same molecule and change chain ID's as you did in the earlier steps to combine this into a single model file. Now if you Real Space Refine across the junction, it should automatically form a bond.

**4.2** Coot made a poly-alanine helix so now we need to change the amino acids to have the correct side chains. For this we need to know our sequence. You can find it in the hCFTRsequence.fasta file. [As a cheat sheet, the missing residues are highlighted in the sequence below]. You can change them one at a time using the mutate buttons in the menu, or use the Calculate>Mutate Residue Range and paste the missing range.





Once you're satisfied with this helix, you can try again by building, merging, attaching and mutating a helix extending from residue 1144.

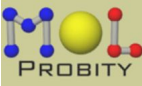
**4.3** Our final step is to add back the ATP molecule we lost from NBD2. You should find empty density for it near Tyr1219 (It may be difficult to recognize. If you can't see where it should go, you can adjust the map level or get extra help by using the initial model 3GD7 or the final published model 6MSM to see where it should be). Common ligands can be added in Coot using File>Get Monomer. Coot uses the Refmac monomer library that contains most common ligands (<http://www.ccp4.ac.uk/html/refmac5/dictionary/list-of-ligands.html>).

Because ATP is common, you can just type in the three letter code: ATP. It will be placed at the cursor center and then you can refine it into the density just as you did for protein atoms. It will have hydrogen atoms, but you can delete these easily using the delete button. By default, ligands added this way are placed into a separate model, so you'll have to merge it with the protein if you want to have them in the same file (but by now, you're a pro at this!).


## Part 5 - Validation

How do I know when my model is done? After rounds of building and refinement (which we did not cover here), I should satisfy some cutoff of validation metrics and decide that I have done my best, for myself and future users of my model. Validation encompasses many metrics. Two major aspects are model quality & fit to map. We will only look at model quality for now.

Ideally, you would visually check your model residue by residue. There are some quick ways to check for big issues within Coot first, using the Validate menu. (Ramachandran, Geometry analysis, rotamer analysis are useful). Once you're done in Coot you can use validation servers to check your work. Molprobity is great one: <http://molprobity.biochem.duke.edu/>



Uploaded PDB file as  
**CFTR\_TMDs\_plus\_NBDs\_noHOH-coot-0.pdb**

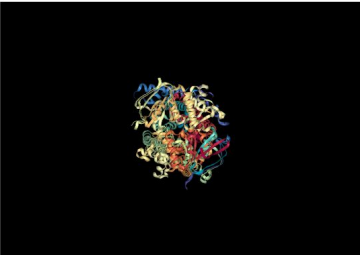


Duke Biochemistry  
Duke University School of Medicine


Your file from local disk was uploaded as CFTR\_TMDs\_plus\_NBDs\_noHOH-coot-0.pdb.

- This compound is identified as **PHOSPHORYLATED, ATP-BOUND HUMAN CYSTIC FIBROSIS TRANSMEMBRAN CONDUCTANCE REGULATOR (CFTR)**
- This structure was solved by ELECTRON MICROSCOPY.
- This structure was solved at 3.20 Å resolution.
- 2 chain(s) is/are present [2 unique chain(s)]
- A total of 1103 residues are present.
- Protein mainchain and sidechains are present.
- No explicit hydrogen atoms are included.
- 8 hetero group(s) is/are present.
- 0 PDBv2.3 atoms were found. Proceeding assuming PDBv3 formatted file.


[Continue >](#)



About MolProbity | Website for the Richardson Lab | Using ecloud x-H | Internal reference



Viewing CFTR\_TMDs\_plus\_NBDs\_noHOH-coot-0FH-multi.table



Duke Biochemistry  
Duke University School of Medicine

When finished, you should [close this window](#) . Hint: Use File | Save As... to save a copy of this page.

All-Atom Contacts	Clashscore, all atoms:	19.54	95 <sup>th</sup> percentile* (N=41, 2.95Å - 9999Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	32	3.43% Goal: <0.3%
	Favored rotamers	847	90.78% Goal: >98%
	Ramachandran outliers	0	0.00% Goal: <0.05%
	Ramachandran favored	1037	96.55% Goal: >98%
	MolProbity score <sup>^</sup>	2.42	97 <sup>th</sup> percentile* (N=1544, 3.20Å ± 0.25Å)
	Cβ deviations >0.25Å	19	1.84% Goal: 0
	Bad bonds:	42 / 8788	0.48% Goal: 0%
	Bad angles:	54 / 11886	0.45% Goal: <0.1%
Peptide Omegas	Cis Prolines:	2 / 28	7.14% Expected: ≤1 per chain, or ≤5%
Low-resolution Criteria	CaBLAM outliers	4	1.8% Goal: <1.0%
	CA Geometry outliers	0	0.00% Goal: <0.5%
Additional validations	Tetrahedral geometry outliers	2	
	Waters with clashes	0/0	0.00% See UnDownser table for details

In the two column results, the left column gives the raw count, right column gives the percentage.  
\* 100<sup>th</sup> percentile is the best among structures of comparable resolution; 0<sup>th</sup> percentile is the worst. For clashscore the comparative set of structures was selected in 2004, for MolProbity score in 2006.  
<sup>^</sup> MolProbity score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution.  
Key to table colors and cutoffs here: [?](#)

Use the feedback from Molprobity to go back and fix up your model in Coot.

## Homework:

Finish refinement of the NBDs into the map, adding as many of the unmodeled residues as you can. Finish rebuilding the TMD helices and connect them to the NBDs if possible. Use validation to judge your progress. When you're "done" report the following (MolProbity will provide most of these):

- 1) How many residues are modelled (out of how many total):
- 2) What percent of residues are Ramachandran favored/Ramachandran outliers:
- 3) What percent of residues have favored rotomers/poor rotomers:
- 4) What is your clash score:

## Advanced resources:

This was just a taste of the basics of model building. There are many other tools at your disposal (this list is not exhaustive):

Making threading models or homology models:

Modeller: <https://salilab.org/modeller/>

SWISS-MODEL: <https://swissmodel.expasy.org/>

I-TASSER: <https://zhanglab.ccmb.med.umich.edu/I-TASSER/>

Chainsaw (part of CCP4): <http://www.ccp4.ac.uk/html/chainsaw.html>

Automated building:

Buccaneer (part of CCP4): <http://www.ccp4.ac.uk/html/cbuccaneer.html>

ArpWarp (part of CCP4 - stated to work for high resolution maps):

<https://arpwarp.embl-hamburg.de/>

Phenix map\_to\_model:

[https://www.phenix-online.org/documentation/reference/map\\_to\\_model.html](https://www.phenix-online.org/documentation/reference/map_to_model.html)

Refinement

Isolde: <https://isolde.cimr.cam.ac.uk/>

Rosetta: <https://www.rosettacommons.org/software> (also available in Phenix)

Phenix refine:

[https://www.phenix-online.org/documentation/reference/refine\\_gui.html](https://www.phenix-online.org/documentation/reference/refine_gui.html)

Chimera/Coot advanced use:

See the new version of Coot (optimized for lower resolution maps) in action:

<https://www.youtube.com/watch?v=hhB8qUBBnJQ>

Oli Clarke's customization scripts: <https://github.com/olibclarke?tab=repositories>

## Appendix

Human CFTR amino acid sequence. Residues missing in our initial model are highlighted in yellow.

MQRSPLEKASVVSCLFFSWTRPILRKGYRQRLELSDIYQIPSVDSADNLSEKLEREWDRELASKKN  
PKLINALRRRCFFWRFMFYGIFLYLGEVTKAVQPLLLGRIIASYDPDNKEERSIAIYLGIGLCLLFIVRTL  
LLHPAIFGLHHIGMQMRIAMFSLIYKTKLKLSSRVLDKISIGQLVSLLSNNLNKFDEGLALAHFVWIAP  
LQVALLMGLIWELLQASAFCLGLFLIVLALFQAGLGRMMMKEYRDQRAGKISERLVITSEMIENIQSV  
KAYCWEEAMEKMIENLRQTELKLTRKAAYVRYFNSSAFFFSGFFVFLSVLPYALIKGIILRKIFTTIS  
FCIVLRMAVTRQF **PWAVQTWYDSLGAINKIQDFLQKQEYKTLEYNLT**TEVVMENVTAFWEEG**FG**  
**ELFEKAKQNNNNRKTSSNGDDSLFFSNFSL**LGTPVLKDINFKIERGQLLAVAGSTGAGKTSLLMVIM  
GELEPSEGKIKHSGRISFCSQFSWIMPGTIKENIIFGVSYDEYRYSVIKACQLEEDISKFAEKDNIVL  
GEGGITLGGQQRARISLARAVYKDADLYLLDSPFGYLDVLTEKEIFESCVCKLMANKTRILVTSKME  
HLKKADKILILHEGSSYFYGTFSSELQNLQ**PDFSSKLMGCDSFDQFSAERRNSILTETLHRFSLEGDA**  
**PVSWTETKKQSFKQTGEFGEKRKNSILNPINSIRKFSIVQKTPLQMNGIEEDSDEPLERRLSLVPDS**  
**EQGEAILPRISVISTGPTLQARRRQSVLNLMTHSVNNQGNHRKTTASTRKVSLAPQANLTEDIYS**  
**RRLSQETGLEISEEINEEDLKECFDDMESIPAVT**TWNTYLYRYITVHKSLIFVLIWCLVIFLAEVAASL  
VVLWLLGNTPL**QDKGNSTHSRNNSYAVIITSTSSYYVFYIYVG**VADTLLAMGFFRGLPLVHTLITVS  
KILHHKMLHSLVQAPMSTLNTLKAGGILNRFSKDIAILDLLPLTIFDFIQLLLVIGAIYVAVLQPYIFV  
ATVPVIVAFIMLRAYFLQTSQQLKQLESEGRSPIFTHLVTSKGLWTLRAFGRPYFETLFHKALNL  
HTANWFLYLSTLRWFQMRIEMIFVIFIAVTFISILTTGEGEGRVGIILTLAMNIMSTLQ**WAVNSSIDV**  
**DSLMRSVSRVFKFIDMPTEGKPTKSTKPYKNGQLSKVMIIENSHVKKD**DIWPSGGQMTVKDLTAK  
YTEGGNAILENISFSISPGQRVGLLGRTGSGKSTLLSAFLRLNTEGEIQIDGVSWDSITLQQWRKA  
FGVIPQKVFIKSGTFRKNLDPYEQWSDQEIWKVADEVGLRSVIEQFPGKLDFVLVDGGCVLSHGH  
KQLMCLARSVLSKAKILLDEPSAHLDPVTYQIIRRTLKQAFADCTVILCEHRIEAMLECCQQFLVIEEN  
KVRQYDSI**QKLLNERSLFRQAISPSDRVKLFPHRNSSKCKSKPQIAALKEETEEEVQDTRL**